

El Diccionario de Euskaltzaindia

Este mismo año, La Real Academia de la Lengua Vasca – Euskaltzaindia publicará su diccionario bajo el título *Euskaltzaindiaren Hiztegia. Adierak eta adibideak (Diccionario de Euskaltzaindia. Significados y ejemplos.)*. Esa obra, de cuyas características les hablaré más tarde, es el último eslabón, por ahora, del proceso de normalización y normativización de la lengua vasca.

La normativización del euskera, a pesar de ser el gran objetivo de la Academia desde su nacimiento en 1918, arranca definitivamente en 1968, en el Congreso celebrado en Aránzazu, con ocasión del quincuagésimo aniversario de la institución.

En ese Congreso, el profesor Koldo Mitxelena presentó una ponencia, que fue la base de todo el desarrollo posterior del euskera unificado. En dicha ponencia estableció una serie de prioridades, en cuanto a las áreas que se debían cultivar, para alcanzar con éxito el objetivo propuesto:

1. Ortografía
2. Fijación de la forma de las palabras antiguas de la lengua
3. Reglas de formación de palabras y adaptación de préstamos
4. Morfología del nombre y del verbo
5. Sintaxis

Las dos primeras áreas quedaron establecidas de forma básica en la propia ponencia de Mitxelena, que poco después enriqueció, proponiendo una lista de cerca de 2.000 formas del acervo léxico.

Ese mismo año de 1968 propuso los criterios que se debían tener en cuenta para elaborar un diccionario en el que las palabras eusquéricas estuvieran descritas en euskera:

1. El diccionario debía ser general, es decir debía contener palabras de todos los dialectos y de todas las épocas.
2. Debía ser una guía para hablantes, escritores y lectores.
3. Se debían fijar los límites en cuanto a qué variantes registrar y cuáles no, qué hacer con los arcaísmos, los préstamos y su adaptación, qué hacer con los neologismos.
4. El diccionario debía ser selectivo, para ayudar al usuario del mismo a distinguir lo bueno de lo no tan bueno.
5. Había que realizar una recopilación sistemática del léxico utilizado en la literatura vasca a lo largo de la historia.

Este impulso inicial no llevó a la Academia inmediatamente a la elaboración de su diccionario. En cambio, fuera de la Academia se fueron poniendo en marcha proyectos que dieron lugar a diversos diccionarios. Así en 1970 se publica un manual titulado *Batasunaren Kutxa*, cuyos autores eran Gabriel Aresti y Xabier Kintana, gozaba del apoyo de numerosos académicos, tanto de número como correspondientes y se ajustaba a los criterios establecidos por Mitxelena dos años antes. Sus autores se mostraban dispuestos a hacer todas las correcciones que fueran oportunas. En esta obra se incluía un diccionario en el que todas las definiciones y demás informaciones estaban en euskera.

En 1974 la Academia publicó un nomenclator comercial, *Merkatalgo izendegia*, destinado sobre todo a la rotulación de los establecimientos comerciales, que venía a responder a una demanda social creciente. Este nomenclator tuvo una edición aumentada en 1982. Por lo demás Euskaltzaindia se limitó a elaborar algunos pequeños vocabularios: uno de Filatelia y otro de Arquitectura.

Esos años son de una actividad muy intensa, en el campo cultural en general y en el de la enseñanza en particular, con el nacimiento de muchas ikastolas. El desarrollo del vocabulario para la expresión de muchos conceptos no usados en vasco hasta entonces resulta ser una necesidad imperiosa. No hay tiempo que perder ni, lamentablemente, el sosiego ni los medios materiales necesarios para hacer un análisis del léxico vasco, tal como el que Koldo Mitxelena deseaba, y que como un titán trataba de llevar a cabo en solitario.

La Academia, presionada por la comunidad educativa, se ve obligada a publicar en 1975 un vocabulario titulado *Zortzi urte arteko ikastola hiztegia* (*Vocabulario para las ikastolas hasta los ocho años*). Tal era la necesidad, que la Academia prefirió cubrirla cuanto antes en vez de esperar a tener una obra de otro nivel más elevado.

