



UNIVERSITAT_{DE}
BARCELONA

**Significado, distribución y frecuencia
de la categoría preposicional en español.
Una aproximación computacional**

Francesc Reina González



Aquesta tesi doctoral està subjecta a la llicència **Reconeixement 4.0. Espanya de Creative Commons.**

Esta tesis doctoral está sujeta a la licencia **Reconocimiento 4.0. España de Creative Commons.**

This doctoral thesis is licensed under the **Creative Commons Attribution 4.0. Spain License.**



UNIVERSITAT DE
BARCELONA

Significado, distribución y frecuencia de la
categoría preposicional en español
Una aproximación computacional

Tesis doctoral

Sr. Francesc Reina González

Departament de Filologia Catalana i Lingüística General
Facultat de Filologia i Comunicació. Universitat de Barcelona
Programa de doctorado *Ciencia Cognitiva y Lenguaje*

Directores

Dra. Irene Castellón Massalles (Universitat de Barcelona)
Dr. Lluís Padró Cirera (Universitat Politècnica de Catalunya)

Enero de 2020

RESUMEN

La categoría preposicional ha sido, tradicionalmente, una clase de palabra provista de rasgos lingüísticos y conductas gramaticales controvertidas. En la tesis, *Significado, distribución y frecuencia de la categoría preposicional del español. Una aproximación computacional*, se examina la naturaleza de esa controversia a la luz de una metodología cuantitativa, computacional y de lingüística de corpus. La brecha más inexplicada en la historia de su análisis gramatical ha sido cómo identificar su significado.

Con frecuencia las nociones descriptivas sobre la semántica de la preposición han sido observadas como subsidiarias de su papel sintáctico, vía caso o asignación de papeles temáticos. Este hecho, sin embargo, no es más que el reconocimiento de que su significado incide, también, en su tarea global dentro de la gramática.

Desde una concepción neodistribucionalista, según la cual el significado de las piezas o elementos lingüísticos se encuentra en su distribución contextual, la hipótesis que se plantea es que esa expresión semántica de las preposiciones en español se produce de manera gradual. La denominada Hipótesis Gradual del Significado, aplicada a la categoría preposicional en español, nos permite establecer subclases preposicionales, desde la funcionalidad hasta la lexicidad, pasado por clases intermedias como semifuncionales y semiléxicas.

La justificación empírica de la Hipótesis Gradual del Significado se realiza a partir de cuatro experimentos.

El primero de ellos se inserta en la metodología del aprendizaje automático (*machine learning*). Para ello, y usando la técnica del *clustering*, observamos un conjunto de 79.097 tripletas de la forma X – P – Z, donde P es una preposición del español –basadas en sintagmas preposicionales complementos-. Estas tripletas dependen de una serie de verbos de movimiento del español para las preposiciones *a*, *hacia* y *hasta* extraídas de cuatro corpus lingüísticos muy reconocidos del español. Una vez obtenidas las agrupaciones automáticas se evalúan porcentualmente a partir de la coincidencia entre las predicciones del anotador humano –las clases preposicionales sugeridas- y de la máquina –los clusters-.

En el segundo y el tercer experimentos utilizamos otra metodología y acudimos a la medición de la entropía –magnitud de la Teoría de la Información-. En el segundo clasificamos los nombres de 3.898 tripletas que dependen de una serie de verbos del

español que representan la mayoría de campos semánticos; y en el tercero son 3903 que complementan a otros nombres.

Esta clasificación de los nombres se realiza a partir de una propuesta de seis tipos de categorías semánticas: *Animado*, *Inanimado*, *Entidad abstracta*, *Locativo*, *Temporal* y *Evento*. Una vez clasificados los nombres se mide su organización entrópica y se verifica que existe una correlación entre el grado de entropía y la clase preposicional: a mayor entropía mayor significado.

El cuarto experimento parte del uso preposicional. A partir de un test de 90 respuestas con las clases preposicionales de la hipótesis se recogen las respuestas de 366 participantes y se analiza el grado de variación de esas respuestas según la clase preposicional. De nuevo volvemos a usar la entropía como índice de identificación del significado. Sometemos los resultados a ciertas pruebas de control estadístico para verificar la fiabilidad de las muestras, de la significancia y de la coincidencia entre observadores (coeficiente kappa de Cohen).

El balance de los cuatro experimentos –a través de los resultados- es favorable a la predicción de la hipótesis. Asimismo, la diversidad de herramientas de análisis es una forma metodológicamente robusta para la investigación y sus conclusiones.

Finalmente, se indica que la hipótesis abre perspectivas de futuro en áreas como el contraste interlingüístico –de familias tipológicamente diversas en la expresión adposicional-, o la afasiología como disciplina que se interroga por las relaciones entre errores y valores gramaticales.

ABSTRACT

The prepositional category has traditionally been a word group endowed controversial traits concerning both its linguistic features and its grammatical behaviors. In this thesis the controversy is examined from a quantitative, computational and linguistic methodology point of view. The most unexplained gap in the story of this difficulty of analysis lies in the fact of how its meaning can be identified.

From a neo-distributionalism conception, according to which the meaning of the linguistic pieces is in their contextual distribution, the hypothesis that arises is that this semantic expression of the prepositions in Spanish is gradual. The so-called *Gradual Meaning Hypothesis* establishes four prepositional subclasses, from functional to lexical, through intermediate phases such as semi-functional and semi-lexical.

The empirical justification of the *Gradual Hypothesis of Meaning* is performed with four experiments.

The first one experiment is inserted in the machine learning methodology. To do this, and using the clustering technique, we observed a set of 79,097 triplets of the form X - P - Z, where P is a preposition of Spanish - based on complementary prepositional phrases. They are triplets with the prepositions *a*, *hacia* and *hasta* of movement verbs, and they are extracted from four well-known linguistic corpus of Spanish. Once the automatic groupings have been obtained, we indicate to what extent, the percentage between the predictions of the human scorer - the suggested prepositional classes - and the machine - the clusters - are confirmed.

In the second and third experiments, we changed our methodology and turned to the measurement of entropy –variable in Information Theory-. In the second one we classify the names of 3,898 triplets that depend on verbs that appear on most semantic fields in Spanish; and in the third one we classify 3903 triplets that complement other names.

This name classification is based on a proposal of six types of semantic categories: *Animate*, *Inanimate*, *Abstract Entity*, *Locative*, *Temporary* and *Event*. Once the names are classified, their entropic organization is measured and it is verified that there is a correlation between the degree of entropy and the prepositional class: the greater the entropy, the greater the meaning.

The fourth experiment starts with prepositional use. From a test the degree of variation of these responses is analyzed according to the prepositional class. Again we use entropy as an index of identification of meaning.

The balance of the four experiments - through the results - is favorable to the prediction of the hypothesis. The diversity of analysis tools is a methodologically robust way for the research and its conclusions.

Agradecimientos

Empecé a interesarme por la investigación lingüística gracias a la profesionalidad docente que despertaba el añorado profesor Jesús Tuson en sus clases de Historia de la Lingüística. Fue él quien, además, nos habló de la Gramática Generativa y de la profesora encargada de un curso sobre esta visión del lenguaje humano y las lenguas, la Joana Rosselló. Con ella inicié mis pesquisas desde la perspectiva teórica sobre cómo funcionan unas palabras tan “pequeñas”, pero que resultan tan “pertinaces e imprescindibles”. El final de aquella primera etapa de estudio de las preposiciones fue, precisamente, con Ramon Ferrer-i-Cancho de la UPC, en el área de la lingüística cuantitativa. Con él elaboramos lo que, según creo, fue el germen de mi investigación: maneras de medir el significado de las partículas en el inglés.

Tras más de una década de diversos avatares que no vienen a cuento he vuelto al cauce preposicional gracias a la recomendación de Faust Diéguez, colega y amigo, hace ya más de treinta años; y a la generosidad de la que luego ha sido una de las directoras de la tesis, la Dra. Irene Castellón.

La Dra. Irene Castellón me propuso trabajar con el Dr. Lluís Padró, y ambos han sido mis directores durante estos últimos tres años. Reconozco en ellos virtudes complementarias: la exigencia investigadora, la generosidad personal y el trabajo constante. En el ámbito en el que ellos investigan, la ciencia computacional, he podido dirigir modestamente mis nuevos pasos empíricos en esta categoría lingüística. Además, tuve la oportunidad, gracias a Irene, de explicar mi hipótesis en el Workshop del GRIAL.

También ha querido el azar que en el último tramo de mi investigación haya podido impartir docencia como profesor asociado en el Departament de Filologia Catalana i Lingüística General de la Universidad de Barcelona, donde yo me licencié. A mis colegas

que me han hecho más fácil estos meses de tesis, por supuesto, gracias; en particular a mi compañera de despacho, Lourdes Romera; y mis amables y generosas vecinas, Maria Antònia Martí y Mariona Taulé.

En este tiempo, y por otras casualidades, también he podido compartir charlas con el lingüista y profesor Dr. Ángel J. Gallego del departamento de filología hispánica de la UAB, quién me presentó el ADIBA y me facilitó algunos de sus recursos. Fue él quién me permitió intercambiar observaciones y correos electrónicos con el lingüista, Dr. Juan Uriagereka. Asimismo he podido consultar cuestiones bibliográficas con el académico Dr. Ignacio Bosque. Gracias a los tres por su disposición y sus agudas observaciones.

Jaume Llopis, estadístico de la Facultad de Biología de la Universidad de Barcelona me recibió y me ayudó a entender algunos asuntos básicos de las demostraciones estadísticas. Asimismo, el amigo, profesor y físico Dr. Bartolo Luque me situó el valor de los cálculos estadísticos en un contexto científico más general. Mis deudas con él solo se pagan con un cálido afecto que compartimos desde hace casi cuarenta años.

También tuve oportunidad de comentar ciertos aspectos iniciales hace algunos años con el profesor retirado Henk Van Riesdijk, y recientemente Ray Jackendoff me recordó alguna bibliografía sobre preposiciones espaciales.

Igualmente, el profesor y colega del instituto, Santiago Arrondo, me ha ayudado con prontitud y claridad en la traducción de las citas del latín; y Núria Cáceres, también colega, en algunas correcciones del inglés.

Mi familia me ha acompañado, entendido y soportado en estos más de cien fines de semana de trabajo y otros tantos días de vacaciones, más otras fiestas de guardar. A los tres caminos generosos de mi corazón solo puedo expresarles mi reconocimiento.

Per vosaltres, Montserrat Ferrando, Albert Reina i Júlia Reina. Moltes gràcies per la vostra comprensió, el vostre suport i la vostra paciència.

Y acabo con la referencia, *in memoriam*, a mi padre, Francisco Reina Estrada, que siempre quiso que fuésemos mejores contra cualquier adversidad y sin tregua alguna; y a mi madre, Rosa González, que sigue alegrándose de lo que alegra a los suyos.

Sin todos ellos, sin todos vosotros no hubiese conseguido acabar mi trabajo, y es tan justo como necesario dejarlo por escrito. Gracias, de todo corazón. *Gràcies, col·legues i amics, de tot cor.*

ÍNDICE

Resumen.....	ii
Abstract.....	iv
Agradecimientos.....	v
Índice.....	vii
Lista de cuadros, gráficos, tablas, figuras y esquemas.....	xii
INTRODUCCIÓN.....	1
CAPÍTULO 1. HISTORIA, DESCRIPCIÓN Y EXPLICACIÓN DE LA CATEGORÍA PREPOSICIONAL.....	4
1.1 Objetivos del capítulo	4
1.2 De la antigüedad grecolatina hasta el siglo XIX: gramática general y gramáticas del español	5
1.3 Algunas consideraciones de la lingüística general del siglo XX: de Edward Sapir a Bernard Pottier	16
1.4 Gramáticas, estudios y propuestas descriptivas y normativas para la preposición española durante el siglo XX	20
1.5 La teoría lingüística y la preposición	32
1.5.1 La hegemonía de la teoría sintáctica. De Chomsky (1965) a Emonds (1985)	33
1.5.2 Sobre las categorías funcionales y el problema de la correspondencia entre sintaxis y léxico	37
1.5.3 La gradualidad categorial: un camino hacia el significado	42
1.6 Contrastes y evidencias interlingüísticas	48
1.7 Perspectivas en el estudio del significado de las preposiciones: de los usos concretos a las teorías cognitivas	52

1.7.1 La semántica descriptiva y/o composicional	53
1.7.2 La semántica formal	56
1.7.3 La gramática cognitiva: prototipos y espacio	57
1.8 A modo de síntesis	60
CAPÍTULO 2. ASUNCIONES DESCRIPTIVAS, TEÓRICAS Y METODOLÓGICAS.....	62
2.1 Justificación y formulación de la hipótesis	62
2.2 Objetivos de la investigación	64
2.3 Descripción y ejemplificación del objeto de análisis	65
2.3.1 El conjunto de piezas preposicionales en español	66
2.3.2 Los valores léxico-semánticos de las preposiciones en español. Relación y justificación de las clases preposicionales	69
2.4 Marco teórico y metodológico	78
2.4.1 El marco computacional	78
2.4.2 Asunciones semánticas	80
2.4.3 Herramientas de análisis y procesamiento	87
2.4.3.1 El aprendizaje automático.....	87
2.4.3.2 La teoría de la información y la entropía.....	89
2.4.4 Procedencia de los datos lingüísticos: los corpus	91
2.4.4.1 Composición de los corpus.....	92
CAPÍTULO 3. LA GRADUALIDAD SEMÁNTICA PREPOSICIONAL. APRENDIZAJE AUTOMÁTICO, ENTROPÍA Y USO	95
3.1 PRIMER EXPERIMENTO. La gradualidad semántica preposicional a través del aprendizaje automático (<i>clustering</i>). Un universo específico: los verbos de movimiento en español.....	95

3.1.1	Objetivos y justificación del experimento	95
3.1.2	Clases preposicionales y valores semánticos de <i>a, hacia y hasta</i>	97
3.1.3	Procedimientos: el <i>clustering</i> en el aprendizaje automático	102
3.1.4	Datos lingüísticos	106
3.1.5	Experimento	106
3.1.6	Análisis de los resultados	108
3.1.6.1	La clase funcional.....	110
3.1.6.2	La clase semifuncional	121
3.1.6.3	La clase léxica.....	126
3.1.7	Conclusiones parciales	131
 3.2. SEGUNDO EXPERIMENTO. Entropía y significado preposicional en el ámbito verbal		134
3.2.1	Objetivos y justificación del experimento	134
3.2.2	Las clases de la gradualidad semántica preposicional.....	135
3.2.3	Selección y procedencia de los datos.....	136
3.2.4	Metodología: el cálculo de la entropía y las categorías semánticas.....	143
3.2.5	La realización del experimento	146
3.2.6	Análisis de los resultados	148
3.2.6.1	La clase funcional.....	153
3.2.6.2	La clase semifuncional.....	156
3.2.6.3	La clase léxica.....	159
3.2.7	Conclusiones parciales.....	162
 3.3. TERCER EXPERIMENTO. Entropía y significado preposicional en el ámbito nominal		166
3.3.1	Objetivos y justificación del experimento	166

3.3.2 Los valores léxico-semánticos en estructuras nominales.....	167
3.3.3 Procedencia de los datos y procedimientos: entropía y categorías semánticas.....	171
3.3.4 La realización del experimento	172
3.3.5 Análisis de los resultados	173
3.3.5.1 La clase funcional.....	174
3.3.5.1.1 Frecuencia, polisemia y significado. El caso de la preposición <i>de</i>	179
3.3.5.2 La clase semiléxica.....	183
3.3.5.3 La clase léxica.....	187
3.3.6 Conclusiones parciales.....	193
3.3.7 Comparación de resultados entre estructuras verbales (experimento segundo) y estructuras nominales (experimento tercero).....	197

**3.4 CUARTO EXPERIMENTO. Los grados semánticos a partir del uso
de las preposiciones** 201

3.4.1 Objetivos y justificación del experimento.....	201
3.4.2 El uso de las preposiciones: de la ambigüedad al significado	202
3.4.3 Metodología: diseño y características del test de uso preposicional	204
3.4.3.1 Procedencia de los datos, procedimiento de realización del test y rasgos de las participantes.....	206
3.4.3.2 El test: información sobre las respuestas.....	208
3.4.4 Descripción de los resultados del test	212
3.4.5 Análisis de los resultados: acuerdo, dispersión y entropía	227
3.4.5.1 La clase funcional	236
3.4.5.2 Las clases semifuncional y semiléxica	239
3.4.5.3 La clase léxica.....	242

3.4.6 Consideraciones estadísticas sobre los resultados del test	245
3.4.7 Conclusiones parciales	252
CAPÍTULO 4. Conclusiones y perspectivas futuras.....	254
Referencias y bibliografía.....	268
Apéndice.....	315
A. Tabla de datos del test del cuarto experimento	315
B. Tabla de cálculo de la frecuencia teórica para el χ^2	318
C. Cuadro de respuestas comparativas entre los participantes que han repetido en dos momentos distintos (para calcular el coeficiente kappa de Cohen).....	319
D. Tablas de contingencia de las comparaciones entre 11 participantes que han respondido en dos ocasiones (para calcular el coeficiente kappa de Cohen).....	321
E. Tablas de datos sobre el coeficiente kappa de Cohen para la comparación de 11 participantes en dos momentos distintos.....	336

Listas de cuadros, gráficos, tablas, ilustraciones y esquemas
(Por orden de páginas)

Cuadros

Cuadro 1. Contribuciones gramaticales al estudio preposicional desde la antigüedad grecolatina hasta el siglo XIX.....	15
Cuadro 2. Comparación entre las propuestas de partículas de Pavón (2003).....	29
Cuadro 3. Selección de aportaciones de la lingüística contrastiva relacionadas con la preposición	51
Cuadro 4. Valores léxico-semánticos de las preposiciones en español.....	70
Cuadro 5. Organización de las clases preposicionales por experimentos.....	75
Cuadro 6. Disposición de rasgos y subdivisiones del Top Concept Ontology (TCO)....	84
Cuadro 7. Síntesis de correspondencias entre las categorías lingüísticas, Top Concept Ontology (TCO) y las categorías semánticas propuestas	86
Cuadro 8. Descripción y ejemplos de los valores semánticos de <i>a</i>	99
Cuadro 9. Descripción y ejemplos de los valores semánticos de <i>hacia</i>	99
Cuadro 10. Descripción y ejemplos de los valores semánticos de <i>hasta</i>	100
Cuadro 11. Organización de las clases preposicionales por experimentos.....	101
Cuadro 12. Simplificación de la agrupación automática para la preposición <i>a</i>	110
Cuadro 13. Tipos y ejemplos de nombres de expresión espacial en clusters de <i>a</i>	117
Cuadro 14. Tipos y ejemplos de nombres de expresión espacial en clusters de <i>hacia</i> ...	123

Cuadro 15. Expresión de la ubicación y la localización en el espacio del clúster léxico de <i>hasta</i>	128
Cuadro 16. Organización de las clases preposicionales por experimentos.....	135
Cuadro 17. Clasificación de verbos procedentes de http://adesse.uvigo.es/data/clases.php	138
Cuadro 18. Distribución de subdivisiones en la clase funcional para la categoría semántica <i>Animado</i>	154
Cuadro 19. Distribución de subdivisiones para la categoría semántica de <i>Entidad abstracta</i>	157
Cuadro 20. Subdivisiones de la categoría semántica de <i>Locativo</i>	160
Cuadro 21. Valores léxico-semánticos de las preposiciones en estructuras nominales...	170
Cuadro 22. Organización de las clases preposicionales por experimentos.....	171
Cuadro 23. Subdivisiones de la categoría semántica <i>Animado</i> de la clase funcional	175
Cuadro 24. Ilustración de equivalencias nominal-verbal en la clase funcional con <i>de-a</i>	176
Cuadro 25. Ejemplificación de rasgos del Top Concept Ontology (TCO) para la categoría semántica <i>Inanimado</i>	178
Cuadro 26. Subdivisiones de la categoría semántica <i>Animado</i> en la clase semiléxica	185
Cuadro 27. Subdivisiones de la categoría semántica <i>Entidad abstracta</i> en la clase semiléxica.....	186
Cuadro 28. Subdivisiones de la categoría semántica <i>Locativo</i> en la clase léxica.....	190
Cuadro 29. Ejemplos de las subdivisiones de la categoría semántica <i>Temporal</i> en la clase léxica.....	192

Cuadro 30. Descripción de las variables sociolingüísticas de los participantes del test...	207
Cuadro 31. Correlaciones observadas entre los resultados del test y la hipótesis.....	228

Gráficos

Gráfico 1. Comparación porcentual entre las tres preposiciones del experimento.....	101
Gráfico 2. Porcentajes de pureza (azul), y pureza inversa (marrón) por el número de clúster para la preposición <i>a</i>	111
Gráfico 3. Porcentajes de valores semánticos en clases y agrupaciones de <i>a</i> . Clúster 0	114
Gráfico 4. Porcentajes de valores semánticos en clases y agrupaciones de <i>a</i> . Clúster 1	119
Gráfico 5. Porcentajes de valores semánticos en clases y agrupaciones de <i>a</i> . Clúster 2	120
Gráfico 6. Porcentajes de valores semánticos en clases y agrupaciones de <i>a</i> . Clúster 3	120
Gráfico 7. Porcentajes de valores semánticos del clúster 0 de <i>hacia</i>	124
Gráfico 8. Porcentajes de valores semánticos del clúster 1 de <i>hacia</i>	125
Gráfico 9. Porcentajes de valores semánticos del clúster 2 de <i>hacia</i>	125
Gráfico 10. Porcentajes de valores semánticos del clúster 0 de <i>hasta</i>	129
Gráfico 11. Porcentajes de valores semánticos del clúster 1 de <i>hasta</i>	130
Gráfico 12. Porcentajes de valores semánticos del clúster 2 de <i>hasta</i>	130
Gráfico 13. Entropía para las clases preposicionales según las categorías semánticas...150	
Gráfico 14. Distribución porcentual de categorías semánticas de la clase funcional (verbal).....	153
Gráfico 15. Distribución porcentual de categorías semánticas de la clase semifuncional (verbal).....	156
Gráfico 16. Distribución porcentual de categorías semánticas de la clase léxica (verbal).159	

Gráfico 17. Comparación de las categorías semánticas por clases preposicionales (verbal).....	162
Gráfico 18. Distribución porcentual de las categorías semánticas en la clase funcional (nominal).....	175
Gráfico 19. Distribución porcentual de las categorías semánticas en la clase semiléxica (nominal).....	183
Gráfico 20. Distribución porcentual de las categorías semánticas en la clase léxica (nominal).....	181
Gráfico 21. Comparación de las categorías semánticas por clases preposicionales (nominal).....	193
Gráfico 22. Comparación de las categorías semánticas para las tres clases preposicionales (verbal)	199
Gráfico 23. Comparación de las categorías semánticas para las tres clases preposicionales (nominal)	199
Gráfico 24. Porcentaje de acuerdo y acierto en las preposiciones funcionales.....	224
Gráfico 25. Porcentaje de acuerdo y acierto en las preposiciones semifuncionales.....	225
Gráfico 26. Porcentaje de acuerdo y acierto en las preposiciones semiléxicas.....	226
Gráfico 27. Porcentaje de acuerdo y acierto en las preposiciones léxicas.....	227
Gráfico 28. Comparación entre las cuatro clases preposicionales de las cuatro partes del test	229
Gráfico 29. Comparación de la acuerdo de las clases preposicionales.....	229
Gráfico 30. Promedio de dispersión por respuestas en las cuatro partes y en las cuatro clases	231
Gráfico 31. Comparación de la dispersión global por clases preposicionales.....	232
Gráfico 32. Resultados de entropía de la dispersión de las cuatro partes del test.....	234
Gráfico 33. Resultados de entropía de la dispersión del test completo.....	235
Gráfico 34. Comparación de purezas por clases preposicionales para la <i>a</i>	256
Gráfico 35. Comparación porcentual de purezas por clases para <i>hacia</i>	258

Gráfico 36. Comparación porcentual de purezas por clases para <i>hasta</i>	259
Gráfico 37. Comparación de entropías en los ámbitos nominal y verbal.....	261
Gráfico 38. Comparación porcentual de acuerdo entre el modelo y los participantes...	262
Gráfico 39. Medida de la entropía de la dispersión en las cuatro clases preposicionales	263

Tablas

Tabla 1. Agrupación en 4 clusters y 4 valores de la preposición <i>a</i>	112
Tabla 2. Porcentajes de clases semánticas del clúster semiléxico correspondiente a la preposición <i>a</i>	115
Tabla 3. Agrupación en 3 clusters y 3 valores de la preposición <i>hacia</i>	121
Tabla 4. Agrupación en 3 clusters y 3 valores de la preposición <i>hasta</i>	126
Tabla 5. Comparación de frecuencias de aparición de las siete preposiciones más comunes en el experimento con tres corpus representativos	142
Tabla 6. Clasificación de categorías semánticas y clases preposicionales (con repeticiones).....	148
Tabla 7. Clasificación de categorías semánticas y clases preposicionales (sin repeticiones).....	148
Tabla 8. Comparación entre los totales de entropía con repeticiones y sin repeticiones...	149
Tabla 9. Resultados de la prueba de azar (ámbito nominal)	152
Tabla 10. Clasificación de clases semánticas y clases preposicionales en el ámbito nominal.	173
Tabla 11. Ocurrencias de las preposiciones en la clase semiléxica.....	184
Tabla 12. Comparación de las categorías semánticas minoritarias de la clase funcional y semiléxica.....	186
Tabla 13. Ocurrencias de las preposiciones de la clase léxica.....	188
Tabla 14. Diferencias numéricas entre las categorías semánticas de la clase léxica.....	191

Tabla 15. Comparación de las diferencias por casos entre categorías semánticas de las tres clases preposicionales	191
Tabla 16. Resultados de entropía con criterios aleatorios de organización de las clasificaciones (ámbito nominal).....	195
Tabla 17. Comparación de resultados entrópicos en clases preposicionales en los ámbitos verbal y nominal	196
Tabla 18. Diferencia de entropía entre clases preposicionales y estructuras sintácticas	198
Tabla 19. Distribución de ocurrencias, valores y clases preposicionales en el test.....	205
Tabla 20. Parte 1 del test. Resultados de las respuestas identificadas por número, respuestas y porcentaje de acuerdo	215
Tabla 21. Parte 2 del test. Resultados de las respuestas identificadas por número, respuestas y porcentaje de acuerdo	217
Tabla 22. Parte 3 del test. Resultados de las respuestas identificadas por número, respuestas y porcentaje de acuerdo	220
Tabla 23. Parte 4 del test. Resultados de las respuestas identificadas por número, respuestas y porcentaje de acuerdo	222
Tabla 24. Ejemplo de tabla de cálculo de la entropía (preposiciones funcionales de la primera parte del test)	233
Tabla 25. Entropía de la dispersión en el uso de las clases preposicionales.....	233
Tabla 26. Algunos ejemplos representativos de dispersión de la clase léxica en el test...	244
Tabla 27. Tamaño de las muestras del test para cada clase preposicional en relación a tres estándares de error y confianza estadística.....	246
Tabla 28. Porcentajes de formación académica en España y correspondencia respecto del total en el test y con el 50 % de los datos reales de participantes del test	247
Tabla 29. Porcentajes de franjas de edad en España y correspondencia respecto del total en el test y con el 50 % de los datos reales de participantes del test.....	247
Tabla 30. Datos estadísticos para el cálculo del valor χ^2	248

Tabla 31. Coeficientes kappa (k) de Cohen para la comparación de resultados entre 11 participantes distintos	250
Tabla 32. Coeficientes kappa (k) de Cohen para la comparación de resultados de los mismos 11 participantes en dos momentos distintos	250
Tabla 33. Datos de agrupación por <i>clustering</i> 0, 1, 2 y 3 para la preposición <i>a</i>	256
Tabla 34. Datos de agrupación automática para <i>hacia</i> y <i>hasta</i>	257
Tabla 35. Resultados de la entropía de las clases preposicionales (ámbito verbal y nominal)	260

Figuras

Figura 1. La ubicación del aprendizaje automático en relación a otras áreas de conocimiento (Gómez, Jurista, Montes y Pazos, 1997).....	103
Figura 2. Cronología de los procedimientos y heurísticas sobre el aprendizaje automático.....	104

Esquemas

Esquema 1. Preposición <i>de</i> . Correspondencia entre su adscripción a las clases de preposiciones y sus valores léxico-semánticos.....	181
--	-----

1. INTRODUCCIÓN

La preposición, como clase de palabra y categoría lingüística, ha sido y es un objeto de investigación controvertido. La naturaleza de ese debate procede de los términos de su definición, de los rasgos de su caracterización, de la determinación de su significado en cualquiera de sus estructuras y contextos, y de sus relaciones con otras categorías.

