



Pontificia Universidad Católica de Chile
Facultad de Letras
Doctorado en Lingüística

ANÁLISIS DESCRIPTIVO DE ADJETIVOS COMPUESTOS CON PARTICIPIO EN INGLÉS

Tesis para optar al grado de Doctor en Lingüística

Alumna: Vanja Ostoic Marroquín
Profesor Guía: Dr. César Aguilar
Santiago, Chile
2018

1. Resumen

La composición como rama de la morfología involucrada en la formación de palabras, es el proceso mediante el cual se da origen a nuevos vocablos combinando ítems léxicos ya existentes en la lengua. En un proceso de por sí muy productivo en inglés, los adjetivos compuestos con participio constituyen una categoría bastante fecunda, tanto en términos de frecuencia de ocurrencia como de variedad de patrones morfológicos. Las descripciones los dividen en dos grandes grupos: aquellos en los que el núcleo es una forma de participio pasado (terminados en *-ed*) y aquellos que tienen un participio presente como base (terminados en *-ing*). Adicionalmente, dentro de los primeros, es posible encontrar estructuras compuestas cuyo núcleo se deriva de un sustantivo (*N-ed*), las que, a su vez, conforman un subgrupo de considerable tamaño.

En esta investigación hemos buscado determinar los distintos patrones morféimicos de este tipo de estructuras compuestas, teniendo en cuenta no solo la categoría gramatical del elemento inicial (sustantivos, adjetivos, adverbios, cuantificadores y números) sino que también la relación que se establece entre los dos constituyentes del compuesto. Es así que los sustantivos pueden ser interpretados como argumento de la base verbal o como adjuntos (*time-consuming, Nazi-occupied*). Para esto hemos analizado un corpus de prueba y muestras sacadas del corpus electrónico *enTenTen12* (11.191.860.036 palabras) usando una plataforma de análisis textual (*Sketch Engine*) con la que identificamos y cuantificamos los patrones para verificar si existen correspondencias numéricas entre los dos tipos de corpus. Además, damos cuenta de las regularidades encontradas en cada patrón en cuanto a frecuencia y a usos específicos para lograr una descripción detallada y una clasificación actualizada de estas unidades compuestas.

Índice

| | |
|---|----|
| 1. Resumen | 1 |
| 2. Introducción | 9 |
| Capítulo 1. Procesos morfológicos de la lengua inglesa | 12 |
| 1.1. La flexión..... | 13 |
| 1.2. La derivación | 15 |
| 1.2.1. Los afijos derivacionales en los diccionarios..... | 17 |
| 1.3. Similitudes entre derivación y composición..... | 21 |
| 1.4. Restricciones que afectan a la derivación y la composición | 22 |
| 1.4.1. Restricciones morfológicas | 22 |
| 1.4.2. Restricciones sintácticas | 23 |
| 1.4.3. Restricciones semánticas | 23 |
| 1.5. Los límites entre derivación y composición..... | 24 |
| 1.5.1. Afijos y semi-afijos..... | 25 |
| 1.5.2. Elementos compositivos | 28 |
| 1.6. Procesos de formación de palabras en inglés | 31 |
| 1.6.1. Acuñación léxica..... | 32 |
| 1.6.2. Préstamo lingüístico..... | 32 |
| 1.6.3. Entrecruzamiento | 33 |
| 1.6.4. Truncamiento | 35 |
| 1.6.5. Formación regresiva..... | 36 |
| 1.6.6. Analogía..... | 37 |
| 1.6.7. Conversión | 40 |
| 1.6.8. Acronimia | 42 |
| 1.6.9. Reduplicación | 43 |

| | |
|---|----|
| 1.7. Productividad en procesos de formación de palabras..... | 46 |
| 1.7.1. Enfoques y mediciones | 47 |
| 1.7.2. Restricciones a la productividad | 50 |
| Capítulo 2. La composición | 55 |
| 2.1. Criterios de identificación | 55 |
| 2.2. La interfaz entre morfología y sintaxis..... | 58 |
| 2.3. La composición en lengua inglesa..... | 60 |
| 2.3.1. ¿Qué es un compuesto en inglés? | 66 |
| 2.4. Principales clasificaciones de compuestos | 68 |
| 2.4.1. Compuestos endocéntricos y exocéntricos | 69 |
| 2.4.2. Compuestos subordinativos, atributivos y coordinativos | 71 |
| 2.4.3. Compuestos primarios y sintéticos | 72 |
| Capítulo 3. Adjetivos compuestos con participio en inglés | 76 |
| 3.1. Características generales..... | 76 |
| 3.2. Descripciones..... | 77 |
| 3.3. Adjetivos compuestos terminados en <i>-ed</i> | 80 |
| 3.3.1. Adjetivos compuestos del tipo N- <i>ed</i> | 80 |
| 3.3.2. Patrones morfológicos | 86 |
| 3.4. Adjetivos compuestos con participio presente | 88 |
| 3.4.1. Patrones morfológicos | 88 |
| 3.5. La teoría de Roeper y Siegel (1978)..... | 89 |
| Capítulo 4. Objetivos e hipótesis..... | 94 |
| 4.1. Objetivos Generales..... | 94 |
| 4.2. Objetivos específicos | 94 |
| 4.3. Hipótesis | 95 |
| Capítulo 5. Enfoque metodológico y metodología..... | 96 |
| 5.1. Lingüística de corpus..... | 96 |

| | |
|--|-----|
| 5.2. Corpus lingüístico..... | 97 |
| 5.2.1. Corpus electrónicos..... | 99 |
| 5.2.2. Corpus <i>enTenTen12</i> | 99 |
| 5.2.3. Corpus de prueba | 100 |
| 5.3. Sketch Engine | 101 |
| 5.4. Investigaciones guiadas o basadas en un corpus | 103 |
| 5.5. Metodología..... | 104 |
| 5.5.1. Búsqueda y extracción de datos | 105 |
| 5.5.2. Depuración del corpus | 108 |
| Capítulo 6. Resultados por patrones morfológicos individuales..... | 121 |
| 6.1. Corpus <i>enTenTen12</i> : Adjetivos compuestos terminados en <i>-ed</i> | 121 |
| 6.1.1. Sustantivo como primer elemento | 121 |
| 6.1.2. Adjetivo como primer elemento | 124 |
| 6.1.3. Adverbio como primer elemento | 125 |
| 6.1.4. Adjetivo comparativo como primer elemento | 126 |
| 6.1.5. Adverbio comparativo como primer elemento | 127 |
| 6.1.6. Adjetivo superlativo como primer elemento..... | 127 |
| 6.1.7. Adverbio superlativo como primer elemento | 128 |
| 6.1.8. Número como primer elemento | 129 |
| 6.2. Adjetivos compuestos con participios irregulares | 129 |
| 6.3. Corpus <i>enTenTen12</i> : Adjetivos compuestos terminados en <i>-ing</i> | 130 |
| 6.3.1. Sustantivo como primer elemento | 130 |
| 6.3.2. Adjetivo como primer elemento | 131 |
| 6.3.3. Adjetivo comparativo como primer elemento | 132 |
| 6.3.4. Adverbio como primer elemento | 133 |

| | |
|---|-----|
| 6.3.5. Adverbio comparativo como primer elemento | 133 |
| 6.3.6. Adjetivo superlativo como primer elemento..... | 134 |
| 6.3.7. Adverbio superlativo como primer elemento | 134 |
| 6.3.8. Número como primer elemento | 135 |
| 6.4. Resultados obtenidos en los diferentes corpus..... | 135 |
| 6.4.1. Corpus de prueba: Compuestos terminados en <i>-ed</i> | 136 |
| 6.4.2. Corpus de prueba: Compuestos terminados en <i>-ing</i> | 140 |
| 6.4.3. Corpus <i>enTenTen12</i> : Compuestos terminados en <i>-ed</i> | 142 |
| 6.4.4. Corpus <i>enTenTen12</i> : Compuestos terminados en <i>-ing</i> | 145 |
| 6.4.5. Resultados comparados..... | 146 |
| Capítulo 7. Propiedades morfosintácticas de los adjetivos compuestos con participio | 150 |
| 7.1. Formas flexivas como primer elemento: Plurales internos | 150 |
| 7.2. Formas derivadas en los adjetivos compuestos | 154 |
| 7.3. Elementos compositivos cultos | 156 |
| 7.3.1. <i>Self</i> | 157 |
| 7.4. Procesos de formación de palabras en los adjetivos compuestos..... | 158 |
| 7.4.1. Acuñación | 158 |
| 7.4.3. Truncamientos..... | 159 |
| 7.4.4. Hipocorísticos | 160 |
| 7.4.6. Reduplicación | 161 |
| 7.4.7. Compuestos como parte de otros compuestos | 162 |
| 7.5. Números en los compuestos | 163 |
| 7.6. Partículas | 165 |
| Capítulo 8. Análisis de los patrones morfológicos individuales | 167 |
| 8.1. Compuestos con núcleo nominal (N- <i>ed</i>) | 167 |

| | |
|--|-----|
| 8.1.1. Adj + N- <i>ed</i> / N + N- <i>ed</i> en la descripción de rasgos físicos y de carácter | 167 |
| 8.1.2. Adj + N- <i>ed</i> / N + N- <i>ed</i> en la descripción y nombres de animales | 171 |
| 8.1.3. Adj + N- <i>ed</i> / N + N- <i>ed</i> en la descripción de vestimentas | 175 |
| 8.1.4. Participios adjetivos + N- <i>ed</i> | 177 |
| 8.1.5. Adjetivo comparativo / superlativo + N- <i>ed</i> | 178 |
| 8.1.6. Número + N- <i>ed</i> | 178 |
| 8.1.7. Cuantificador + N- <i>ed</i> | 180 |
| 8.1.8. Adverbio + N- <i>ed</i> | 181 |
| 8.2. Compuestos con núcleo verbal (V- <i>ed</i>) | 182 |
| 8.2.1. Sustantivo + participio pasado | 182 |
| 8.2.2. Nombres propios | 188 |
| 8.2.3. Adverbio + participio pasado..... | 189 |
| 8.2.4. Adverbio comparativo / superlativo + participio pasado..... | 193 |
| 8.3. Adjetivos compuestos del tipo - <i>ing</i> | 194 |
| 8.3.1. Adjetivo + participio presente..... | 194 |
| 8.3.2. Adjetivo comparativo / superlativo + participio presente..... | 196 |
| 8.3.3. Sustantivo + participio presente..... | 197 |
| 8.3.4. Adverbio + participio presente..... | 199 |
| 8.3.5. Adverbio comparativo / superlativo + participio presente..... | 200 |
| 8.3.6. Número + participio presente..... | 201 |
| 8.4. Singularidades | 202 |
| 8.4.1. Pronombre como primer elemento..... | 202 |
| 8.4.2. Compuestos con - <i>shaped</i> | 202 |
| 8.4.3. Elementos iniciales coordinados..... | 204 |

| | |
|--|-----|
| 8.4.4. Compuestos creados por analogía..... | 205 |
| 8.4.5. Compuestos que eliden sustantivo..... | 205 |
| 9. Conclusiones | 207 |
| 9.1. Frecuencia de ocurrencia de patrones morfológicos en los corpus | 207 |
| 9.2. Identificación de patrones morfológicos | 209 |
| 9.3. Ambigüedad morfológica y semántica..... | 210 |
| 9.4. Consideraciones finales | 210 |
| Bibliografía..... | 212 |

“Although compounding is the most productive type of word-formation process in English, it is perhaps also the most controversial one in terms of its linguistic analysis and I must forewarn readers seeking clear answers to their questions that compounding is a field of study where intricate problems abound, numerous issues remain unresolved, and convincing solutions are generally not so easy to find.”

(Plag, 2003: 132)

2. Introducción

Las diferentes lenguas del mundo poseen diversos mecanismos que les permiten acrecentar su caudal léxico y entre ellos destaca la composición por ser un proceso particularmente productivo y extendido. El problema para los especialistas en morfología radica en la dificultad para lograr consenso sobre lo que constituye un término compuesto: “[...] there has always been much discussion of exactly what a compound is, and even of whether compounds exist as a distinct species of word formation” (Lieber y Stekauer, 2009: 4). Una de las razones para la disparidad de puntos de vista estriba en que no siempre es posible establecer una distinción clara entre palabras compuestas y palabras derivadas por un lado, y entre compuestos y frases, por el otro.

Si bien gran parte de las investigaciones sobre la formación de compuestos se ha realizado en lengua inglesa, dentro de esos estudios morfológicos, los adjetivos compuestos, en particular, han recibido poca atención. Las investigaciones sobre composición morfológica han estado mayoritariamente centradas en las combinaciones de sustantivos, puesto que, en inglés, su frecuencia de aparición es de 12 a 1 en relación a los adjetivos. Este tipo de unidades compuestas en particular constituye una de las formas más productivas de creación léxica en inglés aun cuando dicha productividad varía considerablemente dependiendo del patrón combinatorio del adjetivo. La creación de estos compuestos es un proceso de formación de palabras que combina composición con aspectos derivacionales y flexivos dando origen a descripciones y clasificaciones de naturaleza muy variada, lo que ha contribuido a que la sistematización de los patrones morfológicos y la distribución sintáctica de los adjetivos compuestos sea un área relativamente poco explorada todavía.

Estas unidades compuestas son muy utilizadas por los hablantes nativos de lengua inglesa al momento de realizar descripciones de muy diversa índole. El hecho de contar con un corpus de prueba representó una ventaja, pues nos permitió tener un conjunto apreciable de datos (casi 4.000 compuestos) representativos del fenómeno en su generalidad seleccionados de diversas fuentes. Efectivamente, en el corpus de prueba se recogieron ejemplos tomados tanto de textos literarios como de revistas populares y científicas y en todos ellos se observó una gran ocurrencia de estas

formas. Además, su análisis nos facilitó la identificación de patrones morfológicos que luego fueron comparados con aquellos presentes en el corpus de análisis.

Los grandes corpus electrónicos representan fuentes muy ricas para la recolección de datos pero la búsqueda se vuelve poco práctica cuando no hay un elemento determinado que buscar. Por otra parte, las fuentes más pequeñas, las que pueden ser revisadas manualmente, ayudan a restringir el campo de investigación pero pueden dejar fuera algunas formas y patrones entregando, por ende, una visión parcial del fenómeno. En consecuencia, la posibilidad de analizar estos dos tipos de corpus nos aseguraba una descripción de mayor alcance y detalle. Tomando en cuenta lo anterior, me propuse analizar un corpus de prueba de naturaleza léxica, así como muestras tomadas aleatoriamente de un gran corpus de análisis (*enTenTen12*) para determinar si los patrones morféimicos muestran correspondencias en estructura y frecuencia de ocurrencia entre los dos tipos de corpus lo que permitiría generar una descripción y clasificación actualizada de estas estructuras compuestas.

Esta es una investigación guiada por el corpus cuyo objetivo principal es profundizar en la descripción de los patrones morfológicos de los adjetivos compuestos con participio en inglés tomando en cuenta las regularidades y singularidades encontradas en los distintos volúmenes de datos.

El trabajo está presentado de la siguiente manera: en el capítulo 1 se hace una revisión general de los diferentes procesos de formación de palabras en lengua inglesa (derivación, composición, truncamientos, conversión, etc.), con especial énfasis en las similitudes y diferencias entre los dos primeros. En el capítulo 2 abordamos primeramente la composición y las complejidades que rodean la definición de compuesto para después presentar las principales clasificaciones. En el capítulo 3 se describen los adjetivos compuestos con participio y sus patrones morfológicos en las tres categorías principales (*N-ed*, *V-ed* y *V-ing*) y se resume la propuesta de Roeper y Siegel (1978). A continuación, se establecen los objetivos e hipótesis de la investigación (capítulo 4). En el capítulo 6 se muestran los resultados obtenidos del análisis de los datos contenidos en los corpus. En el

capítulo 7 se pasa revista a las propiedades morfosintácticas de los adjetivos compuestos para, finalmente, en el capítulo 8 analizar los patrones morfológicos individualmente. Finalmente, se presentan las conclusiones del estudio.

Capítulo 1. Procesos morfológicos de la lengua inglesa

La formación de palabras (*word formation*) es un área de la morfología derivacional que estudia la creación de nuevas palabras y cuyas dos categorías principales son la composición y la derivación. Además de ellas, existe una serie de procesos marginales (*minor word formation processes*) que también contribuyen a expandir el vocabulario en el idioma inglés. Bauer (2004: 284) delimita las áreas de la morfología de la siguiente manera: “Word-formation deals with derivational morphology, compounding and some rather less central processes, but not with inflection; anything that changes word-class is derivation.”

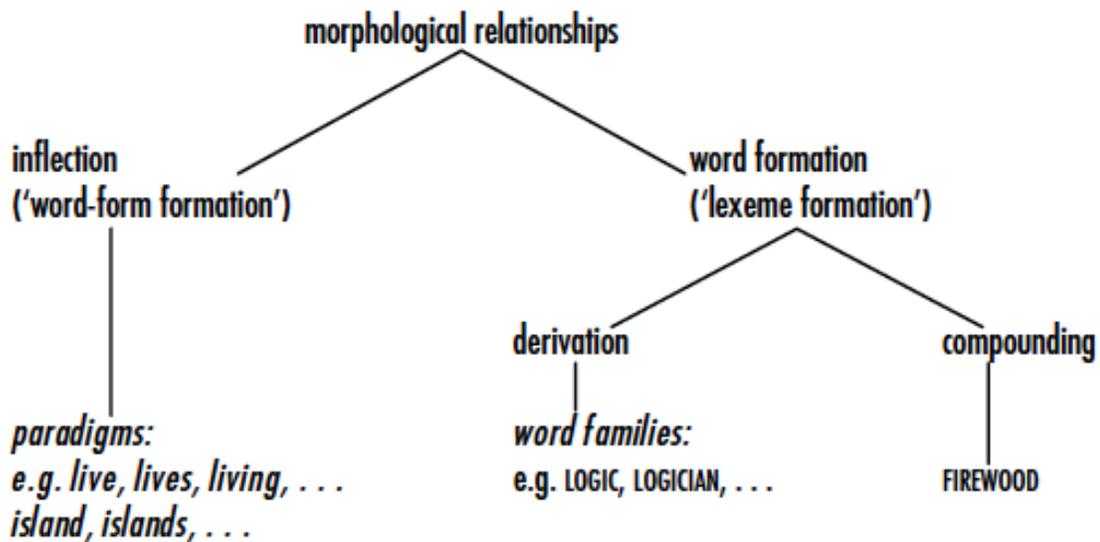


Figura 1. Principales subdivisiones de la morfología (Haspelmath y Sims, 2010: 19)

Según Bauer (2004), la formación de palabras tiene dos funciones principales: a) el enriquecimiento del léxico mediante la acuñación de términos que denotan conceptos nuevos o relevantes; b) la transposición de los lexemas a una categoría gramatical diferente para que un mismo significado

pueda cumplir nuevas funciones dentro de la oración (por ejemplo, que un sustantivo se convierta en un verbo y/o un adjetivo).

1.1. La flexión

La morfología flexiva es tradicionalmente descrita como aquella área que estudia las diferentes formas de una palabra que son determinadas por la sintaxis. Es así que también ha sido descrita como un proceso gramatical en el que se combinan palabras y afijos para producir formas gramaticales alternativas adecuadas al contexto sintáctico en el que aparecen. La flexión es obligatoria, expresa nociones como número, persona, tiempo, aspectualidad, modo (*mood*), grado y género entre otros. Además, se distingue de la derivación en que la incorporación de afijos no cambia la categoría gramatical de la palabra. Por ejemplo, al agregar el sufijo plural *-s* a un sustantivo, este no altera su categoría: *dog* → *dogs*. Lo mismo ocurre con las formas verbales: *work* → *works*–*worked*. En inglés, la flexión es siempre expresada por sufijos, de los que existen ocho: dos para sustantivos; cuatro para verbos y dos para adjetivos. La siguiente tabla los detalla:

Tabla 1. Morfemas flexivos del inglés

| | | | |
|-------------|------------------|-----------------------------------|-------------------------------|
| SUSTANTIVOS | <i>-s, -es</i> | Pluralidad | <i>dogs, horses, boxes</i> |
| | <i>-’s, -’</i> | Posesión | <i>Tom’s, the boys’ shoes</i> |
| VERBOS | <i>-s, -es</i> | 3era persona singular presente | <i>works, changes, cries</i> |
| | <i>-ed, -d</i> | Pasado simple / participio pasado | <i>worked, changed, cried</i> |
| | <i>-ing</i> | Participio presente | <i>working, changing</i> |
| | <i>-en, -n</i> | Participio pasado | <i>worked, eaten, taken</i> |
| ADJETIVOS | <i>-er, -r</i> | Comparativo | <i>shorter, whiter</i> |
| | <i>-est, -st</i> | Superlativo | <i>shortest, whitest</i> |

En cuanto a las formas irregulares, los sufijos varían bastante: incluyen plurales de origen clásico (*cactus* → *cacti*; *criteria* → *criterion*; *formula* → *formulae*) y otros como *child* → *children*; *ox* →

oxen. Casos como *brother* → *brethren*; *tooth* → *teeth* reflejan una modificación interna muy común también en verbos irregulares (*grow* → *grew*; *meet* → *met*). Las alteraciones a la raíz de la palabra que implican un cambio vocálico se describen como *ablaut* o apofonía. Los adjetivos registran muy pocas formas comparativas y superlativas irregulares: *good* → *better/best*, *bad* → *worse/worst* y *far* que tiene dos posibles realizaciones: *farther/farthest* o *further/furthest*.

El orden en que los sufijos derivacionales y flexivos se agregan a la base reflejan la relevancia semántica que tienen para la raíz. En general, las flexiones se añaden después de las derivaciones puesto que su relevancia es sintáctica. Por ejemplo, el verbo *agree* primero se nominaliza mediante el sufijo *-ment* (*agreement*) y luego incorpora el sufijo plural *-s* (*agreements*). Lo mismo ocurre con las verbalizaciones: primeramente se deriva el verbo a partir del adjetivo añadiendo el sufijo *-ify* (*pure* → *purify*) y posteriormente se agregan las flexiones verbales correspondientes (*purifies-purifying-purified*). Esto se explica porque primero hay que derivar el nuevo lexema para después poder crear sus distintas formas gramaticales. Se observa, entonces, que flexión y derivación son dos procesos morfológicos que están muy relacionados: “Inflection and derivation are two different, yet related and even overlapping concepts” (Naumann y Vogel, 2000: 929).

Esto último queda de manifiesto en el hecho de que, no obstante lo anterior, hay casos en los que los elementos flexivos preceden en orden a los derivativos. El sustantivo *connectedness* se obtiene al agregar el sufijo nominal *-ness* a una base que ya tiene una flexión (*-ed*). Algunos teóricos postulan que los procesos morfológicos tienen lugar en distintos niveles o estratos (*Level Ordered Morphology*) y distinguen entre dos tipos de afijos. En consecuencia, un sufijo clase I como *-ness* puede agregarse a una base que contiene un sufijo clase II como *-ed* (Lang, 1990: 50). Por su parte, Booij (2006) lo atribuye al hecho de que el participio pasado en inglés posee características tanto verbales como adjetivales, lo que le permite servir de base para nominalizaciones que añaden *-ness*. De esta manera queda demostrado que estas formas verbales –los participios– pueden servir de input en procesos derivacionales, que involucran tanto bases complejas flexionadas (*excited* → *excitedly*) como compuestas (*close-mouthed* → *close-mouthedness*).

1.2. La derivación

La derivación es un proceso morfológico que consiste en crear palabras complejas mediante la incorporación de un afijo a una base léxica. En otras palabras, se combinan morfemas léxicos con morfemas gramaticales dependientes (*bound morphemes*) lo que produce cambios sistemáticos y regulares de forma y significado en un mismo lexema. Al igual como sucede con los otros procesos de formación de palabras, la función principal de la derivación es contribuir a la expansión del vocabulario de una lengua, lo que consigue modificando el significado de lexemas ya existentes. Desde una perspectiva histórica, Naumann y Vogel (2000) distinguen varios procesos derivacionales, a continuación mencionamos tres de ellos que son relevantes para el inglés:

- 1) La adición de afijos, también llamada “derivación explícita”, incluye las operaciones más comunes como la prefijación y la sufijación. La reduplicación es considerada un tipo especial de derivación de tipo aditivo puesto que incorpora la totalidad (*job-job*) o parte de la base léxica (*walkie-talkie*).
- 2) La derivación cero o conversión en la que la transposición de una categoría gramatical a otra no entraña ningún cambio en la forma de la palabra base, como es el caso del sustantivo *look* y el verbo (*to*) *look*.
- 3) La sustracción de elementos (*backformation*) es un proceso diacrónico bastante raro que consiste en la eliminación de una parte de la base que es erróneamente percibida como un sufijo: *editor* → *edit*.

Tanto la reduplicación como la conversión y la sustracción de elementos son descritas en más detalle en la sección sobre procesos de formación de palabras (cf. 1.6.).

Mientras que la flexión es importante para la sintaxis, la derivación lo es para el lexicón. Los afijos derivacionales poseen contenido léxico, lo que les permite expresar una gran variedad de nociones

semánticas. Por ejemplo, los sufijos pueden denotar acción/estado (*stupidity*), característica/propiedad (*careful*), agente (*trainer*), paciente (*trainee*), instrumento (*blender*), modo/manera (*stupidly*), ausencia/carencia (*careless; childless*), similitud (*greenish; businesslike*) y varios otros. Los prefijos, por su parte, pueden denotar cuantificación (*bi-, tri-, mega-, micro-, poly-*, etc.), negación u oposición (*dis-, mis-, un-, anti-, counter-*, etc.), espacialidad (*sub-, trans-, circum-, intra-*, etc.), temporalidad (*mid-, pre-, post-, retro-*, etc.) y otras como repetición (*re-*) y reflexividad (*auto-*). Por esta razón son más relevantes para el significado del lexema base y, por lo tanto, se ubican más cerca de la raíz. Por el contrario, los afijos flexivos son más periféricos. Con respecto a los adverbios, cabe destacar que algunos especialistas cuestionan las descripciones más tradicionales, al sostener que el sufijo *-ly* en función adverbial no debería ser clasificado como derivacional ya que despliega características propias de la flexión. En consecuencia, argumentan que adverbios como *quickly* o *stupidly* son variantes de sus formas adjetivas determinadas por la posición que ocupan dentro de la oración (Plag, 2011: 15).

En inglés, la derivación se restringe principalmente a las tres grandes categorías sintácticas: sustantivos, adjetivos y verbos, los que pueden incorporar sufijos de manera bastante libre para derivar nuevas palabras. Por otra parte, las transposiciones hacia adverbios son menos frecuentes. Como se ha mencionado, en la mayoría de los casos, la derivación supone una transposición de una categoría gramatical a otra. Sin embargo, esto parece circunscribirse a los sufijos. En general, los prefijos no afectan la clase sintáctica de las palabras a las que se agregan: pares de adjetivos como *clean/unclean*, de verbos como *write/rewrite* y de sustantivos como *belief/disbelief* así lo demuestran. La excepción son dos prefijos que forman verbos: *-en* y *-be* (*slave* → *enslave*; *friend* → *befriend*), los que son considerados improductivos en inglés moderno. De igual modo existen sufijos nominales que convierten un sustantivo concreto en uno abstracto como es el caso de *-hood*, *-dom* y *-ship* (*childhood; martyrdom; membership*) y otros que marcan género como *-ess* (*host/hostess; tiger/tigress*), o expresan tamaño y/o evaluación como *-let*, *-ette* (*booklet; kitchenette, starlet*). Este tipo de sufijos introduce una especificación semántica al lexema manteniendo inalterable la categoría base.

Los sufijos, junto con determinar la categoría gramatical de la forma derivada, producen ciertos cambios a nivel morfológico y fonológico como la elisión de segmentos finales en la palabra base y la modificación de la sílaba acentuada, respectivamente (Plag: 2011). La sufijación es un mecanismo recursivo, por lo tanto, cada nueva forma puede servir de base para otras derivaciones. Por ejemplo, el sustantivo *care* (cuidado; atención) puede incorporar el sufijo *-less* y transformarse en el adjetivo *careless* (descuidado/-a; imprudente) el que, a su vez, puede agregar el sufijo nominal *-ness* y producir el sustantivo *carelessness* (descuido; imprudencia). Desde una perspectiva cognitiva, las bases se procesan antes que las derivaciones, pero existe una “preferencia sufijativa” dado que al momento de procesar afijos, el orden de reconocimiento es final, principio y mitad de palabra, es decir, sufijo, prefijo e infijo (Naumann y Vogel, 2000: 941).

En clasificaciones formales, se distinguen diferentes tipos de procesos derivacionales según la naturaleza de las bases léxicas, las que pueden ser simples o complejas:

- a) Bases simples (*arrive* → *arrival*; *happy* → *unhappy*)

- b) Bases complejas:
 - a. Formas derivadas (*careless* → *carelessness*; *accessible* → *inaccessible*)
 - b. Formas flexivas (*connected* → *connectedness*; *excited* → *excitedness*)
 - c. Formas compuestas (*open-minded* → *open-mindedness* / *open-mindedly*)
 - d. Frases lexicalizadas (*good-for-nothing* → *good-for-nothingness*)

1.2.1. Los afijos derivacionales en los diccionarios

La derivación (*derivation*) es por lejos el proceso de formación de palabras en lengua inglesa más activo y productivo. Yule (2010) destaca que, aun cuando existe un gran número de afijos (*mis-*, *pre-*, *un-*, *-less*, *-ful*, *-ness*, *-ish*, etc.), estos generalmente no figuran por separado en los diccionarios aunque algunos de ellos tienen uno o más significados bien definidos. Sin embargo, una revisión

somera de algunos diccionarios modernos en línea demuestra que estos sí incorporan las definiciones de algunos prefijos, las que varían en complejidad y detalle:

1. **mis-** (prefix): *added to the beginning of a verb or a word formed from a verb, to show that the action referred to by the verb has been done wrongly or badly.*

(<http://dictionary.cambridge.org/>)

2. **mis-** prefix

1 a: badly: wrongly <*misjudge*>

b: unfavorably <*misesteem*>

c: in a suspicious manner <*misdoubt*>

2: bad: wrong <*misdeed*>

3: opposite or lack of <*mistrust*>

4: not <*misknow*>

(<http://www.merriam-webster.com/>)

Como se puede observar, la primera definición es bastante general y solo da cuenta del uso del prefijo con formas verbales. La segunda definición, por otra parte, no solo detalla varios significados sino que los asocia tanto a verbos (1a, 1b, 1c, 4) como a sustantivos (2, 3).

Algo similar ocurre con respecto a los sufijos. Tomemos como ejemplo *-ful*:

1. **-ful** suffix (HAVING)

Having the stated quality to a high degree, or causing it:

colourful

powerful

painful

beautiful

(<http://dictionary.cambridge.org/>)

2. **-ful**

1. adjective suffix

: full of <prideful>

: characterized by <peaceful>

: having the qualities of <masterful>

: tending, given, or liable to <helpful>

2a. noun suffix

: number or quantity that fills or would fill <roomful>

(<http://www.merriam-webster.com/>)

3. **-ful** [in adjectives]

1 full of something:

an eventful day

2 having the quality of something or causing something:

restful colours

Is it painful?

-fully suffix [in adverbs]

shouting cheerfully

3a. **-ful** [in nouns]

1 the amount of a substance needed to fill a particular container:

two cupfuls of milk

2 as much as can be carried by or contained in a particular part of the body:

an armful of flowers

(<http://www.ldoceonline.com/>)

4. *-ful*

suffix

1 (forming adjectives from nouns) full of: *sorrowful*

-having the qualities of: *masterful*

2 forming adjectives from adjectives or from Latin stems with little change of sense: *grateful*

3 (forming adjectives from verbs) apt to; able to; accustomed to:

forgetful, watchful

4 (plural *-fuls*) forming nouns denoting the amount needed to fill the specified container, holder, etc.: *bucketful, handful*

Origin: from FULL

(<http://www.oxforddictionaries.com/es/>)

Una vez más, son apreciables las diferencias en cuanto al detalle de la descripción y uso del sufijo. De las cuatro definiciones, la primera es la más básica puesto que ofrece el significado general y distintos ejemplos, sin distinguir la categoría gramatical de los lexemas a los que puede agregarse. Las otras tres, sin embargo, van complejizando la descripción tanto a nivel semántico como morfo-sintáctico, especificando los significados según el tipo de palabra a la que se agrega el sufijo (adjetivo o sustantivo). De hecho, el *Longman Dictionary of Contemporary English* (definición 3) tiene dos entradas para el sufijo *-ful*, una para adjetivos y otra para sustantivos. La definición número 4 entrega una descripción morfológica bastante prolija en cuanto a:

- El tipo de lexemas que pueden incorporar *-ful*
- El origen de las bases a las que puede agregarse (latinas)
- El tipo de palabras que se crean mediante su incorporación

Además de listar sus diferentes significados, especifica el origen léxico del sufijo (el adjetivo *full*). El proceso de gramaticalización de *-ful* es detallado más adelante. (cf. sección 1.5.1.)

1.3. Similitudes entre derivación y composición

Desde el punto de vista de su estructura, existen similitudes entre los dos procesos: ambos son de naturaleza combinatoria, por lo tanto en ambos hay un encadenamiento de elementos que pueden ser analizados sintagmáticamente para determinar las relaciones entre ellos. También se asemejan en que, a veces, el significado global puede ser derivado directamente del significado de sus constituyentes como en el caso de palabras que contienen los sufijos *-ful* y *-worthy* (*helpful*, *trustworthy*). En el aspecto semántico, la derivación puede expresar una amplia gama de categorías semánticas, otro rasgo que la aproxima más a la composición que a la flexión (Booij, 2006).

Derivación y composición comparten una serie de propiedades formales:

- a) Poseen un núcleo
- b) Son estructuras binarias
- c) Son recursivas

La regla que establece que el núcleo de una palabra compleja se ubica a la derecha se aplica tanto a palabras derivadas como compuestas. En el caso de las formas derivadas, el elemento a la derecha es un afijo. Parece lógico asignarle un rol central a los sufijos dado que determinan la categoría sintáctica del lexema, lo que es aplicable incluso a aquellos sufijos que no efectúan ningún cambio. De esta manera *childhood* es clasificada como sustantivo sobre la base de la categoría asignada al sufijo *-hood*. La estructura interna de un derivado como *carelessness* revela que los elementos se agrupan jerárquicamente en forma binaria:

$$[\textit{carelessness}] \rightarrow [[\textit{careless}] \textit{ness}] \rightarrow [[\textit{care}] \textit{less}] \textit{ness}]$$

Finalmente, la presencia de una forma derivada como base de otra demuestra que son recursivas. Por su parte, el análisis de un compuesto como *peanut butter sandwich* también deja de manifiesto las mismas características: núcleo a la derecha (*sandwich*) que determina la categoría léxica de toda la

estructura (sustantivo), constituyentes agrupados en pares (binariedad) y un compuesto dentro de otro (recursividad):

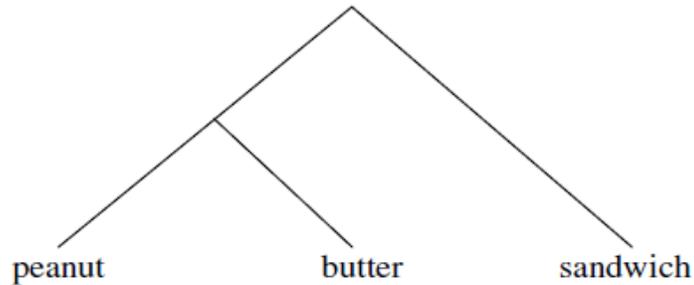


Figura 2. Recursividad y binariedad en un compuesto
(Arcodia, Grandi y Montermini, 2009: 16)

1.4. Restricciones que afectan a la derivación y la composición

Otro rasgo que comparten la derivación y la composición es el hecho de verse limitadas por una serie de restricciones de distinta naturaleza. A continuación se mencionan algunas de ellas, con especial énfasis en las limitaciones que conciernen a los adjetivos compuestos con participio.

1.4.1. Restricciones morfológicas

En el proceso de derivación hay casos en los que la estructura morfológica de la base determina el sufijo que puede añadirse. Esto es lo que sucede con virtualmente todos los verbos terminados en el sufijo *-ize* los que pueden ser sustantivados mediante la incorporación del sufijo *-ation*. Esta restricción descarta otras posibilidades por lo que los sustantivos derivados de *colonize* y *sterilize* solo pueden ser *colonization* y *sterilization* excluyendo formas potenciales como **colonizement* o **sterilizage*. De igual manera, los verbos que contienen un morfema (raíz) dependiente (*bound root*) como *-ceive* (*receive*, *perceive*), *-mit* (*permit*, *admit*) y *-duce* (*reduce*, *produce*) solamente pueden formar sustantivos terminados en *-ception* (*reception*, *perception*), *-sion* (*permission*, *admission*) y *-duction* (*reduction*, *production*), respectivamente. En lo que respecta a las estructuras compuestas,

Olsen (2000: 898-899) señala que no aceptan la inserción de otro constituyente ni tampoco flexiones plurales en el primer elemento. Sin embargo, otros autores (Spencer, 1991, Bauer, 2006 y Blevins, 2006) discrepan de esta última aseveración. De hecho, en los corpus analizados se verificó la ocurrencia de plurales internos (regulares e irregulares) en compuestos como *rights-respecting*, *sales-related* y *teeth-whitening products* (cf. sección 7.1). Además de lo anterior, no parece haber restricciones respecto a las características morfológicas del primer elemento. Como veremos más adelante, la composición interactúa con prácticamente todos los otros procesos de formación de palabras en la generación de nuevos ítems léxicos (cf. sección 7.4).

1.4.2. Restricciones sintácticas

Los procesos de formación de palabras también se restringen a categorías sintácticas específicas. Entre los ejemplos más conocidos está la regla que impide que sustantivos y adjetivos añadan el sufijo *-able*, por lo que derivaciones como *heart* → **heartable* o *happy* → **happiable* son imposibles. Este sufijo solamente puede ser incorporado a verbos como *read* → *readable*, *wash* → *washable*, etc. Entre las restricciones que afectan la construcción de adjetivos compuestos con participio, Roeper y Siegel (1978) sostienen que las categorías sintácticas que pueden intervenir en la creación de un compuesto de base verbal son aquellas que, a nivel oracional, ocupan la posición inmediatamente contigua al verbo (*First Sister Principle*). Por lo tanto, estas solo pueden ser sustantivos que actúan como objeto directo del núcleo (*plant-eating*) o como adjunto (*rain-washed*, *sea-going*), adjetivos que son complementos del sujeto (*fragile-looking*, *nice sounding*) y adverbios (*badly-played*, *fast-growing*).

1.4.3. Restricciones semánticas

En restricciones de este tipo, un rasgo semántico específico se transforma en un prerrequisito para un proceso de formación de palabras. Un ejemplo es el sufijo *-ee* el que solo puede referirse a entidades conscientes (*sentient*) que denotan al objeto de la acción o paciente (cf. agente): *interviewee* (entrevistado/a), *employee* (empleado/a). Esta regla hace que *amputee* solo pueda ser interpretado como la persona a la que se la ha amputado una extremidad y nunca como la

extremidad misma. En el caso de adjetivos compuestos como *three-legged* (de tres piernas) o *red-roofed* (de techo rojo) el prerequisite es que la base del segundo elemento del compuesto denote posesión no enajenable, es decir, que sea un atributo inherente del sustantivo modificado por ese adjetivo (cf. sección 3.3.1.)

1.5. Los límites entre derivación y composición

El estatus de los elementos constituyentes parece marcar la diferencia principal entre los dos grandes procesos de formación de palabras en inglés. Mientras la derivación es la combinación de raíces o bases con morfemas gramaticales (afijos), la composición es el proceso que combina lexemas independientes. Olsen (2014: 27) lo explica así:

[...] the primary difference between affixation and compounding lies in the status of the constituent parts: if at the relevant level of analysis both constituents are lexemes belonging to the open word classes of the language, the result is a compound, if one constituent is a formative, that is, a bound morpheme belonging to a finite class of elements in the language, the structure is an affixation”.

Entonces, una diferencia fundamental entre derivación y composición es la independencia de los elementos constituyentes. Por otra parte, al especificar que las palabras al interior de un compuesto deben pertenecer a una categoría léxica (sustantivos, adjetivos, verbos o adverbios), lexemas como *extra-terrestrial*, *post-election* y *semi-detached*, cuyo primer elemento es un afijo, debieron ser eliminados. Entre los datos recolectados para esta investigación encontramos muchísimos casos que incluían elementos de esta naturaleza, es decir, compuestos cuyo primer constituyente es un prefijo, por lo tanto, se hizo necesario realizar una revisión y depuración acabada del corpus electrónico. Sin embargo, durante el proceso de revisión y limpieza de ambos corpus verificamos la ocurrencia de un par de compuestos sinónimos cuyas estructuras incorporan tanto un prefijo como un lexema: *multi-colored* y *many-colored*. Ateniéndonos al criterio de independencia, se descartó el primero de ellos (*multi-colored*) y se mantuvo el segundo. Por otra parte, en casos como *downright*, *overreact*, *undergo*, *backfire* y *upcoming* el estatus del primer constituyente es más confuso puesto que *down*, *over*, *under*, *back* y *up* pueden ser clasificados como prefijos y en otras ocasiones como adverbios o

partículas preposicionales. Nos enfrentamos a un problema parecido con los verbos frásticos (*phrasal verbs*) en inglés, los que, a pesar de comportarse semánticamente como unidad, poseen una serie de rasgos que complican su clasificación. En primer lugar, uno de sus constituyentes puede ser ambiguamente interpretado como partícula o adverbio. En segundo lugar, estos verbos son morfológicamente irregulares ya que verbo y partícula pueden ser separados como en *fill in the form / fill the form in*, lo que viola una propiedad fundamental de una palabra, ya sea compuesta o no: la indivisibilidad. Además, su núcleo se ubica a la izquierda, lo que no es un rasgo propio de las estructuras compuestas. En consecuencia, representan un problema particularmente difícil de resolver puesto que se encuentran en los márgenes entre sintaxis y morfología.

En nuestra investigación nos ceñimos al criterio de pertenencia a una clase gramatical de tipo léxico y no funcional (preposiciones, conjunciones, artículos, etc.) para determinar qué palabras deben ser clasificadas como compuestos. Estos son solo dos ejemplos de las dificultades que los analistas enfrentan al momento de distinguir entre composición y derivación. A continuación se discuten otros problemas de diversa complejidad y envergadura.

1.5.1. Afijos y semi-afijos

Otro problema representativo de lo difusos que pueden ser los límites entre composición y derivación se observa en construcciones como *businesslike* y *trustworthy*, las que a primera vista parecen ser compuestos dado el paralelismo en forma y significado con los lexemas *like* y *worthy*. Sin embargo, el hecho de que pueden incorporar un prefijo negativo (*unbusinesslike*, *untrustworthy*), algo que los compuestos no permiten, comprueba que son formas derivadas y demuestra claramente el estatus sufijal de *-like* y *-worthy*. Así como hay afijos que todavía son percibidos como formas independientes, en inglés también existen muchas palabras que son elementos recurrentes en la creación de compuestos (*-craft*, *-free*, *-proof*, *-friendly*,) que comienzan a ser percibidas como sufijos, lo que dificulta su descripción morfológica. Estos son los llamados semi-afijos o afixoides (*affixoids*). Olsen (2000: 905) explica que históricamente muchos afijos comienzan siendo palabras independientes que, al funcionar como elementos de un compuesto, pueden dar origen a una serie de

compuestos similares. Entonces, gradualmente empiezan a desviarse en forma o significado del lexema original y a tomar características de afijo. Un afixoide es definido como un elemento léxico que se encuentra en una etapa de transición entre constituyente de un compuesto y un morfema derivacional. Los especialistas concuerdan en que para ser considerado como tal, un semi-afijo debe cumplir con dos requisitos básicos:

- 1) Debe ser parte de una serie de formas similares.
- 2) Su significado debe experimentar algún grado de variación semántica respecto de su equivalente léxico independiente.

Un ejemplo de semi-afijo bastante productivo es *-free* el que participa de palabras como *tax-free*, *stress-free*, *smoke-free*, *fat-free*, *pain-free*, etc. No obstante lo anterior, esta es una categoría morfológica que no del todo aceptada. En un artículo posterior, esta misma autora (Olsen, 2014) reconoce que así como no existe consenso entre los especialistas al momento de clasificar una determinada forma como afijo o lexema, tampoco lo hay respecto a lo que constituye un afixoide y cuestiona derechamente su utilidad. Además, los criterios postulados no parecen ser suficientemente específicos. Por una parte, el hecho de contribuir productivamente en la formación de palabras no significa que un lexema haya comenzado a desarrollar características de afijo. Simplemente, ciertos lexemas son más combinables que otros, por lo que los compuestos tienden a agruparse en familias que tienen un mismo lexema central, como es el caso de *-free*. Por otra parte, es ciertamente difícil decidir en qué medida deben distanciarse semánticamente el semi-afijo y la variante independiente. Olsen (2014: 31-33) pone como ejemplo *-friendly* el que puede aparecer en muchas combinaciones del tipo *US-friendly* (que muestra simpatía hacia los Estados Unidos) en la que se retiene el significado central del vocablo independiente. Sin embargo, hay casos como *user-friendly*, *consumer-friendly* o *environment-friendly* cuya interpretación semántica es diferente (fácil para, favorable a, respetuoso con). No obstante, esta extensión semántica sigue estando asociada al significado original, por lo que *-friendly* tampoco amerita ser clasificado como un afixoide.

En efecto, las alteraciones de orden semántico son frecuentes al combinar lexemas, es decir, en la composición. No es extraño que una estructura compuesta desarrolle características propias y pierda algo de transparencia, lo que puede conducir a una independencia semántica de uno de los constituyentes del compuesto (*-friendly*) con respecto al adjetivo *friendly*. De esta manera, el lexicón mental debe incorporar una entrada separada para la variante dependiente al interior del compuesto (*positionally bound compound constituent*). A nivel cognitivo, esto facilita el procesamiento de compuestos en los que un elemento constituyente ha establecido una separación semántica ya que hace innecesaria la activación de significado de su equivalente léxico. El distanciamiento entre los significados de constituyentes y lexemas es un fenómeno común en la composición y no es indicativo del comienzo de un proceso de lexicalización que lleve a la aparición de un nuevo afijo. En conclusión, es mejor evitar esta categoría intermedia (afixoide) puesto que no ayuda a definir los límites entre afijos y lexemas y, por lo además, representa una complicación adicional para estudios de naturaleza sincrónica, ya que solo una perspectiva diacrónica puede resolver el estatus morfológico de estas formas. En consecuencia, el término “afijo” debe reservarse para formantes ligados que no compiten con lexemas independientes: “[...] the term affix should be reserved for reference to a pure formative, that is, a bound morpheme for which there is no competition with a free lexeme in the mental lexicon, and the term semi-suffix is best avoided.” (Olsen, 2014: 34)

Un ejemplo de evolución histórica de lexema a sufijo lo constituye *-ful* en unidades de medida como *spoonful* (cucharada), *cupful* (taza), *handful* (puñado), etc. Estas construcciones partitivas se desarrollaron a partir de frases nominales del tipo *two cups full of rice* cuyo núcleo designa un contenedor o recipiente (*cups*) post-modificado por una frase adjetiva (*full of rice*). Con el tiempo, la ortografía del adjetivo se redujo a *ful* y la flexión plural se trasladó del sustantivo al adjetivo (*fuls*). Finalmente, sustantivo y adjetivo comenzaron a escribirse juntos (*cupfuls*). De esta manera *full* dio origen a un elemento dependiente con rasgos nominales que resultó particularmente productivo y que cumple una función léxica específica (Olsen, 2014: 28).

1.5.2. Elementos compositivos

Una categoría de formas dependientes (*bound forms*) que solo ocurren como parte de palabras complejas –ya sean derivadas o compuestas– son los denominados elementos compositivos (*combining forms*). Se diferencian de los afijos en cuanto a origen histórico puesto que muchos de ellos son préstamos lingüísticos del latín o el griego (*electr(o)* o *cephal*), lenguas en las que, además, solían ser palabras independientes. Es por esta razón que en morfología española también se les llama temas cultos o formantes clásicos o grecolatinos. Sin embargo, tampoco pueden ser considerados raíces ligadas (*bound roots*) ya que la mayoría de ellos no puede agregar afijos directamente, sino que al entrar en combinación. En efecto, uno de sus rasgos distintivos es la capacidad que tienen para combinarse entre ellos como en *philosophy*, *necropsy* o *telescope*, vocablos en los que ninguno de los constituyentes puede funcionar de manera autónoma. Esta propiedad establece una diferencia crucial con los otros elementos morfológicos: así como no es posible combinar dos raíces ligadas, un afijo tampoco puede combinarse con otro afijo para formar una palabra nueva. En otras palabras, un lexema no puede estar compuesto solo de afijos careciendo de una raíz. Es por esto que en morfología inglesa estas combinaciones son denominadas “compuestos neoclásicos” (*neo-classical compounds*) o bases compositivas cultas. Por añadidura, los elementos compositivos poseen mayor contenido semántico que los afijos y, de hecho, algunos de ellos se comportan como raíces pudiendo incorporarlos directamente, como en el caso de *electro* y *cephal*, dando origen a *electricity*, *electrify*, *macrocephalia* y *acephalous*, respectivamente. Bauer (1983), lista algunos ejemplos de elementos compositivos que han alcanzado estatus de independencia como es el caso de *-naut* en *nautic* y *phobe* en *phobia*, *phobic* y *phobism*, por lo que a veces se les considera raíces ligadas. Sin embargo, constituyen excepciones a la regla y no una tendencia, por lo tanto propone mantener su descripción como elementos compositivos:

“My own preferred solution is to accept these combining forms for what they are etymologically: elements of the classical languages which are used in English word-formation. [...] It is because these elements are put together by speakers of English that it is possible for coiners to mix Greek and Latin as in *television*.” (Bauer, 1983: 216)

Cabe destacar que la mayoría de estas formas no se originaron en la antigüedad clásica sino que siglos después. Un ejemplo emanado del latín del medioevo es *Anglo*, el que generalmente aparece en compuestos referidos a grupos étnicos (*Anglo-Saxon*, *Anglo-American*, *Anglo-Japanese*, etc.) y en otros como *anglophone* (de habla inglesa), *anglophile* (anglófilo) y *anglophobe* (anglófobo). También puede agumentarse que *Anglo* es una raíz independiente en formas como *Anglia*, *Anglic* y *Anglo-allied* aunque es preciso señalar que esta forma no debe confundirse con el lexema independiente *Anglo* (sustantivo, plural *Anglos*) que data de 1800 y que hace referencia a los habitantes de raza blanca de los Estados Unidos. Por esta razón, solo el contexto nos ayudará a interpretar correctamente los compuestos *Anglo-looking* y *Anglo-allied* encontrados en el corpus electrónico.

