

## **Sintaxis y formación de palabras en los diccionarios electrónicos: los compuestos nominales del islandés**

Rafael García Pérez, Universidad Carlos III (Spain)

### **Índice**

#### [1 Introducción](#)

#### [2 Los compuestos nominales del islandés](#)

##### [2.1 Los compuestos sintagmáticos](#)

##### [2.2 Los compuestos ortográficos del islandés](#)

###### [2.2.1 Los compuestos lexicalizados](#)

###### [2.2.2 Compuestos institucionalizados](#)

###### [2.2.3 Los sintagmas composicionales](#)

#### [3 Conclusión](#)

#### [Bibliografía](#)

### **1 Introducción**

Una de las grandes dificultades, cuando se trata de describir una lengua natural desde el punto de vista del tratamiento automático del lenguaje, es el reconocimiento de las estructuras composicionales. Este problema resulta especialmente significativo en el caso de las lenguas románicas, donde los llamados “compuestos sintagmáticos”, por su propia forma externa, no siempre se distinguen con facilidad de los sintagmas libres.<sup>1</sup> La separación entre las unidades que pueden combinarse libremente en el discurso y las que presentan un cierto grado de fijación sintáctica constituye una actividad lingüística imprescindible para el lexicógrafo que decide confeccionar un

---

<sup>1</sup> Bustos Gisbert (1986: 69) ya ponía de manifiesto esta dificultad al definir los compuestos sintagmáticos como “aquéllos que, en su estructura formal no ofrecen diferencia alguna respecto a los sintagmas nominales, constituidos de dos sustantivos unidos por medio de la preposición *de*, o por un sustantivo y un adjetivo”. Lang (1992: 93) señala específicamente que el principal problema de los compuestos sintagmáticos “radica en decidir qué sintagmas de entre los muchos que podrían considerarse compuestos deben ser clasificados indudablemente como tales y cuáles son los sintagmas ‘libres’ en una estructura sintáctica normal”. Es evidente que, aunque el compuesto sintagmático puede considerarse una construcción sintáctica fija, “ello no significa que necesariamente pierda todos los rasgos de una construcción sintáctica. La fijación no altera la configuración acentual y segmental de la construcción, ni la forma de la estructura sintáctica—por ejemplo, sintagma nominal con complementos preposicionales o adjetivos (Val Álvaro 1999: 4763). Gross (1996a: 29), desde el punto de vista de la traducción automática, habla de “lexías complejas” o “unidades poliléxicas”, términos que incluyen tanto los compuestos propios como los compuestos sintagmáticos o impropios: “La composition, elle, concerne des éléments lexicaux, c’est à dire, susceptibles d’un emploi autonome, constituant entre eux des unités polylexicales, séparées par des blancs ou par d’autres séparateurs”. Los trabajos del léxico-gramática y las clases de objetos en Francia han revelado que los compuestos nominales presentan estructuras sintagmáticas muy diversas, desde las tradicionales N prep. N (“Pomme de terre”) o NA (“Table ronde”) hasta las expansiones sintagmáticas más largas, como N de Det. N Adv. Pp (“La clause de la nation la plus favorisée”). Mathieu-Colas (1996) ha establecido la tipología más completa para el francés. En el presente trabajo, adoptaré el término de “compuestos nominales” para referirme tanto a los propios como a los impropios o sintagmáticos. Entiendo el concepto de “compuesto sintagmático”, siguiendo los avances en el campo del tratamiento automático del lenguaje, en su sentido más amplio: “aquellas agrupaciones sintácticas fijadas por el uso con un significado unitario” (Pieira y Varela 1999).

diccionario electrónico.<sup>2</sup> Aunque ese grado de fijación puede ser mayor o menor, como bien ha demostrado, entre otros, Gross (1996a)—y, por tanto, más o menos difícil de determinar—de su correcta detección y definición depende enormemente la calidad del programa informático que el ordenador utilice y su posterior eficacia en la interpretación de los textos. En los ejemplos siguientes, sólo (1.b), por su falta de libertad combinatoria, podría considerarse propiamente un compuesto nominal fijo:

(1.a) El coche de mi hermano

(1.b) Mano de santo (\*mano de santos, \*mano de ese santo, \*mano de los santos, \*mano del santito, etc.)

La exclusión de (1.b) como “lexía compleja” generaría errores de interpretación y traducción por parte de un ordenador. El lingüista-lexicógrafo debe, por tanto, delimitar rigurosamente cuándo nos hallamos ante estructuras fijas y cuándo ante verdaderas estructuras libres e incluir las primeras como lemas. La criba a la que han de someterse los sintagmas nominales, en el caso de las lenguas románicas, tiene su justificación en el hecho de que los compuestos impropios se hallan camuflados entre ellos, a pesar de constituir una excepción a su naturaleza sintáctico-semántica. Los compuestos propios, como unidades ortográficas, presentan claras estructuras no sintagmáticas y, por lo tanto, carentes de ambigüedad con un alto grado de restricción transformacional (“pelirrojo”, “sacapuntas”, etc.), lo que los hace fácilmente localizables.