Las pruebas de esos conflictos se resiguen en la historia de la gramática universal, no solo la del español; en las descripciones gramaticales y en buena parte de los desarrollos de la lingüística teórica desde mediados del siglo XX. De hecho, desde que tenemos memoria historiográfica, la ocupación de la lingüística sobre esta categoría ha mostrado interés en ese carácter difícil, y ha buscado las causas de esa complejidad.

Como vamos a ver, en la mayoría de perspectivas teóricas y desde metodologías lingüísticas muy dispares, el tratamiento de la preposición muestra ciertas incomodidades o encajes que suceden en su tarea funcional y en su valor léxico-semántico, simultáneamente.

La observación y el análisis de esos valores semánticos, en convivencia en la misma preposición como si fuesen grados, fases o clases en sus contextos sintácticos, ha sido el camino que hemos identificado como una oportunidad de sondeo de nuestra investigación.

La dirección de ese camino se gobierna, por asunción metodológica, hacia la identificación del significado preposicional en su distribución contextual. Esos contextos son observados en el seno de gigantescos almacenes de datos (corpus) y con las herramientas de la observación y la cuantificación que hacen posible el desarrollo y la verificación de generalizaciones empíricas. Se trata de instrumentos que se proceden de la Inteligencia Artificial (el aprendizaje automático) y la Teoría de la Información (la medida de la entropía). Además, también abordamos el uso de esa convivencia gradual a partir de las respuestas de un conjunto de usuarios o hablantes del español respecto de un modelo real, planteado en forma de test.

Esperamos que esa triple perspectiva, aprendizaje automático, medición de la entropía (orden u organización de los contextos de las preposiciones) y relación entre uso y significado, permita consolidar los pronósticos de nuestra hipótesis. Consideramos que el sentido heurístico de esa experimentación múltiple garantiza que la investigación apoya sus predicciones en fuentes empíricas complementarias y solidarias entre sí.

Nuestra hipótesis, en consecuencia, aborda el hecho de la convivencia de fenómenos funcionales con evidencias léxicas en la categoría a partir del concepto de gradualidad semántica. Pensamos que existe un universo de representaciones semánticas donde los valores de las preposiciones van desde el significado nulo, funcional o gramatical hasta un significado activo o léxico. La caracterización y la identificación de esos rasgos se producen en la diversidad de contextos de las piezas preposicionales. Esta perspectiva, conocida como neodistribucionalismo, abre los términos de la polémica sobre el significado preposicional y los sitúa en relación –no ya a prescripciones teóricas- sino al análisis de datos que justifican cómo medirlo (Boleda y Herbelot, 2017; y Baroni et alia, 2012).

Los resultados obtenidos en los cuatro experimentos de la investigación son examinados, por lo tanto, bajo el prisma de esa fluctuación natural que se produce en las realizaciones preposicionales para el español. Esa oscilación de lo que signifiquen las preposiciones a partir de lo que las rodea se puede reconocer como el punto de llegada de nuestra hipótesis, y, asimismo, como el punto de arranque en otras áreas, en otras iniciativas y perspectivas.

Creemos, así, que una de las direcciones que abre nuestra propuesta es el contraste interlingüístico (no en vano, una parte considerable de los cambios en el rumbo categorial del análisis de las preposiciones procede de su contraste con lenguas que poseen otras clases de categorías adposicionales como posposiciones o circunposiciones). Junto a la dirección interlingüística, otra fuente de datos que consideramos prometedora –en la medida que los trabajos hasta la actualidad así lo anticipan (Friederici, 1981; Grodzinsky, 1988, 1989 y 1990; Kaplan, 1993; Branchenau y Nespoulos, 1989; y Froud, 2001)- es la afasiología. En las conclusiones y las perspectivas indicaremos el valor de esas prometedoras opciones.

Adentrémonos, para empezar, en la historia de la controversia gramatical y lingüística donde hemos buscado una brecha de oportunidad en el significado preposicional. Y desde ahí, y así lo esperamos, iniciaremos la justificación de cómo se puede medir su gradualidad semántica y qué clases preposicionales se podrán proponer.

CAPÍTULO 1. HISTORIA, DESCRIPCIÓN Y EXPLICACIÓN DE LA CATEGORÍA PREPOSICIONAL

1.1 Objetivos del capítulo

La preposición ha sido una de las partes del habla, clases de palabras o categoría lingüística que más ha incrementado su estudio y, por tanto, su conocimiento gramatical en el último medio siglo. Hasta entonces su presencia entre las otras categorías de la gramática había sido tradicionalmente aceptada y reconocida, aunque con un carácter menos relevante o decisivo.

El objetivo fundamental de esta sección es, precisamente, realizar una crónica analítica hasta nuestros días de los términos en que se ha producido ese reconocimiento, y cuáles fueron los motivos que abrieron nuevos caminos para alcanzar un renovado protagonismo e interés en su estudio lingüístico.

Nuestra exposición se ordenará entorno a tres cuestiones que aglutinan la mayoría de obras, autores, escuelas y metodologías que la abordan. Las cuestiones se recogen entre (1) y (3).

(1) ¿Qué criterios definitorios se usan en la descripción y/o explicación sobre su definición? (o su intensión).

(2) ¿Qué estatus categorial posee y cuáles son las piezas que componen la clase? (o su extensión).

(3) ¿Cuáles son sus relaciones de semejanza con otras categorías gramaticales?

De este modo, la diversidad de épocas, autores y perspectivas, se conocerán a partir de estos tres focos de interés.

Tal y como veremos, a medida que nos aproximamos en el tiempo a ese periodo de un cierto renacimiento de su interés, el estudio y el análisis de las preposiciones (a partir de 1960, aproximadamente) se irá convirtiendo –de manera particular en las teorías sintácticas- en más problemático y controvertido.

Es, también, una selección comentada de aquellos hitos destacados, de las líneas de investigación, de los asuntos relevantes y los trabajos más útiles en la comprensión de las respuestas a las preguntas anteriores.

La sección se organiza en seis partes. Los dos primeros representan la historia clásica o tradicional sobre la categoría: desde la antigüedad grecolatina (1.2) hasta algunos ejemplos de lingüística general del siglo XX (1.3). Los otros cuatro versan sobre la gramática descriptiva del español del siglo XX (1.4), la teoría lingüística y sintáctica del siglo XX (1.5), los contrastes y las evidencias interlingüísticas (1.6), y tres perspectivas paradigmáticas del significado preposicional: la descriptiva, la formal y la cognitiva (1.7).

1.2 De la antigüedad hasta el siglo XIX: gramática general y gramáticas del español

En el periodo clásico de la historia del pensamiento lingüístico occidental, la historiografía (Koerner, 1978; Parret, 1976; y Hüllen, 1990) ha reconocido en el libro de Dionisio de Tracia, *Téchne Grammatiké*¹, una de las primeras presentaciones casi sistemática sobre las partes de la oración para una lengua natural. La definición que ofrece sobre la preposición se cita en (4).

1. Utilizo el texto que recoge Arens (1969): *La lingüística: textos desde la antigüedad hasta nuestros días*.

(4) *Parte de la oración que puede figurar delante de todas las categorías, lo mismo en composición que en asociación o sintaxis* (Arens, 1969, pág. 44).

Esta primera enunciación posicionalista será perpetuada en muchas de las consideraciones de los manuales de la antigüedad griega y latina: la preposición como una palabra que exige una posición determinada en la oración, o bien en su morfología, o bien en la construcción de varias de ellas.

Desde la *De Lingua latina* de Varrón hasta las *Institutio* de Quintiliano, pasando por la sintaxis de Apolonio Díscolo, la *Ars grammatica* de Donato o las *Institutiones Grammaticae* (Campbell, 2002) la constante definicional es que la preposición “ocupa” siempre lugares precisos en la cadena sintáctica. Luego el criterio clásico que identifica la conducta preposicional es sintáctico-estructural, uno de los más comunes y extendido, por otra parte.

Por lo que respecta al lugar de la preposición dentro del resto de las categorías gramaticales, la cuestión es diferente debido a ciertas variaciones estrictamente lingüísticas, según nos hallemos en la gramática del griego o del latín; o variaciones filosóficas, dependiendo de los supuestos que asumieran los autores. Así, por ejemplo, Diógenes Laercio, en las cinco partes del discurso: nombre, apelativo, verbo, conjunción y artículo, no incluye la preposición. Varrón, en cambio, las agrupaba en cuatro: las que tienen casos, las que tienen tiempo, las que tienen casos y tiempo, las que no tienen ni casos ni tiempos. Llama la atención, la importancia que han tenido después las nociones de caso y de tiempo en la discusión teórica sobre el componente categorial de las gramáticas. Como síntesis del periodo clásico resulta útil el esquema comparativo de Robins (1966, pág.18) que recoge estas diferencias desde Platón hasta Prisciano.

Situados ya en la Edad Media, comprobamos que la escolástica, a través de la definición modística de Thomas de Erfurt, incide en el valor posicional que se cita en (5).

(5) *Est praepositio pars orationis significans per modum adiacentis alteri causali ipsum contrahens et ad actum reducens* (Trad. de Maestre et alía, 2008: La preposición es la parte de la oración que significa por medio de construcción sintáctica con una palabra declinada, limitándola y relacionándola con una acción, pág. 871).

La construcción sintáctica está afectada por el caso de la palabra que interviene en ella. La condición causal, aunque específica para la gramática de la lengua latina, siguió y sigue como rasgo básico en la identidad preposicional.

El despliegue del programa intelectual del Renacimiento supondrá un interés, una preocupación y una voluntad de conocer la idiosincrasia de las lenguas vulgares o vernáculas. En ese contexto, se inserta la labor de Antonio Elio de Nebrija en relación al castellano², cuya importancia y relación con la tradición gramatical medieval ha sido destacada (Rico, 1978). También se ha comentado la calidad que tuvo su obra de *Introductiones Latinae* (1481) respecto de sus contemporáneos y de sus detractores, así como el rechazo de estos humanistas a los catálogos medievales como un canon nefasto. (Codoñer, 1983; Codoñer y González 1994).

En la primera gramática española de Antonio Elio de Nebrija (1492) se nos ofrece en (6) la indicación posicional de la preposición y, la novedad, la referencia a su posibilidad de componerse.

(6) *Es una de las diez partes de la oración, la cual se pone delante de las otras, por aiuntamiento, o por composición* (Nebrija, 1492, págs. 195-196).

De este modo se inicia la descripción de sus rasgos gramaticales. Nebrija advierte que poseen dos atributos: la *figura* y el *caso*. La figura, accidente que comparte con los nombres, divide las preposiciones en simples (*dentro*) o compuestas (*dedentro*). Y a su vez,

2. Calero Vaquera (1986) escribe una historia de la gramática del español desde Nebrija hasta Bello, pasando por el Anónimo Lovaina, Villalón, Minerva del Brocense, Correas, el padre Benito de San Pedro, Hermosilla y Salvá. Algunos de ellos son comentados en nuestro itinerario.

dos son los casos que se unen con ellas: el genitivo y el acusativo.

El tercer tipo de comentarios que introduce la tradición renacentista es la apreciación del significado preposicional a partir de una comparación con las preposiciones latinas. Según el gramático, la polisemia de *cerca* o de *por* se debe, como se apunta en (7) a su diversidad.

(7) *Los latinos abundan en preposiciones por las cuales distinguen muchas maneras de significar; e por que nuestra lengua tiene pocas es forçado que confunda los significados'* (Nebrija, 1492, pág. 196).

Respecto al caso como identificador funcional de la preposición Nebrija observa en (8).

(8) *Sirven para demostrar la diversidad de la significación de los casos, como de, para demostrar cuia es alguna cosa, que es el segundo caso; a, para demostrar a quién aprovechamos o empecemos, que es el tercero caso; a esso mesmo, para demostrar el cuarto caso en los nombre propios, y aún algunas vezes en los comunes* (Nebrija, 1492, pág. 196).

El gramático de Alcalá toma nota de la transición del sistema causal latino al sistema preposicional español que luego será uno de los campos de análisis sobre la génesis diacrónica de la clase. Efectivamente, se ha estudiado el origen diacrónico de nuestro sistema preposicional (Menéndez Pidal, 1904; Lapesa, 1942 y 1964), así como los procesos de transformación (Brea, 1985) y la dinámica sincrónica del acusativo y el dativo entre latín y castellano (Echarte, 1996), por ejemplo.

Un siglo más tarde, Sánchez de las Brozas (1578) en su *Minerva, seu de causis linguae latinae*³ compone una más de las versiones de las gramáticas, llamadas en la época generales, filosóficas o lógicas, que irían surgiendo en la Europa tardo-renacentista bajo el predominio de las especulaciones empiristas y racionalistas sobre la naturaleza del lenguaje humano. Se ha señalado la influencia de César Escalígero y su *De causis linguae Latinae*

3. Edición de Sánchez-Salor y Chaparro (1995).

(1540) en su propósito de búsqueda racional en la glosa gramatical. Uno de los apuntes más característicos del propio Escalígero es la advertencia en su obra de que las “preposiciones expresan el modo de las causas” (Galán, 2004).

Aunque se ocupe del latín nos interesa la aportación de Francisco Sánchez de las Brozas, El Brocense, respecto de las partes de la oración (1587). Propone tres categorías: nombres, verbos y partículas, en la línea de otros gramáticos racionalistas de la Europa de los siglos XVI y XVII. La noción de partícula será muy querida entre los investigadores del siglo XX, como se verá en la sección 1.5 de este capítulo.

Los libros segundo y tercero de su obra, dedicados a la *constructio*, observan formaciones con preposiciones. La novedad de este gramático consistiría, en incorporar, junto a Cristóbal Villalón, Gustavo Correas y Bartolomé Jiménez Patón, simplificaciones en el número de categorías. Los dos primeros adoptan tres categorías y el último, cinco. En el capítulo de su obra ya citada *Minerva sive de causis linguae latinae* (1587) indica que el orden es uno de los objetivos sintácticos de esta clase de palabras, como se lee en (9).

(9) *Praepositio, etiam si postponatur, non desinit uocari praepositio, ut Caesar Scaliger uidetur innuere, nam si grammaticum ordinem spectes semper praeponi debet* (Sánchez de las Brozas, 1578, pág. 112. Trad. de Santiago Arrondo: La preposición, incluso si se antepone, no deja de llamarse preposición según lo que César Escaliger parece afirmar, ya que si observas el orden gramatical (correcto), esta siempre debe anteponerse).

El hecho de que El Brocense señale el carácter generalizado y aplicable a otras lenguas ha sido visto como una influencia de lingüistas árabes y hebreos, aunque insisten en la impronta de las teorías de Apolonio Díscolo y Prisciano para su sintaxis.

La *Grammaire générale et raisonnée contenant les fondemens de l'art de parler, expliqués d'une manière claire et naturelle* (1660), conocida como la gramática de Port-Royal, en referencia a la conocida abadía cisterciense al sudoeste de París, incluye también la preposición entre las partes del discurso junto a nombres, artículos, pronombres, verbos, adverbios, conjunciones, interjecciones y participios.

De vuelta al ámbito español, el lapsus de tiempo que va de El Brocense a la GRAE de (1771) ha sido analizado con detalles desde las comparaciones sobre los criterios de definición, los oficios y los valores, hasta las clasificaciones que hicieron los lingüistas de la época sin encontrarse ningún aspecto nuevo a destacar (Calero, 1986).

Tendremos que esperar a la primera gramática de la Real Academia Española (1771) para aumentar el caudal de pormenores y novedades. Allí proponen la inclusión de artículos, pronombres, artículos, participios, adverbios, preposiciones, conjunciones e interjecciones, es decir, la misma serie que la de Port Royal. No en vano, ya se están produciendo debates intensos sobre cuáles son las categorías gramaticales y cuáles sean los criterios más pertinentes⁴ para establecer esos grupos.

Si bien es cierto que la definición académica no avanza respecto de la óptica lógico-sintáctica, incorpora en su descripción una característica nueva: la dependencia preposicional, otro hecho de la controversia contemporánea. Así se confirma en (10).

(10) El oficio de la preposición por sí sola es indicar en general alguna circunstancia que no se determina sino por la palabra que se le sigue; pero junto ya con ella, denota la diferente relación ó respeto que tienen unas cosas con otras (RAE, 1771, pág. 319).

Esta propiedad, la dependencia contextual o distribucional, será muy atractiva en el desarrollo de la semántica preposicional en las teorías del siglo XX y en la actualidad. A continuación en (11) advierten sobre cuáles sean las “verdaderas” preposiciones.

(11) Verdaderas preposiciones son las que constan de una sola dicción, y se usan sencillamente. Las que no se usan sino en composición, no se deben reputar como preposiciones, sino como parte de aquellas voces compuestas con ellas. Las que constan de dos o más dicciones separadas tampoco se deben reputar como preposiciones, sino como modos o frases adverbiales (RAE, 1771, pág.320).

4. La edición de Sarmiento (1987) de la GRAE recoge dos cuadros comparativos de la tradición grecolatina sobre las partes de la oración, que prefería entre 8 y 10 clases, y las reducciones de los racionalistas, que optaban por tres, págs. 57-8. En el prólogo de esta primera gramática académica se hacen eco de esta diversidad de opiniones de los antecedentes de los autores académicos, págs. 102-105.

Se detienen en cada uno de los elementos que forman la clase, como ya hiciera Nebrija, y deslindan los usos y las expresiones particulares de cada una. Quizás el aspecto que nos interese de la GRAE sea la advertencia que hace sobre la dificultad en la elección del régimen preposicional (Gutiérrez Araus, 1986); la primera gramática académica ya identificaba uno de los asuntos más espinosos de las construcciones preposicionales. Así, se observa en (12).

(12) *Otro punto difícil de nuestra Gramática, y en que consiste la parte principal de la sintaxis es saber qué preposiciones piden después de sí algunos verbos y otras partes de la oración* (RAE, 1771, pág.107)⁵.

La cuestión del inventario preposicional en español será muy dispar en los textos, tanto en los ítems que lo forman como en las razones o causas para elegirlos; ni siquiera el modelo académico será homogéneo en los siglos XVIII y XIX⁶. De hecho, la lista de preposiciones de 1771 incluye la palabra *como* y están ausentes *cabe*, *bajo*, *mediante*, y *durante*. Entre la gramática académica y el Romanticismo se reconoce (Gómez Asensio, 1981 y 1985) que se dedica poca atención a las preposiciones, aunque surge la separación entre preposiciones propias e impropias, y se organizan a partir de dos tipos de clasificación: uno casual y otro semántico.

En el siglo XIX, tres lingüistas, Vicente Salvá, Andrés Bello y Rufino José Cuervo, caracterizan el proceder sintáctico y el valor semántico de las preposiciones con renovados matices (Mourelle, 1968).

De la obra de Vicente Salvá, *Gramática de la lengua según ahora se habla* (1831) destacaremos tres aspectos que continuarán en el futuro: las descripciones semánticas sobre las piezas preposicionales, la inclusión de ejemplos de fenómenos sintácticos (la elipsis, la redundancia, el régimen y la concordancia) y la discusión sobre cuáles son los miembros

5. Para el seguimiento de las doctrinas académicas, véase Sarmiento (1978): *La gramática de la Academia: historia de una metodología*, BRAE, 59, págs. 435-446; y la edición del mismo Sarmiento (1987).

6. Para una comparación de los inventarios en Villalón, Jiménez Patón, GRAE (XIX), GRAE (XX), y Salvá véase Lázaro Mora (1985).

de la clase. Sorprende el detalle descriptivo tan minucioso de las piezas preposicionales (Peñalver, 1992).

Como en otras cuestiones, la influencia y la actualidad de Andrés Bello en la gramática española también se observa en ciertas apreciaciones sobre la definición y el comportamiento preposicional. La recepción de las pesquisas de Bello así como su doctrina gramatical, sus antecedentes filosóficos, el carácter innovador de su obra, su herencia y la modernidad de su propuesta taxonómica de palabras han ocupado abundantes investigaciones (Alonso, 1963; Torres, 1966; Calderón 1967; Urrutia 1984; García 1997; y Obediente y D’Introno, 2001).

La preposición es, para Bello, un procedimiento de modificación que sirve para todas las palabras. Esta noción se perpetuará en la definición académica y en la mayoría de gramáticas del siglo XX que comentaremos. Le interesa más la naturaleza sintáctica que la aportación semántica de los elementos de la clase en tanto que deja, como se afirma en (13), la tarea explicativa a los diccionarios.

(13) *En la enumeración de los varios significados que toma cada preposición, y de los verbos que las rigen, nos limitaremos a unas pocas observaciones generales sobre el modo de usarlas* (Bello, 1847, pág. 700).

Es, en consecuencia, un asunto léxico-semántico. Dos de esas observaciones nos parecen muy sugerentes desde la perspectiva contemporánea. La primera tiene que ver con el régimen. Asevera Bello en (14):

(14) *Las preposiciones castellanas no tienen propiamente régimen, porque régimen supone elección; así un verbo rige un modo o un complemento particular, porque hay varios modos y multitud de complementos; al paso que con todas las preposiciones lleva el término de forma invariable* (Bello, 1847, págs. 702-703).

Esta distinción sobre el régimen, como notaremos, ha ilustrado muchas descripciones actuales, no solo del denominado complemento de régimen verbal,

preposicional o suplemento, sino de otras tareas sintáctico-semánticas (Chevalier, 1968; y Martínez, 1983).

La segunda indicación es sobre el origen de ciertos elementos de la clase. Así, tras referir la imperfección de palabras como *excepto, salvo, durante, mediante, obstante* y *embargante*, advierte en (15).

(15) *Muchas preposiciones, y acaso todas, han sido en su origen palabras de otra especie, particularmente nombres* (Bello, 1847, pág. 698).