Desde los siglos XVII y XVIII la productividad de este tipo de constituyentes se ha visto particularmente incrementada en áreas científicas y tecnológicas, en las que nuevas áreas de estudio, descubrimientos e inventos muchas veces conllevan la incorporación de elementos compositivos neoclásicos al momento de acuñar un término. En realidad, continúan siendo una fuente activa de creación léxica en muchas lenguas europeas además del inglés. Otro ejemplo, mucho más moderno y de origen desconocido, es *jazzophile* (fanático del jazz) que ilustra el hecho de que todavía son productivos y, por otra parte, la capacidad que tienen los elementos compositivos de dar origen a nuevos vocablos mediante analogía con otras palabras. De hecho, *jazzophile* demuestra la flexibilidad que poseen estas formas para combinarse incluso con lexemas de la lengua nativa como *astro-dog* y *musicology*. Plag (2003: 156) ilustra sus propiedades combinatorias de la siguiente manera:

Elemento compositivo + raíz ligada = *glaciology*, *scientology*

Dos elementos compositivos = *morphology*, *hydrology*

Elemento compositivo + lexema = *morpho-syntax*, *bio-warfare*, *lazyitis*

El último ejemplo (*lazyitis*) ilustra una tendencia que se ha verificado en varios idiomas respecto a una extensión de significado de las bases compositivas *-itis* y *-ology* en usos cotidianos y/o

humorísticos. Esta “banalización” de elementos cultos de origen clásico se evidencia en *cellphon(e)itis* y su equivalente alemán *Telefonitis*. En español usamos términos como “mamitis” y “suegritis” para describir un apego excesivo y de igual forma ocurre en inglés, aunque solo el primero de ellos está descrito (*motheritis, momitis*). En inglés, *-ology* se utiliza para referirse a áreas de conocimiento práctico como *beerology, wineology* y *sandwichology*. En español, esta tendencia se puede apreciar en vocablos como “rumerología” y “discotecología”.

A esta descripción Bauer (1998) agrega otros patrones combinatorios como:

Truncamiento o reducción + elemento compositivo = *Eurocrat, Amerophile, chocophile*

Acrónimo + elemento compositivo = *ufology, UFOmania*

Para Plag (2003: 74) la única diferencia entre compuestos “nativos” y los neoclásicos estriba en que los elementos compositivos siempre son dependientes: “The only difference between neoclassical forms and native compounds is that the non-native elements are obligatorily bound.”

En cuanto a los elementos compositivos de origen local (*native*), las descripciones no siempre concuerdan. Por ejemplo, *-fold*, un multiplicador que se agrega a números (*fivefold, tenfold, sixtyfold, hundredfold, thousandfold, millionfold, etc.*) es también descrito como sufijo. Sin embargo, el hecho de que pueda incorporarse a algunos cuantificadores (*manifold, severalfold*) y, más aun, que tenga la capacidad de adicionar un prefijo como en *multifold*, son claros indicadores de que todavía retiene características compositivas.

Entre los adjetivos compuestos con participio existe una categoría cuyo segundo elemento no es verbal sino que es derivado de un sustantivo que incorpora el sufijo *-ed* como *stone-walled* (con muros de piedra), *eight-petalled* (de ocho pétalos), *long-sleeved* (de manga larga). Desde una perspectiva semántica, muchos de los sustantivos que conforman este tipo de compuestos describen partes de un todo, el que puede ser una persona, un animal, una casa o parte de una construcción y otros objetos. En la mayoría de estos casos, el núcleo del compuesto es un elemento compositivo y no una forma libre. Esto se comprueba en los adjetivos compuestos que denotan una cualidad o

característica física, y/o un rasgo de carácter. En otras palabras, todos ellos describen una propiedad inherente (*inalianable*) como *almond-shaped eyes* (ojos almendrados), *bad-tempered* (de mal carácter) o *long-legged* (de piernas largas). El segundo elemento de estas estructuras es considerado un elemento compositivo, por lo tanto no puede ser usado de manera independiente como adjetivo. Debido a esto **shaped eyes*, **tempered men* o **legged dancers* no son posibles ya que la descripción es tan obvia que se vuelve inútil. Solamente la incorporación de un primer elemento que especifique una cualidad, un tipo o un número hace posibles esta clase de compuestos (cf. sección 3.3.1.).

1.6. Procesos de formación de palabras en inglés

Aun cuando algunos procesos como la acuñación y el préstamo lingüístico son universales, las lenguas poseen mecanismos propios para crear nuevas palabras, los cuales son bastante regulares dado que su funcionamiento obedece, la mayoría de las veces, a reglas precisas y claras. El hecho de que la lengua inglesa posea muy poca morfología flexiva se ve compensado por un amplio abanico de procedimientos de creación léxica que se valen principalmente de la composición (*compounding*), de la derivación (mediante el uso de prefijos y sufijos, rara vez de infijos) y de una serie de procesos morfológicos que excluyen totalmente los afijos (*non-affixational processes*) tales como la conversión, la síntesis y el entrecruzamiento (Plag, 2011).

A continuación se reseñan y ejemplifican procesos de formación de palabras¹ en inglés denominados “menores” (*minor word formation processes*) puesto que son menos frecuentes y productivos que la derivación y la composición. Se mencionan también algunos problemas de demarcación que surgen del análisis de algunas de estas formas, las que o bien despliegan características híbridas o se encuentran en los límites entre dos procesos:

¹ Para las traducciones de los diferentes procesos de formación de palabras se consultó el libro *Morfología Léxica: la formación de palabras*, de Soledad Varela y Santiago Fabregat (2005).

1.6.1. Acuñación léxica

La acuñación léxica (en inglés *coinage* y también *word manufacture*) es la creación de términos totalmente nuevos. Es particularmente común en áreas como la industria, la tecnología, la publicidad y el comercio las que actúan como verdaderas fuentes para la invención de nombres y marcas de productos, términos que luego se vuelven de uso cotidiano. Entre los ejemplos más antiguos están *aspirin*, *nylon* y *zipper*. Ejemplos más modernos incluyen *kleenex*, *xerox* (verbo) y *teflon*. Aunque en menor medida que la ciencia y la tecnología, la literatura también ha contribuido a ampliar el léxico del idioma inglés con formas tales como: *sensuous* (John Milton), *chortle* y *portmanteau* (Lewis Carroll), *blatant* (Edmund Spenser), *witticism* (John Dryden), *intensify* (Samuel Taylor Coleridge), *boredom* (Charles Dickens), *dreamscape* (Sylvia Plath), *beep* (Arthur C. Clark) y *cyberspace* (William Gibson). Algunos autores se valieron de su conocimiento de las lenguas clásicas para acuñar términos como *utopia* (Thomas More) y *pandemonium* (John Milton). Para describir al personaje principal de su novela *Lolita*, Vladimir Nabokov tomó la palabra *nymph* (de origen clásico) y agregó el diminutivo *-et* (del francés) para crear *nymphet* que el diccionario *Merriam-Webster* define como *a sexually precocious girl barely in her teens; also: a sexually attractive young woman*².

Esta categoría incluye los epónimos (*eponyms*) o palabras nuevas que hacen referencia a una persona o lugar tales como *hoover* (verbo y sustantivo), *fahrenheit*, *watt* y *sandwich*, este último en honor al Earl of Sandwich que vivió en el siglo XVIII y dio origen a este tipo de alimento (Yule, 2010).

1.6.2. Préstamo lingüístico

El préstamo o adopción lingüística (*borrowing*) consiste en tomar palabras de otras lenguas sin traducirlas. En toda lengua existen “vacíos” léxicos respecto a nuevas tecnologías y conceptos, por lo tanto los préstamos cumplen una doble función: demuestran que existe un vacío léxico en la

² Consultar <http://blog.oxforddictionaries.com/2016/02/invented-words/> y <https://www.dailywritingtips.com/25-words-coined-by-twentieth-century-authors/>

lengua de llegada y a la vez suplen esa carencia. El idioma inglés se distingue por haberse apropiado de palabras de muy diversa procedencia como: *lilac* (persa), *croissant* (francés), *tycoon* (japonés), *sofa* (árabe), *zebra* (bantú), etc.

El calco lingüístico (*calque*), por otra parte, representa un tipo particular de préstamo en el que una palabra extranjera es traducida literalmente al ingresar al otro idioma. Un ejemplo que cruza diferentes lenguas es *skyscraper*, el que posee varios calcos, algunos más fieles que otros: *gratte-ciel* (francés → “*scrape-sky*”), *wolkenkrabber* (holandés → “*cloud scratcher*”), *Wolkenkratzer* (alemán → “*cloud scraper*”) y la versión española → *rascacielo*. (Yule, 2010: 55)

1.6.3. Entrecruzamiento

El entrecruzamiento (*blending*) es la combinación de partes de dos palabras diferentes para crear un término nuevo. En inglés, generalmente se toma el principio de una palabra y el final de otra como en *smog* (*smoke* + *fog*), *brunch* (*breakfast* + *lunch*), *motel* (*motor* + *hotel*), *dramedy* (*drama* + *comedy*) y *Spanglish* (*Spanish* + *English*). En español tenemos ejemplos como *docudrama* (documental + drama) y *eurocracia* (europea + burocracia). En general, el entrecruzamiento puede ser descrito por la siguiente regla tomada de Plag (2003: 155):

$$\begin{array}{l} \text{Blending rule} \\ \text{A B + C D} \quad \rightarrow \quad \text{A D} \end{array}$$

Si las letras A, B, C y D representan las distintas partes que conforman las bases léxicas antes de ser “fusionadas”, observamos que en el resultado final hay una pérdida importante de material fonético y ortográfico. Sin embargo, se registran algunos casos cuya estructura difiere de la proyectada por la regla, como *beervana* (*beer* + *nirvana*) en la que la primera parte permanece intacta o *guesstimate* (*guess* + *estimate*) en el que la primera palabra está completa y la segunda pierde una mínima parte. También están los casos de *modem* (*modulator* + *demodulator*) y *sitcom* (*situational comedy*) en el que se combinan los comienzos de las bases, por lo tanto, su estructura es del tipo AC y no AD

como lo indica la regla (Plag, 2003: 155-157). Con posterioridad, el mismo autor (Plag, 2011: 19) amplió la descripción estableciendo dos tipos de entrecruzamientos que dan cuenta de todos los casos anteriores:

- a. $AB + CD \rightarrow AD$ (*breakfast + lunch \rightarrow brunch*)
- b. $AB + CD \rightarrow AC$ (*modulator + demodulator \rightarrow modem*)

Sin embargo, para Bauer (2006) las combinaciones del tipo AC no son entrecruzamientos sino que compuestos creados por truncamiento (*clipping*).

Olsen (2014) releva el hecho de que sin importar qué parte o cuánto se reduzca, es fundamental que los entrecruzamientos retengan el significado de los lexemas originales para poder interpretar correctamente el nuevo término. En otras palabras, el proceso de reducción debe calibrarse de tal manera que asegure que los oyentes o lectores puedan reconstruir ambos elementos constituyentes al encontrar ejemplos como *Twitizens* (*Twitter + citizens*), *stalkarazzi* (*stalker + paparazzi* = reporteros que “persiguen” a una celebridad a todas partes) o *rurban* (*rural + urban*). Plag (2011: 20) lo explica con claridad:

“The formal relationship between the bases and the blend is regulated by a compromise of two competing forces. One is the deletion of material to form a single word, the other the necessity to preserve as much material as possible to relate the blend to its bases.” (Plag, 2011: 20).

Sin embargo, la recuperación de las bases léxicas no siempre permite comprender con exactitud neologismos como *screenager* (*screen + teenager* = adolescente que pasa mucho tiempo en Internet o que muestra gran habilidad al usar una computadora) o *gundamentalist* (*gun + fundamentalist* = término peyorativo para referirse a un defensor acérrimo del derecho a portar armas en los Estados Unidos) puesto que incorporan elementos culturales o sociales no siempre deducibles en primera instancia.

1.6.4. Truncamiento

El truncamiento (*clipping*) también es denominado síntesis, síncope, acortamiento, reducción, o sustracción morfológica y, al igual que en el proceso anterior (*blending*), su característica más notoria es la reducción morfológica, puesto que consiste en la creación de palabras a partir de la apócope de otras. Este proceso generalmente comienza en el habla casual y se aplica a palabras polisilábicas las que, en la mayoría de los casos, se acortan a la sílaba inicial: *gas* (*gasoline*), *fan* (*fanatic*), *ad* (*advertisement*), *flu* (*influenza*), *plane* (*aeroplane*). También afecta a frases nominales *perm* (*permanent wave*), *pub* (*public house*) y nombres propios tanto de personas: *Ed* (*Edward*) o *Sam* (*Samuel* o *Samantha*) como de lugares: *Frisco* (*San Francisco*), *Derry* (*Londonderry*). Algunos ejemplos en español son: *metro*(politano), *bici*(cleta), *cole*(gio). De hecho, Yule (2010) hace notar que el entorno escolar es uno de los más prolíficos en este tipo de palabras: *chem*, *exam*, *gym*, *lab*, *math*, etc.

Bauer (2006) destaca el hecho de que el acortamiento de palabras no entraña la creación de nuevos lexemas con nuevos significados sino que produce lexemas con un valor estilístico nuevo que responde a variaciones de registro, grupo etario y/o de dialecto. Además, menciona la capacidad de unir estas formas y crear “compuestos recortados” (*clipping compounds*) en los que solo permanece la primera parte de ambas bases léxicas tales como: *psy ops* (*psychological operations*), *phys-ed* (*physical education*), *poli sci* (*political science*), *kidvid* (*kid's video*) o *cyborg* (*cybernetic organism*). También hay casos en los que la reducción se aplica solo a uno de los constituyentes como en *op art* (*optical art*) y *teletext* (*television text*). Debido a su similitud, los compuestos recortados muchas veces son confundidos con los entrecruzamientos (*blends*) por lo que a menudo son tratados y descritos como un mismo fenómeno. Bauer (2006: 501) los distingue con precisión:

“In a clipping compound, the first part of both words in the original is represented in the new form; in a blend the first part of the first word in the original and the last part of the second word in the original are represented in the new form.”

Los hipocorísticos (*hypocorisms*) representan un tipo particular de reducción muy extendida en el inglés de Australia y el Reino Unido. Consiste en acortar las palabras y las frases a una sílaba y agregar los sufijos *-ie* o *-y*. Este es el proceso que dio origen a *Aussie* (*Australian*), *telly* (*television*), *barbie* (*barbecue*), *movie* (*moving pictures*) y *hankie* (*handkerchief*). Bauer (2006) los denomina “*embellished clippings*” y restringe el término “hipocorístico” al acortamiento de nombres propios: *Chuck* (*Charles*), *Hal* (*Harry*), *Ned* (*Edward*), etc. De acuerdo a esta descripción, *Liz* sería la forma hipocorística de *Elizabeth* y *Lizzie* su variante “embellecida”. En español tenemos ejemplos como *Pato* (*Patricio*), *Pancho* (*Francisco*), *Pepe* (*José*) y *Charo* (*Rosario*) por mencionar algunos y podemos observar que, al igual que en inglés, no existen reglas fijas al momento de acortar los nombres y crear apodos cariñosos. De hecho, en el caso de *Pepe* y *José* el número de sílabas es incluso el mismo.

1.6.5. Formación regresiva

La formación regresiva (*backformation*) es un tipo de reducción especializada en el que un lexema, generalmente un sustantivo, es acortado para formar una palabra de otra categoría gramatical, usualmente un verbo, como en *television* → *televise*, *donation* → *donate*, *emotion* → *emote*, *babysitter* → *babysit*. Este último caso es un ejemplo representativo de uno de los patrones más comunes para la creación de verbos en inglés en el que sustantivos terminados en *-er* / *-or*, que hacen referencia a personas que desempeñan un trabajo u oficio, se reducen mediante este proceso para crear los verbos que denotan la actividad que realizan: *editor* → *edit*, *burglar* → *burgle*, *sculptor* → *sculpt*, etc. (Yule, 2010: 56-57). Lo interesante es que la formación regresiva es un proceso morfológico de naturaleza diacrónica muy escaso por lo que, desde una perspectiva sincrónica, a menudo es interpretada como un proceso de adición de afijos.

Olsen (2014) describe la formación regresiva como un proceso derivacional en el que, paradójicamente, en vez de agregar un afijo, el nuevo lexema es el resultado de la sustracción de lo que se percibe como un sufijo: *cohesion* → *cohes*, *seizure* → *seize*. Esto se debe a que la reducciones que afectan a las formaciones regresivas se realizan por analogía con palabras creadas

por derivación como *close* → *closure*. Además, tanto Olsen (2014) como Plag (2011) destacan que existe una relación muy productiva entre compuestos y la creación de formas regresivas verbales:

anger-management → *to anger-manage*

chain-smoker → *to chain-smoke*

house-cleaning → *to houseclean*

tailor-made → *to tailor-make*

Aunque los nuevos lexemas pueden ser denominados compuestos verbales, en estricto rigor son formas regresivas creadas a partir de un compuesto del cual son semánticamente dependientes. Para Olsen (2014: 45), esta restricción semántica es determinante al momento de establecer el límite entre ambos procesos de formación de palabras:

“A back-formed verb is a reduction of an existing complex form, hence its semantics will be based strictly on the meaning of the complex word; the constituents are not freely related to one another as in a compound.”

1.6.6. Analogía

La analogía (*analogy*) consiste en la creación de nuevas palabras tomando como modelo formas complejas ya existentes. A veces una palabra puede dar origen a toda una serie de nuevos lexemas como es el caso de *hamburger* (forma reducida de *Hamburger steak* = carne al estilo de Hamburgo), que al ser erróneamente analizada como un compuesto ha producido *beefburger* y *cheeseburger*, sus equivalentes marinos *fishburger* y *shrimpburger* (Plag, 2011: 18) y otros como *chickenburger* y *vegeb主rger/veggieburger*. Los entrecruzamientos (*blends*) también pueden servir como base para crear analogías: la combinación de *chocolate* y *alcoholic* produjo *chocoholic* (adicto al chocolate) el que, a su vez, a dado pie a diversas formas como *workaholic*, *foodaholic*, *bookaholic*, *shopaholic* y *spendaholic* (comprador y gastador compulsivo) y posiblemente continúe siendo un modelo para futuras locuciones como lo demuestran *loveholic*, la que sin estar lexicalizada registra 22

ocurrencias en el corpus *enTenTen12* y 628.000 resultados en *Google* o *coffaholic* que por ahora registra solo dos ocurrencias en dicho corpus pero 98.200 en internet. Respecto a su estatus morfológico, ambas formas (*-aholic/-oholic*) son clasificadas como sufijos en la mayoría de los diccionarios consultados en línea. El *Webster's New World College Dictionary* en línea (5ta edición, 2014), especifica que estos sufijos con frecuencia son usados en construcciones únicas o creadas especialmente para la ocasión (*occasionalisms* o *nonce constructions*). En efecto, una búsqueda simple en el corpus electrónico demuestra la libertad y originalidad con que los hablantes usan esta analogía para producir formas tan idiosincráticas como *sushi-aholic*, *pizza-aholic*, *Formula 1-aholic*, *LOST-aholic* (por la serie de televisión), *TV-aholic*, *jewelry-aholic*, y probablemente varias otras. Un ejemplo elocuente de este tipo de creación analógica lo constituye *medaholic*³ que es el nombre de un blog dedicado a profesionales de la medicina apasionados por su trabajo. El nombre fue cuidadosamente ideado, como lo detalla su creador:

“What is the definition of *medaholic*? The term “medaholic” was created by combining the word *medicine* with the *-holic* suffix, to imitate other *-holic* ending words (*alcoholic*, *workaholic*, *shopaholic*). The ending *-holic* usually suggests an addiction, compulsive need, or love for whatever subject matter. Therefore, *medaholic* is a person who is obsessed or focused on medicine.”

El significado conferido parece un poco arbitrario puesto que ya existe una definición, aunque quizás poco conocida: *Persona adicta a medicamentos con receta retenida*, por lo que el autor confía en popularizar la definición propuesta en su blog:

[...] an added bonus was the word *medaholic* was original. It returned only a few hits on Google. The term had no connotations associated with it. I could define the word as I liked and what it would mean.

³ <http://www.medaholic.com/the-meaning-behind-the-name-medaholic/>

Incluso explica la ortografía escogida:

Why is it *medaholic* and not *mediholic*? I actually gave the spelling considerable thought. [...] I ultimately settled with using an “a” because of convention. The words *workaholic* and *shopaholic* both use an “a” to connect the terms, so I eventually settled with the current spelling, *medaholic*.

No podemos predecir si en el futuro el uso del término *medaholic* se masificará o se usará para denotar algo diferente, quizás “*alguien adicto a los procedimientos quirúrgicos*”. Aun cuando no es usual encontrar el origen preciso de un neologismo y mucho menos una explicación pormenorizada de cómo se le dio forma y significado, creemos que los usuarios intuitivamente usan la analogía para crear palabras que pueden o no volverse populares.

Otro ejemplo de una forma emergente es *clothaholic* (4.090 resultados en *Google*) cuyo significado pareciera bastante simple: “*persona que compra ropa de manera excesiva*”. Sin embargo, no existe todavía una definición estable y en la web se encontraron al menos otros dos usos. Uno de ellos es un significado no composicional, específico a una comunidad de personas: madres con conciencia ecológica que prefieren y fomentan el uso de pañales de tela, mientras que el otro se refiere a una persona con un gusto excesivo por coleccionar diferentes tipos de telas. Un análisis morfológico simple revela que las diferencias de significado se deben a las interpretaciones semánticas atribuidas a la primera base léxica de *clothaholic*:

- a) *Clothes* (ropa) + *alcoholic* (se reduce la última parte del primer elemento y la primera parte del segundo). Significado composicional (“*persona que compra ropa de manera excesiva*”).
- b) *Cloth* (tela) + *alcoholic* (el primer elemento permanece intacto, solo se reduce el segundo). Dos últimos significados.

El análisis también demuestra lo difuso de los límites entre los distintos procesos de formación de palabras. Por un lado está el entrecruzamiento (*blending*) como resultado de la reducción de ambas bases léxicas. Por otro están la analogía y la derivación si consideramos que *clothaholic* “copia” formas ya establecidas como *workaholic* o *shopaholic* agregando a la palabra base lo que se percibe como un sufijo. Para Olsen (2014) mientras el significado del lexema original *alcoholic* deba ser recuperado al momento de decodificarlo, la palabra debe ser considerada una forma analógica. Un caso similar al anterior lo constituyen las formas que incorporan *-scape* a partir del entrecruzamiento original *seascape* (*sea* + *landscape* = paisaje marino) tales como *cityscape*, *urbanscape*, *streetscape*, *dreamscape*, *spacescape*, *starscape*, *moonscape*, etc. Bauer (2006) denomina *splinters* (literalmente “astillas”) a las partículas o fragmentos de este tipo cuyo uso frecuente y productividad con el tiempo los lleva a establecerse como sufijos propiamente tales.

En cuanto a creación de compuestos por analogía con otros, hay ejemplos como *house-husband* (“amo de casa”) basado en el compuesto nominal *house-wife*, *First Daughter* (hija del presidente) modelado a partir de *First Lady* (Primera Dama), *whitemail* (persuasión basada en los efectos positivos de una acción) como contraparte de *blackmail* (chantaje, extorsión) o *slow food* como respuesta a la sobreabundancia de *fast food* (comida rápida) en la vida actual. Estos neologismos representan una variación del significado del compuesto original, al cual están necesariamente ligados. Si su interpretación depende de otro elemento entonces carecen de un rasgo esencial de las formas compuestas: la libertad semántica con la que se combinan sus elementos: “The openness of meaning characteristic of the compounding process is lacking” (Olsen, 2014: 46). Por lo tanto deben ser consideradas formas derivadas y no compuestos.

1.6.7. Conversión

La conversión (*conversion*) implica un cambio en la función gramatical del lexema, no así en su forma fonológica. Es un proceso morfológico bastante común y productivo en inglés moderno mediante el cual muchos sustantivos de uso frecuente adquieren estatus verbal: *bottle* → *to bottle* (embotellar), *butter* → *to butter your toast* (poner mantequilla), *chair* → *to chair* (dirigir una

reunión; presidir un comité), *a text* → *to text* (enviar un mensaje de texto). De igual forma, los verbos, especialmente aquellos acompañados de una partícula, pueden “convertirse” en sustantivos: *must* → *it’s a must*, *to guess* → *make a guess*, *to print out* → *a printout* (una copia impresa), *to take over* → *a takeover* (adquisición de un empresa por otra) *to cheat* → *a cheat* (engaño, fraude), *to cook* → *a cook* (un cocinero/a) o adjetivos: *to see through* → *a see-through dress* (un vestido con transparencias). Adjetivos como *crazy*, *empty* y *obscure* pueden transformarse en sustantivos (*a crazy*) o verbos (*to empty* = vaciar, *to obscure*). Muchos otros adjetivos pueden transformarse en sustantivos para referirse a personas como en *an intellectual*, *the subconscious*, *the rich/poor*, *the jobless* (los desempleados) o *my superiors* (Plag, 20011).

Los compuestos nominales constituyen un gran recurso para la creación de verbos compuestos como lo ejemplifican *to blackmail* (chantajear, extorsionar), *to blindfold* (vendar los ojos), *to honeymoon* (ir de luna de miel), *to spotlight* (destacar), etc. Es evidente que el lexema base y la forma derivada por conversión tienen una estrecha relación semántica por lo tanto vuelve a plantearse la necesidad de distinguir entre formas compuestas y derivadas. Olsen (2014) plantea que el verbo *to instant-message* no es un verdadero compuesto sino que un derivado del sustantivo *instant-message*, que sí es producto de la libre combinación de los lexemas *instant* y *message*. Olsen (2014: 44) lo ejemplifica con el compuesto nominal *white-wash* y sus distintas transformaciones:

Tabla 2. Secuencia de procesos morfológicos en un compuesto

| Lexemas base | Proceso morfológico | Significado |
|-----------------------------------|--|-------------------------------|
| 1. <i>to wash</i> → <i>a wash</i> | conversión de verbo en sustantivo | líquido para lavar |
| 2. <i>white</i> + <i>wash</i> | composición (adjetivo + sustantivo) | líquido blanqueador |
| 3. <i>to whitewash</i> | conversión (de sustantivo compuesto a verbo compuesto) | blanquear, encubrir |
| 4. <i>a whitewash</i> | conversión (de verbo compuesto a sustantivo compuesto) | blanqueamiento, encubrimiento |

El proceso de composición solo ocurre en el punto 2 al combinar de manera libre el adjetivo *white* con el sustantivo *wash*. Las subsecuentes conversiones (puntos 3 y 4) derivan una nueva palabra a partir de la totalidad del compuesto base sin combinar nuevos elementos. Esto demuestra que los procesos de composición y derivación pueden interactuar en diferentes niveles o etapas, añadiendo en cada una de ellas un matiz diferente al significado del concepto base.

1.6.8. Acronimia

Los acrónimos (*acronyms*) se forman a partir de las iniciales de un conjunto de vocablos y, a diferencia de las siglas (*initialisms*), se leen secuencialmente, es decir, se pronuncian como una palabra, como es el caso de AIDS /eɪdz/ y su contraparte española SIDA. La tecnología es fuente de muchos de ellos aunque por la rapidez con que avanza en algunas áreas, ciertos términos ya están pasados de moda como es el caso de VCR (*video cassette recorder*) y CD (*compact disk*), los que en estricto rigor son siglas dado que cada letra se pronuncia por separado. Organismos internacionales e instituciones también constituyen un caudal fecundo para este tipo de formaciones: NASA, UNESCO, NATO, TESOL (*Teaching of English to Speakers of Other Languages*), FBI, IRS (*Internal Revenue Service*), etc. No está del todo claro por qué algunas se pronuncian como acrónimos y otras como siglas: MIT (*Massachusetts Institute of Technology*) podría ser pronunciada /'mɪt/ pero se pronuncia como una sigla /,em aɪ 'ti:/. También hay casos en que el acrónimo debe incorporar sonidos vocálicos para facilitar su pronunciación: SCSI (*Small Computer System Interface*) → /'skʌzi/. En la creación de estas formas abreviadas tampoco existe una regla única para determinar cuáles palabras son escogidas y cuáles quedan fuera. Además hay ejemplos en los que se ocupan las dos primeras letras ID (*identity*) y otros en los que solo una parte de la palabra se reduce a la letra inicial como en *e-mail*. Es difícil decidir si el primer caso (ID) constituye un acortamiento (*clipping*) y si el segundo (*e-mail*) es un compuesto recortado o un híbrido de ambos procesos (Bauer, 2006).

Los ejemplos antes mencionados todavía son vistos como abreviaciones por lo tanto aún son escritos con mayúsculas pero existen otros que han pasado a formar parte del vocabulario común por ende su

ortografía ya no es indicativa de su origen como acrónimos: *radar* (*radio detecting and ranging*), *laser* (*light amplification by stimulated emission of radiation*), *scuba* (*self-contained underwater breathing apparatus*), etc.

1.6.9. Reduplicación

La reduplicación (*reduplication* o *echo word*) consiste en la repetición total (*bye-bye, hush-hush, so-so, goody-goody*) o parcial del lexema base. Este último tipo incluye la reduplicación por apofonía o alternancia vocálica (*ablaut*) con ejemplos como *zig-zag, chit-chat, flip-flop, criss-cross, hip-hop*, etc. y la reduplicación con rima: *walkie-talkie, razzle-dazzle, teenie-weenie, super-duper, okey-dokey*, etc. Ejemplos de este tipo en el español de Chile son las expresiones “mucho Lucho” y “ojo piojo”. Cabe mencionar que, en cuanto a origen, muchos casos de reduplicación en inglés se originan en el lenguaje infantil (*night-night, pee-pee/wee-wee, choo-choo*). Aunque estos dos tipos de reduplicación son frecuentemente denominados compuestos, Bauer (2006: 503) enfatiza el hecho de que ni los motivados por rima (*rhyme-motivated compounds*) ni aquellos motivados por apofonía (*ablaut-motivated compounds*), cumplen con el requisito de contener dos lexemas independientes: “*Namby-pamby* and *shilly-shally* are equally not compounds, because they fail to meet the part of the definition which states that a compound must contain bases of two independent lexemes.”

Un tipo bastante productivo de duplicados por rima se originó en el Yiddish usado por la comunidad judía de Nueva York, que luego se traspasó al inglés de la costa este de Estados Unidos y al resto del país, y que consiste en la adición de *sch-/shm-* al inicio de la segunda parte para expresar una actitud despectiva, irónica o burlona hacia lo denotado por el primer elemento. Algunos ejemplos son: *baby-schmaby, fancy-shmancy, moon-schmoon* e incluso *Yiddish, Schmiddish!*

Un caso limítrofe entre reduplicación y composición (Olsen, 2014) lo constituye un tipo de duplicados exactos con función enfatizadora como se ha verificado en varias lenguas y que algunos analistas describen como compuestos idénticos (*identical constituent compounding*). Según Fabb (2001), lo que hace difícil la categorización es el hecho de que ambos constituyentes son ítems

léxicos independientes, por lo que, en estricto rigor, estas formas pueden ser consideradas compuestos. Tanto en inglés como en español estos duplicados enfatizan la referencia prototípica de la palabra escogida. Entonces el término *job-job* especifica que se habla de un trabajo de verdad, con horarios regulares, oficina, etc. así como *milk-milk* aclara que se trata de leche de vaca y no de otros sustitutos como leche de soya (*soy milk*), de almendras (*almond milk*) o de arroz (*rice milk*). Para mayor ilustración, Bauer, Lieber y Plag (2013: 457) ofrecen los siguientes ejemplos:

1. *Are you leaving-leaving?* (¿te vas definitivamente?)

2. A: *Are you nervous?*

B: *Yeah, but, you know, not nervous-nervous.*

Por ejemplo, en Chile ofrecemos “café café” para diferenciarlo del café instantáneo y mucha gente aspira a vivir en una “casa casa” y no en una mediagua. Este recurso también se aplica a adjetivos, por lo que una persona puede ser “mala mala” y algún producto puede ser calificado de “bueno bueno”, como se recalca en una serie de comerciales. La reduplicación como recurso lingüístico satisface una necesidad discursiva puntual y por ende estas formas, ya sean descritas como duplicados o compuestos, son de carácter transitorio, por lo tanto no están sujetas a lexicalización o fijación lingüística.

Como algunos de los ejemplos han puesto de manifiesto, a veces es difícil establecer límites precisos entre los distintos procesos de formación de palabras. Además, cabe destacar que aun cuando estos procesos son bastante frecuentes, entre las diferencias más notorias con la composición, los analistas destacan la productividad y regularidad que caracterizan a esta última en contraste con la “libertad creativa” propia de procesos morfológicos que no están gobernados por reglas (Bauer, 2006).

Aunque la mayoría de los ejemplos incluidos en el listado anterior son el resultado de un solo proceso de creación léxica, Yule (2010: 60) también destaca el hecho de que es posible encontrar casos en los que se combinan dos procesos:

- a. **deli**: *a store where you can buy foods (such as meats, cheese, salads, and sandwiches) that are already cooked or prepared.* (<http://www.merriam-webster.com/>)

Proceso 1: préstamo lingüístico del alemán (*borrowing*) → *delicatessen*

Proceso 2: acortamiento o síntesis (*clipping*) → *deli*

- b. **waspish**: adjetivo derivado de **WASP**: *an American of Northern European and especially British ancestry and of Protestant background; especially : a member of the dominant and the most privileged class of people in the United States.* (<http://www.merriam-webster.com/>)

Proceso 1: creación del acrónimo **WASP** a partir de las iniciales de *White Anglo-Saxon Protestant*.

Proceso 2: derivación → adjetivación mediante la incorporación del sufijo *-ish*. Adicionalmente, al cambiar su estatus morfológico de acrónimo a adjetivo, su ortografía varía de mayúsculas a minúsculas.

- c. **cyber-bullying**: *The use of electronic communication to bully a person, typically by sending messages of an intimidating or threatening nature.*
(<https://en.oxforddictionaries.com/definition/cyberbullying>)

Proceso 1: acortamiento de *cybernetics* → *cyber*

Proceso 2: composición → combinación de *cyber* con el lexema *bullying*.

Sin embargo, si desglosamos ambas partes del compuesto final, nos daremos cuenta de que hay más procesos morfológicos involucrados. Por ejemplo, primero el matemático estadounidense Norbert

Wiener⁴ acuñó el término *cybernetics* en 1948, el que fue posteriormente acortado. Por otra parte, es posible describir *bullying* como una nominalización derivada del verbo *to bully* (acosar, intimidar) mediante el sufijo *-ing*, aunque también ha sido analizado como un ejemplo de conversión (participio presente → sustantivo). Cabe mencionar que *cyber* se ha transformado en un elemento compositivo bastante productivo como lo demuestran *cyberspace*, *cybersex*, *cyberwar*, *cybercrime*, *cyberterrorism*, etc.

1.7. Productividad en procesos de formación de palabras

Los hablantes de una lengua pueden crear nuevas palabras sirviéndose de otras palabras o elementos y para ello utilizan diferentes mecanismos. No solo varía el tipo de reglas, patrones o procesos morfológicos sino que la frecuencia con la que se emplean, lo que hace que algunos de ellos sean considerados productivos mientras que otros son catalogados como menos productivos o derechamente improductivos. La noción de productividad presenta problemas teóricos e interrogantes de envergadura y uno de los principales tiene que ver con la naturaleza misma del fenómeno ¿es cualitativo o cuantitativo? La productividad ha sido definida cualitativamente en los siguientes términos: para Lyons (1977, en Plag, 2006) es uno de los rasgos que definen el lenguaje humano y la define como la habilidad que tienen los hablantes para producir y comprender un número infinito de oraciones que nunca han sido producidas antes. La distingue de la CREATIVIDAD en que esta última es la habilidad para expandir el lenguaje de forma motivada pero impredecible ya que no está gobernada por reglas, lo que impide hacer generalizaciones teóricas sobre ella. Así como los hablantes pueden formar nuevas oraciones, también pueden formar y entender nuevas palabras. Adams (1973) define un patrón productivo como aquel que puede ser usado como modelo para crear nuevos ítems cuando la ocasión lo requiere. Para Bauer (1983) un proceso de formación de palabras es productivo cuando puede ser empleado de manera sincrónica en la producción de nuevas formas. Por su parte, Spencer (1991) considera que una regla es productiva si es usada activa y regularmente

⁴ <https://www.etymonline.com/word/cybernetics>

en la creación de palabras completamente nuevas. Finalmente, Plag (2003) la define como la propiedad que tiene un afijo de ser utilizado para acuñar nuevas palabras complejas. Bauer (2001) además la divide en dos fenómenos de distinta índole: DISPONIBILIDAD, definida como la capacidad para producir formas nuevas. Un proceso morfológico está disponible o no, por lo tanto, es una propiedad cualitativa. La PROFITABILIDAD/RENTABILIDAD (*profitability*) se refiere al grado o medida en que un proceso morfológico puede emplearse en la creación de nuevas palabras por lo tanto es una noción cuantitativa.

1.7.1. Enfoques y mediciones

En términos simples, un proceso morfológico o regla de formación de palabras se entiende como un patrón reconocible, “In general one would say that we can speak of a rule if there is a sufficient number of regular form-meaning correspondences of individual items, i.e. a recognizable pattern.” (Plag, 2006: 539). La definición, sin embargo, no ha estado exenta de problemas y algunos teóricos piensan que las reglas y procesos morfológicos son simplemente mecanismos de analogía mediante los cuales es posible relacionar palabras en grandes grupos. En un enfoque cualitativo un patrón morfológico supuestamente improductivo no puede dar origen a otras formas, sin embargo a veces los hablantes toman formas derivadas ya existentes para crear nuevas palabras por analogía. Si esto ocurre una o dos veces, la regla o patrón morfológico sigue siendo improductivo pero es difícil decidir cuándo deja de serlo y se vuelve productivo. Por este motivo, algunos investigadores optan por agregar un tercer tipo, al que designan como semi-productivo, para dar cuenta de aquellos procesos que desafían una clasificación tan tajante. En esta área gris se encuentran patrones morfológicos “incierto” en cuanto a su disponibilidad.

Los enfoques cuantitativos, por otra parte, no tratan la productividad en términos de “todo o nada” (cualitativamente) sino que la conceptualizan como un continuum en uno de cuyos extremos se encuentran procesos totalmente productivos y en el otro aquellos catalogados como improductivos. La productividad también es vista como un tipo de probabilidad, es decir, como la facilidad estadística con la que un elemento se combina con otros (Plag, 2006). No obstante, existen

dificultades empíricas y teóricas al decidir qué mediciones de productividad deben aplicarse. En el caso de los sufijos, se ha sugerido contar las formas derivadas que aparecen listadas en un diccionario pero eso solo demuestra que existe una cantidad determinada de palabras con ciertos sufijos, pero no comprueba si los hablantes siguen usándolos en formaciones nuevas. Esta es la razón por la que se propone un método alternativo: contar los derivados (neologismos) que se acuñaron en un periodo de tiempo específico. Los diccionarios históricos generalmente contienen este tipo de información pero también es probable que muchas formas hayan quedado fuera. Además, está el hecho de que solamente una pequeña proporción de formas posibles se vuelven formas establecidas (Bauer, 1983). Con respecto a los adjetivos compuestos, algunos de ellos están lo suficientemente establecidos para aparecer como entradas léxicas, pero la mayoría de ellos no figura en los diccionarios, por lo que este tipo de mediciones (*dictionary-based measures*) resultan poco confiables.

Una forma alternativa de medir la productividad involucra calcular la proporción entre palabras “reales” (existentes y ya establecidas) dividiéndolas por aquellas “posibles” (potenciales). Mientras más alto es el índice (*ratio*), más alta es la productividad de una determinada regla. Sin embargo, esta propuesta hace cálculos erróneos cuando trata con procesos altamente productivos o completamente improductivos. Por ejemplo, el número potencial de palabras que pueden incorporar el sufijo *-ness* es, teóricamente, infinito, por lo tanto el cálculo final arrojará un índice de productividad comparativamente bajo. A esto se agrega el problema de contabilizar las palabras posibles: si el número de formaciones potenciales derivadas de la aplicación de una regla morfológica es, en principio, infinito ¿cómo es posible cuantificarlo? De la misma manera, el número de palabras potenciales asociadas a un sufijo nominal improductivo como *-th* será cero, lo que imposibilita realizar cálculo alguno (Plag, 2006). De aplicar esta medida a los adjetivos compuestos, el resultado final implicaría una baja productividad puesto que el número de adjetivos compuestos posibles terminados en los sufijos *-ing* y *-ed* sin duda debe ser altísimo. La discusión es muy relevante al intentar determinar la productividad de este tipo de estructuras compuestas y nos lleva a plantearnos otras preguntas como ¿debe el cálculo incluir el segundo o el primer elemento

del compuesto? ¿Deben contabilizarse tomando en cuenta los ítems léxicos en su interior o el patrón morfológico subyacente?

Finalmente, las mediciones de productividad basadas en grandes corpus electrónicos (*corpus-based measures*) producen listados de palabras que conforman una muestra representativa de lenguaje que puede ser utilizada en diferentes investigaciones. El primer tipo de medición busca determinar la frecuencia de los tipos o casos (*types*), es decir, las diferentes palabras que tienen un sufijo específico. Las otras dos mediciones involucran la noción de HAPAX LEGOMENA o palabras que registran una sola ocurrencia en el corpus. En relación a los procesos productivos, entre los hapaxes es posible encontrar una mayor proporción de neologismos. Puesto que el número de hapaxes de una categoría morfológica debe mostrar una correspondencia con el número de neologismos dentro de esa categoría, los hapaxes son considerados un indicador de productividad (Plag, 2006: 542). Cabe destacar que el tamaño del corpus tiene mucha importancia al momento de determinar la naturaleza de los hapaxes: en un corpus pequeño los hapaxes serán palabras conocidas pero a medida que aumenta el tamaño del corpus, la proporción de neologismos entre los hapaxes también se incrementa.

Además de la anterior, existe una medición de productividad en “sentido estricto” que se obtiene dividiendo el número de hapaxes de una estructura o palabra determinada por el número de ocurrencias (*tokens*) registradas en el corpus. Esta fórmula permite calcular la probabilidad de encontrar una palabra nueva entre todas las palabras de una categoría morfológica específica. Todas estas medidas presentan complicaciones que los investigadores deben resolver al aplicarlas. Entre los problemas metodológicos que deben ser tomados en cuenta al usar este enfoque, está la cuestión de cuán grande debe ser el corpus. Determinar cuáles son las formas léxicas que deben incluirse en el estudio también conlleva difíciles decisiones teóricas y metodológicas ya que muchas veces las categorías no son nítidas y el análisis morfológico escogido puede incidir en la clasificación de las palabras y, por ende, en los resultados.

“In sum, we can say that researchers have a number of different measures at their disposal to assess the productivity of word-formation processes. Each measure highlights different aspects of productivity and brings with it special methodological problems of data sampling and data analysis.” (Plag, 2006: 546)

Todo lo anterior hace que la productividad continúe siendo una de las áreas más debatidas en el estudio de los procesos de formación de palabras aun cuando es innegable que estos procesos son productivos y, por lo tanto, responsables en gran medida del extenso vocabulario que el idioma inglés ha acumulado a través de los siglos.

1.7.2. Restricciones a la productividad

Existen una serie de factores fonológicos, pragmáticos, morfológicos, sintácticos y semánticos que tienen incidencia directa en los procesos de formación de palabras los que pueden operar de manera individual o al unísono. A continuación revisaremos algunos de ellos.

1.7.2.1. Restricciones pragmáticas

Entre las restricciones de naturaleza pragmática está el requisito de existencia que impide la creación de una palabra que denote un ítem, acción o cualidad que no existe. Por supuesto que “existencia” debe ser interpretada de manera flexible para dar cabida a realidades y universos ficticios o mitológicos en los que la aceptabilidad de ciertos términos obedece al contexto en el que se producen. Ese es el caso del compuesto *Adamantium-clawed* utilizado en la descripción de un personaje de comic (*Wolverine*). A veces no es la existencia de un término la que se pone en duda sino que su interpretación. Por ejemplo, el significado de un sustantivo compuesto como *garbage man* (basurero; persona que recoge la basura) no puede ser extendido mediante una analogía con *snowman* (mono de nieve) para denotar “hombre hecho de basura”. Sin embargo, hay que ser muy cuidadosos al decidir si una palabra o un significado diferente existe ya que es posible que simplemente nunca los hayamos visto. Tampoco el hecho de que no aparezcan en diccionarios es

prueba suficiente, dado que solo indica que no son formas establecidas (Bauer, 1983: 85). Teniendo esto en cuenta, entonces es posible imaginar contextos en los que *garbage man* pueda tener una interpretación diferente. Un segundo requisito es el de nombrabilidad (*nameability*): un lexema debe denotar algo que puede ser nombrado. Como las lenguas resaltan distintos aspectos de la realidad este requisito no solo es específico a cada lengua sino que podrá variar dependiendo de la percepción del mundo que determinado grupo cultural posee. Otro factor pragmático guarda relación con modas que hacen que ciertas palabras o elementos morfológicos se vuelvan populares y, por lo tanto, productivos. A raíz de la crisis de Watergate que forzó la renuncia del Presidente de Estados Unidos a mediados de la década de 1970, el sufijo *-gate* ha ganado popularidad en Chile como mecanismo para nombrar escándalos políticos, grandes y pequeños. Recordados son el *Piñeragate* (1992) y el *MOP-Gate* (2003), durante la administración de Ricardo Lagos. No hace mucho se descubrió el *Milico-Gate* y actualmente se investiga el *Pacogate*.

La analogía se distingue de los procesos de formación de palabras en su productividad. Una formación analógica es aquella que se acuña por imitación de un lexema existente pero que no da origen a una serie productiva de otras formas sino que es un fenómeno acotado a unos pocos ejemplos. La analogía, entonces, se relaciona más con la noción de creatividad aunque no es posible excluir completamente la posibilidad de que una forma analógica en algún momento adquiera impulso y comience a generar otros lexemas de manera más productiva (cf. 3.6.6).

Por último, está el fenómeno del BLOQUEO (*blocking*) que describe la no ocurrencia de una forma determinada debido a la preexistencia de otra que ya está aceptada por la comunidad de hablantes. También se la ha denominado “la regla del casillero lleno” (*rule of the filled slot*) y es la responsable de bloquear la acuñación de una palabra como **stealer* puesto que *thief* (ladrón) denota lo mismo y tiene una alta frecuencia por lo que ya está bien establecida. Algo similar ocurriría con un par como **arrivement* y *arrival* (llegada). Algunos teóricos argumentan que lo que se bloquea es la lexicalización o institucionalización de la palabra rival, no la creación de un nuevo término, el que puede existir en contextos restringidos o ser usado de manera idiosincrática. De hecho, existe una relación inversa entre productividad y lexicalización: los patrones más productivos no están

lexicalizados y los que sí lo están son poco productivos. Cognitivamente, el bloqueo de un potencial sinónimo se explica porque la palabra más frecuente está almacenada en la memoria de largo plazo y se activa con más facilidad impidiendo la formación de una palabra rival. Por el contrario, una palabra poco frecuente es de difícil acceso por lo que se registran ocurrencias de pares como *passivate/passivize*. Sin embargo, cuando se bloquea la aplicación de una regla, como sería el caso de sufijos rivales, es necesario tener en cuenta que a veces formas consideradas “anómalas” (*ill-formed*) ocurren varias veces tanto en corpus electrónicos como en Internet lo que hace difícil decidir si deben ser consideradas agramaticales. Ese es el caso del par *decency* y **decentness*: si bien la segunda es una forma que a algunas personas les pueden parecer inusual, registra más de 12.000 ocurrencias en la Web y aparece listada en varios diccionarios. De hecho, existen muchos casos de pares similares en los que ambos elementos tienen el mismo grado de aceptabilidad:

destructiveness / destructivity

discoursiveness / discoursivity

exclusiveness / exclusivity, etc.

Como hemos visto, no es fácil determinar el estatus gramatical de algunas formas, por ende la noción de bloqueo debe ser manejada con prudencia. (Plag, 2006: 552-553)

En resumen, la productividad debe ser vista como un continuum en cuyo extremo menos productivo se ubica la analogía, la que puede generar una sola forma. Al medio debemos ubicar procesos semi-productivos los que pueden verse limitados por una serie de restricciones. Por el contrario, el extremo superior es mucho más complicado puesto que determinar si un proceso de formación de palabras es totalmente productivo es ciertamente difícil y conlleva resolver muchas decisiones metodológicas y teóricas.

