

## TÉCNICAS DE REPRESENTACIÓN EN LA LEXICOGRAFÍA PLURILINGÜE<sup>1</sup>

ALBERTO ÁLVAREZ LUGRÍS  
*Universidad de Vigo*

**RESUMEN.** *En este artículo presentamos una propuesta de codificación para los diccionarios electrónicos plurilingües. Tras un breve panorama de las características de los diccionarios tradicionales y los electrónicos, estableceremos un modelo de estructuración y de artículo lexicográfico básicos a partir del cual empezar a trabajar. De entre los diferentes estándares de marcación, elegiremos SGML para crear un diccionario inglés-gallego-inglés, del que presentamos una DTD provisional.*

**PALABRAS CLAVE.** *Lexicografía, diccionario electrónico, SGML.*

**ABSTRACT.** *This paper presents a proposal for the codification of multilingual electronic dictionaries. After a brief review of the main features of traditional and electronic dictionaries, we propose a basic dictionary structure and entry. Among the different markup standards, SGML is chosen to create an English-Galician-English dictionary, for which a provisional DTD is presented.*

**KEYWORDS.** *Lexicography, electronic dictionary, SGML.*

### 1. INTRODUCCIÓN: INFORMACIÓN Y FORMATO DE LOS DICCIONARIOS

Cuando hablamos de diccionarios plurilingües nos referimos tanto a los diccionarios bilingües como a los multilingües (Haensch *et al.* 1982: 134). Aunque vamos a trabajar con un proyecto de diccionario bilingüe inglés-gallego-inglés, esta exposición y sus conclusiones se pueden aplicar también al resto de los diccionarios bilingües y multilingües, por lo menos a los de nuestro ámbito lingüístico. Por otra parte, entendemos por *representación de la información* la forma en que se codifica información de cualquier tipo en formato electrónico; es decir, el modo en el que esa información se representa para que un ordenador pueda entenderla, manipularla y

producir unos resultados que sean útiles para el usuario. La representación de la información remite siempre a la estructura lógica de la cosa representada, que en nuestro caso es un diccionario electrónico o, más generalmente, un texto electrónico. Simplificando un poco, podemos decir que la estructura lógica básica de un diccionario es su división en entradas o artículos, la división de estos en acepciones y la división de las acepciones en usos. La estructura lógica de otro texto electrónico sería, en el caso de una novela, su división en partes, la división de las partes en capítulos, de los capítulos en párrafos, de los párrafos en oraciones, etc. No importa si los capítulos son más o menos largos, si se utiliza un tipo u otro de letra, si hay efectos tipográficos (negritas, cursivas, etc.). Lo que nos interesa es la información, el contenido informativo de los textos. Otra forma de definir la parte informativa de un texto, en oposición a su formato, es decir que es aquella parte de un texto que no varía de una edición a otra, que permanece inalterable a pesar de los cambios físicos (de formato) que el texto pueda sufrir. La estructura lógica de un texto electrónico, por lo tanto, refleja exclusivamente la ordenación de la información, no el formato con el que esta se imprime o se envía a la pantalla del ordenador.

Dijimos antes que un diccionario electrónico es un tipo de texto electrónico; para completar esta pequeña introducción sobre la información y el formato de los textos, veamos una definición de texto electrónico (básicamente, la que presentamos en Álvarez Lugrís 1996a: 87-89). Llamamos texto electrónico a una versión de otro texto<sup>2</sup> que es manipulable mediante un ordenador y ciertos programas, lo que implica tres grandes e importantes diferencias con el concepto tradicional de texto:

- el texto electrónico no es más que una cadena de caracteres ASCII, una sucesión de bytes almacenada en un disco.
- el texto electrónico, por lo tanto, no tiene entidad física, no es un objeto material y por eso no lo podemos ver ni tocar.
- la única forma de acceder al contenido de un texto electrónico es a través de la pantalla del ordenador o bien mediante impresión; pero debemos tener en cuenta que de esta forma no estamos accediendo al texto electrónico en sí, sino a una representación del mismo.

Aparte de estas características, los diccionarios electrónicos poseen otras, derivadas de su naturaleza lexicográfica. Se ha definido el diccionario tradicional (Haensch et al. 1982: 461-462) como un extracto estandarizado del conjunto o de un subconjunto del léxico de una lengua presentado en forma de lista ordenada según un determinado criterio, normalmente alfabético. Este extracto se divide en artículos o entradas, que son la unidad autónoma más pequeña de los diccionarios. De esta definición podemos deducir dos consecuencias fundamentales para la concepción de los diccionarios electrónicos:

- a) un diccionario es un tipo de texto con una estructura extraordinariamente formalizada, repetitiva y predecible.
- b) la estructura lógica del diccionario es más evidente que la de cualquier otro texto por estar basada en una unidad claramente identificable: el artículo lexicográfico o entrada.

Esto quiere decir que, en principio, la formalización de la estructura lógica de un diccionario de cualquier tipo, podría ser más fácil que la de determinados textos narrativos, por ejemplo, en los que es difícil determinar sus unidades porque combinan voces narrativas, tiempos, estructuras e incluso géneros literarios.

## 2. EL DICCIONARIO TRADICIONAL Y EL DICCIONARIO ELECTRÓNICO

Todas las definiciones tradicionales de diccionario coinciden en que se trata de una colección de palabras (también llamadas voces o dicciones —de donde viene su nombre) explicadas (traducidas a otra lengua o lenguas, en el caso de los diccionarios plurilingües) y ordenadas según algún criterio, normalmente el alfabético. Aunque muchos de los diccionarios electrónicos a los que hoy podemos recurrir son representaciones electrónicas de ediciones impresas hechas a partir de las cintas de composición utilizadas por las editoriales<sup>3</sup>, no debemos ver los diccionarios electrónicos como una simple alternativa de presentación. Existen diccionarios electrónicos concebidos y realizados únicamente en formato electrónico que poseen características definitorias propias diferentes de los diccionarios tradicionales.

La diferencia principal entre estos dos tipos de diccionarios es la total independencia de los electrónicos con respecto a cualquier medio físico, ya sea el papel y los efectos tipográficos o las plataformas y programas informáticos. Quiere esto decir, en primer lugar, que un mismo diccionario electrónico puede tener formatos de salida por pantalla muy diferentes según los gustos o necesidades del usuario. El OED, por ejemplo, nos permite mover las ventanas y variar su tamaño, cambiar los tipos, tamaños y colores de las letras, etc. En segundo lugar, un diccionario electrónico permite, por regla general, exportar en formato electrónico partes de su contenido que luego podemos imprimir. En tercer lugar, la independencia de plataformas y programas informáticos permite que la información lexicográfica de un proyecto pueda ser compartida con otro proyecto, lexicográfico o no, que se beneficiaría fácilmente de un trabajo ya realizado. Pensemos, por ejemplo, en la redacción de nuevos diccionarios o en su actualización, o el aprovechamiento de información semántica o sintáctica para la construcción de bases de datos léxicas<sup>4</sup>.

Por el contrario, los diccionarios tradicionales, tienen un solo formato físico, único e invariable, sea de utilidad o no, nos guste o no. La única forma de extraer su información para reutilizarla es copiándola a mano, por lo que la posibilidad de

reutilización en otros proyectos es mucho menor y tiene el inconveniente de que si en las sucesivas copias se comenten errores se puede llegar a modificar sensiblemente el contenido o la forma de los artículos.

Por otra parte, los diccionarios tradicionales tenían una total, y comprensible, dependencia de la ordenación alfabética: al carecer de otros medios técnicos, el uso de fichas lexicográficas manuscritas imponía una cuidadosa disposición de las entradas antes incluso de comenzar el proceso de redacción. Una vez comenzada la labor del cajista en la imprenta, la más mínima alteración suponía un considerable retraso, ya que el texto del diccionario era copiado por los tipógrafos palabra a palabra, entrada a entrada. Por el contrario, los diccionarios electrónicos imponen sólo la restricción de la coherencia: no podemos redactar dos artículos al mismo tiempo; pero podemos redactarlos sin seguir ninguna ordenación concreta y podemos alterar la ordenación en cualquier momento porque no depende de la secuencia de introducción de la información y se realiza de forma totalmente automática.

Por todo lo anterior, podríamos definir el diccionario electrónico como un conjunto de artículos lexicográficos independientes y almacenados en formato electrónico entre los que se establecen una serie de relaciones hipertextuales.

Que los artículos sean independientes, autónomos, está relacionado con la posibilidad de introducirlos aleatoriamente, sin orden concreto, para someterlos, si es preciso, a un ordenamiento automático posterior. La autonomía de los artículos nos lleva a reflexionar una vez más sobre la independencia con respecto a los formatos físicos, ya que dentro del diccionario electrónico no existe el concepto de página. La página es un tipo más de formato físico que se le puede dar a un texto, como la negrita, las mayúsculas, etc.: no forma parte de la estructura lógica de ningún texto electrónico y tampoco de los diccionarios electrónicos, como lo demuestra el hecho de que la división en páginas puede variar de una edición a otra, a diferencia de elementos como “transcripción fonética” o “acepción 1”, etc. Por supuesto, no existe tampoco el concepto de línea. Recordemos una vez más que las páginas y líneas que vemos en la pantalla de nuestro ordenador o en las hojas que imprimimos son sólo una de las muchas representaciones posibles de la información del diccionario, no una parte de su estructura. Si no existe la página, lo que tenemos es, como decíamos antes, una sucesión de artículos lexicográficos de longitud variable; la longitud se mide, desde luego en bytes o en palabras, no en líneas ni en páginas.

En nuestra definición de diccionario electrónico dijimos que entre los diferentes artículos lexicográficos se establecen una serie de relaciones hipertextuales<sup>5</sup>. Estas relaciones hacen posible acceder rápidamente de un artículo a otro u otros relacionados:

- referencias cruzadas entre entradas; por ejemplo, entre un participio y su infinitivo o entre dos variantes ortográficas. El OED, por ejemplo, en la entrada *legacy* nos remite a *legace* y *legasy*:

† **'legacy**, *v. Obs.* Also 6 **legace**, **-asy**.

[f. prec.]

1. *trans.* To send as a legate.

1563 Foxe *A. & M.* 1373/2 You are legasyd by thauthoritie of the Pope.

2. a. To give or leave as a legacy.

b. To bequeath a legacy to.

1546 *Wills & Inv. N.C.* (Surtees 1835) 126 The reste of all my goodes not beinge legaced nor gyuen.

1594 Nashe *Unfort. Trav. Wks.* (Grosart) V. 185 Where yet liuing, hee might behold his flesh legacied amongst the foules of the aire.

Figura 1. *Entrada legacy en el OED.*

En otra entrada nos advierte de la forma alternativa de una palabra:

**'leftmost**, *a.* Also **leftermost**.

[f. left *a.* + -most.]

Situated furthest to the left.

1863 Kinglake *Crimea* II. 443 The Grenadiers..were making good use of that delicate bend in the formation of their leftmost company.

1875 *Ibid.* (1877) V. i. 269 The leftermost portion of them, under the direction of Serjeant O'Hara.

1894 O. O. Howard in *Voice* (N.Y.) Sept., Mansfield..pushed out toward Lee's leftmost troops.

Figura 2. *Entrada leftmost en el OED.*

- incluso es posible en algunos diccionarios buscar automáticamente cualquiera de las palabras que forman parte de una entrada.

Esto significa que la independencia y autonomía de los artículos no implican aislamiento de los mismos cuando buscamos una información.