La cuestión del contacto entre preposiciones y otras categorías formará parte de los debates en los estudios del siglo XX que promoverán soluciones como la recategorización, la neutralización y el reanálisis (Ramat, 1999). La voluntad reductora de las categorías sintácticas -a propósito del acercamiento entre determinantes, pronombres y adjetivos- ha sido considerada y comparada (Luján, 1999) en relación a las estipulaciones del reciente Programa Minimista de la tradición de la Gramática Generativa sobre los rasgos y las clases categoriales.

Por su parte, Rufino Cuervo en la introducción de su *Diccionario de construcción y régimen de la lengua española* habla de las preposiciones como exponentes (Muñoz, 2017). Esta aproximación se da dentro de una posición sintagmática y asegura que se emplea con la intención apuntada en (16).

(16) *Completar la significación del verbo señalando los objetos con que el agente se pone en relación mediante el acto de que se trata, y han de llevar un exponente que especifique tal relación* (Cuervo, 1872-1994, pág. 8).

Cuervo juzga el criterio posicional como insuficiente ya que es sólo *una circunstancia externa* e insiste en la importancia que tienen las preposiciones en la composición del significado (Folgar de la Calle, 1983).

Como balance de lo traído y comentado hasta aquí, reunimos en el cuadro 1 las contribuciones y las consideraciones que se han producido en este extenso periodo. Están en relación con las tres preguntas que guían nuestro recorrido, recogidas en la página 4. Cuando algún autor no se ha ocupado de alguna de las cuestiones queda en sombra.

Síntesis de aportaciones sobre la categoría desde la antigüedad grecolatina hasta el siglo XIX			
Época y autores	La intensión de la categoría (definición)	La extensión. Sus miembros y su estatus categorial	Relaciones con otras categorías
Antigüedad grecolatina Dionisio de Tracia Diógenes Laercio Varrón (Quintiliano, Apolonio Díscolo, Donato y Prisciano)	La posición es crucial. La composición y la morfología son necesarias. La caracteriza a partir del caso y otras informaciones gramaticales.	No la incluye entre las clases de palabras.	
Edad Media Thomas de Erfurt	La posición y el caso son sus dos rasgos definitorios.		
Renacimiento y Barroco Elio Antonio de Nebrija	Adopta una definición posicional y señala su sintaxis	Establece una relación de piezas y las compara con el latín.	Las relaciona con locuciones y adverbios.

<p>Sánchez de las Brozas, “El Brocense” (Villalón, Jiménez Patón y Correas)</p>	<p>en relación con el caso y la figura. Trata sobre el significado.</p> <p>Ahonda en la información semántica. Destaca el orden como decisivo.</p>	<p>[Redacted]</p>	<p>[Redacted]</p>
<p>Ilustración</p> <p>Gramática de Port-Royal</p> <p>GRAE</p>	<p>[Redacted]</p> <p>Incorpora la noción de relación. Introducen la cuestión del régimen verbal.</p>	<p>Las incluye junto a otras partes del discurso.</p> <p>Establece una lista de preposiciones.</p>	<p>[Redacted]</p> <p>Añaden consideraciones en relación con los adverbios.</p>
<p>El siglo XIX</p> <p>Vicente Salvá</p> <p>Andrés Bello</p> <p>Rufino Cuervo</p>	<p>Continúa con las descripciones semánticas e incorpora aspectos sintácticos.</p> <p>Señala aspectos sintácticos y semánticos. Puntualizaciones sobre el régimen.</p> <p>Indicación sobre el valor de especificación y de exponentes.</p> <p>Insistencia en el valor léxico.</p>	<p>Aparece una discusión sobre cuál sea la lista definitiva.</p> <p>Propuesta de una lista (a, ante, bajo, con, contra, de, desde, en, entre, hacia, hasta, para, por, según, sin, so, sobre, tras)</p> <p>[Redacted]</p>	<p>[Redacted]</p> <p>Dudas sobre el valor preposicional de algunas piezas en comparación con otras categorías.</p> <p>[Redacted]</p>

Cuadro 1. Contribuciones gramaticales al estudio preposicional desde la antigüedad grecolatina hasta el siglo XIX.

1.3 Algunas consideraciones de la lingüística general del siglo XX: de Edward Sapir a Bernard Pottier

La investigación teórica en el estudio de las lenguas naturales supuso, con el inicio del estructuralismo y a partir de la publicación del *Cours de Linguistique Générale* (1916) de Ferdinand de Saussure, una disposición más sistemática y global frente a los hechos, los fenómenos y las categorías lingüísticas. A pesar de este carácter metodológicamente más científico los puntos de vista que revisamos reiteran muchos criterios, observaciones y rasgos que proceden de la tradición moderna, desde Nebrija a Cuervo.

Es imprescindible, sin embargo, situar alguna de esas asunciones y valoraciones teóricas por lo que tuvieron de precursoras e inspiradoras para ciertos gramáticos afanados en la descripción del español y otras lenguas en la segunda mitad del siglo XX.

Saussure (1916) no se ocupa de asuntos teóricos relacionados con la categoría, al margen de una breve indicación diacrónica sobre la aparición de las preposiciones como una nueva clase de palabras.

Deberemos esperar cinco años a la monografía del estadounidense Edward Sapir (1921) para encontrar una referencia explícita a las clases de palabras.

El autor admite que el problema de las partes de la oración es poco menos que un falso problema, o mejor, una cuestión irrelevante en el estudio de la lengua. Sapir advierte en (17).

(17) The upshot of such an examination would be feel convinced that the "part of speech" reflects not much our intuitive analysis of reality as our ability to compose that reality into a variety of formal patterns. A part of speech outside of the limitations of syntactic form is but a will o the wisp. For this reason no logical scheme of the parts of speech -their number, nature, and necessity confines- is of the slightest interest to he linguist. Each language has its own scheme. Everything depends on the formal demarcations which it recognizes (Sapir, 1921, pág. 119).

La actualidad de esta observación sapiriana es sorprendente. Seguramente, el conocimiento que el lingüista poseía de ciertas lenguas amerindias (el fragmento anterior procede de un comentario de la lengua Chinook, perteneciente a la familia de las penutías, habladas en Oregón y California) y su comparación con las gramáticas de lenguas como el inglés llevó a Sapir a adoptar una posición contracategorial y relativista. En ese sentido, y como veremos en la sección 1.5, muchos datos actuales proceden de lenguas africanas con organizaciones gramaticales muy distintas a las de las familias lingüísticas occidentales. La perspectiva tipológica e interlingüística sigue siendo una de las vetas de enriquecimiento empírico inexcusable en el debate y la investigación sobre la preposición.

Todavía en EUA, Bloomfield (1933) es crítico con la correlación entre partes del discurso / clases de palabras, y clases sintácticas, como se desprende de (18).

(18) *The term parts of speech is traditionally applied to the most inclusive and fundamental word-classes of a language, and then, in accordance with the principle just stated, the syntactic form-classes are described in terms of the part of speech that appear in them. However, it is impossible to set up a fully consistent scheme of parts of speech, because the word-classes overlap and cross each other* (Bloomfield, 1933, pág. 19).

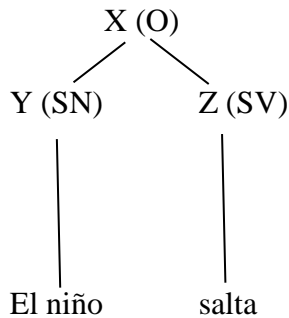
El solapamiento del que habla está en la base, también, de algunas revisiones más recientes sobre las semejanzas entre las preposiciones y los nombres; y las preposiciones y los adverbios.

De todos modos, la contribución de Bloomfield -que fue descriptivamente estable y aceptada durante cerca de cuarenta años- es la determinación de que los sintagmas preposicionales que proceden del binomio estructural relación-axis son exocéntricos⁷. Así, considera que expresiones inglesas como *beside John, with me, in the house, o by running away* forman complementos exocéntricos y no forman un constituyente inmediato. En el esquema (19) se representa un sintagma exocéntrico clásico, con el símbolo O(ración),

7. Lingüistas como Bloch y Tragger (1942), Hockett (1958), Halliday (1961) aceptan esta descripción exocéntrica. Charles Hockett, consideró la distinción endocéntrico-exocéntrico como reguladora de la disposición de los tipos de construcciones sintácticas para las lenguas naturales.

donde el núcleo del sintagma (O) no es de la misma categoría que sus componentes (SN) y (SV).

(19)



De vuelta a Europa es la obra de Otto Jespersen, *The Philosophy of Grammar* (1924), la que más atención dedica a estudiar las propiedades preposicionales. Tres cuestiones resultan reveladoras décadas después: su interés por la complejidad semántica de las preposiciones, su reducción de adverbios, conjunciones, interjecciones y preposiciones a una única categoría, y el estudio distribucional de los grupos introducidos por las preposiciones. Sobre el primer aspecto se admite en (20).

(20) *Es perfectamente normal que los diccionarios expliquen los diferentes usos de at, for, in, etc., de la misma forma que tratan a fondo los diferentes significados de los verbos put y set. Pero, por otro lado, las preposiciones merecen ocupar un lugar en la gramática en la medida en que hay que citar hechos generales en relación con ellas* (Jespersen, 1924, pág. 14).

Parece que no son suficientes las indicaciones aisladas del diccionario; de manera que los aspectos sintagmáticos y los “hechos generales” aportan matices y percepciones nuevas de la conducta preposicional.

Su segunda contribución reconocida es la crítica al sostenimiento de cuatro categorías diferenciadas (preposiciones, adverbios, conjunciones e interjecciones). Aduce Jespersen en (21) respecto a las similitudes y diferencias entre ellas.

(21) *Se exageran demasiado las diferencias que existen entre dichas palabras y en consecuencia sus evidentes semejanzas quedan oscurecidas* (Jespersen, 1924, pág.90).

El grado de similitud que indica Jespersen es muy alto a partir de criterios como la invariabilidad morfológica o la dificultad para distinguir significados.

La última cuestión que nos interesa del autor danés es su simplificación y crítica sobre la terminología que dibuja los grupos introducidos por preposición, en particular, sobre los tipos de dativo en inglés, alemán y francés. Señala que algunas distinciones de los complementos a partir de la presencia de tal o cual preposición inducen a error.

Todavía en el ámbito danés, será la glosemática de Louis Hjelmslev, en *La categorie du cas* (1935)⁸, la que aporte una de las monografías más recurridas en la escuela estructuralista europea. Hjelmslev atribuye al caso el papel de identificador de la categoría, como se muestra en (22).

(22) *Si las preposiciones constituyen una categoría particular en el seno del sistema, es porque constituyen un paradigma definido por la facultad de sus miembros de regir casos; además las preposiciones se ordenan naturalmente en subcategorías según los casos que rigen* (Hjelmslev, 1935, pág.188).

La relación formal y la correlación sintáctica entre preposiciones y caso se recogen en la cita (23).

(23) *Si las preposiciones y sus subcategorías quedan definidas por el hecho de definir ciertos casos, esos casos constituyen evidentemente una categoría por el hecho de admitir la función de régimen preposicional. La categoría de los casos puede definirse como un paradigma en el que algunos miembros pueden estar sometidos a la acción de las preposiciones y los verbos* (Hjelmslev, 1935, pág.188).

8. Hjelmslev (1935) *La catégorie du cas. Étude de grammaire générale*, y también en sus *Estudios lingüísticos* (1959). También son destacables, Bally (1941) : *Les prépositions dans leur rapport avec les verbes transitifs*, y Benvéniste (1958), en *Catégories de pensée et catégories de langue*. Este último aporta observaciones generales sobre el vínculo entre pensamiento y lenguaje, y el papel que las categorías lingüísticas juegan en estas estipulaciones.

Sin embargo, y a diferencia de Jespersen, parece reconocer diferencias suficientes entre preposiciones y conjunciones, tal y como se indica en (24).

(24) *Es más bien conveniente constatar que hay en la preposición un elemento casual y en la conjunción un elemento modal que rigen a un elemento idéntico contenido en su régimen* (Hjelmslev, 1935, pág.191).

Todavía en el paradigma estructuralista debemos citar la influencia de la obra de Viggo Brondal (1940), *Théorie des prépositions. Une introduction à une sémantique rationnelle*. En la línea de los postulados estructuralistas, reitera que la comprensión del sistema lingüístico pasa por las unidades, sus usos y sus relaciones con otros elementos. Opta por una visión conceptual según la cual las preposiciones se pueden explicar con tres propiedades para las relaciones que establecen: simetría, transitividad y conectividad (Hoekstra, 1984; y Gutiérrez Araus, 1987). A partir de esas características se desgranar otras, como la variabilidad, la pluralidad, la generalidad, la continuidad y la totalidad.

Concluimos la sección reseñando el gramático francés Bernard Pottier, y su título sobre la semántica de los elementos de relación, *Systématique des éléments de relation* (1962). Identifica que ciertas preposiciones pueden funcionar, a su vez, como prefijos (volar *sobre* el mar / *sobrevolar* el mar). Además se interesa en los procesos de gramaticalización en los que participan estas partículas. Al entrar en la vertiente semántica, en la sección tercera, volveremos sobre alguna de sus consideraciones dentro de la tradición francófona.

1.4 Gramáticas, estudios y aportaciones descriptivas y normativas para la preposición española durante el siglo XX

El volumen de referencias, trabajos y propuestas que se produjeron durante el siglo XX ha sido excepcional, en particular, y como indiqué al inicio, en la segunda mitad de la centuria (Guimier, 1981). La extensión social de la educación en la población y el aumento de la oferta universitaria así como la emergencia de los centros de investigación han

supuesto una acentuación del interés por el funcionamiento de las lenguas. Este hecho se produce de manera exponencial a partir de los años 70 del siglo XX.

La selección de unas contribuciones y no otras parte de su interés metodológico, de la influencia en otros y/o la repercusión en el progreso descriptivo, empírico-analítico o explicativo sobre la categoría. Las tres cuestiones iniciales, recogidas en (1) – (3), sección 1.1, página 4, seguirán ayudando en nuestra exposición.

Desde el punto de vista de la organización del recorrido alternamos aquí el criterio cronológico a partir de la publicación de gramáticas generales sobre el castellano, junto con visiones globales sobre la categoría, análisis de casos particulares, estudios de construcciones sintácticas y miradas diacrónicas.

Fue Rodolfo Lenz, en *La oración y sus partes. Estudios de gramática general y castellana* (1920), el primero del siglo XX en dedicar un extenso capítulo a tratar sobre la categoría. Asume la definición relacional y transformadora (o modificadora), y a partir de ella, analiza los vínculos entre la preposición y el caso, a la luz de la historia casual de la familia lingüística indoeuropea. Por otra parte, se detiene en su función. Si sumamos estos tres intereses, su definición de (25) es bastante elocuente respecto de lo visto hasta ahora.

(25) *Palabra invariable que sirve para transformar un sustantivo en atributo complemento de otro elemento de la misma proposición* (Lenz, 1920, pág.509).

Respecto a las clasificaciones para las preposiciones considera dos criterios: el formal y el lógico. El segundo de ellos maneja concepciones que son de actualidad en la semántica de la deixis. Efectivamente, se ocupa de preposiciones locativas, de movimiento, temporales y de relación general. Estos términos descriptivos seguirán activos durante décadas y se interrelacionarán con otras áreas como la lógica y la filosofía del lenguaje.

Este análisis semántico que refleja las denotaciones de la preposición también se observa en el *Manual de gramática española* (1930) de Rafael Seco. De hecho, el gramático insiste en *la relación ideológica que existe entre los conceptos* que están representados por las palabras que une la preposición.

Desde la perspectiva sintáctica, Samuel Gili Gaya en su *Curso superior de sintaxis española* (1943) incluye la preposición entre *uno de los signos gramaticales diversos* con los que se puede expresar una relación entre núcleos y complementos, como el orden de colocación, la concordancia y la flexión. Se aventura ya una conexión entre la categoría preposicional y los rasgos de la funcionalidad (como veremos en la sección siguiente). Reconoce una zona borrosa en la delimitación entre conjunciones y preposiciones. Se incorporan por aquel entonces gramáticas pedagógicas (Alonso y Ureña, 1938), descripciones del sistema verbal en relación con la preposición (Criado de Val, 1951) y la publicación de una gramática estructural con la introducción de los conceptos de la glosemática danesa (Alarcos Llorach, 1951).

En esta última, y a través de la divulgación y aplicación de esa teoría se incluyen definiciones y terminologías nuevas (26).

(26) *Tienen la función privativa de establecer relaciones entre distintos planos sintácticos (conectores del modificador con el núcleo e indicadores de la inclusión de una construcción en otra mayor* (Alarcos Llorach, 1951, pág. 23).

La noción de partícula, tan querida en los siglos XVIII y XIX, reaparece en la *Introducción a la gramática* (1960) de Roca Pons cuando se ocupa de preposiciones y conjunciones. También remarca en (27) la ligazón con los adverbios. Las conexiones intercategoriales han sido constantes en este repaso descriptivo.

(27) *Las preposiciones proceden, con frecuencia, de antiguos adverbios y, siempre, en último término, de antiguos elementos nominales. El origen nominal directo es claro en casos como durante o excepto, antiguos participios, el uno de presente y activo, y el otro, de paso y pasivo* (Roca Pons, 1960, pág. 261).

Recoge la vecindad diacrónica en las funciones de los casos latinos y las preposiciones españolas. Sus clasificaciones son listas de criterios empleados con anterioridad, citando algunos gramáticos comentados en este itinerario, y usando ejemplos de lenguas como el alemán o el griego.

Otra aproximación estructuralista y funcional (Barrenechea y Manacorda de Rosetti, 1969) introduce la distinción de las clases de palabras como tipos funcionales. La gramática estructural europea hace décadas que ha introducido las nociones de relacionante, coordinante y subordinante. Solo un año antes, se había propuesto el concepto de suplemento (Alarcos Llorach, 1968) para referirse a los tradicionales complementos de régimen, y se inicia una escrupulosa veta de investigación sobre el papel de la preposición en tales estructuras.

En 1970 se publica una de las primeras monografías sobre la preposición para el español (López, 1970). Es una extensa panorámica que recoge desde la historia de la categoría hasta distintos problemas empíricos relacionados. Sus fuentes metodológicas son estructuralistas y representa una recapitulación bastante adecuada de las contribuciones de esta tendencia. Dos capítulos muestran un interés particular para nosotros: el dedicado a la rección y el que se ocupa de los prefijos. Introduce el concepto de rección como equivalente al de concordancia, analogía que tiene un valor especial para las descripciones del momento.

La segunda monografía es, con toda probabilidad, uno de los mejores compendios descriptivos de los que disponemos (Trujillo, 1971). Su primera consideración, recogida en (28), es bastante alentadora respecto de la intensidad de la categoría.

(28) *El conjunto de las preposiciones presenta un perfil bastante definido y deslindado de otras clases de unidades lingüísticas* (Trujillo, 1971, pág. 345).

Su propuesta de definición es un conglomerado de términos estructuralistas, de manera que la presenta como *morfema intenso hipotáctico independiente*, es decir, una palabra independiente con carácter relacional. Y, contra el pronóstico de observadores anteriores, mantiene una posición homogénea sobre la clase. Más adelante, reanuda la caracterización de dependencia semántica a favor de los términos que relaciona, tal como se indica en (29).

(29) *Para el análisis del valor de las preposiciones es necesario tener en cuenta las características semánticas de los dos términos, regente y regido* (Trujillo, 1971, pág. 354).

Con una pregunta, que se lee en (30), plantea uno de los argumentos que más fortuna ha tenido en las discusiones sobre la categoría.

(30) *¿Posee un valor semántico cada preposición, o se trata sólo de elementos vacíos cuyo valor está a merced de los términos relacionados?*
(Trujillo, 1971, pág. 364).

Los criterios que aduce para respetar su variable carga semántica serán examinados más adelante; en particular el de la conmutación.

Los usos preposicionales y sus valores también se presentan al público en la década de los setenta (Luque, 1973). Son dos volúmenes dedicados a recoger algunos asuntos ya conocidos en la tradición gramatical. Destaquemos la presencia tan evidente que tienen los valores idiomáticos, es decir, aquellos significados fijados a lo largo del tiempo. No en vano, esos procesos de gramaticalización y fosilización de estructuras semánticamente complejas serán uno de los fenómenos más difíciles en el tratamiento de estas partículas en lenguas diferentes (Hopper y Clos, 1993; Amiot y De Mulder, 2003; y Girón, 2014).

En el mismo año 1973 se publica el texto gramatical de la Academia, *Esbozo de una nueva gramática de la lengua española*. La obra académica reúne las atenciones analíticas de cierto estructuralismo de la lingüística anterior e insiste en la posición de enlace relacional, respecto del criterio sintáctico para la definición. A su vez, dejan la “vaguedad” del significado en manos del contexto que determinará el sentido adecuado de la partícula: un lugar común en nuestra investigación. Compara, en (31), el carácter nexivo de la preposición con el de la conjunción y sugiere el doblete funcional.

(31) *Ya que no son, pues, las preposiciones nexos enteramente vacíos, sino que apuntan con más o menos vaguedad a la clase de relaciones que pueden expresar* (RAE, 1973, págs. 435)⁹.

Están irrumpiendo, incluso desde la óptica normativa, alguna de las cuestiones explicativas: ¿son meros índices funcionales o ayuda a perfilar el sentido de sus relaciones?

9. La sección dedicada a las preposiciones es la 3.11, págs. 434-443.

Otra gramática de la época (Alcina y Blecua, 1975) sigue a la Academia en las tres características señaladas. Una de las novedades, que también recuperaremos después, es la introducción de los adverbios prepositivos como parientes necesariamente destacables. El contacto entre preposiciones y adverbios es una constante todavía hoy, y una fuente de evidencias empíricas para el mejoramiento analítico de la categoría. Recogen las anomalías y excepciones que producen las piezas como *excepto*, *salvo*, *durante*, *mediante* y *embargante*.

Junto a las gramáticas y los estudios generales de sintaxis o de la preposición en general, en la década de los 70 del siglo XX, empiezan a producirse los primeros trabajos sobre usos particulares de cada una de ellas. Desde un artículo sobre la preposición *de* (Alarcos, 1972), una investigación sobre la alternancia de *para* y *por* (Amícola, 1975) y el estudio de la oposición iberrorrománica de *por* y *para* (Riiho, 1979) hasta el estudio del valor relacional y su sentido déctico en algunas de ellas (Carbonero, 1975 y 1979) pasando por las relaciones entre complementos directos e indirectos (Barrenechea y Orecchia, 1970), el papel diferenciado de las preposiciones en directos, indirectos y circunstanciales y en el gerundio preposicional (Gutiérrez Araus, 1978; González Ollé, 1979; y Mozos 1973). Queda claro que el predominio de la sintaxis es sustancioso en este primer aluvión de acercamientos (Bosque, 1980).

En la década siguiente no se publica ninguna gramática de referencia. En cambio, sí que aparecen nuevas visiones globales de la categoría y líneas de investigación exhaustivas. Ven la luz caracterizaciones muy exhaustivas sobre la sintaxis relacional (Porto Dapena, 1987) y sobre la semántica desde la óptica estructural (Morera, 1988 y 1990). En esos dos bloques de interés se ahonda en el suplemento de Alarcos (Martínez, 1986, 1992 y 1995; Gutiérrez Ordóñez, 1996; y Cuesta, 1997) y la centralidad de su sentido relacional (Rylov, 1989). En el rastreo empírico de los perímetros semánticos se comparan datos de sintagmas nominales y preposiciones para la expresión de causas, lugares y tiempos (Bosque, 1989), uno de los tópicos más recurrentes décadas después; o se cotejan con otras categorías a partir de las palabras *atrás*, *adentro*, *afuera* (posposiciones), *bajo*, *tras*, *sobre* (preposiciones), y *detrás*, *cerca*, *lejos* o *encima* (substantivos) (Plann, 1988).

La primera gramática de los años 90 del siglo XX es la Alarcos Llorach (1995), que patrocinada por la RAE, y en consonancia con las asunciones estructuralistas que ya hemos venido apuntando, sugiere la siguiente tesis (32).

(32) *Se trata de unidades dependientes que incrementan a los sustantivos, adjetivos o adverbios como índices explícitos de las funciones que tales palabras cumplen bien en la oración, bien en el grupo unitario nominal* (Alarcos Llorach, 1995, pág. 214).

Una vez más, dependencia, incremento e índices son los términos que convergen en la identidad de la preposición.

El siglo XX se cierra con la publicación de la completísima *Gramática descriptiva de la lengua española* (1999), dirigida por Ignacio Bosque y Violeta Demonte. Desde el punto de vista metodológico estamos ante la obra más completa, en referencias, ejemplos y suma de perspectivas, que alcanza a nuestra categoría. No en vano, es obra de 73 autores diferentes. La preposición se estudia en los capítulos 9, de María Victoria Pavón, y el 10, de Jacques de Bruyne. Además, y siguiendo los criterios expositivos y metodológicos de la obra, los lingüistas encargados de fenómenos donde participan preposiciones, como el acusativo preposicional (cap. 28 por Esther Torrego), el dativo (cap. 30 por Salvador Gutiérrez), el régimen (cap. 29 por Rafael Cano) o las perífrasis verbales (caps. 51 por Leonardo Gómez, y 52 por Alicia Yllera), se ocupan con mayor o menor intensidad de ellas. No hay, por lo tanto, una posición interpretativa única, exclusiva y homogénea.

