



Tipos complejos: Una redefinición a partir de su comportamiento en corpus

Redefining dot objects through their corpus behavior

Marina Berri

CONICET

UNIVERSIDAD DE BUENOS AIRES

UNIVERSIDAD NACIONAL DE GENERAL SARMIENTO

ARGENTINA

marinaberri@gmail.com

Recibido: 14-I-2018 / **Aceptado:** 28-XI-2018

DOI: 10.4067/S0718-09342019000200481

Resumen

El trabajo revisa la noción de tipo complejo del Lexicón Generativo (Pustejovsky, 1995) a partir del análisis contextual contrastivo de nombres tradicionalmente considerados tipos complejos ('libro', 'cena', 'escuela') y de nombres considerados tipos unificados ('botella') y pseudo-tipos complejos ('cerdo'). Se concluye que tienen un comportamiento dispar respecto de dos variables, la copredicación y la ambigüedad, pero que también hay diferencias entre los mismos nombres clasificados como tipos complejos. Se discute la importancia de estas variables en base a la redefinición de los mecanismos generativos (Pustejovsky, 2008, 2011). Finalmente, se propone una articulación de la noción de tipo complejo con la Teoría de las Normas y Explotaciones (Hanks, 2013), que tome en cuenta las características del perfil combinatorio de estas palabras.

Palabras Clave: Polisemia, tipos complejos, tipos unificados, Lexicón Generativo, Teoría de las Normas y Explotaciones.

Abstract

We aim to revise dot objects (Pustejovsky, 1995) through a contrastive contextual analysis of nouns traditionally classified as dot objects ('libro', 'book'; 'cena', 'dinner'; 'escuela', 'school') and nouns classified as unified types ('botella', 'bottle') and as pseudo-dot objects ('cerdo', 'pig/pork'). We show that these classes have a different behavior regarding copredication and ambiguity, but that there are also differences between the nouns traditionally classified as dot objects. The importance of these variables is discussed taking into account the changes introduced in the generative mechanisms (Pustejovsky, 2008, 2011). Finally, an articulation between the notion of dot objects and the Theory of Norms and Exploitations (Hanks, 2013) is proposed; this articulation takes into account the combinatory profile of these words.

Key Words: Polysemy, dot objects, unified types, Generative Lexicon, Theory of Norms and Exploitations.

INTRODUCCIÓN

Los tipos complejos se definen como palabras con cierto grado de polisemia que tienen simultáneamente dos tipos semánticos vinculados mediante una relación (Pustejovsky, 1995; Rumshisky, Grinberg & Pustejovsky, 2007; Asher & Pustejovsky, 2013). La polisemia que presentan se denomina inherente, dado que el potencial de interpretaciones múltiples es propio del ítem léxico mismo y la referencia se realiza mediante la conjunción de ambos tipos (Pustejovsky & Jezek, 2009). Así, los tipos complejos pueden manifestar uno u otro de sus tipos semánticos, o bien ambos, según el contexto. En la bibliografía, 'libro' constituye un ejemplo clásico, ya que combina los tipos INFORMACIÓN y OBJETO FÍSICO mediante la relación 'contener': un libro es un objeto físico que contiene información. Existen predicados que modifican independientemente a cada uno de estos tipos, como se observa en (1) y (2):

(1) Un **libro** difícil → INFORMACIÓN

(2) Un **libro** que pesa dos kilos → OBJETO FÍSICO

Se considera que la relación que une a los tipos es parte de su significado. Así, en 'libro', la relación 'contener' forma parte del significado de 'libro'. Otros ejemplos clásicos son 'almuerzo' (ALIMENTO/ EVENTO en que se consume ese alimento) y 'hospital' (INSTITUCIÓN/ EDIFICIO en el que funciona esa institución).

Tradicionalmente, la prueba utilizada para determinar si una palabra es un tipo complejo ha sido la copredicación (Pustejovsky & Jezek, 2009). En la copredicación, dos sentidos diferentes de un ítem léxico son actualizados simultáneamente al aplicársele dos predicados, cada uno de los cuales requiere un tipo semántico diferente e incompatible. La copredicación (Pustejovsky, 1995; Brandtner, 2011) consiste en que dos predicados o modificadores que se aplican a diferentes dominios semánticos, como estados, objetos o locaciones, pueden ser predicados simultáneamente del

mismo nombre, aun cuando indican diferentes lecturas de ese nombre. Se ha afirmado que no se presenta en homónimos y ha sido usada como un test para determinar en qué grado dos lecturas/sentidos están relacionados (Cruse, 2000). En (3), que ilustra la copredicación sobre el ítem 'libro', 'memorizando' actualiza el tipo semántico INFORMACIÓN mientras que 'pesa un kilo' actualiza el tipo OBJETO FÍSICO:

(3) El **libro** que estoy memorizando INFORMACIÓN pesa un kilo OBJETO FÍSICO.

Sin embargo, este tipo de tests ha sido criticado desde la semántica léxica por proporcionar resultados contradictorios (Geeraerts, 1993). Por otra parte, en la bibliografía no se ha tenido en cuenta la ambigüedad en corpus, es decir, la imposibilidad de determinar qué significado se actualiza en contexto (Edmonds, 2006). En este trabajo, me propongo caracterizar el funcionamiento de los tipos complejos en corpus. Parto del supuesto de que la estructura semántica de los nombres impacta sobre su funcionamiento contextual. Por lo tanto, si existen diferencias entre la estructura semántica de un tipo complejo y otros ítems que no presentan polisemia inherente esto debería reflejarse en contexto. El estudio en corpus de estos nombres permitirá elaborar una metodología que permita distinguir entre tipos complejos y otros nombres con diferentes grados de polisemia.

En este sentido, considero que este estudio contextual puede basarse en la Teoría de las Normas y Explotaciones (TNE), propuesta por Hanks (2013). La TNE postula que las palabras tienen usos normales, es decir, significados asociados a patrones de uso típicos, pero que a su vez las voces pueden explotarse, es decir, adquirir nuevos usos que se desvían de esos patrones. Por lo tanto, esta teoría provee un fundamento para determinar si las copredicaciones y otros índices de polisemia, como los contextos ambiguos (Berri, 2014), pueden considerarse como correspondientes a una norma o bien como explotaciones esporádicas con respecto a un nombre. De hecho, la TNE incorpora explícitamente aportes del LG, como la noción de coerción de tipos (Hanks, 2013), y Hanks y Pustejovsky (2005) han colaborado en la elaboración del *Pattern Dictionary of English Verbs*. Sin embargo, los tipos complejos no han sido estudiados a partir de la articulación de estas teorías.

El objetivo del presente trabajo es caracterizar el funcionamiento contextual de los tipos complejos a partir del análisis exhaustivo de concordancias. La hipótesis sostenida es que los tipos complejos se caracterizan por mostrar copredicación de un modo más frecuente que los tipos unificados, pero que la ambigüedad desempeña un papel decisivo en lo que respecta al funcionamiento de los tipos complejos. En particular, cabe preguntarse:

- ¿Copredicación y ambigüedad son variables independientes? Es decir, a más casos de copredicación para un nombre, ¿hay más casos de ambigüedad?

- ¿La copredicación es un fenómeno lo bastante frecuente para caracterizar el comportamiento distribucional de los tipos complejos?