Contrariamente a lo que sucede en las lenguas románicas, las lenguas germánicas recurren con mayor frecuencia a la formación de compuestos ortográficos.<sup>3</sup> El primer impulso nos llevaría a considerar que la tarea de los lexicógrafos se ve así facilitada, ya que una gran parte de los compuestos sintagmáticos románicos tiene como correlato, en las lenguas germánicas, unidades ortográficas de carácter pluriverbal.<sup>4</sup> Así sería si el conjunto de dichas unidades coincidiera con la suma de los compuestos propios y los impropios de las lenguas románicas, es decir, si todas ellas presentaran algún tipo de restricción sintáctico-semántica. Desgraciadamente, estos compuestos no encubren sólo construcciones más o menos fijas, lexicalizadas o institucionalizadas, sino también sintagmas libres. Por otro lado, la reducción del número de compuestos sintagmáticos, como consecuencia de la preferencia de estas lenguas por la composición ortográfica, no implica su total desaparición. En el presente artículo se pretende mostrar, desde el punto de vista de las necesidades de un diccionario electrónico, los problemas que plantean la selección y clasificación de los compuestos nominales del islandés moderno.

---

<sup>2</sup> El término “diccionario electrónico” podría resultar ambiguo. En este trabajo no me refiero, evidentemente, a la variante en soporte informático de las obras lexicográficas tradicionales, sino a la herramienta que debe utilizar un ordenador para llevar a cabo una correcta interpretación textual e, incluso, generar textos. Para la forma de un diccionario electrónico de estas características, me remito a los estudios de Gaston Gross (1992) y Maurice Gross (1994).

<sup>3</sup> Por “compuesto ortográfico” hay que entender las unidades pluriverbales que no se hallan separadas por blancos en la escritura (Lang 1992: 91).

<sup>4</sup> Como ejemplo, bastaría mencionar que “molino de viento” en español o “moulin à vent” en francés se traducen como “windmill” en inglés, “Windmühle” en alemán, o “vindmyllur” en islandés.

## 2 Los compuestos nominales del islandés

El islandés, como ya he apuntado, tiene origen germánico, aunque se ha clasificado, más concretamente, dentro de la rama nórdica de esta subfamilia, conocida como “germánico septentrional” (Ramat y Ramat 1995: 491). Comparte con el resto de la familia el recurso frecuente a los mecanismos de composición para la formación de nuevas palabras,<sup>5</sup> mecanismos que, según la gramática tradicional, producirían los siguientes resultados:

- a) Los compuestos del islandés constituyen siempre una unidad gráfica.<sup>6</sup>
- b) Por lo general, los compuestos presentan una estructura bimembre (Bráinsson 1995: 254). Incluso las combinaciones de más de dos unidades se reagrupan en torno a un elemento nuclear o regidor y un elemento regido, y no es rara la lexicalización—por ejemplo, “ljósmynd” (“fotografía”) es un compuesto lexicalizado que puede combinarse con el sustantivo “vél” (“máquina”) para formar “ljósmyndavél” (“cámara fotográfica”), que a su vez puede considerarse lexicalizado. Enseguida veremos los criterios capaces de determinar el grado de lexicalización del compuesto.
- c) La conservación de un sistema casual bastante rico permite distinguir tres tipos de compuestos desde un punto de vista morfológico: a) compuestos por yuxtaposición de dos sustantivos—“ljós” (“luz”) + “mynd” (“imagen”) > “ljósmynd” (“fotografía”); b) compuestos en que el primer elemento presenta un morfema de genitivo—“rannsókn” (genitivo plural) + “stofnun” > “rannsóknastofnun” (“centro de investigaciones”); c) compuestos en que aparece una vocal de unión—“ráð” (“consejo”) + “nautur” > “ráðunautur” (“consejero”).<sup>7</sup>

De aquí se desprende que los criterios que ha seguido la gramática tradicional son, fundamentalmente, gráficos y morfológicos, hasta tal punto que se considera compuesto cualquier grupo de palabras que aparece como un bloque, sin blanco en la escritura. Tales criterios parecen tener su correlación inmediata en una definición implícita de la palabra como espacio entre dos blancos. Según este punto de vista, serían compuestos los conjuntos léxicos (2a), (2b), (2c) y (2d):

- (2.a) Ráðunautur, “consejero”.
- (2.b) Menntamálaráðuneyti, “Ministerio de Educación”.
- (2.c) Timburkirkja, “iglesia de madera”.
- (2.d) Bókasala, “venta de libros”.