Quizás por el efecto de la extensión y la divulgación escolar, o cierto mal uso en el lenguaje periodístico, en ese periodo temporal se publican tres repertorios normativos y prescriptivos sobre el uso de la preposición (García Yebra, 1988; Nández, 1991; y Fernández, 1999). Desde el punto de vista descriptivo estos manuales no aportan novedades sustanciales, como tampoco lo hará el más reciente de Moliner (2012), *Uso de las preposiciones*, donde ya se incluye una presentación de *versus*, *vía*, *allende*, *aquende*, *ápud* y *circa* como miembros de la serie.

El buen uso y adecuación ha venido preocupando más recientemente y se ofrecen misceláneas y diccionarios de combinaciones y construcciones (Slager, 1997 y 2004; y Bosque, 2006). Nuevamente, son contribuciones con escaso interés teórico pero con una

manifiesta voluntad orientadora y prescriptiva, recursos inexcusables para los profesionales del área de ELE.

Al abrigo de esta extensa y completa tradición normativa y didáctica, eclosionan y se multiplican diversas perspectivas que se concretan en artículos, tesis doctorales y estudios de toda condición (Hernando, 2002). El predominio en el interés sobre la preposición es todavía eminentemente sintáctico (Trujillo, 1993; y Alarcos Llorach, 1990).

Se nos ofrecen estudios desde la gramática de valencias (García-Miguel, 1991 y 1995a), y se profundiza en los aspectos distributivos (Kailuweit, 1993; y Gutiérrez Ordóñez, 1994) y con diversidad de complementos y contextos sintácticos donde las preposiciones ejercen sus tareas (Masiá, 1995; Pensado, 1996; Serradilla, 1997; Laka, 1995; Hernando Cuadrado, 1998; y Bosque, 1998); a la vez que se examinan construcciones especiales como las dobles preposiciones (Bosque, 1993).

Igualmente se abunda en los casos particulares de preposiciones concretas (Lenarduzzi, 1996; y Crego, 1997), en las locuciones prepositivas (Koike, 1997) y surgen los primeros trabajos desde la gramática cognitiva (Cifuentes, 1997), así como la consideración de las preposiciones como una versión de las adposiciones, variante tipológicamente universal emparentada con formas adverbiales (Chapado, 1996; Haspelmath, 1997; Tammenga-Helmantel, 1998; y Levinson y Meira, 2003).

El último conjunto de aportaciones en esos años son las incursiones de carácter diacrónico. En el II Congreso Internacional de Historia de la Lengua Española, celebrado en Sevilla en 1992, se presentan estudios sobre los orígenes históricos de *hasta* (Martínez, 1992), sobre la expresión de causa a través de sintagmas prepositivos en la Edad Media; (Mendoza, 1992), el suplemento propio e impropio en el *Cantar de Mío Cid* (Millán, 1992), y una explicación histórica para la oposición *para / hacia* (Llitas, 1992). El interés histórico continúa con la recopilación de ejemplos de preposiciones ante la conjunción *que* (Serradilla, 1995).

Se abre el siglo XXI con la *Nueva gramática de la lengua española* (2009), abreviada *NGLE*. Es el referente normativo y descriptivo más actual que asume buena parte de los tópicos de la controversia que venimos exponiendo.

En el capítulo XXIX, “La preposición y el grupo preposicional” se definen, tal y como se lee en (33).

(33) *Palabras invariables y por lo general átonas que se caracterizan por introducir un complemento que en la tradición gramatical hispánica se denomina término (NGLE, 2009, Vol. II, pág. 2223).*

El sintacticismo que ha predominado, prácticamente desde Nebrija, sigue vigente. Tres rasgos más siguen en la última edición académica: rigen su término, algunas (*a* y *de*) son marcas de función y su término pueden ser grupos nominales, pronominales, adjetivales, adverbiales y preposicionales. La caracterización distribucional completa su sintaxis.

Por otra parte, se da por hecho, en el epígrafe 29.1 de la *NGLE*, una distinción básica en las exégesis teóricas, y que transcribimos en (34).

(34) *La distinción entre preposiciones de contenido gramatical o funcional y preposiciones de contenido léxico se considera gradual (NGLE, 2009, Vol II, pág. 1289).*

Esta manera de entender la clase es, como veremos, uno de los puntos cardinales del debate sobre la preposición: ¿hasta dónde es un grupo homogéneo?

El significado preposicional se considera abstracto y relacional; y se nos remite a tres apartados extensísimos, el 29.6, 29.7 y 29.8, donde se detallan los valores, los contenidos y los usos de cada una de las piezas.

Por último, es importante reseñar que la *NGLE*, dedica un capítulo íntegro al complemento de régimen preposicional. El examen de este peculiar complemento verbal ha crecido de manera exponencial pero ya empieza a incorporar las indicaciones semánticas como imprescindibles y sustanciales en su análisis.

Una de las propuestas más exhaustivas del momento (Hornos, 2002) formula cuatro cuestiones entre (35) y (38) que estarán en el centro del debate teórico contemporáneo.

(35) ¿En qué consiste el concepto de argumentalidad?

(36) ¿Cuáles son los límites entre la argumentalidad y la adjunción en el seno del Sintagma Verbal?

(37) ¿Qué entendemos por sintagma preposicional?

(38) ¿Cuál es la naturaleza de la preposición?

Parte de una propuesta lexicalista según la cual los principios estructurales se pueden prever desde la información del predicado en el lexicón. Se acepta como empíricamente justificada la distinción entre preposiciones funcionales y léxicas y, además, establece una gradación de unidades: léxicas predicativas con rasgos de naturaleza fuerte (sustantivos y verbos), léxicas predicativas con rasgos de naturaleza débil (adjetivos); léxicas de tipo ligero con rasgos de naturaleza fuerte (verbos y nombres soporte), léxicas de tipo ligero con rasgos de naturaleza débil (*preposiciones léxicas* y afijos tipo *-oso*), unidades no predicativas con rasgos de naturaleza fuerte (verbos y nombres cópula) y unidades no predicativas con rasgos de naturaleza débil (*preposiciones funcionales* y afijos tipo *-al*). Añade cuatro clases que no afectan a nuestra exposición. De hecho, estos continuum graduales ya habían sido presentados en otros contextos empíricos, para la descripción de las locuciones preposicionales en italiano (Casadei, 2001).

Otra apuesta importante, por lo que respecta a la gestión unitaria de las relaciones entre preposiciones, adverbios y conjunciones (Pavón, 1995 y 2003), estudia en profundidad las similitudes distributivas entre ejemplos como los que se recogen en el cuadro 2.

Adverbios locativos transitivos	Adverbios locativos intransitivos	Preposiciones
Encima	Arriba	Sobre
Debajo	Abajo	Bajo
Delante	Adelante	Ante
Detrás	Atrás	Tras
Dentro	Adentro	En

Cuadro 2. Comparación entre las propuestas de partículas de Pavón (2003)

En el ámbito de la complementación verbal y la acción de la preposición se desarrollan algunas investigaciones basadas en corpus. El complemento de régimen verbal fue el objeto de una iniciativa en la que se recopilaron datos de 15.000 núcleos para el análisis de las preposiciones afectadas sobre los verbos de régimen y los tipos de argumentos (Alcoba y Del Teso, 2000). Más tarde, desde una visión de la lingüística no discreta, y de nuevo con base de datos para esta clase de complementos, se intenta detallar todos los predicados de verbos de régimen (Cuadros, 2005).

El concepto de categoría vacía motiva algunas causas empíricas en relación, una vez más, con el régimen verbal y el significado de esta clase de palabras (Candalija, 2006). En liza con este concepto de vaciedad (Barrajón, 2014) se discutió el carácter delimitador de ese criterio, en la diferenciación sintáctica de la complementación nocional y local regida, y en la distinción entre la complementación preposicional argumental y la circunstancial (Boer, 1988).

Todavía dentro de la complementación se explicitan las relaciones entre el suplemento y el complemento circunstancial (Tornel, 2006a y 2006b). Avanzan las opciones que versan sobre las relaciones entre clíticos, preposiciones y verbos en la comparación de complementos directos y circunstanciales, mezclándose consideraciones sintácticas y semánticas (Aijón, 2009).

Igualmente, surgen planteamientos originales procedentes de la matemática (Alonso, 2008) que usan el concepto de función para dar cuenta de los vínculos entre los verbos de régimen y las dos preposiciones más comunes, la *de* y la *a*.

Los panoramas globales se suceden (Rodrigues, 2007 y 2008) aunque con una incidencia destacada de la semántica preposicional en perspectiva diacrónica, junto con algunas novedades sobre las preposiciones en perífrasis (Talby-Boulhais, 2003).

Como en décadas anteriores se persevera en la exégesis de piezas concretas, ocupándose de *por* y *para* con valor de finalidad a través de cuestionarios de uso (Bakken, 2014), de *mediante* y *durante* y su gramaticalización (Castro, 2010; y Garachana, 2019) o de *hasta* y su función pragmática (Miyoshi, 2015).

Algunas contribuciones optan por la lengua hablada en zonas concretas (Barrajón, 2006; y Mendizábal, 2012), determinándose los usos locativos prepositivos con los verbos *salir* y *entrar* desde una perspectiva funcionalista, o de locuciones prepositivas en registros orales (Ojeda, 2013).

La orientación diacrónica encuentra vetas de indagación nuevas en los últimos años. La Edad Media y los verbos de régimen (Serradilla, 2001; y Del Barrio, 2004), los textos literarios de esa época (Coello, 2004), las locuciones prepositivas (Castro, 2013; y Granvik, 2014), la complementación (García Martín, 1991) o las frases preposicionales (Báez Montero, 1996) representan intentos de reconstruir desde otros momentos de la lengua la significación de esta clase de palabras. Las relaciones con otras categorías vecinas (adverbios y locuciones adverbiales) sigue aportando materiales valiosos, sobre todo en algunas piezas con un marcado carácter locativo como *ante*, *bajo*, *sobre*, *tras* (De Toledo y Huerta, 2016; y Gallego, 2011). En esa misma línea de casos concretos se ha persistido en una historia de las preposiciones *de* y *por* a partir de un corpus diacrónico desde el latín (Nieuwenhuijsen, 2013; y Trabelsi, 2014); y en una crónica de los relacionantes locativos incluido en la categoría preposicional (Octavio de Toledo, 2016).

Las reconstrucciones desde el latín (Portilla, 2011), su seguimiento etimológico y su proximidad con los adverbios ha llevado a proponer una gradación como la de (39), en consonancia con lo que ya se había apuntado (Horno, 2002).

(39) Nombre – adverbio –preposición léxica – preposición gramatical – prefijos.

Como explicaremos a continuación en la perspectiva teórica (sección 1.5), la gradación categorial ha sido una de las respuestas explicativas, no solo a la polisemia preposicional sino también a su carácter heterogéneo y polimórfico, sintáctica y semánticamente. De hecho, el último binomio de la serie (39): preposiciones-prefijos ha sido objeto de atención y discusión (Amiot, 2004; Ashbury et alia, 2006; y Fábregas, 2010).

Todavía en la línea diacrónica se estudia el paso de las funciones del latín al español tomando las preposiciones en el latín clásico y desde una orientación funcional (Agud,

1980; Hernández Cabrera, 1998; y Martínez García, 2012). También tenemos algún trabajo de referencia en el caso de las preposiciones del griego clásico (Rodríguez Adrados, 1986).

La ampliación descriptiva que ha supuesto el último medio siglo es extraordinaria. Esa riqueza, junto con la sistematización en el trabajo metodológico, ha permitido conocer cuáles son los límites de la categoría.

En la siguiente sección, vamos a comprobar cómo se produjo el protagonismo categorial de la preposición y cómo se fueron tensando esos límites desde la voluntad explicativa de la teoría lingüística. Y comprobaremos, como indicaremos al final del repaso, que ha sido el abordaje del significado de las preposiciones, tanto en el lugar donde deba representarse, cómo en las condiciones y las asunciones que se acepten en su identificación y caracterización, el que ha permitido resolver esas estrechuras explicativas.

1.5 La teoría lingüística y la preposición

A partir de finales de los años 50 del siglo XX se presentará y se desarrollará un nuevo paradigma de la investigación lingüística, conocido como Gramática Generativa y Transformacional (GGT) y liderado por el lingüista estadounidense Noam Chomsky. Sus apuestas, metodológicas y conceptuales, van a suponer un cambio de perspectiva en los objetivos del quehacer del investigador en el lenguaje humano.

El elemento que nos interesa a nosotros en esa nueva forma de observar los hechos lingüísticos es lo que se denominó la “adecuación explicativa de la teoría lingüística”, recogido por primera vez en su obra *Aspects of Theory of Syntax* (1965), del propio Chomsky.

De manera que la tarea del lingüista no es solo observar, describir y analizar los fenómenos de las lenguas (adecuación descriptiva) sino dar cuenta, con ellos y a partir de ellos, de su capacidad predictiva para una teoría general sobre cuáles sean los principios rectores de la facultad lingüística humana (adecuación explicativa), lo que se denominó Gramática Universal. De aquí, la inclusión del concepto de *teoría* en nuestro recorrido. Sin

embargo, y sobre todo hasta la conformación del Modelo de Principios y Parámetros, las contribuciones generativistas al debate descriptivo son imprescindibles para situar el estado del arte en su correcta dimensión.

Como en la sección anterior, el repaso alternará, dentro de un eje cronológico aproximado, la presencia de trabajos punteros o referentes de la comunidad científica con microobservaciones que presentan y explican el papel de la preposición en el seno de las gramáticas humanas. Es cierto, sin embargo, que la mayor parte de las referencias y los lingüistas usan como lengua objeto o de análisis, el inglés. Este hecho no disminuye ni resta interés a sus valoraciones en relación a nuestra exposición porque, además y como diremos, a medida que se articulen sus propuestas (a partir de los años 80 del siglo XX), se irán incorporando otras lenguas tipológicamente más disímiles a las que conforman la familia indoeuropea. Y ese contraste será beneficioso para la descripción y la explicación.

1.5.1 La hegemonía de la teoría sintáctica.

De Chomsky (1965) a Emonds (1985)

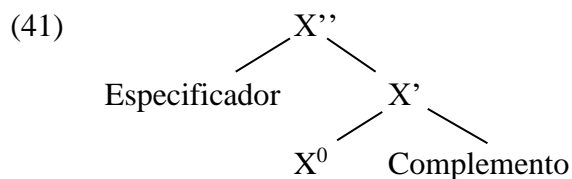
Es, precisamente en la obra chomskiana *Aspects of Theory of Syntax* (1965), donde se introducen consideraciones sobre el componente categorial de las gramáticas de las lenguas naturales. Es cierto, también, que en esa obra no hallamos más que análisis o referencias subsidiarias sobre las preposiciones. Y, de hecho, todavía no se recogen entre las categorías que denomina *primordiales*.

Tendremos que esperar a su trabajo de 1970 para encontrarnos la caracterización más fructífera durante casi veinte años. Lo que se ha dado en llamar el sistema binario de definición de las categorías léxicas [+/-N] y [+/-V]. Según esta, recogida en (40), existen cuatro núcleos: el nombre, el verbo, el adjetivo y la preposición con la siguiente distribución de rasgos.

- (40) N = [+ N, -V]
 V = [- N, + V]
 A = [+ N, + V]
 P = [- N, - V]

Junto a la estipulación categorial, se elabora la teoría de la X' (X-Barra), parte de la gramática que regula las estructuras sintagmáticas o estructuras de frase. De acuerdo con esta teoría, todas las frases o sintagmas de esas cuatro categorías (descritas en (40)) se encuentran encabezadas por un núcleo, y por tanto son endocéntricas. No olvidemos que el estructuralismo americano de Bloomfield, Hockett y Halliday había considerado el sintagma preposicional como exocéntrico (véase la sección 1.3, página 17, nota 7). El núcleo de la proyección es una proyección cero X⁰, restando a derecha e izquierda las posiciones de especificador y de complemento.

La teoría de la X-Barra distingue dos niveles más de proyección tal y como se observa en (41).



Con la asunción categorial y esta estructura proyectiva, la preposición y el sintagma preposicional pasan a tener un protagonismo insólito hasta el momento; no solo por el apartamiento de adverbios y conjunciones, sino por la afirmación del carácter endocéntrico de los sintagmas de las lenguas con preposiciones.

La insistencia en la necesidad de afrontar la preposición de un modo más sistemático y exhaustivo (Jackendoff, 1973) creará una exigencia renovada en su estudio. En (42) se recoge una de las causas de ese desinterés.

(42) *The neglect of prepositions arises from the assumption that prepositional phrases invariably take the form P-NP: if this were the case, prepositions would indeed be dull* (Jackendoff, 1973, pág. 345).

Como vemos, paradójicamente, ha quedado caracterizada en términos negativos, ni sustantiva o nominal [-N], ni verbal [-V]. La preeminencia que tendrá la preposición como núcleo de una estructura sintáctica, se verá sujeta a un cierto desconcierto por el valor negativo de su naturaleza léxica, que será reanalizado años más tarde.

El rasgo que permitió incorporar nuevos datos empíricos en la posición de la teoría lingüística fue un fenómeno clásico a lo largo de la historia: su capacidad en la asignación del caso gramatical (Giorgi, 1991b; y Jolly, 1993). Las distinciones entre los tipos de caso han sido objeto de escrutinio en sucesivos y recientes modelos generativistas (Woolford, 2006; y Bošković, 2006).

Una de las lenguas que mejor refleja la riqueza y la variedad de asignación de caso es el holandés (su gramática dispone de preposiciones, posposiciones y partículas). Una de las fuentes más influyentes del momento (Van Riemsdijk, 1978), y desde la metodología de la Gramática Generativa, indagará en el estatus de la categoría dentro de la teoría sintáctica y recogerá las disputas del momento. Se destaca que las preposiciones son categorías de ligamiento, una noción en boga por aquel entonces (dentro de la denominada Teoría de la Rección y el Ligamiento). La idea es, sencillamente, que se produce una correlación sintáctica y estructural entre ellas, las partículas, y otros elementos de la cadena sintagmática (pronombres referenciales, por ejemplo).

Desde la lingüística italiana (Rizzi, 1984 y 1988) se elabora una de las primeras aproximaciones globales a la categoría dentro, todavía de la Gramática Generativa. Se reconocen etiquetas identificativas en la gramática descriptiva: primarias, propias, secundarias, impropias y adverbiales. Contemporáneamente se asigna el concepto de preposición vacía a ejemplos del italiano (De Boer, 1988). Una de las observaciones interesantes, recogida en (43), es la vinculación en el seno de los sintagmas nominales entre los núcleos y la elección de una preposición determinada.

(43) *La grande maggioranza delle funzioni grammaticali e semantiche degli elementi nominale e quindi codificata tramite la scelta della preposizione appropriata* (Rizzi, 1988, pág. 509).

La comparación, como estrategia heurística de la Gramática Generativa, será presente en otras contribuciones (Kayne, 1984) que se ocupan de las diferencias, por ejemplo, entre el francés y el inglés.

La orientación que están siguiendo las propuestas del momento sobre el componente categorial, donde ha sido encajada la preposición, revuelve dos cuestiones que son cardinales (Emonds, 1985) y que reproduzco en (44) y (45).

(44) ¿Cuáles son las propiedades esenciales de N, V, A y P?

(45) ¿Realizan todas las lenguas el mismo inventario de categorías?

En ambos interrogantes la atención se dirige hacia nuevos terrenos de la investigación: cómo se relaciona la sintaxis con el léxico (lo que en la siguiente sección denominamos “El problema de la correspondencia”) y la necesidad de disponer de datos empíricos interlingüísticos. Se están generando dos conceptos decisivos en la caracterización de la preposición: los rasgos y las proyecciones (Muysken, 1981; y Muysken y Van Riemsdijk, 1985). En el horizonte de los objetivos están los rasgos sintácticos, la emergencia de nuevas categorías (las funcionales) y la distribución de los rasgos en árboles de representación sintáctica. Efectivamente se afianza la búsqueda en las estructuras, y no en las piezas o ítems (Larson, 1988a, 1988b y 1990; y Wunderlich, 1991).

Eran años de debate y de discusión sobre la interpretación semántica de las reglas de transformación y de las categorías (Corbert, Fraser y McGlasham, 1993), y sobre el lugar donde debían incorporarse esas reglas interpretativas. La sintaxis, en cierto sentido, se está quedando corta y se investiga intensamente en el corazón del léxico y de las relaciones de los predicados.

1.5.2 Sobre las categorías funcionales y el problema de la correspondencia entre sintaxis y léxico

La confirmación de la Teoría de la Rección y el Ligamiento (más tarde articulada en el modelo de Principios y Parámetros) como iniciativa explicativa de las condiciones de la Gramática Universal necesita, a su vez, de un conocimiento a fondo de qué está incluido en esas leyes y qué está sujeto a la variación paramétrica (diferencias entre lenguas).

Además, el modelo se aplica a decenas de lenguas distintas al inglés y se empiezan a acumular otras evidencias empíricas que irán nutriendo de alcance descriptivo las generalizaciones teórico-explicativas. Se añaden, por ejemplo, las lenguas semíticas y sus evidencias sobre categorías funcionales en relación con las románicas (Borer, 1984). Se está apostando por el confinamiento de la variación paramétrica en las diferencias morfológicas de estas categorías.

Poco después irrumpirá la distinción entre categorías léxicas y categorías funcionales con todas sus consecuencias (Abney, 1987). Es cierto que esta separación es heredera de la vieja distinción entre clases mayores y clase menores de palabras. Sin embargo, la argumentación es más refinada porque se enmarca –como venimos señalando– en un propósito explicativo global (la llamada teoría de la Gramática Universal).

El punto de partida de la distinción léxico / funcional es el sintagma nominal donde diríamos que es la categoría funcional SD (determinante) la que da cuenta de ciertos hechos empíricos que suceden en su seno, no solo para el inglés sino para otras lenguas (Giorgi, 1991a; y Coene y D’Hulst, 2003).

La caracterización que se hace sobre las categorías funcionales y cómo se distinguen de las categorías léxicas se presenta en (46).

(46) *Rasgos de las categorías funcionales*

(a) Pertenecen a una clase léxica cerrada.

(b) Pueden ser fonológica o morfológicamente dependientes de otros materiales.

- (c) Adoptan solamente un tipo de complementos.
- (d) Tienen complementos que no son argumentos, es decir, no son marcadores θ .
- (e) Son inseparables de sus complementos (no pueden ocurrir de manera intransitiva o ser abandonadas).
- (f) Tienen ausencia de contenido descriptivo. Apenas tiene una contribución semántica de segundo orden.

Las categorías funcionales ayudan a la interpretación de las categorías léxicas que toman como complementos, y sus proyecciones son extensiones imprescindibles para comprender las dependencias internas de los predicados (Cormack, 1994). La nómina de categorías funcionales sufrió una progresiva inflación en la década de los noventa del siglo XX tanto en el ámbito nominal (SNúm, SGrado, SPoss...), como en el verbal (SCaso, SConc, SFlex...).

Asistimos a unos años de exploraciones que descubren las relaciones entre las categorías léxicas y las funcionales (Corbett, Fraser y McGlasham, 1993). Muchos de los análisis se detienen en la tarea preposicional en esos vínculos. Así, se identifican asimetrías restrictivas a favor de las funcionales (Fukui y Speas, 1986) y se relacionan con la estructura temática, el criterio temático y la jerarquía temática (Fillmore, 1968; Vez, 1982; y Speas, 1990).

Se inician exploraciones sobre cómo debe ser la relación entre las proyecciones funcionales y las léxicas (Bartra y Suñer, 1992; Borer, 1993; Williams, 1994; Gelderen, 1993; Hoekstra, 1995; y Belletti y Rizzi, 1996), situando la lupa en las relaciones argumentales entre núcleos, complementos y cómo se asignan los papeles temáticos en toda suerte de estructuras: combinaciones preposicionales y locuciones (Bosque, 1993; Martínez Linares, 1999; y Jiménez Juliá, 2002). Uno de los fenómenos que había sido usado, años antes, en la descripción de estas relaciones fue el *stranding* (abandono) preposicional (Hornstein y Weinberg, 1981). Ejemplos como *¿De qué edificio está cerca? (la facultad)* llevan a microobservaciones (Plann, 1989; y Campos, 1990) sobre este fenómeno.

Precisamente, y bajo la influencia del análisis anglosajón, se considera a *cerca* preposición cuando no figura en la nómina de preposiciones de la gramática normativa y descriptiva. Esta división y/o convivencia de las propiedades funcionales junto a las léxicas se investiga, también, desde lenguas tipológicamente distintas, como el ruso u otras lenguas eslavas (Yadroff y Franks, 1999; Zajara-Nekrasova, 1993; y Biskup, 2019).

Cómo se exprese el significado en el entramado sintáctico se irá indagando en aquellos años (Hout, 1989). Se responde a qué papeles temáticos asigna la preposición, en qué posiciones se descargan los papeles temáticos y si surgen diferentes tipos de sintagmas preposicionales a partir de esas averiguaciones. El papel del caso sigue ocupando el interés teórico sobre todo en la distinción entre los casos morfológicos y las relaciones de caso de las relaciones gramaticales (Cook, 1989; y Blake, 1994). No en vano, se va intentar simplificar o reducir la tarea preposicional desde el caso (Solías, Pisabarro y Cardeñoso, 2002).

Se observa, por tanto, como se produce la transmisión casual –un marcador sintáctico- y ese interés, se comparte con la asignación de los papeles temáticos o semánticos, una noción de índole sintáctico-semántica.