Otras instituciones surgidas en esa década de los 70 del siglo XX, tales como Elhuyar y UZEI van respondiendo a esa necesidad de modernizar el léxico, publicando diccionarios de diversas especialidades, siendo fieles a las normas de la Academia, pero adelantándose a ella, mediante estructuras organizativas de gran eficacia.

En el campo de los diccionarios de equivalencias vasco-español/español-vasco cabe destacar Hiztegi Teknikoa de L. M. Mujika y sobre todo Euskal Hiztegi Modernoa (1977) de X. Kintana y otros. Este último diccionario tuvo muchos seguidores a lo largo de las décadas de 1980 y 1990, constituyéndose en una especie de norma *de facto*.

Se está "poniendo el carro delante de los bueyes", en expresión de Mitxelena. Se está generando nuevo léxico sin conocer suficientemente lo que ha sido el léxico vasco a lo largo de la historia. La situación crea inquietud y polémica. La Academia, preocupada, celebra unas jornadas de reflexión en diciembre de 1981 para fijar la estrategia que va a desarrollar en cuanto al léxico.

La estrategia queda definida en los siguientes términos:

- a) El Diccionario General y el moderno son parte de los quehaceres de la Academia.
- b) La elaboración de diccionarios de especialidad no entra dentro de su ámbito de actuación.
- c) Euskaltzaindia ve con buenos ojos que cualesquiera otros los elaboren, ya sea individualmente o de forma colectiva.
- d) Euskaltzaindia ve la necesidad urgente de crear una Comisión de Lexicografía.

El Diccionario General era precisamente el sueño aún incumplido de Mitxelena. En cuanto al diccionario moderno, el propio Mitxelena advirtió a

la Academia, en un informe de 1982, de que no se podía dar un solo paso sin antes fijar la grafía de los préstamos. La urgencia que tenía la fijación de criterios para adaptar los préstamos fue planteada a la Academia también desde otras instancias, como la Asociación Elhuyar y la Consejería de Educación del Gobierno Vasco. Inmediatamente se nombró una comisión para que propusiera un plan de acción con esa finalidad. A raíz de esa decisión se celebraron dos jornadas en 1982, en las que se identificaron los problemas.

En febrero de 1982 la Academia empieza a dar los primeros pasos para crear esa Comisión de Lexicografía y la nombra en 1983, encomendándole los siguientes cometidos:

1. El Diccionario General o diccionario de Mitxelena
2. Otros diccionarios de la Academia
3. La fijación de criterios para la adaptación de préstamos
4. La investigación acerca de la normativización del léxico

En 1984 llega por fin la posibilidad de elaborar el Diccionario General Vasco, gracias a que el Gobierno Vasco y las Diputaciones Vascas deciden apoyar económicamente el proyecto que había estado esperando un cuarto de siglo.

Lamentablemente, el profesor Mitxelena, no pudo ver publicado ni tan siquiera el primer tomo de la obra, puesto que falleció en 1987 y el primer volumen vio la luz poco después. El equipo de trabajo continuó su labor bajo la dirección del académico Ibon Sarasola.

Ese mismo año de 1984 la Comisión de Lexicografía propuso un programa de trabajo para el período 1984-1996, en el que se preveían dos fases:

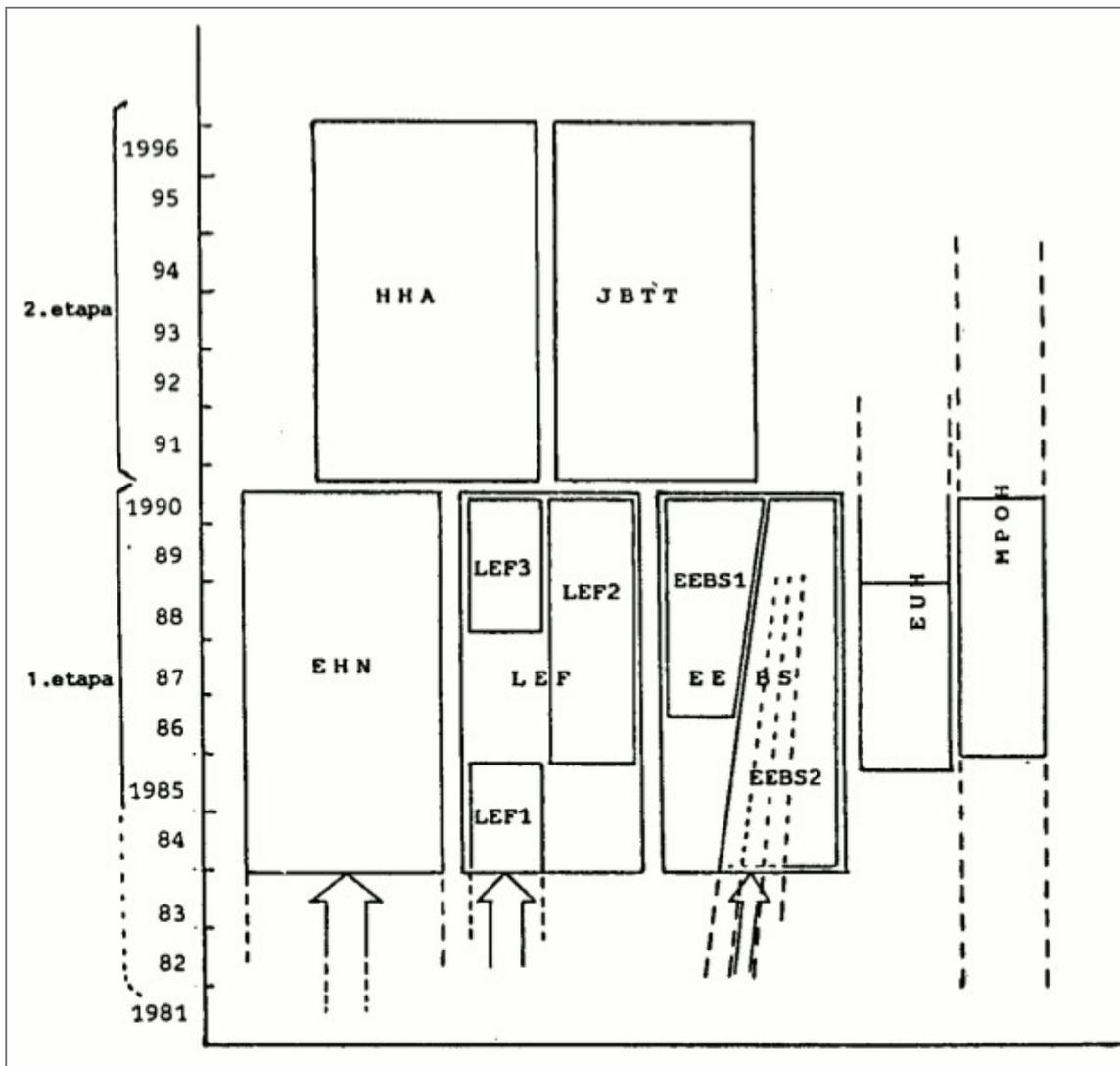
1ª fase 1984-1990 con tres objetivos:

1. La elaboración del Diccionario General Vasco
2. La fijación de los criterios lexicológicos
3. Realizar la recopilación sistemática del léxico actual

2ª fase 1990-1996 con dos objetivos:

1. Elaborar el diccionario normativo moderno
2. Crear terminologías técnicas de distintas disciplinas

En el informe que presentó el programa se incluyó el esquema adjunto, en el que se pueden apreciar los tres pilares básicos: el Diccionario General Vasco (EHN), la fijación de los criterios lexicológicos (LEF) y la recopilación sistemática del léxico actual (EEBS). En la parte superior, sustentados por los tres pilares anteriores aparecen el diccionario moderno (HHA) y las terminologías técnicas (JBTT).



Este plan, como veremos, no se cumplió, ni mucho menos, en cuanto a plazos, pero fue el primer planteamiento estratégico coherente, y marcó muy positivamente el futuro de la labor lexicográfica de la Academia.