Quedarían excluidos, por tanto, (3.a), (3.b) y (3.c):

---

<sup>5</sup> La tendencia se encuentra ya en el proto-germánico (Ramat y Ramat 1995: 522-523).

<sup>6</sup> Al estudiar el fenómeno de la composición, las gramáticas y los tratados de morfología del islandés sólo tienen en cuenta los procedimientos de formación de nuevas unidades ortográficas, a partir siempre de dos o más raíces diferentes (Röngvaldsson 1990: 37-40; Bráinsson 1995: 253-255; Pétursson 1996: 170-171).

<sup>7</sup> Llamados por la gramática tradicional, respectivamente, “fast samsett” o “stofnsamsett”, “laust samsett” o “eignarfallssamsett” y “bandstafssamsetning” (Röngvaldsson 1990: 37).

- (3.a) Grænlands sléttbakur, “ballena de cabeza arqueada”<sup>8</sup>
- (3.b) Þjóð Guðs, “pueblo de Dios”
- (3.c) Hús nágrannans, “casa del vecino”

Desde el punto de vista del tratamiento automático del lenguaje, esta interpretación del concepto “composición nominal” presenta varios problemas:

- a) No tiene en cuenta la posible existencia de una composición sintagmática (similar a la estudiada en las lenguas románicas) o, al menos, la considera irrelevante.
- b) Oculta el carácter sintáctico de la construcción composicional y le atribuye una entidad uniforme según su imagen puramente gráfica.

El desarrollo de estos dos puntos, a continuación, me parece esencial para entender el verdadero alcance del fenómeno de la composición en islandés moderno.

## 2.1 Los compuestos sintagmáticos

Si el reconocimiento de los compuestos sintagmáticos resulta imprescindible en la descripción de las lenguas románicas, no lo es menos en el caso de las lenguas germánicas y, concretamente, del islandés. Estos compuestos se caracterizan por presentar un cierto grado de fijación sintáctica a pesar de su forma lingüística. Dada la frecuencia con que se utiliza el mecanismo de la composición cabría preguntarse si tales compuestos sintagmáticos existen en realidad. En este sentido, la aplicación de las pruebas mencionadas más arriba constituye la única vía para la obtención de una prueba satisfactoria.

Parece claro que un ejemplo como (3.c) responde adecuadamente a los criterios de un sintagma libre, ya que ninguno de los elementos que lo integran se halla sometido a restricciones sintáctico-semánticas. Todas las manipulaciones recogidas en (4) son posibles.

- (4.a) Hús þessa nágranna, “la casa de este vecino”
- (4.b) Hús franska nágrannans, “la casa del vecino francés”
- (4.c) Hús nágranna míns, “la casa de mi vecino”
- (4.d) Rauða hús nágranna þíns, “la casa roja de tu vecino”
- (4.e) Hús og bill nágrannans, “la casa y el coche del vecino”
- (4.f) Hús nágrannans og föður síns, “la casa del vecino y su padre”
- (4.g) Hús nágrannans → Húsið sem nágranni á, “la casa que posee el vecino”.

Si nos fijamos en (3.a), por el contrario, nos daremos cuenta de que no existe posibilidad alguna de alterar sintácticamente la estructura del sintagma:

- (5.a) \*Spánar sléttbakur
- (5.b) \*Grænlands góður sléttbakur
- (5.c) \*Grænlands og Noregs sléttbakur
- (5.d) \*Sléttbakur þess lands
- (5.e) \*Hvíta Grænlands sléttbakur
- (5.f) \*Grænlands sléttbakur og síld

---

<sup>8</sup> Tipo de ballena de cuerpo cilíndrico, más estrecho hacia la cola, y cuya cabeza, levemente triangular, mide un tercio, más o menos, de la longitud total del animal.

Evidentemente, se trata de un compuesto nominal que, salvo por la anteposición del sustantivo en genitivo, poco frecuente en el habla cotidiana, no se distingue fácilmente de otros sintagmas. La traducción de un ordenador que desconociera el carácter composicional de la estructura sería completamente inadecuada.<sup>9</sup> Ejemplos como éste ponen de manifiesto que la composición sintagmática tiene importancia en la descripción de las lenguas germánicas y que su estudio sistemático es una condición “sine qua non” para elaborar, con rigor, un buen diccionario electrónico. Ahora bien, si casos como el anterior pueden parecer claros para la lexicografía tradicional, tanto especializada como no especializada,<sup>10</sup> el fenómeno se complica cuando, como sucede en las lenguas románicas, la fijación de los compuestos es sólo una cuestión de grado. El ejemplo (3.b) plantea dudas a simple vista, porque ni los rasgos morfológicos ni la colocación de sus elementos (posposición del genitivo como en los sintagmas libres más frecuentes) parecen aislarlo como compuesto. Sólo el recurso al análisis sintáctico de sus componentes puede ayudarnos a disiparlas:

- (6.a) \*Þjóð þessa Guðs, “pueblo de este dios”
- (6.b) \*Þjóð guða, “pueblo de dioses”
- (6.c) \*Guðsþjóð / Guðþjóð
- (6.d) \*Þjóð og land Guðs, “pueblo y tierra de Dios”
- (6.e) \*Þjóð Guðs og konungs, “pueblo de Dios y del Rey”

Parece claro que el sintagma se acerca más a una expresión fija que a una combinación libre, a pesar de su forma, pues ha sufrido un proceso de institucionalización bastante acusado. Este tipo de compuestos no puede integrarse en los lexicalizados, porque las restricciones guardan más relación con el uso que con la pura gramática. Si analizáramos “Þjóð Guðs” fuera de los contextos en los que normalmente aparece, nos percataríamos de que muchas transformaciones son, en cierta medida, posibles. Si las transformaciones propuestas no son aceptables, ello se debe a la aplicación concreta que se hace del término en una cultura determinada.

A veces, algunos de estos compuestos resultan aún más difíciles de reconocer, pues los elementos integrantes pueden alterar su posición, omitir algunos elementos en el contexto, e incluso presentar distintas variantes ortográficas, lo que daría una impresión de aparente libertad sintáctica.

- (7) Alþjóðlegur baráttudagur verkalýðsins / alþjóðlegur dagur verkalýðsins / dagur verkalýðsins / verkalýðsdagur, “día del trabajo”.

No obstante, las posibilidades combinatorias, dejando aparte estas estructuras canónicas, son escasas. En el eje paradigmático, por ejemplo, sería imposible sustituir

---

<sup>9</sup> De hecho, la imprecisión de una traducción literal del compuesto queda confirmada al analizar algunos de sus posibles equivalentes en otras lenguas. La traducción en español, como he señalado, sería “ballena de cabeza arqueada” o “ballena boreal”; en francés, podemos encontrar “baleine boréale” o “balaine franche du Groenland”; en alemán “Grönlandwal” o “Polarwal”; y en inglés “arctic right whale”, “bowhead whale”, o incluso “Greenland right whale”. Un ordenador tiene que contar con la denominación en la lengua de partida y todos sus equivalentes en las distintas lenguas meta.

<sup>10</sup> Los diccionarios monolingües y, en menor medida, los diccionarios bilingües suelen recoger compuestos sintagmáticos con un alto grado de fijación: “öffentliche Hände” en alemán; “main de Fatma” en francés; o “beggja handa járn” en islandés, por poner algunos ejemplos al azar contruidos con el sustantivo “mano”. Los diccionarios especializados, monolingües y bilingües recogen muchos términos que pueden considerarse compuestos sintagmáticos, aunque no son exhaustivos. Un diccionario electrónico para la traducción automática debe ir más lejos y lematizar todas las unidades si pretende obtener los mejores resultados.

el sustantivo "verkalýður" por "vinna", tratando de aproximar el término islandés a su traducción española:

(7.a) \*Dagur vinnunnar / \*Dagur vinnu / \*Vinnudagur<sup>11</sup>

Estos compuestos, con escasa representación en las obras lexicográficas tradicionales, pueden generar errores de traducción si no se registran como entradas de un diccionario electrónico. No bastaría, sin embargo, con recoger una sola variante o hacer una selección de las variantes más extendidas, sino que habría que incorporarlas todas. Un ordenador, por extraño que nos parezca a los hablantes de una lengua, no es capaz de remitirnos de unas variantes a otras si éstas no se han descrito antes como términos equivalentes y no se han puesto en relación con su traducción o sus traducciones potenciales en otras lenguas.

## 2.2 Los compuestos ortográficos del islandés

De igual modo que sucede con los sintagmas, los compuestos ortográficos del islandés esconden relaciones sintácticas muy diversas. En realidad, desde el punto de vista de su fijación, no conforman un conjunto homogéneo. Cabría distinguir, al menos, tres subgrupos:

- a) Los compuestos lexicalizados, es decir, las expresiones fijas en sentido estricto, en las que no es posible alteración alguna de la estructura. Enseguida veremos qué tipo de manipulaciones pueden aplicarse a este tipo de construcciones.
- b) Los compuestos institucionalizados, o compuestos no lexicalizados en sentido estricto que tienen un uso específico dentro del sistema lingüístico.
- c) Los falsos compuestos o, si se prefiere, los sintagmas composicionales. Se trata, como veremos, de sintagmas libres que revisten la forma tradicional del compuesto nominal.