De hecho, ejemplos tan usados y reutilizados en la descripción, como los de (47) serán objeto de escrutinio ahora desde la labor de la preposición en la construcción.

(47)

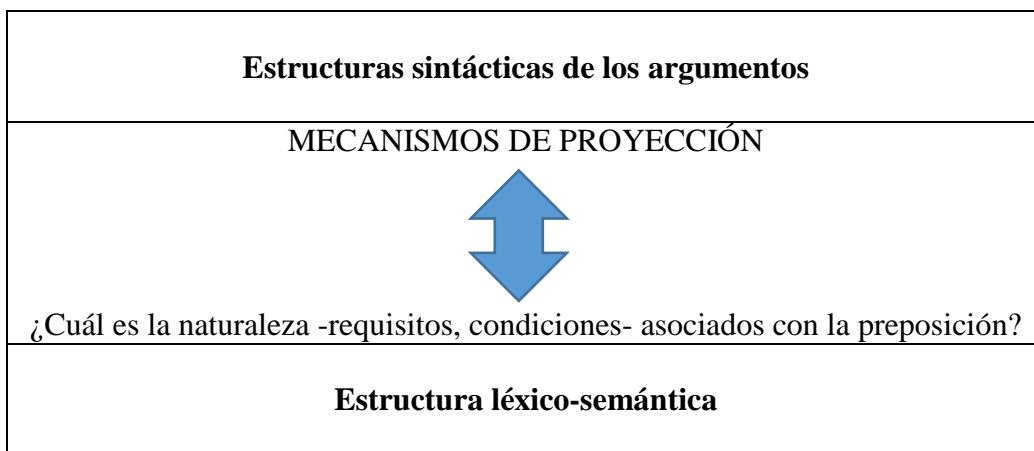
(a) Veo *a* Miguel

(b) Pedro le dio el libro *a* Pablo

Así, en la Gramática Generativa para el español y el catalán se empiezan a caracterizar ciertos usos preposicionales, asociados con la asignación de caso acusativo o dativo y sus efectos argumentales como los de (47) (Royo, 2018). Las relaciones entre el acusativo y la naturaleza locativa de algunos de esos verbos también ha sido comentada (Fábregas, 2015). De otra parte, y desde la perspectiva de las categorías funcionales (Brugè y Brugger, 1993), se están valiendo de rasgos como [+ Animado] y de la proyección funcional [+/- Acusativo].

El estudio de la preposición en el seno de las relaciones predicativo-argumentales supondrá uno de los espacios de análisis más fecundos del momento (Demonte, 1991). Estamos asistiendo a un esfuerzo analítico por enfrentarse a uno de los problemas más sofisticados de la teoría gramatical, lo que se denominó el problema de la correspondencia es decir, cómo se proyectan las relaciones entre las estructuras léxico-semánticas y las estructura sintáctico-argumentales y cuál sea el papel transmisor de la preposición en ese camino, tal y como representamos en (48).

(48)

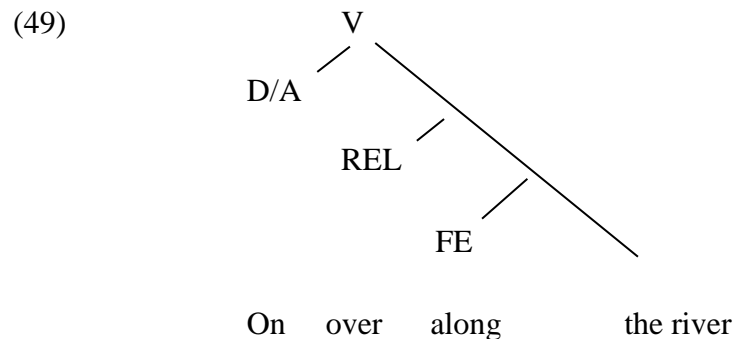


Las investigaciones punteras en su momento (Hale y Keyser, 1991 y 1993) sientan las bases sobre cómo se producen las relaciones entre la sintaxis, la estructura argumental y los rasgos de las categorías. Más adelante (Hale y Keyser, 2002; y Reuland, Bhattachanya y Spathas, 2007) se reconoce que en el esquema de (48) la estructura argumental está determinada por las propiedades del ítem léxico. De hecho, ya se había intervenido en esta relación con una propuesta minimizadora del vínculo (Baker, 1997). El examen de la naturaleza de esa correspondencia ha supuesto una mejora en las estipulaciones léxico-conceptuales de las interficies (Mateu, 1999 y 2002; y Alsina, 1996), de la sintaxis relacional y de la semántica relacional.

En consonancia con esa perspectiva, al investigar en esa *relación* y/o esas *propiedades*, se está indicando (Torrego, 1998) que el marcaje manifiesto de objetos

acusativos obedece a patrones regulares de comportamiento enraizado en la semántica de predicados. Efectivamente, el interés por determinar qué clase de preposiciones sean las *a* de acusativo y dativo (Demonte, 1991), el estudio de los verbos preposicionales (Demonte, 1994; y Gallego, 2010b), los microparámetros derivados del contenido de las preposiciones (Demonte, 2002), o las consecuencias de la interacción con el aspecto de los verbos con expresión eventiva en (Demonte, 2011) son algunas de las propuestas que, desde la sintaxis generativa, incluyen la articulación de asunciones semánticas.

El ahondamiento empírico de esa articulación incorpora, de nuevo, datos procedentes del holandés y su complejísima diversidad adposicional (Cuyckens, 1995; y Neeleman, 1997a y 1997b). Se realiza una mejora descriptiva de la sintaxis de los complementos preposicionales, y se concluye que el verbo y el sintagma nominal contenido del sintagma preposicional están temáticamente relacionados. En ese vínculo, se propugnan (Klippel, 1997) proyecciones semánticas como functor espacial (FE), relación locativa (REL) y dirección/aspecto (D/A); y las tres tendrán una posición proyectiva en el esquema del SV, como se aprecia en (49)



De hecho, la conexión y la dependencia que se produce entre la preposición y su contexto se valoran como determinantes en (50)

(50) Determining the semantic properties of prepositional elements is complicated by the fact that their interpretation depends on the other constituents with which they occur. They vary in interpretation according to both their complements (usually an NP) and to the elements they complement or modify (Neeleman, 1997a, pág. 100).

Esa senda de expectativas siguió activa muchos años. De hecho, tanto la adscripción funcional de la preposición como cuál sea su papel en la transmisión de caso, los papeles temáticos o estructuras conceptuales *ad hoc*, llevaron a dibujar una serie de criterios semánticos para caracterizar las categorías funcionales: significado abstracto, significado no independiente, asignación- θ , relación con la ontología, relaciones de sentido, coerción de significado, significado flexible, proyección semántica, y restricciones de selección (Muysken, 2008). Por último, se propusieron modelos construccionistas que vinculasen las preposiciones con los contenidos de sus sintagmas (Tungseth, 2008).

La puerta de acceso al significado se está abriendo poco a poco; antes, sin embargo, vamos a ver cómo se intenta, en paralelo a la búsqueda del significado dentro de la sintaxis, flexibilizar primero y abandonar más tarde el binarismo sintáctico-categorial que arrancó nuestra sección.

1.5.3 La gradualidad categorial: un camino hacia el significado

La necesidad de flexibilizar los requisitos categoriales de las preposiciones no procede, solamente, de las evidencias sobre funcionalidad y cómo se adapten a los desarrollos teóricos, sino también de la dificultad para explicar el licenciamiento temático en determinados predicados (Stewart, 2015). Así, en el ejemplo de (51), algunos lingüistas se fijan en cómo se realice el traspaso del papel temático (Neeleman, 1997a).

(51) John has always believed in Bill's honesty

Complementos preposicionales como *in Bill's honesty* de (51) muestran que existe un desencaje entre su sintaxis y su semántica. Parecería que la preposición *in* y el sintagma determinante *Bill's honesty* forman un solo constituyente; sin embargo, en la interpretación semántica el verbo *believe* y la preposición irían juntas. No en vano, este problema sucede en los ejemplos de complementos de régimen del español. ¿Quién determina la composición semántica de los sintagmas preposicionales de (52)? ¿El significado de *por*, *en*, *de* y *por*, o el de sus correspondientes parejas *abogar por*, *insistir en*, *abusar de* u *optar por*?

(52)

- (a) Abogan *por* la supresión de la ley
- (b) Insisten *en* la vieja teoría
- (c) Abusa *de* su confianza
- (d) Optó *por* el camino más corto

Se requería un paso nuevo que estipulase otro grado o fase en el componente categorial. A través de diversidad de ejemplos empíricos para lenguas muy diferentes, se propusieron las denominadas categorías semiléxicas (Van Riemsdijk, 1996).

Con la multiplicación de categorías funcionales crece el número de especificadores, produciendo dominios redundantes de endocentricidad (o posiciones vacías). Los núcleos sintácticos tendrán tres tipos de rasgos: categoriales, de nivel y de funcionalidad; y se propondrán tres tipos de núcleos: léxicos, funcionales y semiléxicos. Asimismo, se consideran las dependencias particulares entre nombres y preposiciones. Estas relaciones se basan en una dependencia ascendente, gracias al conocimiento léxico sobre la información del nombre regido elegimos la preposición (en el caso del V-P es al revés). Por último, se concluye que las preposiciones ocuparían el nivel intermedio, la semilexicidad.

Se abandona la clasificación de las piezas y se las localiza en otra proyección sintáctica de acuerdo a una serie de requisitos de carácter léxico-semántico.

La caracterización que se hace de esta nueva clase de núcleos se sintetiza en (53).

(53)

- (a) Los núcleos semiléxicos N, V, A y P no tienen rasgos semánticos puros.
- (b) Son elementos morfológicamente complejos que tienen un nudo léxico y un sufijo funcional.
- (c) Son ranuras por las que se mueve el material léxico, sin selección argumental.
- (d) Son núcleos léxicos sin contenido semántico intrínseco, y sirven para expresar sintácticamente determinadas posiciones.

En este nuevo escenario, ni léxico, ni funcional, y como sucediera con las categorías funcionales, una parte de las pruebas empíricas y de los datos son del holandés. La insuficiencia del binarismo categorial de la Gramática Generativa y la necesidad de buscar en el universo semántico de la preposición ciertas explicaciones sintácticas (Zwarts, 1992 y 1997) constituyen las dos direcciones en las que avanza la flexibilización de los rasgos categoriales.

Con la combinación de prescripciones de semántica composicional y cognitiva se desarrolla una tipología preposicional que incluye locativas (proyectivas y no proyectivas) y direccionales (Zwarts, 1999; y Terzi, 2007). Esa senda construye mapas de significado espacial con nociones como el aspecto preposicional y la noción de la convexidad (Rice, 1999; Ursini y Akagi, 2013; y Zwarts, 2016 y 2017).

Otra de las vías en la que se adentra la perspectiva teórica son las pruebas sobre los parecidos y las diferencias de la preposición con otras categorías, la que fue nuestra pregunta (3) de la introducción: ¿Cuáles son las relaciones de semejanza –de la preposición- con otras categorías gramaticales? (Jacobsson, 1977).

Para ello asistimos a la comparación con lenguas de otras familias distintas a las indoeuropeas (Déchaine, 1989 y 1993). En esa búsqueda de analogías se recopilan evidencias del haitiano, el igbo -hablada en Nigeria-, el yoruba –localizable en Benín, Togo y Nigeria-, o el kabiyè –también en Togo- (Kézie, 1996). En esas lenguas, muchas de las funciones asociadas a preposiciones y adverbios ingleses se expresan con verbos. Más recientemente se ha seguido con estudios preposicionales en la familia lingüística kwa (Aboh, 2010), o en el kinyaruanda (ruandés) (Zeller y Ngoboda, 2018). Esta misma equivalencia sucede con algunas lenguas criollas como el krio de Sierra Leona (Van de Vate, 2006), o el kru y el kamtok, pidgin ingleses de Camerún y Ghana, respectivamente. En el continente americano se examinó el particular comportamiento de las preposiciones en otomí, lengua hablada por unos 200.000 hablantes en México, y con correspondencias entre marcas morfológicas y nuestras preposiciones de objeto (Lastra, 1992).

Estas diferencias interlingüísticas plantean la cuestión sobre cómo se determina el estatus predicativo o referencial de una proyección (¿de nuevo el problema de la

correspondencia!) (Romeu, 2017). Se sugiere que las preposiciones sean núcleos semánticos de los verbos. Este hecho no pasó inadvertido realizándose reanálisis de verbos como preposiciones, fenómeno extendido en muchas lenguas germánicas y románicas (Kortmann y König, 1992); o señalándose el necesario vínculo entre nombres y preposiciones a causa del factor casual (Schumacher, 2018).

Otra vez serán los rasgos semánticos, incorporados durante la computación sintáctica, los que acuden en la caracterización preposicional (Déchaine, 1993). El rasgo [+Referencial] se defiende como sustancial en la discriminación de unos ejemplos y otros.

El cuestionamiento del marco categorial, a través de estipulaciones semánticas, continúa activo (Tremblay, 1998). Se aboga por consideraciones donde el recurso a categorías *light* permitiría considerar las preposiciones como casi funcionales (Déchaine, 2005) a partir de conjuntos de rasgos como [+/- Relación], reduciendo las diferencias categoriales a rasgos morfosintácticos (Brugman, 2001), o a otra clase de categorías funcionales (Gallego, 2010a y 2016).

Una parte de la discusión relativizadora del marco categorial se está produciendo dentro del Programa Minimista de inspiración generativista donde la herramienta explicativa del rasgo es fundamental en la comprensión del sistema de computación lingüística humana (Chomsky, 1995; y Brody, 1998).

En el Programa Minimista se determina la existencia de dos tipos de rasgos: los rasgos interpretables y los no interpretables. Estos últimos no tienen un valor semántico, interpretable por el Sistema Conceptual-Intencional (Mendikoetxea, 2007). De manera que serán suprimidos durante la derivación, es decir, durante la construcción de los objetos lingüísticos apropiados. Los rasgos están en el centro de los mecanismos de derivación de las cadenas, desplazando las estipulaciones preliminares sobre categorías (Varas, 2008). Desde la perspectiva minimista las categorías funcionales poseen uno o varios rasgos no interpretables paralelos a los rasgos interpretables de las categorías léxicas que seleccionan o con las que crean proyecciones extendidas (Kayne, 1999, 2000 y 2007). Estos hechos fueron observados en lenguas alejadas del patrón anglosajón, como el ruso (Yadroff, 1999).

No olvidemos que no han cesado de recogerse evidencias sobre este tipo de partículas (ni funcionales, ni léxicas), desde predicados y verbos *light* (*hacer, poner, llevar*) hasta clasificadores o cuantificadores pasando por ciertas preposiciones para lenguas tan diversas como el inglés, el vietnamita, el bengalí, el checo o el urdu (Löbel, 1999; Corver y Van Riemsdijk, 2001; y Butt y Geuder, 2001).

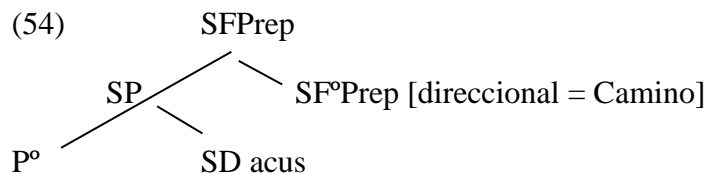
Esta flexibilización del componente categorial lleva a distinguir tres tipos de preposiciones pero recuperando términos anteriores: léxicas, regidas y gramaticales (Rauh, 2002). Asimismo, se establecen subclasificaciones superpuestas como espaciales, temporales, escalares, transitivas, intransitivas. Crecen, en ese escenario de relativización y gradualidad categorial, las adscripciones a diversos grupos, semifuncionales (Botwinik-Rotem, 2004 y 2008), o semiléxicas y semifuncionales para el alemán (Bayer y Bader, 2007). De otra parte, se observaron núcleos semiléxicos no preposicionales tales como locuciones del tipo *As for as* para el inglés (Rafel, 1999); se incluyeron las preposiciones entre el lexicón funcional (Borer, 2005); o se las consideró abiertamente categorías semiléxicas con datos del rumano (Zegrean, 2007; Mardale, 2011; y Fagard y Mardale, 2019). Recientemente se ha dedicado un monográfico a tratar sobre la semilexicidad en algunos casos de la gramática del español (Sanromán y Carrasco, 2019). Y aunque menos habitual, también se sostenía que las adposiciones no forman parte de las categorías léxicas sino del sistema funcional existiendo más terreno en común con pronombres, determinantes y complementizadores que con nombres, verbos y adjetivos (Baker, 2003). De hecho, las adposiciones han representado un problema categorial dentro de la morfología distributiva (Deacon, 2014).

Junto a la figura del rasgo, en cualquiera de sus variantes explicativas y metodológicas, otro concepto relevante fue el de *partícula* (Baker, 1997). La noción, más amplia y versátil, permitió revelar más hechos semánticos y perseverar en la crítica a la rigidez categorial (Dehé, Jackendoff y McIntyre, 2002), echando mano de las combinaciones de verbo con partículas en los verbos frasales, y de modelos diversos como la *Lexical Functional Grammar* (Bresnan, 1982) o el Programa Minimista (Chomsky, 1995). Vuelven las conexiones entre categorías, como verbos y preposiciones (Guéron, 1990) considerándose las partículas como preposiciones intransitivas y subrayándose que

el potencial semántico de las preposiciones depende de los núcleos y complementos que enlaza (Arad, 1995 y 1999).

Pero la noción de partícula no había de dirigirse solo al caudal semántico (Berenguer, 1992), también sirvió para revisar ciertos hechos sintácticos sobre la causatividad (Den Dikken, 1995), las adposiciones locativas (Den Dikken, 2003; y Romeu, 2012a) y las estrechísimas relaciones entre la morfología y la sintaxis sobre todo en alemán y holandés (Zeller, 1997 y 1999).

Más tarde, se vuelve sobre las proyecciones sintácticas funcionales (Zeller, 2001) y reaparecen más rasgos como se propone en (54).



El esquema de (54) es una de las muchas opciones de encajamiento para el holandés y que combinan, con relativa frecuencia, espacios representacionales con rasgos sintácticos y semánticos.

Se logran algunas síntesis muy representativas de la batería de soluciones que han venido sugiriéndose para encajar el puzzle semántico en las estructura sintácticas (Svenonius, 2006, 2007, 2010 y 2012). Y así, con renovados datos empíricos de lenguas muy ricas en estructuras adposicionales, se investiga en la incidencia sobre la expresión espacial con rasgos lingüísticos específicos y el uso de nociones semánticas de la gramática cognitiva. El objetivo que se persigue es dar cuenta de la elasticidad que presenta la conducta de las preposiciones (Hottenroth, 1991; Saint-Dizier y Vázquez, 2001; y Evans, 2007).

Incluso en investigaciones realizadas en el seno de la HPSG (Pollard y Sag, 1994; y Tseng, 2005) se acata la premisa de que los usos preposicionales son como espectros, entendidos como una secuencia continua, reiterándose la dicotomía entre funcional y léxica.

Un recurso más ha sido acudir a la emergencia de rasgos como *límite* –en las configuraciones que se producen en los sistemas de computación, siguiendo las asunciones minimistas-, y prescindiendo de proyecciones privativas o exclusivas (Real, 2013).

Estas miradas sobre los procesos computacionales conducen a la creación de ciertas cartografías en las estructuras sintácticas. La tarea analítica consiste en ofrecer mapas que den cuenta de posiciones en ejemplos concretos, como los sintagmas preposicionales espaciales (Cinque y Rizzi, 2010), las adposiciones en español (Romeu, 2012b y 2014) o los complementos direccionales (Fábregas, 2007).

Por otra parte, y ahora desde la lingüística computacional, comprobamos que la noción de gradualidad también se ha utilizado en el análisis del aspecto verbal desde la experimentación psicolingüística y en contextos perifrásticos del español (Aparicio, Castellón y Coll-Florit, 2011 y 2013). Ha servido, por ejemplo, para informar sobre el etiquetaje a verbos como *enfriar*, *ensanchar*, *envejecer* y *engordar*. Es posible la analogía de la gradualidad en la medida en que tanto verbos como preposiciones ofrecen evidencias empíricas, aunque en el caso del significado preposicional sus valores sean contextuales y dependientes. La flexibilidad que se reconoce en los significados se relaciona con la polisemia y otras características de los rasgos y las clasificaciones de las piezas verbales que fueron tratados desde la filosofía del lenguaje (Vendler, 1967).

1.6 Contrastes y evidencias interlingüísticas

Uno de los impulsos metodológicos, tanto de la descripción como de la teoría lingüística en relación con el estudio de la preposición, procede de la aportación y la observación de nuevos datos empíricos recaudados en multitud de lenguas y en comparaciones entre lenguas afines tipológicamente o no. Qué sea y cómo funcione una preposición no dependerá, en exclusiva, de una sola gramática sino de una compleja red de fenómenos intergramaticales. Igualmente, y desde el área de la traducción automática y la lingüística contrastiva, se han venido realizando toda suerte de cotejos de casos, sistemas

y estructuras con el propósito de mejorar las técnicas y los resultados en la conversión de una lengua en otra.

Ambos motivos nos llevan a incluir en esta panorámica algunas referencias que sean útiles –no solo por la voluntad de completar el balance enciclopédico- sino también por el valor que tendrá en la presentación de la hipótesis.

Hemos organizado algunos trabajos por lenguas y/o familias lingüísticas. En todos ellos, de un modo u otro, se explican fenómenos, construcciones o piezas de la categoría preposicional o versiones semejantes como posposiciones, circunposiciones o partículas.

Empezamos por las misceláneas que reúnen una variedad de casos y ejemplos. El interés por los casos dativo y elativo, sus marcadores y sus clases en lenguas muy diferentes tipológica y geográficamente: chino mandarín, húngaro, camboyano, somalí, inglés español, inglés o swahili ha sido objeto de comentario (Blansitt, 1988; Blume, 1998; Delbecque y Lamiroy, 1998; y Surányi, 2009), estableciéndose correlaciones entre los tipos de marcadores y el orden de los casos en esas lenguas.

El concepto de adposición, englobador de estructuras antepuestas y pospuestas, cuenta con un interesante estudio tipológico de carácter universal (Chapado, 1966). Otros análisis menos globales (Kurzon y Adler, 2008) también incluyen locuciones, partículas y frases adverbiales en lenguas disímiles como francés, inglés, hebreo, coreano, turco, ruso y ucraniano (Borillo, 1993). Asimismo, se eligen contextos como los predicados de movimiento del inglés, alemán, holandés, checo y ruso (Gerhke, 2008).

La síntesis más completa de contrastes interlingüísticos contiene 434 lenguas, con preposiciones, posposiciones y circunposiciones, que nos permite observar, apreciar y comparar con mucho detalle un espectro muy completo de posibilidades gramaticales, en su organización, su distribución y sus repertorios (Hagège, 2010).

En nuestro contexto tipológico, el de la familia románica, y ocupándonos ya de obras para una sola lengua, hallamos una revisión de la historia de la categoría en catalán, junto a una aplicación de la semántica de prototipos o algunos casos concretos de esa lengua (Martí y Miñana, 1988; Sancho, 1994; y Alturo, 2004). Del mismo modo se

comparan también las preposiciones comitativas (*con*) y participativas entre español, occitano, catalán, italiano y gallego (Rigau, 1990). En esa misma tradición geográfica se presentan aspectos sobre la dirección y el espacio en la preposición francesa (Zwicky, 1987), y más recientemente; del universo espacial dentro de la cultura francesa (Sirvent, 2005). La comparación entre piezas de dos lenguas distintas, la *até* portuguesa y la *hasta* española, muestran la calidad de los argumentos comparativos (Grolla, 2004). Los complementos directos e indirectos del gallego han requerido de atención (López Martínez, 1993), así como las llamadas preposiciones articuladas del italiano *alle*, *della*, *nel* o *sul*, (Napoli y Nevis, 1987), o el acusativo preposicional en esa misma lengua (Zamboni, 1992). Todavía en Italia se han abordado las preposiciones combinadas, *tradi noi*, *sudi noi* (Franco, 2018).

Todavía en Europa, y dentro de la gran familia anglogermánica, tenemos descripciones genéricas para el inglés (Lindstromberg, 1997; e Infante, 2005) o microobservaciones para *in* y *on* (Guardón, 2003) que utilizan el concepto de trayectoria (Jackendoff, 1990). Cuando las preposiciones disponen de flexión, como en el escocés, se descubren comportamientos sintácticos como el abandono preposicional (*stranding*) (Borsley, 1986), que ya citamos en la sección 1.5.2, página 38 de este capítulo.

Dentro de la familia germánica se ha investigado en las lenguas nórdicas o escandinavas y los objetos dobles en sueco e islandés (Holmberg, 1990); junto al laberíntico sistema de 25 casos del finlandés (Holmberg y Nikanne, 1993; y Nikanne, 1993).

Volviendo al continente tenemos que mencionar una contribución que versa sobre la representación del espacio en el sistema de preposiciones en eslovaco (Altman, Dömötöv y Riska, 1968). En polaco también se revelan las sofisticadas relaciones entre la asignación del caso y las preposiciones (Borsley y Jaworska, 1988).