En base a estas respuestas, intentaré establecer una norma de los tipos complejos, tomando en cuenta los predicados que seleccionan ambos, o bien uno u otro tipo, de los tipos complejos.

En primer lugar, describiré brevemente el Lexicón Generativo (LG) y la TNE. En segundo lugar presentaré la metodología propuesta para clasificar estos nombres, que se basa en corpus y que parte del supuesto de que la estructura semántica debe establecerse a partir de un análisis contextual exhaustivo. Luego analizaré diez nombres que han sido clasificados de diferente manera en la bibliografía a partir de la metodología propuesta. Más adelante presentaré una discusión, en la que estableceré la definición de tipo complejo y discutiré la copredicación y la ambigüedad, articulando el LG con la TNE. Finalmente, esbozaré las conclusiones y las líneas futuras.

1. Marco teórico y antecedentes

1.1. El Lexicón Generativo

El LG (Pustejovsky, 1995, 1996, 2001, 2005, 2011; Pustejovsky, Bouillon, Isahara, Kanzaki & Lee, 2013) es una teoría centrada en explicar la polisemia, que sostiene que el léxico almacena gran parte de la información que subyace al uso creativo del lenguaje y que el sentido se produce en contexto. Específicamente, el LG se distancia de una concepción del léxico estática, según la cual existirían significados listados y grupos de contextos que licenciarían de antemano el uso de las palabras.

Pustejovsky (1995, 2001) postula que las unidades del léxico tienen una estructura semántica que consta de cuatro niveles de representación: la estructura argumental, la estructura eventiva, la estructura de *qualia* y la estructura de herencia léxica. En el primer nivel se especifican el número y el tipo de argumentos codificados en una unidad léxica y su expresión en la sintaxis. En el segundo nivel, el de la estructura eventiva, se especifica el tipo de evento de una expresión y la estructura subeventiva. La estructura de *qualia* representa los atributos del referente del ítem léxico según cuatro modos de explicación. Para Pustejovsky (2001), el *quale* formal (QF) es la categoría básica que distingue una entidad en un dominio más amplio, el *quale* constitutivo (QC) es la relación entre un objeto y sus partes constituyentes, el *quale* télico (QT) es su función y el *quale* agentivo (QA) está conformado por aquellos factores involucrados en su origen. Esta estructura da cuenta de la fuerza predicativa de un ítem. Pustejovsky (2008, 2011) señala que el tipo del *quale* formal constituye el tipo nuclear (*head type*), el tipo que define la naturaleza de la entidad. Resultaría, entonces, el menos costoso de seleccionar, dado que se selecciona sin presión contextual extra. Los otros *qualia* constituyen la cola (*tail*) y definen las causas explicativas de la entidad del tipo nuclear. De acuerdo con Batiukova (2006), la

estructura de *qualia* define los atributos esenciales de objetos, eventos y relaciones asociados a la unidad léxica que determinan su polisemia y su uso en ciertos contextos sintácticos. Así, los valores de los *qualia* son predicados y relaciones prototípicas asociadas con las palabras. Finalmente, en la estructura de herencia léxica se establecen las relaciones entre los ítems léxicos, específicamente entre los tipos semánticos. La Figura 1 ilustra la estructura semántica de los nombres:

EA= Arg1: x
EE= Ev1: e1
EQ= Formal: what x is
Constitutive: what x is made of
Telic: function of x
Agentive: how x came into being

Figura 1. ES de los nombres (Pustejovsky, 2001).

Dentro del LG, se distingue entre tipos simples, unificados y naturales. Los tipos simples son los que no cuentan con QA y QT, por ejemplo, ‘árbol’. En cambio, los tipos unificados cuentan con QA y QT, por ejemplo, ‘café’. Los tipos complejos han sido estudiados en diferentes lenguas, entre ellas inglés (Pustejovsky, 1995, 2005; Asher & Pustejovsky, 2013), italiano (Jezek & Melloni, 2011) y español (Berri, 2014). Como ya se ha afirmado, cuentan con dos tipos en su QF unidos por una determinada relación (‘libro’) y han sido caracterizados porque permiten la copredicación. Sin embargo, su determinación es problemática. Así, Asher y Pustejovsky (2013) señalan que no solo los tipos complejos, sino también los unificados admiten la copredicación. Así, los autores citan el siguiente ejemplo, en el que *wine*, una palabra tradicionalmente considerada como un tipo unificado, es objeto de copredicación:

(4) She opened[CONTENEDOR] the **wine** and poured[BEBIDA] some into the glass.

Pustejovsky y Jezek (2009) señalan que el sentido de ‘bebida’ sería el sentido inherente, pero el de ‘contenedor’ sería introducido contextualmente mediante el verbo *open*. No obstante, los autores señalan que la existencia de estos ejemplos pone de relieve que la distinción entre tipos complejos y unificados es difícil y no ha sido establecida con claridad en la bibliografía.

1.2. La Teoría de las Normas y las Explotaciones

Hanks (2013) propone basar la investigación sobre el significado léxico no en la especulación acerca de lo posible, realizada mediante tests y ejemplos inventados, sino en aquello que es típico y se evidencia mediante el análisis de corpus textuales. A partir de la identificación de los usos típicos, resulta posible explicar otros usos que dan cuenta de la no convencionalidad y explican en parte la creatividad lingüística.

Según Hanks, existen dos sistemas de reglas interrelacionados. Un primer sistema de reglas está conformado por las normas, que son los modos en que las palabras se

usan de modo usual e idiomático. Un segundo sistema lo constituyen las explotaciones de esas normas, es decir, el modo en que los hablantes transgreden las convenciones para decir cosas nuevas. Así, los hablantes explotan los estereotipos, típicamente para lograr un efecto retórico. Específicamente, una norma constituye un patrón de uso convencional –una estructura argumental con un conjunto de colocaciones– que tiene asociado un significado. El siguiente artículo de *Verbario* (Renau & Nazar, en proceso) ilustra este concepto a partir de las normas asociadas al verbo ‘acuñar’:

verbo ‘acuñar’

1. Patrón: [[Institución = Estado | Persona = Gobernante]] acuñar [[Artefacto = Moneda]]

Implicatura: [[Institución = Estado | Persona = Gobernante]] fabrica y pone en circulación [[Artefacto = Moneda]] para su uso oficial.

2. Patrón: [[Humano]] acuñar [[Parte de Acto de Habla = Palabra | Expresión]]

Implicatura: [[Humano]] crea [[Parte de Acto de Habla = Palabra | Expresión]] que se hace célebre.(...)

Figura 2. Normas de ‘acuñar’ (*Verbario*).

Como muestra el artículo, una primera norma está asociada al significado de ‘fabricar y poner en circulación una moneda’. El patrón se especifica mediante los tipos GOBERNANTE/ESTADO y MONEDA. Una segunda norma se asocia al sentido de ‘crear una palabra’ e incorpora los tipos semánticos HUMANO y PALABRA/EXPRESIÓN. Las siguientes concordancias ilustran ambas normas:

(5) Es una moneda de oro **acuñada** en tiempos de la Colonia.