1.7.2.2. Factores psicolingüísticos

La mayoría de los modelos de procesamiento morfológico establecen dos maneras en que los hablantes de una lengua pueden acceder y recuperar las palabras complejas almacenadas en el léxico mental: a) por descomposición y b) por palabra completa, como lo ejemplifica Plag (2006: 547) con este esquema:

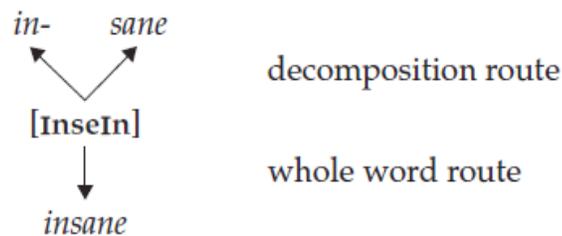


Figura 3. Formas de acceso a palabras complejas en el léxico

Esto significa que cada palabra derivada es procesada simultáneamente siguiendo dos “rutas”: por acceso a los elementos constituyentes para lo que es descompuesta en sus partes (afijos y base), cada una de las cuales es accedida individualmente. Esto implica un alto costo en cuanto al procesamiento puesto que cada morfema debe ser recuperado, calculado su significado para poder, finalmente, calcular el significado total. La otra ruta es más económica de procesar ya que accede directamente a la representación mental de la palabra como un todo.

La frecuencia relativa de la palabra derivada y de su base determina cuál ruta será activada. Este cálculo se realiza al dividir la frecuencia de la derivación por la frecuencia de la base:

$$f_{\text{relative}} = \frac{f_{\text{derivative}}}{f_{\text{base}}}$$

Una palabra derivada como *business*, por ejemplo, está altamente lexicalizada por lo que es fonológica y semánticamente opaca. Además, tiene una frecuencia mayor que su base *busy* lo que, en términos psicolingüísticos, implica que su representación mental es la de una palabra unitaria, por lo tanto, al momento de acceder al lexicón mental la ruta preferida será la de palabra completa. Por el contrario, si la base está más fuertemente representada, como en *cuteness*, se accederá a ella por la vía de la descomposición, puesto que los hablantes perciben mejor las partes que conforman el lexema (base + sufijo). Cabe destacar que en la mayoría de las palabras complejas, la base es más frecuente que su forma derivada (Plag, 2006: 547-548).

Los procesos productivos muestran una preponderancia de formas que son transparentes tanto semántica como fonológicamente. El hecho de que las categorías morfológicas productivas sean semánticamente transparentes es resultado del procesamiento y almacenamiento por descomposición. Si cada uno de los morfemas constituyentes es almacenado individualmente, se refuerza tanto su representación mental como su activación lo que, a su vez, evita la opacidad y la variación semántica.

“Decomposition and individual storage of the constituent morphemes thus leaves little room for semantic drift and opacity, which arise easily under whole word access and storage, where the meanings of the parts are less likely to be activated” (Plag, 2006: 548-549).

Así como se verifica, entonces, una correlación entre productividad y frecuencia, también es posible establecer una relación entre transparencia y productividad. Los especialistas concuerdan en que la economía de almacenamiento se equilibra con la economía de procesamiento lo que justifica la existencia de los dos tipos de almacenamiento (de elementos separados y de palabras completas) ya que ambos son necesarios para lograr máxima efectividad.

Capítulo 2. La composición

2.1. Criterios de identificación

La capacidad para adquirir o dar forma a nuevas palabras es uno de los atributos propios de las diferentes lenguas del mundo, sin importar las características tipológicas específicas de cada una. Algunas de ellas prefieren incorporar nuevos vocablos mediante préstamos, mientras que otras – adoptando el principio de economía de recursos– se inclinan por combinar o recombinar elementos lingüísticos ya existentes en el sistema. A grandes rasgos, la composición es el procedimiento de combinación de dos o más morfemas libres para dar origen a una nueva unidad léxica. Estos morfemas “se funden en una nueva palabra que puede o bien denotar algo completamente nuevo, o bien ser una asociación peculiar de los conceptos ya existentes” (Rodríguez, 1993: 22). Las lenguas se diferencian en cuanto a los tipos de categorías léxicas que pueden formar compuestos. En inglés, existe una gran variedad de compuestos con sustantivos como núcleo aunque también se permite que el núcleo sea un adjetivo y, en menor medida, un verbo.

La gran diversidad de posturas teóricas respecto a lo que constituye un término compuesto hace muy difícil la determinación de un conjunto de criterios generales que ayuden a determinar cuándo estamos ante un compuesto. Aun cuando las estructuras compuestas y los procesos compositivos han sido investigados extensamente en muchas lenguas, prácticamente no existen criterios universalmente aceptados para determinar qué es un compuesto. Donelies (en Lieber y Štekauer, 2009), basado en el análisis de varias familias de lenguas europeas, postula diez criterios:

1. Son estructuras complejas
2. No incluyen afijos usados en la formación de palabras
3. Su ortografía es unitaria
4. Tienen un patrón de acentuación específico
5. Incluyen elementos de unión (*linking*)
6. Su núcleo se ubica a la derecha

7. La flexión se aplica al total del compuesto
8. Son sintácticamente inseparables
9. Son islas sintáctico-semánticas
10. Son unidades conceptuales

Para Lieber y Štekauer (2011), algunos de estos criterios merecen ser revisados cuidadosamente y otros son menos plausibles o, por lo menos, menos generalizables a otras lenguas. En cuanto al primer criterio, es evidente que son construcciones complejas pero esta característica no los distingue necesariamente de las palabras derivadas, las que también son complejas. El segundo criterio parece apuntar a que la composición es diferente a la derivación pero tampoco resulta muy útil puesto que determinar en qué consiste esa diferencia es lo que está en discusión. Los criterios de ortografía (3) y ubicación del núcleo (6) varían de una lengua a otra e incluso dentro de un mismo idioma. Lo mismo ocurre con los elementos de unión (5). En consecuencia, la utilidad de estos criterios puede restringirse a una lengua o grupo de lenguas en particular, pero desde una perspectiva interlingüística no pueden ser considerados criterios definitivos. Esto es especialmente cierto en lo que respecta a la ortografía de los compuestos, como queda claro al comparar el eslovaco y el checo con el inglés. En las dos primeras lenguas todas las estructuras compuestas se escriben como si fueran una sola palabra. En inglés, en cambio, no existe una convención ortográfica aplicable a la totalidad de los compuestos sino que, muy por el contrario, es notoria la inconsistencia a nivel de ortografía. Esto hace posible que algunos compuestos puedan llegar a tener tres variantes escritas como lo ilustra *flowerpot* (macetero), *flower-pot* y *flower pot*. Sin embargo, poner demasiado énfasis en la ortografía y considerarla un criterio válido de composición, en cualquier idioma, parece inadecuado si tenemos en cuenta que la escritura es solo un reflejo secundario de lo que ocurre en el lenguaje oral. Con respecto al último criterio, no es fácil decidir qué constituye una unidad conceptual. Un criterio que no se menciona y que puede ser potencialmente útil es la lexicalización pero también es cierto que mientras más productivo es el proceso de composición en una lengua, es menos probable que los compuestos sean lexicalizados y listados en un diccionario.

Los criterios restantes: la acentuación (4), el comportamiento de las palabras compuestas ante la impenetrabilidad, inseparabilidad e inalterabilidad sintáctica (2, 8, 9) y la flexión (7) son los más importantes. Las lenguas hacen uso de diferentes medios fonológicos para diferenciar compuestos de frases sintácticas, por ejemplo patrones tonales distintivos, armonía vocal, patrones de acentuación, elisión o reducción de vocales, etc. Sin embargo, los estudios de los que se dispone actualmente no permiten establecer la consistencia con que estos criterios ayudan a distinguir la composición como un tipo particular de formación de palabras. Es necesario profundizar en la investigación de estos fenómenos para determinar si en algunas lenguas la composición posee rasgos fonológicos propios.

En inglés, la posición del acento es un criterio de composición relevante aunque ambiguo. La noción más generalizada es que las palabras compuestas se acentúan en el constituyente izquierdo mientras que las frases lo hacen en el elemento de la derecha o bien adoptan un acento doble o simétrico (*level stress*). En realidad, la situación es bastante más compleja y problemática puesto que los patrones de acentuación pueden divergir. Algunas variaciones son simplemente el resultado de diferencias idiosincráticas entre hablantes individuales o entre grupos específicos de hablantes. En el primer caso, el contexto y las condiciones pragmáticas al momento de pronunciar una palabra compuesta pueden influenciar su acentuación. Con respecto a las diferencias entre grupos, algunos especialistas sostienen que en el inglés de Estados Unidos existe una marcada tendencia a acentuar el componente izquierdo de los compuestos. Sin embargo, esto no está del todo resuelto, por lo tanto una afirmación de esta envergadura y alcance necesita ser corroborada por más estudios. Lo que sí está claro es que la acentuación de las unidades compuestas en inglés presenta gran variabilidad y los teóricos se han esforzado por encontrar explicaciones sistemáticas para estas divergencias.

En cuanto a los criterios sintácticos que permiten distinguir compuestos de frases, algunos son específicos a cada lengua. En danés solamente los sustantivos simples pueden tener un artículo definido pospuesto, por lo tanto si una secuencia de dos raíces va seguida de un artículo definido es clara evidencia de que se trata de un término compuesto. En francés, el orden de las palabras es un indicio de composición ya que una secuencia de lexemas que muestre un orden que no puede ser

aplicado a frases sintácticas probablemente sea una construcción compuesta. Entre los criterios de naturaleza sintáctica más usados para reconocer estructuras compuestas en lengua inglesa están la inseparabilidad, la imposibilidad de modificar el primer elemento del compuesto o de reemplazar el segundo sustantivo de un compuesto nominal por una pro-forma como *one*. Este último criterio queda claro en un ejemplo como *black box* (caja negra) que permite responder a la pregunta *What kind of box are you looking for?* (¿Qué tipo de caja buscas?) mediante un pronombre (*a black one*) cuando se trata de una frase nominal pero no cuando se refiere a la caja negra de un avión, término que funciona como un compuesto y que, por lo tanto, no acepta este tipo de modificación.

2.2. La interfaz entre morfología y sintaxis

La composición es el proceso de formación de palabras que muestra más paralelismo con la sintaxis, lo que la ha convertido en un campo de batalla para diferentes enfoques teóricos. Al comparar estructura de frase y estructura de palabra, se observan innegables similitudes entre estos dos niveles de análisis y los paralelos más grandes se encuentran precisamente en las estructuras compuestas. La relación entre formación de palabras y sintaxis es vista de muy diferentes maneras, dependiendo del enfoque que se adopte. Algunos de ellos se atienen a un principio de integridad léxica según el cual las reglas sintácticas no son aplicables a la estructura interna de las palabras, por lo tanto no pueden participar de su creación. En otras palabras, los principios sintácticos no son responsables de la construcción de palabras derivadas o flexionadas. Para los lexicalistas, sintaxis y léxico están claramente separados puesto que el lexicón constituye un módulo pre-sintáctico que posee sus propias reglas para generar y almacenar palabras. En consecuencia, las reglas sintácticas no tienen acceso a la estructura interna de las palabras y no tienen poder para crear nuevos lexemas. Por otra parte, aquellos que provienen de una tradición generativista consideran la estructura de la palabra como si fuera un tipo de estructura de frase o cláusula, por lo tanto la formación de palabras es una forma de sintaxis. Los teóricos que adhieren a un enfoque radicalmente sintáctico no creen en la separación entre morfología y sintaxis puesto que esta última es capaz de derivar estructuras morfológicas mediante transformaciones de las estructuras sintácticas. En las lenguas romances, así como en el holandés, es posible encontrar ejemplos de construcciones sintácticas al interior de

construcciones morfológicas. En las primeras son comunes los compuestos del tipo verbo finito seguido de un sustantivo que actúa como el objeto del verbo (*matamoscas*) y en el segundo se pueden incorporar frases completas a compuestos con núcleo nominal. Esto puede ser indicativo de una tendencia de las palabras compuestas a reflejar la estructura sintáctica de una lengua (Spencer, 2005).

Para Spencer (1991) la composición posee dos tipos de propiedades características. Algunas de ellas tienen similitudes con la formación de palabras y otras con la sintaxis:

- a. La composición se parece a la sintaxis en que es típicamente recursiva.
- b. Al igual que la sintaxis, los elementos al interior de un compuesto están relacionados como los constituyentes de una oración. Las tres relaciones más importantes son núcleo-modificador, predicado-argumento y aposición.

Dentro de los modelos generativistas, la idea de que las palabras y las frases tienen la misma estructura sintáctica ha sido adscrita a modelos denominados *sintaxis de palabra* (*word syntax*). De acuerdo a estos enfoques, las palabras son construcciones endocéntricas dado que poseen un núcleo, al igual que las frases. En el caso de la morfología derivacional, el núcleo se identifica con el afijo que realiza la derivación y que es responsable del cambio de categoría gramatical asociado a este proceso. Aun cuando en el modelo original se hacía la distinción entre representaciones morfológicas y sintácticas y los principios que se aplicaban a ellas, se sostenía que los principios morfológicos y sintácticos podían superponerse e incluso ser compartidos. Los enfoques de naturaleza radicalmente sintáctica parten del supuesto implícito de que la sintaxis es lo único que debe ser investigado en relación a la estructura de la palabra. Sin embargo, la mayoría de los especialistas en morfología piensan que la estructura de la palabra está gobernada por principios que se diferencian en parte de los principios sintácticos. En resumen, el proceso de formación de palabras puede ser abordado desde dos perspectivas. Desde un punto de vista léxico, aun cuando las palabras derivadas se forman en el nivel morfológico pueden heredar ciertas propiedades sintácticas como la estructura argumental. Desde un enfoque sintáctico radical, la formación de palabras tiene

lugar completamente, o casi, en la sintaxis. Aunque pocos teóricos asumen un enfoque totalmente sintáctico, es un hecho innegable que existen aspectos importantes del proceso de formación de palabras que se ven influidos por la sintaxis. Si morfología y sintaxis están gobernadas por principios que convergen en ciertos puntos, entonces parece cobrar mayor validez la adopción de un enfoque teórico “híbrido” que destaque la complementariedad de los diferentes puntos de vista:

“In other words, ‘splitting’ is a perfectly reasonable research strategy for the field as a whole to adopt. [...] suppose there really are differences between morphology and syntax and we adopt one or other version of the monolithic approach to research. If we assume a ‘syntax-all-the-way-down’ approach we will never hypothesize possible independent morphological principles and therefore we will never find them. The monolithic strategy, in contrast to the ‘splitting’ strategy, uniquely brings with it the risk that it will seriously impede progress.” (Spencer, 2005: 94)

En conclusión, la composición representa la interfaz entre morfología y sintaxis aun cuando se considera que la sintaxis es la concatenación de palabras para formar frases, mientras que la composición, en su forma más prototípica, es la concatenación de palabras para formar otras palabras. Sin embargo, muchas veces no existe una manera inequívoca de distinguir entre una palabra compuesta y una frase. Esto quiere decir que cuando la composición es un proceso generativo libre es difícil decidir si estamos tratando con morfología, sintaxis o ambos. O quizás algo diferente. Cuando estamos ante procesos de composición, nos enfrentamos al problema de cómo se define la noción de palabra (Spencer, 1991).

2.3. La composición en lengua inglesa

El inglés es un idioma notoriamente analítico y depende grandemente de la posición que ocupan las palabras dentro de la oración como medio para determinar las relaciones sintácticas establecidas entre ellas. Este orden más estricto favorece una coocurrencia relativamente estable de ciertos lexemas, los que eventualmente pueden llegar a combinarse en un solo elemento. Esta característica

favorece el uso de la composición como mecanismo para generar nuevas unidades léxicas en las que se compacta la información. Olsen (2000: 898) ofrece la siguiente definición:

“In summary, composition has come to be viewed in current linguistic work as the process of concatenating two existing stems from the lexicon of a language to form a new, more complex stem which has the potential to enter the lexicon as a stable morphological unit.”

En comparación con el español y otras lenguas romances, las palabras compuestas han recibido mayor atención en las lenguas germánicas. Este mismo hecho ha traído como consecuencia que en inglés existan muchas y muy variadas clasificaciones de compuestos las que, sin embargo, se ciñen a ciertos criterios comunes. El primero de ellos es la inseparabilidad. Esta característica permite que los compuestos se comporten como una sola unidad y establece que sus constituyentes no pueden ser separados por la inserción de algún otro elemento léxico. Conti (2007) ejemplifica esta característica con el compuesto *bittersweet* (agridulce), el que no puede ser modificado por un adverbio en una de sus partes sino como un todo: *rather bittersweet* pero no **bitter rather sweet*. Los otros dos criterios que se aplican son la movilidad y la estabilidad interna. El criterio de movilidad establece que un compuesto puede moverse dentro de una estructura sintáctica con todos sus constituyentes. El segundo criterio –estabilidad interna– especifica que los elementos que conforman el compuesto lo hacen en un orden establecido que no puede ser alterado, a menos que su significado sea radicalmente modificado como en los casos de *armchair* → **chairarm* o *wallpaper* → *paperwall*. Sin embargo, Bauer (1998) considera que este criterio puede ser flexibilizado en el caso de compuestos nominales del tipo N+N en los que sí es posible que el primer elemento del compuesto sea pre-modificado por un adjetivo como en *instant noodle salad* y *Serious Fraud Office*.

No obstante lo anterior, no resulta fácil encontrar criterios que describan de manera inequívoca las palabras compuestas en inglés. Incluso, algunos especialistas, como Bauer (2005), sostienen que no existen criterios, ni individuales ni colectivos, que sirvan para distinguir claramente los compuestos de las frases. Entre los criterios propuestos para la composición, podemos mencionar la acentuación,

la ortografía, el significado lexicalizado, la inseparabilidad de los dos elementos y la imposibilidad de someter al primer elemento del compuesto a procesos sintácticos como la flexión, la anáfora y la coordinación.

Como hemos mencionado anteriormente, un criterio ortográfico no es confiable dado que la ortografía inglesa es muy poco sistemática en su tratamiento de los términos compuestos. Una misma unidad puede ser escrita como dos palabras separadas; sus componentes pueden tener un guión entre ellos o simplemente ser representada como un solo elemento léxico como en el caso de *flower pot*, *flower-pot* y *flowerpot*. La lexicalización tampoco entrega una base sólida para distinguir compuestos de frases. Si bien es cierto que muchas estructuras compuestas poseen un significado lexicalizado, como *green house* (invernadero), la composición es un proceso tan productivo que existen muchos nuevos compuestos que todavía retienen un significado composicional, sobre todo si se toma en cuenta el contexto como en el caso de *black box*, cuya interpretación variará muchísimo si el tema de conversación tiene que ver con accidentes de aviación.

En inglés, existe también un criterio acentuación para distinguir unidades compuestas de frases. Este hace referencia a un proceso fonológico propio de la composición llamado Regla de Acentuación de Compuestos (*Compound Stress Rule*) según la cual el acento recae en el elemento izquierdo de un compuesto. Esto está en directa oposición con la Regla del Acento Nuclear (*Nuclear Stress Rule*), la que ubica el acento en el constituyente que está más a la derecha en una frase sintáctica (Spencer, 1991). Aun cuando la acentuación con frecuencia ayuda a distinguir estructuras compuestas de frases –de hecho para Bloomfield (1933) solo la acentuación del constituyente izquierdo era el criterio a considerar al momento de determinar un compuesto– ha quedado establecido que existen grandes diferencias en este aspecto y que diversos factores pueden incidir en su variación. Tanto los compuestos primarios como los sintéticos se acentúan generalmente en el elemento de la izquierda (*apple cake*, *truck driver*) pero los compuestos que constan de dos adjetivos a menudo tienen acentuación doble, es decir, se acentúan ambos constituyentes. Incluso hay compuestos del tipo sustantivo-sustantivo que son acentuados en el componente derecho (*apple pie*).

Aun cuando la afirmación de que la flexión solo es aplicable a la totalidad del compuesto excluye la posibilidad de agregar una flexión al primer elemento de una estructura compuesta, existen casos como *children's hour* y *girls' club* en los que el genitivo inglés es perfectamente aceptable. Ocasionalmente, el primer constituyente también puede aparecer en coordinación con otro elemento como en *cat and dog show* y *medical and life insurance*. Con respecto a la anáfora, se han documentado algunos casos como el citado por Bauer (1998) en el que el primer miembro del compuesto actúa como antecedente discursivo para un pronombre elidido: *So I hear you're a real cat-lover. How many (cats) do you have now?*

Para Lieber (2005), la inseparabilidad parece ser la prueba más confiable de que estamos ante una estructura compuesta. Las dos raíces (*stems*) del compuesto son relativamente inmunes a ser separadas por un modificador (**truck fast driver*). En otras palabras, no puede insertarse un modificador de forma arbitraria sin que se pierda la composicionalidad (*compoundhood*). Sin embargo, esto no quiere decir que el segundo constituyente no pueda ser complejo como en *Chinese [jade figure]* pero sí que la introducción de un elemento adicional (*dirty*) como en **Chinese dirty [jade figure]* la vuelva inaceptable. En este punto, la línea divisoria entre compuestos y frases sintácticas en inglés comienza a desdibujarse.

A lo anterior hay que agregar que una de las consecuencias de la falta de criterios claros y definidos para la identificación de los compuestos ha permitido que ítems como *a one-size-fits-all solution* alimenten una controversia en la que algunos teóricos han optado por considerarlos compuestos “frásticos” (*phrasal compounds*) dado que el primer elemento del compuesto es una frase sintáctica. Por otra parte, especialistas como Bauer y Renouf (en Štekauer y Lieber, 2005) optan por describirlo como una frase que actúa como pre-modificador. No existe una respuesta sencilla que ayude a resolver esta controversia. Este tipo de estructuras es muy productivo, sobre todo en inglés oral, aun cuando se observan cada vez con más frecuencia en su forma escrita. No obstante, es difícil decidir cómo pueden ser identificadas positivamente como compuestos. La acentuación no es un criterio válido puesto que en las unidades compuestas de tipo primario el acento en el constituyente izquierdo no es obligatorio. Además, los compuestos frásticos muestran mucha variación respecto a

la acentuación. Al intentar incorporar un adjetivo entre el primer y segundo elemento, queda demostrado que el criterio de inseparabilidad parece ser el más sólido puesto que el ejemplo (b) es bastante más aceptable que (a):

(a) *A *one-size-fits-all standard* solution

(b) A *standard one-size-fits-all* solution

Debemos reconocer, sin embargo, que estos juicios todavía tienen gran parte de intuición por lo que es necesario ahondar en el estudio de estas formas “límitrofes” las que, precisamente por lo controvertidas, tienden a ser dejadas de lado por los investigadores. Paradojalmente, esa misma característica “híbrida” puede ayudar a clarificar la relación entre composición y sintaxis.

Lieber (2005) sostiene que de los procesos de formación de palabras en inglés, la composición es uno de los más productivos e investigados, aunque los manuales de morfología inglesa prestan especial atención a los compuestos nominales por ser los más comunes. Generalmente se distinguen dos categorías de compuestos: los primarios y los sintéticos. A los primeros también se los denomina compuestos primarios (*root compounds*) y son descritos como palabras concatenadas (*file cabinet, sky blue, blackboard*). Los compuestos sintéticos, por otra parte, son aquellos en que el segundo elemento y núcleo es de origen verbal mientras que el elemento de la izquierda cumple la función de argumento interno del verbo del que se deriva el núcleo. Por este motivo, son llamados compuestos verbales, deverbales (*deverbal*) o de nexos verbal (*verbal nexus compounds*). Algunos ejemplos son: *truck driver, waste disposal, gift-giving* y *windblown*. Los compuestos sintéticos son altamente productivos en inglés, al igual que los compuestos de raíz del tipo sustantivo-sustantivo. Las combinaciones sustantivo-adjetivo (*sky blue*), adjetivo-sustantivo (*blackboard*) y adjetivo-adjetivo (*red hot*) son menos productivas. Otros patrones morféimicos más difíciles de encontrar son verbo-verbo (*stir-fry*) o sustantivo-verbo como *babysit*.

Por último, existen tipos de compuestos que son la simple conjunción de dos elementos sin que exista otra dependencia entre ellos. Estos son los llamados compuestos copulativos o coordinativos

y algunos ejemplos son: *mother-child*, *Austria-Hungary*. Los tres tipos de relaciones recién descritas tienden a alinear la composición con la sintaxis. En inglés, aun cuando los ejemplos de frases lexicalizadas como *Jack-in-the-box* (tipo de juguete) y *forget-me-not* (flor) son bastante frecuentes, para muchos analistas no constituyen casos de verdadera composición. Estas expresiones multi-léxicas representan un problema al momento de definir los límites entre sintaxis y composición puesto que se comportan como palabras: son unidades léxicas ininterrumpidas que pueden ser modificadas por un adjetivo, pueden tomar artículo y ser pluralizadas. Sin embargo, se diferencian de los compuestos en rasgos esenciales:

“[...] such multi-word words do not have the usual internal structure of compounds, but have the internal structure of syntactic phrases. Thus, they lack a right-hand head, and they do not consist of two elements that meet the criteria of our above definition of compound. For example, under a compound analysis *jack-in-the-box* is headless, since a *jack-in-the-box* is neither a kind of box, nor a kind of *jack*.” (Plag, 2003: 136)

Por otra parte, los compuestos también poseen rasgos que los asemejan a las palabras. Primero, a menudo están lexicalizados, por lo que su significado, con el tiempo, puede volverse no composicional, es decir, totalmente idiosincrático como *penknife* (cortaplumas). La no referencialidad es otra característica que emparenta a los compuestos con las palabras. Esto quiere decir que los elementos no nucleares en *student film society* no se refieren a ningún estudiante o película específicos y es por ello que pueden ser usados atributivamente. Otra característica de estos constituyentes es que no aceptan flexiones por lo que *pickedpocket* no es posible. Sin embargo, se ha verificado el hecho de que algunos compuestos en inglés aceptan flexión interna irregular en el primer constituyente como es el caso de *teethmarks* (marcas de dientes) mientras que la incorporación de un marcador plural regular (-s) en un sustantivo similar como **nails-marks* (marcas de uñas) es considerada anómala. Aunque menos frecuente, también se ha comprobado la ocurrencia de flexión regular como en el caso de *systems analyst* y *sports car* (Spencer, 1991). La ocurrencia de formas flexivas plurales en los adjetivos compuestos se describe con mayor detalle en la sección 7.1.

2.3.1. ¿Qué es un compuesto en inglés?

Entre los especialistas existen opiniones divergentes respecto a cuáles son los elementos que configuran un compuesto. Conti (2007) prefiere usar el término “material léxico” para dar cabida a todos los puntos de vista respecto a qué es lo combinado en una palabra compuesta. Algunos estudiosos (Adams, 1973 y Fabb, 2001) consideran que solo las formas libres (*free forms*) pueden combinarse para crear una construcción compuesta. Otros autores definen un compuesto como un lexema que contiene al menos dos raíces (Bauer, 1983 y McCarthy, 1991) entendiendo “raíz” como la parte de la palabra que queda una vez que se han quitado todos los afijos. La definición de Bauer es bastante sencilla y clara: “the formation of a new lexeme by adjoining two or more lexemes” (en Lieber y Stekauer, 2009: 4). Esta última definición es la menos restrictiva, pudiendo dar cuenta de una mayor variedad de términos compuestos. En lo que respecta a la lengua inglesa, la definición más abarcadora parece ser la de Plag: “...a compound is a word that consists of two elements, the first of which is either a root, a word, a phrase, the second of which is either a root or a word (2003: 135) quien además incluye otros tipos de compuestos como aquellos que contienen frases (*phrasal compounds*): *do-it-yourself (equipment)*, *good-for-nothing* cuya estructura interna es idéntica a la frase que les da origen (Guevara, 2012: 176). Sin embargo, este enfoque no es compartido por todos los especialistas ya que implica aceptar que reglas sintácticas y morfológicas interactúan de alguna manera. Dressler (2005), por ejemplo, considera que estas formaciones pertenecen a una morfología “extra gramatical” que trata con formas lexicalizadas de construcciones sintácticas, las cuales no están gobernadas por reglas ni son el resultado de procesos morfológicos propiamente tales. Más recientemente, Plag (2011: 3) especifica la definición de términos compuestos de la siguiente manera:

“Compounds can be defined as words that consist of two or more bases. Bases in turn can be defined as elements that can serve as input to further word-formation, i.e. to compounding and suffixation. Bases can be words (as in *breath test*), bound roots (*astro-physics*), or even phrases (*strawberries-in-July talk*).”

Esta definición incluye palabras de origen neoclásico como *astronaut* y *agriculture*, las que están constituidas por un prefijo y un sufijo (griegos y latinos, respectivamente) no por lexemas o formas libres como lo especifican la mayoría de las definiciones. En cuanto a formas compuestas de tipo frástico como *strawberries-in-July talk*, ya hemos visto que representan una categoría por lo menos controvertida (Dressler, 2005, Bauer, 2006). Por su parte, Olsen (2015) acota un poco más la definición al describir la composición como la creación de palabras complejas nuevas (compuestos) mediante la combinación de dos o más lexemas ya sean raíces, bases o palabras independientes. En cuanto a la longitud de los compuestos, en la mayoría de las lenguas generalmente se restringen a dos constituyentes. A este rasgo se le debe agregar el de recursividad, es decir, la posibilidad de reiterar el proceso de composición usando el primer compuesto como *input* para un segundo proceso de composición. Al expandir el compuesto más allá de los dos constituyentes iniciales emerge la noción de binariedad que establece que los compuestos largos deben ser divididos y analizados en estructuras binarias progresivas:

- (a) [N-N] *child care*
- (b) [N-[N-N]] *child care center*
- (c) [N-[N-[N-N]]] *child care center administration*

Si bien la recursividad es potencialmente ilimitada, los compuestos excesivamente largos son difíciles de producir y procesar. Es un rasgo frecuente en alemán y bastante común en inglés y otras lenguas germánicas pero no es una característica esencial del proceso de composición. La recursividad está limitada por una restricción estructural: la posición del núcleo dentro del compuesto. Los idiomas en los que el núcleo aparece al final (a la derecha) como el inglés y el alemán permiten amplia recursividad mientras que los idiomas en los que el núcleo aparece al inicio (español e italiano) tienden a no ser recursivos (Conti, 2007).

En resumen, las características principales de las palabras compuestas regulares en inglés son:

- Son combinaciones de lexemas independientes.
- Los elementos constituyentes pertenecen a clases léxicas abiertas (sustantivos, verbos, adjetivos y adverbios).
- Poseen un núcleo formal ubicado a la derecha (segundo constituyente) que determina la categoría sintáctica del compuesto.
- Este elemento funciona como núcleo semántico respecto del cual el otro constituyente funciona como modificador o complemento.
- El significado global puede derivarse de los significados individuales de cada componente.
- Son estructuras binarias y recursivas (*bone marrow-derived*)
- No incluyen frases, por lo que *blackbird* no puede ser expandido a **black-as-coal bird*. Tampoco aceptan post-modificación como en **slightly-used-car salesman*. (Spencer, 1991: 319)

2.4. Principales clasificaciones de compuestos

En esta sección haremos una revisión de los principales tipos de compuestos comenzando por la clasificación propuesta por Bloomfield en 1933, cuyas categorías siguen vigentes. También incorporamos otras más modernas como la distinción entre compuestos primarios y sintéticos.

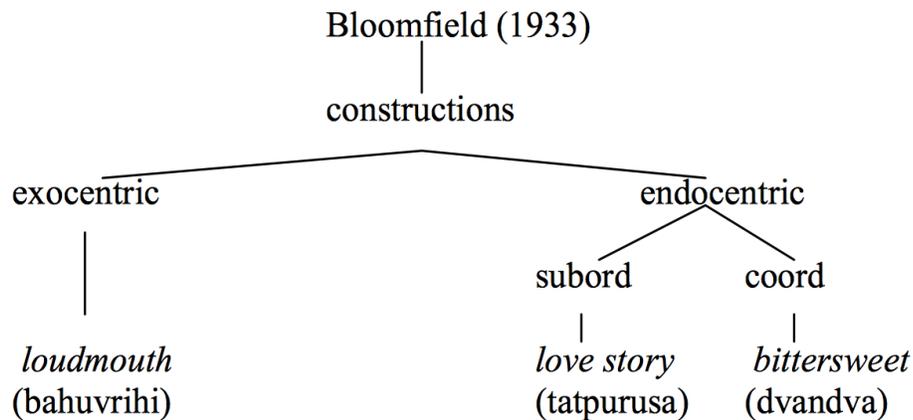


Figura 4. Clasificación de compuestos de Bloomfield (en Scalise y Bisetto, 2005)

2.4.1. Compuestos endocéntricos y exocéntricos

La noción de núcleo pertenece tanto a la sintaxis como a la morfología y es particularmente relevante para la identificación de los diferentes tipos de compuestos y su subsecuente categorización. Este elemento determina las propiedades semánticas y gramaticales (número y género) del compuesto. La propiedad más importante del núcleo es la de asignar categoría gramatical a todo el compuesto: *hummingbird* es un sustantivo puesto que su núcleo (*bird*) lo es. Esta propiedad ha sido llamada “transmisión de rasgos” (*feature percolation*) del núcleo a la totalidad del compuesto (Lieber, 1980, en Conti, 2007).

En los compuestos endocéntricos uno de los elementos funciona como núcleo y el otro actúa como modificador, atribuyéndole una propiedad. Además, el compuesto funciona como un hipónimo del núcleo gramatical (Bauer, 1983). Por ejemplo, *drawbridge* (puente levadizo) representa un tipo particular o subconjunto de puente (*bridge*). Sin embargo, en inglés, las relaciones de tipo predicado-argumento también se observan en los compuestos endocéntricos como *truck driver* en el que *driver* es el núcleo y *truck* actúa como su objeto directo. Este tipo de unidades compuestas conforman el tipo más común de compuestos en lengua inglesa y su núcleo tiende a ubicarse a la derecha, lo cual corresponde a una preferencia universal de acuerdo al enfoque de la Morfología

Natural (Dressler, 2005). Sin embargo, también existen casos de compuestos cuyo núcleo está a la izquierda, como en el caso de las lenguas romances. De hecho, la mayoría de los compuestos sintagmáticos N + N del español presentan el núcleo a la izquierda (*asunto clave, ciudad dormitorio*) mientras que en palabras compuestas como *vasodilatador* o *drogadicción* se ubica a la derecha.

En el caso de la lengua inglesa, se ha identificado una regla llamada *Right-hand Head Rule* (o RHR), propuesta por Williams (1981), la cual sostiene que el núcleo de una palabra morfológicamente compleja lo constituye el elemento ubicado más a la derecha de esa palabra y que es modificado por el otro constituyente. Además, las unidades compuestas heredan todas las propiedades sintácticas y semánticas de sus correspondientes núcleos. Esta regla se aplica tanto a palabras derivadas como a la mayoría de los compuestos. Cabe destacar que la RHR representa una preferencia y no un rasgo universal a todas las lenguas. De hecho, resulta bastante útil en descripciones del inglés y el alemán pero no en las lenguas romances.

En lo que concierne a los compuestos exocéntricos, estos se caracterizan por carecer de un constituyente nuclear (*headless*). Al no poder identificar un núcleo formal, la categoría léxica del compuesto no está determinada por ninguno de sus elementos constituyentes, por lo tanto el núcleo debe ser recuperado o inferido al exterior del compuesto dado que su referente no se corresponde con aquel denotado por el núcleo. Este es el caso de términos compuestos como *pickpocket* (carterista) o *piel roja*, los cuales no describen un tipo de bolsillo (*pocket*) o de piel, sino que hacen referencia a un sustantivo implícito, usualmente una persona. Esta referencia generalmente es hecha de manera peyorativa (*pickpocket, lazybones*). Los compuestos de tipo exocéntrico no son comunes en el idioma inglés puesto que, en general, las lenguas naturales muestran preferencia por los compuestos endocéntricos. Una de las razones es la facilidad para acceder al elemento focal del compuesto (núcleo) el que no requiere ser inferido o recuperado del contexto. De hecho, esto puede no resultar del todo fácil en casos como *buttercup* (tipo de flor) o *spoilsport* (aguafiestas) los que han perdido su significado composicional y han ganado mayor opacidad semántica.

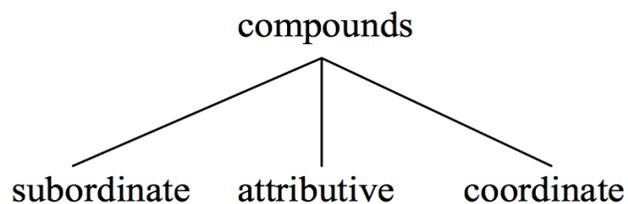


Figura 5. Clasificación de compuestos de Scalise y Bisetto (2005)

2.4.2. Compuestos subordinativos, atributivos y coordinativos

En un compuesto subordinativo los constituyentes están ordenados jerárquicamente de manera tal que un elemento es semántica y estructuralmente dependiente del otro. En *killjoy* (aguafiestas), “*joy*” actúa como complemento directo de “*kill*”, por lo tanto está sintácticamente subordinado al verbo. Los compuestos atributivos, por otra parte, son aquellos en los que el primer elemento actúa como modificador del núcleo (*doghouse*, *red skin*). Ambas clasificaciones se aplican tanto a compuestos endocéntricos como exocéntricos.

También existen compuestos coordinativos o copulativos en cuya interpretación semántica se le asigna igual peso a ambas raíces: *producer-director* designa a una persona que se desempeña tanto como productor como director de una película. Esto quiere decir que al interior de un compuesto también puede existir una relación de igualdad en la que sus constituyentes poseen idéntico estatus jerárquico. Un compuesto endocéntrico como *sofa-bed* se denomina coordinativo puesto que es un hipónimo de ambos elementos, es decir, representa tanto un tipo de sofá como un tipo de cama. Se considera que este tipo de compuestos poseen dos núcleos. Un compuesto exocéntrico como *user-system (interface)* también es coordinativo porque al ser sus constituyentes jerárquicamente iguales, ninguno es el núcleo.

Bauer, Lieber y Plag (2103: 465-466), por otra parte, clasifican los compuestos en argumentales y no-argumentales. Por ejemplo, los compuestos subordinativos son denominados argumentales porque el primero de sus elementos funciona como argumento del otro ya sea como sujeto,

complemento directo o adjunto. Asimismo, los compuestos no-argumentales son equivalentes a los atributivos y coordinativos puesto que la relación que se establece entre los primeros es de modificación mientras que en los segundos ambos elementos tienen igual peso semántico.

2.4.3. Compuestos primarios y sintéticos

Los compuestos primarios (*root compounds*) consisten en combinaciones de las categorías gramaticales abiertas como sustantivos (N), verbos (V) y adjetivos (Adj). Lieber (2005) da algunos ejemplos como *steelmill, towel rack* (N-N), *leaf green, rock hard* (N-Adj), *icy cold, wide awake* (Adj-Adj), *hard hat, bluebird* (Adj-N), *sweet talk, double coat* (Adj-V), *spoonfeed, machine wash* (N-V), *drawbridge, pickpocket* (V-N), *blow-dry, jump start* (V-V). Los primeros cuatro conforman patrones morfémicos muy productivos, especialmente la combinación sustantivo-sustantivo que constituye el tipo de compuestos más común en lengua inglesa. Por el contrario, los patrones que contienen formas verbales son los menos productivos de todos.

En su gran mayoría, el núcleo de los compuestos primarios es el constituyente de la derecha, tanto sintáctica como semánticamente, por lo tanto son de naturaleza endocéntrica lo que determina que, en términos semánticos, generalmente puedan ser interpretados como hipónimos del concepto denotado por el núcleo. Sin embargo, en inglés también existen compuestos primarios de tipo exocéntrico como *red head* el que denota, mediante metonimia, a una persona con cabello rojo y no a un tipo específico de cabeza. Curiosamente, este compuesto puede ser descrito como endocéntrico en cuanto a su sintaxis puesto que la categoría sintáctica está determinada por la segunda raíz, en este caso un sustantivo.

Entre los especialistas existe relativo consenso respecto a la estructura interna de los compuestos primarios: la raíz derecha es el núcleo y, por lo tanto, determina la categoría gramatical y los rasgos morfosintácticos del compuesto. Lieber (2005: 380) lo ejemplifica en los siguientes diagramas:

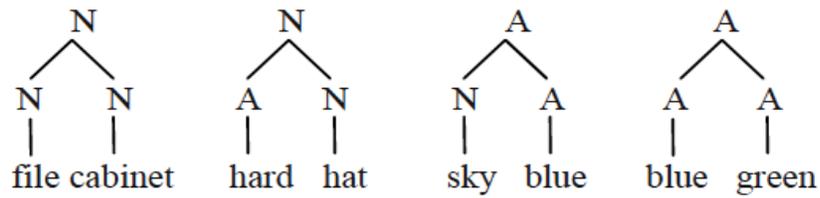


Figura 6. Estructura interna de los compuestos primarios

Con respecto a la interpretación de los compuestos primarios, en el último tiempo se ha llegado a la conclusión de que no existe una relación semántica determinada entre las dos raíces que conforman un compuesto. Sin embargo, hay dos observaciones generales que deben tomarse en cuenta:

- a. A la raíz izquierda se le asigna una interpretación no referencial o genérica. En el caso del compuesto *henhouse* (gallinero), la raíz *hen* no se refiere a ninguna gallina en particular, por lo tanto el compuesto se refiere a cualquier lugar donde se guardan gallinas.
- b. El compuesto denota un subconjunto (hipónimo) de la denotación de la raíz derecha. En el ejemplo anterior, *henhouse* debe ser interpretado como un tipo de “refugio” o “lugar para” gallinas en oposición a otros términos relativos a la tenencia de animales como *doghouse* o *birdhouse*, los que, a su vez, conforman un subconjunto de la denotación más genérica de *house* (casa, vivienda, edificio).

Por su parte, los compuestos sintéticos también tienen el núcleo a la derecha y son altamente productivos en inglés. El segundo elemento de estas estructuras compuestas se deriva de un verbo. Olsen (2000) sostiene que la interpretación semántica de este tipo de compuestos de raíz verbal (*deverbal*) se basa en su estructura argumental puesto que se caracterizan por tener una estructura interna en la que el núcleo hereda, de manera sistemática, los complementos argumentales de su base verbal. Lieber (1992, en Stekauer y Lieber, 2005) los agrupa de la siguiente manera:

- a. *truck driver, meat eater, cat lover, flower grower, gift giver*
- b. *truck driving, meat eating, cat loving, flower growing, gift giving*
- c. *power driven, moth eaten, well loved, home grown, law given*
- d. *tax evasion, revenue enhancement, school closure, waste disposal*

En aquellos compuestos que agregan los sufijos *-er* e *-ing* (a, b) o nominalizadores menos productivos como *-ion*, *-ment*, *-ure* y *-al* (d), el primer elemento es interpretado como argumento interno de la raíz verbal, por lo tanto *flower grower* denota a una persona que cultiva flores. En el caso de estructuras compuestas que carecen de un argumento interno, la primera raíz es interpretada como un adverbial (un adjunto) tal y como lo demuestran los ejemplos agrupados en (c) en los que *moth eaten* hace referencia a algo que ha sido comido por las polillas (agente) y *home grown* especifica el lugar donde algo ha sido cultivado (locativo). Como se puede observar, esto es muy frecuente cuando el sufijo toma la forma pasiva del participio pasado (*-ed / -en*) aunque también ocurre en compuestos terminados en *-ing* como *lake-swimming*. Varios teóricos generativistas han tratado de describir la relación establecida entre los dos constituyentes de un compuesto sintético. Lieber (2005) comenta que entre los modelos propuestos destacan el Principio de Contigüidad (*First Sister Principle*) de Roeper y Siegel (1978), la Hipótesis de Proyección de Primer Orden (*First Order Projection Hypothesis*) de Selkirk (1982) y el Principio de Vinculación Argumental (*Argument Linking Principle*) de Lieber (1983). La diversidad de modelos se debe a que la estructura e interpretación de los compuestos sintéticos ha producido algo de controversia entre los generativistas, quienes, según Lieber (2005: 380) proponen dos estructuras básicas:

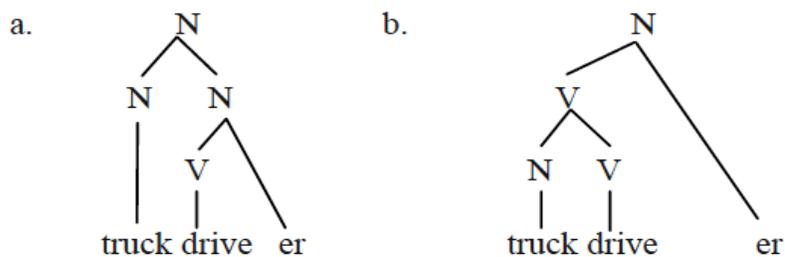


Figura 7. Estructura interna de los compuestos sintéticos

En el caso de (a), el sustantivo *driver*, nominalización del verbo *drive*, es combinado con el sustantivo *truck* (camión). En (b), el compuesto se forma combinando el sustantivo *truck* y el verbo *drive* al que posteriormente se le incorpora el sufijo nominal *-er*. La mayoría de los morfólogos generativistas prefieren la estructura (a) como descripción base de los compuestos sintéticos. Sin embargo, la interpretación de la estructura interna de los compuestos sintéticos dependerá del tipo de relación que el especialista considera que se establece entre los elementos del compuesto. Aquellos que piensan que *truck* actúa como el complemento directo de *drive* preferirán (b).

Capítulo 3. Adjetivos compuestos con participio en inglés

3.1. Características generales

Los adjetivos compuestos en inglés constituyen una de las formas más productivas de creación léxica aun cuando dicha productividad varía considerablemente dependiendo del tipo de adjetivo compuesto. Dentro de esta categoría, los compuestos adjetivos con participio pasado y presente (*participial adjectives compounds*) configuran algunos de los patrones compositivos que mayor rendimiento exhiben.

Existen dos grandes grupos de adjetivos compuestos en inglés cuya morfología se caracteriza por la adición de los sufijos *-ing* o *-ed*, o su variante *-en*, a la palabra base, ya sea ésta un verbo o un sustantivo. En estricto rigor, estos compuestos son el resultado de un proceso híbrido de composición, flexión y derivación, tal como lo indican los afijos contenidos en el segundo componente. En las estructuras compuestas es el segundo elemento o núcleo el que generalmente incorpora un marcador flexivo o derivativo. En el caso de compuestos cuyo núcleo es un participio presente (*-ing*) o pasado (*-ed/-en*) de un verbo, estamos ante una forma flexiva como lo ilustran *fast-flowing*, *city-planned* y *long-forgotten*. Sin embargo, muchos compuestos combinan un sustantivo (*eye*) y un adjetivo (*blue*) para luego agregar un sufijo derivacional al núcleo (*blue-eye* → *blue-eyed*). Esta característica ha causado que su estatus morfológico sea ambiguo por lo que no siempre han sido clasificados como compuestos. Algunos investigadores proponen que esta clase de adjetivos es el resultado de procesos de transformación sintáctica y que la estructura subyacente es una frase nominal con un post-modificador que puede ser una cláusula relativa (Rel Cl) o una frase preposicional (PP). Es por esta razón que los adjetivos compuestos con participio pueden ser considerados formas elípticas u frases condensadas:

(Rel Cl) *a river that flows fast → fast-flowing*

(PP) *a person with blue eyes → blue-eyed*

3.2. Descripciones

Los manuales de morfología inglesa presentan una síntesis de todos los procesos de formación de palabras activos en lengua inglesa prestando especial atención a los compuestos nominales por ser los más comunes. Conti (2007), quien revisa la bibliografía más relevante sobre el tema, menciona algunos autores que ofrecen una breve clasificación de los adjetivos compuestos, entre ellos están: Jespersen (1942), Bauer (1983), Quirk y otros (1985) y Plag (2003). Según esta autora, las únicas descripciones detalladas son las de Marchand (1969) y Adams (1973). Marchand propone una clasificación con nueve tipos de adjetivos compuestos, los cuales corresponden a diferentes patrones de acentuación. Los adjetivos compuestos son tratados como un subgrupo y a su vez subdivididos de acuerdo a la clase léxica a la que pertenece el modificador (sustantivo o adjetivo). Por su parte, Adams los clasifica en diez grupos pero rechaza el criterio de la acentuación para identificar los subgrupos. Su clasificación adopta un criterio mixto que mezcla aspectos morfológicos, funcionales y semánticos, aunque no siempre de manera consistente. Ambos autores destacan la productividad morfológica como un aspecto relevante al tratar este tipo de formaciones. Cabe hacer notar, además, que una gran parte de la investigación sobre estos adjetivos ha sido realizada en estudios contrastivos inglés-alemán, dadas las similitudes de los procesos compositivos en ambas lenguas.

Conti (2007) estima que existen muy pocas monografías sobre adjetivos compuestos en inglés y que la mayoría tiene ya bastantes años. Entre los trabajos que la autora menciona, destaca a Meys (1975), quien llevó a cabo una extensa investigación, la única dirigida a crear una clasificación global de estas formaciones en inglés, en la que analizó un corpus de cerca de 2.000 compuestos sacados de periódicos y revistas. Su enfoque es de tipo transformacional y su clasificación se basa en la sintaxis, por lo tanto considera que los adjetivos compuestos son formaciones elípticas u oraciones condensadas que han pasado por sucesivas transformaciones. “Compound adjectives can, as a rule, be derived in a fairly straightforward way from underlying structures; they are in the true sense only condensed sentences” (en Conti, 2007: 120). Meys (1975) es el único especialista que basó sus hallazgos en un corpus léxico de prueba compilado por él mismo. Coincidentemente, parte de mi investigación consistirá en analizar los patrones morfológicos de compuestos participios que he recogido en un corpus bastante más extenso que el de este autor.