Las opciones que hemos denominado intermedias, como las categorías semiléxicas, se han localizado en lenguas asiáticas, como los clasificadores en blanga, lengua indoiriana hablada por los Bengalís en el sur de Asia (Bhattacharya, 1999); y se han examinado las preposiciones en persa con una metodología funcional (Bubenik, 2006; y Rezai, 2012).

Otro ámbito tipológico esencial para comprender este abanico heterogéneo de las evidencias interlingüísticas lo encontramos en África. Retengamos un caso paradigmático: las lengua kwa, una rama de la familia de lenguas Níger-Congo habladas en el sudeste de la Costa de Marfil, en Ghana, Togo y Benín y en la parte sudoeste de Nigeria, y usadas por unos 20 millones de personas. Una de esas lenguas, el igbo, usa recursos como sufijos extensionales, complejos verbales y serialización verbal (Déchaine, 1989; e Ifeganyi, 2016) para aumentar el escaso número de preposiciones en esa lengua.

Algunos de los materiales elegidos por la lingüística contrastiva y la traductología se sintetizan en el cuadro 3 que recoge las lenguas que se comparan y los datos de los trabajos. En todos ellos se ponen de manifiesto, no ya las evidentes variaciones entre lenguas diferentes, sino la casuística de adaptaciones y modificaciones de los sistemas gramaticales en relación a las preposiciones u otros instrumentos expresivos con valores léxico-semánticos semejantes. Los contrastes comparativos, precisamente, contribuyen a iluminar rasgos específicos de la categoría y su materialización en sistemas lingüísticos distintos.

Datos de la investigación (autor y año)	Lenguas en contraste
Palet (1987)	Español y catalán
Cienki (1989)	Inglés, polaco y ruso
Folli y Ramchand (2005)	Inglés e Italiano
Waluch-de la Torre (2007)	Español, polaco y portugués
González Hernández (2007)	Español y francés
Izquierdo (2008)	Español e inglés
Jiménez Julià y Lübke (2012)	Adposiciones alemanas y preposiciones españolas
Trovato (2013).	Régimen preposicional entre italiano y español
Johansson (2014)	Inglés y sueco
San Vicente (2015)	Gramática comparada entre español e italiano
Rodriguez (2015)	Español e inglés
Luo (2016)	Español y chino
Mikelenic (2017)	El complemento preposicional entre croata y español
Biyun Li (2018)	Chino y español (construcciones de régimen)

Cuadro 3. Selección de aportaciones de la lingüística contrastiva relacionada con la preposición

1.7 Perspectivas en el estudio del significado de las preposiciones: de los usos concretos a las teorías cognitivas

Como hemos señalado, la necesidad de abordar el significado de las preposiciones ha sido una constante en cualquier periodo cronológico del estudio gramatical y desde cualquier perspectiva. El énfasis, las asunciones y las metodologías se han ido matizando y modulando. En esta sección dirigimos nuestra atención a las perspectivas que, explícitamente, intentan responder a tres cuestiones que se presentan en (55), siendo, en cierto sentido, una prolongación de la pregunta (1) de la sección 1.1, página 4.

(55)

- (a) ¿Qué es el contenido léxico de una preposición y dónde se interpreta?
- (b) ¿Cómo se relacionan los diferentes significados de una preposición, la polisemia preposicional?
- (d) ¿Cuál es el papel del contexto en la identificación del significado preposicional?

Vamos a organizar la sección en tres bloques, recogidos en los rótulos: (i) la semántica descriptiva y/o composicional, (ii) la semántica formal y (iii) la gramática cognitiva, con el propósito de recoger las consideraciones más notables entorno a esas preguntas centrales que se resumiría en qué modo se caracteriza y se interpreta la contribución semántica de las preposiciones (Coventry y Ludwig, 1991).

1.7.1 La semántica descriptiva y/o composicional

Los manuales de uso y descripción de la semántica preposicional para el español suelen organizar los valores de significado entorno a tres grupos de modalidades: los locativos o espaciales, los temporales y los nocionales (modo o manera, finalidad, causa, medio, tema, condición...). Se acepta, por tanto, que el significado puede ser atribuible, directamente, a las piezas preposicionales.

De hecho, cualquiera de los trabajos reconocidos y ya citados en la sección descriptiva (Osuna, 1991; Morera, 1991; Fernández López, 1999; Slager, 2004 y 2010; y Moliner, 2012), beben de una concepción donde el uso del hablante restringe, orienta y determina el significado (Barretta, 1979). Así, en la muestra de ejemplos de (56), se sintetizan algunos de los ejemplos con los valores que son usados en las descripciones semánticas.

(56)

- (a) Lanzaron agua *contra* los manifestantes
- (b) Tengo clase *por* la mañana
- (c) El centro se encuentra *hacia* las afueras
- (d) Pablo sale *dela* oficina *a* las ocho
- (e) Comprarán un regalo *por* su cumpleaños
- (f) *Con* asistir a la visita tienes suficiente
- (g) El enchufe está *a* la derecha del armario
- (h) Nos dejaron el piso *en* un millón de pesetas

En una visión compositiva del enunciado cualquiera de las piezas y/o usos recogidos en (56) es susceptible de ser caracterizada como una parte más del significado global o total del enunciado. Este punto de partida, el principio de composicionalidad, según el cual el significado de una expresión compleja como las de (56) es una función del significado de los elementos que la compone y de las relaciones sintácticas que se dan entre ellas, está en la base de la mayoría de investigaciones, usen una u otra metodología.

En su quehacer semántico, han sido muy sustanciales las aportaciones que sitúan las preposiciones como asignadoras de papeles temáticos en la predicación (Napoli, 1989; Rauh, 1991, 1993 y 1994). Así, sobre el traspaso temático, se afirman en (57) tres hechos que son indiscutibles.

(57)

(a) Son la expresión de los papeles temáticos que un verbo atribuye en su contexto a los sintagmas nominales.

(b) Son las transmisoras de los papeles temáticos.

(c) Son ellas las que atribuyen los papeles temáticos.

Son tres estadios diversos respecto de cómo interviene la partícula en la relación semántica que se establece dentro del predicado. Pero quedan muchas preguntas que se desplegarán, con más detalles, en las asunciones semánticas de nuestra tesis, en el Capítulo 2, página 62 y siguientes.

El acercamiento a las cuestiones teóricas sobre el significado preposicional arranca, como en tantas otras cuestiones, en la tradición anglosajona. Desde el análisis de las 38 preposiciones del inglés en el seno de una semántica estratificacional hasta la importancia del contexto en la variación, la identificación y la interpretación de sus significados han sido presentadas a lo largo de casi un cuarto de siglo (Lehrer, 1974; Bennet, 1975; Bolinger, 1977; Taylor, 1989; y Zelinsky-Wibbelt, 1993).

Las herramientas de análisis para la semántica preposicional han sido diversas. Partiendo de la visión léxico-conceptual se ha utilizado el cálculo- λ para señalar los alcances de su significado (Saint-Dizier y Vázquez, 2001). Se han construido, también, redes semánticas dentro de proyectos muy ambiciosos como PrepNet, en la línea de WordNet, VerbNet y EuroWordNet, donde se elaboran marcos clasificatorios que permitan incluir todos los tipos y subtipos expresivos (Saint-Dizier, 2006c).

En una de las panorámicas sobre semántica preposicional de carácter interlingüístico (vasco, inglés y español), más sobresalientes (Saint-Dizier, 2006a) se considera el contexto como el marco imprescindible de la interpretación. Esta idea, lejos de ser nueva, refuerza la necesidad de vencer el corsé sintáctico-categorial y rastrear en el entorno con exigencia y perseverancia. En (58) se sintetiza esa visión.

(58) *Perhaps a large part of the reason for this is that the semantics of a transitive preposition can be bleaded and determined largely by the semantics of the head noun it governs, or its governing verb. However, many prepositions also have predicative usages* (Saint Dizier, 2006, pág. 197).

Esta perspectiva nos dirigirá al distribucionalismo de Zellig Harris (1951) que nos ocupará en las asunciones semánticas de nuestra hipótesis y dentro de nuestras estrategias metodológicas para acceder a esos significados.

En el entorno francófono, debemos aludir a una tradición semántica que, iniciada a mediados del siglo XX (Pottier, 1961) –comentado este capítulo en la sección 1.3, página 16- insiste en la tarea analítica con denominaciones, algunas ya conocidas, como preposiciones vacías (*vides, a*), para las funcionales; y otras nuevas como preposiciones incoloras (*incoloras, de*), para las semifuncionales (Marque-Pucheu, 2008).

Se disponen de diversas metáforas explicativas, entorno al color y las restricciones categoriales para entender cómo significan las preposiciones (Spang-Hanssen, 1965; y Gaatone, 2004). Es un itinerario, el francófono, muy completo, y que apuesta por descripciones de los valores léxicos dentro de universos de uso muy abiertos (Cervoni, 1991; Vandeloise, 1993 y 1994; Cadiot, 1997; Leeman, 2008; Tremblay, 1999; y Aurnague y Col, 2017).

1.7.2 La semántica formal

Desde la lógica y la filosofía del lenguaje en el siglo XX la preposición se ha visto como la mediadora en las tareas predicativo-argumentales en contextos verbales (Davidson, 1967). En ese análisis de oraciones de acción se parte de la premisa de que

conocer el significado de una oración de esa clase implica conocer sus consecuencias lógicas, expresadas en su forma lógica. Una de las propuestas del análisis es separar los verbos de las preposiciones. Pero el interés en este tipo de mediaciones está sujeto a problemas particulares en la medida en que, una vez más, parece que la interpretación entre verbos y preposiciones está condicionada y/o correlacionada de modos diversos (Laur, 1993). Se ha articulado una posible solución, cuyo comentario exhaustivo ultrapasa, por su especialización técnica, la intención de nuestra crónica (Barrero, 2012).

A partir del análisis seminal de Davidson se han sucedido diferentes incursiones que relacionan la categoría preposicional con la cuantificación y los adverbios (Verkuyl, 1973). En otras ocasiones (Creswell, 1985) se indagan las preposiciones inglesas en la determinación de los puntos de vista, dentro de una semántica de condiciones de verdad.

Dentro de otra metodología formal, en la semántica de situaciones (Barwise y Perry, 1984; y Gawron, 1986) se identifican las preposiciones como una clase semántica propia para cualquiera de los tipos sintácticos: complementos, argumentos subcategorizados o adjuntos. Asimismo se ha utilizado la semántica de Montague (Dowty, 1989) para correlacionar el marcaje de las preposiciones con el género gramatical en la medida en que se asientan en razones de significado.

La semántica formal no podía ser ajena a la expresividad espacial de las preposiciones (Crangle y Suppes, 1989). En este caso, se nos ofrece una teoría para las preposiciones espaciales a partir de una visión geométrica; y se nos vuelve a recordar que sin contexto no es posible una interpretación completa de estas piezas.

Del mismo modo, y junto a la expresión espacial, el tiempo y su compleja concreción lingüística también ha sido objeto de interés en esta perspectiva (Pratt y Francez, 2001; Brée y Pratt-Hartmann; y Pratt-Hartmann, 2005), junto a la determinación de variables contextuales y dentro de dominios de cuantificación.

Sean, por tanto, la predicación, la cuantificación o las expresiones locativas y temporales desde la óptica de la semántica formal, siguen activas dos incertidumbres: con qué relacionamos el significado preposicional y, sobre todo, qué herramientas, protocolos o metodologías usamos en su identificación.

1.7.3 La gramática cognitiva: prototipos y espacio

Precisamente, fue la gramática cognitiva una de las primeras versiones del significado lingüístico que apeló, de manera programática, al contexto como realidad de interpretación de los hechos cognitivos vinculados al lenguaje humano.

Esta concepción del lenguaje considera que es una habilidad indisociable del sistema cognitivo general. La gramática cognitiva estipula que la información del mundo real concierne a la clase de informaciones de cualquier hecho o acto comunicativo y lingüístico (Langacker, 1987 y 2008). Se acepta que las convenciones lingüísticas se regulan dentro de un contexto global, físico, social y pragmático. Desde el punto de vista descriptivo se amparan en la noción de *relación*, y consideran que las preposiciones tienen un significado en la situación que representan del mundo.

Desde los pioneros, precursores de la lingüística cognitiva (Jackendoff, 1983; y Lakoff, 1987) serán la cognición, el lenguaje, el significado, el contexto y la situación los cinco pilares con los que abordar los objetivos de la investigación de los mecanismos gramaticales de las lenguas (Inchaurrede y Vázquez, 2000; Lee, 2001; y Moya, 2007).

Algunas de los instrumentos y de los supuestos teóricos previos del cognitivismo son el binomio fondo-figura, el pensamiento metafórico como operación básica en la creación de los hablantes, el carácter primitivo de la concepción del espacio, los puntos de vista (Chilton, 2005), y el concepto fundamental de prototipo.

Detengámonos en la teoría de los prototipos de la gramática cognitiva (Osherson y Smith, 1981; Geeraerts, 1989, 2006 y 2012). La crisis categorial clásica se ejemplifica con bastante naturalidad en el tratamiento del significado de las preposiciones (Moure, 1994; y López Rúa, 2003). La pertenencia de todas las preposiciones al mismo conjunto con rasgos semejantes –necesarios y suficientes– queda en entredicho y han sido necesarias estipulaciones previas a los hechos empíricos, tal y como hemos visto en los desarrollos de la teoría sintáctica en la sección 1.5.1.

La categorización prototípica, en cambio, establece categorías continuas con fronteras difusas donde una de ellas (la más representativa) sirve para captar el ejemplar más idóneo de una categoría (Kleiber, 1981). Este carácter borroso de sus límites encaja con una de las nociones que ya se han ido aventurando en la necesidad explicativa: la gradualidad entre fases distintas de significado. Más tarde, en las asunciones semánticas de la hipótesis (Capítulo 2, sección 2.4.2, página 80), volveremos sobre algunas de las propuestas cognitivas sobre las categorías. Ahora es momento de recuperar ciertos trabajos que, simultáneamente a otras visiones y metodologías, se ocupan de las preposiciones como fuente empírica de sus predicciones (Taylor, 2004).

La nómina de autores o la relación de referencias es, en esta tradición, densa y abundante; y se enmarca, por mandato conceptual, en la psicolingüística. Elegimos algunas muestras que representan la riqueza descriptiva que, ajenos ya a las prescripciones teórico-estructurales de la sintaxis, ha aportado esta escuela.

Desde sus orígenes, a principios de los 80 del siglo XX, es el estudio de los mecanismos expresivos para la cognición del espacio una de las vetas que articulan las pesquisas cognitivas (Talmy, 1983 y 2010). Los esquemas, la idealización semántica y los patrones de lexicalización son algunos de los recursos heurísticos.

De hecho, aseguran que la psicopercepción del espacio es uno de los fundamentos del razonamiento humano. El tratamiento de la cognición espacial a través de los conceptos que proponen para sistemas preposicionales deberá abordar el uso hasta la polisemia, las situaciones tipo o los significados idealizados, una versión de los prototipos (Cuyckens, 1984; Herskovits, 1981 y 1986; y Bierwisch, 1988).

Todas las herramientas conceptuales del diseño cognitivo: forma del cuerpo humano, continente-contenido, eje vertical, acceso a la percepción, se ponen al servicio de un análisis pormenorizado, minucioso y rico de preposiciones en inglés, en francés o en alemán (Vandeloise, 1986 y 1991; Coventry y Ludwig, 1991; Benware, 1993; Wunderlich, 1993; y Coventry, Carmichael y Garrod, 1994).

Se usan estrategias para justificar que la información semántico-conceptual de las preposiciones se puede organizar en redes de densidad informativa, relacionando usos y

lecturas (Cuyckens, 1997), y vinculando esos datos semánticos con datos procedentes de corpus (Gries, 1999 y 2006).

Las aproximaciones empíricas desde las preposiciones del inglés son muy frecuentes (Feigenbaum y Kurzon, 2002; y Evans y Tyler, 2003), siendo objeto de atención particularmente *over*, *above*, *under*, *below*, *up* y *down* (Brugman, 1981).

El contexto español no es ajeno al impacto y el desarrollo de la gramática cognitiva en relación a nuestra categoría (Ruiz de Mendoza, 2001). Desde el dativo (Llopis, 1996) hasta el papel de las preposiciones en el área de la ELE (Minguell, 2010; y Llopis, 2015), pasando por el análisis individual de algunas preposiciones como *por* y *para* con valores causales (Funes, 2010) con la descripción radial de la polisemia preposicional, o de *bajo* y *sobre* según las redes semánticas que representan sus significados, son algunas de las investigaciones que se ocupan de la categoría.

También se incorporan esquemas descriptivos en la caracterización de locuciones como *dentro de* (Hernández, 2013), y se usan cuestionarios de hablantes y de textos literarios en la reconstrucción histórica de la preposición *de* (Granvik, 2012).

Por último, la perspectiva visual de estas partículas se ha usado recientemente, de manera sistemática, en el área de la didáctica ELE con herramientas cognitivas (Romo, 2016).

1.8 A modo de síntesis

Hemos comprobado, en nuestra crónica, que las razones descriptivas y explicativas sobre cómo abordar la identidad lingüística de la preposición son de distinta naturaleza y con justificaciones argumentativas, apoyadas en evidencias empíricas y procesos de análisis diversos.

Parece claro que las teorías lingüísticas de preeminencia sintáctica, en sus consideraciones sobre cómo funcionan las preposiciones, han ido aportando matices y sugerencias, y abriéndose a utensilios explicativos que apelan a las tareas semánticas de esta categoría en sus múltiples y variados contextos y estructuras, no solo en español sino en la mayoría de lenguas que hemos referido. Este hecho, además, venía avalado por un aluvión incesante de microobservaciones sobre los aspectos más diversos de las piezas que componen la clase, en una variedad de perspectivas: sincrónicas, diacrónicas, desde la pragmática o desde la cognición, por ejemplo.

Por otra parte, junto a esta búsqueda en el significado para describir y/o explicar cómo se comporta la preposición también hemos identificado una progresiva crisis en los sustentos categoriales, es decir, cómo caractericemos su pertenencia a un conjunto más o menos homogéneo de piezas. Desde la visión categorial aristotélica, que se extendió durante siglos, los gramáticos y los lingüistas han sostenido su filiación a un grupo definido a partir de unos rasgos necesarios y suficientes (podemos consultar el cuadro 1 del Capítulo 1, sección 1.2, página 15). Incluso un poco después de mediados del siglo XX, se introdujeron rasgos exclusivos para seguir con su inclusión entre las categorías gramaticales mayores o principales. Pero pronto iban a empezar los conflictos sobre cómo adscribir ciertas conductas preposicionales según ese nuevo estatus y su semejanza con otras categorías adyacentes o emparentadas (Hilferty, 1993). A partir de ese momento hemos comprobado que se han producido dos hechos en la teoría categorial: o bien se dividía la categoría en subtipos y se justificaba a posteriori los requisitos representacionales de su sintaxis, o se acudía a nuevos rasgos de caracterización con motivación estrictamente semántica.

Hemos comprobado, asimismo, que el surgimiento de la gramática cognitiva (véase la sección 1.7.3 del Capítulo 1) ayudó a ese énfasis semántico en la medida en que esta nueva mirada sobre el lenguaje humano adoptaba procedimientos y principios más holísticos y globales donde las estructuras gramaticales no son autónomas de su significado (Rudzka-Ostyun, 1987). En particular la teoría categorial de los prototipos que determina que la caracterización de las unidades se produce por la emergencia de la pieza más representativa, frente a los más periféricos (Kamp y Partee, 1995). De este modo, muchas líneas de trabajo cognitivas se preguntan por los valores semánticos preposicionales destacando su flexibilidad, vinculándolos, a menudo a su carácter polisémico y relacionándolos también con la frecuencia de uso y aparición (Taylor, 1988).

Y por último, y es en esta nueva perspectiva metodológica que se sitúa nuestra hipótesis, el uso de herramientas de observación y análisis del procesamiento del lenguaje natural en conjuntos de datos masivos, primero en corpus y más tarde en las colecciones conocidas como Big Data, ha permitido que aumente y mejore la manera de localizar e identificar su significado, junto a los recursos estadísticos en la medición de algunos fenómenos asociados con ellas (Litkowski, 2017).

Veamos, en el próximo capítulo, la formulación de nuestra hipótesis para seguir la senda semántica de la investigación, cuál es la naturaleza de esas herramientas y qué asunciones son necesarias para nuestra propuesta.

CAPÍTULO 2. ASUNCIONES DESCRIPTIVAS, TEÓRICAS Y METODOLÓGICAS

2.1 Justificación y formulación de la hipótesis

Hemos comprobado como la descripción de las características lingüísticas de la preposición ha sido, predominantemente, sintáctica o posicional. En efecto, la distribución respecto de otras palabras, el lugar en la cadena lingüística o el sintagma y su quehacer estructural han sido el centro de atención, observación y análisis durante muchos siglos, en muchos autores y escuelas lingüísticas.

Con los desarrollos teóricos de la lingüística y la construcción de teorías semánticas generales (formales, cognitivas o computacionales) su condición categorial ha requerido de aportaciones empíricas más sólidas y distintas para poder explicar cómo significan y cómo accedemos a sus valores semánticos. Hemos visto que las restricciones de las sucesivas teorías categoriales han impulsado una atención distinta a la caracterización de sus significados, desplazando los requisitos posicionales, sintácticos o estructurales. Esa pérdida de protagonismo de las condiciones categoriales ha ido en paralelo a la emergencia de herramientas de análisis empírico con datos masivos. Así se genera una coyuntura nueva en la investigación sobre cómo dirigirse a la naturaleza de sus significados.

Un tratamiento global de las preposiciones exige, en cualquiera de sus aproximaciones metodológicas más recientes, la determinación de cómo medir, evaluar y especificar su significado. Formularemos, en lo que sigue, las condiciones necesarias para nuestra hipótesis, los objetivos y las asunciones necesarias para su consecución.

Esas condiciones necesarias pasan por responder, a lo largo de nuestro trabajo, a tres cuestiones, recogidas entre (59) y (61).

(59) ¿Qué aporta la preposición al significado de su contexto?

(60) ¿Cómo se produce la interacción semántica en el contexto?

(61) ¿Son los métodos estadísticos suficientes para captar la gradualidad semántica?

La hipótesis de nuestra investigación se formula en (62).

(62) Dado un conjunto de valores semánticos descriptivos asociados a las preposiciones y procedentes de la gramática del español, predecimos que la identificación y la distribución de estos valores ocurre gradualmente desde una clase más funcional (con poco o nulo contenido semántico, y función meramente gramatical) a una clase léxica (con un peso semántico relevante e interpretable en el contexto).

A modo de subhipótesis, diremos en (63).

(63) En esta progresión gradual, se propugnan clases intermedias que se llaman semifuncionales y semiléxicas según la estructura y/o la construcción en la que participan, susceptibles de identificarse gracias ciertas herramientas estadísticas.

El significado preposicional oscila entre aquellos valores y usos más funcionales o vacíos como la preposición *a* o la *de*, en las perífrasis verbales o los complementos directos e indirectos descritos en la gramática tradicional, pasando por el *por* del complemento agente en las oraciones en voz pasiva, y llegando hasta los usos locativos de *a*, *ante*, *bajo*, *de*, *en*, *por*, *hacia* o *hasta*, o los temporales de *a*, *entre*, *bajo*, *desde*, *durante* o *tras*, o los modales de *por* y *para*; estos tres últimos considerados léxicos o llenos.

Tomamos como entorno metodológico la lingüística computacional (Confortiová, 1987; Church y Mercer, 1993; y Bolsakov y Gelbukh, 2004). Al abrigo de ese marco usaremos herramientas clasificatorias del aprendizaje automático (primer experimento) y de métodos estocásticos de la teoría de la información, como la medida de entropía, (segundo, tercer y cuarto experimento).

A continuación, detallamos los objetivos y las asunciones teóricas y metodológicas que subyacen y se desarrollan en los tres experimentos.

2.2 Objetivos de la investigación

La relación y la presentación de los objetivos pretenden guiar la comprensión y la lectura del proceso que hemos seguido en nuestro trabajo.

Primero. Estudiar, presentar y describir los periodos historiográficos, los hechos descriptivos y teóricos sobre la categoría preposicional a lo largo de la tradición gramatical y de los modelos de investigación lingüística hasta la actualidad, caracterizando las limitaciones, las dificultades, los problemas, las bondades y los debates que han suscitado las propuestas generales y los análisis de casos concretos para la preposición en español, tal y como hemos realizado en el Capítulo, páginas 1-61.

Segundo. Proponer, formular y verificar la hipótesis de la gradualidad semántica para las preposiciones del español.

Tercero. Contrastar empíricamente la hipótesis mediante tres experimentos (3.1, 3.2 y 3.3) a partir del uso de herramientas del campo del aprendizaje automático y del concepto de entropía, procedente de la teoría de la información (Shannon, 1948), aplicado a la distribución semántica de conjuntos de sintagmas preposicionales. Los datos proceden, como se indicará, de distintos corpus lingüísticos.

Cuarto. Someter la hipótesis de la gradualidad semántica a contraste a partir de la experimentación con hablantes sobre el uso preposicional con un test (experimento 3.4).

Quinto. Establecer conclusiones y posibles perspectivas sobre la hipótesis de la gradualidad semántica a partir del análisis de los resultados obtenidos en los tres experimentos.

El desarrollo de cada uno de estos cinco objetivos se distribuye en cuatro capítulos diferentes.