(6) La palabra *ecología* fue **acuñada** en 1869 por el biólogo alemán Ernest Hackel. (RAE, CORPES)

Las explotaciones son modos dinámicos del lenguaje y se usan para decir algo de modos nuevos. Las siguientes concordancias ilustran la explotación de las normas anteriores. En ellas, el objeto del verbo ‘acuñar’ corresponde a otros tipos semánticos. Así, en (7) y (8) ‘acuñar’ tiene el sentido de ‘crear’:

(7) La tragedia está presente aquí como la más sublime de las formas escénicas **acuñadas** por los helenos.

(8) Luis Rafael Sánchez **acuñó** la aspiración a escribir en puertorriqueño. (RAE, CORPES)

La noción de norma está asociada fundamentalmente a la categoría verbal, porque un patrón se define precisamente como la estructura argumental de un predicado y los tipos semánticos (o sets léxicos) asociados a esa estructura argumental. Sin embargo, en Hanks (2004) se presenta un modelo de norma para explicar la explotación metafórica de nombres, muchos de los cuales no poseen estructura argumental. Para el autor, las normas de los nombres están constituida fundamentalmente por sus colocaciones, algunas de las cuales están en una relación sintáctica específica respecto de ellos. Así, por ejemplo, respecto de *storm*, *protest* se encuentra en una relación

sintáctica específica (*storm of protest*), mientras que *rain* y *hurricane* suelen coocurrir en el contexto, pero no en una relación específica respecto de *storm*. La descripción de las colocaciones conforma el perfil combinatorio que indica cómo es usado el nombre *storm* normalmente. Hanks (2004) presenta el siguiente perfil combinatorio para los usos literales de *storm*:

- WHAT DO STORMS DO? Storms **break, blow, rage, lash** coastlines, **batter** ships and places, **hit** ships and places, **ravage** places.
- BEGINNING AND END OF A STORM: Before it breaks, a storm is **brewing, gathering, or impending**. (...)
- WHAT HAPPENS TO PEOPLE IN A STORM? People can **weather, survive, or ride (out)** a storm. Ships and people may get **caught in** a storm.
- WHAT KINDS OF STORMS ARE THERE? There are **thunder storms, electrical storms (...)**. Storms are **violent, severe, raging, howling** (...). Storms, especially snow storms, may be **heavy**. An unexpected storm is a **freak** storm.
- The centre of a storm is called the **eye of the storm**.
- STORMS ARE ASSOCIATED WITH **rain, wind, hurricanes, gales, and floods**.

Figura 3. Norma de *storm* (Hanks, 2004).

Estos antecedentes ponen de manifiesto que la noción de norma puede extenderse a piezas léxicas que no son predicados. Aquí proponemos una extensión a los tipos complejos. Dado que estos nombres se definen en base a su capacidad de aparecer en contextos que requieren tipos contradictorios, proponemos caracterizarlos tomando en cuenta su perfil combinatorio en relación con la capacidad de que sus tipos se activen simultáneamente en contexto. Los tipos complejos serían aquellas piezas léxicas que tienen una única norma que cubre ambos tipos y que da cuenta precisamente del hecho de la activación simultánea en contexto. En cambio, los tipos unificados y los pseudo-tipos complejos contarían con más de una norma, es decir, con más de un set de colocaciones, dado que los tipos se encontrarían en competencia más que en colaboración y corresponderían dos sentidos claramente distinguibles. Sería posible pensar que los tipos complejos son aquellas unidades cuya norma está compuesta por predicados que requieren ambos tipos semánticos y por predicados que interactúan en la copredicación. Los tipos complejos tendrían así una norma vinculada a la actualización de ambos tipos en contexto y esto debería observarse en un mayor número de copredicaciones y de contextos ambiguos. Estas variables indicarían que los tipos semánticos en el caso de la copredicación se están activando de manera conjunta o, en el caso de la ambigüedad, son indiscernibles.

2. Metodología

Para realizar el análisis, se estudió el funcionamiento contextual en corpus de diez nombres: ‘acuarela’, ‘almuerzo’, ‘botella’, ‘cena’, ‘cerdo’, ‘escuela’, ‘hospital’, ‘libro’, ‘pinar’, ‘vals’. La elección de estos nombres obedece a que sus equivalentes ingleses y/o las polisemias regulares de las que participan han sido clasificados en la

bibliografía de manera diversa: algunos como tipos complejos (‘almuerzo’, ‘cena’, ‘libro’, ‘pinar’) (Pustejovsky, 1995, 2005; De Miguel 2009; Berri, 2014); otros (‘escuela’, ‘hospital’) han sido clasificados alternativamente como tipos complejos (De Miguel, 2009) y como tipos unificados (Rumshisky et al., 2007), o bien como pseudo-tipos complejos, como ‘cerdo’ y ‘botella’ (Pustejovsky, 2005). Finalmente se han seleccionado dos nombres, ‘acuarela’ y ‘vals’, que aun no han sido clasificados en la bibliografía. Se han tomado estos nombres, cuya clasificación en la bibliografía es diversa, para establecer contrastivamente cuál es el comportamiento de los tipos complejos y cuál el de los tipos unificados y pseudo-tipos complejos en relación con las variables de copredicación y ambigüedad. El corpus utilizado fue el *Spanish Web Corpus* (esTenTen), disponible en el *Sketch Engine* (para una presentación general del *Sketch Engine*, véase Kilgarriff, Baisa, Bušta, Jakubíček, Kovář, Michelfeit, Rychlý & Suchomel, 2014; para una presentación del esTenTen, véase Kilgarriff & Renau, 2013), que cuenta con aproximadamente diez mil millones de palabras. En primer lugar, se tomó una muestra al azar de 100 concordancias. Siguiendo los lineamientos de Berri (2014) se clasificaron los contextos según presentaran uno u otro sentido, fueran ambiguos o mostraran copredicación.

En segundo lugar, se constituyó un algoritmo de búsqueda específica de copredicaciones usando el *corpus query language* (CQL), un lenguaje del *Sketch Engine* que permite buscar patrones gramaticales o léxicos complejos que no pueden buscarse mediante la interfaz estándar. Para esto se tomó como insumo los resultados arrojados por la función *Word Sketch*, función que permite extraer automáticamente las palabras que coaparecen con un determinado ítem, descritas según las relaciones gramaticales que mantienen con ese ítem. En palabras de Kilgarriff et al. (2014: 9), un word sketch es “*a one page summary of a word’s grammatical and collocational behaviour*”. El primer paso consistió en seleccionar del perfil combinatorio generado por el *Word Sketch* las colocaciones que activaran un determinado tipo semántico, que llamaremos indicadores (Brandter, 2011; Adelstein & Berri 2013; Berri 2014). Las voces que no constituían indicadores, ya sea por ser ambiguas, o por constituir errores del Word Sketch, se descartaron manualmente luego de analizar las concordancias en las que figuran. Se obtuvieron, por ejemplo, los siguientes indicadores de los respectivos tipos semánticos de ‘cena’:

- ALIMENTO: ‘delicioso’, ‘rico’, ‘ligero’, ‘copioso’; ‘servir’, ‘preparar’, ‘degustar’, ‘cocinar’...
- EVENTO: ‘íntimo’, ‘informal’, ‘transcurso’, ‘anual’; ‘terminar’, ‘concurrir’, ‘transcurrir’, ‘asistir’...