Otro trabajo dedicado a la investigación de adjetivos participios compuestos en inglés es el de Hellinger (1969), escrito en alemán. El interés del autor por el tema se debe a dos factores a) la alta frecuencia y productividad de estas construcciones y b) la falta de una descripción completa tanto en la gramática tradicional como la transformacional. A partir de un corpus propio recogido de *Time Magazine*, Hellinger llega a la conclusión de que los compuestos de tipo participio tienen un bajo grado de lexicalización como lo demuestra el limitado número de formaciones establecidas registradas en los diccionarios. Al igual que Meys (1975) considera que estas formaciones están más relacionadas con la sintaxis que con el lexicón y que el modelo generativo-transformacional es el único capaz de dar cuenta de ellas. Su clasificación descansa en un criterio semántico-sintáctico en el que determina los distintos subgrupos de acuerdo al rol que tiene el componente izquierdo en relación al participio (a la derecha) y al contexto sintáctico en general. Hellinger identifica tres tendencias en el comportamiento de estos compuestos:

- Salvo contadas excepciones, no pueden ser graduados (*non-gradable*).
- Generalmente ocurren en posición atributiva (antes del sustantivo).
- Difícilmente pueden ser modificados por un intensificador.

Otra monografía en línea con los trabajos anteriores es la de Todenhagen (1974), también escrita en alemán. Es un estudio contrastivo entre los participios usados de manera atributiva delante de sustantivos tanto en inglés como en alemán y señala las diferencias de las formaciones complejas en ambas lenguas. Cabe destacar lo perceptivo de su análisis al enfrentarse al problema de la ambigüedad que surge al intentar analizar un compuesto como *a silver-mounted (cartridge belt)*, dado que no es posible determinar si el compuesto deriva de un verbo o de un sustantivo:

a silver-mounted (cartridge belt) → *a cartridge belt which has been mounted with silver*
→ *a cartridge belt which has a silver mounting*

El autor señala que este tipo de ambigüedad es más común en inglés que en alemán y que no puede ser resuelta fácilmente ya que no existe una regla para distinguir entre los dos posibles orígenes.

Otro estudio contrastivo de los adjetivos compuestos en inglés y alemán es el realizado por König (1971) quien hace notar que en alemán la premodificación tiene menos restricciones que en inglés, lengua en que los adjetivos compuestos generalmente aparecen en posición atributiva, a menos que se trate de formaciones bien establecidas, como en *Mary is good-looking* (entrada de diccionario), las cuales pueden aparecer en posición predicativa. En alemán, la mayoría de los compuestos pueden ocupar libremente ambas posiciones. Otra investigación contrastiva inglés-alemán es la de Welte (1982), quien ofrece una visión panorámica de los adjetivos compuestos en ambos idiomas e indaga en casos problemáticos que hacen que la clasificación se vuelva compleja.

Otro autor que enfocó sus investigaciones en los adjetivos compuestos con participio es Gram-Andersen. Sus contribuciones son bastante más recientes y se dividen en dos etapas. Primeramente, en 1992, describió los que él define como *N-ed compounds*, es decir, combinaciones del tipo Adj-N y N-N seguidas del sufijo *-ed*, las que generalmente expresan una idea de posesión como en *cucumber-nosed* (nariz de pepino). Coincidentemente, esta es una de las grandes categorías presentes en el corpus de prueba, incluso con el mismo nombre (*N-ed compounds*). Además, presenta un gran corpus de datos sacados de periódicos y obras literarias, la mayoría del inglés moderno aunque algunos ejemplos pertenecen a otras épocas. Ofrece una detallada categorización de un fenómeno muy vasto, cuya productividad es evidente en varios registros lingüísticos, concentrándose en las relaciones semánticas establecidas entre los compuestos y el sustantivo al que se refieren, sobre todo en sus usos figurativos y connotativos.

En 1995, Gram-Andersen publicó un segundo estudio sobre adjetivos compuestos en el cual aborda las formaciones que contienen un participio presente como segundo constituyente (e.g. *long-suffering, nail-biting*). Al igual que en su anterior trabajo, el autor presenta una gran cantidad de datos, los que abarcan desde el inglés antiguo hasta el moderno, los que incluyen tanto lenguaje cotidiano (revistas y diarios) como lenguaje poético (contemporáneo y antiguo). Lleva a cabo un análisis exhaustivo en el que clasifica este tipo de construcciones de acuerdo al papel que desempeña el primer elemento constituyente en relación al segundo (participio presente). Adopta un criterio dual (semántico y sintáctico) mediante el cual identifica nueve patrones básicos que tienen

sustantivos, pronombres y adverbios como primer constituyente. También investiga la naturaleza del núcleo al que modifica el adjetivo compuesto y distingue ocho tipos de acuerdo a sus rasgos semánticos (+/- animado, +/- humano, etc.).

3.3. Adjetivos compuestos terminados en *-ed*

La mayoría de las unidades compuestas analizadas en esta investigación pertenecen al grupo de compuestos sintéticos puesto que su núcleo es verbal y, como ya hemos visto, en su formación intervienen tanto la composición (combinación de dos lexemas) como la flexión (los participios incorporan sufijos flexivos *-ed* / *-ing*). Sin embargo, dentro de la categoría terminada en *-ed*, también hay un grupo numeroso de compuestos que se obtienen mediante la derivación de una base nominal. Comenzaremos describiendo este último tipo de adjetivos compuestos.

3.3.1. Adjetivos compuestos del tipo N-*ed*

Estos adjetivos compuestos se distinguen por tener como segundo elemento no el participio de un verbo sino que un sustantivo al que se le añade el afijo *-ed*, el que en este caso es un marcador adjetival. Según Crocco-Galeas (2003) es necesario distinguir entre el sufijo *-ed* característico de la flexión verbal de los verbos regulares del inglés en pasado simple y participio pasado (*end* → *ended*) y su contraparte derivacional. Esta última opera en los procesos de formación de adjetivos de origen nominal (*denominal adjectives*) como *lipped* y *walled* creados a partir de los sustantivos *lip* (labio) y *wall* (pared; muro), respectivamente. Sin embargo, los adjetivos derivados de sustantivos son poco productivos de manera independiente puesto que representan una forma literaria usada para atribuir una característica específica: *lipped bowl* (un recipiente con “boca”), *walled garden* (un jardín “cerrado”). También Ljung (1976) estima que la creación de adjetivos denominales se ve limitada por consideraciones estilísticas: son formas típicas del lenguaje técnico y científico, los que tienen necesidades y estándares descriptivos propios. Por ese motivo, adjetivos que son perfectamente legítimos en discursos especializados pueden sonar pretenciosos, e incluso ridículos, en el lenguaje cotidiano. Esto hace que los hablantes prefieran construcciones alternativas (frases

preposicionales o cláusulas relativas) lo que incide directamente en su renovación. Por el contrario, su productividad es muy alta en la formación de compuestos, como lo demuestran los ejemplos registrados en el corpus de muestra:

| | | | |
|-----------------|------------------------|-----------------|-----------------------|
| <i>lipped</i> → | <i>black-lipped</i> | <i>walled</i> → | <i>crystal-walled</i> |
| | <i>dry-lipped</i> | | <i>double-walled</i> |
| | <i>full-lipped</i> | | <i>hard-walled</i> |
| | <i>hard-lipped</i> | | <i>high-walled</i> |
| | <i>long-lipped</i> | | <i>metal-walled</i> |
| | <i>loose-lipped</i> | | <i>mud-walled</i> |
| | <i>luscious-lipped</i> | | <i>plank-walled</i> |
| | <i>pillow-lipped</i> | | <i>stone-walled</i> |
| | <i>platter-lipped</i> | | <i>thin-walled</i> |
| | <i>scarlet-lipped</i> | | <i>white-walled</i> |
| | <i>thin lipped</i> | | |
| | <i>tight-lipped</i> | | |

Los patrones morfémicos más comunes de estos adjetivos compuestos son N + N como en *pillow-lipped* (con labios “mullidos como almohadas”) o *stone-walled* (con muros de piedra) y, sobre todo, Adj + N como en *scarlet-lipped* (de labios rojos) y *thin-walled* (de paredes delgadas). Cabe destacar que, en muchos casos, el segundo constituyente de estas estructuras compuestas no es aceptable como palabra independiente puesto que su significado denota una parte integral de algunos sustantivos por lo que su uso como adjetivo resultaría pleonástico: **faced girl* o **roofed house*. Todos los adjetivos denominales terminados en *-ed* transmiten la idea de **posesión no enajenable** (*inalienable possession*) del concepto denotado por el sustantivo base lo que vuelve inaceptables ejemplos como **carred man* o **wifed man* ya que ni *car* ni *wife* cumplen con esta condición (Hudson, 1975). En consecuencia, la restricción principal al momento de formular estos adjetivos es de naturaleza pragmática: no enuncie aquello que su interlocutor ya sabe (Ljung, 1976). Por esta

Las dos primeras son denominadas *de primer grado* e incluyen partes del cuerpo o capacidades físicas –en el caso de las personas y animales– y partes como ruedas, respaldo y lados en el caso de vehículos, sillas y formas geométricas, respectivamente. No obstante lo anterior, la posesión no enajenable *absoluta* es rara ya que a veces se registran excepciones a la regla:

“A much more common situation is that we know there to be exceptions to the rule, but we choose to ignore them. Thus we all know that not all men have legs, hands, eyes, etc., but we claim that men normally do, that there is something odd about men who do not. The same kind of relation is found in innumerable other cases, for example between wings and planes, wheels and cars, cruelty and tyrants. In these cases it is no longer true that the relevant part-of noun is DEFINITIONALLY part of all and every possessor noun”. (Ljung, 1976: 162)

Al estimarse que el sustantivo ya no es, por definición, parte de todos y cada uno de los posibles poseedores, se establece una distinción entre posesión no enajenable *absoluta* y otra que es *relativa*. Esta última se aplica a las partes y propiedades inherentes que “se espera” que estén presentes en otros sustantivos pero que no siempre se constatan. La posesión no enajenable de segundo grado se establece cuando solo un subconjunto de miembros de una clase léxica comparte una propiedad o característica inherente. Este es el caso de *bearded* o *blue-eyed* y *men* puesto que no todos los hombres tienen barba o los ojos azules. Finalmente, la posesión enajenable ocurre cuando no existe una relación parte-todo: un auto no es parte de un hombre y tampoco su esposa, lo que excluye **carred man* o **wifed man*, respectivamente. Ljung (1976) aclara que dentro de los principios que rigen los procesos de adjetivación aparecen excepciones como *moneyed* (con dinero), *landed* (que posee tierras) y *kingdomed* (que posee un reino), adjetivos cuyas bases son sustantivos que no cumplen con el requisito de ser parte inherente de los referentes que los poseen.

En resumen, los tres tipos de posesión no enajenable permiten la creación de adjetivos compuestos con pre-modificación (*blue-eyed*) mientras que los adjetivos denominales simples (*bearded*) no ocurren en casos de posesión absoluta pero sí cuando es relativa o de segundo grado. Respecto a este último tipo de posesión discrepamos con Ljung (1976) quien incluye sustantivos como *hat*, *scarf* y

sweater, como lo demostrarían *yellow-hatted girls* (muchachas de sombrero amarillo), *red-scarfed girl* (chica de bufanda roja) o *blue-sweatered students* (estudiantes con chalecos azules). Nos parece evidente que las prendas de vestir y los accesorios no son “parte inherente de” o “inherentemente poseídas” por el subconjunto de personas que los usan por lo que deberían ser incluidos en la última categoría. Consideramos que ejemplos como estos comprueban fehacientemente que sustantivos a los que se atribuye *posesión enajenable* sí participan en la formación de compuestos adjetivos aunque de manera limitada. En el corpus de prueba encontramos construcciones similares como *black-bereted* (de boina negra), *red-coated* (de abrigo rojo) y *flower-crowned* (con una corona de flores). Un caso especial de posesión no enajenable de segundo grado es *black-spectacled* ya que los lentes se relacionan estrechamente con una discapacidad física, la que sí es parte inherente de algunas personas (cf. 8.1.3.).

La discusión anterior es reflejo de las dificultades que se ven inevitablemente asociadas a un concepto fluido y variable como la posesión no enajenable, la que está determinada por las muy diversas capacidades de conceptualización de los seres humanos en las que inciden factores cognitivos y culturales como el conocimiento de mundo y la percepción de la realidad: “It is clear, however, that serious difficulties are created for any attempt to specify 'what can inalienably possess what' by the freedom speakers have to create new domains of inalienability”. (1976: 167) A esto debe agregarse que la formación de adjetivos compuestos también responde a usos literarios (*robin-breasted sailing boats*), metafóricos (*milk-blooded coward*), idiomáticos/lexicalizados (*backhanded / harebrained*), e incluso idiosincráticos (*strange-eyed*) y/o con referencias culturales (*Cookie-monster-haired*).

3.3.1.1. Estructura interna

Como ya hemos visto, el sufijo *-ed* puede participar en la creación de adjetivos denominales y adjetivos compuestos de base nominal, es decir, puede ser agregado tanto a un sustantivo como a una forma compuesta. En el segundo caso, el sufijo se agrega a todo el compuesto, no solo al segundo elemento (Crocco-Galeas, 2003: 35), por lo tanto se obtiene una estructura morfológica del tipo:

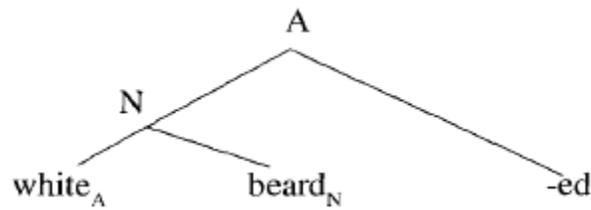


Figura 8: Estructura interna de compuestos de base nominal 1

El diagrama representa la estructura $[[N/Adj + N]_N + ed]_{ADJ}$ la que corresponde a un compuesto que está sufijado (*suffixed compound*). Esto quiere decir que la terminación *-ed* puede combinarse con diferentes bases: verbos (flexión), sustantivos y compuestos (derivación). Estos últimos son llamados adjetivos compuestos derivacionales y se construyen de una frase nominal, como en el ejemplo *white beard* (barba blanca), a la que se le agrega el sufijo *-ed* para crear el adjetivo *white-bearded* (de/con barba blanca). En consecuencia, la frase primero pasa por una regla de composición y luego una regla derivacional. En este análisis el sufijo es considerado el núcleo del compuesto – determina la categoría léxica– el que, a su vez, posee una estructura interna que combina un núcleo nominal con un pre-modificador (N/Adj). Un argumento a favor de esta descripción es que, en inglés, los compuestos con estructuras morfológicas del tipo N + N y Adj + N son más productivos y naturales que los adjetivos denominales (N-*ed*) al estar sujetos a pocas restricciones semánticas y de selección que limiten su co-ocurrencia con ciertos sustantivos. Para Hudson (1975) las restricciones de selección se basan en la noción de posesión no enajenable la que determina la “normalidad” (*normalcy*) de las combinaciones que se obtienen. En consecuencia, el criterio de naturalidad o normalidad prevalece sobre el de diferenciación (*markedness*), por lo que los compuestos constituyen candidatos viables para servir como bases morfológicas para posteriores procesos de formación de palabras.

Por otra parte, la segunda representación posible de la estructura de los adjetivos compuestos en inglés (Crocco-Galeas, 2003: 35) asigna el rol de núcleo al segundo componente, el que es una palabra derivada (*bearded*):

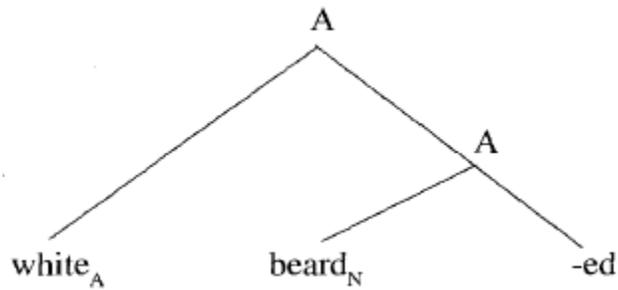


Figura 9: Estructura interna de compuestos de base nominal 2

Desde una perspectiva morfológica ambas representaciones son aceptables pero tienen diferentes grados de plausibilidad. Como ya hemos visto, los adjetivos denominales que ocurren como segundo elemento no siempre corresponden a una palabra real. Hudson (1975) discrepa con la noción de que la estructura pre-modificada (*blue-eyed*) presupone la existencia, en una etapa anterior en el proceso de derivación, de la forma semi-productiva terminada en *-ed* (*eyed*) puesto que la gramaticalidad de esta última es cuestionable. Crocco-Galeas (2003: 36) concuerda con esta apreciación: “The suffix *-ed* is preferable as head of compound whereas the complex word in *-ed* does not always represent an actual word”. Por ejemplo, en el corpus electrónico las formas “*bellied*” y “*brained*” no registran ocurrencias como adjetivos denominales independientes pero sí aparecen al interior de compuestos como *big-bellied* (panzón), *white-bellied* (de vientre blanco) o *hare-brained* (descabellado), *pea-brained* (tonto) y otros. En atención a razones semánticas, entonces, debiera preferirse un análisis estructural que incorpora formas léxicas verificables y no aquellas que todavía están en un estado intermedio de aceptabilidad y/o gramaticalidad. Teniendo en cuenta los argumentos presentados, Lieber (2009) concede validez a ambos análisis en la descripción general que hace de los compuestos en lengua inglesa.

3.3.2. Patrones morfológicos

Los adjetivos compuestos terminados en *-ed* configuran un grupo numerosísimo puesto que, además de poderse combinar con elementos de diversas clases gramaticales, las bases pueden ser de

naturaleza verbal o nominal. Estas unidades complejas presentan un rango de variación que va desde patrones morfémicos estándares a formas menos estandarizadas:

Tabla 3. Patrones morfológicos compuestos terminados en *-ed*

| Patrón combinatorio | Formas estándar |
|----------------------------|--|
| Adj + N- <i>ed</i> | <i>long-sleeved, silver-haired, big-boned, long-legged</i> |
| N + N- <i>ed</i> | <i>angel-faced, lion-hearted, bell-voiced</i> |
| N + V- <i>ed</i> | <i>mineral-enriched, rain-swept, sun-burned</i> |
| Adv + N- <i>ed</i> | <i>commercially-minded, nobly-profiled (man), tightly-bodied (dress)</i> |
| Adv + V- <i>ed</i> | <i>freshly-picked, finely-developed, newly-wedded</i> |
| Patrón combinatorio | Formas semi-estándar |
| Núm + N- <i>ed</i> | <i>one-sided, two-edged, three-legged, eight-eyed</i> |
| Adj + V- <i>ed</i> | <i>British-born, soft-boiled, bridal-inspired (dress)</i> |
| Cuant + N- <i>ed</i> | <i>many-faceted, few-paged, half-hearted</i> |
| Adv sup + V- <i>ed</i> | <i>best-dressed, most-imitated, worst reviewed</i> |
| Patrón combinatorio | Formas emergentes |
| Adj sup + N- <i>ed</i> | <i>widest-winged, worst-tempered, keenest-eyed</i> |
| Adj comp + N- <i>ed</i> | <i>better-humored, lighter-boned, lighter-skinned</i> |
| Adv comp + V- <i>ed</i> | <i>lesser-known (sister), better-dressed, lower-born</i> |
| Núm + V- <i>ed</i> | <i>1-ranked</i> |

Con respecto a la acentuación de los compuestos de base nominal (N-*ed*), Lorenz (2012) señala que en posición predicativa como en “He’s quite **absent-minded**”, “that’s quite **short-sighted**” “the girl is **blue-eyed**” ambos componentes son acentuados (acentuación simétrica), pero que al adoptar una posición atributiva, como pre-modificadores del sustantivo, la acentuación es inicial, trasladándose al primer componente: “an **absent-minded** profesor”, “a **short-sighted** person” y “such a **blue-eyed** boy”.

3.4. Adjetivos compuestos con participio presente

Los adjetivos compuestos con participio presente (-*ing*) también constituyen un grupo productivo, aunque en menor medida que el anterior. Aunque los patrones morfológicos muestran menos variedad que los de su contraparte, igualmente se ha constatado la ocurrencia de formas menos estandarizadas y emergentes.

3.4.1. Patrones morfológicos

Los tres patrones más numerosos son los que incorporan un sustantivo, un adjetivo o un adverbio como constituyente inicial. En el caso del patrón Adj + V-*ing*, aunque productivo, el núcleo está acotado a los verbos de percepción *look*, *smell*, *sound* y *taste*.

Tabla 4. Patrones morfológicos compuestos terminados en -*ing*

| Patrón combinatorio | Formas estándar |
|----------------------------|---|
| Adj + V- <i>ing</i> | <i>bitter-looking, fragrant-smelling, sweet-tasting</i> |
| N + V- <i>ing</i> | <i>man-eating, power-producing, woman-hating</i> |
| Adv + V- <i>ing</i> | <i>long-suffering, well-meaning, ever-changing, low-lying</i> |
| Patrón combinatorio | Formas semi-estándar |
| Adv comp + V- <i>ing</i> | <i>better-performing, higher-paying, harder-hitting</i> |
| Adv sup + V- <i>ing</i> | <i>best-selling, fastest-talking, highest-grossing</i> |
| Patrón combinatorio | Formas emergentes |
| Adj comp + V- <i>ing</i> | <i>lighter-looking, better-sounding, younger-looking</i> |
| Adj sup + V- <i>ing</i> | <i>roughest-looking, best-tasting, sternest-seeming</i> |
| Núm + V- <i>ing</i> | <i>third-ranking, eleventh-leading, second-leading</i> |

3.5. La teoría de Roeper y Siegel (1978)

Las estructuras compuestas fueron abordadas a principios de los años 60 por teorías desarrolladas dentro de la Gramática Generativa, la que distinguía dos tipos de compuestos: verbales y primarios. En el caso de los primeros, el segundo constituyente (núcleo) es un elemento de naturaleza verbal (*bus-driver, steel-production, meat-eating*) por lo tanto el constituyente de la izquierda es interpretado como un argumento del núcleo. Los compuestos primarios, por otra parte, están constituidos por sustantivos (*car radio, honey-bee, suspense film*) y la relación semántica que se establece entre ellos es bastante heterogénea y difícil de determinar desde una perspectiva gramatical. Por ejemplo, el compuesto *suspense film* puede ser interpretado como una “película que contiene suspenso” o “que produce suspenso” (Graña, 1994).

Una de las primeras descripciones generativistas de compuestos sintéticos es la de Lees (1961) quien propuso una derivación transformacional en la que los compuestos emergían de la estructura profunda de las oraciones mediante una serie de transformaciones que permitían derivar construcciones como *apple eater* y *apple eating* a partir de *the man eats apples*. Esta propuesta recibió duras críticas en su momento debido a la arbitrariedad de las estructuras subyacentes y de las transformaciones propuestas. Marchand (1969) realizó un análisis basado en la semántica el que, sin embargo, tenía similitudes con el de Lees (1961) puesto que la clasificación que ofrece de los compuestos sintéticos está basada en relaciones gramaticales que son un reflejo de aquellas que ocurren a nivel oracional (sujeto, objeto, verbo). Según Lieber (2005), dentro de la teoría generativista, el trabajo de Lees (1961) continuó siendo reformulado, como se puede apreciar en los análisis propuestos por autores como Levi y Bauer (ambos en 1978). Las subsecuentes reformulaciones teóricas del modelo generativista y la aparición del enfoque lexicalista a mediados de la década de 1970 contribuyeron todavía más al refinamiento de los análisis. La hipótesis lexicalista postulaba la existencia de un componente generativo independiente de formación de palabras, por lo tanto los compuestos debían ser derivados mediante reglas léxicas y no transformacionales. Es así como, también en 1978, Roeper y Siegel proponen un modelo de transformación léxica para los compuestos sintéticos que combina postulados léxicos (*compound*

rule, affix rule) con un enfoque sintáctico basado en la estructura argumental y adverbial proyectada por el núcleo verbal del compuesto.

En inglés, los adjetivos compuestos con participio presente y pasado representan un caso complejo para las diferentes corrientes teóricas puesto que no existe consenso al momento de analizarlos y clasificarlos. Poseen una naturaleza dual que los hace ser en parte sintácticos y en parte léxicos. Roeper y Siegel (1978), al describir este tipo de compuestos, proponen que el paralelismo existente entre el compuesto y su correspondiente frase verbal debe ser representado incorporando aspectos de la estructura sintáctica a la representación léxica del compuesto. Para ello proponen una serie de reglas transformacionales que operan sobre las entradas léxicas y que:

- a. Transforman un verbo en un adjetivo o sustantivo mediante la generación de una posición (*slot*) delante del verbo donde se insertará el elemento no nuclear. Esto constituye un marco de subcategorización.
- b. Insertan el sustantivo no nuclear en la posición del complemento sintáctico del verbo original.
- c. Mueven el sustantivo no nuclear a la posición inicial.

La Regla de Composición que refleja el Principio de Contigüidad se expresa de la siguiente manera:

$$\begin{array}{ccccccccc}
 [[\text{empty}] + \text{verb} + \text{affix}] & [X+N + \text{word}] & W & \rightarrow & [[+ \text{word}] + \text{verb} + \text{affix}] & & W \\
 1 & 2 & 3 & 4 & 5 & 4 & 2 & 3 & \emptyset & 5
 \end{array}$$

Ejemplo: $[[\text{empty}] + \text{make} + \text{er}] [N \text{ coffee}] W \rightarrow [[\text{coffee}] + \text{make} + \text{er}] W$ (*coffee-maker*)

Sustantivos, adverbios, adjetivos y, quizás partículas (preposiciones), parecen ser los únicos complementos que pueden ser incorporados a un compuesto ocupando el marco “vacío” [empty] ubicado a la izquierda del verbo base. El objetivo de su propuesta es demostrar que las transformaciones léxicas dan cuenta de forma económica de los compuestos verbales (aquellos que

toman los afijos: *-er*, *-ing* y *-ed*). De acuerdo a la transformación propuesta, esta clase de compuestos se genera a partir de un verbo y sus elementos subcategorizados. Como la regla mueve adverbios, adjetivos y sustantivos nos permite obtener:

oven-cleaner a partir de *clean* [_N *oven*]

Los compuestos verbales son descritos como a) predecibles, puesto que su significado es composicional y b) extremadamente productivos. Los compuestos verbales distinguen claramente entre formas “posibles” e “imposibles”:

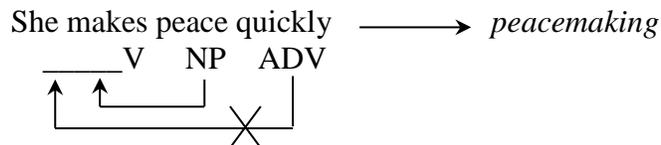
| POSIBLE | IMPOSIBLE |
|----------------------|-----------------------|
| <i>grim-acting</i> | * <i>grim-wanting</i> |
| <i>nice-sounding</i> | * <i>man-sounding</i> |
| <i>sea-going</i> | * <i>huge-going</i> |

Esto puede ser explicado mediante un solo principio: FIRST SISTER (FS) PRINCIPLE (*Principio de Contigüidad*) y quiere decir que todos los compuestos verbales se forman por la incorporación de una palabra inmediatamente contigua al verbo. Al considerar los siguientes pares vemos que algunos no son posibles porque el primer elemento no aparece a continuación del verbo en una oración:

| | |
|-------------------------|---------------------------|
| * <i>peace-thinking</i> | <i>quick(ly)-thinking</i> |
| <i>peacemaking</i> | * <i>quick-making</i> |

En el caso de **peace-thinking* el problema radica en que el verbo *think* necesita una preposición (*about*) antes del complemento directo (*peace*) y **quick-making* resulta anómalo porque el verbo *make* es transitivo y requiere de un complemento directo para que poder dar origen a compuestos como *history-making* o *money-making (game)*.

Esta correspondencia responde al hecho de que tanto las oraciones como los compuestos se forman a partir de marcos de subcategorización (*subcategorization frames*): Un sustantivo, un adjetivo o un adverbio pueden combinarse con un verbo solamente si ocurre en posición FS (*First Sister*):



Roeper y Siegel (1978) estipulan una serie de restricciones para la formación de compuestos. Por ejemplo, muchos de ellos son el resultado de una transformación que involucra una frase preposicional que funciona como adjunto pero estos solo pueden expresar ciertas nociones semánticas como agente, locativo e instrumento:

- grown at home* (locativo)
- built by Indians* (agente)
- made by hand* (instrumento)

A continuación, se elide la preposición y luego se inserta el remanente a la izquierda del núcleo verbal, obteniendo *home-grown*, *Indian-built* y *hand-made*. Los autores restringen el número de frases preposicionales susceptibles de formar compuestos a aquellas que comienzan con *in*, *to*, *with*, *at* y *by*.

Puesto que solo ciertas clases gramaticales (principalmente sustantivos, adverbios y adjetivos) son los únicos complementos que pueden ser incorporados a un compuesto, se excluyen los complementos oracionales, aunque aparezcan en posición inmediatamente contigua.

They prayed for her to return → **for-her-to-return-praying*

Los verbos, aun cuando son ítemes léxicos y no frases, también son excluidos de la formación de compuestos. Incluso en el caso de los verbos subcategorizados para actuar como complementos verbales. Por ejemplo, de una oración como *He hates to eat* podríamos pensar que es posible derivar **eat-hating*, sin embargo, en inglés no se encuentran compuestos con la forma [[V] + verb + affix].

Si bien las formulaciones de Roeper y Siegel (1978) han causado disenso y crítica, ofrecen un marco teórico de naturaleza formal bastante claro y sólido que permite dar cuenta del tipo de adjetivos compuestos que conforman gran parte de los corpus analizados (se excluyen los compuestos de base nominal). Por otra parte, hay que tener en cuenta que el volumen de datos manejados por los autores era muy reducido en comparación con las cantidades disponibles en los corpus electrónicos actuales, por lo tanto, es probable que en los corpus aparezcan compuestos cuya forma no se ajuste a lo postulado en su teoría.

Capítulo 4. Objetivos e hipótesis

4.1. Objetivos Generales

—Abordar la composición como proceso de formación de palabras mediante el estudio de una clase particular de adjetivos compuestos en inglés.

—Describir morfológicamente la clase de adjetivos compuestos con participio (*participial adjective compounds*) a partir de un corpus de prueba y de un corpus de análisis.

4.2. Objetivos específicos

— Analizar un corpus de prueba y cuantificar los patrones morfológicos de los dos grandes grupos de adjetivos compuestos con participio en inglés: *-ed / -ing compounds*.

— Revisar muestras de datos recogidos de un corpus electrónico de uso general en lengua inglesa (*enTenTen12*) para verificar y cuantificar la ocurrencia de estos adjetivos compuestos.

— Reconocer y clasificar los patrones morféimicos de los adjetivos compuestos con participios hallados en ambos tipos de corpus.

— Comparar los datos (numéricos) arrojados por el corpus de prueba con los datos obtenidos del análisis de los corpus generales

— Describir las regularidades de forma y uso encontradas en los diferentes patrones morféimicos.

4.3. Hipótesis

—Los patrones más comunes serán aquellos cuyos constituyentes pertenezcan a las categorías gramaticales básicas, es decir, que combinen sustantivos, adjetivos y adverbios con los dos tipos de participio en inglés.

—En ambos grupos de adjetivos compuestos (*-ed / -ing*), Se verificará la ocurrencia de formas no estándar o emergentes.

—Se observará cierta correspondencia en la frecuencia de ocurrencia de los patrones morfémicos encontrados en ambos tipos de corpus.

Capítulo 5. Enfoque metodológico y metodología

5.1. Lingüística de corpus

La lingüística de corpus, a diferencia de otras áreas de la lingüística, no se trata del estudio de un aspecto particular del lenguaje sino que es un conjunto de procedimientos y métodos para su estudio. Las herramientas y procedimientos que se utilizan están en constante evolución aunque algunos de ellos –como los *concordancers*– ya están bien establecidos y son de amplia aplicación. El desarrollo de la lingüística de corpus y de las tecnologías aplicadas al lenguaje, junto con la creación de corpus de tamaño considerable y de procesadores con la capacidad suficiente para explotarlos, han hecho posible una exploración más acabada de teorías que antes era difícil estudiar de manera empírica. De hecho, la escala de los corpus modernos es de tal magnitud que el análisis manual es sencillamente imposible.

“The set of texts or corpus dealt with is usually of a size which defies analysis by hand and eye alone within any reasonable timeframe. It is the large scale of the data used that explains the use of machine-readable text. Unless we use a computer to read, search and manipulate the data, working with extremely large datasets is not feasible because of the time it would take a human analyst, or team of analysts, to search through the text.” (McEnery y Hardie, 2012: 2)

Las herramientas desarrolladas para el procesamiento de ingentes cantidades de datos le permiten a los analistas revisar los corpus de manera rápida y confiable, permitiendo así determinar con mayor precisión la frecuencia de ocurrencia de un lexema o estructura determinado. Los *concordancers* —que permiten observar cómo funciona una palabra en su contexto lingüístico— y los listados de frecuencia son ejemplos de análisis cualitativos y cuantitativos, respectivamente, lo que los hace importantes para la lingüística de corpus.

Si bien la lingüística de corpus es una disciplina heterogénea en la que no existe un consenso monolítico respecto a los métodos y procedimientos que deben emplearse, una cuestión fundamental

que se aplica a todos los tipos de estudios conducidos en este campo es la correspondencia que debe existir entre la pregunta de investigación y los datos o corpus escogidos para explorarla. Si bien este último punto parece evidente, sus implicaciones son varias. Al momento de escoger el corpus se debe tener presente que no puede consistir en un solo tipo de texto o contener un número limitado de tipos textuales. El grado de homogeneidad de un corpus debe estar acorde con las preguntas de investigación y así garantizar que los descubrimientos realizados puedan generalizarse, en lo posible, a la totalidad de una lengua determinada.

5.2. Corpus lingüístico

La lingüística de corpus busca trabajar con datos reales, en cantidades más o menos grandes, que permitan acercarnos al fenómeno que se busca investigar. Es por este motivo que los corpus escogidos deben ser lo suficientemente representativos de la realidad que se quiere observar. Para Torruella y Llisterri (1999: 2) “la función principal de un corpus, tanto textual como oral, es establecer la relación entre la teoría y los datos; el corpus tiene que mostrar a pequeña escala cómo funciona un lenguaje natural”. Para que el corpus sea un modelo de parte de la realidad lingüística, es de esperar, entonces, que sea neutro, es decir, que esté constituido por muestras proporcionales de todos los aspectos de dicha realidad. De esto se deduce que el análisis de los enunciados contenidos en el corpus “debe permitir mejorar el conocimiento de las estructuras lingüísticas de la lengua que representan” (1999: 8). Es necesario admitir que, aunque el especialista aspira a esa neutralidad, muchas veces esto es más una tendencia que un hecho concreto puesto que resulta inevitable enfocarse en aquello que se quiere demostrar. Cuanto más grande sea el corpus, más posibilidades habrá de que incorpore diferentes aspectos de la lengua (niveles, temas, registros, etc.). Por otra parte, sin embargo, un corpus debe ser necesariamente selectivo ya que es imposible abarcar todo lo escrito o hablado en una lengua.

Un corpus lingüístico computarizado entraña una recopilación de textos seleccionados siguiendo criterios lingüísticos, codificados de modo estándar y homogéneo a los que se pueden aplicar procesos informáticos de diversa índole. El origen y proveniencia de los textos allí contenidos está

debidamente documentado. Los datos sobre los autores, el nivel sociocultural de los participantes, los medios de transmisión usados, etc. son criterios lingüísticos externos al corpus. Si los datos hacen referencia a patrones lingüísticos presentes en los textos, entonces hablamos de criterios internos.

Biber (2010: 159) define el concepto de corpus como “a large and principled collection of natural texts” lo que pone de manifiesto que se construyen siguiendo criterios específicos entre los que destacan la naturaleza de los textos —los que son reales ya sean orales y/o escritos— así como los géneros que incluyen (periodísticos, literarios, científicos, publicitarios, etc.) y, sobre todo, la cantidad empleada, la que se expresa en número de palabras. Un concepto clave, sobre todo para las investigaciones basadas en el corpus, es el de REPRESENTATIVIDAD (*corpus representativeness*) la que depende de dos factores:

- a) Composición: Los textos que componen el corpus deben ser muestreados (*sampled*) de forma deliberada para representar diferentes registros, géneros y/o variedades dialectales, dependiendo de la naturaleza de la investigación que se busca realizar. Por otra parte, los estudios guiados por el corpus muchas veces prefieren ocupar muestras aleatorias (*random sampling*).
- b) Tamaño: El corpus debe ser lo suficientemente grande para representar de manera fidedigna la distribución de los rasgos lingüísticos en estudio.

En relación a este último aspecto, cabe destacar que la lingüística de corpus necesariamente debe utilizar herramientas computacionales que permiten recolectar, almacenar y procesar cantidades ingentes de información. Es por ello que esta disciplina ha experimentado un gran desarrollo en las últimas décadas con la aparición de grandes corpus electrónicos de uso general y la masificación de Internet como fuente inagotable de datos. Por último, sin importar el enfoque que se escoja, la lingüística de corpus depende de técnicas cuantitativas y cualitativas de análisis.

5.2.1. Corpus electrónicos

Los estudios descriptivos que buscan dar cuenta de algún aspecto específico de la lengua han encontrado en los corpus electrónicos una fuente invaluable de datos que pueden ser explotados de muy diversas maneras. Los corpus informatizados ponen a disposición de los investigadores bases reales “muy provechosas para comparar diferentes variedades de una lengua o para explotar sus aspectos cuantitativos y probabilísticos” (Torruella y Llisterri, 1999: 47). Para una investigación como la nuestra, enfocada en profundizar en la descripción de los patrones morfémicos de los adjetivos compuestos con participio en inglés, los grandes corpus como el *British National Corpus* (BNC), el *Corpus of Contemporary American English* (COCA) y el *enTenTen12* el representan herramientas extraordinarias en cuanto a la extensión y variedad de los textos contenidos en ellos. El BNC cuenta con 100 millones de palabras tomadas de muestras de inglés escrito y oral de fuentes y géneros muy variados. El COCA, por su parte, contiene 450 millones de palabras tomadas de textos de cinco géneros distintos, los que incluyen textos orales (de TV y radio), ficción, periódicos y revistas populares y académicas. Además, cuentan con la ventaja de haber sido diseñados por especialistas siguiendo criterios rigurosos que permiten mantener un equilibrio entre las secciones que conforman cada corpus.

5.2.2. Corpus *enTenTen12*

En el campo de la investigación lingüística, los grandes corpus electrónicos han servido para dar un nuevo impulso a los estudios descriptivos puesto que constituyen excelentes herramientas que permiten validar empíricamente diversos aspectos de una lengua. En el caso específico de este estudio, se escogió un corpus del inglés lo suficientemente abarcador y representativo que nos permitiera comparar la frecuencia de ocurrencia de los distintos patrones morfológicos identificados en el corpus de muestra. Para esos efectos seleccionamos el *enTenTen12* que es uno de los corpus contenidos en *Sketch Engine*. Este corpus está compuesto de la siguiente manera:

-Palabras: 11.191.860.036
-Tokens: 12.968.375.937
-Oraciones: 594.433.009
-Párrafos: 228.918.174
-Documentos: 22.223.066

El *enTenTen12* es un corpus etiquetado gramaticalmente siguiendo las convenciones del *Penn TreeBank*. *Sketch Engine* pone a disposición de los usuarios una gran cantidad de corpus en muchas lenguas (inglés, chino, español, árabe, francés, ruso, alemán, turco, etc.) y, de todos ellos, los más grandes son los llamados *TenTen* con más de un billón de palabras cada uno: *deTenTen* 2013 (alemán) 19.918.263.493, *esTenTen* 2011 (español) 10.994.616.207 o *svTenTen* 2014 (sueco) 3.900.846.988.

Sin embargo, no debemos olvidar que, aunque los corpus nos permiten observar cómo funciona el lenguaje, no son el lenguaje mismo. Tampoco debemos pasar por alto que existen otros métodos de recolección de datos que siguen siendo útiles para la investigación lingüística y, sobre todo, que ninguno de ellos puede reemplazar la introspección del analista. Sin esta habilidad, el lingüista no podría formular las preguntas adecuadas para guiar la revisión del corpus, ni reformularlas en caso de encontrar ejemplos que falseen la teoría inicial.

5.2.3. Corpus de prueba

Para llevar a cabo los objetivos planteados, se contó con un corpus de prueba con alrededor de 3.900 ejemplos de adjetivos compuestos con participio, cuyos datos fueron recolectados, en un periodo promedio de cinco años, de muchas fuentes diferentes, entre las que destacan obras literarias como *Wuthering Heights* de Emily Brontë, *Dracula* de Bram Stoker, *Lolita* de Vladimir Nabokov y los volúmenes compilatorios de cuentos cortos de Scott Fitzgerald y W. Somerset Maugham, versiones impresas y electrónicas de revistas como *Vanity Fair* y *Empire* y las versiones electrónicas de *National Geographic*, *The New York Times* y *Daily Mail*. También se incluyen colecciones de

cuentos de ciencia ficción y de suspenso que abarcan prácticamente todo el siglo XX. Resulta imposible dar cuenta acabada de las fuentes de datos puesto que se recopilaron ejemplos de todo el material impreso y digital al que se tuvo acceso durante ese tiempo. Sin embargo, podemos afirmar que, en términos diacrónicos, las fuentes son representativas de la lengua inglesa escrita desde mediados del siglo XIX hasta la primera década del siglo XXI. Una vez que el corpus de muestra tuvo suficientes datos, se analizó la estructura interna de cada compuesto y se identificaron y clasificaron los patrones morfológicos.

5.3. Sketch Engine

*Sketch Engine*⁵ es una plataforma de análisis textual en línea diseñada para realizar diferentes tipos de consultas y búsquedas (concordancias, extracción de términos, listas de palabras, etc.) que permiten comprobar cómo se comporta una palabra colocacional y gramaticalmente. Pero, sobre todo, esta interfaz computacional nos posibilita confeccionar, entre otras cosas, nuestros propios corpus.

En el procesamiento del lenguaje natural (PLN) es de suma importancia definir claramente los límites de las unidades que se van a estudiar. Esta labor presenta diferentes grados de complejidad dependiendo del idioma que se procese dada la gran variedad de lenguas y los diferentes sistemas de escritura que utilizan. El proceso de segmentación de un texto consiste en dividirlo en unidades lingüísticamente significativas y consta de diferentes niveles de análisis. Para nuestra investigación, dos procesos son particularmente relevantes: la tokenización y el etiquetado sintáctico (*part-of-speech-tagging*). La tokenización es el proceso mediante el cual se dividen las secuencias de caracteres en un texto al identificar los límites de las palabras (*word boundaries*), es decir, el punto donde termina una palabra y comienza otra. En lingüística computacional, las palabras identificadas son denominadas *tokens*. En aquellos idiomas en que las divisiones entre palabras no están definidas

⁵ Para mayor información ir a <http://www.sketchengine.co.uk>

con claridad, este proceso se conoce como segmentación léxica (*word segmentation*) y se usa como sinónimo de tokenización. Las diferentes convenciones ortográficas usadas en inglés para las unidades compuestas (presencia o ausencia de guión) constituyen rasgos que deben ser tomados en cuenta para segmentar el texto de manera exitosa y que los *tokens* así obtenidos guarden correspondencia con los compuestos que conforman el corpus de prueba (Palmer, 2000).

Por otra parte, al analizar los grandes corpus, también debemos tener presente que la ortografía de un compuesto ya identificado puede variar debido a múltiples factores que van desde factores idiosincráticos como el desconocimiento de las “reglas” de ortografía o la preferencia de un autor por una forma sobre otras, hasta las imposiciones emanadas de las editoriales. Hay que tener en cuenta que los corpus actuales también contienen muestras tomadas de la Web que incluyen blogs y comentarios de los usuarios, los que al hacer uso del lenguaje de manera bastante libre y espontánea pueden emplear formas “anómalas” (*ill-formed*) de las estructuras buscadas. Además de estos factores, debemos recordar que en inglés el uso del guión en construcciones compuestas no es siempre uniforme por lo que al utilizar un *tokenizer* se debe tratar de especificar muy bien el uso del guión que sirve a nuestras búsquedas.

En cuanto al segundo proceso, el etiquetado sintáctico, la mayoría de los grandes corpus electrónicos ya poseen algún tipo de anotación. Esto quiere decir que ya se ha introducido información sintáctica que ayuda a categorizar y reconocer el tipo de palabras. El etiquetado le asigna una categoría gramatical (sustantivos, verbos, adjetivos, etc.) a las palabras contenidas en el corpus y agrega información relevante como persona y número. También incorporan rasgos morfológicos como sufijos, prefijos, guiones, etc. Aun cuando este proceso es realizado por un etiquetador automático (*automatic tagger*) es posible refinar las consultas para que la búsqueda y reconocimiento de ítems sean todavía más precisos.

5.4. Investigaciones guiadas o basadas en un corpus

Mi investigación se enmarca dentro de la corriente de estudios descritos como “guiados por un corpus” (*corpus-driven*) puesto que no está basada en una descripción estandarizada de los adjetivos compuestos con participio en inglés. Por el contrario, mi punto de partida es un corpus de muestra recolectado por la propia analista del cual surgieron los patrones morfológicos que luego serían comparados con aquellos encontrados en un gran corpus electrónico. Creo que es necesario explicar cómo se inscribe este enfoque metodológico en la lingüística de corpus y qué alcances y aplicaciones tiene dentro de este tipo de estudios.

Uno de los principales objetivos de la lingüística de corpus es apoyar empíricamente investigaciones descriptivas de muy diversa índole para lo que se analizan patrones reales de uso en textos naturales. Existen dos enfoques metodológicos que buscan cumplir con ese propósito y que guían todos los estudios: *corpus-based* y *corpus-driven research*. Las investigaciones basadas en corpus (*corpus-based*) buscan validar, refinar o refutar una teoría sobre un fenómeno lingüístico y para ello se valen de un corpus del cual extraen los datos necesarios para ahondar en la teoría escogida / subyacente. Según esta definición, la lingüística de corpus es una metodología utilizada para explorar un modelo teórico determinado.

“Corpus-based research assumes the validity of linguistic forms and structures derived from linguistic theory. The primary goal of research is to analyse the systematic patterns of variation and use for those pre-defined linguistic features”. (Biber, 2010: 162)

Las investigaciones basadas en un corpus son altamente confiables puesto que estos se construyen de acuerdo a criterios definidos que aseguran su representatividad. Por ende los resultados son generalizables, lo que demuestra que poseen validez externa. Aun cuando los análisis basados en un corpus son considerados más tradicionales ya que emplean categorías reconocidas por otras teorías y no tienen como objetivo descubrir nuevas estructuras o rasgos lingüísticos, con frecuencia se producen descubrimientos que no pueden ser anticipados por los analistas ya que se apartan de las teorías lingüísticas estándar.

Por otra parte, el término *corpus-driven* se refiere a aquella rama de la lingüística de corpus que lo considera como la fuente de la cual se extraen no solo datos sino que las hipótesis que dan forma a una teoría del lenguaje. En otras palabras, este segundo tipo de estudios no descansa sobre una teoría lingüística determinada sino que las hipótesis se construyen inductivamente a partir del análisis de un corpus.

“Corpus-driven linguistics rejects the characterisation of corpus linguistics as a method and claims instead that the corpus *itself* should be the sole source of our hypotheses about language. It is thus claimed that the corpus itself embodies a theory of language”. (Tognini-Bonelli 2001: 84-5)

A diferencia del enfoque anterior, las investigaciones guiadas por un corpus exploran y explotan la posibilidad de descubrir formas emergentes y/o desconocidas puesto que las categorías lingüísticas son derivadas sistemáticamente de patrones recurrentes que surgen del lenguaje en contexto y de su frecuencia de distribución. (Biber, 2010: 168). En términos generales, los análisis guiados por un corpus son más innovadores en tanto apuntan a descubrir formas lingüísticas que no son reconocidas por las teorías tradicionales.

5.5. Metodología

Esta investigación se propuso realizar una comparación entre los patrones morfémicos de los adjetivos compuestos con participio encontrados en un corpus de prueba recolectado por la propia analista con aquellos extraídos de un gran corpus electrónico. Para ello se debía utilizar una herramienta de búsqueda adecuada para primero extraer muestras lo suficientemente abarcadoras de adjetivos compuestos para luego proceder a analizarlos mediante recursos computacionales que nos permitieran determinar y cuantificar los distintos patrones morfológicos al interior de cada gran grupo. Todo ello con la finalidad de comprobar la frecuencia de aparición de cada patrón y poder comparar estos datos con los obtenidos a partir del corpus de prueba y verificar las similitudes en cuanto a ocurrencia y frecuencia de los patrones combinatorios de estas unidades compuestas. Esto tenía como objetivo no solo corroborar cuáles son más comunes dentro de las formas regulares o estándar sino que constatar si dentro de los menos frecuentes se registran ejemplos de compuestos

con formas no canónicas o emergentes. A continuación se describen los aspectos más relevantes de la metodología empleada.

5.5.1. Búsqueda y extracción de datos

Como ya se ha establecido, los adjetivos compuestos con participio, además de poseer una gran diversidad morfológica, conforman una categoría muy productiva, lo que hace que el número de compuestos lexicalizados sea bastante bajo. Por esta razón, uno de nuestros principales objetivos era poder recolectar grandes muestras de las dos grandes categorías de adjetivos compuestos (*-ing* y *-ed*) para luego determinar su estructura interna y así verificar, mediante datos fidedignos, la frecuencia de ocurrencia de los distintos patrones morféimicos. La recolección de datos, sin embargo, resultó ser un proceso complicado y laborioso.