Y al hilo de las relaciones hipertextuales entre artículos, podemos recordar que el origen de esta característica de los diccionarios electrónicos tenemos que buscarlo precisamente en los diccionarios tradicionales y sus juegos tipográficos. Cuando los diccionarios comenzaron a ser más completos y complejos y sus entradas se hicieron más largas, se vio la necesidad de facilitar su consulta (Cf. Haensch et al. 1982: 477) mediante el uso de diversos artificios tipográficos: negritas, cursivas, versalitas, barras, números, etc. El cometido de estos juegos era facilitarle al lector la localización de las diversas informaciones dentro de una entrada pero también indicarle qué palabras o expresiones debía buscar en otro artículo.

En la página del diccionario de John Walker (1852:311) que presentamos a continuación, por ejemplo, podemos ver una referencia a la forma moderna de un

adjetivo (*lorn*), la referencia cruzada desde el participio pasado y el pretérito de un verbo a su infinitivo (*lost*), y numerosas referencias numéricas a las tablas de pronunciación que preceden al cuerpo del diccionario.

LOO	LOU
nör 167, nôt 163—tûbe 171, tûb 172, bûll 173—ân 299—pând 313—âm 466, rusia 469.	
any thing fastened; to relax; to free from any thing painful; to disengage.	<b>LOTION</b> , lôshân, <i>s.</i> A lotion is a form of medicine compounded of aqueous liquids, used to wash any diseased parts; a cosmetic.
<b>To LOOSE</b> , lôse, <i>v. n.</i> To set sail, to depart by losing the anchor.	<b>LOTTERY</b> , lôttûr-ê, <i>s.</i> 557. A game of chance, distribution of prizes by chance.
<b>LOOSE</b> , lôse, <i>a.</i> Unbound, untied; not fast; not tight; not crowded; wanton; not close, not concise; vague, indeterminate; not strict, unconnected, rambling; lax of body; disengaged; free from confinement; remiss, not attentive; to break Loose, to gain liberty; to let Loose, to set at liberty, to set at large.	<b>LOUD</b> , lôd, <i>a.</i> 312. Noisy, striking the ear with great force; clamorous, turbulent.
<b>LOOSE</b> , lôse, <i>s.</i> Liberty, freedom from restraint; dismissal from any restraining force.	<b>LOUDLY</b> , lôd'lê, <i>ad.</i> Noisily, so as to be heard far; clamorously.
<b>LOOSELY</b> , lôse'lê, <i>ad.</i> Not fast, not firmly; without bandage; without union; irregularly; negligently; meanly; unchastely.	<b>LOUDNESS</b> , lôd'nês, <i>s.</i> Noise, force of sound; turbulence, vehemence or fierceness of clamour.
<b>To LOOSEN</b> , lôs'an, <i>v. n.</i> 103. To part, to separate.	<b>To LOVE</b> , lûv, <i>v. a.</i> 165. To regard with passionate affection; to regard with tenderness of affection; to be pleased with, to like; to regard with reverence.
<b>To LOOSEN</b> , lôs'an, <i>v. a.</i> To relax any thing tied; to make less coherent; to separate a company; to free from restraint; to make not coercive.	<b>LOVE</b> , lûv, <i>s.</i> 165. The passion between the sexes; kindness, good-will, friendship, affection; courtship, tenderness; liking, inclination to; object beloved, lewdness; fondness, concord; principle of union; picturesque representation of love, a cupid; a word of endearment; due reverence to God; a kind of thin silk stuff.
<b>LOOSENESS</b> , lôse'nês, <i>s.</i> State contrary to that of being fast or fixed; criminal levity; irregularity; lewdness, unchastity, diarrhoea, flux of the belly.	<b>LOVEAPPLE</b> , lûv'âp-pl, <i>s.</i> 405. A plant; the fruit of a plant.
<b>LOOSESTRIFE</b> , lôse'strife, <i>s.</i> An herb.	<b>LOVEKNOT</b> , lûv'nôt, <i>s.</i> A complicated figure, by which affection is figured.
<b>To LOP</b> , lôp, <i>v. a.</i> To cut the branches of trees; to cut off any thing.	<b>LOVELETTER</b> , lûv'lêt-tûr, <i>s.</i> Letter of courtship.
<b>LOP</b> , lôp, <i>s.</i> That which is cut from trees; a diseased.	<b>LOVELY</b> , lûv'lê-lê, <i>ad.</i> Amiably.
<b>LOPEKA</b> , lôp'kâ, <i>s.</i> 98. One that cuts trees.	<b>LOVELINESS</b> , lûv'lê-nês, <i>s.</i> Amiability; qualities of mind or body that excite love.
<b>LOQUACIOUS</b> , lô-kvâ'shûs, <i>a.</i> 414. Full of talk; babbling, not secret.	<b>LOVELORN</b> , lûv'lôr'n, <i>a.</i> Forsaken of one's love. See <i>Forlorn</i> .
<b>LOQUACITY</b> , lô-kvâ'sê-lê, <i>s.</i> Too much talk.	<b>LOVELY</b> , lûv'lê, <i>a.</i> Amiable; exciting love.
<b>LORD</b> , lôrd, <i>s.</i> 167. The Divine Being, Jehovah; monarch, ruler; master; a tyrant; a husband; a nobleman; a general name for a peer of England; an honorary title applied to officers, as, lord chief justice, lord mayor.	<b>LOVEMONGER</b> , lûv'mûng-gûr, <i>s.</i> One who deals in affairs of love.
<b>To LORD</b> , lôrd, <i>v. n.</i> To domineer, to rule despotically.	<b>LOVER</b> , lûv'ûr, <i>s.</i> 98. One who is in love; a friend, one who regards with kindness; one who likes any thing.
<b>LORDING</b> , lôrd'îng, <i>s.</i> Lord in contempt or ridicule.	<b>LOUVER</b> , lôv'ûr, <i>s.</i> An opening for the smoke.
<b>LORDLING</b> , lôrd'lîng, <i>s.</i> 410. A diminutive lord.	<b>LOVESECRET</b> , lûv'sê-krit, <i>s.</i> Secret between lovers.
<b>LORDLINESS</b> , lôrd'lê-nês, <i>s.</i> Dignity, high station; pride, haughtiness.	<b>LOVESICK</b> , lûv'sîk, <i>a.</i> Disordered with love, languishing with amorous desire.
<b>LORDLY</b> , lôrd'lê, <i>a.</i> Befitting a lord; proud, imperious, insolent.	<b>LOVesome</b> , lûv'sûm, <i>a.</i> Lovely. A word not used.
<b>LORDLY</b> , lôrd'lê, <i>ad.</i> Imperiously, proudly.	<b>LOVESONG</b> , lûv'sông, <i>s.</i> Song expressing love.
<b>LORDSHIP</b> , lôrd'shîp, <i>s.</i> Dominion, power; signiory, domain; title of honour used to a nobleman not a duke; titular compellation of judges, and some other persons in authority.	<b>LOVESUIT</b> , lûv'sûte, <i>s.</i> Courtship.
<b>LORE</b> , lôre, <i>s.</i> Lesson, doctrine, instruction.	<b>LOVETALE</b> , lûv'tâle, <i>s.</i> Narrative of love.
<b>To LORICATE</b> , lôr'rê-kâte, <i>v. a.</i> 168. To plate over.	<b>LOVETHOUGHT</b> , lûv'thâwt, <i>s.</i> Amorous fancy.
<b>LORIMER</b> , lôr'rê-mûr, } <i>s.</i> 98. 168. Bridle cutter.	<b>LOVEVOYS</b> , lûv'vôis, <i>s.</i> Small presents given by lovers.
<b>LOBINER</b> , lôr'rê-nûr, }	<b>LOVETRICK</b> , lûv'trîk, <i>s.</i> Art of expressing love.
<b>LOBN</b> , lôm, <i>a.</i> Forsaken, lost. Obsolete.—See <i>Forlorn</i> .	<b>LOUGH</b> , lôk, <i>s.</i> 392. A lake, a large inland standing water.
<b>To LOSE</b> , lôse, <i>v. a.</i> 164. To forfeit by unlucky contest, the contrary to win; to be deprived of; to possess no longer; to have any thing gone so as that it cannot be found or had again; to be wiser; to throw away, to employ ineffectually; to miss, to part with so as not to recover.	<b>LOVING</b> , lûv'îng, <i>part. a.</i> Kind, affectionate; expressing kindness.
<b>To LOSE</b> , lôse, <i>v. n.</i> Not to win, to suffer loss; to decline, to fail.	<b>LOVINGKINDNESS</b> , lûv'îng-kyînd'nês, <i>s.</i> Tenderness, favour, mercy.
<b>LOSEABLE</b> , lôs'â-bl, <i>a.</i> 405. Subject to privation.	<b>LOVINGLY</b> , lûv'îng-lê, <i>ad.</i> Affectionately, with kindness.
<b>LOSER</b> , lôs'ûr, <i>s.</i> 98. One that is deprived of any thing, one that forfeits any thing, the contrary to winner or gainer.	<b>LOVINGNESS</b> , lûv'îng-nês, <i>s.</i> Kindness, affection.
<b>LOSS</b> , lôs, <i>s.</i> Forfeiture, the contrary to gain; damage; deprivation; fault, puzzle; useless application.	<b>LOUIS-D'OR</b> , lô-ê-dôre, <i>s.</i> A golden coin of France, valued at about twenty shillings.
<b>Loss</b> , lôst, <i>part. of To Lose.</i>	<b>To LOUNGE</b> , lôdnje, <i>v. n.</i> To idle, to live lazily.
<b>LOST</b> , lôst, <i>part. of To Lose.</i>	<b>LOUNGER</b> , lôdn'jûr, <i>s.</i> An idler.
<b>LOT</b> , lôt, <i>s.</i> Fortune, state assigned; a chance; a die, or any thing used in determining chances; a portion, a parcel of goods as being drawn by lot; proportion of taxes, as, to pay scot and lot.	<b>LOUSE</b> , lôse, <i>s.</i> 312. A small animal, of which different species live and feed upon the bodies of men, beasts, and perhaps of all living creatures.
<b>LOVE-TREE</b> , lôv'trê, <i>a.</i> The Lotus.	<b>To LOUSE</b> , lôse, <i>v. a.</i> 437. To clean from lice.
	<b>LOUSEWORT</b> , lôse'wûrt, <i>s.</i> The name of a plant.
	<b>LOUSILY</b> , lôs'ê-lê, <i>ad.</i> In a paltry, mean, and scurvy way.
	<b>LOUSINESS</b> , lôs'ê-nês, <i>s.</i> The state of abounding with lice.
	<b>LOUSY</b> , lôs'ê, <i>a.</i> Swarming with lice, over run with lice; mean, low born.

Figura 3. Diccionario de Walker (1852).

TÉCNICAS DE REPRESENTACIÓN EN LA LEXICOGRAFÍA PLURILINGÜE

En esta otra página del diccionario de Velázquez de la Cadena (1861, vol 2:271), vemos referencias que apuntan a sinónimos (*fellon*), abreviaturas (*fellow*), formas ortográficas alternativas (*fellies*), al infinitivo de un participio (*felt*) o a la forma más moderna de la palabra (*feminity*).