Luego se seleccionaron siete indicadores de cada uno de los sentidos y se completó el algoritmo, que arrojó como resultado contextos en los que coaparecía el nombre buscado con un indicador cualquiera de cada uno de los grupos en un rango de cinco palabras. Así, por ejemplo, para el ítem ‘cena’ se obtuvieron resultados en que

apareciera ‘cena’ y también ‘delicioso’, que selecciona el tipo semántico ALIMENTO, e ‘informal’, que selecciona el tipo semántico EVENTO. Así, mediante estas búsquedas que llamaremos búsquedas dirigidas, diseñadas a partir de los indicadores y el algoritmo de búsqueda, se obtuvo un mayor número de contextos en los que hay copredicación, como (9); pero también otros como (10) en los que los indicadores seleccionados no se aplican al ítem léxico (‘copiosa’ se aplica a ‘cena’, pero ‘íntima’, a pesar de estar dentro del rango especificado, se aplica a otro ítem, a ‘ceremonia’) y fueron descartados (manualmente) para el análisis por considerarse ruido:

(9) **Copredicación:** El Nuevo Lunes entregó también los trofeos (...) en el transcurso[EVENTO] de una exquisita[ALIMENTO] y cuidada **cena servida**[ALIMENTO] en el restaurante madrileño La Dorada, propiedad de Félix Cabeza.

(10) **Ruido:** Todo salió como estaba esperado: el tiempo fue soleado y cálido, (...); la **ceremonia** fue íntima y muy emotiva[EVENTO]; la **cena**, copiosa[ALIMENTO] a más no poder y el baile divertido e inacabable.

Luego de un análisis manual de estos resultados arrojados por el algoritmo, se conservaron para el análisis las copredicaciones y se descartó el ruido. Cabe aclarar que la cantidad de contextos arrojados por el algoritmo fue variable y también la cantidad de copredicaciones obtenidas; pero en principio estas búsquedas posibilitan examinar el comportamiento de la palabra en relación con la copredicación. Se consideraron hasta doscientos contextos, que fueron sumados a las copredicaciones detectadas de manera manual. A partir de estos datos, se buscó analizar contrastivamente el comportamiento contextual de los ítems léxicos. Luego se recogieron las concordancias en una base de datos para cada nombre. En las concordancias se marcaron los indicadores que activan uno u otro tipo semántico y se clasificaron los contextos según se activara uno u otro tipo semántico del nombre, resultarían ambiguos o mostrarían copredicación. Finalmente, a partir del análisis de las concordancias y de los Word Sketch se propusieron las normas de los nombres que se presentan en la ‘Discusión’.

3. Análisis

3.1. Copredicación

Un primer paso del análisis consistió en hacer una detección manual de nombres en relación con la copredicación. Tal como se observa en la Tabla 1, el estudio reveló que, de 100 contextos analizados manualmente, la copredicación se presentaba siempre en menos del 10% de los casos. Esto pone de manifiesto que resulta difícil postular a la copredicación como un fenómeno lo suficientemente frecuente como para permitir caracterizar el comportamiento contextual de los tipos complejos, dada su baja frecuencia en general. En estudios anteriores (Pustejovsky, 1995; Cruse, 2000;

Croft & Cruse, 2008; Pustejovsky & Jezek, 2009; Asher & Pustejovsky, 2013) y al ser empleada únicamente como test, no se tuvo en cuenta la relevancia total, sino meramente en relación con la dicotomía posible / no posible. Por otro lado, como ya se ha señalado, la copredicación se realiza también sobre nombres que han sido clasificados como pseudo-tipos complejos y tipos unificados, como por ejemplo ‘cerdo’ y ‘botella’, y esto se evidencia también en el corpus analizado.

Tabla 1. Copredicación - Análisis manual sobre 100 contextos.

Nombre	Porcentaje decopredicaciones
‘pinar’	7%
‘libro’	4%
‘almuerzo’	5%
‘cena’	5%
‘escuela’	1%
‘hospital’	2%
‘cerdo’	3%
‘botella’	1%
‘vals’	0%
‘acuarela’	0%

Sin embargo, este primer análisis permite realizar algunas observaciones, que luego se profundizan al realizar búsquedas dirigidas específicamente para obtener contextos de copredicación. Como se explicitó en la ‘Metodología’, a partir del Word Sketch de cada palabra se obtuvieron concordancias en las que aparecen indicadores contradictorios, cada uno de los cuales apunta a un tipo diferente del tipo complejo. Se pudo observar así que, de los nombres analizados, ‘acuarela’ no muestra copredicación en los resultados de estas búsquedas; para ‘vals’ tampoco se encuentran mayormente resultados. En cambio, para algunos nombres, resulta fácil encontrar contextos de copredicación. Así, para ‘pinar’, ‘libro’, ‘almuerzo’ y ‘cena’ han podido encontrarse más de 40 casos de copredicación. En cambio, continúa siendo relativamente difícil encontrar copredicación ejercida sobre otros nombres. Así, para ‘escuela’ y ‘hospital’ se encuentran, respectivamente, solo 8 y 13 casos de copredicación. De hecho, estas palabras han sido consideradas tipos complejos a partir del test de la copredicación; sin embargo, resulta igual de frecuente encontrar casos de copredicación para ‘cerdo’ (12 casos), considerado como un pseudo-tipo complejo, y es incluso más frecuente encontrar casos de copredicación para ‘botella’, considerado un tipo unificado con información sobre su contenido en el *quale* tético (20 casos). Los siguientes ejemplos ilustran casos de copredicación:

(11) En fin, después[EVENTO] de un **almuerzo ligero**[ALIMENTO], durante[EVENTO] el cual pondremos especial atención para no hablar de trabajo, volvemos a nuestras traducciones.

(12) Nelly abrió[CONTENEDOR] la **botella** y la bebió[CONTENIDO] a sorbos pequeños.

(13) Según las autoridades, el **hospital** (construido[EDIFICIO] en 1884 y que atiende[INSTITUCIÓN] a pacientes de ambas comunas), evidencia el daño que ha sufrido a través del tiempo (...)

(14) (...)Nacho Rubio marchará por esposos[CONJUNTO DE ÁRBOLES] **pinos** y otras zonas[LUGAR] en las que la diversidad botánica está presente en cada paso.

(15) “Disfruto de esto tanto como cualquier otra cosa y ahorro mucho dinero”, dice Bishop, de 46 sosteniendo una salchicha de [ALIMENTO] **cerdo criado en una granja familiar** [ANIMAL], metida en un panecillo y cubierta de pepinillos.

(16) (...) podéis aprender a bailar los mejores **vals** que se puedan escuchar[MÚSICA], y demostrarlo en cualquier restaurante o pub de la ciudad, donde siempre tienen un hueco para este baile[DANZA].

En la Tabla 2 se observan los casos de copredicación encontrados mediante las búsquedas dirigidas:

Tabla 2. Copredicación mediante búsquedas dirigidas.