La primera dificultad consistió en que este tipo de lexemas no están identificados como tales por el etiquetador sintáctico del corpus escogido, el *Penn TreeBank*⁶, el cual no posee una etiqueta (*tag*) específica para compuestos. El listado de etiquetas (*tagset*) contempla tres para adjetivos: JJ (adjetivo), JJR (adjetivo comparativo) y JJS (adjetivo superlativo). Esto quiere decir, que los adjetivos compuestos con participio están considerados dentro de la categoría general adjetivo (JJ) junto con millones de otros tipos de adjetivos. El primer problema a resolver era cómo lograr extraer los datos correctos. Como las categorías sintácticas no eran de utilidad, nuestras primeras búsquedas incluyeron expresiones regulares que tenían al guión como rasgo distintivo central:

```
[tag = ".*" & word = ".*-.*ing"]
```

Como esta consulta describe cualquier secuencia de dos palabras que contenga un guión y cuyo segundo elemento termine en *-ing*, los datos recolectados incluían muchísimos elementos que no eran adjetivos, así como otros que ni siquiera eran estructuras compuestas sino palabras derivadas

⁶ Para mayor información ir a <https://www.sketchengine.co.uk/penn-treebank-tagset/>

con prefijación (*off-spring*, *de-beaking*, *co-ordinating*, etc.). Otros resultaron ser sustantivos compuestos cuya apariencia es exactamente igual a la de los adjetivos (*tree-planting*, *well-being*, *fund-raising*) y que fueron excluidos al revisar el contexto en el que ocurrían. Tomando en cuenta que los adjetivos pueden aparecer tanto en posición atributiva (delante del sustantivo) como predicativa (después de un verbo), las búsquedas posteriores se perfilaron de mejor manera utilizando las siguientes expresiones regulares:

Posición atributiva: [tag = "DT"] [tag = "JJ" & word = ".*-.*ed|.*-.*ing"] [tag = "NN"]

Posición predicativa: [tag = "V.*"] [tag = "JJ" & word = ".*-.*ed|.*-.*ing"] (todos los tiempos)

Cabe destacar que al especificar la posición, los resultados demostraron categóricamente que los adjetivos compuestos en inglés ocurren con mucha mayor frecuencia como pre-modificadores de un sustantivo con 1.876.146 instancias comparadas con las 1.041.970 ocurrencias en posición predicativa.

Estos resultados fueron corroborados por los datos obtenidos respecto a la ocurrencia de los dos grupos principales de adjetivos compuestos con participio en ambas posiciones, como se detalla en la siguiente tabla:

Tabla 5. Totales por posición de adjetivos compuestos

| Tipo de compuesto | Posición atributiva | Posición predicativa |
|-------------------|---------------------|----------------------|
| <i>-ed</i> | 1.411.374 | 840.084 |
| <i>-ing</i> | 464.772 | 201.886 |

Más tarde buscando realizar una consulta unitaria estas dos expresiones se fusionaron en una sola en la que se especifica que el determinante (DT) y el sustantivo (NN) son optativos, mediante la instrucción {0,1}, así como el verbo, el cual no contiene especificación alguna:

```
[tag = "DT"] {0,1} [tag = "V.*"] {0,1} [tag = "JJ" & word = ".*-.*ed|.*.*ing"]  
[tag = "NN"]{0,1}
```

Sin embargo, esta consulta resultó ser demasiado general ya que expandió la búsqueda a 57.515.333 resultados. Al restringir el elemento verbal a “ser/estar” (be) en 3era persona singular (VBZ) o en cualquiera de las otras formas del presente (VBP) la consulta arrojó 21.645.144 instancias, lo que todavía era un número extraordinariamente alto y que, además, contenía redundancias como la identificación de un compuesto varias veces en un mismo entorno sintáctico con y sin los elementos optativos.

Como estas modificaciones expandieron considerablemente la recolección de datos, decidimos realizar búsquedas individuales para cada categoría de adjetivos compuestos. El tener los datos separados en sub-corpus más pequeños hizo mucho más manejable el procesamiento de los datos. Debemos señalar que aun cuando el número adjetivos compuestos identificados en nuestras búsquedas iniciales alcanzaba varios millones, *Sketch Engine* solo permite extraer muestras aleatorias de 100.000 ítems como máximo.

Al poco tiempo nos dimos cuenta que en nuestras consultas estábamos dejando fuera aquellos compuestos cuyo segundo elemento es un participio pasado irregular como en *well-known*, *long-forgotten* o *wind-swept*. Probamos una expresión bastante general que tenía como primer elemento un adjetivo (JJ) y cuyo segundo constituyente era simplemente un verbo (V): [tag = "JJ"] "-" [tag = "V.* "]. La falta de precisión de esta consulta (*query*) produjo muchos ítems que no se ajustaban a las especificaciones (*olive-brown*, *full-tilt*, *now-absent*) o que lo hacían parcialmente (*un-spoilt*, *not-written*, *out-spoken*).

Aun cuando muchos participios irregulares tienen la terminación *-en* (*taken*, *seen*, *written*, etc.), una búsqueda acotada arrojó datos como *even-steven*, *microwave-oven* que simplemente muestran una similitud superficial con los adjetivos en estudio. Para lograr que la herramienta de búsqueda identificara compuestos que contienen participios como *taught* o *brought* se incluyó la

instrucción `.*.*ught` en la expresión regular. En otros casos fue necesario especificar la totalidad de la forma verbal (*done, beat, held*). El resultado final no fue conciso pero sí bastante efectivo:

```
[tag = "DT"] [tag = "JJ" & word = ".*.*en|.*.*ught|.*.*held|.*.*eant|.*.*done|.*.*sent|.*.*hid|.*.*bent|.*.*spent|.*.*tood|.*.*eft|.*.*shot|.*.*shod|.*.*beat|.*.*clad|.*.*ground|.*.*bound|.*.*pt|.*.*uck|.*.*cast|.*.*wn|.*.*orn|.*.*nk|.*.*set|.*.*ung|.*.*ade|.*.*ead|.*.*lit|.*.*un|.*.*lt|.*.*aid|.*.*cut|.*.*girt"] [tag = "NN.*"]
```

Los resultados de esta búsqueda también confirmaron la predominancia de la posición atributiva por sobre la predicativa:

Tabla 6. Totales por posición de adjetivos compuestos irregulares

| Tipo de compuesto | Posición atributiva | Posición predicativa |
|-------------------|---------------------|----------------------|
| irregulares | 336.142 | 180.550 |

A pesar del detalle de las instrucciones dadas, en los datos recolectados todavía se observaban ítems como *drop-down* y *half-dozen* los que para la herramienta de búsqueda se ajustaban a los criterios especificados pero que evidentemente debían ser eliminados (el primero es una forma verbal y el segundo un sustantivo) así como otros muchos cuyo segundo elemento solo coincidía en forma con los participios irregulares (*-open, -porn, -adult, -belt, -fun, -bankrupt, -lung, etc.*). Todos ellos tuvieron que ser identificados y eliminados manualmente.

5.5.2. Depuración del corpus

Una vez recolectados los datos, se eliminó la mayor parte del contexto en que aparecían los compuestos con el propósito de facilitar la remoción de las repeticiones, puesto que un mismo compuesto puede aparecer en diferentes contextos y la plataforma los considera instancias distintas, como se puede apreciar en la siguiente captura de pantalla:

The screenshot shows the Sketch Engine search interface. At the top, the search term 'well-known' is entered in the search bar, and the corpus is identified as 'English Web 2012 (enTenTen12, sample 40M)'. A sidebar on the left contains navigation options such as 'Inicio', 'Buscar', 'Listas', 'Word sketch', 'Tesaurus', 'Sketch dif', 'Corpus info', 'Mis tareas', 'Guía de usuario', 'Guardar', 'Make subcorpus', 'Opciones de visualización', 'KWIC', 'Oración', 'Ordenación', 'Contexto a la izquierda', 'Contexto a la derecha', 'Nodo', and 'Referencias'. The main content area displays search results for 'well-known', showing 915 occurrences across 46 pages. The first page is selected. The results list various URLs followed by text snippets where 'well-known' is highlighted in red. Examples include:

- mhra.org.u... undertake its production". It is based on two well-known episodes in Pacific exploration : the voyage
- peskythera... the planet. Indian cuisine in Glasgow are well-known amongst people there. From road foods to
- singaporet... For fast food, McDonalds and KFC are very well-known , alongside a scattering of Dairy Queens
- checkpoint... Kay Klein's 'Disclave 68' report. JKK is well-known for his excellent con photography and the
- meyerweb.c... conversation with an editor at a relatively well-known and respected publisher about a CSS book
- bostonhera... combative. Santorum's prickly personality is well-known in western Pennsylvania, where he grew
- creativewr... lead many of the discussions, and bring well-known writers and industry professionals in to
- mysticfayr... since the Angel Raphael appeared to her. Well-known for starting the Cornwall New Age Festival
- esri.mmu.a... perspectives. In addition to corroborating well-known unintended consequences (eg increased test
- gimv.com... cardiovascular disorders, and is especially well-known for its TEE (Trans Esophageal Echocardiography
- rightwingw... Parenthood is one of the largest and most well-known groups working in the extensive field of
- helmets.or... several years ago at about 15mph. There is a well-known curve in the road with a large bump in
- ctwhale.co... Hartford for the 1974-75 season. Barnes was well-known for his marketing prowess, even being considered
- mht.maryla... Odd Fellows Lodge. In fact, the lodge was well-known for its dances because it was one of the
- wd40.com... Once voted in, the new members will join well-known celebrities on the Board, including: Daytona
- blogcampai... to outshine their much larger, older and well-known competitors by understanding the emotional
- modperl.pl... Later, they were also ported to Microsoft's well-known Windows Operating System, and we have included
- pokerstars... one on day two against David Bach, a very well-known mixed game cash player. He plays very high-stakes
- content-8... pen name, Ann Fiery , in addition to your well-known children's books, the Ivy and Bean Series
- oriona.biz... to start is in affiliate marketing... A well-known foreboding statistic is that 75 percent

Figura 10. Ejemplos de concordancias del compuesto *well-known*

En la imagen se puede apreciar que el adjetivo compuesto *well-known* registra 915 ocurrencias en el corpus electrónico (ver **búsqueda** arriba a la izquierda) el que provee los ejemplos con su contexto inmediato (alrededor de siete palabras a izquierda y derecha del término buscado). Esta es la primera página de un total de 46. Para terminar de depurar los datos, se utilizó *Python*, un lenguaje de programación que nos permite desarrollar programas o códigos para realizar tareas específicas como pre-procesar los textos (eliminar duplicados considerando palabras o patrones particulares). Finalmente, se separaron los elementos al interior del compuesto incorporando un espacio antes y después del guión para permitir el re-etiquetado de cada componente al interior del compuesto.

Uno de los objetivos principales de esta investigación es verificar y cuantificar la ocurrencia de los patrones morfémicos de los adjetivos compuestos en un gran corpus electrónico, para luego compararlos con los patrones encontrados en el corpus de prueba en el que se habían registrado ocurrencias individuales (*types*). Por ende, habiendo realizado las consultas y recolectado los datos,

el siguiente paso era depurarlos para equiparar ambos corpus en términos de confiabilidad de la información compilada. Por lo tanto, había que eliminar tanto los errores de selección como las repeticiones. Para esto se utilizó un código (*script*) en *Python* que redujo considerablemente la cantidad de datos. Sin embargo, nuestras listas todavía contenían un sinnúmero de elementos que debían ser identificados y eliminados:

- 1) Compuestos no adjetivales cuyos segundos elementos consistían en sustantivos con terminaciones similares a los adjetivos buscados (*-citizen, -oxygen, -teen, -evening, -read, -something, -lawn, -string, -wing, -need, -bed, etc.*). Algunos ejemplos son *small-breed, tall-building, go-karting, left-wing, matching-substring, twenty-something, etc.*
- 2) Palabras compuestas que contenían un prefijo como primer constituyente (*semi-, pre-, sub-, extra-, micro-, anti-, -mid-, counter-, etc.*) como *ex-Viking, hyper-ventilating, semi-detached house, etc.*
- 3) Combinaciones de palabras en las que el constituyente de la derecha es un nombre propio como *Roosevelt, Doreen, Egypt* o *Connecticut*. También nombres compuestos o coordinados: *COSMO-SkyMed, Hopkins-Nanjing, etc.*
- 4) Compuestos coordinativos: *browsing-lending, teaching-learning, oscillating-rotating, etc.*
- 5) Formas compuestas que contenían errores de ortografía (*well-kinown, taylor-made, tailer-made, etc.*).
- 6) Palabras unidas por un guión que no eran compuestos (*stroke-struck, year - - brought, while-read, f-cking, gosh-fricking, Boing-Boing, zig-zagging, etc.*).
- 7) La presencia de mayúsculas al inicio de compuestos cuyo primer elemento no las necesita (*Well-known, Locally-made, etc.*), así como al interior de los compuestos, producía repeticiones, por lo que fue necesario verificar manualmente si un compuesto aparecía listado de otra forma para evitar repeticiones. Por ejemplo, el sustantivo *earth* (tierra) puede

escribirse con mayúscula si se refiere al planeta, por lo tanto, dependiendo de la ortografía es posible encontrar el mismo compuesto dos veces: *Earth-saving* y *earth-saving*.

- 8) Las minúsculas en compuestos cuyo primer elemento era un nombre propio, una nacionalidad o un país (*soviet-born*, *london-based*, *swiss-made*, *ottawa-run*, etc.) fueron corregidas o bien sirvieron para eliminar repeticiones si los compuestos en cuestión aparecían también con mayúsculas (*Soviet-born*, *London-based*, *Swiss-made*, etc.).
- 9) Las diferencias ortográficas entre inglés británico y norteamericano (*colour/color*, *travelled/traveled*, *energise/energize*) produjeron pares como *light-coloured/ light-colored*, *much-travelled/much-traveled* y *self-energising/ self-energizing*, en los que uno de los “duplicados” debió ser eliminado.

Para los casos 1) y 2) se realizaron listas de paro (*stop lists*) en *Python* que contenían palabras y prefijos específicos para eliminar. A modo de ejemplo, este es el listado de elementos detectados que debían ser removidos de la muestra de compuestos con participios irregulares:

Palabras que deben ser excluidas como segundo elemento del compuesto:

['down', 'Down', 'open', 'dozen', 'green', 'even', 'uneven', 'present', 'exempt', 'town', 'seven', 'descript', 'concept', 'Concept', 'porn', 'belt', 'instant', 'own', 'sun', 'teen', 'thread', 'fun', 'anointment', 'background', 'ten', 'brown', 'Brown', 'rebound', 'bankrupt', 'lung', 'adult', 'speed', 'make', 'tilt', 'chicken', 'citizen', 'screen', 'young', 'cult', 'dept', 'breakdown', 'heighten', 'downtown', 'difficult', 'platen', 'Egypt', 'absorbent', 'then', 'script', 'Script', 'Zen', 'Kindergarten', 'oven', 'Sept', 'absent', 'slidden', 'Nineteen', 'between', 'when', 'lawn', 'Tibeten', 'drought', 'corn', 'taylor', 'often', 'meltdown', 'Meltdown', 'contempt', 'stun', 'wooden', 'lorn', 'corrupt', 'nondescript', 'prompt', 'dung', 'stubborn', 'fifteen', 'unknown', 'fraught', 'keen', 'attempt', 'golden', 'token', 'adopt', 'manuscript', 'beat', 'gahung', 'den', 'noot', 'inept', 'result', 'lockdown', 'crown', 'broadcast', 'unsung', 'Ampt', 'peppercorn', 'mindset', 'toten', 'apt', 'elven', 'evergreen', 'heathen', 'norn', 'recumbent', 'Héroult', 'sudden', 'accept', 'forlorn', 'Born', 'conementeen', 'Opt', 'opt', 'gluten', 'pollen', 'Starspaw', 'Paid', 'maven', 'oxgyen', 'oxygen', 'consult', 'consent', 'raid', 'underground', 'Pashtun', 'maiden', 'ought', 'unpaid', 'Ofelt', 'roosevelt', 'versicherung', 'killeramateur', 'kimora', 'transcript', 'Uncut', 'uncut', 'Plaid', 'plaid', 'pulldown', 'Watertown', 'omnipresent', 'yadun', 'guilt', 'Locally', 'localy', 'london', 'Rakuten', 'rung', 'drunken', 'Lycra', 'Connecticut', 'Medicaid', 'dihydropsoralen', 'braid', 'popcorn', 'dawn', 'eighteen', 'heaven', 'milgrain', 'forecast', 'Doreen', 'renown', 'Native', 'Natural', 'NAVY', 'Comcast', 'Widespread', 'widespread', 'draught', 'Heaven', 'hometown', 'preen', 'eleven', 'touchdown', 'rethread', 'seatbelt', 'Simulcast', 'biden', 'unspoken', 'ken', 'olden', 'tread', 'ottawa',

'multicast', 'pakistani', 'playground', 'gilt', 'Teen', 'quilt', 'delt', 'fen', 'outspoken', 'polyster', 'interrupt', 'ebuilt', 'Empt', 'Telecast', 'membersown', 'laiden', 'adept', 'midtown', 'arthen', 'Benteen', 'rewell', 'korn', 'dissent', 'haven', 'sculpt', 'thirteen', 'Upaid', 'Oden', 'vixen', 'foamgreen', 'tought', 'takedown', 'adapt', 'kitten', 'clown', 'unicast', 'Baught', 'sorbent', 'newborn', 'upsidedown', 'soviet', 'rotten', 'studden', 'bun', 'yen', 'unborn', 'unbroken', 'unproven', 'unwritten', 'except', 'shorten', 'showdown', 'dread', 'sixteen', 'queen', 'clubseventeen', 'dropdown', 'incumbent', 'unbeaten', 'knockdown', 'kreen', 'breakingsudden', 'unbound', 'ratedhid', 'tighten', 'asset', 'podcast', 'ptnosent', 'anddown', 'repellen', 'theborn', 'gown', 'hidded', 'kempt', 'nown', 'writen', 'xnown', 'bread', 'woollen', 'reknown', 'melt', 'unseen', 'thicken']

Prefijos y otras palabras a eliminar como primer elemento del compuesto:

(semi|pre|re|un|pro|non|one|six|post|sub|to|co|off|on|top|mid|in|E|super|per|anti|extra|mis|over|mag|not|micro|ps
eudo|multi|four|he|she|screen|Gof|tax|dead|two|five|seven|three|apparent|common|starburst|grozen|GMail|Hom
e|Iron|israeli|jagged|javascript|jerseyWell|jokey|Man|memberane|Menzel|Mid|min|mini|miss|mostwell|muggle|
MUST|neo|newcient|obviously|olfld|Photo|pneunatically|polish|Portugese|striken|precision|proto|should|so|str
oe|ultra|very|weell|wel|wll))(-\w+)+(\w+|\$)'

El proceso de limpieza de datos requirió mucho tiempo y trabajo puesto que se necesitaron varias revisiones manuales por parte de la analista para confeccionar las listas de paro y así poder asegurar que los datos finales fueran realmente fidedignos. Una vez ejecutado el procedimiento en *Python*, se realizó una revisión final en la que se eliminaron todos aquellos elementos que no fueron detectados anteriormente. Solo de esta manera podrían obtenerse resultados confiables al momento de analizar los patrones morfológicos de los adjetivos compuestos recolectados a partir del corpus electrónico.

La importancia y magnitud de la depuración realizada queda de manifiesto en las reducciones que sufrieron los tres grupos de datos originales:

Tabla 7. Totales depurados iniciales

| Tipo de compuesto | Muestra original | Sin repeticiones | Resultado final |
|-------------------|------------------|------------------|-----------------|
| <i>-ed</i> | 77.287 | 22.556 | 20.586 |
| <i>-ing</i> | 25.094 | 5.604 | 4.830 |
| irregulares | 7.358 | 6.951 | 6.422 |

La tabla anterior pone de manifiesto uno de los mayores problemas encontrados al momento de recolectar los datos: la ingente cantidad de repeticiones contenidas en las muestras. Al estar enfocada la investigación en determinar los patrones morfológicos contenidos en estas tres categorías y no la frecuencia de palabras, nuestro interés radicaba en recolectar/muestrear un solo ejemplar de cada forma compuesta encontrada en el corpus electrónico, es decir un tipo (*type*) de cada una de ellas. Para lograr esto era indispensable deshacerse de los duplicados los que, como es de esperar, representan un gran porcentaje de todo corpus recolectado.

En el caso de los dos grandes grupos de adjetivos compuestos (*-ed / -ing*), quitar los ítems repetidos resultó en una disminución del 70,8 % y 77,6%, respectivamente. La muestra de compuestos irregulares se redujo tan solo en un 5,5% probablemente debido a que la expresión regular especificaba cada terminación buscada, lo que ayudó a identificar datos de manera mucho más precisa. Esto lleva a plantearse la necesidad de continuar refinando las consultas (*queries*) usadas para obtener las muestras.

Como la categoría de compuestos terminados en *-ed* era mucho más numerosa que las otras dos, se extrajo una muestra aleatoria de 5.000 unidades. De igual manera se hizo con las formas irregulares dado que, dentro del total de adjetivos compuestos, aquellos cuyo núcleo es un participio irregular (*spoken, fought, grown, etc.*) representan un porcentaje menor. En este caso, la muestra aleatoria consistió de 1.000 ejemplos. A excepción de esta última, las muestras continuaron siendo depuradas como parte del proceso de identificación y análisis de los distintos patrones morfológicos hasta llegar a las cifras finales. A continuación se describen algunas de las dificultades encontradas en las etapas superiores de dicho proceso.

Tabla 8. Totales depurados finales

| Tipo de compuesto | Muestra original | Resultado final |
|-------------------|------------------|-----------------|
| <i>-ed</i> | 5.000 | 4.808 |
| <i>-ing</i> | 4.830 | 4.279 |
| irregulares | 6.422 | 1.000 |

5.5.2.1. Depuración final: dificultades adicionales

En la etapa final del análisis de datos, nos percatamos de la presencia de términos compuestos que debían ser excluidos de las muestras tomadas del corpus *enTenTen 2012*. Por su forma e incluso por su naturaleza, su detección en listas de palabras de este tamaño no es una tarea fácil. Por este motivo, muchos de ellos solo pudieron ser identificados al realizar las consultas individuales en *Sketch Engine*, las que nos permitieron revisar los datos de manera parcelada. A continuación explicamos las dificultades que rodearon el proceso de depuración final según el tipo de palabras involucradas.

5.5.2.1.1. Elementos compositivos como primer elemento

En general, los elementos compositivos se combinan entre ellos o con un lexema independiente para crear una palabra como es el caso de *tele-* en *television*, *eco-* en *ecology* y de *geo-* en *geophysics* (cf. sección 3.5.2.). Sin embargo su estatus no está del todo claro cuando aparecen en solitario como parte de *tele-operated*, *eco-approved* y *geo-referenced*. ¿Estamos en presencia de estructuras compuestas o de palabras que incorporan un elemento combinante de origen clásico (*classical combining form*) que funciona como prefijo? También cabe preguntarse si *eco-* y *geo-* son en realidad truncamientos de *ecologically* y *geographical*, respectivamente, lo que facilitaría clasificar estas unidades como compuestos.

5.5.2.1.2. Adjetivos como núcleo

En lo que respecta a los adjetivos compuestos terminados en *-ed*, se detectaron algunos compuestos cuyo núcleo es un adjetivo (*half-naked, self-interested, nationally-renowned*) por lo que debían ser eliminados de los corpus. En aquellos terminados en *-ing*, las pocas ocurrencias registradas incluían *likely-interesting, vaguely-interesting, throughout-interesting* y *gender-nonconforming*. Su forma hace muy difícil que el etiquetador automático contenido en *Sketch Engine* distinga estos adjetivos de un participio pasado regular, de un sustantivo del tipo N-*ed* o de un participio presente. A veces es necesario analizar semánticamente el compuesto para determinar si el segundo elemento es un adjetivo o un participio. A modo de ejemplo, aun cuando existe el verbo *to preoccupy* (preocupar; ocupar o tomar posesión de algo con antelación) en el caso de *self-preoccupied* (= *self-interested*) éste debe ser interpretado como un adjetivo y no una forma verbal. Roeper y Siegel (1978) en su estudio de compuestos de base verbal distinguen entre formas pasivas lexicalizadas (*lexical passives*) y participios verbales. Los primeros corresponden a participios que se comportan como adjetivos (también llamados *participle adjectives*) como *educated* y *respected* los que, por lo tanto, pueden ser pre-modificados por intensificadores (*really educated, very respected*). Los participios, por otra parte, solo permiten el cuantificador *much*, lo que explica compuestos tales como *much-visited* o *much-loved*. Si bien estos especialistas entendían las implicaciones que esta distinción tiene para la descripción de estas estructuras compuestas y sus correspondientes patrones morfológicos, no podían anticipar las dificultades que presentan para una investigación de corpus que utiliza herramientas tecnológicas.

5.5.2.1.3. Formas derivadas mediante *un-* y *under-* como segundo elemento

El prefijo *un-* en el segundo elemento del compuesto (núcleo) generalmente es un indicador de que la palabra es un adjetivo, como lo demuestra *previously-unknown* cuya estructura es Adverbio + Adjetivo y, por lo tanto, debe ser eliminado del corpus. La revisión manual permitió identificar otros compuestos similares, los que fueron debidamente removidos:

unchanged (otherwise-unchanged)

undiscovered (hereto-undiscovered)

unlimited (almost-unlimited)

unexplained (still-unexplained)

unconfirmed (still-unconfirmed)

unfulfilled (often-unfulfilled)

undecided (still-undecided)

untitled (still-untitled)

unresolved (still-unresolved)

uncensored (best-uncensored)

unparalleled (near-unparalleled)

Cabe mencionar que dos de estos adjetivos han dado origen a sendos verbos mediante formación regresiva (*backformation*): *to unfulfill* (incumplir) y *to uncensor* (quitar la censura), por lo tanto, es posible que en el futuro tanto *unfulfilled* como *uncensored* sean interpretados como participios. En cuanto a los compuestos terminados en *-ing*, se registró un solo caso de un adjetivo con el prefijo *un-* en posición nuclear: *unyielding* (inflexible) en *most-unyielding*.

No obstante lo anterior, debemos evitar generalizar y descartar todos los compuestos cuyos núcleos contienen el prefijo *un-* dado que existen algunos verbos que lo incorporan (*to unpack*, *to uncover*, *to undress*) y cuyos participios pueden ser la base de un adjetivo compuesto. Por ejemplo, *to unveil* (revelar), *to unhinge* (trastornar, desquiciar) y *to unfurl* (desplegar) son parte de *self-unveiled*, *hormone-unhinged* y *slowly-unfurling*, respectivamente.

También fue preciso revisar los compuestos en los que el prefijo *under-* es parte del núcleo, ya que podían tratarse de adjetivos como *underpopulated* y *underpowered*. Sin embargo, debemos tener presente la existencia de verbos que contienen dicho prefijo, tales como *to undermine*, *to underline* y *to undersell*, cuyos participios (*undermined*, *underlined*, *undersold*) podrían ser parte del proceso de composición de nuevos adjetivos. Por lo pronto, en los corpus se verificó la ocurrencia de *often-underexploited* y *medically-underserved*, en los que el núcleo es una forma verbal.

5.5.2.1.4. Adverbios sin sufijación

La presencia de adverbios como elemento inicial en estos compuestos es muy frecuente y conforma uno de los patrones morfológicos más productivos en ambas categorías. La mayoría de ellos incorporan formas adverbiales derivadas de adjetivos mediante el sufijo *-ly*:

| | |
|----------------------------------|------------------------------|
| <i>wirelessly-connected</i> | <i>newly-developing</i> |
| <i>historically-grounded</i> | <i>luxuriously-relaxing</i> |
| <i>professionally-presented</i> | <i>contractually-binding</i> |
| <i>constitutionally-mandated</i> | <i>universally-appealing</i> |

Además de los adverbios que tienen una forma derivada estándar, existen algunos que comparten forma con un adjetivo (*long, hard, fast, high, low*) y otros que, al interior de un compuesto, pueden no incorporar el sufijo como ocurre con *newly* y *freshly* y sus correspondientes variantes *new* y *fresh* en *new-started* y *fresh-killed*. Los etiquetadores sintácticos tienen problemas para identificarlos correctamente y, por lo tanto, un alto número de ellos son clasificados como adjetivos (cf. sección 8.2.3.).

Tabla 9. Distinción entre adverbios y adjetivos

| Adjetivos | Adverbios |
|----------------------|----------------------|
| <i>long-fingered</i> | <i>long-divorced</i> |
| <i>fresh-feeling</i> | <i>fresh-made</i> |
| <i>new-looking</i> | <i>new-found</i> |
| <i>high-spirited</i> | <i>high-paid</i> |
| <i>low-spirited</i> | <i>low-paid</i> |
| <i>hard-bodied</i> | <i>hard-working</i> |

5.5.2.1.5. Verbos compuestos

En los corpus se encontraron formas conjugadas de verbos compuestos como *jump-starting*, *deep-fried*, *dry-cleaned*, *sandblasted*, etc. las que deben ser identificadas y removidas manualmente. Asimismo, muchos compuestos dan origen a formas verbales mediante formación regresiva y conversión como lo ejemplifican *chain-smoker* → *to chain-smoke* y *blackmail* → *to blackmail* (cf. secciones 1.6.5. y 1.6.7.), cuyas conjugaciones en presente continuo (*chain-smoking*, *blackmailing*) tampoco deben ser confundidas con adjetivos compuestos. De hecho, Roeper y Siegel (1978) advierten sobre la confusión entre compuestos verbales (*verbal compounds*) y verbos compuestos (*compound verbs*) a los que denominan “compuestos verbales aparentes.” El problema se complica en casos como *house-cleaning* y *tailor-made*, cuyo estatus es ambivalente: pueden ser verbos conjugados o adjetivos compuestos. Solo el contexto aclarará si deben ser eliminados del corpus.

5.5.2.1.6. Formas derivadas

Como hemos señalado anteriormente, los límites entre derivación y composición pueden ser bastante difusos y, además, ambos procesos pueden afectar a una misma palabra en etapas sucesivas (cf. 1.6.6.). Debemos tener presente que los constituyentes de un compuesto establecen una relación “libre”, que no depende de la existencia de otra forma. Esto significa que las estructuras creadas por analogía con una combinación previa de elementos, no son en realidad formas compuestas. “No son palabras compuestas, sino derivadas, las que se obtienen mediante algún procedimiento derivativo aplicado a una palabra compuesta.” (RAE, 2009: 748). En los corpus se registraron ejemplos como *gas-masked*, *company-logoed* y *sideburned* que se desprenden de *gas mask*, *company logo* y *sideburn*, respectivamente.

5.5.2.1.7. Nominalizaciones del participio presente

En las muestras aleatorias tomadas del corpus electrónico se registraron innumerables casos de sustantivos compuestos que, por su forma, se ajustaban a las especificaciones usadas en la búsqueda. Entre los datos recolectados en el grupo de adjetivos compuestos terminados en *-ing*, encontramos ejemplos que son fácilmente identificables como sustantivos compuestos (*church-wedding*, *summer-*

reading, global-warming), como la combinación de un adjetivo y un sustantivo unidos por un guión (*early-morning, financial-planning, sustainable-living*).

Roeper y Siegel (1978) en su análisis de adjetivos compuestos de base verbal mencionan que el participio presente —la forma del verbo terminada en *-ing*— puede ser nominalizado “*-ing* is a simple nominalization of a verb” (1978: 215). Cuando esto ocurre se le llama **gerundio** (*gerund*). En inglés abundan, por lo tanto, sustantivos compuestos cuyos núcleos tienen esta forma tales como *back-packing* (“mochilear”), *insider-trading* (uso de información privilegiada), *power-walking* (caminata rápida) y *voice-acting* (doblaje), etc.

5.5.2.1.8. Compuestos nominales vs adjetivos compuestos

Los casos más difíciles de depurar son aquellos que involucran una doble interpretación. Esto es lo que sucede con términos como *witch-hunting*, *champagne-making* y *wife-beating*, los que pueden funcionar como los sustantivos “caza de brujas”, “fabricación de champaña” y “violencia contra la mujer”, respectivamente. Sin embargo, en otros contextos también pueden funcionar como adjetivos compuestos con participio. A continuación se ejemplifican ambos usos con ejemplos tomados del corpus original.

Compuestos nominales

[...] *The **witch-hunting** of the Medieval Age ...*

[...] *The art of **champagne-making** might be an old tradition ...*

[...] *Canada, regards **wife-beating** as illegal ...*

Adjetivos compuestos

[...] *And now, **witch-hunting** assassins, necromancers, half-demons ...* (asesinos que cazan brujas)

[...] *The letter is from our **champagne-making** friends* (de nuestros amigos que fabrican champaña)

[...] *a drunken, **wife-beating** cowboy* (un vaquero borracho que golpea a su esposa)

En la revisión manual se identificaron varios casos de compuestos con doble lectura por lo que se tuvo que verificar su potencial funcionamiento como adjetivos en los contextos originales de ocurrencia en el corpus electrónico. Sin embargo, la cantidad de datos hace imposible una revisión tan cuidadosa para cada compuesto de naturaleza dual.

Los criterios anteriormente descritos, nos ayudaron a eliminar cierto número adicional de formas que no podían considerarse compuestos. Sin embargo, es justo afirmar que en muestras de este tamaño es imposible asegurar la total detección y eliminación de elementos no deseados, sobre todo si el análisis está a cargo de una sola persona.

Capítulo 6. Resultados por patrones morfológicos individuales

En este capítulo se entregan los resultados obtenidos en las dos muestras tomadas del corpus *enTenTen12* (adjetivos compuestos terminados en *-ed* y en *-ing*). También se detallan algunas de las dificultades encontradas al usar la herramienta *Sketch Engine*.

6.1. Corpus *enTenTen12*: Adjetivos compuestos terminados en *-ed*

6.1.1. Sustantivo como primer elemento

El patrón morfológico que incorpora un sustantivo como primer constituyente resultó ser el más numeroso. Además, al compilar el corpus en *Sketch Engine*, el etiquetador sintáctico de esta plataforma nos permitió realizar consultas diferentes según el tipo de sustantivo, las que se describen a continuación.

La búsqueda [tag = "NN"] "-" [tag = "V.*"], en la que NN es un sustantivo común, produjo 2.364 resultados. Los sustantivos identificados incluyen acrónimos (*VA-guaranteed*, *SBA-designated*), acortamientos (*ad-based*, *croco-embossed*), hipocorísticos (*indie-minded*, *veggie-based*), apodos cariñosos (*mom-assisted*, *dad-related*, *kid-targeted*), y letras (*U-shaped*, *X-rayed*).

Sin embargo, la exactitud del etiquetador no es óptima puesto que se detectaron 79 errores de clasificación. La mayoría de ellos involucran palabras que tienen más de una función gramatical como *half*, la que fue etiquetada como sustantivo aunque al interior de un adjetivo compuesto solo puede operar como cuantificador (*half-hearted*) o adverbio (*half-formed*, *half-closed*). De igual manera, los vocablos relativos a formas son adjetivos y no sustantivos en compuestos como *round-necked* (de cuello redondo), *square-handled* (de manillas cuadradas) y *oblong-faced*. Por ende, el total final de esta consulta en particular es de 2.285 compuestos. Este resultado no implica que no haya un número indeterminado de compuestos no identificados por el etiquetador, ya que, debido a

la gran cantidad de datos contenidos en este patrón morfológico, no se pudo realizar una revisión manual completa de la muestra original que corroborara las cifras obtenidas.

Adicionalmente, *Sketch Engine* nos permite buscar sustantivos plurales (NNS) por lo que realizamos la búsqueda [tag = "NNS"] "-" [tag = "V.*"] la que arrojó 39 resultados entre los que se cuentan formas plurales regulares (*numbers-based*, *sales-related*) e irregulares (*alumni-sponsored*, *cattle-related*). Aun cuando se registraron solo dos errores de identificación (*SMS-enabled*, *comb-toothed*), el análisis manual contabilizó 23 compuestos más que contienen un elemento plural y que no fueron identificados por el etiquetador sintáctico, con los que el total se eleva a 60 ejemplos:

| | | |
|----------------------------|-----------------------|---------------------------|
| <i>claims-related</i> | <i>points-tested</i> | <i>seeds-oriented</i> |
| <i>Morrison-owned</i> | <i>points-based</i> | <i>sports-themed</i> |
| <i>gods-damned</i> | <i>needs-based</i> | <i>sports-based</i> |
| <i>gods-related</i> | <i>frames-based</i> | <i>motorsports-themed</i> |
| <i>singles-oriented</i> | <i>means-tested</i> | <i>watersports-based</i> |
| <i>results-oriented</i> | <i>arts-based</i> | <i>projects-based</i> |
| <i>results-based</i> | <i>arts-focused</i> | <i>bases-loaded</i> |
| <i>sweepstakes-related</i> | <i>people-focused</i> | |

En este patrón, otro grupo de sustantivos que está etiquetado aparte son los nombres propios, tanto en forma singular (NP) como plural (NPS). La consulta [tag = "NP"] "-" [tag = "V.*"] resultó ser bastante productiva (676 ejemplos) ya que en esta categoría se incluyen nombres de personas (*Oppenheimer-backed*, *Scorsese-inspired*), personajes literarios (*Dracula-themed*), películas (*Twilight-themed*), países (*Egypt-negotiated*, *Israel-inspired*), ciudades (*Tokyo-based*, *Washington-sponsored*), compañías (*FOX-affiliated*, *Pepsi-sponsored*) y productos (*Lumia-related*, *TiVo-dominated*) así como acrónimos (*Ni-depleted*, *IRS-assessed*), nombres de grupos o asociaciones (*Beatles-inspired*, *Hezbollah-controlled*) y festividades (*Thanksgiving-sized*).

No obstante, dado que solo hubo un sustantivo con minúscula identificado como nombre propio (*chardonnay*), la búsqueda parece guiarse principalmente por la presencia de mayúscula inicial. Sin embargo, en el compuesto *chardonnay-fuelled* la ausencia de mayúscula en el sustantivo parece recalcar su carácter genérico, lo que se ve corroborado por un compuesto muy similar (*merlot-dominated*) en el que *merlot* fue clasificado como sustantivo común. Esta consulta presenta 67 errores, por lo que este criterio no parece muy confiable. De estos errores, cuatro son resultado de una mala identificación dado que son adjetivos como *Brown-eyed* y *Blue-headed* y el resto son sustantivos comunes como “vinagre” (*Vinegar-based*), “jirafa” (*Giraffe-inspired*) o “tribunal” (*Court-appointed*). Asimismo, compuestos en los que sí aparecen nombres propios de ciudades (*Seattle-bred*, *Oakland-based*), países (*Finland-based*), asociaciones (*Hamas-appointed*) y de personajes históricos (*Edison-licensed*, *Picasso-sized*, *Elgar-themed*, *Bach-inspired*) no fueron identificados en esta categoría sino que en el grupo de sustantivos comunes. Por su parte, la búsqueda de nombres propios plurales (NPS) solo arrojó tres casos, todos incorrectamente clasificados.

Finalmente, aunque *Sketch Engine* también ofrece la posibilidad de identificar palabras de origen extranjero (*foreign words* = FW), la consulta [tag = "FW"] "-" [tag = "V.*"] solo arrojó tres resultados de los cuales ninguno es un término foráneo (*me-shaped*, *Sweden-based*, *needle-nosed*). Sin embargo, al revisar la muestra manualmente, se identificaron algunos compuestos que incorporan préstamos lingüísticos como *matzo-flavored*, *prosciutto-wrapped* o *miso-glazed*, los que fueron pasados por alto al momento del etiquetado automático.

La revisión cuidadosa de cada una de las búsquedas individuales realizadas en *Sketch Engine* implica no solo la detección de errores de etiquetado sino que también la reclasificación de cada uno de ellos en el patrón morfológico correcto. Por este motivo, al contabilizar los totales de este grupo deben agregarse los compuestos que tienen un sustantivo como constituyente inicial pero que al ser erróneamente etiquetados aparecen en otras búsquedas. Por ejemplo, 13 de ellos figuran entre los compuestos con adjetivo como primer elemento, tres como palabras extranjeras, dos como preposiciones, uno como determinante, otros tres como nombres propios plurales y dos como

pronombres personales, con un total de 24 ejemplos. Como se aprecia en la tabla de abajo, también de adicionaron los 23 compuestos con una forma plural que no fueron detectados automáticamente.

Tabla 10. Resultados de las consultas individuales sobre compuestos con sustantivo

| | Resultado inicial | Errores | Resultado final |
|--|-------------------|---------|-----------------|
| Sustantivo común como primer elemento (NN) | 2.364 | 79 | 2.285 |
| Sustantivo plural como primer elemento (NNS) | 39 | 2 | 37 |
| Nombre propio como primer elemento (NP) | 676 | 67 | 609 |
| Otros compuestos con sustantivo plurales | | | 23 |
| Compuestos agregados de otras búsquedas | | | 24 |
| | | | 2.978 |

Esto resulta en un proceso lento y laborioso, probablemente no exento de equivocaciones dado el volumen de datos manejados. Desafortunadamente, aun después de realizar todas las búsquedas posibles, de revisar acuciosamente los resultados y de reclasificarlos, no es posible identificar de manera cabal los patrones morfológicos contenidos en la muestra. Los etiquetadores sintácticos no están diseñados para distinguir entre una forma verbal terminada en *-ed* (*controlled, based, etc.*) y un sustantivo que agrega *-ed* (*-eyed, hearted, etc.*). Esto significa que el segundo grupo más grande de adjetivos compuestos en esta categoría no puede ser identificado automáticamente. Por ende, la correcta identificación y clasificación de los distintos patrones morfémicos de base nominal es total responsabilidad de la analista.

6.1.2. Adjetivo como primer elemento

La búsqueda [tag = "JJ"] "-" [tag = "V.*"] arrojó 652 resultados. Sin embargo, también registró una gran cantidad de desaciertos (142). Esto se explica porque, al igual como sucedió al analizar el

corpus de prueba, muchas palabras funcionan tanto como adjetivos como adverbios (*hard, fast, ill, long, little, low, high, etc.*), registrando 112 adverbios etiquetados como adjetivos. Otro factor a considerar es el hecho de que el etiquetador sintáctico de *Sketch Engine* no contempla una etiqueta para cuantificadores, por lo tanto deben ser identificados y separados manualmente. Finalmente, debemos identificar los compuestos cuyo núcleo es de naturaleza nominal (*Noun + -ed*) y clasificarlos separadamente.

Tabla 11. Resultados de la consulta sobre compuestos con adjetivo

| | Resultado inicial | Errores | Resultado final |
|---|-------------------|---------|-----------------|
| Adjetivo como primer elemento (JJ) | 652 | 141 | 511 |
| Compuestos agregados de otras búsquedas | | | 47 |
| Compuestos no identificados | | | -- |
| | | | 558 |

Cabe mencionar que el hecho de que en la tabla no figuren compuestos no identificados por el etiquetador automático no significa que no los haya. Dado el volumen de datos en este patrón, no se pudo llevar a cabo una revisión manual.

6.1.3. Adverbio como primer elemento

Al realizar la consulta sobre adverbios (RB) en los compuestos, se obtuvieron 632 resultados, todos correctamente etiquetados. Entre los adverbios identificados están aquellos con sufijo (*-ly*) como *newly-appointed* (recién nombrado) y *highly-qualified* (altamente calificado) así como otros que no tienen un rasgo morfológico distintivo (*just-cleaned, well-trained, seldom-discussed, never-solved, etc.*). Sin embargo, no se encontró ningún adverbio cuya forma se parece a un adjetivo como *long* (*long-established*), *high* (*high-valued*) o *little* (*little-debated*) ya que fueron clasificados en la otra

categoría. Asimismo, se encontraron otros 27 adverbios etiquetados como sustantivos, a los que se agregaron 14 identificados manualmente.

Tabla 12. Resultados de la consulta sobre compuestos con adverbio

| | Resultado inicial | Errores | Resultado final |
|---|-------------------|---------|-----------------|
| Adverbio como primer elemento (RB) | 632 | -- | 632 |
| Compuestos agregados de otras búsquedas | | | 139 |
| Compuestos no identificados | | | 14 |
| | | | 785 |

6.1.4. Adjetivo comparativo como primer elemento

La identificación de compuestos que contienen un adjetivo comparativo (JJR) como elemento inicial se realizó mediante la consulta [tag = "JJR"] "-" [tag = "V.*"], la que resultó en 17 ejemplos. De estos, 14 son del base nominal como *larger-screened*, *browner-backed* y *tighter-fisted*. Los otros tres corresponden al patrón adverbio comparativo + participio pasado (*lower-ranked*, *lower-rated* y *lower-valued*).

Tabla 13. Resultados de la consulta sobre compuestos con adjetivo comparativo

| | Resultado inicial | Errores | Resultado final |
|---|-------------------|---------|-----------------|
| Adjetivo comparativo como primer elemento (JJR) | 17 | 3 | 14 |
| Compuestos agregados de otras búsquedas | | | 1 |
| Compuestos no identificados | | | 4 |
| | | | 19 |

6.1.5. Adverbio comparativo como primer elemento

Se encontraron 30 compuestos, todos ellos correctamente etiquetados, usando la consulta [tag = "RBR"] "-" [tag = "V.*"] la que apunta a la identificación de un adverbio comparativo (RBR) como elemento inicial (*later, earlier, better, etc.*). Sin embargo, los compuestos *lesser-used* y *lesser-discussed* fueron omitidos pero recuperados en la revisión manual.

Tabla 14. Resultados de la consulta sobre compuestos con adverbio comparativo

| | Resultado inicial | Errores | Resultado final |
|---|-------------------|---------|-----------------|
| Adverbio comparativo como primer elemento (RBR) | 30 | -- | 30 |
| Compuestos agregados de otras búsquedas | | | -- |
| Compuestos no identificados | | | 2 |
| | | | 32 |

6.1.6. Adjetivo superlativo como primer elemento

La búsqueda de compuestos con adjetivo superlativo como primer constituyente, [tag = "JJS"] "-" [tag = "V.*"], dio como resultado 62 ejemplos. Una vez más, el etiquetador sintáctico no pudo distinguir entre adjetivos y adverbios, registrando el índice más alto de errores de clasificación. De hecho, la totalidad de los compuestos que incorporan un adverbio superlativo fueron erróneamente incluidos en esta categoría.

No obstante lo anterior, el etiquetador sí pudo distinguir entre algunas formas adjetivas y verbales descartando aquellos compuestos cuyos núcleos son *educated* y *skilled*, puesto que, aun cuando muy similares en forma a un participio verbal, funcionan como adjetivos en los compuestos *most-skilled*, *least-skilled* y *least-educated* por lo que deben ser removidos del corpus. Sin embargo, esto implica que también quedaron fuera aquellos adjetivos compuestos en los que *educated* funciona como participio pasado del verbo *to educate* tales como *Harvard-educated*, *Australian-educated* y *grant-*

educated, lo que vuelve a poner de relieve la importancia de las revisiones manuales en la correcta identificación de este tipo de estructuras compuestas.

Tabla 15. Resultados de la consulta sobre compuestos con adjetivo superlativo

| | Resultado inicial | Errores | Resultado final |
|---|-------------------|---------|-----------------|
| Adjetivo superlativo como primer elemento (JJS) | 62 | 55 | 7 |
| Compuestos agregados de otras búsquedas | | | -- |
| Compuestos no identificados | | | -- |
| | | | 7 |

6.1.7. Adverbio superlativo como primer elemento

La consulta [tag = "RBS"] "-" [tag = "V.*"] no arrojó ningún resultado puesto que, como ya vimos, los adjetivos compuestos que contienen un adverbio superlativo fueron reconocidos en la búsqueda de adjetivos superlativos, registrando 55 casos. A estos se agregó un compuesto identificado en la consulta de nombres propios.

Tabla 16. Resultados de la consulta sobre compuestos con adverbio superlativo

| | Resultado inicial | Errores | Resultado final |
|---|-------------------|---------|-----------------|
| Adverbio superlativo como primer elemento (RBS) | -- | -- | -- |
| Compuestos agregados de otras búsquedas | | | 56 |
| Compuestos no identificados | | | -- |
| | | | 56 |

6.1.8. Número como primer elemento

En el anotador sintáctico incluido en Sketch Engine, la etiqueta CD sirve para identificar números cardinales y ordinales. La búsqueda [tag = "CD"] "-" [tag = "V.*"] obtuvo siete resultados. En seis de ellos el número es cardinal y actúa como especificador de un núcleo de base nominal como *ten-toed* (con diez dedos en los pies) y *eight-legged* (de ocho patas). El etiquetador no reconoció ningún número ordinal, de los que se registraron 13 ocurrencias. (*seventh-placed*, *fourth-rated*, etc.).

Tabla 17. Resultados de la consulta sobre compuestos con número

| | Resultado inicial | Errores | Resultado final |
|---|-------------------|---------|-----------------|
| Número como primer elemento (CD) | 7 | -- | 7 |
| Compuestos agregados de otras búsquedas | | | 7 |
| Compuestos no identificados | | | 13 |
| | | | 27 |

Una vez obtenidos estos resultados, se procedió a una exhaustiva revisión manual de la muestra para la identificación de los compuestos de base nominal, los que totalizaron 785 ejemplos. Los compuestos cuyo núcleo es verbal sumaron 3.683. Al sumar ambas cifras se obtiene un total de 4.468. La muestra original consta de 4.808 compuestos por lo tanto hay 340 unidades sin clasificar (7,1 %). Esto probablemente se debe a que el etiquetador sintáctico no pudo identificarlas y, por lo tanto, no fueron incluidas en ninguna de las búsquedas.