La filosofía que subyacía tras este plan era la siguiente: el Diccionario General Vasco proporcionará una visión de lo que ha sido el léxico vasco desde la antigüedad hasta el siglo XX, la recopilación sistemática del léxico actual ofrecerá la visión del uso moderno, la fijación de los criterios lexicológicos permitirá hacer un uso apropiado de las reglas de formación de palabras para crear nuevos vocablos, los criterios de adaptación de préstamos servirán para el mismo fin en su caso, y, una vez hecho todo esto, se estará en condiciones de elaborar el diccionario normativo moderno y de crear las terminologías técnicas. Todo eso debía ocurrir de modo secuencial en dos fases.

En realidad, el Diccionario General Vasco terminó de publicarse en formato papel en 2005. Consta de 16 tomos con un total de 14.000 páginas. Se basa en un corpus informatizado de 6 millones de palabras texto, que acoge de forma bastante exhaustiva las obras de los autores clásicos y abarca hasta 1970 aproximadamente, pero a medida que se acerca a este año resulta ser más selectivo. Ese corpus no fue etiquetado, ni anotado lingüísticamente, porque en el tiempo en el que se emprendió el proyecto no se disponía de herramientas para ello.

Hoy en día se puede consultar en el sitio web de la Academia (<http://www.euskaltzaindia.net/oe>) y se puede descargar en formato PDF.

Un fragmento del mismo (parte de la entrada *zahi*) nos puede dar una idea de su contenido:

lema

Tradición lexicográfica

zahi (V, G, AN, L, BN, Ae, S, R; Ht *VocGr* 425, Arch *VocGr*, *VocBN*, Gèze, Dv, H; *zai* Lcc, Lar, Añ, Izt 50r, Dv, H, Zam *Voc*), *zagi* (L, B, Sal; *VocB*), *sahi* (V-gip; SP; *sai* Mic 9r, Dv (G)), *sei* (V-arr). Ref.: Bon-Ond 151; A (*zai*, *zagi*); Lh (*zagi*); Lrq; Arin *AEF* 1960, 69; Satr *VocP*; Iz *Utz*, *Als* 84, R 295; Etxba *Eib(saixa)*; *CEEN* 1970, 349; Ibarra *Dima (seye)*; Elexp *Bèrg*; Izeta *BHizt2 (zagi)*.

Acepciones

1. Salvado. "Soma y salvado, *saia*" Mic 9r. "Afrecho, salvado" Lar. "Allegadores de la ceniza y derramadores de la harina, *zaietan zur ta iriñetan ero*" Ib. "*Zaietan ziogi, irunetan ero*, mezquino en salvado y pródigo en harina" Izt 50r. "*Ogi arinak zagi aunitz (B)*" A. "*Zaietan urun, urunetan zaikoak gara gu (V-ger)*" Ib. "*Zaiatz uruna egin ezinik*, sin poder conseguir a fuerza de regalitos [...] (V-ple)" A *EY* III 334. "El salvado mayor" Iz *Utz*. Cf. *Zai-Zaku*, apodo en Urruz *Urz* 61. *Gure zaiok iriñ eske*. RIs 46. *Zaietan zur eta iriñetan ero*. Ib. 19. *Florea, irina, hoberena, gatzetasuna, eramaiten du etsai gaixtoak, eta gero nahi duzu kontenta dadin lainkoa zahiaz eta zahartzeko hondarratz*. Ax 182 (V 123). *Ez horak zahia jan, ez oilloer ützi*. Bela 22 (tbn., con alguna variación, en Saug 25 y O *Pr* 248). *Hobe da zahi hutsa ezi ez aho hutsa*. O *Pr* 246. *Lastasakuba baininzan ala, / kendu jeuste zaiak*. Acto 295. *Zai apur bat galdu ez dedin bildurrez, ta urun ederra galduarren ardurarik ez*. Mg *CO* 20. *Eioten da, eralgi, alderatu zaia*. Añ *LoraS* 55. *Irten zaite bada errota orretatik, bestela, zai, ta iriñ usuraz kargaturik ondatuko zera su eta garren lize izugarrian*. AA II 190. *Ogijari karia, zaija edo beste gauzarik ezarten deutsenak*. Astar II 165. *Bein nagusiena egin ezkerro, eta esan daroagun legez, lanari zaia kendu ezkerro, erraz izaten da zerbait geiago aziten*. Zav *Fab RIEV* 1907, 91. *Irina behar eta pagatzen du zahiz*. AstLas 74.