Nombre	Copredicación manual	Copredicación mediante búsquedas dirigidas
‘pinar’	7%	56 casos (sobre 200 contextos, 28%)
‘libro’	4%	42 casos (sobre 200 contextos, 21%)
‘almuerzo’	5%	66 casos (sobre 200 contextos, 33%)
‘cena’	5%	79 casos (sobre 200 contextos, 39%)
‘escuela’	1%	8 casos (sobre 200 contextos, 4%)
‘hospital’	2%	13 casos (sobre 200 contextos, 6%)
‘cerdo’	3%	12 casos (sobre 63 contextos, 19%)
‘vals’	0%	2 casos (sobre 84 contextos, 2%)
‘botella’	1%	20 (sobre 200 contextos, 11%)
‘acuarela’	0	0

Por otro lado, cabe preguntarse si existe una relación específica entre un nombre y el mayor porcentaje de copredicaciones y, más en particular, entre una clase específica de polisemia y la cantidad de copredicaciones que los nombres correspondientes a esas clases presentan. Así, para ‘escuela’ y ‘hospital’, que muestran la polisemia regular INSTITUCIÓN/EDIFICIO, se registran manualmente uno o dos casos de copredicación y mediante las búsquedas dirigidas, 8 y 13, mientras que para ‘almuerzo’ y ‘cena’, que tienen la polisemia EVENTO/ALIMENTO, se registran manualmente 5 casos de copredicación, mientras que mediante búsquedas dirigidas el número trepa a 66 para ‘almuerzo’ y 79 para ‘cena’. La evidencia obtenida en corpus permite probar que los nombres tienen, a gran escala, diferente comportamiento en lo que respecta a

su posibilidad de ser objeto de copredicación si se parte de búsquedas que incorporan voces que forman parte de su perfil combinatorio (Word Sketch).

3.2. Ambigüedad

Una segunda parte del análisis consistió en detectar aquellos contextos en los que no se pudiera determinar el tipo actualizado para el nombre analizado. Así, por ejemplo, sucede en los siguientes ejemplos:

(17) La donante es ingresada en un **Hospital**, donde sólo permanecerá un día.

(18) En Chile, la IGVZ está disponible en el stock crítico de medicamentos de la farmacia del **Hospital** de Urgencia (...)

En (17), ‘ingresar’ puede tomar como argumento tanto un LUGAR como una INSTITUCIÓN; de hecho, creemos que en este caso selecciona ambos. Por otra parte, en ‘la farmacia del Hospital’, la farmacia puede ser parte de la institución y/o estar ubicada en el edificio. En estos casos, el contexto no requiere de una desambiguación, dado que los sentidos están en colaboración, más que en competencia. Considero que una mayor dificultad para distinguir el tipo semántico que se actualice en un contexto es indicio de que nos encontramos frente a un tipo complejo, dado que también constituyen una evidencia de polisemia inherente. De este modo, relevar a partir del Word Sketch de los nombres analizados qué predicados seleccionan ambos predicados del tipo complejo (como ‘ingresar’, para el caso de ‘hospital’, y ‘colección’, para el caso de ‘libro’) contribuiría a confeccionar la norma de estos nombres.

La Tabla 3 permite observar el comportamiento de los nombres respecto de la ambigüedad. También manifiesta que no existe una relación directa entre cantidad de contextos ambiguos y copredicación. Hay nombres como ‘escuela’ y ‘hospital’ que tienen un bajo porcentaje de copredicación y un alto porcentaje de ambigüedad. ‘Cerdo’, que muestra una cantidad de contextos de copredicación similar a ‘escuela’ y ‘hospital’, tiene en cambio un porcentaje de contextos ambiguos muy bajo (1%). Por otra parte, nombres que presentan frecuentemente contextos de copredicación, como ‘cena’ y ‘almuerzo’, evidencian bajos porcentajes de ambigüedad (17% y 7%, respectivamente).

Tabla 3. Copredicación y contextos ambiguos.

Nombre	Copredicación manual	Copredicación (búsquedas dirigidas)	Contextos ambiguos
‘pinar’	7%	56	55%
‘libro’	4%	42	37%
‘almuerzo’	5%	66	17%
‘cena’	5%	79	7%
‘escuela’	1%	8	38%
‘hospital’	2%	13	38%
‘cerdo’	3%	12	1%
‘vals’	0%	2	23%
‘botella’	1%	20	17%
‘acuarela’	0	0	9%

4. Discusión

El análisis realizado permite establecer que el comportamiento contextual de los nombres clasificados como tipos complejos, como pseudo-tipo complejos y como tipos unificados difiere. Sin embargo, la copredicación no alcanza para distinguirlos, dado que se presenta tanto en unidades como ‘cena’, tradicionalmente consideradas tipos complejos, como en unidades como en ‘cerdo’, consideradas pseudo-tipos complejos (Pustejovsky, 2005). Por otro lado, nombres que en la bibliografía son considerados tipos complejos tienen en corpus un comportamiento también diferente. Así, mientras nombres como ‘cena’ muestran usualmente copredicación pero son más raramente ambiguos, otros como ‘escuela’ muestran más raramente copredicación, pero en cambio son muy frecuentemente ambiguos. Esto permite afirmar que copredicación y ambigüedad son dos variables hasta cierto punto independientes. A la vista de estos resultados, creemos que la definición misma de tipo complejo debe ser revisada tomando en cuenta el comportamiento contextual, de modo que se puedan obtener criterios claros que permitan establecer con fiabilidad qué es un tipo complejo. Por otro lado, una base de evidencia empírica otorgará sustento para determinar qué es –y qué no– una referencia compleja construida mediante dos tipos.

La redefinición de la noción de tipos complejos puede realizarse en base a la revisión de la que han sido objeto los mecanismos generativos en el LG. En Pustejovsky (1995) se postulaban tres mecanismos generativos: la coerción de tipos, el ligamiento selectivo y la co-composición. En trabajos posteriores (Pustejovsky, 2008, 2011) se revisaron esos mecanismos y se distinguió entre selección, acomodación – mecanismo que no abordaremos aquí por no resultar relevante para la discusión– y coerción de tipos. La selección tiene lugar cuando el tipo que requiere una función es directamente satisfecho por el argumento. En la coerción de tipos, el predicado impone un determinado tipo semántico al argumento. Pustejovsky (2008, 2011) distingue dos variantes, explotación e introducción. La primera, la explotación, ocurre

cuando se accede a un subcomponente del tipo del argumento para satisfacer los requerimientos del predicado. Por ejemplo, en ‘Juan empezó su novela’, se accede al evento ‘escribir’, presente en el *quale* agentivo del nombre ‘novela’ y en ‘Juan manchó el libro’, se explota uno de los dos tipos semántico del tipo complejo ‘libro’, el tipo OBJETO FÍSICO. En cambio, en la segunda variante, la introducción, el argumento se “envuelve” con el tipo requerido por la función. Esta operación introduce nuevos materiales conceptuales que no eran parte del significado original de la palabra, es decir, que a diferencia de lo que ocurre en la explotación, no se encuentran en la estructura semántica. Así, en ‘Juan leyó la pared’, el tipo INFORMACIÓN es introducido por ‘leer’. El tipo unificado ‘pared’ es convertido en contexto en un tipo complejo, ya que el predicado introduce el tipo INFORMACIÓN y se lo vincula al tipo OBJETO FÍSICO mediante la relación ‘contener’. En el caso de los tipos complejos, la selección es del tipo complejo completo y se considera que cuando en contexto un predicado requiere uno solo de los tipos se emplea el mecanismo –más costoso que la simple selección– de explotación de tipos complejos.