FEL	FEL	FEN
<p>quiseda; y yerba; w guapo; h fáco. — q chico; j yema; t zapá; d dedo; x zèle (Fr.); j chéz (Fr.); g Jean; y saagre.</p> <p>guna cosa como las cabezas, los árboles, etc. <i>Felling of wood</i>. Corta ó corte de monte ó leña.</p> <p><b>FELLER</b> [fe'l-ur] s. Derribante.</p> <p>+<b>FELLIFLUOUS</b> [fe-li'f-lu-us] a. Lo que d'uye hiel.</p> <p><b>FELLMONGER</b> [fe'l-muŋ-gur] s. Pellejero, el que trata en pellejos.</p> <p><b>FELINESS</b> [fe'l-nes] s. Crueldad, ferocidad, barbarie.</p> <p><b>FELLOES</b> [fe'l-oz] s. pl. Pinas, los maderos curvos que forman el círculo de la rueda. Escríbese á veces <i>fellies</i>.</p> <p><b>FELTON</b> [fe'l-un] s. Postemillas. V. <b>FELON</b>.</p> <p><b>FELLOW</b> [fe'l-o] s. 1. Compañero, camarada, sócio; cómplice. 2. Compañero, hermano, la cosa que hace juego con otra. <i>My buckles are not fellows</i>. Mis hebillas no son compañeras, ó no son iguales. 3. Expresion familiar usada para denotar unas veces cariño y otras desprecio. A' veces es expletivo cuando se unc á los adjetivos, como, <i>A brave fellow</i>. Un valiente. (Fam.) Un buen chico, una buena alhaja ó maula. <i>A worthless fellow</i>. Un para nada, un pícaro. 4. Sócio ó individuo de alguna colegio, sociedad ó academia. 5. <i>Fellow being ó creature</i>. El prójimo, nuestro semejante. V. <i>F.</i> como abreviatura.</p> <p>To <b>FELLOW</b>, va. Hermanar, igualar una cosa con otra.</p> <p><b>FELLOW-CITIZEN</b> [fe'l-o-si'ti-zen] s. Ciudadadano, compatriota.</p> <p><b>FELLOW-COMMONER</b> [fe'l-o-co'm-un-ur] s. El que tiene los mismos derechos que otro.</p> <p><b>FELLOW-COUNSELLOR</b> [fe'l-o-er'n-sel-tr] s. Individuo del mismo consejo.</p> <p><b>FELLOW-CREATURE</b> [fe'l-o-cre'-gur] s. Criatura de la misma especie.</p> <p><b>FELLOW-FEELING</b> [fe'l-o-fe'l-in] s. 1. Simpatía, compasión. 2. Interés comun.</p> <p><b>FELLOW-HEIR</b> [fel-o-a'r] s. Coheredero.</p> <p><b>FELLOW-HELPER</b> [fe'l-o-help-ur] s. Coadjutor, coadyuvador.</p> <p><b>FELLOW-LABORER</b> [fe'l-o-lá-bur-ur] s. Colaborador, compañero en algun trabajo.</p> <p>+<b>FELLOWLIKE</b> [fe'l-o-ljé].</p> <p>+<b>FELLOWLY</b> [fe'l-o-li] a. Igual, como compañero.</p>	<p><b>FELLOW-MAIDEN</b> [fe'l-o-ma'd-n] s. Doncella que vive ó está con otra.</p> <p><b>FELLOW-MEMBER</b> [fe'l-o-me'mbur] s. Individuo de la misma sociedad; miembro del mismo cuerpo; compañero.</p> <p><b>FELLOW-MINISTER</b> [fe'l-o-mi'n-is-tur] s. El que sirve el mismo oficio.</p> <p><b>FELLOW-PARTNER</b> [fe'l-o-pq'rt-nur] s. Consocio en algun negocio ó casa de comercio, donde hay mas de dos compañeros.</p> <p><b>FELLOW-PERR</b> [fel-o-pe'r] s. El que goza los mismos privilegios de nobleza, como sucede en Inglaterra, donde los pares tienen títulos y precedencia diferentes, aunque son iguales en los demas privilegios.</p> <p><b>FELLOW-PRISONER</b> [fe'l-o-priz-un-ur] s. Preso en la misma cárcel ó por la misma causa.</p> <p><b>FELLOW-SCHOLAR</b> [fe'l-o-sco'l-ur] s. Condiscípulo.</p> <p><b>FELLOW-SERVANT</b> [fe'l-o-su'r-vant] s. Criado ó criada que tiene el mismo amo.</p> <p><b>FELLOWSHIP</b> [fe'l-o-sip] s. 1. Compañía, 2. Asociación. 3. Igualdad. 4. Compañía, convenio ó contrato entre comerciantes. 5. Sociedad, compañía, reunion social. 6. Buen humor. 7. Beca, plaza de colegial en algun colegio. 8. (Arit.) Regla de compañía.</p> <p><b>FELLOW-SOLDIER</b> [fe'l-o-sol-jur] s. Comiliton; úsase por cariño.</p> <p><b>FELLOW-STUDENT</b> [fe'l-o-stu'dent] s. Condiscípulo.</p> <p><b>FELLOW-SUBJECT</b> [fe'l-o-su'b-ject] s. El que vive bajo el mismo gobierno.</p> <p><b>FELLOW-SUFFERER</b> [fe'l-o-su'f-ur-ur] s. El que sufre por la misma causa ó al propio tiempo que otro.</p> <p><b>FELLOW-TRAVELLER</b> [fe'l-o-tra'v-el-ur] s. Compañero de viaje.</p> <p><b>FELLOW-WORKER</b> [fe'l-o-wu'rc-ur] s. El que trabaja en el mismo intento ó objeto que otro.</p> <p><b>FELLOW-WORKMAN</b> [fe'l-o-wu'rc-man] s. Artesano que trabaja en la misma obra que otro.</p> <p><b>FELLOW-WRITER</b> [fe'l-o-ri't-ur] s. El que escribe al mismo tiempo ó sobre el mismo asunto que otro.</p>	<p><b>FELLY</b> [fe'l-li] ad. Cruelmente, bárbaramente, ferozmente. — s. V. <b>FELLOWS</b>.</p> <p><b>FELLO-DE-SE</b> [fe'l-o-de-se'] s. (For.) Suicida.</p> <p><b>FELON</b> [fe'l-un] s. 1. Reo de algun delito capital. 2. Panarizo, panarizo, uñero. — a. Crúel traidor.</p> <p><b>FELONIOUS</b> [fe'l-o-ni-us] a. Malvado, perverso, traidor, villano, maligno, pérfido.</p> <p><b>FELONIOUSLY</b> [fe'l-o-ni-us-li] ad. Traidoramente, alvadosamente, malvadamente.</p> <p><b>FELONY</b> [fe'l-o-ni] s. 1. Felonia. 2. El delito que tiene pena capital.</p> <p><b>FELSPAR</b> ó <b>FELDSPAR</b> [fe'l-spqr] s. (Min.) Feldspato, una especie de piedra dura.</p> <p><b>FELT</b> [felt] s. 1. Fieltro, lana no tejida, sino unida ó incorporada á fuerza de agua caliente, lejía y goma. 2. Cuero ó pellejo. — pref. del verbo To <b>FELT</b>.</p> <p>To <b>FELT</b>, va. Trabajar el fieltro para los sombreros.</p> <p><b>FELTMAKER</b> [fe'lt-mac-ur] s. El oficial de sombrero que trabaja el fieltro.</p> <p>To <b>FELTER</b> [fe'lt-ur] va. Unir alguna cosa como si fuese fieltro.</p> <p>* <b>FELUCCA</b> [fe-lu'ca] s. Faluca, falúa, embarcacion de seis remos.</p> <p><b>FEMALE</b> [fe'-mal] s. Hembra. — a. Femenino. <i>Female screw</i>. Hembra de tornillo.</p> <p><b>FEME COVERT</b> [fem-co'v-urt] s. (For.) Mujer casada. <i>Feme sole</i>. Soltera.</p> <p><b>FEMINALITY</b> [fem-i-nu'l-i-ti] s. Femenidad.</p> <p>+<b>FEMINATE</b> [fe'm-i-nat] a. Femenino.</p> <p><b>FEMININE</b> [fe'm-i-nin] a. 1. Femenino, femineo. 2. Femenil, tierno, delicado. 3. Mujeril, afeminado. — s. Hembra.</p> <p>+<b>FEMINITY</b>, s. V. <b>FEMINALITY</b>.</p> <p>+To <b>FEMINIZE</b> [fe'm-i-niz] va. Afeminar.</p> <p><b>FEMORAL</b> [fe'm-o-ral] a. Femoral, perteneciente al muslo.</p> <p><b>FEN</b> [fen] s. Marjal, pantano.</p> <p><b>FENBERRY</b> [fe'n-ber-i] s. Especie de zarzamora.</p> <p><b>FEN-BORN</b> [fe'n-bern] a. Nacido en pais pantanoso.</p> <p><b>FENCE</b> [fens] s. 1. Defensa, reparo, resguardo. 2. Cerca, palizada, vallado que se pone alrededor de algun terreno para dividirlo de otro y resguardarlo. 3.</p>

Figura 4. Diccionario de Velázquez de la Cadena (1861)

En los diccionarios electrónicos, las hiperreferencias son más fáciles de hacer, se pueden hacer en mayor cantidad y no hay peligro de que queden referencias cruzadas sueltas: es más fácil asegurarse de que si A remite a B, B remita a A.

## 2. LA ESTRUCTURA DE LA INFORMACIÓN EN LOS DICCIONARIOS PLURILINGÜES

Entendemos por estructura de la información la disposición lógica de la información lexicográfica dentro de cada entrada, es decir, la forma en la que esa información se distribuye en una jerarquía de elementos como lema, transcripción fonética, clase de palabra, acepciones, ejemplos, etc. Todos estos elementos tienen distinta importancia dentro del artículo, por lo que su distribución no es completamente libre, aunque puede variar de diccionario a diccionario; pero, por ejemplo, sería extraño encontrar casos en los que los ejemplos precediesen a las explicaciones o traducciones del lema.

Una vez más tenemos que recordar la diferencia existente entre la información contenida en el diccionario electrónico y sus distintos formatos de salida por pantalla; mientras este último puede ser muy semejante al de los diccionarios tradicionales, el “aspecto”, si así le podemos llamar, de la información contenida en la base de datos electrónica es muy distinto, ya que la base está diseñada para ser “entendida” por un ordenador, no por un usuario humano (cf. Boguraev 1991: 227-228). Por este motivo, y como ya apuntamos más arriba, debemos tener presente la gran diferencia entre los datos (la información lexicográfica) y el formato de los datos<sup>6</sup>.

Cuando un equipo de lexicógrafos se propone hacer un diccionario, ya sea monolingüe o plurilingüe, lo primero que tiene que decidir y dejar bien fijado es la estructura de la entrada: qué información se va a incluir sobre cada palabra y, sobre todo, cuál será la mejor disposición posible de los datos. En la decisión influyen factores como la facilidad de localización de determinados componentes como por ejemplo la(s) acepción(es) más común(es) o la(s) traducción(es) más corriente(s) a otra(s) lengua(s). Ese es, después de todo, uno de los criterios con los que los usuarios potenciales van a decidir si lo compran o no. Pero en la concepción general de un artículo lexicográfico entra también la concepción que de la lengua o lenguas implicadas tiene el autor (cf. Haensch *et al.* 1982: 297-328). Hay determinados diccionarios que optan por separar lemas homónimos (como gallego *pota*, calamar o recipiente) según criterios etimologistas:



**pota**<sup>1</sup> (do lat. *pottu* = bebida, vaso [?]) *s. f.*  
1. Recipiente redondo, de barro ou metal, máis ben baixa, con unha ou dúas asas, que se usa para cocíñar; ola. 2. Caldeira do alambique.  
FAM. **Potada.**  
**pota**<sup>2</sup> (de *pota*<sup>1</sup>, ou cat. *pota* = pata) *s. f.*  
Molusco cefalópodo, semellante á lura.  
FAM. **Poteira.**

Figura 5. *Diccionario Xerais da Lingua*

Otros optan por agrupar todas los homónimos dentro de un solo artículo dándole así más importancia a la polisemia, pues defienden que hoy en día el recurso a la etimología, desconocida por la mayor parte de los hablantes, no tiene sentido:

**Pota**, *s. f.* Vasilla de ferro, porcelana ou aluminio, de menor tamaño que o pote // Caldeira do alambique // Nome aplicado a diversos moluscos cefalópodos.