Ejemplos de uso extraídos del corpus

Se trata de un diccionario descriptivo. Se puede observar que la palabra objeto de la entrada aparece escrita de varias formas en los ejemplos: *zai*, *zagi*, *zahi*

El segundo pilar del programa, el de la fijación de los criterios lexicológicos, lo desarrolló un grupo de trabajo (LEF: Lexikologi Erizpideen Finkapena) que comenzó su andadura en 1986. El plenario de este grupo lo componían 23 personas, entre académicos de número y correspondientes, y otras cinco personas formaban el comité de gestión.

Este grupo de trabajo, se ocupó primeramente de fijar los criterios de adopción de préstamos y después dedicó sus esfuerzos al estudio de los distintos modos de formación de palabras (composición y derivación), desde 1986 hasta 1994. Como consecuencia de ello, el Pleno de la Academia fue aprobando en ese tiempo diversas normas al respecto.

La recopilación del léxico del siglo XX (EEBS: Egungo Euskararen Bilketa-lan Sistematikoa) fue un proyecto que no desarrolló la propia Academia, ya que en 1987 decidió no crear un equipo de trabajo específico para el mismo, sino que fuese alguna otra entidad la que lo llevase a cabo. La entidad adjudicataria fue UZEI, que suscribió un convenio con el Gobierno Vasco para redactar y ejecutar el proyecto, con financiación del Gobierno Vasco y tutela académica de Euskaltzaindia.

El objetivo de ese proyecto fue la creación de un corpus de textos en euskera publicados en el siglo XX. Lo primero que se hizo fue un inventario completo de todo lo publicado hasta entonces y su clasificación según cuatro criterios: el período en el que se publicó el texto, el dialecto utilizado, el tipo de texto y el tamaño del mismo.

Epea		Euskakia		Testu-mota	
1	1900-1939	1	Bizkaiera	01	Egunkarietako idazkiak
2	1940-1958	2	Gipuzkera	02	Hamabostekarietako idazkiak
3	1960-1987	3	Zuberera	03	Saio-aldizkarietakoak
		4	Lapurtera/nafar.	04	Administrazio-idazkiak
		5	Euskara batua	05	Ikasliburuak
		6	Bestelakoak	06	Salakera
		7	Nahasiak	07	Literatur prosa
				08	Poesia
				09	Antzerkia
				10	Bertsoak
				11	Ikerketa-idazkiak
				12	Haur- eta gazte-literatura
				13	Ahozkoak (transkribapenak)
				14	Liturgi idazkiak

Habida cuenta de que el volumen de los textos publicados en ese período era excesivamente grande para ser abarcado completamente (mas de 200 millones de palabras texto), se elaboró un modelo estadístico, con el objeto de seleccionar una muestra representativa de tamaño razonable para ser procesada (2 millones de palabras texto).

Se optó por un muestreo estratificado y se asignó una ponderación a los distintos tipos de texto, para finalmente seleccionar los textos que pasarían a formar parte del corpus, a través de un sorteo en dos etapas (primero la obra y después el fragmento).

Los textos fueron en una gran parte digitalizados mediante escáner, procesamiento OCR y corrección manual. Los que por su conservación no permitían utilizar ese procedimiento se introdujeron manualmente. Todo ello se ejecutó a lo largo de 1989, y a lo largo de los dos años siguientes se lematizó por completo manualmente. En ese momento el corpus abarcaba lógicamente, el siglo XX solo hasta 1987. El proyecto tuvo luego continuidad y, siguiendo el mismo modelo estadístico fue ampliado hasta abarcar todo el siglo XX.