### 2.2.1 Los compuestos lexicalizados

Son compuestos lexicalizados aquéllos que no admiten manipulación alguna desde el punto de vista sintáctico. Frecuentemente evocan un concepto unitario y, entre ellos, cabe incluir tanto las construcciones N + N con vocal de unión como muchas construcciones por yuxtaposición o en las que el primer elemento se declina en genitivo. Como tendremos ocasión de comprobar, no existe una relación directa entre los rasgos morfológicos del compuesto y su posible fijación. Solamente el grupo de los compuestos formados con la ayuda de una vocal epentética constituyen claramente, y sin excepción, estructuras lexicalizadas. En realidad, la epéntesis de vocal constituye un mecanismo muy poco productivo de creación de palabras y los ejemplos no son numerosos.

La prueba principal para determinar el grado de fijación del compuesto ortográfico implica comprobar, en primer lugar, si sus elementos son separables y, en caso de

---

<sup>11</sup> "Dagur" significa literalmente "día" y "verkalýður", "clase trabajadora". La sustitución de "verkalýður" por el sustantivo "vinna" ("trabajo") generaría estructuras inaceptables, al menos en lo que se refiere a su supuesta equivalencia con el término original. "Vinnudagur" sería totalmente imposible como traducción de "Día del trabajo", porque ya existe en islandés para designar un "día laborable".

que así sea, averiguar si la estructura sintagmática resultante puede someterse a transformaciones idénticas a las de los sintagmas libres.

- (8.a) Ráðuneyti, “ministerio” → \*Neyti ráðs
- (8.b) Ljósmynd, “fotografía” → \*Mynd ljóss
- (8.c) Krossfiskur, “estrella de mar” → \*Fiskur kross
- (8.d) Samkeppnispróf, “oposición” → \*Próf samkeppnis

La traducción confirma la opacidad semántica del compuesto y, en consecuencia, el bloqueo de sus propiedades sintácticas. La posibilidad de separar sus dos partes es razón “sine qua non” para proceder a la aplicación de cualquier otra prueba, de modo que, en los ejemplos de (8), nos hallaríamos ante compuestos plenamente lexicalizados que un diccionario electrónico debe recoger como lemas.

## 2.2.2 Compuestos institucionalizados

Los compuestos institucionalizados suponen, en principio, y desde un punto de vista teórico, un cierto grado de libertad sintáctica, de modo que sus elementos podrían ser separables. No obstante, presentan restricciones que tienen que ver con el uso cotidiano y que impiden importantes transformaciones. Tomemos como ejemplo el compuesto “Parísarháskóli” (“Universidad de París”).

“Parísarháskóli” está formado por dos sustantivos: “París”, declinado en genitivo, y “háskóli” (“universidad”). Del mismo modo que éste, se han formado también los compuestos de (9.a):

- (9.a) Kaupmannahafnarháskóli, “Universidad de Copenhague”; Uppsalaháskóli, “Universidad de Uppsala”; Oxfordháskóli, “Universidad de Oxford”.

No parece descabellado intentar dictar una regla según la cual el nombre de la ciudad se antepone, declinado en genitivo, al sustantivo “háskóli”. Tal presunción podría resultar errónea si tenemos en cuenta los ejemplos de (9.b):

- (9.b) Háskólinn í Kaupmannahöfn, “Universidad de Copenhague”; Háskólinn í Uppsölum, “Universidad de Uppsala”.<sup>12</sup>

Como puede comprobarse, muchos de los compuestos ortográficos formados por la suma del topónimo y el sustantivo “háskóli” se transforman fácilmente en estructuras sintagmáticas en las que el primero recupera plenamente su sentido locativo. Esta transformación no siempre resulta aceptable, sin embargo:

- (9.c) Parísarháskóli VIII → \*Háskóli VIII í París / \*Háskólinn París VIII

El desplazamiento del topónimo, declinado en genitivo, a la derecha del núcleo no se produce prácticamente en ningún caso:

- (9.d) Kaupmannahafnarháskóli → \*Háskóli Kaupmannahafnar

En ocasiones, la estructura como compuesto ortográfico es muy poco frecuente o imposible. Las combinaciones sintagmáticas alternativas, por su parte, no resultan predecibles:

---

<sup>12</sup> La traducción literal sería “\*Universidad en Copenhague” y “\*Universidad en Uppsala” respectivamente.