Figura 6. *Diccionario Ir Indo* (Feixó Cid ed. 1986)

Finalmente, hay otras obras que optan por la solución intermedia de aplicar los dos criterios según los casos, como el *Diccionario Oxford Español-Inglés Inglés-Español* (cf. Álvarez LUGRÍS 1996b: 241):

- row<sup>1</sup>** *n* I /rəʊ/ 1(a)(straight line) hilera *f*; (of people) fila *f*; (of seats) fila *f*; **the trees were planted in a ~** los árboles estaban plantados en hilera; **a ~ of houses** una hilera de casas; **they lined up in ~s** hicieron filas *or* se formaron en fila; **a seat in the front/fifth ~** un asiento en primera fila/en la quinta fila; **to have a hard ~ to hoe** tener\* una vida muy dura (b) (in knitting) pasada *f*, carrera *f*, vuelta *f* (Arg), corrida *f* (Chi)
- 2 (succession) serie *f*; **he experienced a ~ of failures** sufrió una serie de fracasos, sufrió fracaso tras fracaso; **four times in a ~** cuatro veces seguidas; **three days in a ~** tres días seguidos
- 3 (Leisure, Sport): **we went for a ~** fuimos *or* salimos a remar *or* a dar un paseo en bote
- II /raʊ/ (a)(noisy argument) pelea *f*, riña *f*; **a family ~** una rencilla familiar; **to have a ~ with sb** pelearse *or* reñir\* con algn (b) (about a public matter) disputa *f* (c) (noise) (no pl) ruido *m*, bulla *f* (fam); **they were making a hell of a ~** (colloq) estaban armando *or* haciendo tremenda bulla (fam), estaban armando mucho jaleo (fam)
- row<sup>2</sup>** *vt* /rəʊ/: **he ~ed the boat towards the shore** remó hacia la orilla; **we ~ed them/the supplies across the river** los llevamos/llevamos las provisiones hasta la otra orilla a remo; **they ~ed her out to the ship** la llevaron en bote hasta el barco
- ~ *vi* I /rəʊ/ remar; **to go ~ing** salir\* *or* ir\* a remar; **to ~ across a river** cruzar\* un río a remo
- II /raʊ/ pelearse, reñir\*

Figura 7. Diccionario Oxford.

En la lexicografía alemana es frecuente agrupar en un solo bloque entradas que pertenecen a una misma familia léxica o que comparten una parte del lema. Es una forma de ahorrar espacio, pero también de dejar patente la tendencia del alemán a la derivación y a la composición:

Gu'mmi|faden *m* hilo *m* de goma || -feder *f*  
 (Tech) muelle *m* de caucho || -geschwulst *f* →  
 Gumma || -gurt *m* correa *f* de goma || -gutt *n*  
 [-s] goma guta, gutagamba *f* || <sup>2</sup>haltig *adj*  
 gomífero, gomoso || -hammer *m* martillo *m* de  
 goma od caucho || -handschuh *m* guante *m* de  
 goma || -harz *n* resina de goma, gomorresina *f* ||  
 -höschen *n* bragas *fpl* de goma || -industrie *f*  
 industria *f* gomera || -kabel *n* cable *m* bajo  
 (envoltorio de) goma || -kissen *n* cojín *m* de  
 goma || -knoten *m* → Gumma || -knüppel *m*  
 porra *f* || -lack *m* goma *f* laca || -linse *f* (Fot)  
 objetivo de focal variable, zoom *m* || -lösung *f*  
 solución *f* de goma || [für Temperafarben] agua  
 [fel] goma, temple *f* || -manschette *f* puño *m* de  
 goma od impermeable || -mantel *m* (abrigo)  
 impermeable, abrigo *m* de caucho bzw  
 impregnado od engomado || -matte *f* estera bzw  
 [größer] alfombra *f* de goma || -paragraph *m*  
 (fam) artículo *m* (de interpretación) flexible (en  
 una ley) || -puffer *m* (Tech) [Anschlag] tope *m*  
 de goma, [Dämpfer] amortiguador *m* de goma ||  
 -puppe *f* muñeca *f* de goma || -räder *npl* ruedas  
*fpl* de goma || -reifen *m* neumático *m* || -ring *m*  
 anillo *m* de caucho || -sauger *m* [für Säuglinge]  
 chupete, chupador *m* || -scheibe *f* disco *m* od  
 arandela *f* de goma || -schlauch *m* manga od  
 manguera *f* od tubo *m* de goma od caucho ||  
 -schuh *m* chancho *m* de caucho, galocha *f* de  
 goma, (Am) zapatón, zapato *m* de goma ||  
 -schürze *f* delantal *m* de goma || -schwamm *m*  
 esponja *f* de goma || -seil *n* cable *m* de goma ||  
 -sohle *f* suela *f* de goma || -sonde *f* (Med)  
 sonda de goma, candelilla *f* || -stempel *m* sello  
*m* de goma || -stiefel *m* bota *f* de goma || -stoff  
*m* [Tuch] tela *f* impregnada od engomada ||  
 -strumpf *m* media *f* elástica od de goma || -tier  
*n* [Spielzeug] animal *m* bzw figura *f* de goma ||  
 -tragant *m* goma tragacanto, alquitira *f* ||  
 -überschuh *m* → -schuh || -überzug *m*  
 [Präservativ] preservativo *m*, goma *f* higiénica,  
 (pop) condón *m* || -unterlage *f* [für Säuglinge]  
 impermeable *m* || → -kissen || -walze *f*  
 [Straßenbau] rodillo *m* apisanador de caucho ||  
 -waren *fpl* → -artikel || -wasser *n* solución *f* de  
 goma, agua [fel] goma || -wulst *m* pestaña *f* od  
 burlete *m* de goma || -zelle *f* celda *f* de  
 seguridad od acolchada || -zeug *n* tela *f* de  
 goma impermeable || -zug *m* [am Stiefel usw.]  
 elástico *m*

Figura 8. Diccionario Slaby-Grossman (1991).

En el caso de la lexicografía árabe, por ejemplo, la agrupación de las entradas viene motivada por la estructura de la lengua: a partir de una serie relativamente corta de raíces, se forman todas las otras palabras del idioma mediante una serie de reglas de derivación. Obsérvese cómo bajo el lema principal (fa'ala, hacer) se agrupan en primer lugar otros verbos como fa'ala, tafa'ala, efa'ala, etc.; después encontramos los sustantivos, como fe'ala, adjetivos como fe'eli, etc.

activo, eficaz. <sup>2</sup> فَعَالٌ نَفْعًا	hacer; obrar; ejercer; <sup>2</sup> فَعَلَهُ فَعَلٌ *
pies [pros]. أَفَاعِيلٌ <sup>2</sup>	ejecutar.
impresión. excitación. اِنْفِعَالٌ جَات	escandir, medir un verso. فَعَلٌ
excitable. efectivo. اِنْفِعَالِيٌّ	influenciarse. reaccio- نَفَاعَلٌ N
excitabilidad; recep- اِنْفِعَالِيَّةٌ	nar; combinarse.
tividad.	ser hecho. ser impresio- اِنْفَعَلٌ
pie [pros]. تَفَاعِيلٌ ج تَفَاعِيلٌ <sup>2</sup>	nado; ser excitado.
influencia, acción تَفَاعَلٌ جَات	falsificar, alterar. in- اِنْفَعَلٌ
mutua. reacción; combinación; proceso.	ventar. كَذِبًا عَلَى - calumniar, inventar infundios contra.
hecho. efecto, مَفْعُولٌ ج مَفَاعِيلٌ <sup>2</sup>	acción, acto, obra. فَعْلَةٌ جَات
resultado. objeto, complemento [gr]. اِسْمٌ اِل - participio	acción, acto. fun- اِفْعَالٌ - اِفْعَالٌ
no-agentivo/pasivo. تَعَرَّى - تُعَرَّى tener/hacer efecto. رَجَعِيٌّ - efecto retroactivo. بِه - objeto directo.	ción. verbo. بِال - efectivamente, en efecto. بِ - a consecuencia de.
impresionado; excitado. مَنَعِلٌ	hazaña, gesta. ج اِفَاعِيلٌ <sup>2</sup>
impresionable; excitable.	actual; efectivo. práctico. فَعْلِيٌّ
artificial. falso; fal- تَفَعَّلَ	objetivo. verbal.
sificado. جَا بِال - hacer algo extraordinario.	costumbre, hábito. فَعْلَةٌ
	acción (buena), actuación فَعَالٌ
	de una persona.
	¡haz!. فَعَالٌ
	acción [colectiva]. فَعَالٌ
	mango. ج فَعَلٌ
	efec- فَاعِلٌ ج وِن / فَعْلَةٌ م ة جَات / فَوَاعِيلٌ <sup>2</sup> -
	tivo. agente, actor. operario activo. اِسْمٌ اِل - participio agentivo/activo. رَادِيُوْمٌ - radioactivo.
	eficacia. acti- فَاعِلِيَّةٌ & فَعَالِيَّةٌ N
	vidad.

Figura 9. *Diccionario árabe-español de Corriente* (1988)

Los diccionarios multilingües, al ocuparse de más de dos lenguas, suelen ofrecer menos información, por lo que su estructuración es más sencilla y simple. Como ejemplo, veamos una página del *Wörterbuch der Handels- und Geschäftssprache* (Bruker et al. 1867).

**Legen,** 1) *v. a.*, Geld in ein Geschäft l. — *h.* geld in eene zaak steken. — *e.* invest money in a concern. — *f.* verser de l'argent dans un commerce. — *i.* mettere danaro in un negozio. — *s.* contribuir fondos á un negocio. — Geld auf Interessen l. — *h.* geld op interest uitzetten. — *e.* put money upon interest. — *f.* mettre de l'argent en rente. — *i.* mettere danaro a interesse. — *s.* colocar dinero á reditos. — An's Herz l. — *h.* aanbevelen. — *e.* recommend urgently. — *f.* recommander. — *i.* raccomandare. — *s.* recomendar eficazmente. — 2) *v. pron.*, *s. v. a.* sich mit Etwas beschäftigen: sich auf einen Geschäftszweig l. — *h.* zich aan een handelstak wijden. — *e.* devote one's self to a line of business. — *f.* s'occuper d'une branche. — *i.* applicarsi ad un ramo del commercio. — *s.* dedicarse á un ramo de comercio. — 3) *s. v. a.* sich vermindern: die Nachfrage hat sich gelegt. — *h.* de vraag heeft opgehouden. — *e.* the demand has subsided. — *f.* la demande s'est calmée. — *i.* la domanda si è calmata. — *s.* el pedido ha aflojado (cesado.)

Figura 10. *Diccionario comercial de Bruker et al. (1867).*

Sea cual sea el tipo de diccionario o los criterios que se siguen en su realización, la información que la mayoría de ellos suele contener es la siguiente:

1. lema o cabecera de la entrada: la palabra por la cual se ordena alfabéticamente el artículo. Es la forma básica de cada palabra y suele consignarse en lo que se denomina *forma de cita*: el infinitivo de los verbos, el masculino singular de los sustantivos y adjetivos (en el caso de lenguas con flexión en estas clases de palabras, como el gallego o el castellano).
2. transcripción fonética en el caso de lenguas que, como el inglés, no tengan unas normas de pronunciación fijadas y sin excepciones (o casi). Para la

transcripción suele emplearse el alfabeto de la IPA (Asociación Fonética Internacional), aunque ciertos diccionarios emplean otras notaciones.

3. acepciones, usos o significados establecidos y aceptados de cada palabra.
4. acotaciones diafásicas, diastráticas, diatópicas, etc., es decir, los diferentes niveles de uso y variedades o registros: familiar, culto, vulgar; literatura, química, teatro; inglés australiano, francés de Quebec, español de Colombia; obsoleto, arcaico, etc.
5. ejemplos de las acepciones o de parte de ellas. Pueden ser inventados o tomados de textos reales, normalmente de un corpus de lengua, como el diccionario *Collins COBUILD English Language Dictionary* (Sinclair (ed.) 1987).