La distinción entre selección, explotación e introducción parece requerir que se considere en cierta medida la diferencia entre aquello que es normal y aquello que no. Si bien en el caso de ‘pared’ es claro que INFORMACIÓN no es un tipo que esté en su *quale* formal, en otros casos resulta difícil definir el estatuto y la ubicación en la estructura semántica de un tipo. Por ejemplo: ¿es el tipo locativo parte del *quale* formal de ‘escuela’, o más bien corresponde a algún otro *quale* de su estructura (por ejemplo, el télico o el constitutivo)? ¿MÚSICA y DANZA se encuentran unidos en el *quale* formal de ‘vals’, o bien DANZA se encuentra en el *quale* télico? Una posible respuesta surge de examinar cuán frecuentemente ambos tipos se actualizan en contexto, es decir, si es parte del perfil combinatorio o no de ‘escuela’ y de ‘vals’ la actualización conjunta de los tipos semánticos, para lo cual puede tomarse el concepto de norma de la TNE. Si ambos tipos semánticos se actualizan conjuntamente de manera frecuente, sería posible determinar que esos tipos se encuentran en el *quale* formal: dada esa frecuencia, la selección, la operación menos costosa, es la que prevalece. De esta manera, se probaría que se trata de una referencia compleja, dado que usualmente los tipos se actualizan de manera conjunta.

Los últimos avances del LG parecen apuntar también en esta dirección, es decir, en considerar el comportamiento en corpus para determinar cuál es la estructura semántica de un nombre. En Pustejovsky y Jezek (2016) se afirma que la teoría ha hecho adelantos en base a la lingüística de corpus y que este tipo de análisis proporciona una nueva evidencia para determinar cuál es la estructura semántica de los ítems léxicos. Así, los atributos convencionalizados¹, una de las últimas innovaciones del modelo (Pustejovsky & Jezek, 2016), se determinarían en base al comportamiento distribucional y a ciertos usos elípticos.

En principio, Pustejovsky (2011) postula que en los tipos complejos el mecanismo de selección –el menos costoso– selecciona ambos tipos del tipo complejo. Por lo tanto, desde esta reformulación de los mecanismos es posible afirmar que los tipos complejos se caracterizan por presentar ambos tipos en contexto, i.e, por ser típicamente ambiguos. De ahí que consideremos que los tipos complejos presentan una alta tasa de ambigüedad: un significado compuesto por dos tipos, pero que no suele especificarse en contexto. Como ya ha sido expuesto, Pustejovsky (2011) señala que el tipo del *quale* formal constituye el tipo nuclear (*head type*), el tipo que define la naturaleza de la entidad. Resultaría, entonces, el menos costoso de seleccionar o explotar. Los otros *qualia* constituyen la cola (*tail*) y definen las variadas causas explicativas de la entidad del tipo nuclear. Así, los valores de *qualia* distintos del *quale* formal pueden verse como un completamiento estructural del *head type*. Postular un tipo nuclear implica afirmar que ese tipo se corresponde con el tipo general de la unidad –i.e., el del *quale* formal, que captura la relación ‘es-un’– y que no se requiere presión contextual adicional para activarlo, lo que se realiza mediante la selección pura.

La ambigüedad sería un aspecto fundamental para determinar si un nombre es o no un tipo complejo. Por lo tanto, los nombres para los que es habitual que en contexto no se especifique un sentido, deberían ser tipos complejos. Esto bastaría para determinar que ‘pinar’, ‘libro’, ‘escuela’ y ‘hospital’ son tipos complejos. En este sentido, consideramos que es la variable más indicada para determinar que un nombre es un tipo complejo, dado que se vincula con la referencia doble y con el hecho de que sea menos costoso seleccionar el tipo complejo que explotar una de las partes que lo constituyen.

Una segunda variable en el análisis ha sido la copredicación. Si bien es un fenómeno relativamente raro en las concordancias examinadas, implica que se exploten activamente ambos tipos semánticos.

La TNE de Hanks proporciona herramientas para describir el comportamiento contextual habitual de las piezas léxicas. Si bien la noción de norma ha sido específicamente fructífera respecto de la categoría verbal, es posible construir normas para los nombres analizados de modo similar a la norma planteada por Hanks para *storm* (Figura 3), fundamentalmente a partir de las voces con las que suele combinarse. Sin embargo, es conveniente tener en cuenta que, de acuerdo con Hanks (2004), las normas de los nombres se construyen de manera diferente a aquellas de los verbos, dado que no se basan en estructuras argumentales sino que se centran en las colocaciones, “*making statements about prototypical usage*” (Hanks, 2004: 254). Algunas de las colocaciones se encuentran en una relación sintáctica específica con la palabra que es objeto de la norma, pero otras se asocian libremente. De acuerdo con Hanks (2004), estas afirmaciones recogen el perfil combinatorio de un nombre y dan cuenta

del modo en que suele ser usado. En base a la norma de *storm*, al *Word Sketch* de ‘libro’ y al análisis realizado, elaboramos la siguiente propuesta para la norma de ‘libro’:

<p>‘libro’</p> <ul style="list-style-type: none">• ¿QUÉ SE HACE CON UN LIBRO? Un libro se lee, escribe, compra, edita, publica, presenta, imprime, comenta, dedica, traduce, firma, abre, cierra, dona, consulta, adquiere• ¿QUÉ HACEN LOS LIBROS?: abordan, analizan, describen, enseñan, hablan, narran, ofrecen, proponen, recopilan, relatan• PARTES DE UN LIBRO: capítulo, página, portada, título, tapa• CONJUNTOS DE LIBROS: colección, lote• UN LIBRO ES DE / SOBRE / ACERCA DE [TEMA]: libro de historia, libro sobre arte, libro acerca de la flora• ¿QUÉ CONTIENE UN LIBRO? cuentos, relatos, poesía (un libro de cuentos)• ¿CÓMO SON LOS LIBROS? Los libros suelen ser grandes, voluminosos, interesantes, famosos, magníficos, maravillosos, antiguos, impresos, ilustrados• ¿CON QUÉ SE ASOCIAN LOS LIBROS? revistas, artículos, cuentos, películas, documentos, folletos, cuadernos, discos, periódicos, novelas, papeles, textos, ensayos, apuntes
--

Figura 4. Propuesta para la norma de ‘libro’.