El primer capítulo ha versado sobre el estado de la cuestión y la revisión de la historia, la descripción y la explicación vinculada a los rasgos lingüísticos sobre las preposiciones y sus problemas de análisis sobre su identificación y su distribución.

En segundo lugar, la que constituye el núcleo de la propuesta, es el planteamiento de la hipótesis sobre cómo se produce, se expresa y se identifica el significado preposicional junto con las asunciones teóricas (semánticas) y metodológicas (computacionales), así como los instrumentos de observación y análisis de los datos.

En el tercer capítulo, el más extenso de la investigación, es la presentación del análisis de los datos empíricos y las pruebas que avalan nuestra investigación. Se compone de un núcleo de tres experimentos (objetivo tercero), un análisis cualitativo de datos procedentes de los datos del uso preposicional en relación con la hipótesis de la gradualidad (objetivo cuarto, y cuarto experimento).

Por último, estableceremos las conclusiones a las que llegamos en relación con la solidez empírica y argumental de la hipótesis.

2.3 Descripción y ejemplificación del objeto de análisis

Todavía hoy, tras algunos siglos, metodologías y perspectivas de análisis de la categoría, siguen activos tres aspectos de la gramática descriptiva del español sobre la preposición. Esas cuestiones enmarcarán nuestro contorno descriptivo respecto de las piezas lingüísticas que son objeto de nuestra hipótesis.

El primero es que se considere una categoría gramatical cerrada, como otros dos grupos que también podrían ser enumerados, como los adverbios y las conjunciones. Este hecho, a efectos de la propuesta, no representa ningún obstáculo ni inconveniente en la medida en que no necesitamos ni prescripciones ni requisitos categoriales para validar su eficacia y su consistencia.

La segunda cuestión es la insistencia en las diversas gramáticas de que no existe acuerdo sobre cuáles sean sus miembros, o la extensión de la categoría. Efectivamente, tal y como hemos visto en nuestra sección sobre la historia de la categoría (secciones 1.2, 1.3 y 1.4), esta inestabilidad ha sido la norma a lo largo muchos siglos. Ni tan siquiera el seguimiento actual de los usos garantiza una cierta seguridad en las fronteras de sus componentes. En la siguiente sección detallaré cuál es nuestra propuesta y los criterios descriptivos que seguiremos.

Y por último, que algunas preposiciones posean solo algunas de las propiedades que caracterizan dicha clase de palabras, y no todas. Esta descomposición de las palabras-piezas-ítems en rasgos o propiedades ha estado en la base de los últimos desarrollos teóricos, no solo de la metodología de la Gramática Generativa (Brucart, 1993) sino también de la Gramática Cognitiva.

La propuesta de un conjunto central de preposiciones así como sus valores léxico-semánticos toma en cuenta la representatividad de usos y piezas desde una visión cuantitativa, dejando al margen los elementos más residuales, anecdóticos o excepcionales, sujetos a cualquier clase de variación, diacrónica, diatópica o diafásica, para el español. (Gutiérrez Araus, 1991).

2.3.1 El conjunto de piezas preposicionales en español

Existen, según las últimas clasificaciones gramaticales descriptivas, entre 19 y 30 piezas individuales consideradas preposiciones, o con usos y comportamientos preposicionales en español. La relación estándar procede de la *Nueva Gramática de Lengua Española* (2009). La lista propuesta por la Academia se recoge en (64).

(64) a, ante, bajo, cabe, con, contra, de, desde, durante, en, entre, hacia, hasta, mediante, para, por, según, sin, so, sobre, tras, versus y vía.

De las 23 unidades académicas se ha propuesto soslayar alguna, como *cabe*, *so*, *versus*, *vía*; o añadir otras como *excepto*, *salvo*, *menos*, *pro*, *como*, *cuando* y *donde*.

Veamos los motivos descriptivos de ambas opciones, excluir unas e incluir otras, y cuál será nuestra opción definitiva.

La preposición *cabe* significa ‘cerca de, junto a’, y procede de *cabo* en el sentido de ‘orilla, extremo de algo’, como se interpreta en (65).

(65) Se tumbó *cabe* la puerta principal del Castillo

No tiene ningún uso en la actualidad, salvo de manera esporádica y residual en algunos poemas u otros textos literarios. El DRAE reconoce su situación de desempleo.

Le sucede algo parecido a *so*. Es equivalente a *bajo*, y hoy solo se utiliza en expresiones fijas como *so capa de, so pena de...*

La partícula *versus* se incorpora recientemente al español por influencia del inglés con el significado de *contra*. La palabra *versus* es un participio latino que significa ‘orientado hacia’, ‘encarado a’ o ‘en dirección a’. Indica un contraste o enfrentamiento entre dos personas, entidades, conceptos u objetos, como reflejan los casos de (66).

(66)

Protección de individuos *versus* ideas
Actitudes hostiles *versus* actitudes amables
Partido Francia *versus* Noruega

Esta palabra aparece en el DRAE (2014) como voz española aunque en las versiones anteriores no estaba registrada. En el *Diccionario Panhispánico de Dudas* (2005) se mostraba que su uso no era censurable, puesto que palabras españolas como *adversario*, procedentes en latín de la misma raíz que *versus*, presentan el rasgo semántico de confrontación. Recomendaba, sin embargo, sustituir ese latinismo anglicado por la preposición española *contra* o por la locución preposicional *frente a*. De nuevo la *NGLE* (2009), en el epígrafe 29.2, también reconoce el uso de *versus*, pero señala que se prefiera *contra* y *frente a*.

Del mismo modo, la Fundación Fundéu, patrocinada por la Agencia EFE y el BBVA, y asesorada por la RAE, en su *Manual del español urgente* (2005) aconseja emplear fórmulas más propias del español, como en (67).

(67)

La protección de testigos *frente* a la prensa

Actitudes hostiles *frente* a espacios cerrados

Tratando sobre el deporte, otra vez la Fundación Fundéu propone separar los adversarios con un guion, mientras que en gran parte de Hispanoamérica está asentado el uso de *versus*, y reconoce otra posibilidad como en *partido entre... y...*, así se aprecia en (68).

El partido Madrid-Barça fue el más visto desde 2014, con 320 millones de espectadores

Holanda *versus* España en el Sánchez Pizjuán

El partido entre el Racing y el Osasuna ha sido considerado de alto riesgo

La partícula *vía*, procedente de un sustantivo, indica el lugar por el que se pasa o transita o el medio, como en (69), y tiene unos contextos muy determinados casi siempre.

(69) Fue a Varsovia *vía* Berlín

Será emitido *vía* satélite

En último lugar, los usos y valores preposicionales de *excepto*, *salvo*, *menos*, *pro*, *como*, *cuando* y *donde*, recogidos en los ejemplos de (70 a-g), no serán tenidos en cuenta por su excepcionalidad y el carácter restringido tanto de sus entornos estructurales como de sus significados.

(70) (a) Le gusta toda la pasta *excepto* la fresca

(b) Iremos todos *salvo* el vecino

(c) Acudirán las elegidas *menos* la experta

(d) Acto *pro* defensa de los derechos civiles

(e) Usaron la escoba *como* pata

(f) No había que comer *cuando* la guerra

(g) Ve *donde* la señora Patricia

La restricción distribucional y la fijación interpretativa de estas preposiciones menos frecuentes no propicia ni favorece su elección para nuestros propósitos. Dada la fijación de los contextos de estas preposiciones y su baja frecuencia, optamos por no incluirlas en el estudio. En ese sentido, ya se ha investigado la historia de las etimologías de estas preposiciones (Portilla, 2011). De manera que en nuestra investigación sigue la propuesta mayoritaria de la *NGLE* (2009) con excepción hecha de *so* y *cabe*, *versus* y *vía*, tal y como queda recogido en (71), lista definitiva de uso para nuestra experimentación.

(71) A, ante, bajo, con, contra, de, desde, durante, en, entre, hacia, hasta, mediante, para, por, según, sin, sobre y tras.

Todos los ejemplos de sintagmas preposicionales que hemos usado en los tres experimentos y el test que utilizaremos para contrastar la hipótesis proceden exclusivamente de piezas de la serie bajo (71). La participación de estas piezas en las estructuras locucionales y su funcionamiento queda excluida de nuestra investigación.

2.3.2 Los valores léxico-semánticos de las preposiciones en español. Relación y justificación de clases preposicionales.

La descripción léxico-semántica que utilizamos para presentar los valores de las preposiciones toma aquellos términos que son más aceptados en la literatura gramatical. La propia naturaleza del significado preposicional, de carácter relacional (Morera, 1994), conduce el sentido de las denominaciones.

El consenso terminológico de los manuales, normativos, descriptivo-gramaticales o descriptivo-semánticos sobre las preposiciones suele usar descriptores como *locativo*, *temporal* o *modal*. A su vez, esas etiquetas suelen incluir otras como *dirección*, *límite*, *manera* o *situación*, en el caso de la localización; o *causa*, *finalidad*, en el caso de la modalidad, por citar dos ejemplos. La relación de referencias que constituyen nuestro marco descriptivo se encuentra en (72).

(72)

López (1970), Trujillo (1971), Luque (1973), Morera (1984), García Yebra (1988), Osuna (1991), Nández (1995), Cifuentes (1996b), Fernández (1999), Moreno y Tuts (1998), Bosque y Demonte (1999), Martínez (2002), Slager (2004), Laguna (2004), *NGLE* (2009), y Moliner (2012).

Nuestra descripción recoge verbos como *introduce*, *expresa*, *indica*, *localiza*, *señala* o *sitúa* en alusión a cómo se produce el significado de la preposición en los sintagmas o estructuras donde participa. Tal y como acabamos de indicar, el significado preposicional no es referencial, de modo que su contribución a la interpretación global del sintagma en el que aparece está vinculada a esas maneras de participación. Respecto a las cuestiones formales o gramaticales utilizo denominaciones descriptivas de la sintaxis tradicional como complemento directo, indirecto, preposicional o circunstancial; así como otros conceptos como sujeto paciente.

De este modo, en el cuadro 4, se reúne la colección de los valores léxico-semánticos que constituyen el universo semántico que intentamos medir para establecer la validez de la hipótesis. A partir del elenco descriptivo estandarizado procedente de las referencias de (72) y sintetizado en el cuadro 4, y con las agrupaciones de las propuestas teóricas sintácticas y categoriales que hemos revisado en el estado del arte (Capítulo 1, páginas 58-60), propugnamos la organización de los valores léxico-semánticos de las preposiciones del español en cuatro clases, y cuya denominación es funcional, semifuncional, semiléxica y léxica.

CUADRO SOBRE LAS ASUNCIONES DESCRIPTIVAS DE LOS VALORES LÉXICO-SEMÁNTICOS		
Clases preposicionales propuestas	Descripción gramatical de los valores léxico-semánticos de las preposiciones	Algunos ejemplos
Funcional (A, para y hacia y por)	Introduce el objeto, el destinatario y el beneficiario	Enviamos flores a María

	<p>de la acción o proceso verbal (los llamados CD – complemento directo u objeto-, CI -complemento indirecto-).</p> <p>Introduce el sujeto paciente de la construcción oracional pasiva.</p> <p>Introduce el complemento agente, destinatario o del proceso.</p> <p>Forma parte de las perífrasis verbales o de las locuciones prepositivas, adverbiales o conjuntivas.</p>	<p>Han redactado una felicitación para él Siente un gran respeto hacia el alcalde</p> <p>Hemos sido recibidos por el Presidente El puente fue construido por el equipo de ingenieros</p> <p>La edición de Cervantes La comida de mañana</p> <p>Echaron a correr Lo hicieron a lo loco</p>
<p>Semifuncional (A, ante, con, contra, de, en, por, entre, hacia, sobre...)</p>	<p>Introduce el Complemento de Régimen Verbal o Preposicional.</p>	<p>Tratan sobre la política agrícola Abogan por el libre comercio Abusan de la confianza Contaban con su fidelidad Han deliberado sobre la cuestión jurídica Pienso en los asuntos comunes</p>
<p>Semiléxica (A, con, según, en, por)</p>	<p>Expresa el medio, el instrumento, el modo o la costumbre con la que se ejecuta o hace algo.</p>	<p>Lo hemos hecho a máquina Lo construyeron con una grúa hidráulica Se encuentran con dificultad Abrió con su llave Lo hicieron según la moda del momento Habla en serio</p>

(a, de, en, para, por, según)	Expresa la causa, la consecuencia o la finalidad de una acción.	Murió de un ataque al corazón Supimos que algo iba mal en el tono de voz Ahorran para la casa Se levantó por no llegar tarde
(a, de, en, por)	Expresa el precio o la materia de cualquier producto.	Los hemos vendido a un euro Un pastel de manzana Lo alquilaban por 600 euros al mes Un adorno de oro
(a, por)	Expresa la distribución, la proporción, el reparto o la sucesión de las cosas.	Tocamos a cinco por cabeza Tres juguetes por niña
(a, de, hasta, por, sobre)	Expresa la velocidad, la cantidad límite, el límite de una acción (verbos de desplazamiento) o las medidas de los objetos, procesos o entidades.	Conduce a más de cien kilómetros por hora Descanso sobre las seis Llegan hasta la locura
(de)	Expresa una relación de propiedad o posesión.	Ese reloj es de mi hermano La mitad de la naranja Esa camisa de Pablo
(de, con, sin y contra)	Indica el encuentro, desencuentro, oposición, compañía o ausencia, o el contenido de los objetos, procesos o entidades.	Un vaso de agua Estamos con el profesor agredido Firmamos contra la nueva ley Agua sin gas
(de, en y sobre)	Introduce el tema, el ámbito, el asunto o la materia de qué trata alguna entidad o actividad.	El libro de física Una tesis sobre preposiciones
(de, con, entre y por)	Introduce y presenta los rasgos, las partes, el estado medio, los elementos	Me apetece un poco de agua Los colores de la camisa Las mangas del chaleco

(entre)	integrantes de algo o de una serie.	Una porción de pastel
(para y según)	Expresa la asociación, la reciprocidad, la inclusión, la participación y la relación de varias personas o cosas.	Me honra entre sus amigos Lo compramos entre los dos
(hacia)	Expresa el punto de vista o la opinión de las personas.	Según Lluís todo irá bien Para Irene el asunto es más sencillo
	Indican el receptor humano o abstracto de un movimiento real o figurado (con verbos de movimiento).	Se desplazan hacia el público Se dirige hacia la democracia
Léxica		
(A, ante, bajo, entre, hacia, por, sobre, tras...)	Localiza en el espacio objetos, personas o conceptos respecto de un punto (sin desplazamiento, como una mera orientación), en sentido literal o figurado, e indica el punto, la zona o el lugar concreto en el que sucede o se localiza algo.	Está bajo la mesa Bajo la dirección de Francisco todo resultó más dinámico Se situaron entre los dos Déjalo sobre el mármol Próximo a la carretera Corre hacia el norte
De, desde, a, entre	Señala el origen, la procedencia, el punto de partida o el inicio de un movimiento o una observación.	De aquí a Barcelona hay diez kilómetros Desde tu casa iremos en coche Lejos de aquí
(a, contra, entre, hacia, hasta, para, por, sobre)	Señala y expresa la dirección, la orientación, el destino, el término, el lugar intermedio, el curso o los límites respecto de un punto o de una acción.	Viajaremos a Roma La casa está contra el sur Se dirigen hacia Mataró Acércate hasta el bar y le avisas El mercancías partirá para la frontera hoy Todos los días camina por la ribera norte

		Marcharon sobre la capital
(a, de)	Indica la distancia exacta respecto de un punto o la localización en una escala.	Está a diez manzanas
(en)	Indica el medio de transporte en el que se llevan a cabo los movimientos.	Nos moveremos en avión
(a, de, entre, hacia)	Introduce, indica y sitúa la hora concreta, el momento, la situación o un suceso en distintas fases del desarrollo temporal o de un acontecimiento de un modo concreto o aproximado.	A las cinco nos acercaremos a tu casa Del 1 al 10 se realizan las matrículas Estarán entre las dos y las tres de la tarde Asientan hacia siglos
(a, de y durante)	Expresa la distancia en el tiempo, el tiempo invertido en la realización de una actividad y su duración, o que sucederá en el futuro.	Lo trajeron en tres días Vivimos a dos horas de aquí
(a, de, hasta)	Expresa el punto, el término, la dirección o el límite temporal de una acción, en combinación o no con <i>de</i> , u otras preposiciones.	Hemos ido de Lugo hasta Santiago en tres horas Fuimos a Toledo
(a, bajo, desde, en, hacia, para, por, sobre y tras)	Expresa, señala o localiza la aproximación, la frecuencia, la duración, la posteridad y la periodicidad (en meses, años, estaciones) con que se lleva a cabo una acción o acontecimientos, o cuando empieza.	En tres meses sucederá todo. Hacia mayo empezará a llover Lo tendremos para el lunes Para mañana es imposible Tras la primavera será todo más fácil
(de)	Expresa la edad de una persona, o la edad que tiene cuando ocurre un acontecimiento.	Son alumnos de quince años

Cuadro 4. Valores léxico-semánticos de las preposiciones en español

Comprobaremos que se producen zonas de tránsito próximas a los dos extremos, entre la funcionalidad completa y la lexicidad completa. La gradualidad se manifiesta intuitivamente con elasticidad o flexibilidad, y dirigida por la naturaleza de los hechos y de las metodologías empleadas (aprendizaje automático, clasificación semántica y uso de hablantes), tal y como queda representada en el cuadro 5.

CUADRO SOBRE LA ORGANIZACIÓN FLEXIBLE DE LAS CLASES PREPOSICIONALES					
Experimentos	Clases preposicionales	Funcional	Semifuncional	Semiléxica	Léxica
Primer experimento (preposiciones <i>a, hacia y hasta</i>)					<i>L1</i> <i>L2</i>
Segundo experimento Estructuras en el ámbito verbal (todas las preposiciones)					
Tercer experimento Estructuras en el ámbito nominal (todas las preposiciones)					
Cuarto experimento (todas las preposiciones)					

Cuadro 5. Organización de las clases preposicionales por experimentos.

En cada uno de los experimentos explicaremos las razones empíricas que ofrecen los datos para identificar unas clases u otras. Las que están sombreadas son las opciones elegidas para cada experimento, y las que están en blanco están ausentes.

Los estudios empíricos más recientes perseveran y confirman con nuevos datos y evidencias esa dualidad progresiva, fluctuación y transición. De hecho, la génesis de la noción de funcionalidad procede de la teoría sintáctica del último tercio del siglo XX (Abney, 1986; Baker, 1988; Grimshaw, 1990; Ouhalla, 1991; Hale & Keyser, 2002; y Muysken, 2008). Esta profusa tradición explicativa buscaba dar cuenta de fenómenos sintagmáticos en las relaciones entre las categorías léxicas mayores, nombres, verbos y

adjetivos, y categorías como el tiempo, la concordancia o el grado (las funcionales). La estrategia general fue la flexibilización de las proyecciones sintácticas y la estipulación de nuevos rasgos delimitativos de la funcionalidad. En cierto sentido, se estaba gestando una crisis sobre cómo debía expresarse el componente categorial de las gramáticas en las relaciones entre los ítems léxicos y los funcionales.

No en vano, este debate venía de antiguo y una tradición anterior, la estructuralista, que ya expresaba dudas e inconvenientes en el abordaje de la cuestión categorial (Coseriu, 1972), reconociéndole ventajas a otras orientaciones.

La colección de investigaciones con renovadas aportaciones empíricas en la década de los noventa del siglo XX, de carácter interlingüístico y comparativo, ha permitido comprobar que la caracterización negativa [-N, -V] de la preposición era insuficiente. Esa dificultad emerge frente a hechos semánticos no solo en otras lenguas donde las relaciones se expresaban con posposiciones o caso (como serían lenguas como el ruso y el polaco – con caso y preposiciones-, el japonés –con preposiciones y posposiciones-, o el igbo de la familia nigercongolesa –con una sola preposición-), sino en las lenguas más frecuentes en la exploración teórica con diversidad preposicional: las anglogermánicas y las románicas. En ese nuevo retrato investigador sobre cómo se caractericen las categorías lingüísticas, condiciones, requisitos y restricciones, se empiezan a encontrar evidencias y analogías entre las nuevas categorías funcionales y la conducta sintáctica y la contribución semántica de las preposiciones. En algún caso (Baker, 2003) se determinó que la preposición es parte de un sistema diferente puesto que tiene más en común con determinantes, pronombres o complementizadores.

Durante tres décadas conviven, en la perspectiva categorial, tres opciones más o menos estables: la que considera la preposición una categoría mayor o *léxica* (Jackendoff, 1973; Chomsky, 1981; Rizzi, 1988; Van Riemsdijk, 1990; Wunderlich, 1996; Zwarts, 1997; Koopman, 2000; Den Dikken, 2003; y Svenonius, 2004 y 2006); la que considera que es *funcional* (Emonds, 1985; Grimshaw, 1992; Baker, 2003; y Ashbury, 2005); y una opción híbrida que divide la categoría y optan por el término de *semiléxico* (Zwarts, 1992 y 1997; Tseng, 1999; Corver y Van Riemsdijk, 2001; Gaatone, 2007; Mardale, 2007 y 2009; y Genná, 2010).

Como se recoge en (73), se ha justificado empíricamente un espacio explicativo en la línea continua entre las categorías léxicas y las funcionales donde se situaría una parte considerable de lo que hace una preposición (Mardale, 2011).

(73)

Léxicas	Semiléxicas	Funcionales
< ----- >	< ----- >
Nombres	Preposiciones	Determinantes
Verbos		Concordancia
Adjetivos		Grado
Tiempo		
Grado		

De hecho, el cuestionamiento de la visión homogénea e unitarista de cómo definir una categoría lingüística, se estaba produciendo en la gramática cognitiva. Hemos visto en el estado del arte (Capítulo I, sección 1.7.3) cómo su teoría de los prototipos y su desarrollo de las clases lingüísticas conduce a una visión no discreta de las categorías donde la identificación se ejerce a partir del predominio de los prototipos (Ibarretxe-Antuñano, 2016). La historia de esta visión, sus bases psicolingüísticas, junto a las justificaciones epistemológicas de sus argumentos han sido tratadas con precisión y detalla en las últimas décadas (Ungerer y Schmidt, 1996; Stadler y Eyrich, 1999; Talmy, 2000; Croft y Cruse, 2004; y Langacker, 2008).

Además, y casi simultáneamente, surgen propuestas en el seno de las metodologías de la lingüística computacional y el PLN (procesamiento del lenguaje natural) que priorizan el significado preposicional entendido como un sistema de relaciones que puede ser descrito y organizado, semántica y ontológicamente, en un sentido práctico y con posibles aplicaciones (Litkowski, 2005, 2007 y 2017; Srikumar y Roth, 2013a y 2013b; Hartrumpf, Helbig y Osswald, 2006; y Lassen, 2006). Esta perspectiva semántica se ofrece gracias a distintas novedades intrínsecas a esas metodologías, a saber, la disposición y el manejo de cantidades exponenciales de datos, la facilidad en la observación de esos datos y la aplicabilidad en aplicaciones tecnológicas como la traducción automática. La semántica preposicional es cardinal para la comprensión de otros fenómenos de las tareas

computacionales en la medida en que buena parte del uso de esta clase de palabras está vinculado, precisamente, a su acción matizadora, correctora, precisadora o diferenciadora (Baldwin, Kordoni y Villavicencio, 2009; Jorgensen y Lonning, 2008; Litkowski y Hargraves, 2006; y Kordoni, 2006).

2.4 Marco teórico y metodológico

Presentamos aquí los principios teóricos así como los requisitos metodológicos de nuestra investigación, así como los antecedentes y su justificación en relación a nuestro planteamiento y los datos empíricos estudiados.

2.4.1 El marco computacional

La hipótesis presentada en el apartado 2.1 se inserta dentro del área de la lingüística computacional y de la lingüística del corpus, tanto en el origen de los datos, como en sus heurísticas procedimentales y el uso de las herramientas metodológicas para observar, tratar y analizar tales datos (Fernández Pérez, 1996; Thomas & Short, 1996; Habert, Nazarenko y Salem, 1997; Sampson & McCarthy, 2004; Rodrigues, 2009; Tordera, 2011; y O’Keffe y McCarthy, 2019).

Es cierto que en la descripción de las restricciones y/o las limitaciones que han surgido en la tradición lingüística sobre las preposiciones, algunas referencias que procedían de la denominada lingüística teórica preconizan el método hipotético-deductivo. Sin embargo, es dentro de la perspectiva computacional, y con herramientas probabilísticas y estadísticas (Pardo, 1993; y Hastie, Tibshirani y Friedman, 2009), donde se produce, desde nuestra perspectiva, una explicación más consistente de los fenómenos sobre la determinación del significado preposicional. Esta importancia de la dilucidación a partir de los datos y de instrumentos estocásticos ya había sido manifestada desde por algunos lingüistas teóricos (Abney, 1996; Abney y Bird, 2011). En otras ocasiones se indica que

desde los datos no controlados y en el seno de la lingüística cuantitativa se pueden obtener resultados relevantes, precisamente, para algunas cuestiones suscitadas desde la teoría lingüística (Koehler y Rieger, 1993; Bresnan, 2017; y Ouyang y Jiang, 2019).