En el caso de los diccionarios plurilingües, la explicación de los significados de cada palabra se sustituye por su traducción a una o varias lenguas, pero pueden utilizarse también las acotaciones de uso.

Otros elementos que pueden estar presentes en los artículos de los diccionarios son:

6. formas alternativas del mismo lema, es decir, variantes ortográficas, relativamente frecuentes en inglés.
7. la caracterización del lema según la clase de palabra, forma de la flexión verbal o nominal o si la conjugación verbal es regular o irregular.
8. la etimología de cada palabra mediante una pequeña indicación que por lo general suele ser muy breve y concisa.
9. un apartado de fraseología donde se recogen frases hechas, refranes, proverbios, etc. en los que participa el lema.
10. subentradas, como los *phrasal verbs* ingleses; a pesar de tener la estructura de artículos independientes, suelen estar incluidos dentro de la entrada que se ocupa del verbo a partir del cual se forman. Estas subentradas pueden llegar a contener todos los elementos de una entrada menos, lógicamente, la subentrada. Dentro de la subentrada encontraríamos también las palabras compuestas de los diccionarios alemanes, agrupadas dentro del artículo de la palabra principal.
11. además, en el caso de los diccionarios plurilingües, las traducciones del lema a otra(s) lengua(s) pueden llevar también una caracterización según la clase de palabra, género, forma de la flexión verbal o nominal y algunas otras características como la regularidad o irregularidad de la flexión o, en el caso del inglés, la adscripción al grupo de palabras contables o incontables.

Veamos algunos ejemplos:

**terrible** ('tɛrɪb(ə)l), *a.* Also 5-6 **terry-**, 6 **terra-**, **terre-**, **tirre**.

[*a.* F. *terrible* (12th c.), *ad.* L. *terribilis*, *f.* *terrere* to frighten: see -ble.]

1. Exciting or fitted to excite terror; such as to inspire great fear or dread; frightful, dreadful.
  - 1860 Tyndall *Glac.* i. vii. 50 A foe more terrible than the avalanches.
  - 1870 Swinburne *Ess. & Stud.* (1875) 311 Superb instances of terrible beauty undeformed by horrible detail.
2. **a.** Exciting some feeling akin to dread or awe; very violent, severe, painful, or bad; hence *colloq.* as a mere intensive: Very great, excessive. (Cf. the similar use of *tremendous*, *awful*, *frightful*, etc.)
  - 1596 Dalrymple tr. *Leslie's Hist. Scot.* (S.T.S.) I. 128 Thair constant amitie..to thair nychtbouris the Britanis brocht a terrabill feir.
  - 1965 E. J. Howard *After Julius* vii. 100 Her mother had made his life so terrible—not worth living.
- b.** Applied to a person who behaves in a shocking or outrageous manner; **terrible child or infant** = enfant terrible; **terrible twins**, applied *joc.* to a pair of associates whose behaviour is troublesome or outrageous.
  - 1926 A. Huxley *Two or Three Graces* 25 Peddley was not the man to be put out by even the most terrible of terrible infants.
  - 1982 *Financial Times* 10 May 10/3 Since the 'terrible twins', as they were dubbed, were both powerful figures, and did not always see eye-to-eye on matters, this rivalry was reflected down the line.
- c.** As a hyperbolic term of depreciation: of shockingly poor performance or quality; incompetent; defective.
  - 1983 R. Rendell *Speaker of Mandarin* xvi. 190, I got this Hollywood offer and I went to Hollywood and made that terrible *Mind over Matter*.
3. *quasi-adv.* = terribly. (Esp. in sense 2.) Now chiefly *dial.* and *U.S.*
  - 1906 *Westm. Gaz.* 21 Apr. 4/1 There was only one burglar, by no means a terrible-looking fellow.

Figura 11. *Entrada terrible del OED2 en CD-ROM.*

En este extracto del artículo “*terrible*” del *OED2 on CD-ROM*, podemos ver la transcripción fonética, caracterización, grafías alternativas, etimología, acepciones y ejemplos.

Veamos ahora el mismo artículo en el diccionario Oxford Inglés-Español:

**terrible** /'terəbəl/ *adj* **(a)** (very bad) (<movie/singer/weather>) espantoso, atroz, malísimo, fatal (Esp fam); **it will look ~ if we don't go** quedará muy mal que no vayamos; **to feel ~ (ill)** sentirse\* *or* encontrarse\* pésimo *or* muy mal; (guilty, ashamed) sentirse\* muy mal; **she had (a) ~ toothache** tenía un terrible dolor de muelas; **that's ~!** ¡qué terrible!, ¡qué horror!; **you're making a ~ mistake** estás cometiendo un tremendo *or* terrible error; **the ~ thing is ...** lo peor del caso es que ...; **she's a ~ one for exaggerating** es muy exagerada; **you're ~!** ¡qué malo eres! **(b)** (*as intensifier*): **he's a ~ bore** es terriblemente aburrido, es un pelmazo (fam); **what a ~ shame!** ¡qué lástima más grande!; **it's a ~ waste of time** es una pérdida lamentable de tiempo **(c)** (terrifying) (liter) sobrecogedor (liter), imponente; **Ivan the T~** Iván el Terrible

Figura 12. *Entrada terrible del Diccionario Oxford Inglés-Español.*

Aparte de los elementos que vimos en el ejemplo anterior, nos encontramos ahora con indicaciones de uso (Esp fam, liter), indicaciones gramaticales (*as intensifier*), modelos de sustantivos que pueden ser modificados por el adjetivo (<movie/singer/weather>), sinónimos (very bad, guilty, ashamed, terrifying).

Como podemos ver en estos dos ejemplos, a veces las acepciones pueden agruparse si tienen cierta afinidad semántica: **2a.**, **2b.**, **2c.** en el caso del OED o **(a)**, **(b)**, **(c)** en el caso del Oxford bilingüe.

En estos ejemplos podemos comprobar que, como dijimos antes, los efectos tipográficos nos ayudan a localizar ciertos elementos dentro del artículo.

#### 4. PROPUESTA DE ENTRADA PARA UN DICCIONARIO BILINGÜE

Estamos ahora en condiciones de hacer una propuesta de entrada para un proyecto de diccionario electrónico bilingüe o, mejor dicho, una propuesta de estructuración de la información lexicográfica. La propuesta que haremos se basa en la estructura de las entradas que presentamos en el *Diccionario Manual Inglés-*



*Gallego-Inglés* (cf. Iglesias Rábade, en preparación), que se encuentra aún en fase de redacción. El siguiente esquema recoge los elementos de los que consta la entrada:

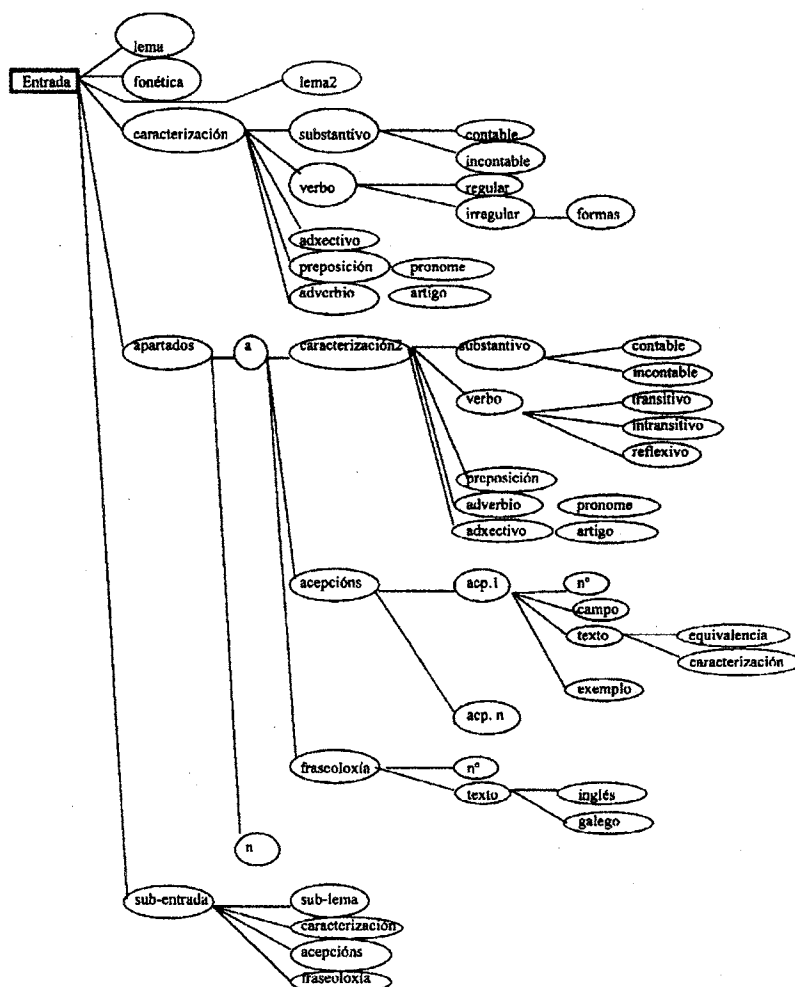


Figura 13. Esquema de entrada del Diccionario Manual Inglés-Galego-Inglés.

Esta sería la entrada del tipo principal, la más numerosa en un diccionario, en la que se nos ofrecen las posibles traducciones de voces de una lengua a otra, pero existen también otros tipos de entradas (como referencias cruzadas, abreviaturas, explicaciones de afijos) que no trataremos ahora.

La entrada o artículo consta de los siguientes elementos:

1. lema
2. transcripción fonética
3. lema2 o variante ortográfica
4. caracterización: clase de palabra (en inglés). Los posibles valores de este elemento son: *noun*, *verb*, *adjective*, *preposition*, *adverb*, *pronoun*, *article*.
5. apartados: división de la entrada en función de caracterización 2.

6. caracterización<sup>2</sup>: mientras que la anterior caracterización nos dice, por ejemplo, que una palabra es un sustantivo, en esta ocasión nos dice que ese sustantivo tiene unas acepciones como contable (*count*) y otras como incontable (*uncount*). En el caso de los verbos: *transitive, intransitive*.
7. Acepciones: dividida en acp.1, acp.2, etc., que recogen las diferentes acepciones de cada lema dentro de cada apartado.
8. acp.: cada una de las acepciones.
9. n°: número de orden de cada def. dentro del apartado.
10. campo: cada una de las acotaciones de uso, variedad, etc.
11. texto: el texto de la acepción, que constará de una o varias equivalencias (o traducciones) al gallego del lema inglés junto con la caracterización de esa equivalencia.
12. equivalencia: la traducción al gallego.
13. caracterización de la equivalencia en gallego.
14. ejemplo: texto con la ejemplificación de cada def.
15. fraseología: relación de frases hechas, giros, modismos, etc. no incluidos en las diferentes acp.
16. n°: que hace referencia a la acp. con la cual la fraseología está relacionada semánticamente.
17. texto de la unidad fraseológica, compuesto por los elementos inglés y gallego.
18. inglés: fraseología inglesa.
19. gallego: fraseología gallega.
20. subentrada: se utilizará para los phrasal verbs.
21. sublema: el lema de la subentrada.
22. caracterización del sublema.
23. acepciones de la subentrada.
24. fraseología de la subentrada.

Optamos con esta entrada por lo que podríamos denominar como opción maximalista: incluir en el artículo tipo el mayor número posible de elementos, para prever todos los casos posibles. Desde luego, habrá muchos artículos que no contengan todos estos elementos. Por este motivo definimos un grupo de elementos obligatorios que son comunes a todas las entradas, desde las más cortas a las más largas: lema, transcripción fonética, apartados (puede ser un solo apartado) y acepciones (una como mínimo).