Actualmente el corpus consta de 4.655.300 palabras texto y se puede consultar en el sitio web de la Academia.

(http://www.euskaracorpora.net/XXmendea/Konts_arrunta_fr.html)

Cuando llegó el año 1991, el que según el plan de 1986, debía ser el del inicio de la segunda fase, la Comisión de Lexicografía propuso al pleno de la Academia un plan de acción para el siguiente lustro. Consideraba que con el corpus del Diccionario General y con el del siglo XX se disponía ya de información suficiente para elaborar un diccionario ortográfico normativo. El Pleno de la Academia adoptó los siguientes acuerdos:

1. Crear la comisión del diccionario ortográfico (Hiztegi Batua)
2. Crear la base de datos lexicográfica del euskera, aunando el corpus del Diccionario General y el del siglo XX
3. Poner en marcha la elaboración del diccionario ortográfico para tenerlo a punto al cabo de cinco años.

La comisión del diccionario ortográfico preparó una lista de palabras en la que incluyó las que habían sido previamente aprobadas por la Academia desde 1968, una serie de voces de uso tradicional bien establecido, de acuerdo con el testimonio del corpus del Diccionario General y las voces que en el corpus del siglo XX tenían una frecuencia de más de 20 apariciones.

La lista constaba de unas 20.000 formas. Cada una de esas formas era analizada por un grupo de trabajo en UZEI, para documentar las características de su uso (forma, frecuencia, permanencia o no a lo largo del tiempo, dialectos en los que ha sido usada, autores que la han usado, categoría gramatical, etc.). Estos informes eran enviados a la comisión del diccionario ortográfico, para que propusiera la forma estándar de cada una de las voces. Las propuestas se dirigían a una red de consulta, formada por usuarios cualificados (escritores, traductores, profesores, etc.), los cuales hacían observaciones, que eran tenidas en cuenta por la comisión, antes de enviar las propuestas definitivas al pleno de la Academia. Los miembros del pleno tenían a su vez ocasión de hacer sus propias observaciones y de defenderlas en el pleno, antes de la aprobación definitiva y de la publicación con carácter de norma.

La primera edición del diccionario unificado, con esas primeras 20.000 formas vio la luz, en formato papel, a finales del año 2000. En la imagen se pueden apreciar las características del diccionario.

The diagram shows a list of dictionary entries for the word 'tobera' and its derivatives. Colored lines connect specific entries to labels on the right side of the page:

- entrada** (red line) points to the entry **tobera**.
- subentrada** (magenta line) points to the entry **tobera-mustra**.
- taxon** (green line) points to the entry **toil 1 *Mustelus mustelus***.
- ejemplo de uso** (blue line) points to the example sentence: *gorputza toil delarik, lou uxatzea zail da.*
- homógrafos** (red line) points to the entries **toil 1** and **toil 2**.
- régimen verbal** (blue line) points to the entry **toildu, toil(du), toiltzen. da/du ad.**
- categoría gramatical** (orange line) points to the entry **toka 1 adlag. 'gizonezko bati hika'**.
- marca de registro** (blue line) points to the entry **toketa 'gizonezkoarekiko hitanoa'**.

The dictionary entries listed are:

- tobera**
- tobera-mustra**
- togoar** (herritarra)
- toil 1** *Mustelus mustelus*
- toil 2** batez ere *Ipar.* izond. 'pisua, motela': *gorputza toil delarik, lou uxatzea zail da.*
- toildu, toil(du), toiltzen. da/du ad.**
- tok pisin** (hizkuntza)
- toka 1** adlag. 'gizonezko bati hika'
- toka 2** iz. 'jokoa'
- tokamen** *Heg.:* irabazietan bakoitzari dagokion tokamena.
- tokatu, toka(tu), tokatzen. Heg. da/zaio ad. Herr.**
- toketa** 'gizonezkoarekiko hitanoa'
- tokhariera** (herritarra)
- toki**
 - en tokian**
 - tokian tokiko**
 - tokitan:** *tokitan dabil!*
 - tokitara:** *tokitara joan da!*
- tokialdaketa**
- tokialdatu, tokialda, tokialdatzen. da/du ad.**
- tokika** adlag. 'han eta hemen'
- tokikotasun**
- tokikotu, tokiko(tu), tokikotzen. da ad. 'bertakotu'**
- tokilabilaso** *Ipar.* 'aitonaren aitona, tokilaso'
- tokilaso** *Ipar.*
- tokilo** *Ipar.* 'moldakaitza, jendetasunik gabea'