6.2. Adjetivos compuestos con participios irregulares

Esta muestra, al ser más pequeña, resultó más fácil de procesar automáticamente y de verificar manualmente. Esto incidió en la obtención de un porcentaje de precisión mayor. La diferencia de 28 unidades compuestas no reconocidas en el etiquetador sintáctico representa 2,8 % de error.

Tabla 18. Patrones morfológicos en compuestos irregulares

| | | |
|---|--|------------|
| 1 | Sustantivo + participio pasado | 706 |
| 2 | Adjetivo + participio pasado | 41 |
| 3 | Adverbio + participio pasado | 207 |
| 4 | Adverbio Comparativo + participio pasado | 6 |
| 5 | Adverbio Superlativo + participio pasado | 11 |
| 6 | Número + participio pasado | 1 |
| | Total | 972 |

6.3. Corpus *enTenTen12*: Adjetivos compuestos terminados en *-ing*

Después de ser revisada y depurada varias veces, la muestra de adjetivos compuestos terminados en *-ing* extraída del corpus *enTenTen12* se redujo a 4.279 ejemplos. A continuación se detallan los resultados encontrados usando el etiquetador sintáctico de *Sketch Engine*:

6.3.1. Sustantivo como primer elemento

En esta muestra también se realizaron consultas separadas utilizando las distintas etiquetas incluidas en el *tagset* para la identificación de sustantivos. Por ejemplo, la búsqueda de adjetivos compuestos con sustantivo común (NN) registró 2.615 resultados, 23 de ellos incorrectamente etiquetados. Aun cuando el número de errores es bajo, debido al volumen de la muestra original esta cifra no pudo ser verificada mediante revisión manual para comprobar si quedaron compuestos de este tipo sin identificar

Por su parte, la consulta para sustantivos plurales (NNS) arrojó 56 ejemplos los que sí fueron revisados manualmente, constatándose la ocurrencia de cinco errores. Las formas plurales

identificadas son tanto regulares (*rights-respecting*, *tasks-providing*) como irregulares (*teeth-gritting*, *people-carrying*), de origen clásico (*bacteria-forming*, *algae-eating*, *media-buying*) incluso un compuesto cuyo primer elemento es semánticamente un sustantivo plural (*roach-crawling*) aunque su forma es singular. Sin embargo, en la revisión se encontraron otros 11 compuestos que contienen un plural regular (*decades-leaping*, *ends-justifying*, etc.).

En cuanto a los nombres propios (NP), se registraron 99 resultados en su forma singular, de los cuales 28 corresponden a errores de etiquetado, y uno solo en forma plural. En esta categoría encontramos compuestos como *Grammy-winning*, *Apple-seeking* y *Libya-bombing*. En la consulta por palabras foráneas (FW) hubo cuatro resultados, ninguno de ellos correctamente identificado. No obstante, al revisar manualmente la muestra, se encontraron compuestos como *fatwa-funding*, *latte-sipping* y *dia-paying* cuyo elemento inicial es de origen extranjero.

Tabla 19. Resultados de las consultas individuales sobre compuestos con sustantivo

| | Resultado inicial | Errores | Resultado final |
|--|-------------------|---------|-----------------|
| Sustantivo común como primer elemento (NN) | 2.615 | 23 | 2.592 |
| Sustantivo plural como primer elemento (NNS) | 56 | 5 | 51 |
| Nombre propio como primer elemento (NP) | 99 | 28 | 71 |
| Otros compuestos con sustantivo plurales | | | 11 |
| Compuestos agregados de otras búsquedas | | | 164 |
| Compuestos no identificados | | | 14 |
| | | | 2.903 |

6.3.2. Adjetivo como primer elemento

La identificación de adjetivos en los compuestos terminados en *-ing* resultó tener el más alto número de errores: de los 1.088 resultados, 713 son desaciertos. Entre los errores, el más común es la

confusión entre adjetivos y adverbios (254), especialmente en compuestos que involucran palabras como *ill*, *hard*, *fast*, *slow*, *free*, *high* y *low*. Además de lo anterior, 184 sustantivos fueron incorrectamente identificados como adjetivos. A modo de ilustración, *fat* aparece etiquetado como adjetivo (gordo) aunque funciona como sustantivo (grasa) en los 15 compuestos en los que aparece: *fat-reducing*, *fat-consuming*, *fat-eliminating*, etc.

Tabla 20. Resultados de la consulta sobre compuestos con adjetivo

| | Resultado inicial | Errores | Resultado final |
|---|-------------------|---------|-----------------|
| Adjetivo como primer elemento (JJ) | 1.088 | 713 | 375 |
| Compuestos agregados de otras búsquedas | | | 7 |
| Compuestos no identificados | | | -- |
| | | | 382 |

6.3.3. Adjetivo comparativo como primer elemento

Al consultar por adjetivos comparativos (JJR) se encontraron 11 ejemplos. Como podía preverse, los resultados mezclan adjetivos y adverbios comparativos, registrando solo cuatro aciertos. Un caso interesante de explicar es el compuesto *blinder-wearing* en el que *blinder* funciona como sustantivo (anteojeras) pero que fue erróneamente interpretado como la forma comparativa de *blind* (ciego).

Tabla 21. Resultados de la consulta sobre compuestos con adjetivo comparativo

| | Resultado inicial | Errores | Resultado final |
|---|-------------------|---------|-----------------|
| Adjetivo comparativo como primer elemento (JJR) | 11 | 7 | 4 |
| Compuestos agregados de otras búsquedas | | | 3 |
| Compuestos no identificados | | | 1 |
| | | | 8 |

6.3.4. Adverbio como primer elemento

Esta consulta produjo 245 resultados, dos de los cuales eran adjetivos (*even-looking, pretty-looking*) mientras que otros 25 errores involucran *all* como elemento inicial, el que funciona como sustantivo en este tipo de estructuras (*all-knowing, all-creating*) por lo que debieron ser reclasificados. Otros 14 ejemplos fueron encontrados al revisar el corpus manualmente.

Tabla 22. Resultados de la consulta sobre compuestos con adverbio

| | Resultado inicial | Errores | Resultado final |
|---|-------------------|---------|-----------------|
| Adverbio como primer elemento (RB) | 390 | 25 | 365 |
| Compuestos agregados de otras búsquedas | | | 254 |
| Compuestos no identificados | | | 14 |
| | | | 633 |

6.3.5. Adverbio comparativo como primer elemento

13 ejemplos, 1 error. Otros seis compuestos fueron identificados en la consulta anterior. Además, los compuestos *lesser-ranking* y *lesser-paying* aparecieron en la consulta de adjetivos.

Tabla 23. Resultados de la consulta sobre compuestos con adverbio comparativo

| | Resultado inicial | Errores | Resultado final |
|---|-------------------|---------|-----------------|
| Adverbio comparativo como primer elemento (RBR) | 30 | 2 | 28 |
| Compuestos agregados de otras búsquedas | | | 8 |
| Compuestos no identificados | | | 20 |
| | | | 56 |

6.3.6. Adjetivo superlativo como primer elemento

La búsqueda de adjetivos superlativos (JJS) obtuvo 110 resultados, de los cuales solo 26 son correctos. Los otros 84 compuestos contienen un adverbio superlativo como constituyente inicial.

Tabla 24. Resultados de la consulta sobre compuestos con adjetivo superlativo

| | Resultado inicial | Errores | Resultado final |
|---|-------------------|---------|-----------------|
| Adjetivo superlativo como primer elemento (JJS) | 110 | 84 | 26 |
| Compuestos agregados de otras búsquedas | | | 1 |
| Compuestos no identificados | | | -- |
| | | | 27 |

6.3.7. Adverbio superlativo como primer elemento

Al igual como ocurrió con la muestra de adjetivos compuestos terminados en *-ed*, la consulta sobre adverbios superlativos (RBS) no arrojó ningún resultado, puesto que fueron clasificados como adjetivos superlativos, registrando 84.

Tabla 25. Resultados de la consulta sobre compuestos con adverbio superlativo

| | Resultado inicial | Errores | Resultado final |
|---|-------------------|---------|-----------------|
| Adverbio superlativo como primer elemento (RBS) | -- | -- | -- |
| Compuestos agregados de otras búsquedas | | | 84 |
| Compuestos no identificados | | | -- |
| | | | 84 |

6.3.8. Número como primer elemento

Aun cuando en *Sketch Engine* podemos encontrar la etiqueta CD para identificar tanto números cardinales como ordinales, estos últimos aparecen identificados como adjetivos. Por su parte, el único número cardinal identificado en la muestra aparece en el compuesto *3d-looking* (parece de tres dimensiones) pero que en realidad es un acrónimo de *three-dimensional*.

Tabla 26. Resultados de la consulta sobre compuestos con número

| | Resultado inicial | Errores | Resultado final |
|---|-------------------|---------|-----------------|
| Número como primer elemento (CD) | -- | -- | -- |
| Compuestos agregados de otras búsquedas | | | 5 |
| Compuestos no identificados | | | 8 |
| | | | 13 |

La identificación y clasificación del primer constituyente del compuesto arrojó un total de 4.106 ejemplos. Teniendo en cuenta que la muestra total contiene 4.279 compuestos, la diferencia es de 173, lo que representa un 4 % de falibilidad.

6.4. Resultados obtenidos en los diferentes corpus

A continuación se presentan las tablas que detallan los resultados obtenidos en cada uno de los corpus analizados. Los datos están desglosados por patrón morfológico y, además de cuantificar los tipos (*type* = cada unidad diferente), se entregan los porcentajes de ocurrencia respectivos. La información también aparece desplegada en los gráficos correspondientes.

6.4.1. Corpus de prueba: Compuestos terminados en *-ed*

Este es el grupo de compuestos más grande (2.926), representando 73,5 % del total de datos recolectados en el corpus de prueba (3.983). A continuación se desglosan los resultados.

6.4.1.1. Compuestos de base nominal

Los patrones morfológicos de los adjetivos compuestos terminados en *-ed* se presentan separadamente. Primero, los de base nominal (*N-ed*) y luego los de núcleo verbal.

Tabla 27. Corpus de prueba: Resultados compuestos de base nominal

| Patrones Morfológicos | | <i>Types</i> | Porcentaje |
|-----------------------|--|--------------|---------------|
| 1 | Adjetivo + Sustantivo(-ed) | 860 | 29,4 % |
| 2 | Sustantivo + Sustantivo(-ed) | 422 | 14,4 % |
| 3 | Adverbio + Sustantivo(-ed) | 11 | 0,4 % |
| 5 | Adjetivo Comparativo + Sustantivo(-ed) | 8 | 0,3 % |
| 6 | Adjetivo Superlativo + Sustantivo(-ed) | 4 | 0,1 % |
| 7 | Cuantificador + Sustantivo(-ed) | 13 | 0,4 % |
| 8 | Número + Sustantivo(-ed) | 61 | 2,1 % |
| | Total | 1.379 | 47,1 % |

En esta categoría, los datos alcanzan los 2.926 adjetivos compuestos: base nominal = 1.379 y base verbal = 1.547. Por lo tanto, los primeros representan 47,1 % del total.

Núcleo nominal con sufijo -ed

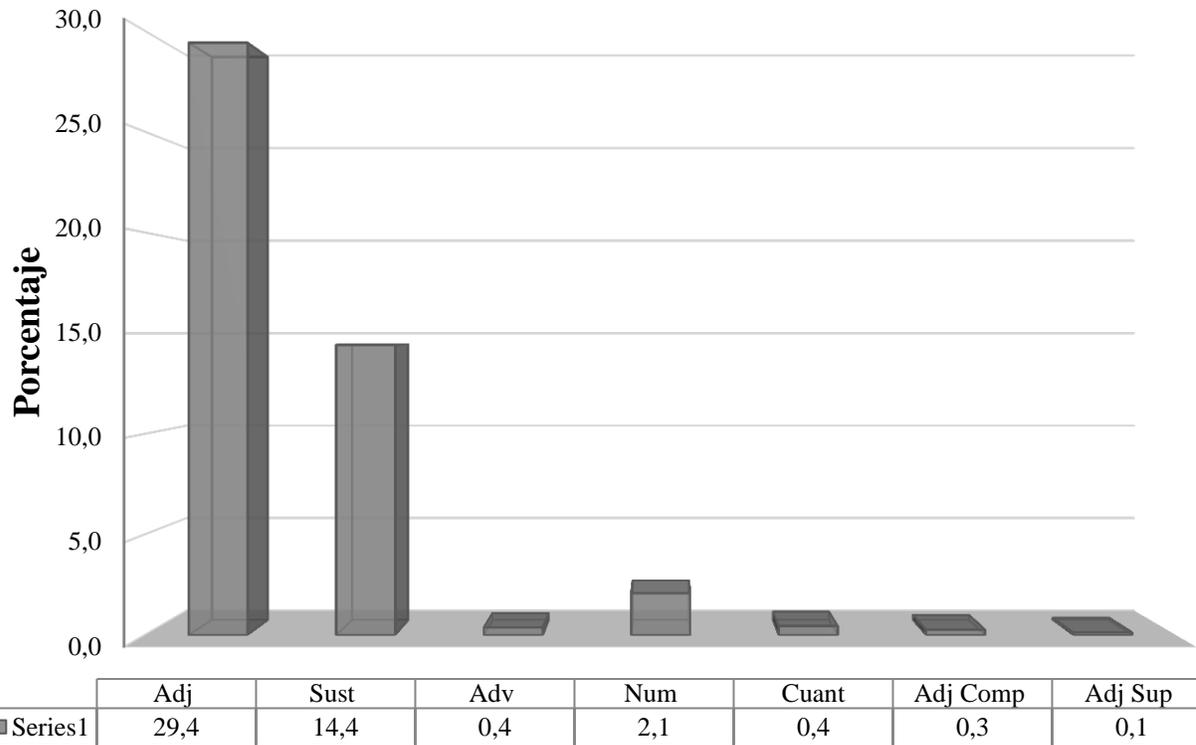


Figura 11. Patrones morfológicos de compuestos de núcleo nominal

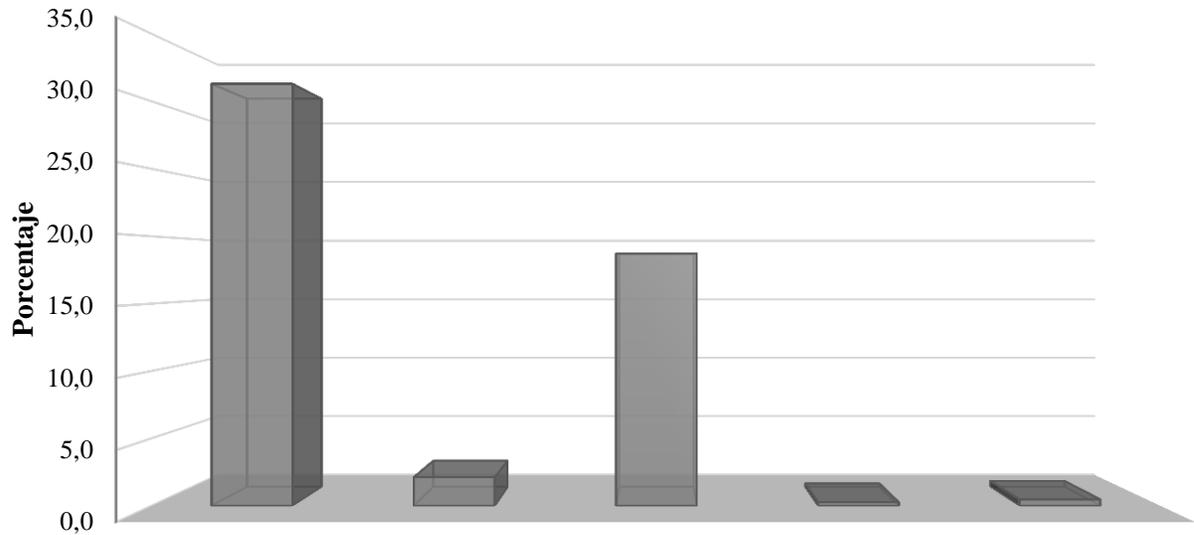
6.4.1.2. Compuestos de base verbal

Tabla 28. Corpus de prueba: Resultados compuestos de base verbal

| Patrones Morfológicos | | <i>Types</i> | Porcentaje |
|-----------------------|--|--------------|---------------|
| 1 | Sustantivo + participio pasado | 915 | 31,3 % |
| 2 | Adjetivo + participio pasado | 62 | 2,1 % |
| 3 | Adverbio + participio pasado | 548 | 18,7 % |
| 4 | Adverbio Comparativo + participio pasado | 8 | 0,3 % |
| 5 | Adverbio Superlativo + participio pasado | 14 | 0,5 % |
| 6 | Número + participio pasado | -- | -- |
| | Total | 1.547 | 52,9 % |

Los compuestos que contienen un núcleo verbal suman 1.547 de un total de 2.926, representando 52,9 % de la categoría.

Núcleo verbal con sufijo -ed



| | Sust | Adj | Adv | Adv Comp | Adv Sup |
|---------|------|-----|------|----------|---------|
| Series1 | 31,3 | 2,1 | 18,7 | 0,3 | 0,5 |

Figura 12. Corpus de prueba: Patrones morfológicos de compuestos de núcleo verbal

6.4.2. Corpus de prueba: Compuestos terminados en *-ing*

Tabla 29. Corpus de prueba: Resultados compuestos terminados en *-ing*

| Patrones Morfológicos | | <i>Types</i> | Porcentaje |
|-----------------------|--|--------------|--------------|
| 1 | Adjetivo + participio presente | 208 | 19,7 % |
| 2 | Sustantivo + participio presente | 643 | 60,8 % |
| 3 | Adverbio + participio presente | 188 | 17,8 % |
| 4 | Adverbio Comparativo + participio presente | 1 | 0,1 % |
| 5 | Adverbio Superlativo + participio presente | 8 | 0,8 % |
| 6 | Adjetivo Superlativo + participio presente | 4 | 0,4 % |
| 7 | Adjetivo Comparativo + participio presente | 4 | 0,4 % |
| 8 | Número + participio presente | 1 | 0,1 % |
| Total | | 1.057 | 100 % |

Compuestos terminados en -ing

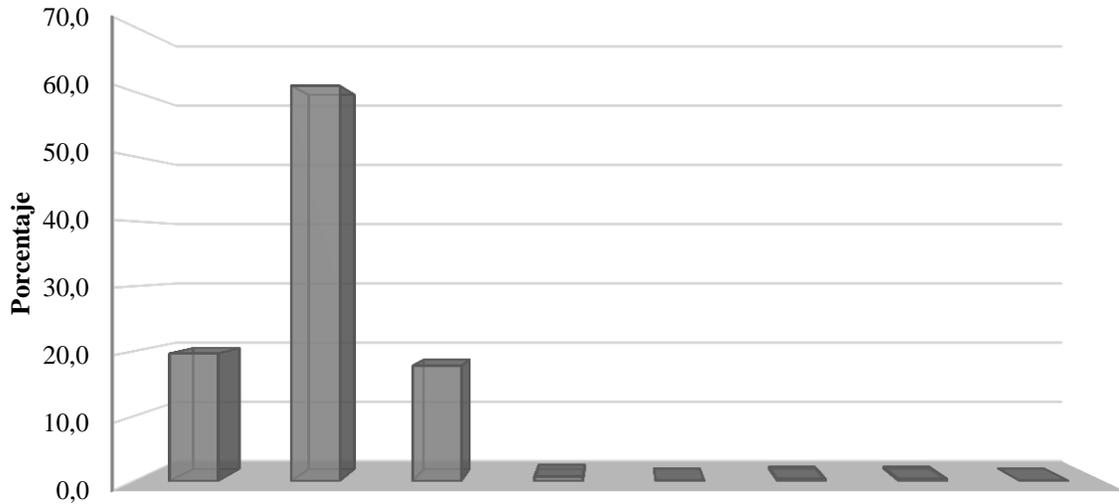


Figura 13. Corpus de prueba: Patrones morfológicos de compuestos con participio presente

6.4.3. Corpus *enTenTen12*: Compuestos terminados en *-ed*

6.4.3.1. Compuestos de base nominal

Tabla 30. Corpus *enTenTen12*: Resultados compuestos de base nominal

| Patrones Morfológicos | | <i>Types</i> | Porcentaje |
|-----------------------|--|--------------|---------------|
| 1 | Adjetivo + Sustantivo(-ed) | 405 | 9,1 % |
| 2 | Sustantivo + Sustantivo(-ed) | 318 | 7,1 % |
| 3 | Adverbio + Sustantivo(-ed) | 14 | 0,3 % |
| 5 | Adjetivo Comparativo + Sustantivo(-ed) | 19 | 0,4 % |
| 6 | Adjetivo Superlativo + Sustantivo(-ed) | 7 | 0,2 % |
| 7 | Cuantificador + Sustantivo(-ed) | 7 | 0,2 % |
| 8 | Número + Sustantivo(-ed) | 15 | 0,3 % |
| | Total | 785 | 17,6 % |

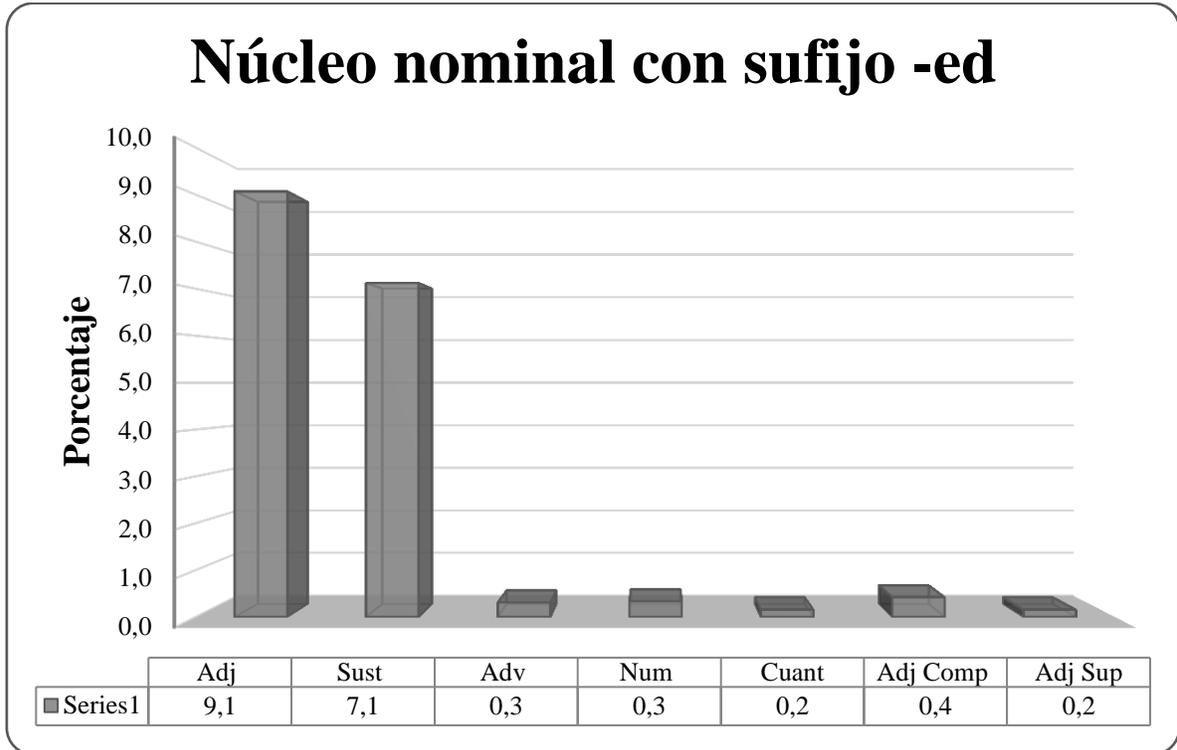


Figura 14. Corpus *enTenTen12*: Patrones morfológicos de compuestos con núcleo nominal

6.4.3.2. Compuestos de base verbal

Una vez procesados los datos de esta muestra, se determinó que 340 unidades quedaron sin clasificar ($4.808 - 4.468 = 340$), lo que representa un margen de error de 7,1 %.

Tabla 31. Corpus *enTenTen12*: Resultados compuestos de base verbal

| Patrones Morfológicos | | Types | Porcentaje |
|-----------------------|--|--------------|---------------|
| 1 | Sustantivo + participio pasado | 2.660 | 59,5 % |
| 2 | Adjetivo + participio pasado | 153 | 3,4 % |
| 3 | Adverbio + participio pasado | 771 | 17,3 % |
| 4 | Adverbio Comparativo + participio pasado | 32 | 0,7 % |
| 5 | Adverbio Superlativo + participio pasado | 56 | 1,3 % |
| 6 | Número + participio pasado | 11 | 0,2 % |
| Total | | 3.683 | 82,4 % |

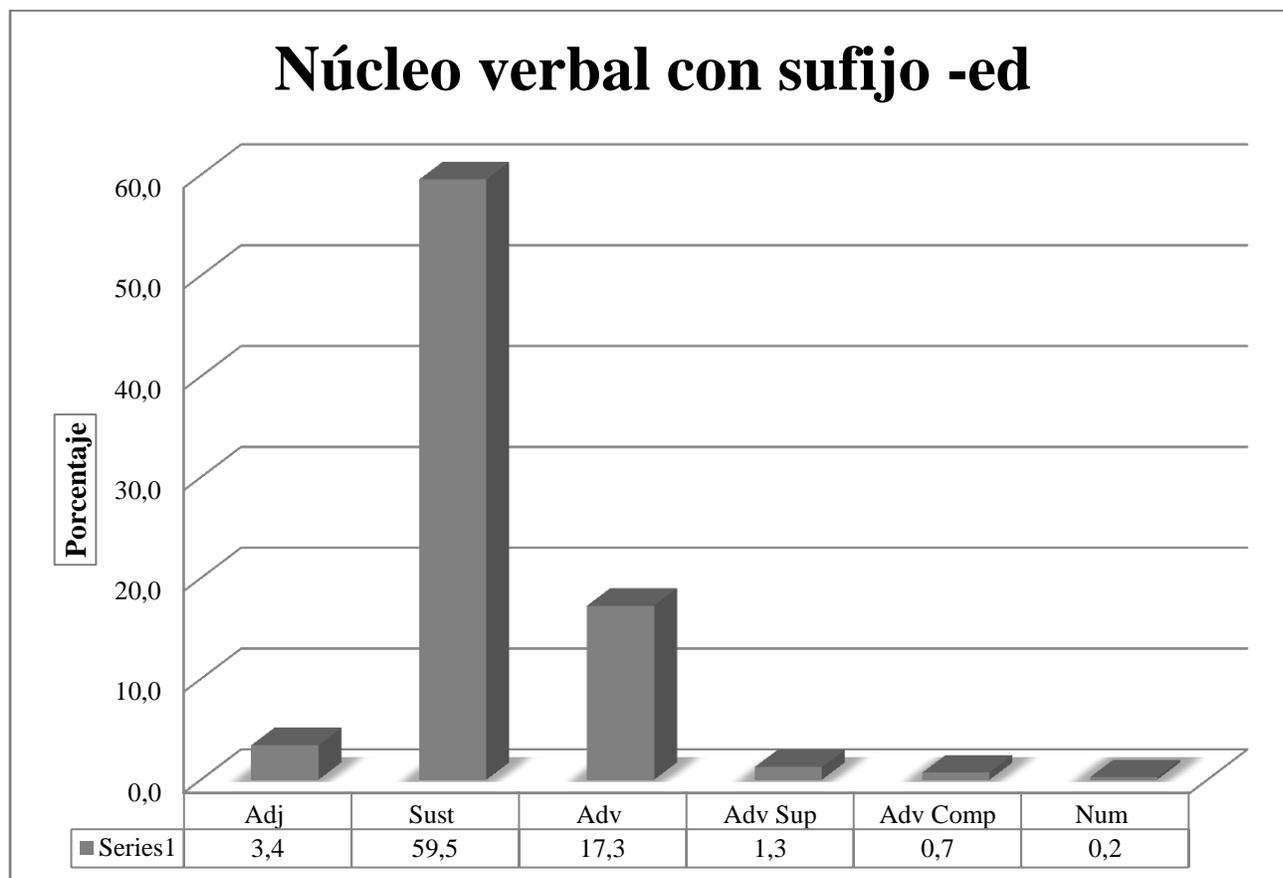


Figura 15. Corpus *enTenTen12*: Patrones morfológicos de compuestos con núcleo verbal

6.4.4. Corpus *enTenTen12*: Compuestos terminados en *-ing*

Tabla 32. Corpus *enTenTen12*: Resultados compuestos terminados en *-ing*

| Patrones Morfológicos | | <i>Types</i> | Porcentaje |
|-----------------------|--|--------------|--------------|
| 1 | Adjetivo + participio presente | 382 | 9,3 % |
| 2 | Sustantivo + participio presente | 2.903 | 70,7 % |
| 3 | Adverbio + participio presente | 633 | 15,4 % |
| 4 | Adverbio Superlativo + participio presente | 56 | 2 % |
| 5 | Adverbio Comparativo + participio presente | 84 | 1,4 % |
| 6 | Adjetivo Superlativo + participio presente | 27 | 0,7 % |
| 7 | Adjetivo Comparativo + participio presente | 8 | 0,2 % |
| 8 | Número + participio presente | 13 | 0,3 % |
| Total | | 1.057 | 100 % |

Al procesar este corpus, quedaron 173 unidades sin clasificar ($4.279 - 4.106 = 173$), lo que representa un 4 % de falibilidad.

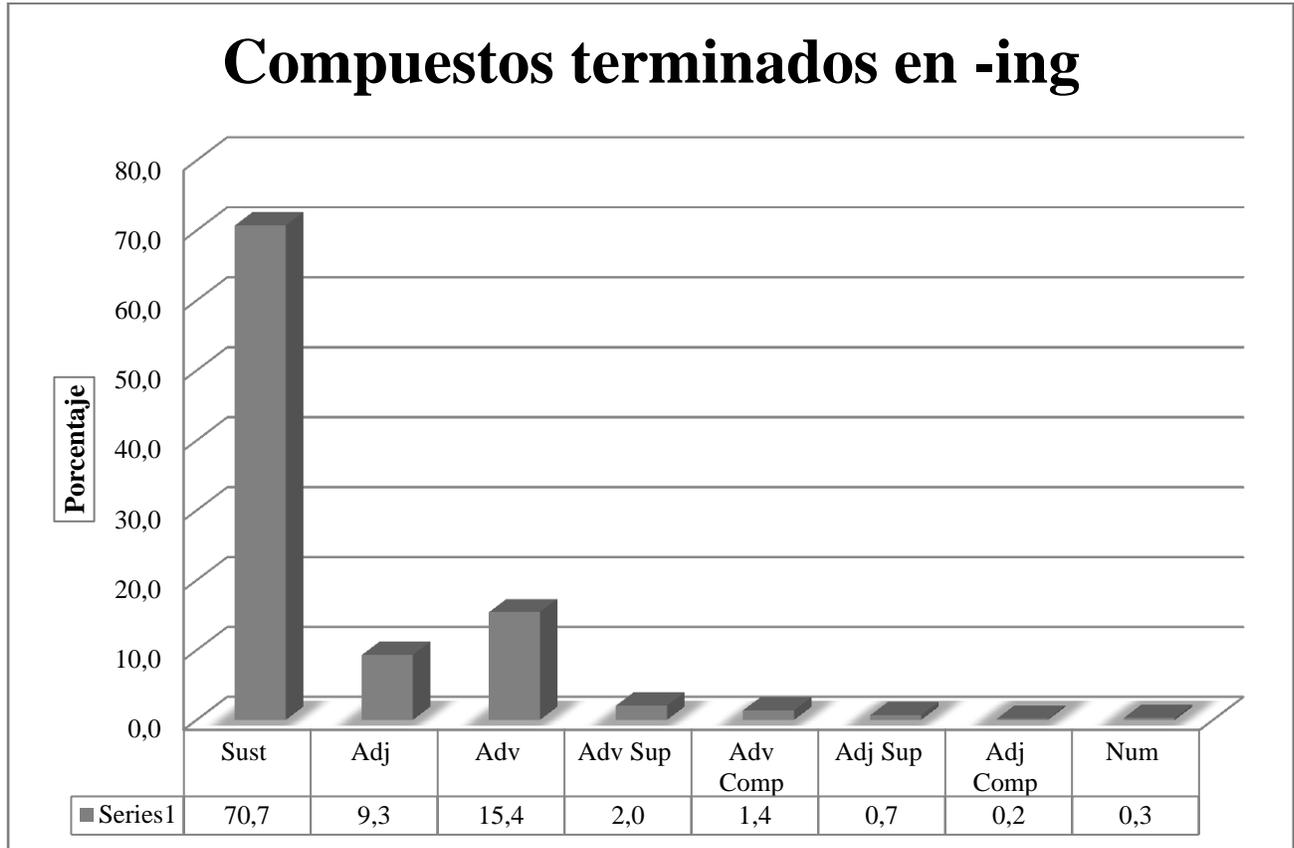


Figura 16. Corpus *enTenTen12*: Patrones morfológicos de compuestos con participio presente

6.4.5. Resultados comparados

6.4.5.1. Adjetivos compuestos de base nominal

La ocurrencia de adjetivos compuestos con núcleo nominal (*N-ed*) mostró grandes diferencias entre los dos corpus. Por una parte, representan casi 50% del corpus de prueba (CP) mientras que en la muestra del corpus electrónico (CE) no alcanzan el 20%. Las diferencias se concentran en algunos de los patrones morfológicos más numerosos, ya que en el CP los compuestos con adjetivos como primer elemento prácticamente triplican a los de su contraparte electrónica. En el caso de los sustantivos y cuantificadores, los duplican. Los compuestos con números también resultaron mucho

más frecuentes en el CP. Los otros patrones morfémicos obtuvieron resultados parecidos como se puede observar en la siguiente tabla:

Tabla 33. Resultados comparados compuestos de base nominal

| Patrones Morfológicos | | Corpus de Prueba | Muestra <i>EnTenTen12</i> |
|-----------------------|--|------------------|---------------------------|
| 1 | Adjetivo + Sustantivo(-ed) | 29,4 % | 9,1 % |
| 2 | Sustantivo + Sustantivo(-ed) | 14,4 % | 7,1 % |
| 3 | Adverbio + Sustantivo(-ed) | 0,4 % | 0,3 % |
| 4 | Adjetivo Comparativo + Sustantivo(-ed) | 0,3 % | 0,4 % |
| 5 | Adjetivo Superlativo + Sustantivo(-ed) | 0,1 % | 0,2 % |
| 6 | Cuantificador + Sustantivo(-ed) | 0,4 % | 0,2 % |
| 7 | Número + Sustantivo(-ed) | 2,1 % | 0,3 % |
| | Total | 47,1 % | 17,6 % |

6.4.5.2. Adjetivos compuestos de base verbal

También se registraron diferencias marcadas en este grupo. Mientras que en el CP los compuestos con núcleo verbal constituyen casi 53 %, su presencia domina en el CE con 82,4 %. Esta vez, el patrón con sustantivos arroja la mayor diferencia, casi duplicando a los encontrados en el CP. Lo mismo ocurre con los adverbios comparativos y superlativos. El patrón Adverbio + participio pasado rompe esta tendencia al registrar un porcentaje levemente mayor de ocurrencias en el CP (1,4 %). Por otra parte, aunque no se registraron ejemplos del patrón con número como primer elemento en el CP, sí hubo un porcentaje menor en el corpus electrónico (0,2 %).

Tabla 34. Resultados comparados compuestos de base verbal

| Patrones Morfológicos | | Corpus de Prueba | Muestra <i>EnTenTen12</i> |
|-----------------------|--|------------------|---------------------------|
| 1 | Adjetivo + participio pasado | 2,1 % | 3,4 % |
| 2 | Sustantivo + participio pasado | 31,3 % | 59,5 % |
| 3 | Adverbio + participio pasado | 18,7 % | 17,3 % |
| 4 | Adverbio Comparativo + participio pasado | 0,3 % | 0,7 % |
| 5 | Adverbio Superlativo + participio pasado | 0,5 % | 1,3 % |
| 6 | Número + participio pasado | -- | 0,2 % |
| Total | | 52,9 % | 82,4 % |

6.4.5.2. Adjetivos compuestos terminados en -ing

Esta categoría de adjetivos compuestos arrojó resultados relativamente similares en dos de los patrones morfológicos más productivos: la diferencia entre los compuestos con sustantivos es de 10 %, mientras que en las unidades con adverbios como primer elemento se reduce a 2,4 %. La mayor disparidad numérica se concentra en los patrones con adjetivos y adverbios comparativos y superlativos.

Tabla 35. Resultados comparados compuestos terminados en *-ing*

| Patrones Morfológicos | | Corpus de Prueba | Muestra <i>EnTenTen12</i> |
|-----------------------|--|------------------|---------------------------|
| 1 | Adjetivo + participio presente | 19,7 % | 9,3 % |
| 2 | Sustantivo + participio presente | 60,8 % | 70,7 % |
| 3 | Adverbio + participio presente | 17,8 % | 15,4 % |
| 4 | Adverbio Comparativo + participio presente | 0,1 % | 1,4 % |
| 5 | Adverbio Superlativo + participio presente | 0,8 % | 2 % |
| 6 | Adjetivo Superlativo + participio presente | 0,4 % | 0,7 % |
| 7 | Adjetivo Comparativo + participio presente | 0,4 % | 0,2 % |
| 8 | Número + participio presente | 0,1 % | 0,3 % |
| Total | | 100 % | 100 % |

Un factor importante a considerar en el análisis de los resultados es el hecho de que en esta categoría los corpus no eran equivalentes en tamaño: 1.057 ejemplos en el corpus de prueba y 4.279 en la muestra extraída del *EnTenTen12*. En las conclusiones se analizan otros factores que pueden haber incidido en los resultados obtenidos y las implicancias para este tipo de investigaciones de corpus.

Capítulo 7. Propiedades morfosintácticas de los adjetivos compuestos con participio

7.1. Formas flexivas como primer elemento: Plurales internos

En general, las palabras compuestas en inglés no aceptan flexión plural interna puesto que el elemento modificador (a la izquierda) es una base lexémica. Bauer (2006: 489) describe la estructura de los compuestos de la siguiente manera: “In the default cases, compounds in English have the structure lexemic-base + lexemic-base (+ inflection). Specifically, this excludes inflections from positions which are compound-internal”.

La preferencia por formas singulares es evidente en un compuesto como *a two-man boat* en el que ni siquiera la presencia del número “dos” (*two*) permite la forma plural *men* (Bauer, 2006: 490). En el corpus de prueba se encontraron ejemplos como *man-hating* (que odia a los hombres) en el que el elemento inicial retiene la forma singular aunque no denota a un hombre en particular. Sin embargo, tanto Bauer (2006) como Spencer (1991) dan cuenta de la ocurrencia de sustantivos compuestos como *teethmarks* y *teeth-ridge* en los que se acepta una forma plural irregular como primer constituyente. Casos de otros sustantivos que contienen desinencia regular como *systems analyst*, *suggestions box*, *arms race* (carrera armamentista) y *sports car* también son posibles aunque son menos frecuentes. De hecho, en los corpus analizados *sports* aparece tal cual en algunos adjetivos compuestos como *sports-based*, *sports-themed*, *motorsports-themed* y *watersports-based*. Sin embargo, se registró el mismo número de ejemplos en los que esa forma plural se había estandarizado a singular: *sport-infused*, *sport-inspired*, *motorsport-honed* y *sport-touring*. Esta dualidad se cristaliza en la ocurrencia de las dos formas posibles en el caso de *sports-loving* y *sport-loving*.

Por su parte, Blevins (2006) concuerda en que los sustantivos plurales creados por alternancia vocálica (*tooth/teeth*; *foot/feet* o *mouse/mice*, *louse/lice*) o por la incorporación de *-en* (*child/children*; *ox/oxen*) ocurren más libremente al interior de estructuras compuestas (*mice-infested*, *feet-elevated*, *teeth-gritting*) que aquellos que tienen desinencia regular (*-s/-es*). Esto se debe a que estas formas irregulares son percibidas como neutras en cuanto a número, por lo que los

hablantes consideran que un compuesto como *lice-infested* (infestado de piojos) es perfectamente aceptable, mientras que un sustantivo plural regular en un adjetivo similar en forma y significado como **fleas-infested* (infestado de pulgas) es percibido como anómalo. Algo similar ocurre con las palabras de origen latino cuyos plurales mantienen la desinencia clásica, los que pueden ocurrir al interior de un adjetivo compuesto como *algae-powered* (que obtiene energía de las algas).

A modo de ejemplificación de la preponderancia de las formas singulares en el elemento inicial, en los corpus analizados se verificó la ocurrencia de 18 adjetivos en que *-infested* funciona como núcleo y, aun cuando este verbo semánticamente implica un gran número de organismos, en solo uno de ellos el sustantivo es un plural irregular (*lice-infested*). En los otros ejemplos el primer constituyente mantiene la forma singular aunque debe ser interpretado como plural, como queda de manifiesto en *shark-infested* (infestado de tiburones), *bug-infested* (infestado de bichos) y *gang-infested* (plagado de pandillas). Sin embargo, también se registraron adjetivos como *malware-infested* (infectado con un virus computacional) y *plague-infested* (infectado con la plaga o peste) cuyo primer elemento es un concepto singular. El compuesto *crime-infested* es ambiguo en lo que respecta a número puesto que *crime* puede ser interpretado como un sustantivo singular incontable (plagado de delincuencia o criminalidad) o bien como la forma singular de actividades criminales (plagado de delitos o crímenes).

En los corpus no se registraron ejemplos de formas plurales irregulares como *men* y *women* mientras que *children* registró solamente una: *children-inhabited (trees)*. De hecho, se registran varios compuestos que incluyen *man* y que claramente denotan no solo “hombres” en plural sino que “personas” en general, como *man-killing machinery* y *man-eating tiger*. Esto también ocurre con “mujeres” en *woman-hating men* (hombres que odian a las mujeres) en la que el compuesto se ciñe a la regla de no aceptar flexiones internas y mantiene la forma singular del primer elemento (*woman*). Por otra parte, el sustantivo plural *people* aparece en varios compuestos (*people-oriented*, *people-focused*, *people-initiated*, *people-sized*). El sustantivo plural *pants* (pantalones) es aceptado como parte de estas unidades compuestas (*pants-wearing*, *pants-busting*) porque esta es la forma

comúnmente usada en este tipo de estructuras compuestas aunque en otros tipos de combinaciones como *pantdress* y *pant legs* se prefiere el singular.

En inglés se denomina plural cero (*∅-plural*) a la ausencia de un marcador de número en algunos sustantivos como *sheep* (oveja/ovejas) y *deer* (ciervo/ciervos). En estos casos son otros elementos del contexto, generalmente el predicado, los que nos ayudan a determinar si estas formas son singular o plural: *the sheep was/were killed by a wolf*. Sin embargo, cuando ocurren al interior de un compuesto deben ser interpretadas como plurales, como lo demuestran *sheep-farming* (que cría ovejas) y *sheep-walking* (que lleva a pastar a las ovejas).

A pesar de que la regla general estipula que en una palabra compuesta el elemento modificador no acepta flexión plural puesto que es la forma base de un lexema (Bauer, 2006), en el caso de los adjetivos compuestos con participio, la presencia de formas plurales regulares también está documentada, especialmente en los patrones que tienen un nombre como elemento inicial. En lo que respecta al patrón sustantivo + participio presente (*-ing*), en el que el primero es un argumento del núcleo verbal, se registraron 45 ocurrencias:

| | |
|------------------------|---------------------------|
| <i>awards-giving</i> | <i>skills-building</i> |
| <i>ends-justifying</i> | <i>events-planning</i> |
| <i>secrets-keeping</i> | <i>rights-respecting</i> |
| <i>Arabs-including</i> | <i>sanctions-busting</i> |
| <i>goods-producing</i> | <i>nations-entangling</i> |

Como se puede observar en los ejemplos, el sustantivo actúa como complemento directo del verbo. En los corpus se encontraron solamente dos instancias de un sustantivo plural con función de adjunto, coincidentemente el mismo en ambos casos: *decades-functioning* y *decades-leaping*.

En cuanto a los compuestos del tipo Noun + past participle (*-ed*), cabe mencionar que en los corpus analizados, la ocurrencia de sustantivos plurales regulares (49 casos) se restringe a un pequeño grupo de verbos, principalmente *-related*, *-based*, *-focused*, y *-oriented*:

| | |
|-----------------------|-------------------------|
| <i>needs-based</i> | <i>strengths-based</i> |
| <i>sales-related</i> | <i>oceans-focused</i> |
| <i>points-tested</i> | <i>options-focused</i> |
| <i>earnings-based</i> | <i>results-oriented</i> |
| <i>numbers-based</i> | <i>singles-oriented</i> |

También se registró un ejemplo de un sustantivo compuesto plural como primer elemento de un compuesto en *nuclear weapons-contaminated (waters)*.

En teoría, los plurales irregulares no funcionan como núcleo en compuestos de base nominal (N-*ed*), es decir, aquellos en los que un sustantivo incorpora el sufijo *-ed* como es el caso de *foot (flat-footed, light-footed)* o *tooth (sharp-toothed, white-toothed)* cuyo segundo elemento se crea a partir de la forma singular y no de su plural. En consecuencia, esto vuelve inviables compuestos como **flat-feeted* o **white-teethed*. Lo mismo debiera esperarse de una forma como *leaf* (hoja) cuyo plural es *leaves*. No obstante, en los corpus se encontraron los siguientes ejemplos:

oak-leaved
fern-leaved
thick-leaved (tree)
broad-leaved (plant)
full-leaved (olive branches)
slender-leaved (mimosa grove)

Como se puede observar, en estos compuestos el sufijo *-ed* se incorpora al sustantivo plural puesto que la forma singular produciría *-leafed*. Este es el único ejemplo registrado en los corpus de una forma plural en posición nuclear. Como elemento compositivo, los diccionarios *Oxford* y *Merriam-Webster* en línea⁷ lo definen como una variante de *-leaved*. De hecho, en el corpus electrónico *enTenTen12*, se verificó la ocurrencia de un número similar de compuestos con estas formas, siendo *-leaved* un poco más frecuente (19 casos) que *-leafed* (16):

tiny-leafed (plants)

broad-leafed (varieties)

Slender-leafed Duck Orchid

7.2. Formas derivadas en los adjetivos compuestos

Se comprobó la presencia de una amplia gama de formas derivadas que incluyen sustantivos, adjetivos y adverbios. Estos pueden ser simples (una derivación) o más complejas con dos derivaciones o un prefijo y un sufijo. A continuación se listan algunos de los sufijos más característicos de los sustantivos en inglés con ejemplos tomados de los corpus:

Tabla 36. Sufijos derivativos nominales presentes en los compuestos

| | |
|---------------------------|--|
| <i>-ness</i> | <i>consciousness-based, darkness-loving</i> |
| <i>-ment</i> | <i>employment-focused, equipment-hauling</i> |
| <i>-ance/-ence</i> | <i>competence-related, performance-shaping</i> |
| <i>-tion</i> | <i>virtualization-related, connection-blasting</i> |
| <i>-ity</i> | <i>sustainability-linked, liquidity-injecting</i> |

⁷ <https://en.oxforddictionaries.com/definition/leafed>, <https://www.merriam-webster.com/dictionary/leafed>

| | |
|-------------------|---|
| -ship | <i>championship-contending, relationship-dominated</i> |
| -er/-or | <i>consumer-directed, foreigner-duping, accelerator-based</i> |
| -ism | <i>metabolism-improving, optimism-charged</i> |
| -ist | <i>imperialist-orchestrated, therapist-led, Marxist-leaning</i> |
| -age | <i>voltage-controlled, language-communicating</i> |
| -dom/-hood | <i>freedom-based, neighborhood-destroying</i> |

Tabla 37. Sufijos derivativos adjetivales presentes en los compuestos

| | |
|------------------|---|
| -ous | <i>vigorous-growing, dangerous-seeming, villainous-looking</i> |
| -ive | <i>progressive-minded, authoritative-seeming, massive-sized</i> |
| -y | <i>hairy-legged, icy-sounding, dreamy-eyed, shiny-faced</i> |
| -ful | <i>beautiful-looking, wonderful-sounding, masterful-looking</i> |
| -less | <i>flawless-looking, timeless-sounding, reckless-eyed</i> |
| -ic/-ical | <i>dramatic-looking, nautical-inspired, musical-sounding</i> |
| -ish | <i>babyish-looking, folkish-sounding, pinkish-hued</i> |

También se encontraron ejemplos de palabras que incorporan prefijos:

Tabla 38. Prefijos en los compuestos

| | |
|---------------|--|
| anti- | <i>antibody-based, antiviral-testing (company)</i> |
| dis- | <i>disability-led, disreputable-looking, time-discoloured, now-disappeared</i> |
| infra- | <i>infrared-controlled, infrastructure-based</i> |
| in-/im | <i>inexpressive-looking, independently-minded, immature-looking</i> |
| mid- | <i>midterm-related</i> |

| | |
|---------------------|--|
| <i>mis-</i> | <i>often-misrepresented</i> |
| <i>non</i> | <i>nondisease-causing, nonprofit-making, nonowner-occupied</i> |
| <i>out-</i> | <i>outstandingly-prepared, well-outfitted, outside-facing</i> |
| <i>over-</i> | <i>oft-overflowing, once-overwhelming, modestly-overstated</i> |
| <i>pre-</i> | <i>prehistoric-looking, preschoolers-rearing</i> |
| <i>re-</i> | <i>recovery-based, reproductive-boosting, life-renewing, debt-rescheduling</i> |
| <i>semi-</i> | <i>semiautomatic-wielding</i> |
| <i>un-</i> | <i>unimportant-looking, unfortunately-sloped, ever-unfolding, unwell-advised</i> |

7.3. Elementos compositivos cultos

Los elementos compositivos cultos también están presentes en los compuestos, principalmente como elemento inicial aunque también pueden ser parte del núcleo:

Tabla 39. Elementos compositivos cultos presentes en los compuestos

| | |
|-----------------------|--|
| <i>adeno</i> | <i>adeno-associated</i> |
| <i>astro</i> | <i>astronomy-themed, astrologically-oriented</i> |
| <i>auto</i> | <i>auto-asphyxiation-loving, automobile-related, autograph-seeking</i> |
| <i>bio</i> | <i>biomass-degrading, biologically-affected</i> |
| <i>cardio</i> | <i>cardio-based</i> |
| <i>eco</i> | <i>eco-themed, eco-approved</i> |
| <i>electro</i> | <i>electromagnetically-loaded, electro-tinged</i> |
| <i>phore</i> | <i>fluorophore-tagged</i> |
| <i>geo</i> | <i>geo-referenced, geo-locating</i> |
| <i>hemi</i> | <i>hemicellulose-based</i> |
| <i>micro</i> | <i>microkernel-based, microphone-toting</i> |
| <i>multi</i> | <i>multimedia-powered</i> |

| | |
|---------------|---|
| neuro | <i>neuroscience-based, neurologically-based</i> |
| photo | <i>photocell-controlled, photographer-trailed, poorly-photoshopped</i> |
| psycho | <i>psychosomatically-activated</i> |
| techno | <i>technology-enhanced, technologically-skilled</i> |
| tele | <i>telecommand-related, television-viewing, internationally-televised</i> |

7.3.1. *Self*

Uno de los elementos compositivos más productivos en inglés es *self*, registrando 108 ocurrencias en el grupo de adjetivos compuestos con participio pasado (-*ed*) y 259 en aquellos con participio presente (-*ing*). En los numerosos compuestos formados con este elemento, hay que distinguir entre aquellos en los que *self-* es complemento directo del núcleo verbal (*self-protected, self-respecting, self-victimizing*) y aquellos en los que funciona como adjunto en el rol de agente (*self-learned, self-closing, self-evolving*). En algunos casos, especialmente en combinación con participio pasado, *self-* puede denotar lugar o dirección (*self-contained, self-directed, self-oriented*). Otros ejemplos son:

Tabla 40. *Self* en los adjetivos compuestos con participio

| <i>Self</i> + pasado participio | <i>Self</i> + participio presente |
|---------------------------------|-----------------------------------|
| <i>self-limited</i> | <i>self-hating</i> |
| <i>self-learned</i> | <i>self-polluting</i> |
| <i>self-propelled</i> | <i>self-consuming</i> |
| <i>self-arrogated</i> | <i>self-replicating</i> |
| <i>self-evaluated</i> | <i>self-vindicating</i> |
| <i>self-contained</i> | <i>self-incriminating</i> |

7.4. Procesos de formación de palabras en los adjetivos compuestos

7.4.1. Acuñaación

En cuanto a los otros procesos de formación de palabras, se constató la presencia de la mayoría de ellos en el primer constituyente de las formas en estudio. Por ejemplo, términos acuñados como *kevlar* (1965) o *nanite* (1990) aparecen en los compuestos *kevlar-reinforced* y *nanite-enhanced*.