- (9.e) Háskóli Íslands, “Universidad de Islandia” / \*Háskólinn á Íslandi / \*Íslandsháskóli  
(9.f) Háskóli Reykjavíkur / Háskóli í Reykjavík, “Universidad de Reykjavík” / ??Reykjavíkurháskóli  
(9.g) Háskóli Akureyrar / Háskóli á Akureyri, “Universidad de Akureyri” / ??Akureyrarháskóli

La traducción española oculta, en cierta medida, el hecho de que la forma adoptada por los dos sustantivos que forman el compuesto ha sido elegida arbitrariamente por la comunidad lingüística. Pero sí nos muestra precisamente que, frente a la aparente libertad del conjunto léxico, la interpretación está bloqueada. Los hablantes han nombrado de modo diferente realidades extralingüísticas que presentan gran similitud desde el punto de vista conceptual. En principio, no existe ninguna razón que justifique la unión gráfica de los constituyentes del compuesto en (9.a) ni su distribución como sintagma tradicional en (9.e), (9.f) o (9.g). Se trata de denominaciones extendidas en su uso habitual, acuñadas por motivos de necesidad, que han sufrido un proceso de institucionalización, hasta tal punto que en la actualidad han forjado en el discurso cotidiano una estructura bastante restringida.

Los compuestos institucionalizados constituyen una zona resbaladiza y no siempre la lexicografía tradicional ha sabido o ha podido dar cuenta de ellos con el rigor necesario. Es evidente que, como sucedía en el caso de los compuestos sintagmáticos del mismo tipo, su reconocimiento es esencial para lograr que un ordenador interprete con éxito los textos escritos. No hay que olvidar que se trata de un conjunto cerrado y, en el caso del islandés, formalmente heterogéneo, cuya función consiste en nombrar elementos determinados, bien establecidos, de la realidad cultural.

### 2.2.3 Los sintagmas composicionales

Por sintagmas composicionales hay que entender aquellos grupos de palabras que adoptan la forma de un compuesto ortográfico, pero no presentan grado alguno de fijación sintáctica. Se trata de conjuntos léxicos que funcionan, en realidad, como sintagmas libres. Esto se pone de manifiesto en la gran libertad combinatoria de sus constituyentes. En muchas ocasiones, ni siquiera forman una nueva palabra, sino que expresan una relación de dependencia sintáctica entre sus elementos. Fijémonos en los ejemplos siguientes:

- (10.a) Bensínverð, “precio de la gasolina”  
(10.b) Tölvunotkun, “uso de ordenadores”  
(10.c) Ávaxtasala, “venta de fruta”

El primer aspecto que conviene destacar es el hecho de que la estructura morfológica de estos compuestos no es idéntica. En (10.a) nos hallamos ante un compuesto por yuxtaposición; en (10.b) y (10.c) descubrimos que el primer constituyente está declinado en genitivo. Pero, mientras en (10.b) el número del sustantivo declinado es singular, en (10.c) es plural. Tal comprobación nos lleva a aceptar que no basta con recurrir a criterios morfológicos para tratar de reconocer los sintagmas composicionales. En ese sentido, no presentan ninguna novedad respecto a los grupos anteriores.

Los compuestos ortográficos de (10.a), (10.b) y (10.c) se caracterizan por ser separables, es decir, por el hecho de poder formar un sintagma nominal libre en el que el primer elemento actúa como complemento. En el caso de que existan construcciones preposicionales equivalentes, este sustantivo puede verse regido por ella, con lo que se pone aún más de manifiesto su naturaleza de complemento

nominal. Es lo que sucede en (11.a) donde el nombre neutro "verð" puede ir acompañado por un sustantivo en genitivo o un sustantivo precedido de la preposición "á":

- (11.a) Bensínverð → verð bensíns / verð á bensíni
- (11.b) Tölvunotkun → notkun tölvu
- (11.c) Ávaxtasala → Sala ávaxta

Naturalmente, todos estos sintagmas pueden ser objeto de las mismas manipulaciones que el resto de los sintagmas libres:

- (12.a) Verð þessa bensíns, "el precio de esta gasolina" / verð bensíns sem hefur verið flutt til Íslands, "el precio de la gasolina que se ha exportado a Islandia" / verð á bensínlitra, "el precio de un litro de gasolina" / verð bensíns → verð þess, "el precio de la gasolina" → "su precio".
- (12.b) Notkun þessarar tölvu, "el uso de este ordenador" / Notkun tölvanna sem fyrirtækið hefur keypt, "el uso de los ordenadores que ha comprado la empresa" / Notkun nýju tölvanna, "el uso de los nuevos ordenadores" / Notkun tölvu → notkun hennar, "el uso de ordenadores" → su uso.
- (12.c) Sala þessara ávaxta, "la venta de esta fruta" / sala ávaxta sem hefur verið fluttir til Íslands, "la venta de la fruta exportada a Islandia" / Sala allra ávaxta, "la venta de toda la fruta" / Sala ávaxta → sala þeirra, "venta de fruta" → su venta.