Veamos como ejemplo un fragmento de una entrada construida según esta estructura:

**bring**/brɪŋ/ v *irr* (*pret & pp brought*) (a) v *tr* 1 *luck, news, object, person, etc*  
traer, levar, conducir; *Jane brought candies for the children Jane trouxo*

caramelos para os cativos. **2** (*cause*) *profit, damage* causar, provocar, traer; *you brought nothing but trouble* non fixeches máis que amolar; *these recollections bring tears to his eyes* estes recordos fan que se lle enchan os ollos de bágoas/auga. **3** (Jur) *charge* facer, formular; *suit, case* principiar, iniciar, comezar, ver; *no charges will be brought* non se fará acusación ningunha; *the case was brought before the court* a causa foi vista polo tribunal. **4** inducir, convencer, facer, *that fact brought me to see that...* aquel feito fíxome decatarme de que...; *he was brought to note his error* convencérono do seu erro.

**FRAS:** **1** *to ~ to a conclusion (matter, business)* concluír, levar a/dar cabo, rematar (asunto, negocio); *~ sth (over) here* traer (algo) aquí, traelo; *~ (sth) closer* achegar, abeirar, acaroar, aproximar (algo). See *house, light, sense*. **2** *to ~ sth to happen* facer que algo suceda, causar. **4** *to ~ sb to do sth* inducir a alguén a facer algo/facer que alguén faga algo.

**(b) ref FRAS:** *to ~ os to do sth + inf* convencerse (a si mesmo) para + inf; ter/darse/reunir ánimos para + inf; resignarse a + inf; contentarse/conformarse con + inf.

•**bring about** *v tr* [v+o+adv] causar, ocasionar, producir, provocar; *change* efectuar *the Council brought about changes in the city's bus service* o concello efectuou cambios no servizo de autobuses da cidade.

•**bring along** *v tr* [v+o+adv] traer/levar consigo *bring your boyfriend along* trae o teu mozo.

•**bring away** *v tr* [v+o+adv] levar *Peter brought all the books away* Peter levou tódolos libros.

•**bring back** *v tr* [v+o+adv, v+adv+o] **1** traer de novo, traer de volta; *monarchy, republic, democracy* restaurar *people were against bringing back the death penalty* a xente estaba en contra de restaura-la pena de morte; *sth borrowed* volver, devolver, restituír *you should bring the book back to the library before next week* debes devolve-lo libro á biblioteca antes da semana que vén. **2** *memories* lembrar, traer á memoria *the sight of the old school brought back to him memories of childhood* a visión da vella escola tróuxolle lembranzas da nenez.

**PHRAS:** *~ to life* volver/devolver á vida.

Figura 14. Fragmento de la entrada *bring* del Diccionario Manual Inglés-Galego-Inglés.

Esta será la apariencia de una entrada tipo del *Diccionario Manual Inglés-Galego-Inglés* en la pantalla del ordenador o en una hoja impresa. Si la comparamos con la representación de la estructura que dábamos en el esquema anterior, nos daremos cuenta de que hay ciertas diferencias. Elementos de la estructura como “caracterización”, “apartados”, “acepción” o “texto” no figuran en el artículo más que

a través de alguno de sus componentes. El usuario final del diccionario podrá sin dificultades identificar inmediatamente y señalar elementos como “lema”, “fonética” o “fraseología”, etc., pero no otros elementos como “apartados”, ya que no son visibles; lo que son visibles son las diferentes realizaciones de este elemento, es decir, apartado (a), (b), etc. Lo mismo sucede con el elemento “acepciones”, del que vemos únicamente la acepción 1, 2, 3, etc. Quiere esto decir que aquello que para el usuario humano es una sucesión continua de datos o elementos sin más ordenación que la lineal, para el ordenador es una estructura lógica formalizada y jerarquizada con elementos y grupos de elementos bien definidos, a pesar de no ser éstos visibles para el ojo humano. Cuando consultamos el artículo *bring*, leemos primero el lema, luego cómo se pronuncia, luego que se trata de un verbo irregular, etc. Cuando un ordenador se enfrenta al mismo artículo, su análisis sería (simplificando): esto es una entrada; esto es el lema, su valor es *bring*; esto es la transcripción fonética, su valor es /brɪŋ/; esto es la caracterización, su valor es v *irr*.

Los diccionarios tradicionales, concebidos sólo para ser impresos, no tenían en cuenta, lógicamente, estas cuestiones. Su redacción era básicamente lineal, con lo que cada entrada es en realidad una sucesión de datos. Un diccionario electrónico, por el contrario, se nutre de una base de datos que responde a una estructura lógica perfectamente definida: los datos no se suceden unos a otros sino que ocupan una serie de espacios predefinidos que corresponden a los elementos de la entrada.

Las ventajas de los diccionarios electrónicos en este sentido son evidentes; por citar dos: sea cuál sea el orden en que se introduzca la información en la base de datos, el formato de salida, bien en hoja impresa, bien por pantalla, puede ser muy variado. Podemos decidir, por ejemplo, que las acepciones que un lema tiene como verbo vayan antes o después de las acepciones que tiene como sustantivo. Además, en cualquier momento se puede optar por aumentar información o eliminarla, ya sea dentro de un mismo elemento o bien introduciendo nuevos elementos o prescindiendo de otros. Por eso también es fácil generar varios diccionarios a partir de una misma base de datos: diccionario infantil, diccionario de bolsillo, etc., e incluso, echando mano de las especificaciones que figuran en el elemento “campo”, hacer vocabularios o glosarios especializados.

Ahora bien, ¿de qué forma se construye una base de datos para un diccionario electrónico y cómo se codifica la información? Uno de los objetivos de los diccionarios electrónicos y en general de todas las grandes proyectos relacionados con la informática y la explotación de fuentes lingüísticas mediante ordenador, es la independencia de los datos con respecto a los formatos (cf. Nathan y Austin 1992: 58), ya que así se facilita, como ya dijimos, la explotación de los datos, su extracción para otros proyectos, etc. El primer paso ya lo hemos dado, y no es otro que el análisis y estructuración de los tipos de datos, de las clases de información que queremos codificar. La propuesta de entrada que hicimos más arriba recoge ese análisis.

Pero además, la información, el trabajo lexicográfico, debe ser totalmente independiente de cualquier plataforma o programa informático. Nos interesa, por lo tanto, un sistema de codificación y almacenamiento de la información que nos permita conseguir esa independencia. Recurrimos entonces a los llamados estándares de marcación.

## 5. LOS ESTÁNDARES DE MARCACIÓN. LA SGML Y LA TEI

Los diccionarios definen estándar con los sinónimos patrón, modelo, referencia, norma, tipo. Es decir, un estándar es un ejemplo o modelo de alguna cosa que está comúnmente aceptado como referencia para medir, juzgar o describir la realidad o una parte de ella (cf. Magán Muñoz 1996: 153). Un estándar de marcación<sup>7</sup> es un sistema que sirve para codificar la información y que es usado por un grupo de personas relativamente amplio. Por ejemplo, ciertos programas de procesamiento de texto, utilizan códigos propios (es decir, un estándar propio) para codificar su formato, como por ejemplo:

.nf	
.ps42	(= tamaño 14 puntos)
.ffam BKO	(= tipo de letra Bookman Oldstyle)
.ft I	(= cursiva)
.ce	(= centrado)
La Lengua en el siglo XII	
para	
<i>La Lengua en el siglo XII</i>	

Figura 15.

Ciertos programas informáticos que nos permiten hacer, por ejemplo, concordancias (listados alfabéticos de las palabras de un texto situadas dentro de su contexto), requieren el uso de otros estándares para definir los parámetros que servirán para localizar las palabras en el texto. Así por ejemplo, el programa OCP (*Oxford Concordance Programme*, Programa de concordancias de Oxford), requiere el uso de un estándar de marcación denominado COCOA que codifica los diferentes parámetros entre ángulos (“<” y “>”):

<obra: Horario de clase> <autor: Christine Nöstlinger> <capítulo: 1>  
<título: Hora do bicho> <párrafo: 1> a madeira do pupitre en tempos fora de cor  
castaña escura. Agora só era de... <párrafo 13> Anika ergueuse, foi cara ó Norte

e... <capítulo: 2> <título: Hora do xantar> <párrafo: 1> a nai non se lucira nada coa comida... <párrafo> 14> no prato había unha masa pardacenta, desfeita, esmagada...

Las ventajas principales de trabajar con estándares de marcación son:

- ya están inventados, por lo cual no hay que invertir tiempo ni esfuerzo en su diseño.
- sus etiquetas (los nombres con que se describen los distintos elementos) suelen ser mnemotécnicas y, por lo tanto, fáciles de recordar y entender.
- suelen estar bien documentados en manuales y literatura científica, por lo que es fácil estudiar su uso y resolver dudas.
- existen programas de análisis que “entienden” estos sistemas, ya que en muchos casos los estándares fueron creados para usar con un determinado programa.

A la vista de este último ejemplo, podríamos decir que un estándar sirve también para describir la estructura (lógica) de un texto, es decir, las partes de las que se compone. Parece, por lo tanto una opción adecuada para nuestras necesidades. Sin embargo, COCOA y otros muchos estándares presentan varios problemas:

- a) la descripción que nos permiten hacer de un documento no es exhaustiva. Podemos definir sus partes, pero no su jerarquía. En el ejemplo que veíamos antes, sabemos que el texto tiene un autor, capítulos, títulos de capítulos y párrafos dentro de cada capítulo, pero, ¿es posible que un capítulo no tenga título? ¿puede haber párrafos antes del título de los capítulos (por ejemplo, citas)? Una descripción completa y exhaustiva de la estructura de los datos es el primer paso indispensable para conseguir la independencia con respecto al formato o al medio físico (Deen 1977: 7). Según la propuesta de entrada que hicimos más arriba, no es suficiente saber que hay un elemento “campo”, sino que debemos hacer constar cuál es su posición y relación jerárquica con los elementos “acp”, “acepciones”, “apartados”, etc.
- b) un segundo problema es la difusión de los estándares. Si pretendemos una verdadera independencia con respecto a los medios físicos e informáticos y el mayor grado de intercambio posible de información entre los distintos investigadores y usuarios, deberemos utilizar un estándar que tenga una amplia difusión, que sea la herramienta de un gran número de usuarios.
- c) incompatibilidad: por el mismo motivo debemos asegurarnos también de que el estándar escogido sea “entendido” (utilizado) por el mayor número posible de plataformas y programas informáticos. Los problemas de incompatibilidad podrían llegar a hacer que una investigación pase desapercibida para el resto de la comunidad científica.

Debido a la creciente importancia de los documentos en formato electrónico, los organismos oficiales que se ocupan de regular los estándares se preocuparon de buscar un estándar que sea verdaderamente universal e independiente de todas las constricciones de los medios físicos, del hardware y del software. Ese estándar es la denominada SGML o Lenguaje de Marcación Estándar Generalizada (*Standard Generalized Markup Language*).

Pero la SGML tal y como la conocemos hoy es el resultado de casi cuarenta años de investigaciones realizadas en un principio, y fundamentalmente, en los Estados Unidos, pero también en Inglaterra, Suecia, Canadá y otros países.