Lo específico de los tipos complejos como ‘libro’ radica en que sus normas reflejarían la activación conjunta de los tipos que lo componen al recoger predicados que requieren ambos tipos semánticos y/o piezas léxicas que interactúan en la copredicación, dado que aquello que caracteriza a los tipos complejos es precisamente la selección o explotación conjunta de los tipos. En efecto, la norma de ‘libro’ incorpora variedad de predicados que requieren ambos tipos, como ‘publicar’, ‘editar’ y ‘adquirir’. Asimismo, incorpora piezas léxicas que entran en interacción en la copredicación, como ‘abrir’, ‘cerrar’, ‘tapa’, ‘voluminoso’, por un lado, y ‘título’, ‘sobre’/ ‘de’/ ‘acerca de’ [TEMA], por otro. Además, las piezas léxicas con las que se asocian, como ‘película’ y ‘documento’, exhiben los mismos tipos semánticos (información. objeto físico). Por otra parte, nombres con un grado de polisemia mayor tendrían, en realidad, normas diferentes, dado que los tipos semánticos se encuentran en competencia más que en colaboración. Así, a partir del análisis presentado, es posible sostener que ‘cerdo’ tiene dos normas, una asociada al sentido de ‘animal’ y otra al de ‘carne’, ya que son escasas las piezas léxicas que requieren ambos tipos y las piezas léxicas que entran en interacción en la copredicación. Así, como se observa en la Figura 5, la mayoría de las piezas léxicas requieren un tipo semántico y no el otro; solo se registran como excepción las palabras que se asocian con los tipos semánticos animal y carne, que constituyen cohipónimos de ambos sentidos aunque existen excepciones (como por ejemplo, ‘ave de corral’, ‘oveja’, ‘ternero’, que son cohipónimos de ‘cerdo’ solo en el sentido de ‘animal’, y ‘pescado’, cohipónimo de ‘cerdo’ solo en el sentido de ‘carne’)²:

‘cerdo’

1. Implicatura: ‘animal’ (contable)

- ¿QUÉ SE HACE CON LOS CERDOS? Se los cría, faena, sacrifica, castra, desteta, alimenta, ceba, engorda, mata, clona
- ¿QUÉ HACEN LOS CERDOS? gruñen, engordan, comen, chillan, ingieren, consumen, sudan, pastan, deambulan, pasean, roban, matan, duermen, gritan, habitan, pesan, mueren, atacan, andan
- ¿CON QUÉ SE ASOCIAN LOS CERDOS? ganado, aves de corral, lechón, pollo, ave, vaca, ternero, gallina, cordero, oveja, cabra, jabalí, pavo, conejo, perro, pato, caballo
- ¿CÓMO SON LOS CERDOS? lactantes, salvajes, cimarrones, neonatos
- CONJUNTO DE CERDOS: piara
- LUGARES EN LOS QUE ESTÁN LOS CERDOS: criadero, granja

2. Implicatura: ‘carne’ (no contable)

- ¿QUÉ SE HACE CON LA CARNE DE CERDO? Se la asa, ahuma, adoba, deshuesa, cuece, come, cocina, guisa, glasea, hornea, fríe
- ¿CON QUÉ SE ASOCIA LA CARNE DE CERDO? pescado, lechón, pollo, cordero, cabra, jabalí, pavo, conejo, pato
- ¿CÓMO ES LA CARNE DE CERDO? rellena, agridulce, afrodisíaca, magra, grasa, tierna, rica
- PARTES Y PRODUCTOS DE LA CARNE DE CERDO: manteca (de cerdo), lomo (de cerdo), costillas (de cerdo), solomillo, chuleta, grasa (de cerdo), pierna, manitas, costillar, pata (de cerdo), jamón, oreja (de cerdo), bondiola, hígado (de cerdo), pernil, chicharrón

Figura 5. Propuesta para la norma de ‘cerdo’.

En el caso de voces como ‘hospital’, la norma recoge una variedad de predicados y modificadores que seleccionan el tipo complejo completo (como ‘inaugurar’, ‘equipar’, ‘colapsar’, ‘abandonar’, ‘fundar’, y ‘municipal’). En cambio, la norma de ‘almuerzo’ recoge pocos predicados que seleccionan ambos tipos (‘compartir’), lo que da cuenta de que la ambigüedad rara, pero su unicidad queda justificada por la existencia de copredicación frecuente:

‘hospital’

- ¿QUÉ SE HACE CON UN HOSPITAL? Se lo inaugura, construye, visita, equipa, recorre, abandona, edifica, funda, remodela, dota, traslada, bombardea, administra
- ¿QUÉ HACE UN HOSPITAL? atiende, funciona, colapsa, brinda, recibe, provee, acoge
- ¿QUÉ TIPOS DE HOSPITALES HAY? universitarios, psiquiátricos, clínicos, pediátricos, infantiles, maternos, generales; públicos, militares; locales, municipales, regionales, zonales, provinciales
- ¿CON QUÉ SE ASOCIAN LOS HOSPITALES? clínicas, centros, escuelas, consultorios, sanatorios, colegios, cárceles, asilos, universidades, laboratorios, policlínicos, postas, farmacias, maternidades, centros, dispensarios
- ¿CON QUÉ PERSONAL CUENTA UN HOSPITAL? enfermeros, pacientes, farmacéuticos, trabajadores, personal, directores
- ¿QUÉ PARTES TIENE UN HOSPITAL? enfermería, maternidad, quirófano, terapia, pasillos, habitaciones, farmacia

Figura 6. Propuesta de norma para ‘hospital’.

‘almuerzo’

- ¿QUÉ SE HACE CON UN ALMUERZO? Se lo sirve, comparte, degusta, prepara, ameniza, cocina, ofrece, disfruta, organiza, come, saborea, celebra, empaca, improvisa, programa, prepara, brinda
- UN ALMUERZO congrega, transcurre, dura, termina, comienza, cuesta, consiste, finaliza
- UN ALMUERZO ES campestre, succulento, criollo, delicioso, frugal, exquisito opíparo, sabroso, liviano, fraterno, multitudinario, gratis, dominguero, ameno, dominical, navideño, típico, campero, comunitario, apetitoso, ligero, tradicional, rico, excelente
- ¿CON QUÉ SE ASOCIA UN ALMUERZO? cena, desayuno, merienda, refrigerero, comida, colación descanso, siesta, break, visita, café, cóctel, reunión, baile, buffet, brindis, misa, lunch, sobremesa

Figura 7. Propuesta para la norma de ‘almuerzo’.

Es posible construir el perfil combinatorio de un nombre polisémico tomando en cuenta su grado de polisemia a partir de una norma que dé cuenta de la activación conjunta de los tipos que lo componen, o bien de más de una norma. Según esta postulación, las normas de los tipos complejos como ‘libro’, ‘pinar’, ‘escuela’, ‘hospital’, ‘cena’ y ‘almuerzo’ implicarían predicados que requieran ambos tipos e indicadores que entren en interacción en la copredicación. Por otra parte, aquellos nombres que no muestran ni un alto porcentaje de contextos ambiguos ni copredican con facilidad, como ‘acuarela’, ‘botella’, ‘cerdo’ y ‘vals’ (y que por lo tanto no son tipos complejos) tendrían diferentes normas según los significados que presentan.

El análisis realizado implica una extensión de la noción de norma a piezas léxicas que no tienen una estructura argumental y exhiben diferentes grados de polisemia. En principio, se postula que el grado de polisemia repercute en el establecimiento de una norma, dado que los contextos ambiguos y la copredicación –más precisamente, los predicados que se ven involucrados– son parte del comportamiento combinatorio de la palabra y dan cuenta de su estructura semántica.

CONCLUSIONES

El trabajo ha permitido establecer una definición de tipos complejos en base a su comportamiento contextual, a partir de dos variables que no necesariamente se dan a la vez a) la ambigüedad; b) la copredicación. Esta definición está construida en base al análisis de corpus. Por otra parte, la redefinición realizada está en consonancia con las nuevas versiones de los mecanismos generativos, que toman en cuenta la noción de costo de procesamiento según el elemento de la estructura semántica que se active, o introduzca, en contexto.