Aunque los ejemplos podrían multiplicarse, parece claro que los sintagmas formados por la separación de los constituyentes de los compuestos ortográficos de (10) no presentan restricciones desde el punto de vista sintáctico. La arbitrariedad de su formación y su carácter puramente sintagmático pueden entenderse perfectamente si sustituimos los complementos nominales por otros sustantivos:

- (13.a) Grænmetisverð → verð grænmetis, "el precio de la verdura" / ávaxtaverð → verð ávaxta, "el precio de la fruta" / landverð → verðlands, "el precio de la tierra" / bókaverð → verð bóka, "el precio de los libros", etc.
- (13.b) Kortanotkun → notkun korta, "el uso de mapas" / Bílanotkun → notkun bíla, "el uso de coches" / orðabókanotkun → notkun orðabóka, "el uso de diccionarios", etc.
- (13.c) Bílasala → Sala bíla, "venta de coches" / bókasala → sala bóka, "venta de libros" / grænmetissala → sala grænmetis, "venta de verdura", etc.

Ninguno de los compuestos así construidos crea una palabra verdaderamente nueva en el sistema lingüístico. A nadie puede pasarsele inadvertido el hecho de que en los ejemplos (10.b) y (10.c) nos hallamos ante predicados nominales (no actualizados) con una amplia estructura argumental.<sup>13</sup> Es curioso comprobar que, en todos los casos, es el argumento N1 el que aparece unido al predicado, a partir de las oraciones de base (14) y (15) previamente transformadas en oraciones pasivas de cierta impersonalidad semántica:

- (14) Einhver notar hlut, "alguien usa un objeto concreto" → Tölvur eru notaðar → Notkun tölvu → tölvunotkun
- (15) Einhver selur einhverjum eitthvað, "alguien vende algo a alguien" → Ávextir eru seldir → Sala ávaxta → Ávaxtasala

---

<sup>13</sup> Para los conceptos de predicado nominal y argumento me baso en el marco teórico de la léxico-gramática y las "Clases de objetos", difundido en Francia principalmente por Maurice Gross y Gaston Gross. Entre otros trabajos de relevancia cabría citar Gross (1981), Gross y Vivès (1986), Giry Schneider (1987) y Gross (1996b).

Esta nominalización constituye un proceso idéntico al que se produce en las lenguas románicas, con la salvedad de que en islandés se avanza un paso más hasta llegar a la tematización del argumento. Ahora bien, las razones de esta tematización no resultan fácilmente deducibles. El sustantivo argumental o complemento del nombre se presenta aislado en esa posición y se requiere la separación previa de los constituyentes para poder aplicar las transformaciones sintácticas:

(16) Ávextasala → \*þessara ávextasala → \*spænsku ávextasala

Este hecho guarda estrecha relación con el carácter de sustantivo genérico que adopta el argumento o el complemento nominal cuando se usa al margen de toda determinación. Se trata de una característica esencial. El desplazamiento del argumento o complemento del nombre a la posición inicial y su integración con el sustantivo que lo rige como unidad gráfica sólo puede entenderse si pensamos que de ese modo el sustantivo argumental alcanza su grado genérico más alto. Cualquier alteración a la que se vea sometido supone una concreción y, por lo tanto, una destrucción del compuesto ortográfico. En todo caso, hemos visto que, con el mismo sentido genérico, los sustantivos pueden presentar una estructura doble: la composicional y la sintagmática, lo que refuerza la idea de que nos hallamos ante una mera cuestión formal.

Desde el punto de vista de la traducción automática, la libertad combinatoria de estos elementos nos obliga a tratarlos como el resto de los sintagmas derivados de procesos transformacionales. El intento de recogerlos todos se revelaría infructuoso, no sólo por el elevado número de los ya existentes, sino por la aparición ininterrumpida de otros nuevos en el discurso cotidiano debido a la creatividad de los hablantes. En ese sentido parece razonable defender, como han propuesto las teorías de la léxico-gramática y las clases de objetos para otras lenguas germánicas,<sup>14</sup> la confección de programas informáticos complementarios dotados de algoritmos capaces de realizar una división en constituyentes de las secuencias ligadas ortográficamente.