En los años 60 se comenzó a pensar en un sistema de codificación genérico y universal que utilizase etiquetas o marcas descriptivas en lugar de códigos de control que dependiesen de los medios informáticos. El objetivo era separar el contenido del documento de su formato. A finales de esta década, Charles Goldfarb, Edward Mosher y Raymond Lorie desarrollaron la llamada GML<sup>8</sup>, que aparte de ser las iniciales de sus apellidos significaba también *Generalized Markup Language* (Lenguaje de Marcación Generalizada). Este lenguaje, que se habría de convertir en la SGML, permitía la edición, formateado y transmisión de información independientemente de los sistemas informáticos. En 1978 el Instituto Estadounidense de Estándares (ANSI, *American National Standard Institute*) creó un comité especial para lenguajes de ordenador y procesamiento de textos en el que colaboró Charles Goldfarb. Este comité publicó tres años más tarde el primer borrador de la SGML, que era ya un estándar coherente, aceptable y completo. Un de los primeros usos prácticos de la SGML fue el proyecto de la Asociación de Editores de los Estados Unidos (AAP, *American Association of Publishers*) de desarrollar una aplicación que facilitase el intercambio de manuscritos entre autores y editores.

La SGML es, antes que nada, una forma de concebir y describir documentos mediante la descripción de su estructura independientemente de cualquier plataforma y, sobre todo, independientemente de cualquier formato. Para retomar un ejemplo que veíamos más arriba, para describir el título de un capítulo, los procesadores utilizan códigos como

.nf	
.ps 42	(= tamaño 42 puntos)
.ffam BKL	(= tipo de letra Bookman light)
.ft I	(= cursiva)
.ce	(= centrado)
La lengua en el siglo XII	

Figura 16.

que sólo recogen información sobre formato. La descripción equivalente en SGML sería:

<TI>La lengua en el siglo XII</TI>

Figura 17. *Marcación de un título en SGML.*

que refleja únicamente la información que es relevante desde el punto de vista de la estructura del texto, es decir, que esto es un título.

Cuando la SGML comenzó a tener difusión, fue un estándar más entre muchos otros. La necesidad de un único estándar que eliminase los problemas que suponía la existencia de varios sistemas, llevó a la creación de la TEI durante un congreso de la ACH (*Association for Computers and the Humanities*) en Nueva York en el año 1987. La TEI (Iniciativa para la Codificación de Textos, *Text Encoding Initiative*), es un proyecto internacional financiado por varios gobiernos e instituciones que tiene los siguientes objetivos:

- documentar los principales sistemas de codificación existentes y crear un metalenguaje que los describa y unifique.
- especificar un sistema de codificación único para los textos electrónicos.
- redactar una serie de recomendaciones para la creación de nuevos textos electrónicos: qué se debe codificar y cómo.

La TEI eligió la SGML para crear ese metalenguaje de descripción debido a su independencia con respecto a sistemas y plataformas. La SGML no es un sistema de codificación propiamente dicho, sino un entorno dentro del cual se pueden crear sistemas de codificación, un lenguaje con el cual se pueden crear sistemas de marcación. Al ser un lenguaje y no un estándar cerrado y acabado, tiene la ventaja de que el número de etiquetas o marcas posibles es infinito: cada usuario puede crear sus propias etiquetas para su uso específico con la única condición de que estén correctamente descritas.

La TEI publica periódicamente unas directrices<sup>9</sup> que están consideradas como recomendaciones de uso, no normas estrictas que haya que cumplir obligatoriamente. Se pretende codificar la mayor cantidad de información posible, pero sólo unas cuantas son obligatorias; la filosofía de la TEI es “si quieres codificar algo, hazlo de esta forma”. Las directrices se centran en la descripción de estructuras y parten del presupuesto de que todos los textos comparten unas características básicas a partir de las cuales se puede generar cualquier tipo de estructura textual.

Las directrices de la TEI describen las estructuras de los diferentes tipos de textos, que están compuestos por unidades denominadas elementos. De esta forma, cada una de las partes de nuestra propuesta de entrada bilingüe (lema, fonética, acepciones, etc.) sería un elemento. Lo que hace la TEI con la SGML no es nombrar todos los posibles elementos y establecer su contenido, sino definir y expresar las



relaciones que hay entre estos elementos; el nombre y el contenido de los elementos particulares dependerá de cada usuario.

Dentro de un texto marcado (es decir, descrito), todos los elementos deben tener una etiqueta o marca que los defina; las etiquetas se consignan entre ángulos (“<” y “>”). Veamos un ejemplo:

```
<obra> A guerra das linguas e as políticas lingüísticas </obra> <autor>
Louis-Jean Calvet </autor> <traductor> Xoán Manuel Garrido Vilariño
</traductor> <capítulo> <tit_cap> Nas orixes do conflito </tit_cap>
<sección> <tit_sec> A cuestión das orixes </tit_sec> <párrafo> <oración> Os
animais non falan porque non teñen nada que dicir... </oración> ... </párrafo> ...
</sección> ... </capítulo> </obra>
```

Como vemos, hay dos tipos de etiquetas, las de comienzo y las de final de elemento; las primeras tienen la forma ángulo + nombre + ángulo, e indican que se abre o comienza un nuevo elemento de la estructura del texto; las de final tienen la forma ángulo + barra + nombre + ángulo e indican el final del elemento en cuestión. Típicamente, entre cada par de etiquetas encontramos texto, pero los elementos pueden también estar vacíos. Imaginemos, por ejemplo, que unos capítulos tengan título y otros no; lo lógico será definir un elemento <tit\_cap> que unas veces tendrá contenido y otras no.

Con este ejemplo nos damos cuenta ya de que la SGML permite una descripción más exhaustiva que COCOA, por ejemplo, ya que podemos ver que <tit\_sec> aparece dentro de <sección> y esta dentro de <capítulo>. Pero esto no llega: ¿puede <tit\_sec> aparecer fuera de <sección>? ¿es posible encontrar la etiqueta <párrafo> dentro de <tit\_cap>? Las respuestas a estas y otras preguntas dependen del tipo de texto que estemos describiendo, es decir, de su jerarquía de constituyentes.

La jerarquía de los elementos de la estructura se describe en lo que la TEI llama una DTD. Una DTD o Declaración del Tipo de Documentos es una serie de declaraciones (es decir, afirmaciones) sobre un texto escritas según la sintaxis de la SGML. La DTD del ejemplo anterior, teniendo en cuenta ciertas simplificaciones por claridad, sería la siguiente:

<!ELEMENT	ensayo	-	-	(autor+, traductor*, capítulo+) >
<!ELEMENT	autor	-	-	(#PCDATA) >
<!ELEMENT	traductor	-	-	(#PCDATA) >
<!ELEMENT	capítulo	-	-	(tit_cap?, sección*) >
<!ELEMENT	tit_cap	-	0	(#PCDATA) >
<!ELEMENT	sección	-	-	(tit_sec?, párrafo+) >
<!ELEMENT	tit_sec	-	0	(#PCDATA) >
<!ELEMENT	párrafo	-	0	(oración+) >
<!ELEMENT	oración	-	0	(#PCDATA) >

Figura 18. Ejemplo de DTD.

Las *declaraciones de elementos* se consignan entre ángulos y comienzan con una exclamación para distinguirlas de las etiquetas. Constan del *identificador genérico* o nombre del elemento (ensayo, autor, capítulo, etc.) seguido de las llamadas *reglas de minimización*, que nos indican la necesidad o no de especificar la etiqueta de comienzo o la de final: dos guiones indican que las dos etiquetas son obligatorias (caso de <capítulo>, p. ej.), un guión y un cero indican que la primera es obligatoria y la segunda opcional (caso de <oración>, p. ej.). Después de las reglas de minimización tenemos el *modelo de contenido*: un elemento puede contener texto (#PCDATA —*parsed character data*—, caso de oración o párrafo) o puede a su vez contener un elemento o más (caso de ensayo o sección). En este último caso, los *indicadores de aparición* nos indican cuántos elementos puede haber dentro de otro: el signo de interrogación “?” indica que puede haber uno o ninguno de los elementos especificados (<tit\_cap> dentro de <capítulo>); el signo más “+” indica que puede haber uno o más elementos (<párrafo> dentro de <sección>); finalmente, el asterisco “\*” indica que puede haber cero, uno o más de uno de los elementos en cuestión (<sección> dentro de <capítulo>). Por último, los *conectores de grupos*, indican el orden en que los elementos aparecen: la coma “,” indica que aparecen en el orden especificado (<tit\_cap> y <sección> dentro de <capítulo>); el ampersán “&” indica que los elementos pueden aparecer en cualquier orden; la barra vertical “|” indica que o bien aparece un elemento o bien aparece el otro.

Por lo tanto, la DTD anterior se debe leer del siguiente modo: un texto del tipo ensayo consta de por lo menos un autor, uno o más traductores o ninguno y por lo menos un capítulo; el capítulo puede tener título o no y puede tener una o más secciones o ninguna; la sección puede tener título o no y debe tener por lo menos, un párrafo; el párrafo debe tener, por lo menos, una oración. El resto de los elementos contiene texto. Todos los elementos tienen que aparecer en este orden necesariamente.

La DTD debe acompañar siempre al texto o textos a los que se refiere, ya que una de sus funciones es facilitar el intercambio de datos. Cuando marcamos un texto para nuestra investigación, creamos las etiquetas que más nos convienen y le ponemos el nombre que nos parezca mejor. Si después queremos que otro investigador utilice nuestro texto, sólo tenemos que entregarle con él la DTD correspondiente: los programas que utilice la otra persona la entenderán sin problemas y podrán interpretar la estructura del texto. En resumen, la DTD nos anticipa lo que vamos a encontrar cuando nos enfrentemos al texto.

Estos son, a grandes rasgos, los rudimentos de la SGML. Como vemos, este lenguaje de marcación o, mejor dicho, metalenguaje, posee una sintaxis muy sencilla y con pocos componentes, pero de grandes posibilidades descriptivas. Pero, por supuesto, la SGML no es sólo esto; podemos encontrar una descripción más detallada de este metalenguaje y de sus elementos en el ya citado documento provisional TEI P3, Sperberg McQueen y Burnard (1994), y en Hockey et al. (1994).

## 6. PROPUESTA DE DTD PARA UN DICCIONARIO BILINGÜE INGLÉS-GALLEGO-INGLÉS

La siguiente propuesta es una versión simplificada de la DTD del Diccionario Manual Inglés-Gallego-Inglés. Por operatividad y para facilitar la exposición suprimimos ciertos elementos, pero mantenemos los elementos básicos de una entrada tipo de la sección inglés-gallego; omitimos todos los detalles referentes a la sección gallego-inglés del diccionario, al prólogo, apéndices, etc. Omitimos también todos aquellos elementos (como atributos, entidades, etc.) a los que no se hizo referencia en la pequeña introducción a la SGML del apartado anterior. A pesar de todas esas omisiones, la DTD es perfectamente válida y respeta la sintaxis de la SGML y las recomendaciones de la TEI.