Por otra parte, se ha demostrado que es posible articular el concepto de tipo complejo con el de norma para la categoría nominal. Determinar específicamente qué es un tipo complejo implica establecer si un ítem léxico tendrá más de una norma (polisemia) o si tendrá una sola norma, que dé cuenta del comportamiento colocacional respecto de sus tipos semánticos (polisemia inherente). Además, se ha demostrado que las variables copredicación y ambigüedad son relativamente

independientes. En este trabajo, la ambigüedad ha demostrado ser la variable más potente.

El análisis realizado ha sido fundamentalmente de orden metodológico, pero tiene un impacto teórico. En principio supone que la estructura semántica se justifique mediante un análisis distribucional. Por otra parte, una buena metodología permitirá, por ejemplo, determinar en un futuro si los tipos complejos se comparten o no entre lenguas, plantear nuevas preguntas relativas al anisomorfismo de las lenguas y aportar a la distinción entre sentido, significado y vaguedad.

REFERENCIAS BIBLIOGRÁFICAS

- Adelstein, A. & Berri, M. (2013). Generación y desambiguación de sentidos en el ámbito nominal: Aportes al Lexicón Generativo. *Revista Filología*, XLV, 95-114.
- Asher, N. & Pustejosky, J. (2013). A type composition Logic for Generative Lexicon. En J. Pustejovsky, P. Bouillon, H. Isahara, K. Kanzaki & C. Lee (Eds.), *Advances in Generative Lexicon Theory* (pp. 39-66). Nueva York / Londres: Springer.
- Berri, M. (2014). *Polisemia regular en nombres locativos del español*. Tesis doctoral, Universidad de Buenos Aires, Buenos Aires, Argentina.
- Batiukova, O. (2006). *Del léxico a la sintaxis: Aspecto y qualia en la gramática del ruso y del español*. Tesis doctoral, Universidad Autónoma de Madrid, Madrid, España.
- Brandtner, R. (2011). *Deverbal nominals in context: Meaning variation and copredication*. Tesis doctoral, Universidad de Stuttgart, Stuttgart, Alemania.
- Croft, W. & Cruse, D. A. (2008). *Lingüística cognitiva*. Madrid: Akal.
- Cruse, D. A. (2000). Lexical 'facets': Between monosemy and polysemy. En S. Beekmann, P. König & T. Wolf (Eds.), *Sprachspiel und Bedeutung: Festschrift für Franz Hundsnurscher zum 60 Geburtstag* (pp. 25-36). Tübingen: Max Niemeyer Verlag.
- De Miguel, E. (2009). La teoría del Lexicón Generativo. En E. De Miguel (Ed.), *Panorama de la lexicología* (pp. 487-516). Barcelona: Ariel.
- Edmonds, P. (2006). Disambiguation, lexical. En K. Brown (Ed.), *Encyclopedia of Language and Linguistics* (pp. 607-623). Nueva York: Elsevier.
- Geeraerts, D. (1993). Vagueness's puzzles, polysemy's vagaries. *Cognitive Linguistics*, 4, 229-272.

- Hanks, P. (2004). The syntagmatics of metaphor and idiom. *International Journal of Lexicography*, 17(3), 245-274.
- Hanks, P. (2013). *Lexical analysis. Norms and exploitations*. Cambridge: The MIT Press.
- Hanks, P. & Pustejovsky, J. (2005). A pattern Dictionary for Natural Language Processing. *Revue Française de linguistique appliquée*, 10(2), 63-82.
- Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P. & Suchomel, V. (2014). The sketch engine: Ten years on. *Lexicography*, 1(1), 7-36.
- Kilgarriff, A. & Renau, I. (2013). esTenTen, a Vast Web Corpus of Peninsular and American Spanish. *Procedia -Social and Behavioral Sciences*, 95, 12-19.
- Jezek, E. & Melloni, C. (2011). Nominals, polysemy, and co-predication. *Journal of Cognitive Science*, 12, 1-31.
- Pustejovsky, J. (1995). *The Generative Lexicon*. Cambridge: MIT Press.
- Pustejovsky, J. (1996). *The semantics of Complex Types*. Conferencia presentada en VIII European Summer School in Logic, Language and Information (ESSLLI' 96), Praga, República Checa.
- Pustejovsky, J. (2001). Type construction and the Logic of Concepts. En P. Bouillon & F. Busa (Eds.), *The language of word meaning* (pp. 91-123). Cambridge: Cambridge University Press.
- Pustejovsky, J. (2005). *A survey on dot objects* [en línea]. Disponible en: <https://pdfs.semanticscholar.org/56df/791fe9e11e3e0b860c0e8930f05a5b740f4c.pdf>
- Pustejovsky, J. (2008). From concepts to meaning. The role of lexical knowledge. En P. van Sterkenburg (Ed.), *Unity and Diversity of Languages* (pp. 73-84). Amsterdam/ Philadelphia: John Benjamins.
- Pustejovsky, J. (2011). Coercion in a general theory of argument selection. *Linguistics*, 49(6), 1401-1431.
- Pustejovsky, J., Bouillon, P., Isahara, H., Kanzaki, K. & Lee, C. (Eds.) (2013). *Advances in Generative Lexicon Theory*. Nueva York / Londres: Springer.
- Pustejovsky, J. & Jezek, E. (2009). Semantic coercion in language: Beyond distributional analysis. *Italian Journal of Linguistics*, 20(2), 181-214.
- Pustejovsky, J. & Jezek, E. (2016). *Integrating Generative Lexicon and Lexical Semantic Resources* [en línea]. Disponible en: http://lrec2016.lrec-conf.org/media/filer_public/2016/05/10/tutorialmaterial_pustejovsky.pdf

Real Academia Española, Banco de datos (CORPES XXI) [en línea]. *Corpus del Español del Siglo XXI (CORPES)*. Disponible en: <http://www.rae.es>

Renau, I. & Nazar, R. (en proceso). *Verbario* [en línea]. Disponible en: <http://www.verbario.com>

Rumshisky, A., Grinberg, V. & Pustejovsky, J. (2007). *Detecting selectional behavior of complex types in text*. Conferencia presentada en 4th International Workshop on Generative Lexicon, Ministère de la Recherche, París, Francia.

Sketch Engine. [en línea] Disponible en: <https://www.sketchengine.co.uk/>

NOTAS

¹ Los atributos convencionalizados (Pustejovsky & Jezek, 2016) son propiedades típicamente asociadas a un objeto a partir de las modalidades perceptuales, por ejemplo, el hecho de que los perros ladran, de que el agua fluye o de que los aviones hacen ruido. De acuerdo con los autores, los atributos convencionalizados pueden activarse en contextos específicos, como ‘oyeron el perro’ (sonido = ladrido), de un modo similar a la información contenida en la estructura de qualia. Estos eventos están asociados convencionalmente con una entidad, pero no se consideraban parte de la estructura de qualia en versiones anteriores del modelo. En la nueva propuesta, se supone que las actividades asociadas se clasifican según una escala de convencionalidad. Solo las más convencionales se codificarían como atributos convencionalizados.

² Véase Adelstein y Berri (2013) y Berri (2014) para un desarrollo acerca de cómo la presencia de cohipónimos favorece la activación de un tipo semántico.