Una vez publicada esta primera edición, la Academia decidió ampliar el diccionario unificado con otras 20.000 entradas, siguiendo el mismo procedimiento. En 2009 el pleno facultó a una comisión para agilizar el procedimiento de aprobación, evitando la discusión detallada en sus reuniones mensuales. Como consecuencia de esta decisión, el segundo lote de 20.000 entradas quedará aprobado en 2012.

Por otro lado, finalizada la publicación del Diccionario General en 2005, la Academia encomendó al equipo de trabajo que había desarrollado ese diccionario un nuevo proyecto, consistente en enriquecer las primeras 20.000 entradas del diccionario incorporando definiciones y ejemplos de uso.

El resultado de ese proyecto es el diccionario titulado *Euskaltzaindiaren Hiztegia. Adierak eta adibideak* (Diccionario de Euskaltzaindia. Acepciones y ejemplos) que fue completado en diciembre de 2011 y está siendo preparado para su publicación este mismo año.

Lo que en la primera edición era básicamente una lista de palabras, ahora tiene las características que se pueden apreciar en este fragmento:

entrada categoría gramatical definición Ejemplo de uso

tobera iz. pl. Andregaiari ezkontza bezperan, gorabeherak dituzten senar-emazteei edo berriro ezkontzen den gizon alargunari egiten zaien musika edo zintzarri joaldia. *Toberak jo ditzagun.*

tobera-mustra Ezkontza hausteren bat-edo gertatzen denean, herriko plazan antolatzen den antzerkia. Ik. **xaribari**. *Tobera-mustra batzuk muntatu.*

togoar 1 izlag. Togokoa, Togori dagokiona. *Togoar agintariak.*
2 iz. Togoko herritarra.

toil1 iz. Marrazoaren eta arraiaren ordenako itsas arraina, handi samarra eta gorputz-luzea, azala latza eta kolore-grisa duena (*Mustelus mustelus*).

toil2 izond. *Batez ere Ipar. Pisua, motela. Gorputza toil delarik, loa uxatzea zail da. Pilota toila. Adimendu toila. Euskara hain beratz eta toil baldin badago, apaizek daukatela enbalditurik.*

-- régimen verbal

toildu, toil/toildu, toiltzen *da/du* ad. Toil edo toilago bihurtu, moteldu. *Bihotza toildu.*

--

toka1 adlag. Gizonezko bati hika. Ik. **noka**. *Eta hasi zitzaion erregeri toka. Hurrei ez zaie toka eta noka hitz egin behar.*

--

toka2 iz. Kaleko jokoa, distantzia jakin batera dagoen metalezko xafla estu baterantz metalezko diskoak jaurtitzean datzana, eta xafla jotzea helburu duena. *Sagardoa galanki edango dugu, tokan jokatzen dugun bitartean. Tokan errege eta maisu.*

Este diccionario será, como decía al principio, el último eslabón de la cadena de proyectos de Euskaltzaindia relacionados con el léxico, por ahora, puesto que se está trabajando en nuevos proyectos que sentarán las bases de ulteriores productos.

Desde 2008 se está construyendo un nuevo corpus denominado *Lexikoaren Behatokia* (Observatorio del léxico), basado en textos de los medios de comunicación y cuyos objetivos son dos: servir de monitor del uso actual del euskera escrito y sentar las bases del futuro Corpus de Referencia del Euskera Actual.

Actualmente el volumen de este corpus es de 17 millones de palabras texto y está etiquetado y lingüísticamente anotado.

2012-3-15
Andoni Sagarna,
director de la Sección de Investigación de Euskaltzaindia