En áreas científicas como medicina y farmacia comúnmente se acuñan nombres para diferentes sustancias y compuestos. Por ejemplo, *xylitol*⁸, nombre de un edulcorante natural, es una palabra creada a partir de la combinación de una raíz griega (*xylō* = madera) y un elemento compositivo culto (*-itol*) usado para denotar diferentes tipos de alcoholes derivados del azúcar (sorbitol, arabitol, eritritol, etc.). Este es un ejemplo de lo difuso de los límites entre los distintos procesos de formación de palabras, puesto que el término *xylitol*, si bien es una acuñaación, también es el resultado de un préstamo lingüístico y una derivación (incorporación de *-itol*). Todos estos procesos confluyen en la creación de un lexema independiente que, posteriormente, puede ser usado en la composición de un adjetivo como *xylitol-sweetened*.

7.4.2. Préstamos lingüísticos

En los corpus analizados también se registraron casos de préstamos lingüísticos de distinta procedencia como primer constituyente:

⁸ <http://www.xlear.com/what-is-xylitol/>

Tabla 41. Préstamos lingüísticos en los compuestos

| | |
|-----------|---|
| francés | <i>lingerie-clad, eau de cologne-shifting, musique concrète-inspired, chardonnay-fuelled, bistre-coloured, merlot-dominated</i> |
| italiano | <i>fascist-cropped, paparazzi-shielding sunglasses, villa-dotted (district), prosciutto-wrapped</i> |
| español | <i>piñata-bursting, salsa-loving, mantilla-decked (women), armadillo-sized, tortilla-making</i> |
| japonés | <i>tsunami-triggered, katana-wielding, miso-glazed, karate-based</i> |
| hebreo | <i>shibboleth-based, kosher-keeping, matzo-flavored</i> |
| portugués | <i>brigadeiro-making, acai-based, pagoda-roofed</i> |
| alemán | <i>zeitgeist-surfing, Führer-fueled</i> |
| árabe | <i>fatwa-funding, hookah-toking, dia-paying</i> |
| malayo | <i>sarong-wrapped (women)</i> |

Se registró un caso en que el préstamo funciona como núcleo: el vocablo *blitz* de origen alemán (*blitzkrieg* = guerra relámpago) es la base del compuesto *budget-blitzing*.

7.4.3. Truncamientos

En los corpus se verificó la presencia de varios truncamientos como primer elemento de estas estructuras compuestas:

Tabla 42. Truncamientos en los compuestos

| | |
|-------------------------------|-----------------------|
| <i>ad (advertisement)</i> | <i>ad-based</i> |
| <i>admin (administration)</i> | <i>admin-avoiding</i> |
| <i>bot (robot)</i> | <i>bot-controlled</i> |
| <i>bra (brassiere)</i> | <i>bra-burning</i> |

| | |
|------------------------------|---|
| <i>fem (female)</i> | <i>female-fronted (bands)</i> |
| <i>gym (gymnasium)</i> | <i>gym-toned (bodies), gym-primped (self)</i> |
| <i>info (information)</i> | <i>info-packed (text)</i> |
| <i>lab (laboratory)</i> | <i>lab-tested</i> |
| <i>nano (nanotechnology)</i> | <i>nano-engineered</i> |
| <i>pram (perambulator)</i> | <i>pram-pushing (mothers)</i> |
| <i>stat (statistics)</i> | <i>stat-tracking</i> |
| <i>shroom (mushroom)</i> | <i>shroom-fueled (stupor)</i> |
| <i>tech (technology)</i> | <i>tech-addled</i> |

Aun cuando el sustantivo *carbohydrate* aparece en el compuesto *carbohydrate-restricted*, las formas reducidas *carb* y *carbo* son más frecuentes, como se aprecia en *carb-controlled*, *carb-starved*, *carb-loading* y *carbo-loading*.

7.4.4. Hipocorísticos

También se registraron ocurrencias de hipocorísticos como *techie* (persona con habilidades y conocimiento sobre tecnología, especialmente computadores), *veggie* (verdura), *lefty* or *leftie* (simpatizante de izquierda), *hankie* (pañuelo), *broolly* (paraguas) y *potty* (pelela) en los compuestos *techie-oriented*, *veggie-loving*, *lefty-propagated* / *leftie-oriented*, *hankie-wringing*, *broolly-breaking* y *potty-trained*, respectivamente. Además, se encontró el hipocorístico del nombre de un personaje de cómics (*Wolverine*) en el compuesto *Wolvie-shaped*.

7.4.5. Acrónimos y siglas

Tanto los acrónimos (NASA) como las siglas (CIA) son frecuentemente usados en este tipo de construcciones, siendo mucho más comunes en adjetivos compuestos terminados en *-ed* (318) que en aquellos con *-ing* (29). Un gran número de ellos tiene como núcleo los verbos *approve*, *sponsor*, *fund* y *base*:

Tabla 43. Acrónimos y siglas en los compuestos

| | | | |
|-----------------------|----------------------|---------------------|------------------|
| <i>FDA-approved</i> | <i>UN-sponsored</i> | <i>CIA-funded</i> | <i>DNA-based</i> |
| <i>USCIS-approved</i> | <i>EU-sponsored</i> | <i>FDA-funded</i> | <i>SF-based</i> |
| <i>OSHA-approved</i> | <i>NCI-sponsored</i> | <i>NSERC-funded</i> | <i>LA-based</i> |

7.4.6. Reduplicación

En la muestra tomada del corpus *enTenTen12* aparecen las formas *flip-flapping*, *zig-zagging* y *criss-crossing* las que fueron erróneamente identificadas como compuestos por la plataforma de búsqueda (*Sketch Engine*). Cabe recordar que los términos duplicados no están conformados por dos lexemas independientes por lo que no pueden ser considerados formas compuestas. Sin embargo, representan una complicación al momento de recopilar datos desde un corpus electrónico dado que por su forma se asemejan a los adjetivos en estudio. De hecho, solo se verificó un caso de reduplicación en el compuesto *goo-goo-eyed*.

7.4.7. Compuestos como parte de otros compuestos

Como hemos visto, los distintos procesos morfológicos no solo coocurren sino que también son recursivos. En consecuencia, no es extraño encontrar palabras compuestas que participan de la creación de otro compuesto, tanto en posición nuclear como primer elemento.

Tabla 44. Compuestos como primer constituyente

| | |
|--------------------------------|--|
| <i>textbook-based</i> | <i>straightforward-looking</i> |
| <i>cupcake-printed</i> | <i>baseball-cap-wearing</i> |
| <i>fiberglass-filled</i> | <i>Triple Crown-winning (horse)</i> |
| <i>horseshoe-shaped</i> | <i>heartstring-pulling (pictures)</i> |
| <i>smartphone-based</i> | <i>bow-tie-wearing (cop)</i> |
| <i>workload-focused</i> | <i>Old World-pining (Southern aristocracy)</i> |
| <i>stormtrooper-led</i> | <i>Sunday-school-teaching</i> |
| <i>touchscreen-based</i> | <i>riverside-dwelling</i> |
| <i>fingerprint-based</i> | <i>classroom-building</i> |
| <i>wheelchair-bound</i> | <i>nightstick-wielding</i> |
| <i>keyboard-controlled</i> | <i>highlight-producing</i> |
| <i>bubble gum coloured</i> | <i>bullwhip-wielding</i> |
| <i>snake-skin-patterned</i> | <i>livestock-keeping</i> |
| <i>gingerbread-trimmed</i> | <i>backpack-emptying</i> |
| <i>whitespace-separated</i> | <i>cornerstone-laying</i> |
| <i>cerebral palsy-stricken</i> | <i>desktop-swapping</i> |
| <i>seafood-based (menu)</i> | <i>laptop-carrying</i> |
| <i>hardware-accelerated</i> | <i>snowman-making</i> |
| <i>comic-book based (film)</i> | <i>cellphone-toting (youth)</i> |
| <i>lily pad-strewn (swamp)</i> | <i>spotlight-basking (mother)</i> |
| <i>seaweed-covered (rocks)</i> | <i>matchbox-dwelling (insect)</i> |

| | |
|--|------------------------|
| <i>grammar-school educated</i> <i>heavy metal-styled (settings)</i> <i>floodwater-carried (sediment)</i> <i>West German-built (factories)</i> <i>venture capital-backed (company)</i> <i>brown paper-covered (guidebooks)</i> <i>nuclear weapons-contaminated (waters)</i> <i>Bond Girl inspired (Versace Oscars dress)</i> <i>Academy Award-nominated (performance)</i> | <i>X-ray-detecting</i> |
|--|------------------------|

Tabla 45. Compuestos en posición nuclear

| | |
|---|--|
| <i>once-dreadlocked</i> <i>white-nightgowned (figure)</i> <i>secretly-videotaped</i> <i>centrally-timetabled</i> <i>four-bedroomed (house)</i> <i>increasingly-networked</i> <i>Germany-headquartered</i> <i>powdery-eyebrowed</i> | <i>syntax-highlighting</i> <i>dolphin-sightseeing</i> <i>energy-benchmarking</i> |
|---|--|

7.5. Números en los compuestos

Es frecuente encontrar números simples (una palabra), especialmente cardinales, como primer constituyente de un adjetivo compuesto de base nominal (N-*ed*) como *one-eyed* (tuerto), *two-roomed* (de dos habitaciones), *three-storied house* (casa de tres pisos). La presencia de números

ordinales es más esporádica y ocurre principalmente en compuestos de núcleo verbal: *fourth-rated*, *third-ranking*, *eleventh-leading*, etc.

Algunos especialistas (Bauer, Lieber y Plag, 2013: 423-424) consideran que los números entre 20 y 99, por ejemplo *twenty-seven* (27) o *eighty-three* (83), semánticamente se comportan como compuestos coordinativos (*mother-child*, *Alsace-Lorraine*) puesto que su significado se deriva de la suma de las partes. Puesto que en inglés los años se escinden al momento de leerlos, es decir, se dividen en dos elementos ($1960 \rightarrow 19 + 60 = \textit{nineteen} + \textit{sixty}$), se asemejan bastante a los compuestos coordinativos, tal como lo plantean esos autores. Adicionalmente, cuando se trata de décadas, existe la costumbre de pluralizar el segundo elemento (*1960s*) y, muchas veces, de omitir el primero *'60s* (*sixties*). Me parece interesante destacar el hecho de que ambas formas son aceptables al interior de un compuesto, como lo demuestran estos ejemplos: *1960s-inspired look / '60s-inspired slabs of chrome*.

Asimismo, Bauer, Lieber y Plag (2013) también sostienen que al aumentar las cifras a cientos, miles y millones, su estructura es más bien sintáctica ya que los elementos pueden agregar la conjunción *and* de manera que *525,003* puede tener dos realizaciones:

five hundred twenty-five thousand and three

*five hundred **and** twenty-five thousand and three*

En el corpus de prueba se encontraron dos casos en los que el primer constituyente es una medida (tres cuartos): *three-quarter-finished jars of cream*, la que también puede ser expresada numéricamente: *3/4-sleeved sheer cream blouse*. La primera realización deja de manifiesto la naturaleza sintáctica de la fracción, la cual debe ser analizada como una frase nominal (*three-quarters*) cuyo núcleo (*quarters*) es pre-modificado por un número (*three*). Además, la forma plural ha sido cambiada a singular para adecuarse a la estructura típica de un adjetivo compuesto. Por su parte, los números ordinales se utilizan para referirse a períodos históricos (siglos) en compuestos como *19th-century-looking*, cuyo constituyente inicial también es una frase nominal (*19th century*).

Ejemplos como estos no están contemplados en las descripciones de este tipo de compuestos ya que al incluir un componente sintáctico dentro de un elemento léxico su análisis se aleja de las descripciones estándar.

7.6. Partículas

Roeper y Siegel (1978) en su descripción de compuestos verbales mencionan las partículas como posibles candidatos a ser incluidas mediante el Principio de Contigüidad. Sin embargo, sus ejemplos las ubicaban al comienzo de la forma compuesta: *ongoing*, *incoming*, *bygone*, etc. Como se explica en la sección 1.5. los verbos frásticos (*phrasal verbs*) representan un problema dado que las partículas *on-*, *in-* y *by-* también han sido descritas como adverbios y prefijos. En el proceso de depuración, se eliminaron la mayor parte de ellos y solo se mantuvieron aquellos en los que el primer elemento cumple una función claramente adverbial (*up-springing*, *up-standing*, *down-going*, etc.).

Por lo pronto, en nuestros corpus se verificó la ocurrencia de varios compuestos en los que la partícula se pospone:

well-thought-out

well-brought-up (*girl*)

well-cared-for (*hands*)

well-thought-of (*person*)

long-looked-for (*tidings*)

much-sought-after (*ticket*)

well-worked-out (*exploration*)

much-buzzed-about (*political thriller*)

En estos casos, el verbo continúa ocupando la posición nuclear pero con un elemento post-modificador, lo que para muchos especialistas no es una característica propia de una palabra compuesta. Podemos observar que aunque al interior de estos compuestos los verbos frásticos muestran integridad léxica al mantener sus componentes unidos, la estructura comienza a manifestar rasgos que la acercan más a la sintaxis.

Capítulo 8. Análisis de los patrones morfológicos individuales

8.1. Compuestos con núcleo nominal (N-*ed*)

8.1.1. Adj + N-*ed* / N + N-*ed* en la descripción de rasgos físicos y de carácter

Aun cuando la descripción estándar de estas unidades compuestas incluye una forma verbal como segundo elemento, en el corpus de prueba (*sample corpus*) se constató que el segundo patrón morfológico más productivo de la categoría terminada en *-ed* resultó ser aquel en el que el núcleo se deriva de un sustantivo con 1.379 casos. En el corpus general también se recolectaron muchos ejemplos (785).

Este patrón se distingue de los otros en que el segundo elemento no corresponde al participio pasado de un verbo regular sino que es de origen nominal. El segundo elemento es un sustantivo que incorpora el sufijo *-ed* para señalar su rol de adjetivo. Dentro de esta categoría de adjetivos compuestos un gran número de ellos denotan características físicas o inherentes según la descripción de Ljung (1976). En general, el segundo constituyente de estos compuestos es un elemento compositivo que no tiene autonomía léxica. En inglés no tiene sentido describir a las personas usando adjetivos como *an *eyed man, a *necked woman, a *footed child, *minded parents or *hearted people* puesto que todos ellos se refieren a rasgos físicos propios de los seres humanos. Constatamos entonces que formas como *-eyed, -necked, -footed, -minded y -hearted* solo pueden existir al interior de una estructura compuesta como lo ilustran *one-eyed* (tuerto), *left-handed* (zurdo), *long-necked* (de cuello largo), *flat-footed* (de pies planos), *open-minded* (de mente abierta) o *kind-hearted* (bondadoso) ya que estas expresiones solo se vuelven posibles al agregar un especificador como primer elemento del compuesto (cf. sección 3.3.1.).

En el corpus de prueba se comprobó que dentro de este grupo prácticamente dos tercios corresponden a la estructura Adj + N-*ed*, con un total de 860 instancias mientras que el patrón S + N-*ed* registró 422.

Tabla 46. Partes del cuerpo / características físicas

| | |
|---|---|
| arm: <i>thin-armed, long-armed</i> | jaw: <i>square-jawed, heavy-jawed</i> |
| ass/arse: <i>big-assed, pansy-arsed*</i> | knee: <i>weak-kneed, bent-kneed</i> |
| back: <i>straight-backed, stiff-backed</i> | knuckle: <i>white-knuckled, callus-knuckled</i> |
| beard: <i>grey-bearded, long-bearded</i> | lash: <i>fair-lashed, long-lashed (eyes)</i> |
| belly: <i>large-bellied, yellow-bellied*</i> | leg: <i>long-legged, stiff-legged</i> |
| blood: <i>warm-blooded, cold-blooded</i> | lid: <i>heavy-lidded, dull-lidded (eyes)</i> |
| body: <i>big-bodied, hard-bodied</i> | limb: <i>large-limbed, long-limbed</i> |
| bone: <i>big-boned, delicate-boned</i> | lip: <i>full-lipped, scarlet-lipped</i> |
| breast: <i>bare-breasted, rich-breasted</i> | mouth: <i>dry-mouthed, loud-mouthed</i> |
| brow: <i>shaggy-browed, white-browed</i> | neck: <i>long-necked, bare-necked</i> |
| cheek: <i>rosy-cheeked, hollow-cheeked</i> | nose: <i>red-nosed, long-nosed</i> |
| chest: <i>bare-chested, flat-chested</i> | palm: <i>sweaty-palmed, wet-palmed</i> |
| chin: <i>sharp-chinned, double-chinned</i> | profile: <i>gaunt-profiled</i> |
| complexion: <i>smooth-complexioned</i> | pulse: <i>warm-pulsed (wrists)</i> |
| ear: <i>big-eared, sharp-eared</i> | shoulder: <i>broad-shouldered, big-shouldered</i> |
| eye: <i>blue-eyed, sad-eyed</i> | sight: <i>short-sighted, long-sighted*</i> |
| face: <i>broad-faced, pleasant-faced</i> | skin: <i>fair-skinned, smooth-skinned</i> |
| feature: <i>broad-featured, gentle-featured</i> | stomach: <i>flat-stomached</i> |
| figure: <i>full-figured</i> | thigh: <i>naked-thighed (beauties)</i> |
| finger: <i>long-fingered, light-fingered*</i> | toe: <i>short-toed, nimble-toed</i> |
| fist: <i>iron-fisted, tight-fisted*</i> | toothed: <i>big-toothed, gap-toothed</i> |
| foot: <i>flat-footed, light-footed</i> | throat: <i>full-throated, rough-throated</i> |
| hair: <i>short-haired, red-haired</i> | vein: <i>rope-veined (hands)</i> |
| hand: <i>left-handed, beast-handed</i> | visage: <i>dark-visaged, grim-visaged</i> |
| head: <i>golden-headed, hard-headed*</i> | voice: <i>husky-voiced, soft-voiced</i> |
| iris: <i>full-irised (eyes)</i> | waist: <i>narrow-waisted</i> |
| hip: <i>slim-hipped</i> | wrist: <i>limp-wristed</i> |

* rasgo de carácter

De este listado, los elementos compositivos más productivos son *-faced*, con 94 ocurrencias, y luego *-eyed*, con 78. Bastante más atrás les siguen *-headed* (60), *-haired* (43) y *-hearted* (39).

A su vez, algunos adjetivos compuestos somáticos son usados como característica de una parte del cuerpo o de un rasgo físico: *full-irised eyes*, *dull-lidded eyes*, *white-knuckled hands*, *fair-lashed eyes*, *full-throated laughter*. Otros, además de su significado literal, agregan un sentido metafórico para referirse a un rasgo de personalidad: *yellow-bellied* (cobarde), *hard-headed* (testarudo), *light-fingered* (“manilargo”), *long-sighted* (con visión de futuro).

Por otra parte, hay un grupo de palabras referidas al cuerpo (*blood*, *brain*, *heart*, *tongue*) que se utilizan casi exclusivamente en formas compuestas que describen rasgos de carácter o personalidad. De esta manera, *blood* (sangre) se usa en compuestos que describen el origen (*blue-blooded*, *pure-blooded*), la pasión (*hot-blooded*), la frialdad (*cold-blooded*, *ice-blooded*) o la cobardía (*milk-blooded*, *thin-blooded*). *Brain* (cerebro) se combina principalmente con ciertos sustantivos con los que se establece una comparación que expresa falta de inteligencia: *pea-brained*, *hare-brained*, *birdbrained*, *feather-brained*. La excepción la constituyen ejemplos como *large brained animals* y *left-brained* en los que literalmente se hace referencia al cerebro (o parte de él) como órgano. El término *heart* (corazón) es muy productivo puesto que entra en combinación con muchas diferentes palabras, mayoritariamente adjetivos, para denotar tanto rasgos positivos de bondad, honestidad y valentía como es el caso de *pure-hearted*, *kind-hearted*, *open-hearted*, *lion-hearted* y también rasgos negativos como tristeza, crueldad y falta de empatía en general como en *heavy-hearted*, *cold-hearted*, *hard-hearted* y *empty-hearted*. Por su parte, las formas compuestas que incorporan *tongue* (lengua) denotan, en general, la rapidez y la rudeza para responder o expresarse: *quick-tongued*, *sharp-tongued*, *rough-tongued*, *vicious-tongued*, así como la facilidad de expresión y poder de persuasión: *smooth-tongued*, *silver-tongued*.

Igualmente existen términos que se refieren exclusivamente a rasgos de carácter que son muy productivos en la formación de adjetivos compuestos. De ellos, el que más ocurrencias registra es *mind* con 79 ejemplos en total, encontrados en ambos corpus. Estos adjetivos denotan diversos tipos

de mentalidad: justa, conservadora, abierta, determinada, débil o maligna: *fair-minded*, *narrow-minded*, *open-minded*, *strong-minded*, *weak-minded*, *evil-minded*. También pueden referirse a los intereses y preocupaciones de una persona: *green-minded* (con conciencia ecológica), *civic-minded* (con conciencia cívica), *bloody-minded filmmaker* (con por las escenas sangrientas), *money-minded* (interesado en el dinero).

Cabe hacer notar que, a diferencia de otros elementos compositivos somáticos, entre aquellos formados a partir de *-minded* como base, el patrón morfológico que incluye un adverbio es bastante regular al momento de describir tanto ciertos intereses como formas de pensar:

analytically-minded

statistically-minded

spiritually-minded

mechanically-minded

pragmatically-minded

educationally-minded

commercially-minded

oppositionally-minded

Aun cuando en los corpus se encontraron las formas equivalentes *independent-minded* e *independently-minded*, no podemos asegurar que los otros compuestos tengan una realización alternativa con un adjetivo como primer constituyente en vez de un adverbio.

Otras palabras usadas en la construcción de adjetivos compuestos que describen rasgos de personalidad son, en orden de productividad:

1) temper: *bad-tempered*, *equal-tempered*, *even-tempered*, *gentle-tempered*, *good-tempered*, *hot-tempered*, *ill-tempered*, *nasty-tempered*, *placid-tempered*, *short-tempered*, *quick-tempered*, *sweet-tempered*, *vile-tempered*.

2) spirit: *free-spirited, fun spirited, high-spirited, light-spirited, low-spirited, mean-spirited, sweet-spirited, tough-spirited.*

3) wit: *dim-witted, dull-witted, keen-witted, quick-witted, razor-witted, sharp-witted.*

4) nature: *bad-natured, good-natured, ill-natured, kind-natured, sweet-natured.*

5) will: *good-willed, strong-willed, weak-willed.*

6) humour/humor: *good-humoured, ill-humored, gallows-humoured.*

7) soul: *tough-souled, whole-souled.*

8.1.2. Adj + N-ed / N + N-ed en la descripción y nombres de animales

Así como es frecuente usar adjetivos compuestos somáticos en la descripción del cuerpo humano, también pueden describir otras criaturas refiriéndose a algunas características anatómicas compartidas (*feet, head, legs, ears, back, etc.*) y otras específicas del animal como pelaje (*black-furred*), plumaje (*red-feathered*), garras (*blunt-clawed*), cuernos (*long-horned cattle*), etc. En consecuencia, la zoología es un área muy proclive a usar este particular tipo de unidades compuestas al momento de nombrar una especie animal. En inglés, muchos nombres de animales incorporan adjetivos compuestos de base nominal para resaltar una característica física que sirve para distinguir al animal de otros. El primer elemento del compuesto tiene una función descriptiva con respecto al núcleo y puede, por lo tanto, tomar la forma de un adjetivo o un sustantivo:

red-winged blackbird (Adj + N-ed) (mirlo de ala colorada)

lion-tailed macaque (N + N-ed) (macaco cola de león)

Durante la confección del corpus de muestra, el sitio web de National Geographic⁹ resultó muy productivo en cuanto a la recolección de este tipo de datos (69 ejemplos) los que pusieron de relieve el uso especializado de ciertos adjetivos compuestos. En la muestra tomada del corpus electrónico se encontraron alrededor de 38 compuestos que describen animales pero no se ha determinado cuántos de ellos corresponden a nombres de especies. Los datos recolectados demuestran que la estructura cuyo primer elemento es un adjetivo es la más productiva y que los nombres que incorporan adjetivos compuestos son más comunes entre las aves. En la siguiente tabla se listan los ejemplos encontrados en el corpus de muestra, los que han sido divididos según la categoría sintáctica del primer elemento y el tipo de animal:

Tabla 47. Adjetivos compuestos de base nominal en nombres de aves

| AVES | Adjetivo + N- <i>ed</i> | AVES | Sustantivo + N- <i>ed</i> |
|------|--------------------------------------|------|---------------------------------------|
| | <i>olive-backed sunbird</i> | | <i>groove-billed ani</i> |
| | <i>red-billed queleas</i> | | <i>ring-billed gulls</i> |
| | <i>lilac-breasted roller</i> | | <i>lappet-faced vultures</i> |
| | <i>orange-breasted sunbird</i> | | <i>racket-tailed drongos</i> |
| | <i>red-breasted geese</i> | | <i>ribbon-tailed bird of paradise</i> |
| | <i>red-breasted passenger pigeon</i> | | <i>wire-tailed manakin</i> |
| | <i>scaly-breasted kingfisher</i> | | <i>club-winged manakin</i> |
| | <i>scaly-breasted munias</i> | | |
| | <i>silver-breasted broadbills</i> | | |
| | <i>white-cheeked barbets</i> | | |
| | <i>white-cheeked bulbul</i> | | |
| | <i>white-cheeked turaco</i> | | |

⁹ National Geographic: <http://www.nationalgeographic.com/>

| | |
|---|--|
| <p><i>red-chested cuckoos</i></p> <p><i>erect-crested penguin</i></p> <p><i>purple-crested turacos</i></p> <p><i>red-crested touraco</i></p> <p><i>long eared owl</i></p> <p><i>short-eared owl</i></p> <p><i>yellow-eyed penguin</i></p> <p><i>blue-faced honeyeater</i></p> <p><i>blue-footed booby</i></p> <p><i>rusty-fronted barwings</i></p> <p><i>black-headed herons</i></p> <p><i>brown-headed gull</i></p> <p><i>golden-headed cisticola</i></p> <p><i>great horned owls</i></p> <p><i>red-legged crakes</i></p> <p><i>rough-legged hawk</i></p> <p><i>black-necked stilt</i></p> <p><i>white-necked raven</i></p> <p><i>red-ruffed fruitcrows</i></p> <p><i>black-rumped flamebacks</i></p> <p><i>yellow-shafted northern flicker</i></p> <p><i>blue-tailed bee-eater</i></p> <p><i>red-tailed hawk</i></p> <p><i>white-tailed eagle</i></p> <p><i>blue-throated barbet</i></p> <p><i>white-throated sparrow</i></p> <p><i>wine-throated hummingbird</i></p> <p><i>black-winged stilt</i></p> <p><i>blue-winged minlas</i></p> | |
|---|--|

| | |
|---|--|
| <i>cobalt-winged parakeets</i> <i>red-winged blackbird</i> <i>single-wattled cassowary</i> <i>double-wattled cassowary</i> | |
|---|--|

Tabla 48. Compuestos de base nominal en nombres de otros animales

| MAMÍFEROS Adjetivo + N- <i>ed</i> | MAMÍFEROS Sustantivo + N- <i>ed</i> |
|--|---|
| <i>white-bellied spider monkey</i> <i>short-eared dog/fox</i> <i>small-eared dog</i> <i>white-faced capuchin monkeys</i> <i>white-fronted capuchin</i> <i>long-haired Dachsund</i> <i>German shorthaired pointer (dog)</i> <i>golden-mantled tamarind</i> <i>snub-nosed monkey</i> <i>golden-snub-nosed monkeys</i> <i>long-tailed macaque</i> | <i>lion-tailed macaque</i> |
| OTROS Adjetivo + N- <i>ed</i> | OTROS Sustantivo + N- <i>ed</i> |
| <i>red-eyed tree frog</i> <i>short-spined porcupinefish</i> <i>yellow-tailed jacks (fish)</i> <i>white-tailed deer</i> | <i>saw-scaled viper</i> <i>knob-tailed gecko</i> |

8.1.3. Adj + N-ed / N + N-ed en la descripción de vestimentas

La descripción de vestimentas es otra área en que los adjetivos compuestos de base nominal son altamente productivos. Algunos incluyen prendas de vestir, calzado o accesorios como núcleo tales como *apron* (delantal), *coat* (abrigo), *nightgown* (camisa de dormir), *shirt* (camisa), *suit* (traje), *slippers* (pantuflas), *gloves* (guantes) o *beret* (boina). Otros describen sus partes como *stiff-collared* (de cuello duro), *metallic-hemmed* (de ruedo metálico), *long-sleeved* (de manga larga) o *wide-lapelled* (de solapa ancha). También pueden describir su diseño, como lo ejemplifican *daisy-patterned* (con diseño de margaritas) y *floral-printed* (con diseño floral):

Tabla 49. Compuestos con núcleos nominales referidos a vestimenta

| |
|---|
| <i>white-aproned (landlord)</i> |
| <i>black-bereted (man)</i> |
| <i>white-bloused</i> |
| <i>round-brimmed, wide-brimmed (hat)</i> |
| <i>black-cassocked</i> |
| <i>grey-cloaked, red-cloaked</i> |
| <i>red-coated, white-coated (doctors)</i> |
| <i>stiff-collared</i> |
| <i>black-dressed</i> |
| <i>black-frocked, white-frocked</i> |
| <i>purple-gloved, leather-gloved (hands)</i> |
| <i>crochet-gowned</i> |
| <i>black-hatted, top-hatted, feather-hatted</i> |
| <i>high-heeled, low-heeled, flat heeled (shoes)</i> |
| <i>brown-helmeted</i> |
| <i>metallic-hemmed</i> |
| <i>yellow-hooded</i> |

wide-lapelled, slim-lapelled (suit)
wide-bottomed (jeans)
floral-lined (jacket)
low-necked, open-necked (shirt)
white-nightgowned
large-patterned (Bermuda shorts), daisy-patterned (bathrobe)
white-pinafores (baby)
floral-printed (outfit)
black-robed (man), violet-robed (figure)
silver-sequined (dress)
flowery-shirted, green-shirted, clean-shirted
white-shoed (Mary)
full-skirted (dress)
long-sleeved, short-sleeved, single-sleeved
red-slippered
white-smocked (man)
flat-soled, thick-soled (shoes)
blue-striped, white-striped (shirt), pin-striped (trousers)
dark-suited, sharp-suited (man), tweed-suited, designer-suited
check-trousered
green-uniformed (soldiers)
red-vested
high-waisted (skirt), shirtwaisted (cashmere dress)

8.1.4. Participios adjetivos + N-*ed*

En los corpus se identificaron algunos compuestos en los que primer constituyente es un participio adjetivo (*participial adjective*), lo que corrobora la función adjetival que muchas de las formas de participio pasado pueden cumplir en inglés.

broken-backed (ship)

swollen-headed (slob)

broken-hearted

shaven-headed

crooked-legged (terrier)

cracked-voiced

extended-sleeved

ragged-edged

spotted-tailed

Además, se encontraron compuestos cuyo primer elemento es un adjetivo que ha sido innecesariamente convertido en participio. Esta alteración afecta incluso a unidades compuestas de uso común listadas en diccionarios. Es probable que esto ocurra por analogía con el patrón anterior, lo que produce ejemplos como:

Forma no estándar

slacked-jawed

opened-ended

singled-sided

doubled-sided

doubled-bladed

wedged-shaped

Forma estándar

slack-jawed

open-ended

single-sided

double-sided

double-bladed

wedge-shaped

8.1.5. Adjetivo comparativo / superlativo + N-*ed*

Puesto que uno de los patrones más productivos tiene un adjetivo como especificador del núcleo nominal, es factible encontrar ejemplos en los que el primer elemento es una forma comparativa o superlativa de dicho adjetivo. Por ejemplo:

Tabla 50. Adjetivos comparativos y superlativos en compuestos de base nominal

| Adj comparativo + N- <i>ed</i> | Adj superlativo + N- <i>ed</i> |
|---|--|
| <i>lighter-boned (than ...)</i> | <i>keenest-eyed</i> |
| <i>bigger-breasted (customers)</i> | <i>cheapest-priced (claret)</i> |
| <i>bigger-busted (women)</i> | <i>(the) worst-tempered</i> |
| <i>harder-edged</i> | <i>(the) widest-winged (condor in the Andes)</i> |
| <i>better-humored (than ...)</i> | <i>best-humoured</i> |
| <i>higher-minded (resources)</i> | <i>softest-toned</i> |
| <i>bigger-named</i> | <i>biggest-nosed</i> |
| <i>larger-beaked</i> | <i>best-flavoured</i> |
| <i>lower-numbered</i> | <i>highest-numbered</i> |
| <i>lighter-skinned (desert peoples)</i> | <i>nimblest-witted</i> |

8.1.6. Número + N-*ed*

El primer elemento de un adjetivo compuesto del tipo N-*ed* también puede ser un número que funciona como especificador o determinante del núcleo. Aunque muy raros en la descripción de personas en la que *one-eyed* y *two-faced (liar)* parecen ser los ejemplos más conocidos, son más comunes al describir animales (*eight-eyed spider* o *eight-legged creepy crawlies*), criaturas fantásticas (*two-headed beings*, *four-taloned hand*) y al nombrar animales como *nine-banded armadillo*, *five-lined skink* (un tipo de lagarto). Mucho más frecuentes son los compuestos que enumeran partes de objetos y viviendas: *one-engined* (monomotor), *six-sided figure* (figura de

seis lados), *two-edged sword* (espada de doble filo), *three-storied house* (casa de tres pisos), *two-bedroomed flat* (departamento de dos dormitorios), *three-windowed Victorian room*, etc.

La ocurrencia de números, especialmente cardinales, como primer constituyente de un adjetivo compuesto de base nominal se verificó en 51 instancias, mientras que la presencia de números ordinales es mucho más esporádica, registrando solo dos ejemplos:

Tabla 51. Números en compuestos de base nominal

| Número cardinal + N- <i>ed</i> | Número ordinal + N- <i>ed</i> |
|---------------------------------|-------------------------------|
| <i>one-eyed</i> | <i>first-named</i> |
| <i>two-handed (backhand)</i> | <i>second-sourced</i> |
| <i>three-cornered</i> | |
| <i>four-wheeled (cab)</i> | |
| <i>five-fingered</i> | |
| <i>six-pointed (structure)</i> | |
| <i>eight-legged</i> | |
| <i>nine-tailed</i> | |
| <i>ten-toed</i> | |
| <i>twelve-faceted</i> | |
| <i>thousand-columned (hall)</i> | |

Además de funcionar como número ordinal, *first* también actúa como adverbio en adjetivos compuestos como *first-mentioned* y *first-poured*, conformando el patrón morfológico adverbio + participio pasado.

8.1.7. Cuantificador + N-*ed*

Un cuantificador es un determinante que cumple una función muy similar a un número pero en términos menos concretos respecto a la cantidad referida (Downing, 2015: 384). Este patrón se reduce básicamente a *many*, el que funciona como una alternativa al prefijo *multi-* como lo ilustran algunos compuestos:

many-armed = *multi-armed*

many-columned (*Iram*)

many-coloured = *multi-coloured* (*myriad-colored*)

many-headed

many-faceted = *multi-faceted*

many-roomed (*mansion*)

many-layered = *multi-layered*

many-sided (*personality*)

many-coursed

many-tiered (*cake*)

Cabe hacer notar que el cuantificador *few*, semánticamente el opuesto de *many*, es mucho menos productivo y solamente registró una ocurrencia en los corpus analizados: *few-paged*.

Por su parte, en este patrón morfológico, *half* es poco productivo, a diferencia de lo que ocurre cuando acompaña a un núcleo verbal (cf. 8.3.2.), registrando solo seis ejemplos:

half-arsed

half-canopied

half-witted

half-lidded (*glance*)

half-hearted

half-timbered (*houses*)

Los adjetivos *double*, *twin*, *single* y *triple*, relativamente productivos en compuestos del tipo N-*ed*, constituyen un caso especial puesto que semánticamente cuantifican un sustantivo, aunque gramaticalmente no sean considerados cuantificadores propiamente tales. De hecho, en algunos ejemplos son equivalentes a un número como en *double-headed* = *two-headed* o *triple-headed* =

three-headed. Por este motivo, en nuestra clasificación, estos compuestos fueron incluidos en el patrón Número + N-*ed*.

| | |
|--------------------------------------|--------------------------------|
| <i>double-bedded (room)</i> | <i>double-tailed</i> |
| <i>double-bladed (axe)</i> | <i>double-headed</i> |
| <i>double-edged (dagger)</i> | <i>double-necked</i> |
| <i>double-breasted (suit/jacket)</i> | <i>double-faced</i> |
| <i>double-chinned</i> | <i>twin-faced</i> |
| <i>double-walled (stone huts)</i> | <i>twin-engined (aircraft)</i> |
| <i>double-sided</i> | <i>single-celled (algae)</i> |
| <i>double-plated (glass)</i> | <i>single-engined</i> |
| <i>double-winged (house)</i> | <i>single-breasted</i> |
| <i>double-screened</i> | <i>triple-headed</i> |

Sin embargo, hay que tener presente que *double* puede expresar no solo número sino que también otras nociones como: que dos capas se superponen (*double-breasted*), que algo tiene un grosor excesivo (*double-chinned*, *double-necked*), o que ha sido reforzado (*double-barred wall*).

8.1.8. Adverbio + N-*ed*

Este patrón pareciera no ser gramaticalmente posible puesto que un adverbio no funciona como determinante de un sustantivo. No obstante, en el corpus de muestra se registraron compuestos como *well-intentioned* y *well-mannered* cuyo núcleo es un sustantivo de base nominal puesto que no existen las formas verbales **to intention* y **to manner*. Algunos compuestos tienen dos realizaciones en usos y contextos similares: *a bright-coloured dress* / *a brightly-coloured dress* y *good-sized* / *well-sized* lo que nos hace pensar que han sido creados por analogía con adjetivos compuestos de base verbal, los que generalmente incorporan un adverbio como primer elemento.

Otros ejemplos son:

respectably-intentioned (lovers)

gravelly-voiced (film star)

well-proportioned

pragmatically-minded

tightly-bodiced (dress)

nobly-profiled (man)

normally-sighted

wonderfully-named

Al igual como ocurre en los otros patrones del tipo N-*ed*, con frecuencia el segundo constituyente es un elemento compositivo y no un lexema por lo que no pueden funcionar de manera independiente (**intentioned lovers, a *bodiced dress, a *voiced film star, etc.*), por lo tanto, solo son posibles en una estructura compuesta.

8.2. Compuestos con núcleo verbal (V-*ed*)

8.2.1. Sustantivo + participio pasado

Este es el patrón morfológico más productivo de los adjetivos compuestos de base verbal y en ellos el sustantivo inicial desempeña una función adverbial de adjunto (*adjunct*). Entre las diferentes roles semánticos que expresan están:

Agente:

mouse-nibbled (pie crust), moth-eaten, parent-arranged (visit), Nazi-occupied (France), British-administered (prison), Spanish-owned (Cuba), slave-built, age-worn (stones), Vatican-endorsed (saints), God-cursed, wind-moulded (dunes), doctor-assisted (suicide), hurricane-obliterated (town), Government-controlled (areas), military-inspired (coat)

Instrumento / medio:

pencil-scrawled (words), candle-lit (room), hand-knitted, machine-tooled (work), diesel-powered (products), napkin-wrapped (bottle of saki), laser-printed, steam-injected

Modo:

custom-built (car), vacuum-sealed (cans), razor-sharpened, oak-gnarled (fingers), cash-settled, precision-guided (weapon)

Locativo:

Australia-based (interactive show), Madrid-based, Harvard-educated, factory-built, hospital-acquired (infections), heaven-sent, wall-mounted (clock), vine-ripened (tomatoes), shop-bought, world-renowned (critic), U.S.-distilled (corn whiskey), family-centered, Edinburgh-set, convent-bred (girl), Brooklyn-born

8.2.1.1. Adjetivos compuestos con *-born*

Los adjetivos compuestos formados con el participio *born* (nacido) constituyen un caso interesante de analizar puesto que al ser el primer elemento el nombre del lugar de nacimiento de una persona, esta información puede ser entregada en diferentes niveles de especificidad, desde una ubicación geográfica general (país), una intermedia (estado) o una específica (ciudad o pueblo). Aun cuando muchas veces se incluyen formas reducidas (abreviaturas y acrónimos) como constituyente inicial, la longitud del nombre parece no ser particularmente relevante, por lo tanto, el compuesto *D.C.-born* coexiste con una alternativa más extensa como *Washington D.C.-born*. En las siguientes tablas se detallan los tipos de compuestos formados con *-born*:

Tabla 52. Compuestos con *-born* referidos a países y estados

| Países | Estados / Provincias |
|-----------------------|-------------------------------------|
| <i>U.S.-born</i> | <i>Ohio-born</i> |
| <i>U.K.-born</i> | <i>Kentucky-born</i> |
| <i>Senegal-born</i> | <i>Alabama-born</i> |
| <i>India-born</i> | <i>New Jersey-born</i> |
| <i>Lebanon-born</i> | <i>Pennsylvania-born</i> |
| <i>Brazil-born</i> | <i>B.C.-born (British Columbia)</i> |
| <i>Sri Lanka-born</i> | <i>D.F.-born</i> |

Tabla 53. Compuestos con *-born* referidos a ciudades y municipios

| Ciudades | Municipios / Distritos |
|---|------------------------|
| <i>Dublin-born</i> | <i>Brooklyn-born</i> |
| <i>London-born</i> | |
| <i>Bagdad-born</i> | |
| <i>Shanghai-born</i> | |
| <i>Barcelona-born</i> | |
| <i>Bethlehem-born</i> | |
| <i>Belfast-born</i> | |
| <i>St Ann's-born</i> (Hay un genitivo presente en el nombre ya que se trata de una reducción de <i>St Ann's Bay</i> en Jamaica) | |

Algunos de estos compuestos pueden volverse bastante complejos al combinar dos locaciones, es decir dos sustantivos, los que, generalmente, toman dos formas distintas: un nombre y una abreviatura (*Chino, Calif.-born*) o un nombre y un acrónimo (*Winnipeg, MB-born*). Es así que un adjetivo como *Connecticut-born* también cuenta con una forma abreviada (*Conn.-born*) la que puede aparecer en una estructura compuesta más compleja como “*the two **Groton, Conn.-born** Donnelly brothers*”. A continuación se listan las combinaciones encontradas en el corpus electrónico:

Tabla 54. Compuestos con *-born* referidos a dos localidades

| Ciudad y estado | Condado y estado |
|----------------------------|---------------------------------|
| <i>Groton, Conn.-born</i> | <i>Chester County, Pa.-born</i> |
| <i>Brooklyn, N.Y.-born</i> | |
| <i>Columbus, Ga.-born</i> | |
| <i>Hamilton, Ont.-born</i> | |
| <i>Chino, Calif.-born</i> | |
| <i>Waltham, Mass.-born</i> | |
| <i>Winnipeg, MB-born</i> | |

-born también forma adjetivos compuestos incorporando el año de nacimiento e incluso la década:

Tabla 55. Compuestos con *-born* referidos a tiempo

| Años | Décadas | Otros períodos |
|------------------|------------------|--------------------|
| <i>1997-born</i> | <i>80's-born</i> | <i>summer-born</i> |
| <i>2005-born</i> | | |
| <i>'93-born</i> | | |

Otra característica del participio *born* es el hecho de que pueden combinarse con adjetivos. Por ejemplo, el compuesto *U.S.-born* registra la ocurrencia de una estructura alternativa como *American-born* en la que el primer elemento denota una característica. Lo mismo sucede con *U.K.-born* y *British-born*. Por ende, es necesario que distinguir entre dos patrones morfológicos diferentes:

Tabla 56. Patrones morfológicos en que participa *-born*

| Sustantivo + participio pasado | Adjetivo + participio pasado |
|--------------------------------|------------------------------|
| <i>U.S.-born</i> | <i>American-born</i> |
| <i>U.K.-born</i> | <i>British-born</i> |
| | <i>Italian-born</i> |
| | <i>Polish-born</i> |

Los compuestos que incluyen nacionalidades o gentilicios (*British-born*, *Scottish-born*) hacen referencia indirectamente al lugar de nacimiento. Los adjetivos más bien denotan el estado o condición en que se nació. En otras palabras, así como se puede nacer “británico”, “escocés” o “polaco”, también se puede nacer en cautiverio (*captive-born*), en libertad (*freeborn*) o muerto (*still-born*).