<!--DTD DO DICCIONARIO MANUAL INGLÉS - GALEGO - INGLÉS-->

<!--versión reducida para demostración-->

<!ELEMENT **diccionario** - - (**prologo**, **corpo**) —Diccionario Manual Inglés-Galego-Inglés-->

<!ELEMENT **p** - - (#PCDATA) —Parágrafos de texto-->

<!ELEMENT **prologo** - - (**p**+) —Indicaciones de uso, etc-->

<!ELEMENT **corpo** - - (**seccionIG**, **seccionGI**) —A parte central do diccionario: prólogo, sección de inglés-galego e de galego-inglés-->

<!ELEMENT **seccionIG** - - (**entradaIG**+) —Entradas Inglés-Galego-->

<!ELEMENT **seccionGI** - - (**entradaGI**+) —Entradas Galego-Inglés-->

<!--AQUI COMEZA A DEFINICIÓN DAS ENTRADAS-->

<!ELEMENT **entradaIG** - - (**lema**, **fonetica**, **apartados**\*) — Entrada da sección Inglés - Galego-->

<!ELEMENT **lema** - - (#PCDATA) — Cada unha das palabras en inglés que constitúe unha entrada independente-->

<!ELEMENT **fonetica** - - (#PCDATA) —Transcripción do lema segundo as normas da Asociación Fonética Internacional-->

<!ELEMENT **apartados** - - (**caract**, **acep**+, **fraseoloxia**) —Cada un dos apartados concretos-->

<!ELEMENT **caract** - - (#PCDATA) —Caracterización na que se basea a en <apartados>-->

<!ELEMENT **acep** - - (**num**, **campo**\*, **texto**+, **exemplo**\*) —Cada unha das definicións concretas-->

<b>&lt;!ELEMENT</b>	<b>num - - (#PCDATA)</b> —Numeración correlativa en números arábigos das diferentes def—>
<b>&lt;!ELEMENT</b>	<b>campo - - (#PCDATA)</b> —Campo do saber e indicacións de variantes—>
<b>&lt;!ELEMENT</b>	<b>texto - - ((pallngles*, gal+)+)</b> —As acepcións propiamente ditas—>
<b>&lt;!ELEMENT</b>	<b>pallngles - - (#PCDATA)</b> —Para centra-lo significado—>
<b>&lt;!ELEMENT</b>	<b>gal - - ((equivalencia, caract.gal) &amp; #PCDATA)</b> —Equivalencias en galego do lema inglés—>
<b>&lt;!ELEMENT</b>	<b>equivalencia - - (#PCDATA)</b> —Cada unha das equivalencias galegas concretas—>
<b>&lt;!ELEMENT</b>	<b>caract.gal - - (#PCDATA)</b> —Caracterización da equivalencia en galego—>
<b>&lt;!ELEMENT</b>	<b>exemplo - - (ingles*, galego*)</b> —Exemplificación da definición entrada—>
<b>&lt;!ELEMENT</b>	<b>ingles - - (#PCDATA)</b> —Texto do exemplo en inglés—>
<b>&lt;!ELEMENT</b>	<b>galego - - (#PCDATA)</b> —Texto do exemplo en galego, que é traducción do exemplo en inglés—>
<b>&lt;!ELEMENT</b>	<b>fraseoloxia - - ((ingles, galego)+)</b> —Apartado de frases e expresións feitas nas que aparece o lema—>

Figura 19. DTD del Diccionario Manual Inglés-Galego-Inglés

Una vez diseñado el modelo de entrada para nuestro diccionario y codificado en la DTD correspondiente (figura 19) estamos preparados para empezar a redactar los artículos lexicográficos. Para ello podemos optar por dos vías. En primer lugar, podremos utilizar un procesador de textos para introducir manualmente (o con ayuda de macros) la información lexicográfica junto con las etiquetas SGML correspondientes. Habrá que tener en cuenta que el archivo o archivos resultantes tendrán que guardarse como archivos .txt o sólo texto para que puedan ser leídos por los programas necesarios para verificar la marcación (*parsers* como SGMLS), bases de datos con las que vayamos a manipular la información, etc.

Los dos únicos (y no pequeños) inconvenientes que presenta el trabajo con procesadores son: a) cuando se hace la conversión a sólo texto, determinados procesadores —como Word para Windows— introducen caracteres que no pertenecen a la tabla básica ASCII, lo que con seguridad provocará interferencias con programas para DOS que no los reconozcan; y b) trabajar con una DTD relativamente larga como la que proponemos puede resultar bastante incómodo: aunque utilicemos macros que nos ahorren mucho trabajo de teclado, tendremos que tener constantemente presente en nuestra memoria en qué lugar exacto de la DTD estamos y saber qué etiquetas están admitidas en él y cuáles no.

Para solventar estos dos problemas recurriremos a un etiquetador, como por ejemplo el *Oxford SGML Tagger* (cf. Light 1994). Se trata de un programa residente para DOS que lee e interpreta una DTD y puede ser utilizado después con un procesador para introducir semiautomáticamente las etiquetas que precisemos. Este etiquetador construye a partir de nuestra DTD un fichero de ayuda que nos recuerda que etiqueta(s) podemos utilizar en cada momento. Si la DTD contiene descripciones sobre las etiquetas (vid. en la figura 19 las indicaciones entre guiones dobles), el etiquetador genera también con ellas un fichero de ayuda que puede ser llamado en cualquier momento para comprobar de que tipo de entrada se trata, su contenido, etc.

Para terminar esta exposición, ofrecemos una entrada del *Diccionario Manual Inglés-Galego-Inglés* marcada en SGML con el etiquetador de Oxford y lista para su inclusión en cuerpo del diccionario electrónico. Hemos resaltado en negrita la información lexicográfica para diferenciarla de las etiquetas SGML.

```
<ENTRADA TIPO = "principal" STATUS = "final" CLAVE = "biscuit">
<LEMA>biscuit
<FONETICA>/'bIskIt/
<APARTAD><APART ID = "a"><CHARACT><SUBS><NOME>(a)n.
<DEFINICI><DEF ID="a1"><NUM>1<CAMPO>(Brit)<TEXTO>
<GAL><EQUIVALE>galleta<CHARACTGA><XENERO>f.</GAL>
<DEF ID="a2"><NUM>2<CAMPO>(US)<TEXTO><GAL>
<EQUIVALE>biscoito<CHARACTGA><XENERO>m.,</GAL>
<GAL><EQUIVALE>bica<CHARACTGA><XENERO>f.</GAL>
<DEF ID = "a3"><NUM>3<TEXTO><GAL><EQUIVALE>panciño
<CHARACTGA><XENERO>m.tostado,</GAL>
<GAL><EQUIVALE>torrada<CHARACTGA><XENERO>f.,</GAL>
<GAL><EQUIVALE>tostada<CHARACTGA><XENERO>f.</GAL>
<DEF ID="a4"><NUM>4<TEXTO><GAL>cor parda clara.</GAL>
<DEF ID="a5"><NUM>5<CAMPO>(Ceram)<TEXTO><GAL>
<EQUIVALE>biscuit<CHARACTGA><XENERO>m,</GAL><GAL>
<EQUIVALE>biscoito<CHARACTGA><XENERO>m.</GAL>
<DEF ID="a6"><NUM>6<TEXTO><GAL><EQUIVALE>disco.</GAL>
<APART ID = "b"><CHARACT><ADXX>(b)adj.<DEFINICI>
<DEF ID="b1"><NUM><TEXTO><GAL><EQUIVALE>de cor parda clara.</GAL>
```

Figura 20. *Entrada marcada del Diccionario Manual Inglés-Galego-Inglés.*

## NOTAS

1. Las investigaciones aquí recogidas han sido posibles gracias a una ayuda de la Consellería de Educación e Ordenación Universitaria de la Xunta de Galicia (XUGA20402A93).
2. Entendemos texto como cualquier acto comunicativo, tanto oral como escrito.

3. Por ejemplo, el *Oxford English Dictionary on CDROM* o el *Diccionario de la Lengua Española* de la Real Academia Española. De hecho, los primeros diccionarios electrónicos se crearon aprovechando los diccionarios impresos, para lo cual se diseñaron estrategias y programas como los que se describen en Byrd et al. (1997).
4. Cf. Boguraev (1991), donde se exponen diversos proyectos de construcción de *léxica* a partir de diccionarios bilingües como el OED o el *Longman Dictionary of Contemporary English*.
5. De hecho, estas relaciones no se establecen sólo entre artículos, sino también entre los artículos y el motor de búsqueda que nos permite localizar una palabra dentro del diccionario.
6. Boguraev (1991:231) habla de estructura (*structure*) en vez de formato, pero parece más correcto reservar el término estructura para la ordenación lógica de la información y utilizar formato para su aspecto físico.
7. Para una descripción de diferentes estándares de marcación de información textual y multimedia, cf. Álvarez Lugrís (1996a:§ 2.4) y Magán Muñoz (1996).
8. Esta breve historia está tomada de SGML User's Group (1985).
9. La tercera versión, corregida y aumentada, es de 1994. Cf. Sperberg-McQueen y Burnard (eds.) (1994).

## REFERENCIAS BIBLIOGRÁFICAS

- Álvarez Lugrís, A. 1996a. "Informática e análise textual". *Lingüística e informática*. Eds. Gómez Guinovart, J. y A. M. Lorenzo Suárez. Santiago de Compostela: Tórculo Edicións. 87-152.
- Álvarez Lugrís, A. 1996b. "Diccionario Oxford Español-Inglés Inglés-Español". *Viceversa. Revista galega de traducción* 2: 239-248.
- Boguraev, B.K. 1991. "Building a Lexicon: The Contribution of Computers". *International Journal of Lexicography* 4 (3): 227-260.
- Bruker, H., et al. 1867. *Wörterbuch der Handels- und Geschäftssprache. Deutsch, holländisch, englisch, französisch, italienisch, spanisch*. Stuttgart: Verlag von Julius Meier.
- Byrd, R.J., N. Calzolari, M.S. Chodorov, J.L. Klavans y M.S. Neff. 1987. "Tools and Methods for Computational Lexicology". *Computational Linguistics* 13 (3-4): 219-240.
- Corriente Córdoba, F. 1988. *Nuevo Diccionario Español-Árabe. Camus 'arabi 'asbani*. Barcelona: Herder.
- Deen, S. 1977. *Fundamentals of Database Systems*. London: McMillan.
- Feixó Cid, X.G., coord. 1986. *Diccionario da lingua galega*. Vigo: Ir Indo Edicións.
- Gómez Guinovart, J. y A. M. Lorenzo Suárez, eds. 1996. *Lingüística e informática*. Santiago de Compostela: Tórculo Edicións.
- Haensch, G., L. Wolf, S. Ettinger y R. Werner. 1982. *La lexicografía. De la lingüística teórica a la lexicografía práctica*. Biblioteca Románica Hispánica, Gredos: Madrid.
- Hockey, S., G. Murphy, L. Horowitz y C. Bohlen, eds. 1994. *Electronic Texts in the Humanities: Methods and Tools. Course Handbook*. Center for Electronic Texts in the Humanities: Princeton.

- Iglesias Rábade, L., coord. En preparación. *Diccionario Manual Inglés-Galego-Inglés*.
- Light, R. 1994. *The Oxford SGML Tagger*. Oxford: Oxford Computing Services.
- Magán Muñoz, F. 1996. "Estándares de representación da información textual e multimedia". *Lingüística e informática*. Eds. Gómez Guinovart, J. y A. M. Lorenzo Suárez. Santiago de Compostela: Tórculo Edicións. 153-186.
- Nathan, D., y P. Austin. 1992. "Finderlist, Computer-Generated, for Bilingual Dictionaries". *International Journal of Lexicography* 5 (1): 32-65.
- Navaza Blanco, G. y X. Lastra Muruais. 1993. *Diccionario Xerais da Lingua*. Cuarta edición. Vigo: Edicións Xerais de Galicia.
- SGML User's Group. 1985. "Standard Generalized Markup Language (International Standard ISO 8879). A Brief History".
- Sinclair, J., ed. 1987. *Collins COBUILD English Language Dictionary*. London y Glasgow: Collins.
- Slaby, R., R. Grossman y C. Illig. 1991. *Diccionario de las lenguas española y alemana*. Cuarta edición. Barcelona: Editorial Herder.
- Sperberg-McQueen, C.M. y L. Burnard, eds. 1994. *Guidelines for Electronic Text Encoding and Interchange*. Draft version 3. TEI document P3. Chicago & Oxford: ACH-ACL-ALLC.
- Velázquez de la Cadena, M. 1861. *A Pronouncing Dictionary of the Spanish and English Languages. In Two Parts*. Segunda edición. Cádiz: Imprenta de la Revista Médica.
- VV.AA. 1994. *Diccionario Oxford Español-Inglés Inglés-Español*. Oxford y Nueva York: Oxford University Press.
- VV.AA. 1994. *OED2 on CDROM*. Versión 1.10. Oxford: Oxford University Press.
- Walker, J. 1852. *A Critical Pronouncing Dictionary and Expositor of the English Language*. London: William Tegg and Co.