### 3 Conclusión

El fenómeno de la composición nominal en islandés resulta más complejo de lo que las gramáticas al uso nos dan a entender. Desde el punto de vista de la elaboración de un diccionario electrónico, la consideración del término compuesto tiene mucho que ver con el grado de fijación que alcanzan las diversas estructuras lingüísticas. En el presente artículo se ha puesto de manifiesto que limitarse a un estudio meramente ortográfico no resuelve los problemas de interpretación textual planteados por los programas de traducción automática. Para que un ordenador pueda “leer” y “producir” textos con las suficientes garantías, es necesario que cuente con una descripción muy fina de todas las estructuras que contienen algún tipo de bloqueo sintáctico-semántico, ya que, al contrario de los seres humanos, las máquinas son incapaces de realizar inferencias y suplir la información no aportada por el lexicógrafo. En ese sentido, se ha establecido una distinción entre combinaciones libres de palabras y combinaciones con un grado más o menos alto de fijación sintáctico-semántica (a las que hemos denominado compuestos), independientemente de su presentación tipográfica. Las soluciones a las que se ha llegado serían las siguientes:

---

<sup>14</sup> Ten Hacken (2002: 61-87), por ejemplo, hace una descripción de algunos programas de formación de palabras que podrían resultar útiles para el reconocimiento de las secuencias libres integradas en los compuestos ortográficos.

- La composición sintagmática es tan importante en islandés como en las lenguas románicas y un diccionario electrónico debe dar cuenta de ella rigurosamente. Mientras las variantes más opacas—aquéllas que presentan significados muy alejados del que produciría la suma de sus componentes y gran rigidez sintáctica—han podido recogerse, con mayor o menor fortuna, en algunos diccionarios tradicionales, generales o especializados, las variantes institucionalizadas (“þjóð Guðs”, “alþjóðlegur baráttudagur verkalyðsins”) no han recibido la atención suficiente por parte de la lexicografía. La lematización y descripción de todos y cada uno de los compuestos sintagmáticos es una condición irrenunciable para evitar los errores de interpretación y generación de textos.
- Los compuestos ortográficos no pueden ser objeto de un tratamiento unitario, pues esconden realidades lingüísticas de diverso tipo: a) expresiones fijas en sentido estricto; b) expresiones institucionalizadas; y c) meros sintagmas nominales cuyo uso puede considerarse impredecible. Mientras las unidades pertenecientes a los dos primeros grupos forman parte como entradas simples de un diccionario electrónico, las del tercero se asimilan a las estructuras sintagmáticas libres y, por tanto, son objeto de análisis por parte de programas específicos de formación de palabras.

## Bibliografía

- Bustos Gisbert, Eugenio de (1986): *La composición nominal en español*. Salamanca: Universidad de Salamanca.
- Giry Schneider, Jacqueline (1987): *Les prédicats nominaux en français*. Genève-Paris: Librairie Droz.
- Gross, Gaston (1992): “Forme d’un dictionnaire électronique”. En: Clas, André y Safar, H. (eds.), *L’environnement traductionnel. La station de travail du traducteur de l’an 2000*. Montréal & Sillery: Presses de l’Université du Québec, 255-271.
- Gross, Gaston (1996a) : *Les expressions figées en français*. Paris: Ophrys.
- Gross, Gaston (1996b): “Prédicats nominaux et compatibilité aspectuelle”. *Langages* 121, 54-72.
- Gross, Gaston y Vivès, Robert (1986): “Les constructions nominales et l’élaboration d’un lexique-grammaire”. *Langue française* 69, 5-27.
- Gross, Maurice (1981): “Les bases empiriques de la notion de prédicat sémantique”. *Langages* 63, 7-52.
- Gross, Maurice (1994): “Constructing lexicon-grammars”. En: Atkins, B.T.S. y Zampolli, A. (eds.), *Computational approaches to lexicon*. Oxford: Oxford University Press, 213-263.
- Lang, Mervyn Francis (1992): *Formación de palabras en español*. Madrid: Cátedra.
- Mathieu-Colas, Michel (1996): “Typologie de la composition nominale”. *Cahiers de lexicologie* 69, 65-118.
- Pétursson, Magnús (1996) : *Manuel d’islandais*, version française revue et augmentée par Emil H. Eyjólfsson. Paris: Klincksieck.
- Pieira, Carlos y Varela, Soledad (1999): “Relaciones entre morfología y sintaxis”. En: Bosque, Ignacio y Demonte, Violeta (eds.), *Gramática descriptiva de la lengua española*. Madrid: Espasa Calpe, 4367-4422.
- Ramat, Anna Giacalone y Ramat, Paolo (eds.) (1995): *Las lenguas indoeuropeas*. Madrid: Cátedra.
- Röngvaldsson, Eiríkur (1990): *Íslensk orðhlutafræði*. Reykjavík: Málvísindastofnun Háskóla Íslands.
- Ten Hacken, A. (2002): “Word formation in computational linguistics”. *Proceedings of TALN 2*, 61-87.

- Þráinsson, Höskuldur (1995): *Handbók um málfræði*. Reykjavík: Prentsmiðjan Oddi.  
Val Álvaro, José Francisco (1999): "La composición". En: Bosque, Ignacio y Demonte, Violeta (eds.), *Gramática descriptiva de la lengua española*. Madrid: Espasa Calpe, 4757-4841.