8.2.1.2. Adjetivos compuestos con *-based*

De igual forma, entre la variedad de adjetivos compuestos que tienen *-based* como núcleo algunos también denotan una localidad geográfica y, al igual que aquellos formados con *-born*, se combinan con diferentes sustantivos dependiendo de lo específico de la información entregada, por lo que resultan muy productivos (652 ocurrencias en la muestra tomada del corpus electrónico y 49 en el de prueba), en especial los que incorporan el nombre de una ciudad:

Tabla 57. Compuestos con *-based* como núcleo

| Países | Estados | Ciudades | Distritos |
|--------------------------|--------------------------|--------------------------|------------------------|
| <i>Brazil-based</i> | <i>Ohio-based</i> | <i>Madrid-based</i> | <i>Brooklyn-based</i> |
| <i>Liberia-based</i> | <i>Tennessee-based</i> | <i>Nairobi-based</i> | <i>Manhattan-based</i> |
| <i>Korea-based</i> | <i>California-based</i> | <i>Prague-based</i> | |
| <i>Finland-based</i> | <i>Texas-based</i> | <i>Sydney-based</i> | |
| <i>Bahrain-based</i> | <i>Colorado-based</i> | <i>Jerusalem-based</i> | |
| <i>Jamaica-based</i> | <i>Illinois-based</i> | <i>Baghdad-based</i> | |
| <i>Australia-based</i> | <i>Connecticut-based</i> | <i>Glasgow-based</i> | |
| <i>Bermuda-based</i> | | <i>Bordeaux-based</i> | |
| <i>Sweden-based</i> | | <i>Lima-based</i> | |
| <i>Netherlands-based</i> | | <i>Atlanta-based</i> | |
| | | <i>London-based</i> | |
| | | <i>Ottawa-based</i> | |
| | | <i>Ankara-based</i> | |
| | | <i>Houston-based</i> | |
| | | <i>Los Angeles-based</i> | |

Curiosamente, el primer constituyente también puede ser un adjetivo y conformar un patrón morfémico diferente, como en el par *Australia-based* / *Australian-based*. Sin embargo, en este último caso (adjetivo + participio pasado), las localidades no incluyen ni estados ni ciudades, mucho menos distritos o municipios. La referencia geográfica es de tipo general (países o continentes):

Tabla 58. Patrones morfológicos en que participa *-based*

| Sustantivo + participio pasado | Adjetivo + participio pasado |
|--------------------------------|------------------------------|
| <i>Leipzig-based</i> | <i>Italian-based</i> |
| <i>Québec-based</i> | <i>African-based</i> |
| <i>Brazil-based</i> | <i>British-based</i> |
| <i>Sarajevo-based</i> | <i>Saudi-based</i> |
| <i>Bahrain-based</i> | <i>Israeli-based</i> |
| <i>Pakistan-based</i> | <i>Scottish-based</i> |
| <i>Saint Louis-based</i> | <i>German-based</i> |
| <i>Massachusetts-based</i> | <i>Taiwanese-based</i> |
| <i>Pyongyang-based</i> | <i>Cambodian-based</i> |

8.2.2. Nombres propios

Además de incluir nombres de locaciones geográficas, en este patrón morfológico (Sustantivo + participio pasado) es común encontrar nombres propios de personas, bandas, personajes y compañías como primer elemento de un compuesto. Estos sustantivos pueden ser simples o compuestos, variando generalmente entre una y tres palabras:

Tabla 59. Nombres propios en los compuestos

| Nombres propios simples | Nombres propios compuestos |
|--|--|
| <i>Beatles-inspired</i> | <i>New York City-based</i> |
| <i>Wagner-scored (helicopter attack)</i> | <i>Barry White-scored (sex scene)</i> |
| <i>Chandler-fuelled (frenzy)</i> | <i>Ian Holm-presented (prologue)</i> |
| <i>Manolete-inspired (tricks)</i> | <i>Jules Verne-inspired</i> |
| <i>Vermeer-inspired (lighting schemes)</i> | <i>Gene Kelly-loving (Sicilian hitman)</i> |
| <i>Cosmopolitan-promised (short stories)</i> | <i>Matthew Vaughn-directed (film)</i> |

| | |
|---|---|
| <i>Obama-endorsed (biography)</i> | <i>Wall Street Journal-hosted (after party)</i> |
| <i>Microsoft-sponsored</i> | <i>Freddie Highmore-shaped (soul)</i> |
| <i>Scorsese-inspired</i> | <i>Seth MacFarlane-hosted (Oscars)</i> |
| <i>Hemingway-based (special dramas)</i> | <i>Olivier Award-nominated</i> |

8.2.3. Adverbio + participio pasado

Este es uno de los patrones más productivos ya que combina dos formas relacionadas sintácticamente a nivel de oración, donde un adverbio puede funcionar como un especificador de modo, tiempo o cantidad de un verbo. En el primer caso, es esperable encontrar *well* en gran cantidad de combinaciones (*well-adjusted*, *well-trained*, *well-articulated*, *well-run*, etc.), de tal manera que el corpus de muestra se registraron 154 compuestos con *well* como elemento inicial, mientras que en el electrónico se contabilizaron 92. Su antónimo (*badly*), por otra parte, rara vez aparece en adjetivos compuestos. Esto lo corrobora el hecho de que en el corpus de muestra se verificaron tan solo cuatro ocurrencias (*badly-dyed*, *badly-behaved*, *badly-handled*, *badly-played*) y en el corpus electrónico solamente una (*badly-damaged*). Los hablantes de inglés prefieren usar el adverbio *ill* (38 y 16 instancias, respectivamente) el que, con frecuencia, forma pares opuestos con *well*, como lo comprueban los siguientes ejemplos:

well-bred / ill-bred

well-timed / ill-timed

well-formed / ill-formed

well-defined / ill-defined

well-dressed / ill-dressed

well-conceived / ill-conceived

well-advised / ill- advised

well-founded / ill-founded

well-adjusted / ill-adjusted

well-equipped / ill-equipped

well-prepared / ill-prepared

well-informed / ill- informed

Por lejos el adverbio de tiempo más productivo es *long* con 49 ocurrencias. Pero, la doble funcionalidad de *long* como adjetivo y como adverbio representa una dificultad al momento de clasificar los datos. En consecuencia, se requiere una revisión manual cuidadosa y rigurosa para asegurar la correcta diferenciación entre patrones:

Tabla 60. Patrones morfológicos en que participa *-long*

| Adverbio + participio pasado | Adjetivo + N- <i>ed</i> |
|------------------------------|-------------------------|
| <i>long-held</i> | <i>long-eared</i> |
| <i>long-exiled</i> | <i>long-nosed</i> |
| <i>long-awaited</i> | <i>long-haired</i> |
| <i>long-promised</i> | <i>long-legged</i> |
| <i>long-forgotten</i> | <i>long-bodied</i> |
| <i>long-established</i> | <i>long-shaped</i> |

El segundo adverbio de tiempo más productivo es *newly* con 31 ocurrencias (*newly-passed*, *newly-released*, *newly-constructed*). Sin embargo, en la creación de adjetivos compuestos muchas veces se prefiere usar una variante más corta. El diccionario *Collins* en línea describe *new* como adverbio en combinaciones con participios como en *new-laid eggs* en los que *new* es sinónimo de *recently* y *freshly*. Esto se ve corroborado por la presencia de pares como *newly-born* / *new-born* (recién nacido). En los corpus analizados se registraron 9 instancias en que *new* opera como adverbio (*new-spilled (blood)*, *new-found*, *new-formed*) pero, a diferencia de lo que sucede con *long*, la dualidad en este caso es meramente formal y no funcional ya que no se encontraron compuestos en los que *new* actúa como adjetivo. El problema se vuelve a presentar con *freshly* que tiene 14 ocurrencias (*freshly-cleaned*, *freshly-opened*, *freshly-glued*), y su variante *fresh* que registra 15 (*fresh-killed*, *fresh-baked (bread)*, *fresh-made*) puesto que, además de encontrar pares como *freshly-brewed* / *fresh-brewed* y *freshly-waxed* / *fresh-waxed*, se verificó una instancia en que *fresh* funciona como

adjetivo al interior de un compuesto de base nominal y por ende conforma un patrón morfológico distinto:

Tabla 61. Patrones morfológicos en que participa *-fresh*

| Adverbio + participio pasado | Adjetivo + N- <i>ed</i> |
|------------------------------|-------------------------|
| <i>fresh-cooked</i> | <i>fresh-faced</i> |

Otros adverbios de tiempo comúnmente usados en la creación de adjetivos compuestos son *recently* con 18 ocurrencias (*recently-reported*, *recently-removed*, *recently-appointed*), *now* con 17 (*now-ended*, *now-restricted*, *now-flooded*), *often* y su variante *oft* con 13 y 14 casos, respectivamente (*often-performed*, *often-confused*, *often-asked*, *oft-stated*, *oft-identified*, *oft-vilified*), y *already* con 11 (*already-determined*, *already-known*, *already-made*).

Menos frecuentes son *rarely* (*rarely-heard*, *rarely-visited*, *rarely-mentioned*), *previously* (*previously-suggested*, *previously-announced*, *previously-owned*), *just* (*just-cleaned*, *just-poured*, *just-mentioned*), *seldom* (*seldom-discussed*, *seldom-smelled*, *seldom-accessed*) y *never* (*never-ratified*, *never-opened* (books), *never-solved*).

En lo que respecta a adverbios de cantidad, *much* lidera en productividad con 55 instancias, mientras que *little* solo registró ocho:

much-cited
much-liked
much-feared
much-quoted
much-visited
much-needed

little-played
little-known
little-expected
little-discussed
little-mentioned
little-frequented

much-debated
much-traveled
much-preferred
much-criticized
much-welcomed
much-contended
much-acclaimed
much-anticipated
much-appreciated

little-debated
little-understood

Entre los adverbios de grado *half* resulta ser particularmente productivo, con 72 ocurrencias en los corpus. Sin embargo, una vez más es importante tener presente que existen unos pocos compuestos de base nominal cuyo primer constituyente es el cuantificador *half* (cf. 8.1.7.), lo que implica un patrón diferente:

Tabla 62. Patrones morfológicos en que participa *-half*

| Adverbio + participio pasado | Cuantificador + N- <i>ed</i> |
|--------------------------------|------------------------------|
| <i>half-created</i> | <i>half-hearted</i> |
| <i>half-abandoned</i> | <i>half-witted</i> |
| <i>half-forgotten (dreams)</i> | <i>half-lidded (glance)</i> |

El adverbio *highly* también resultó ser bastante productivo con 26 instancias. Por su parte, la variante *high* registró 14 ocurrencias. En este último caso, al igual como sucede con *long* y *fresh*, es necesario distinguir entre el adverbio y el adjetivo *high* para determinar el patrón morfológico correcto:

Tabla 63. Patrones morfológicos en que participa *-high*

| Adverbio + participio pasado | | Adjetivo + N- <i>ed</i> |
|------------------------------|------------------------|-------------------------------|
| <i>highly-skilled</i> | <i>high-valued</i> | <i>high-sided</i> |
| <i>highly-qualified</i> | <i>high-placed</i> | <i>high-nosed</i> |
| <i>highly-produced</i> | <i>high-ranked</i> | <i>high-spirited</i> |
| <i>highly-developed</i> | <i>high-demanded</i> | <i>high-ceilinged (hotel)</i> |
| <i>highly-organized</i> | <i>high-distressed</i> | <i>high-terraced (villa)</i> |
| <i>highly-populated</i> | | |
| <i>highly-anticipated</i> | | |
| <i>highly-recommended</i> | | |

En cuanto a adverbios de lugar, *above* y *below* registraron 9 y dos ocurrencias, respectivamente:

above-cited

below-signed

above-mentioned

below-mentioned

above-specified

above-discussed

above-formulated

8.2.4. Adverbio comparativo / superlativo + participio pasado

Al igual como sucede con los adjetivos, las formas comparativas y superlativas de algunos adverbios pueden utilizarse en la creación de unidades compuestas:

Tabla 64. Adverbios comparativos y superlativos en compuestos con participio pasado

| Adv comparativo + participio pasado | Adv superlative + participio pasado |
|-------------------------------------|-------------------------------------|
| <i>better-dressed</i> | <i>best-dressed</i> |
| <i>better-known</i> | <i>best-known</i> |
| <i>better-lined</i> | <i>best-paid (models)</i> |
| <i>better-educated</i> | <i>best-laid (plans)</i> |
| <i>lesser-known (sister)</i> | <i>best-loved (poems)</i> |
| <i>lower-born</i> | <i>earliest-known (hominid)</i> |
| | <i>highest-rated</i> |
| | <i>hardest-learned</i> |
| | <i>lowest-placed</i> |
| | <i>least-acknowledged (fears)</i> |
| | <i>worst-reviewed (movie)</i> |
| | <i>most-imitated</i> |

8.3. Adjetivos compuestos del tipo *-ing*

Los adjetivos cuyo núcleo es un participio presente (*controlling, ending, looking*) pueden incorporar un sustantivo, un adverbio o un adjetivo como modificador para formar compuestos como *mind-controlling, never-ending* y *funny-looking*. A continuación se describen los patrones morfológicos encontrados en los corpus analizados:

8.3.1. Adjetivo + participio presente

Este patrón es relativamente productivo dado que se circunscribe a un reducido grupo de verbos como elemento central. Estos compuestos toman verbos copulativos (*linking verbs*) como núcleo, específicamente aquellos que se refieren a los sentidos. De ellos, *looking* es el más prolífico, seguido muy a distancia de *sounding, smelling* y *tasting*. De hecho, en el corpus de muestra, de 190

ejemplos, 161 tienen *looking* como base, mientras que *seeming*, un casi sinónimo, registró tan solo tres (*safe-seeming (city)*, *genuine-seeming*, *dodgy-seeming*). En la siguiente tabla se ejemplifica la proporción:

Tabla 65. Compuestos con verbos copulativos

| | | | |
|-----------------------------|------------------------------|--------------------------|------------------------------|
| <i>angelic-looking</i> | <i>aristocratic-sounding</i> | <i>bad-smelling</i> | <i>acid-tasting (water)</i> |
| <i>bitter-looking</i> | <i>cool-sounding</i> | <i>damp-smelling</i> | <i>clean-tasting (broth)</i> |
| <i>capable-looking</i> | <i>crappy-sounding</i> | <i>evil-smelling</i> | <i>sour-tasting</i> |
| <i>delicate-looking</i> | <i>darn-sounding (idea)</i> | <i>exotic-smelling</i> | <i>sweet-tasting</i> |
| <i>evil-looking (beast)</i> | <i>desperate-sounding</i> | <i>foul-smelling</i> | |
| <i>foreign-looking</i> | <i>high-sounding</i> | <i>fragrant-smelling</i> | |
| <i>good-looking</i> | <i>low-sounding</i> | <i>friendly-smelling</i> | |
| <i>healthy-looking</i> | <i>hollow-sounding</i> | <i>rancid-smelling</i> | |
| <i>innocent-looking</i> | <i>innocent-sounding</i> | <i>savoury-smelling</i> | |
| <i>jittery-looking</i> | <i>loose-sounding</i> | <i>smoky-smelling</i> | |
| <i>lifeless-looking</i> | <i>nice-sounding</i> | <i>strong-smelling</i> | |
| <i>motherly-looking</i> | <i>similar-sounding</i> | | |
| <i>natural-looking</i> | <i>tired-sounding</i> | | |
| <i>odd-looking</i> | | | |
| <i>painful-looking</i> | | | |
| <i>queer-looking</i> | | | |
| <i>refined-looking</i> | | | |
| <i>sad-looking</i> | | | |
| <i>tired-looking</i> | | | |
| <i>unusual-looking</i> | | | |
| <i>violent-looking</i> | | | |
| <i>warm-looking</i> | | | |
| <i>young-looking</i> | | | |

En la revisión de los corpus se identificó la ocurrencia de una forma alternativa a un compuesto listado arriba. Es el caso de *naturally-looking* (adverbio + participio presente) que registra solo 10 ocurrencias en el corpus *enTenTen12* mientras que el compuesto estándar *natural-looking* (adj + participio presente) registra 2.910. Probablemente, *naturally-looking* sea una forma emergente creada por analogía con otras estructuras compuestas. Entre los compuestos terminados en *-ed*, también es posible encontrar pares de compuestos similares en significado pero diferentes en estructura como *well-sized* y su contraparte *good-sized* o *well-intentioned* y *good-intentioned*.

8.3.1.1. Participios adjetivos + participio presente

De igual manera como ocurre en la categoría de adjetivos terminados en *-ed*, la función adjetival de los participios quedó de manifiesto en un grupo de compuestos cuyo primer elemento es un adjetivo participio (*participle adjective*). Estos solo pueden ocurrir con verbos copulativos:

promising-looking

fascinating-sounding

terrifying-looking

astounding-looking

exciting-sounding

revolting-looking

intriguing-looking

threatening-looking

appealing-looking

tempting-looking

8.3.2. Adjetivo comparativo / superlativo + participio presente

El primer elemento del compuesto puede ser una forma comparativa o superlativa del adjetivo que funciona como especificador, siendo considerablemente más frecuentes los superlativos, como se puede ver en esta tabla:

Tabla 66. Adjetivos comparativos y superlativos en compuestos con participio presente

| Adj comparativo + participio presente | Adj superlativo + participio presente |
|---------------------------------------|---------------------------------------|
| <i>better-looking</i> | <i>best-looking (prospects)</i> |
| <i>lighter-looking (windows)</i> | <i>roughest-looking (men)</i> |
| <i>nastier-looking (man)</i> | <i>sternest-seeming (stoic)</i> |
| <i>younger-looking (breasts)</i> | <i>youngest-looking</i> |
| <i>prettier-looking</i> | <i>fittest-looking</i> |
| <i>fuller-looking</i> | <i>oddest-looking</i> |
| <i>newer-looking</i> | <i>nicest-looking</i> |
| <i>stronger-looking</i> | <i>richest-looking</i> |
| | <i>worst-looking</i> |
| | <i>weirdest-looking</i> |
| | <i>best-smelling</i> |
| | <i>cleanest-smelling</i> |
| | <i>best-sounding</i> |
| | <i>nicest-sounding</i> |
| | <i>dullest-sounding</i> |
| | <i>fullest-sounding</i> |
| | <i>phoniest-sounding</i> |
| | <i>warmest-sounding</i> |
| | <i>sweetest-sounding</i> |
| | <i>best-tasting (tequila)</i> |
| | <i>greatest-tasting</i> |

8.3.3. Sustantivo + participio presente

Al igual como sucede con los compuestos terminados en *-ed*, los adjetivos cuyo primer constituyente es un sustantivo también resultó ser el patrón morfológico más frecuente. La

diferencia estriba en que el sustantivo funciona principalmente como complemento directo de la base verbal y no como adjunto:

life-giving (elixir)

advice-dispensing (trio)

crime-fighting (organization)

energy-producing (part of the cells)

risk-taking

party-loving

heart-stopping

French-speaking

La función de adjunto del primer constituyente del compuesto es muy menor y se restringe a verbos que necesariamente deben ir acompañados de un sustantivo que denota ubicación (*cinema-going, time-travelling, city-dwelling*), medio (*stilt-walking (sorcerer), plane-flying, radio-transmitting*) o tiempo (*night-wandering, day-walking*).

En cuanto a la expresión de modo o manera, ésta se circunscribe a compuestos creados a partir de los verbos *look, seem, sound, smell* y *taste*, en los que se establece una comparación con el sustantivo inicial. Esta función semántica es bastante acotada dentro del patrón morfológico por lo que en el corpus de muestra se registraron solo siete ocurrencias:

bronze-age-looking

cast-iron-looking (brogans)

earth-smelling (passage)

marble-seeming (features)

hay-smelling (fields)

oil-smelling (engine room)

rut-tasting (water)

Por su parte, en la muestra tomada del corpus electrónico (4.279 compuestos) se encontraron 22 ejemplos y de ellos 21 contienen *-looking* como base (*crochet-looking, diamond-looking, vampire-looking*). El otro compuesto registrado es *sulfur-smelling*. Cabe destacar que, en su modelo de generación de adjetivos compuestos, Roeper y Siegel (1978) excluyen la comparación.

Al analizar algunos compuestos nos encontramos con matices semánticos que vale la pena mencionar. Una excepción al análisis anterior la constituye *future-looking* dado que el sustantivo funciona como un adjunto que indica dirección (que ve hacia el futuro) y no comparación. Por otra parte, hay que distinguirlo de *futuristic-looking* que describe apariencia y además corresponde a otro patrón morfológico:

Tabla 67. Matices semánticos y sintácticos

| Sustantivo + participio presente | Adjetivo + participio presente |
|----------------------------------|--------------------------------|
| <i>future-looking</i> | <i>futuristic-looking</i> |

8.3.4. Adverbio + participio presente

Este resultó ser uno de los tres patrones más productivos en este tipo de adjetivos compuestos. El adverbio puede funcionar como adjunto de modo o manera (*hard-working, fast-thinking, well-functioning*), dirección (*backward-turning, down-going, upward-pointing*), tiempo (*long-suffering, ever-increasing, never-ending*) y cantidad (*much-receiving, half-singing*). En este patrón morféxico, los adverbios más productivos en ambos corpus son los siguientes:

Tabla 68. Adverbios más productivos en compuestos con *-ing*

| Adverbio | Ocurrencias | Ejemplos |
|-------------|-------------|---|
| <i>ever</i> | 118 | <i>ever-developing, ever-struggling, ever-worsening</i> |
| <i>fast</i> | 37 | <i>fast-falling, fast-selling, fast-expanding</i> |
| <i>high</i> | 32 | <i>high-paying, high-ranking, high-flying</i> |
| <i>slow</i> | 30 | <i>slow-burning, slow-sinking, slow-moving</i> |
| <i>long</i> | 29 | <i>long-ruling, long-lasting, long-surviving</i> |
| <i>low</i> | 27 | <i>low-lying, low-hanging, low-rolling</i> |
| <i>well</i> | 20 | <i>well-paying, well-judging, well-looking</i> |

A diferencia de lo que ocurre con la otra categoría (-ed) en la que la naturaleza verbal o nominal del núcleo determina la categoría gramatical de palabras bifuncionales como *long*, *fresh* y *high* (cf. 8.2.3.), esto es más complejo de resolver en el caso de los compuestos terminados en -ing. Por ejemplo, para entender la relación sintáctica existente entre los componentes de *down-looking* es necesario recurrir al contexto. Una vez que logramos una adecuada interpretación semántica podemos decidir si el primer elemento es un adverbio o un adjetivo y, solo entonces, es posible asignar el patrón morfológico correcto:

Tabla 69. Ambigüedad semántica y sintáctica

| Adverbio + participio presente | Adjetivo + participio presente |
|--|---|
| <i>down-looking = downward-looking</i> | <i>down-looking = depressed-looking</i> |

Los adverbios de cantidad no resultan muy productivos en este patrón y se circunscriben a *half* con 10 (*half-smiling*, *half-dreaming*, *half-joking*) y *much* con dos (*much-receiving*, *much-enduring*).

Cabe hacer notar que en este patrón morfológico, *all* no funciona como cuantificador sino como alternativa a *everything*, por ende como sustantivo (*all-knowing*, *all-embracing*, *all-consuming*), registrando 44 ocurrencias en los corpus analizados.

8.3.5. Adverbio comparativo / superlativo + participio presente

Las formas comparativas y superlativas de los adverbios también pueden ser parte de un adjetivo compuesto tal como se observa en los pares *well-functioning* / *better-functioning* y *well-paying* / *best-paying*. En los corpus se encontraron 40 ejemplos de comparativos y 83 de superlativos. Estos son algunos de ellos:

Tabla 70. Adverbios comparativos y superlativos en compuestos con participio presente

| Adv comparativo + participio presente | Adv superlativo + participio presente |
|---------------------------------------|---------------------------------------|
| <i>better-performing</i> | <i>worst-performing</i> |
| <i>better-fitting</i> | <i>best-paying</i> |
| <i>cleaner-burning (fuel)</i> | <i>lowest-selling</i> |
| <i>healthier-eating</i> | <i>fastest-growing</i> |
| <i>higher-scoring</i> | <i>worst-polluting</i> |
| <i>lower-growing</i> | <i>oldest-running</i> |
| <i>faster-buying</i> | <i>longest-reigning</i> |
| <i>lesser-ranking</i> | <i>best-selling</i> |
| <i>longer-serving</i> | <i>shortest-serving</i> |
| <i>lower-yielding</i> | <i>quickest-adapting</i> |
| <i>longer-lasting</i> | <i>fastest-spreading</i> |
| <i>higher-earning</i> | <i>lowest-earning</i> |
| <i>slower-moving</i> | <i>slowest-moving</i> |
| <i>faster-growing</i> | <i>fastest-growing</i> |
| <i>lower-melting</i> | <i>largest-selling</i> |
| <i>earlier-ripening</i> | <i>biggest-spending</i> |
| <i>lower-lying</i> | <i>highest-qualifying</i> |
| <i>longer-playing</i> | <i>longest-surviving</i> |

8.3.6. Número + participio presente

Este patrón morfológico es poco productivo, registrando 13 ocurrencias en los corpus analizados y todas ellas incorporan un número ordinal: *second-leading* (que lidera segundo/a), *third-ranking* (en tercera posición), *eleventh-leading* (que lidera undécimo/a).

8.4. Singularidades

“Linguistic idiosyncracies can be defined as knowledge that departs from rules.” (Roepers y Siegel, 1978: 215). Entre los muchos compuestos analizados para este estudio, se identificaron algunos que se distinguen por mostrar algún rasgo morfológico o semántico que los distingue. A continuación listamos algunos de ellos.

8.4.1. Pronombre como primer elemento

Se registró la ocurrencia del compuesto *me-shaped* cuyo primer constituyente es un pronombre, una categoría gramatical inusual en este tipo de estructuras. Además, su interpretación varía en literalidad dependiendo de los contextos en los que ocurre:

[...] *a me-shaped hole in the wall* (un hoyo en la pared con mi forma)

[...] *there is also a 'me-shaped' space in God's plan* (hay un espacio para mí en los planes de Dios)

8.4.2. Compuestos con *-shaped*

Los compuestos que tienen *-shaped* como núcleo además de poder tener distintas interpretaciones, también pueden adscribir a diferentes patrones morfológicos. Es así que, al analizar los muchos ejemplos registrados en los corpus, notamos que *-shaped* tiene características de N-*ed*, es decir, de base nominal en algunos compuestos y de participio pasado en otros. Cuando el primer elemento es un adjetivo identificamos el patrón Adj + N-*ed*, atribuyéndole el significado *with X shape*:

familiar-shaped

funny-shaped

long-shaped

flat-shaped

conical-shaped

contemporary-shaped

Estos compuestos especifican qué tipo de forma tiene algo (conocida, graciosa, alargada, plana, cónica, contemporánea, etc.). Sin embargo, en compuestos con un sustantivo inicial, asignamos el patrón N + V-*ed*, basados en la interpretación *shaped like X*:

woman-shaped

lagoon-shaped

apple-shaped

triangle-shaped

boomerang-shaped

bird-shaped

Para complicar más las cosas, también está la ambigüedad semántica de *God-shaped*:

[...] *He put a **God-shaped** space in the hearts of each that no substitute would be able to fill.*

[...] *followers of God will live into a **God-shaped** reality.*

Además, en los corpus se constató la ocurrencia de varios adjetivos compuestos que incorporan una letra como constituyente inicial en el patrón N + V-*ed*. La mayoría de ellos tienen *-shaped* como núcleo:

X-shaped

C-rated

A-oriented

L-shaped

M-rated

H-terminated

U-shaped

G-rated

X-rayed

V-shaped

X-rated (film)

L-formed

J-shaped

X-carrying

A-shaped

β-blocking

M-shaped

T-shaped

Z-shaped

8.4.3. Elementos iniciales coordinados

En los corpus también se encontraron varios ejemplos en los que dos sustantivos coordinados ocupan la posición no nuclear del compuesto.

woman-and-child-centered birth

cream-and-rose-colored clouds

jazz-and-gospel-inflected original song

shingle-and-granite-faced house

volcano- and glacier-shaped landscape

earthquake-and tsunami-triggered Fukushima Daiichi disaster

heart-and-sky-piercing coordination

up-and-down-moving hand

gas- or diesel-powered products

acorn-, or pumpkin-shaped

Aun cuando el primer elemento ha sido expandido, la estructura de la unidad sigue intacta. Sin embargo, es difícil decidir qué tipo de expansión puede considerarse aceptable dentro de los límites impuestos por la morfología de la palabra. Un compuesto como *never-before-published images* nos parece perfectamente aceptable, mientras que otros como *an ever-more-rarefied aesthetic* o *the many-weeks-deserted parlour* (E. Brontë) plantean preguntas interesantes sobre los límites entre composición y sintaxis: Consideremos el siguiente ejemplo:

Timothy Leary's Tibetan Book of the Dead-inspired writings

Este caso puede parecer complejo dado que el primer constituyente es un nombre compuesto que incluye una frase genitiva como post-modificador. No obstante, teniendo en cuenta que al definir qué es un compuesto, varios autores incluyen las frases como posibles elementos (cf. sección 2.3.1.), la incorporación de un componente de naturaleza sintáctica no debiera afectar su estatus de unidad

compuesta. Sin embargo, en las descripciones no se especifican ni la longitud ni el tipo de frase que puede ocupar esa posición.

8.4.4. Compuestos creados por analogía

Algunos compuestos han sido modelados siguiendo el patrón de otros, dando origen a vocablos nuevos, de forma y uso no canónicos. A modo de ejemplo, el adjetivo *little-known* ha producido dos símiles: *small-known* y *tiny-known*, con 13 y 8 ocurrencias en el corpus electrónico, respectivamente. Lo que resulta interesante es que estos noveles sinónimos reflejan la estructura “aparente” del compuesto original (Adj + V) cuando en realidad *little-known* se ajusta a un patrón diferente (Adv + V). Para ilustrar su uso, hemos seleccionado dos ejemplos del corpus *enTenTen12*.

[...] *Nathan Green, a **small-known** Australian*

[...] *It is a **tiny-known** fact*

Como se puede observar en los ejemplos, ambos funcionan de igual manera y en contextos similares que *little-known*. Sin embargo, al no ser formas estándar, no aparecen listadas en ningún diccionario ni arrojan ningún resultado en *Google*.

8.4.5. Compuestos que eliden sustantivo

En los corpus se registraron algunos casos de compuestos cuyo primer elemento es una frase nominal en la que se ha omitido el sustantivo. Al sobreentenderse el nombre, el adjetivo asume mayor peso semántico por lo que un compuesto como *immune system-stimulating* (10 ocurrencias en el corpus *enTenTen12*) puede reducirse simplemente a *immune-stimulating* (180 ocurrencias). Esta tendencia también se observa en adjetivos como:

| | |
|---------------------------------|---|
| <i>atomic-driven</i> | (<i>atomic energy</i>) |
| <i>sensory-seeking</i> | (<i>sensory stimulation</i>) |
| <i>reproductive-boosting</i> | (<i>reproductive system</i>) |
| <i>endocrine-disrupting</i> | (<i>endocrine system</i>) |
| <i>polar-orbiting satellite</i> | (<i>polar region</i>) |
| <i>adipose-derived</i> | (<i>adipose tissue</i>) |
| <i>Islamic-inspired</i> | (<i>Islamic culture</i>) |
| <i>Edwardian-inspired</i> | (<i>Edwardian period-inspired</i>) |
| <i>Georgian-set</i> | (<i>set in the Georgian period/era</i>) |

Todos estos compuestos se ajustan al patrón Adj + V en una combinación de lexemas no esperable. La excepción parece ser *X-carrying*, en el que *X* puede ser interpretada como un nombre (*X chromosome*).

9. Conclusiones

A grandes rasgos, en este estudio hemos querido abordar un fenómeno tan complejo y extenso como la composición a través de un análisis morfológico y funcional de la clase de adjetivos compuestos con participio. A continuación detallamos algunos de los hallazgos más relevantes en distintos aspectos de la investigación:

9.1. Frecuencia de ocurrencia de patrones morfológicos en los corpus

Inicialmente, las observaciones realizadas en el corpus de prueba nos llevaron a hipotetizar sobre la posibilidad de encontrar una correspondencia entre los patrones morfológicos encontrados en ese corpus y las muestras extraídas del corpus *enTenTen12*. En lo que respecta a los tipos de patrones se comprobó que son los mismos, a excepción de Núm + *V-ed* cuya ocurrencia solo se verificó en el corpus de análisis con 11 ejemplos. Sin embargo, también comprobamos que en algunos casos los porcentajes varían considerablemente. La diferencia más importante guarda relación con los porcentajes de compuestos de base nominal (*N-ed*) presentes en el corpus de prueba y en el electrónico. En el primero de ellos representan casi la mitad con 47,1 % mientras que en el segundo solo alcanzan 17,6 %. Esto puede deberse a principalmente al tipo de textos consultados. En el caso del corpus de prueba, la mayoría de los datos fueron recopilados de textos literarios mientras que en el corpus electrónico las fuentes son mucho más diversas. Esto puede resultar en cierta preferencia o proclividad a usar ciertas estructuras compuestas y no otras. Es así que en algunos textos de corte descriptivo predominan los adjetivos con núcleo nominalizado. Por ejemplo, las revistas femeninas dedican secciones a describir atuendos y para ello se sirven de estos compuestos, los que son muy comunes en la descripción de vestimentas (cf. sección 8.1.3.).

Por otra parte, las críticas cinematográficas también se valen de estos compuestos en sus descripciones, generalmente agrupados en secuencias de patrones morféimicos similares, lo que incide directamente en la cantidad de ejemplos compilados:

N + N-*ed*:

- 1) *'There's always going to be somebody spitting blood about my **wooden-faced toffee-named** crappy acting.'* (Entrevista a Benedict Cumberbatch, Vanity Fair.com, 2014)
- 2) *"Russell Crowe's **black-hatted, black-hearted** outlaw Ben Wade is the sort of rotter who can murder a defenceless man in his sleep."* (Empireonline)
- 3) *"**Milk-skinned, golden-haired** and young enough to be his granddaughter," 'Blood and Beauty: The Borgias,' by Sarah Dunant (NYtimes.com July 5, 2013).*

N + V-*ing*:

*"Make no mistake: for every second that Dennis Hopper's unforgettable, **oxygen-huffing, Heineken-disparaging, ear-severing** brute is on screen, Blue Velvet is a horror movie".*
(Empireonline)

*"A **milk-drinking, plant-tending, Gene Kelly-loving** Sicilian hitman in New York who adopts a young orphan after her family is murdered."* (Empireonline)

N + V-*ed*:

*"And so another spin about the Brontëan moors: **windswept, rain-sodden, winter-chilled**".*
(Empireonline)

Si bien estos ejemplos pueden ayudarnos a entender las diferencias numéricas entre los corpus analizados, solo mediante estudios comparativos acotados a ciertos géneros específicos (académicos, científicos, literarios, periodísticos, etc.) se podrá determinar si la frecuencia de algunos patrones morfológicos tiene que ver con los tipos de textos en los que ocurren.

Otro factor a tener en cuenta es el tamaño del corpus. En lo que se refiere a los compuestos terminados en *-ed*, los dos corpus son bastante similares en volumen de datos con 3.983 en el corpus

de prueba y 4.808 en la muestra sacada del corpus electrónico. No sucede lo mismo con los compuestos terminados en *-ing* con 1.057 y 4.279, respectivamente. Esta desigualdad, sin embargo, no altera el hecho de que los patrones siguen siendo los mismos. Lo mismo ocurrió con la muestra de 1.000 compuestos irregulares, lo que demuestra una regularidad morfológica que no depende de la cantidad de datos manejados.

9.2. Identificación de patrones morfológicos

Como hemos visto, asignar un patrón morféxico a un compuesto implica no solo determinar la categoría gramatical de los elementos que lo componen; también es necesario entender la relación que existe entre ellos. A primera vista, muchos de los ejemplos recopilados en los corpus parecen corresponder a determinado patrón, pero un análisis más detallado pone de manifiesto aspectos fundamentales para su correcta identificación. La primera distinción que hay que tener presente es la que concierne a la naturaleza del compuesto. Distinguir entre una base nominal (*N-ed*) y una verbal (*V-ed*) nos ayuda a entender la función que cumple el primer elemento respecto al núcleo. Por ejemplo, en el estudio, hemos recalcado la importancia que tienen distinciones como adverbio / adjetivo en compuestos que involucran formas como *long*, *hard*, *fast* y *little* en pares de compuestos como:

long-married / long-legged

hard-fought / hard-hearted

fresh-baked / fresh-faced

high-paid y *high-heeled*

En estos ejemplos, la naturaleza de la base (*V-ed* o *N-ed*) ayuda a clarificar la función que cumple el constituyente inicial, lo que también nos permitirá identificar correctamente los adverbios contenidos en compuestos del tipo *new-formed*, *high-valued* y *quick-thinking* evitando confundirlos con adjetivos, como sucedió frecuentemente al utilizar el etiquetador sintáctico de *Sketch Engine*.

9.3. Ambigüedad morfológica y semántica

Aun cuando habiendo identificado los patrones morféimicos, a veces un compuesto puede mostrar ambigüedad estructural como ocurre con los adjetivos que poseen *-shaped* como núcleo como *funny-shaped* y *bird-shaped* (cf. 8.4.2.). Otros ejemplos son:

sunset-colored scarf N + N-*ed* (un pañuelo con los colores del atardecer)

sunset-colored waters N + V-*ed* (aguas coloreadas por la luz del atardecer)

Cuando nos encontramos con el compuesto *Normal-made* supusimos que era el opuesto de *special-made* (Adj + V-*ed*). Sin embargo, *Normal* resultó ser el nombre de una ciudad norteamericana, por lo tanto el patrón correcto es N + V-*ed*. En otros casos la ambigüedad es semántica: *Mango-based* (*Mango* es un sistema operativo de teléfonos celulares) y *mango-based salsa* (fruta).

En todo caso, la ambigüedad no es ajena a las estructuras compuestas, de hecho una de las características que las definen es que poseen significados abiertos (*openness of meaning*). Esto implica que sus componentes se relacionan de manera “libre” o espontánea. Especialistas como Olsen (2014) consideran que este es el criterio fundamental para distinguir compuestos de derivaciones. Estamos de acuerdo en que *blue-jeaned*, *eating-disordered* y *cling-filmed* no cumplen con este criterio puesto que son formas derivadas de sustantivos compuestos pre-existentes (*blue jeans*, *eating disorder* y *cling-film*). En las muestras tomadas del corpus de análisis aparecían muchos términos de este tipo los que debieron ser identificados y eliminados manualmente dado que es una labor que sobrepasa las capacidades de cualquier etiquetador sintáctico. De igual forma, cuando se analizan miles de datos se vuelve una tarea muy difícil de abordar para el analista.

9.4. Consideraciones finales

En esta investigación intentamos entregar una descripción lo más detallada posible de una categoría de palabras que son parte de un fenómeno particularmente productivo y extendido. Aun cuando existen varias clasificaciones de estas formas compuestas, creemos que su descripción no se ha

agotado puesto que, debido a su gran versatilidad y productividad, continúan surgiendo nuevas combinaciones que no siempre corresponden a los patrones ya establecidos o conocidos. Además de cuantificar la ocurrencia de los distintos patrones morfológicos, queríamos dar cuenta de las regularidades de forma y uso comunicativo de los adjetivos compuestos con participio y registrar algunas de sus particularidades (*idiosyncrasies*). En la mayoría de los casos, observamos que su conformación obedece a reglas morfológicas bien definidas: las principales categorías gramaticales que participan de su creación son sustantivos, adjetivos y adverbios en sus distintas realizaciones. De tal suerte que, entre la gran variedad de compuestos están aquellos cuyo primer elemento es un nombre propio (*Sinatra-loving, Obama-endorsed*), acrónimos (*CDA-oriented, US-financed*), palabras de origen hispano (*tortilla-making, piñata-bursting*), adjetivos comparativos y superlativos (*bigger-named, best-dressed*). Agregamos a la descripción los patrones que incluyen cuantificadores (*much-quoted, many-coloured*) y numerales (*one-armed, three-headed*). Entregamos amplias evidencias de la interacción de la composición con los otros procesos de formación de palabras y consigo misma (recursividad). Son precisamente la gran productividad y diversidad desplegadas por estas unidades compuestas las que hacen que relativamente pocas de ellas estén lexicalizadas ya que por su número no pueden ser incluídas en los diccionarios.

Por último, queremos destacar que *Sketch Engine*, como herramienta de análisis textual, nos permitió extraer las muestras y compilar nuestros propios corpus. Sin embargo, como quedó demostrado en el procesamiento de los datos, no mostró mucha exactitud al etiquetar los constituyentes de los compuestos, sobre todo cuando las búsquedas incluían ambos constituyentes. Al analizar el corpus de unidades terminadas en *-ing*, un gran número de datos no fue clasificado, así que enfocamos las consultas en el primer elemento, lo que redujo este porcentaje considerablemente. Desafortunadamente, lo que no varió fue la cantidad de errores que involucraban formas similares, especialmente adjetivos y adverbios en sus diferentes realizaciones. En el estado de cosas actual, una investigación de esta naturaleza implica muchísimas horas de trabajo revisando y depurando los datos recolectados y, posteriormente, verificando y reclasificando los resultados arrojados en las consultas individuales. Esto pone de manifiesto la necesidad de incorporar un mayor y mejor componente morfológico en los etiquetadores de este tipo.

Bibliografía

ADAMS, Valerie, 2013: *Complex words in English*. London & New York: Routledge.

ARCODIA, Giorgio F., Nicola GRANDI y Fabio MONTERMINI, 2009: "Hierarchical NN compounds in a crosslinguistic perspective", *Italian Journal of Linguistics / Rivista di Linguistica* 21(1), 11-33.

BAKER, Paul, Andrew HARDIE y Tony MCENERY, 2006: *A glossary of corpus linguistics*. Edinburgh: Edingburgh University Press.

BAUER, Laurie, 1983: *English word-formation*. Cambridge: Cambridge University Press.

BAUER, Laurie, 1998: "Is there a class of neoclassical compounds, and if so is it productive?", *Linguistics* 36, 403-422.

BAUER, Laurie, 2004: "The function of word-formation and the inflection-derivation distinction" en Henk AERTSEN, Mike HANNAY y Rod LYALL (eds.): *Words and their Places. A Festschrift for J. Lachlan Mackenzie*. Amsterdam: Vrije Universiteit, 283-292.

BAUER, Laurie, 2005: "Productivity: Theories" en Pavol ŠTEKAUER y Rochelle LIEBER (eds.): *Handbook of word-formation*. Amsterdam: Springer, 315-334.

BAUER, Laurie, 2006: "Compounds and Minor Word-Formation Types" en Bas AARTS y April MCMAHON (eds.): *The handbook of English linguistics*. Oxford: Blackwell, 483-506.

BAUER, Laurie, Rochelle LIEBER e Ingo PLAG, 2013: *The Oxford reference guide to English morphology*. Oxford: Oxford University Press.

BIBER, Douglas, 2010: "Corpus-based and corpus-driven analyses of language variation and use" en Bernd HEINE y Heiko NARROG (eds.): *The Oxford handbook of linguistic analysis*. Oxford: Oxford University Press, 159-191.

BISETTO, Antonietta y Sergio SCALISE, 2005: "The classification of compounds", *Lingue e linguaggio* 4 (2), 319-332.

BLEVINS, James, 2006: "English inflection and derivation" en Bas AARTS y April MCMAHON (eds.): *The handbook of English linguistics*. Oxford: Blackwell, 507-536.

BLOOMFIELD, Leonard, 1933: *Language*. New York, Holt.

BOOIJ, Geert, 2006: "Inflection and derivation" en Keith BROWN (ed.): *Encyclopedia of language & linguistics*, volumen 5. Amsterdam: Elsevier, 654-661.

CASTAIRS-MCCARTHY, Andrew, 2002: *An Introduction to English Morphology*. Edinburgh: Edingburgh University Press.

CONTI, Sara, 2007: *Compound Adjectives in English. A descriptive approach to their morphology and functions*. Tesis doctoral. Universidad de Pisa (co-tutela con la Universidad de Viena).

CROCCO-GALÈAS, Grazia, 2003: "Compound adjectives in English: the type lion-hearted and good-natured", *Studi Italiani di Linguistica Teorica e Applicata*. Anno XXXII, número I, 31-43.

DOWNING, Angela, 2015: *English Grammar: A University Course*. New York-London: Routledge.

DRESSLER, Wolfgang, 2005: "Word-formation in natural morphology" en Pavol ŠTEKAUER y Rochelle LIEBER (eds.): *Handbook of word-formation*. Amsterdam: Springer, 267-284.

FABB, Nigel, 2001: "Compounding" en Andrew SPENCER y Arnold M. ZWICKY (eds.): *The Handbook of Morphology*. Oxford: Blackwell, 66-83.

GRAÑA LÓPEZ, Benilde, 1994: "Compound formation in Generative Grammar", *Atlantis XVI*, 149-163.

GUEVARA, Emiliano, 2012: "Spanish compounds", *Probus. International Journal of Latin and Romance Linguistics*. vol 24 (1), 175-195.

GUEVARA, Emiliano y Sergio SCALISE, 2009: "Searching for universals in compounding" en Sergio SCALISE, Elisabetta MAGNI y Antonietta BISETTO (eds.): *Universals of language today*. Berlin: Springer, 101-128.

HASPELMATH, Martin y Andrea D. SIMS, 2010: *Understanding morphology*. London: Hachette.

HUDSON, Richard A., 1975: "Problems in the analysis of ed-adjectives". *Journal of Linguistics*, vol 11 (1), 69-72.

KILGARRIFF, Adam y otros, 2014: "The Sketch Engine: ten years on". *Lexicography*, Volume 1, Issue 1, 1-30.

LJUNG, Magnus, 1976: "-ed adjectives revisited". *Journal of Linguistics*, vol. 12, No. 1, Cambridge: Cambridge University Press, 159-168.

LANG, Mervyn F., 1990: *Spanish word formation. Productive derivational morphology in the modern lexis*. London-NewYork: Routledge.

LIEBER, Rochelle, 2005: "English word-formation processes" en Pavol ŠTEKAUER y Rochelle LIEBER (eds.): *Handbook of word-formation*. Amsterdam: Springer, 375-428.

LIEBER, Rochelle y Pavol ŠTEKAUER, 2009: "Status and definition of compounding" en Rochelle LIEBER y Pavol ŠTEKAUER (eds.): *The Oxford handbook of compounding*. Oxford: Oxford University Press, 3-18.

LORENZ, Frank, 2012: *Basics of phonetics and English phonology*. Berlin: Logos Verlag.

MC CARTHY, M.J. (1995). "Morphology" en Kirsten MALMKJAER y James ANDERSON (eds.): *The Linguistics Encyclopedia*, Routledge: London and New York, 314-323.

MCENERY, Tony y Andrew HARDIE, 2012: *Corpus linguistics: Method, theory and practice*. Cambridge: Cambridge University Press.

MCENERY, Tony y Andrew WILSON, 2001: *Corpus linguistics*. Edinburgh: Edingburgh University Press.

NAUMANN, Bernd y Petra M. VOGEL, 2000: "Derivation" en Geert BOOIJ, Christian LEHMANN y Joachim MUGDAN (eds.): *Morphology. An international handbook on inflection and word-formation*, volumen 1. Berlin, New York: Walter de Gruyter, 929-943.

OLSEN, Susan, 2000: "Composition" en Geert BOOIJ, Christian LEHMANN y Joachim MUGDAN (eds.): *Morphology. An international handbook on inflection and word-formation*, volume 1. Berlin, New York: Walter de Gruyter, 897-915.

OLSEN, Susan, 2014: "Delineating derivation and compounding" en Rochelle LIEBER y Pavol ŠTEKAUER (eds.): *The Oxford Handbook of Derivational Morphology*. Oxford: Oxford University Press.

OLSEN, Susan, 2015: "Composition" en Peter O. MÜLLER, Ingeborg OHNHEISER, Susan OLSEN y Franz RAINER (eds.): *Word-Formation. An international handbook of the languages of Europe*, volumen 1. Berlin, Boston: De Gruyter Mouton, 364-386.

PALMER, David, 2000: "Tokenisation and sentence segmentation" en Robert DALE, Hermann MOISL y Harold SOMERS (eds.): *Handbook of natural language processing*. New York: Marcel Dekker Inc., 11-35.

PLAG, Ingo, 2003: *Word-formation in English*. Cambridge: Cambridge University Press.

PLAG, Ingo, 2006: "Productivity" en Bas AARTS y April MCMAHON (eds.): *The handbook of English linguistics*. Oxford: Blackwell, 537-556.

PLAG, Ingo, 2011: "English" (preprint version) en Peter O. MÜLLER, Ingeborg OHNHEISER, Susan OLSEN y Franz RAINER (eds.): *Word-Formation. An international handbook of the languages of Europe*, 25. Berlin: Mouton de Gruyter.

REAL ACADEMIA ESPAÑOLA, 2009: *Nueva gramática de la lengua española*. Madrid: Espasa Libros.

RODRÍGUEZ, Julián, 1993: "La formación de palabras compuestas en inglés y en español", *Revista de Filología Inglesa*, N° 17, 21-37.

ROEPER, Thomas y Muffy E. A. SIEGEL, 1978: "A Lexical Transformation for Verbal Compounds". *Linguistic Inquiry*, vol. 9, N°. 2, 199-260

SPENCER, Andrew, 2005: "Word-formation and syntax" en Pavol ŠTEKAUER y Rochelle LIEBER (eds.): *Handbook of word-formation*, Springer: Amsterdam, 73-97.

SPENCER, Andrew, 1991: *Morphological theory*. Oxford: Blackwell Publishers.

TORRUELLA, Joan y Joaquim LLISTERRI, 1999: “Diseño de corpus textuales y orales” en José Manuel BLECUA y otros (eds.): *Filología e informática. Nuevas tecnologías en los estudios filológicos*. Barcelona: Editorial Milenio, 45-77.

VARELA, Soledad y Santiago FABREGAT, 2005: *Morfología Léxica: la formación de palabras*. Barcelona: Editorial Gredos.

YULE, George, 2010: *The study of language*. Cambridge: Cambridge University Press.

Diccionarios

Cambridge Dictionary online: <http://dictionary.cambridge.org/>

Collins English Dictionary online: <https://www.collinsdictionary.com/dictionary/english/>

Longman Dictionary of Contemporary English Online: <http://www.ldoceonline.com/>

Merriam Webster Dictionary online: <http://www.merriam-webster.com/>

Oxford Dictionaries online: <https://en.oxforddictionaries.com/>

Webster’s New World College Dictionary online: <http://websters.yourdictionary.com/>

Otras referencias en línea

<https://www.dailywritingtips.com/25-words-coined-by-twentieth-century-authors/>

<http://blog.oxforddictionaries.com/2016/02/invented-words/>

<http://www.medaholic.com/the-meaning-behind-the-name-medaholic/>

<https://www.sketchengine.co.uk/penn-treebank-tagset/>

<http://www.sketchengine.co.uk>

<https://www.etymonline.com>

Referencias corpus de prueba

BRONTË, Emily, 1994: *Wuthering Heights*. Oxford: New Windmill Classics

SHIPPEY, Tom, 1993: *The Oxford book of science fiction stories*. Oxford: Oxford University Press

SOMERSET MAUGHAM, W., 1938: *Altogether. The collected stories of W. Somerset Maugham*. London: William Heinemann Ltd.

SCOTT FITZGERALD, Francis, 1989: *The short stories of F. Scott Fitzgerald*. New York: Scribner.

STOKER, Bram, 2003: *Dracula*. London: The Collector's Library.

Empire Magazine: <http://www.empireonline.com/>

Mail Online: <http://www.dailymail.co.uk/home/index.html>

National Geographic: <http://www.nationalgeographic.com/>

The New York Times: <http://www.nytimes.com/>

Vanity Fair: <http://www.vanityfair.com/>