De este modo asumimos una opción empírico-inductiva (Cheng, 2012) para el proceso de investigación a partir del esquema de (74), respecto de los pasos metodológicos de la lingüística de corpus.

(74)

Observación de hechos lingüísticos



Formulación de hipótesis para explicar estos hechos



Generalización basada en la repetición y evidencia de modelos



Unificación de estas observaciones en enunciados y corolarios teóricos

Sobre la historia, los fundamentos, los objetivos, las metodologías, las herramientas y las aplicaciones de la lingüística computacional disponemos ya de una prolija colección de introducciones, manuales y aproximaciones que detallan, exponen y justifican el marco de nuestra investigación. Veamos, a modo de ilustración, algunas de las referencias generales ordenadas cronológicamente.

En todos ellas el compromiso metodológico pertenece a la opción empírico-inductiva para cualquiera de sus aplicaciones y orientaciones. Así, desde las primeras noticias sobre el área (Bott, 1970), donde se presentan las expectativas de interés de la época: la traducción automática y la recuperación de información (Strzalkowski, 1999; y Vallez y Pedraza, 2007), hasta los últimos manuales (Clark, Fox y Lappin, 2010) que incluyen un capítulo sobre aprendizaje automático –subárea que está creando una gran

expectativa en la actualidad-, pasando por los trabajos en el contexto español (Meya Llopart y Hubert, 1986; Moreno Sandoval, 1998; Martí y Castellón, 2000; Tordera Yllescas, 2012; o Gómez et alia, 2019), o diversos manuales clásicos de los años ochenta (Grishman, 1986; Barton et alia, 1986; Gazdar y Mellish, 1989; y Thomas y Short, 1996), se aprecia la eminencia de los hechos, su cuantificación y la utilización de instrumentos computacionales para la generalización. Entorno a los procedimientos de investigación computacional también se han presentado monografías e investigaciones muy solventes (Barnbrook, 1996; Jurafsky et alia, 2000; Mitkov, 2005; Villayandrade, 2010; Pustejovsky y Stubbs, 2012; y Hausser, 2014), o síntesis que incluyen modelos semánticos que sean implementables computacionalmente (Stanford, 2014), e incluso alguna panorámica en el contexto español (Caravedo, 1999).

2.4.2 Asunciones semánticas

La naturaleza del significado de clases de palabras como las preposiciones no suele estudiarse en la semántica léxica - la encargada de observar y tratar las unidades o piezas lingüísticas individuales - sino en la composicional - la que se ocupa de las expresiones complejas-. Esa distinción atraviesa buena parte de los fundamentos básicos de la semántica que abunda en esta idea (Escandell, 2004; Espinal, 2014; y Saeed, 2016).

Este hecho, no meramente disciplinario, guarda relación con el tipo especial de valor semántico de las partículas preposicionales. Nuestras asunciones semánticas se formulan en (75 a-b).

(75)

(a) El significado de las preposiciones es relacional.

(b) La interpretación del significado preposicional se produce de manera contextual y distribucional.

Entendemos por significado relacional al que se obtiene como consecuencia de establecer un vínculo entre los dos términos que están conectados por la preposición (en la posición de núcleo X, y en la posición de complemento Z), como se identifica en (76).

(76) X - P - Z

Denominamos una tripleta al conjunto de tres elementos cualesquiera donde, y según la variación estructural determinada por las reglas gramaticales del español, X y Z pueden estar encarnados por diferentes categorías gramaticales, y P es cualquiera de las piezas preposicionales del repertorio que usamos como elementos en nuestra hipótesis. No tenemos en cuenta la interacción sintáctica entre los determinantes y los términos de Z.

De manera que, y como se ve en (77 a-b), lo que podamos asegurar sobre el significado de las preposiciones *a*, *en* o *de* está conectado con la relación que se produce entre sus núcleos y sus complementos. A esa relación se le ha atribuido un carácter predicativo (Escandell, 2004).

(77)

(a) Echaron *a* Miguel el otro día.

(b) Se sentaron *en* el banco *de* piedra

En la segunda asunción semántica, recogida en (75b), subyace el principio de contextualidad (Frege, 1892). Según este, las unidades léxicas (y los modos sintácticos de construcción de expresiones complejas a partir de unidades léxicas) sólo tienen significado en el contexto de oraciones.

En esa misma asunción, indicamos el carácter distribucional del significado. Este principio se basa en los desarrollos y las adaptaciones contemporáneas del lingüista estadounidense Zellig Harris (1954). En la última década se ha producido una recuperación de la hipótesis distribucional en los modelos semánticos del procesamiento del lenguaje natural, originándose lo que se conoce como neodistribucionalismo (Sahlgren, 2006) que fue objeto de un monográfico de la publicación *Computational Linguistics* (2009).

En ese número especial se detallan aspectos como que las representaciones distribucionales son vectores, u objetos algebraicos como matrices o tensores junto a aplicaciones diversas (Baldwin, Koldoni y Villavicencio, 2009; y O’Hara y Wiebe, 2009). Los modelos semánticos neodistribucionales y su funcionamiento han sido revisados de manera panorámica recientemente (Martí, 2018).

Una caracterización distribucional, de manera genérica, supone que el significado de una palabra está determinado por el contexto. O de un modo más específico, está en relación con su distribución, es decir, en las palabras que lo rodean (Nazar y Renau, 2012).

El siguiente párrafo (Sahlgren, 2008), transcrito en (78), apunta la correlación entre distribución y significado a través de la noción de similitud.

(78) Distributional approaches to meaning acquisition utilize distributional properties of linguistic entities as the building blocks of semantics. In doing so, they rely fundamentally on a set of assumptions about the nature of language and meaning referred to as “the distributional hypothesis”. The main point of this hypothesis is that there is a correlation between distributional similarity and meaning similarity, which allows us to utilize the former in order to estimate the latter (Sahlgren, 2008, pág. 33).

También se ha caracterizado (Lenci, 2008) en términos de función de los contextos, como se lee en (79).

(79) The degree of semantic similarity between two linguistic expressions A and B is a function of the similarity of the linguistic contexts in which A and B can appear (Lenci, 2008, pág. 3).

Cómo se articula, se mide y se determina la naturaleza de esa similitud se puede encontrar en las propuestas y los experimentos particulares de las investigaciones de la última década, como la distinción entre similitud atribucional o relacional (Turney, 2006), el uso de espacios vectoriales (Turney y Pantel, 2010), o los primitivos semánticos (Ruiz de Mendoza y Pérez, 2000) como unidades mínimas en el análisis computacional (Boleda y Erk, 2015). Constantemente, se ha insistido en la primacía del significado; que ha sido

protagonista destacado en el procesamiento del lenguaje natural (Levy y Goldberg, 2014; y Levy, Goldberg y Dagan, 2015).

En nuestra propuesta optamos por una agrupación basada en seis categorías semánticas. A partir de las palabras que seleccionan las preposiciones en sus contextos de complementación proponemos los siguientes tipos, recogidas en (80).

(80) *Animado, Inanimado, Entidad abstracta, Evento, Temporal y Locativo.*

Cada una de esas categorías semánticas se caracteriza con rasgos que hemos extraído de la Top Concept Ontology de EuroWordnet (Vossen et alia, 1998; y Álvez et alia, 2015).

La TCO organiza los conceptos en tres tipos de entidades, de primero, segundo y tercer orden, apoyándose en una distinción previa (Lyons, 1977).

Las entidades de primer orden son perceptibles por los sentidos y localizadas en el tiempo y en un espacio tridimensional. A su vez, las de segundo orden son situaciones (propiedades o relaciones), estáticas o dinámicas que no pueden ser comprendidas, oídas, vistas o sentidas como una entidad física independiente, básicamente estados, situaciones y sucesos. Pueden estar localizadas en el tiempo y ocurrir o tener lugar, en lugar de existir. Y, por último, las entidades de tercer orden que son proposiciones inobservables las cuales existen al margen del tiempo y del espacio. A partir de aquí establece una jerarquía taxonómica a partir de 63 niveles de etiquetado, tal y como se recoge en el cuadro 6.

Top Concept Ontology	
Entidades de primer orden	Entidades de segundo orden
Origen	Situación tipo
Natural	Dinámica
Vivo	Suceso con límite
Planta	Suceso sin límite

<p>Humano</p> <p>Criatura</p> <p>Artefacto</p> <p>Función</p> <p>Construcción</p> <p>Vehículo</p> <p>Representación</p> <p> Representación económica</p> <p> Representación lingüística</p> <p> Representación visual</p> <p>Software</p> <p>Lugar</p> <p>Ocupación</p> <p>Instrumento</p> <p>Vestuario</p> <p>Mobiliario</p> <p>Cobertura</p> <p>Contenedor</p> <p>Comestible</p> <p>Edificable</p> <p>Objeto</p> <p>Substancia</p> <p> Sólido</p> <p> Líquido</p> <p> Gas</p> <p>Composición</p> <p> Parte – Grupo</p>	<p>Componente de la situación</p> <p>Causa</p> <p> Agentiva</p> <p> Fenomenica</p> <p> Estimulación</p> <p>Comunicación</p> <p>Condición</p> <p>Existencia</p> <p>Experiencia</p> <p>Localización</p> <p>Manera</p> <p>Mental</p> <p>Modal</p> <p>Física</p> <p>Posesión</p> <p>Cantidad</p> <p>Social</p> <p>Tiempo</p> <p>Uso</p>
<p>Entidades de tercer orden</p>	

Cuadro 6. Disposición de rasgos y subdivisiones del TCO (Top Concept Ontology)

La disposición arbórea significa que los rasgos que se extienden hacia la derecha están incluidos o pertenecen a la etiqueta de la izquierda, es decir, que tienen una cierta estructura de inclusión. En nuestra presentación lo anotamos con el signo >. Hemos elegido

algunos rasgos de la TCO para definir nuestras seis categorías (Batet, 2010). Esa elección se justifica por el solapamiento entre los rasgos y los objetos que categorizan, por ejemplo, la función *Construcción* implica *Artefacto*. El objetivo es construir categorías más generales con la ayuda de esos rasgos y de acuerdo con la realidad empírica que hemos encontrado en los datos de nuestros experimentos.

De este modo, los rasgos definitorios que describen nuestra categoría *Animado* son los siguientes de las entidades de primer orden: *Origen > Natural > Animal, Humano y Criatura*. Cualquiera de estas etiquetas se asigna a nombres como *primo, reportero, observador, leopardo, unicornio o perro*.

En el caso de *Inanimado* nos encontramos con *Origen > Artefacto > partes de artefacto*; y *Origen > natural > planta* entre las de primer orden; y *Cantidad* entre las de segundo orden con casos como *aplique, uniforme o rueda*.

Respecto a *Locativo* tenemos un rasgo en las de primer orden, *Función > Lugar*; y *Localización* entre las de segundo orden. Algunas muestras son *ermita, Madrid, afueras, aparcamiento, camarín, iglesia, plaza y calle*.

La categoría de *Evento* pertenece a las entidades de segundo orden y se puede caracterizar con la mayoría de rasgos propuestos que se reconocen en la tabla para describir las situaciones de esos sucesos (comunicación, condición, existencia, experiencia, manera, mental, modal, física, posesión, cantidad, social, tiempo o uso). Encontramos ejemplos con *actuar, fiesta, boda, certamen, reír, charlar o conversar*.

El *Tiempo* se incluye, también, dentro las entidades de segundo orden y sus representantes pueden ser *momento, episodio, mañana, semana, marte o periodo*.

Por último, la *Entidad abstracta*, en su caracterización inicial se nos dice que se refieren a *proposiciones inobservables las cuales existen independientemente del tiempo y del espacio* y se adscriben a las entidades de tercer orden. Igualmente hemos incluido dentro de esta categoría los nombres que expresan *propiedades y relaciones*, rasgos que se encuentran entre las de segundo orden porque su condición inobservable y su existencia no vinculada a concreciones temporales o espaciales resulta evidente en casos como *fundamento, aspecto, caridad, amor, tristeza, rapidez o exagerado*. En determinados

contextos, estas entidades implican transcurso de tiempo y esa propiedad las podría vincular con eventos pero consideremos que es más congruente la clasificación por su carácter abstracto.

En el cuadro 7, y a modo de síntesis, se recogen las agrupaciones del TOC, sus correspondientes categorías lingüísticas y nuestra propuesta organizativa.

Entidades básicas del TOC	Funciones y rasgos conceptuales incluidos del TOC	Categorías lingüísticas	Propuesta de la hipótesis gradual
Primer Orden	Origen, función, composición, parte, grupo, sustancia, objeto, vehículo, software, representación, lugar, ocupación, instrumento, vestuario, mobiliario, cobertura, contenedor, comestible, edificable	Nombres	Animado, Inanimado, Locativo
Segundo orden	Situación tipo, dinámica, suceso limitado, suceso no limitado, estático, propiedad, relación, Componente situacional, causa, agentivo, fenoménico, estimulante, comunicación, condición, existencia, localización, manera, mental, modal, físico, posesión, propósito, cantidad, social, tiempo y uso	Nombres, verbos y Adjetivos	Evento y Temporal
Tercer orden		Nombres	Entidad abstracta

Cuadro 7. Síntesis de correspondencias entre las categorías lingüísticas, el TOC y las categorías semánticas propuestas

2.4.3 Herramientas de análisis y procesamiento

Las herramientas de análisis para los datos empíricos proceden de dos subáreas de las ciencias de la computación y la inteligencia artificial: de un lado, el aprendizaje automático; y del otro, algún aspecto de la teoría de la información.

2.4.3.1 El aprendizaje automático

Una de las decisiones metodológicas para la identificación de los grados semánticos de las preposiciones es el uso de estrategias automáticas que permiten reconocer, en la posición de complemento preposicional, Z en la tripleta de (81), alguna clase de parecido o similitud semántica de los nombres de un conjunto de sintagmas preposicionales elegidos al azar.

(81) X – P – Z

Para ello optamos por las herramientas que proceden del área del aprendizaje automático (*machine learning*). Este campo investiga los procesos computacionales que subyacen en tareas humanas. A partir de la modelización de labores propias de la inteligencia humana se experimenta con el uso de instrumentos matemático-estadísticos que miden los resultados de esas computaciones.

Se suelen señalar dos circunstancias indispensables y concurrentes en el despegue definitivo del uso de esos recursos del análisis matemático-formal; a saber, la disposición de grandes volúmenes de datos y la rapidez en el procesamiento de los ordenadores (Mannig y Schütze, 1999). Este aumento de la cantidad de datos y la celeridad en su tratamiento ha llamado la atención sobre fenómenos que se consideraban irrelevantes observados individualmente o de manera aislada. En el caso de las preposiciones, además, se han venido produciendo pesquisas en y con grandes colecciones de datos para tratar uno de los asuntos que más ha incentivado el área: la polisemia preposicional. La tarea de la desambiguación fue reconocida como uno de los mejores caminos para precisar cómo sea

la propiedad polisémica (Charniak, 1993; Suárez, Palomar y Rigau, 2005). Esa convivencia de significados y su flexibilidad, casi connaturales a la categoría, ha sido descrita, además, como una de las dianas cardinales de la investigación computacional (Baldwin, Kordoni y Villavicencio, 2009; y Hovy et alia, 2011). De la misma manera, la frecuencia de aparición de las preposiciones se ha estudiado en relación con los mecanismos taxonómicos de desambiguación en corpus de diversa naturaleza y desde supuestos metodológicos diversos (Hovy, Tratz y Hovy, 2010; y Hernández-Pastor y Perrián-Pascual, 2016).

La historia del área del aprendizaje automático suele presentarse con las aplicaciones más relevantes y una valoración de los resultados (Cetinsoy et alia, 2015; Ulinic, 2018). El concepto clave, el aprendizaje, se ha definido de un modo estandarizado (Rich y Knight, 1994) indicándose que es un proceso por el cual una entidad adquiere conocimiento.

No olvidemos que uno de los atolladeros de los análisis teóricos de la preposición fue la identificación del significado. Desde el aprendizaje automático se sugiere (Abur-Mostafa, Mugdon-Ismail y Lin, 2012) que en ciertas situaciones son los datos los que promueven soluciones de carácter empírico en relación con la semántica, así como el uso de evidencias y correlaciones estadísticas (Wermater, Riloff y Scheler, 1996).

Se han propugnado dos grandes tipos de aprendizaje automático: supervisado y no supervisado. La diferencia entre supervisado o no supervisado depende si el aprendizaje se realiza con ejemplos anotados o no. La comparación del conjunto de outputs constituye la evaluación. Los denominados semisupervisados son un continuum que va de uno a otro con un cierto grado de supervisión (Chapelle, Schölkopf y Zien; 2006; y Abney, 2007).

En nuestro primer experimento hemos optado por una de las técnicas del aprendizaje no supervisado, el denominado *clustering* (Moisl, 2015). Este procedimiento agrupa a partir de una colección de objetos aquellos que son similares en cuanto a unas características predeterminadas que constituyen los rasgos o atributos. Desde esos rasgos el sistema agrupa casos aplicando medidas de similitud entre ellas. A continuación se produce la clasificación. Dado que las características o atributos corresponden a las palabras del contexto el criterio es la similitud contextual. Se han propuesto una variedad de algoritmos para realizar ese proceso de agrupación, desde los algoritmos divisivos

(construyen varias particiones y luego las evalúan), hasta los jerárquicos (crean una descomposición jerárquica y vertical de los elementos) pasando por los que se basan en la densidad o los que usan un determinado modelo (Alpaydin, 2004).

La aplicación particular que hemos utilizado se denomina CLUTO/GLUTO (Karypis, 2003). Es una aplicación gráfica multiplataforma que agrupa conjuntos de datos de baja y alta dimensión, y analiza las características de los distintos clusters. GCLUTO (el paquete de gráficos) se construye en la parte superior de la biblioteca de clústeres CLUTO. Está considerado muy adecuado para aglomerar conjuntos de datos que surgen en muchas áreas de aplicación, desde la recuperación de información, las transacciones de compra de clientes, la web, o en disciplinas científicas como la biología.

En la justificación del primer experimento detallaremos la clase de algoritmo que hemos utilizado, así como las ventajas clasificatorias que ha mostrado (apartado 3.1.3.1)

2.4.3.2 La teoría de la información y la entropía

El segundo ámbito teórico en el que encontramos una herramienta y medida de observación para el análisis del significado procede de la Teoría de la Información (Shannon, 1948; y Holik, 2016). Se trata del concepto de entropía.

En el caso de los experimentos (3.2 y 3.3) observamos y clasificamos, con una anotación manual doble, las palabras que ocupan la posición de **Z** de la tripleta tipo (X – P – Z) según las categorías semánticas descritas en la sección 2.3.2, página 83 de este capítulo. Una vez anotadas, usamos la medida de la entropía para determinar cómo es la distribución de categorías semánticas asociadas a cada tipo preposicional propuesto (funcional, semifuncional, semiléxico o léxico). En el cuarto experimento (3.4) también usaremos esta medida para determinar la dispersión de las ocurrencias de las respuestas del test de uso.

Desde la publicación de la teoría de la información su rentabilidad explicativa en la ingeniería de las telecomunicaciones ha sido fecunda por motivos y en direcciones muy variadas (Shannon, 1948). En el despliegue de los principios regulatorios sobre la

información a la medida cardinal se la denominó: entropía. El objetivo ingenieril y los esfuerzos iniciales de Shannon iban dirigidos a mejorar la eficiencia de los procesos temporales, energéticos y de costes en los canales técnicos de transmisión.

Hasta entonces en la historia de la ciencia, el concepto pertenecía a la termodinámica, la rama de la física que se ocupa del calor y sus propiedades. Suele decirse que la entropía termodinámica expresa el desorden de un sistema físico. Se caracterizó como el número de estados microscópicos distintos en los que pueden hallarse las partículas que componen un trozo de materia de forma que siga pareciendo el mismo trozo desde un punto de vista macroscópico (Boltzmann, 1877). La función explicativa en termodinámica es cuantificar la irreversibilidad de un proceso (Lemons, 2013), y medir el grado de desorden de un sistema. De hecho, esa asociación de la entropía con la naturaleza irrevocable del tiempo ha dirigido su poder predictivo a otras subdisciplinas de la física teórica (Carroll, 2009). No en vano, la capacidad adaptativa de este concepto lo lleva a encontrarlo también en otras disciplinas del conocimiento, como por ejemplo la biología o la economía.

La historia científica del concepto, así como el traspaso que se produjo de la física a la ingeniería se ha caracterizado profusamente (Abril, 1997). En otras ocasiones se ha descrito la entropía como una medida de ignorancia (Sethna, 2006). Del mismo modo, se ha relacionado esa naturaleza informativa con una noción que procede de la estadística, la impredecibilidad (Seife, 2009; y Ben-Naim, 2011).

Una de las fórmulas estándar de la entropía es la que se recoge en (82)

$$(82) H(X) = - \sum_{x \in X} p(x) \log_2 p(x)$$

Una definición genérica sería que el contenido de información de un mensaje aumenta a medida que la probabilidad de que el mensaje que tome esa forma disminuya, es decir, cuanto menos probable sea el mensaje mayor información contiene. De este modo,

la entropía de (x) se obtiene del sumatorio (\sum) de las probabilidades (p) de los mensajes en el conjunto de los mensajes posibles que una fuente puede enviarnos.

En nuestra propuesta sostendremos que el aumento de la entropía del conjunto de las categorías semánticas a las que pertenecen los sintagmas nominales que dependen de las preposiciones están correlacionado con la clase de preposición que predecimos: a mayor funcionalidad preposicional menor entropía en el grupo de sintagmas nominales, y viceversa, a mayor lexicidad preposicional mayor entropía. Especificaremos que el significado de la preposición se determina por la organización entrópica y semántica de sus contextos.

De este modo, como veremos en el análisis de los resultados del experimento 2, será la distribución contextual la que avale la predicción de las clases que conformen la gradualidad semántica de la categoría.

2.4.4 Procedencia de los datos empíricos: los corpus

La necesidad de usar, como plataforma empírica, grandes cantidades de datos organizados en corpus ha sido subrayada por algunos de los promotores de la semántica distribucional (Baroni y Lenci, 2010). Ese interés, junto al carácter prometedor de los modelos basados en corpus y otros aspectos fundacionales, ha sido referido ya en una larga trayectoria de aportaciones y comparaciones (Leech, 1991; Kennedy, 1998; McEnery, Tray y Wilson, 2001; Tognini-Bonelli, 2001; Hunston, 2002; Parodi, 2008; Kovatchev, Salomó y Martí, 2016; y Rojo, 2017).

El total de sintagmas preposicionales utilizados en los tres experimentos es de 79.097 (71.206 para el primero, 3898 en el segundo, 3903 en el tercero, y 90 para el cuarto), proceden de corpus de naturaleza, tamaño y características bien diferentes.

Los corpus de referencia, presentados en (83), han sido muy bien acreditados, respecto a su composición, sus usos y sus herramientas.

(83)

Corpus Semsem.

<http://grial.uab.es/sensem/corpus?idioma=ca>

Corpus Ancora.

<http://clic.ub.edu/corpus/es/ancora>

Corpus de la RAE.

<http://web.frl.es/CREA/view/inicioExterno.view;jsessionid=32DB08F5203BA9AE3C53FEBF7E3882C9>

Corpus del español (Mark Davies).

<http://www.corpusdelespanol.org/>

Corpus de Wikipedia.

<http://www.lsi.upc.edu/~nlp/wikicorpus/>

Corpus del ADIBA

<http://adiba.cat>

Corpus CUMBRE

(Almela, 2005)

2.4.4.1 Composición de los corpus

El Corpus Semsem alberga textos del mundo del periodismo y, en menor cantidad, del literario. Está formado por cerca de un millón de palabras, expresadas en 30.000 oraciones. Se ha construido a partir de la extracción de 125 oraciones de los 250 verbos más frecuentes en español. Se presenta la anotación manual del nivel sintáctico-semántico: roles semánticos, construcciones, categorías sintagmáticas, funciones sintácticas, modalidad y polaridad (Vázquez, Alonso, Capilla, Castellón y Fernández, 2006).

A su vez, AnCora es un corpus del catalán (AnCora-CA) y del español (AnCora-ES) con diferentes niveles de anotación: lema, categoría morfológica, constituyentes, funciones sintácticas, estructura argumental, papeles temáticos, clase semántica verbal, tipo denotativo de los nombres deverbales, sentidos de WordNet nominales, entidades nombradas y relaciones de correferencia (Martí, Taulé y Recasens, 2009).

El corpus de cada lengua contiene 500.000 palabras y están constituidos fundamentalmente por textos periodísticos: 75.000 del antiguo Lexesp, 250.000 de la agencia EFE y 200.000 de la versión española de El Periódico de Catalunya (Taulé, Recasens y Martí, 2008).

Por lo que respecta al Corpus de la RAE, el Corpus de Referencia del Español Actual (CREA) <http://web.frl.es/CREA/view/inicioExterno.view>, indicaremos que es un conjunto de textos de diverso origen, almacenados en soporte informático, del que es posible extraer información para estudiar las palabras, sus significados y sus contextos. En la versión 3.2, de junio de 2008, dispone de 160 millones de formas. Los textos proceden de todos los países donde se habla español en el período de 1975 a 2004. Respecto a la clase de registros los documentos han sido extraídos de libros, periódicos y revistas de más de cien entornos temáticos diversos, correspondiendo a un 90 % del total. La variedad hablada se representa por la transcripción de grabaciones de radio y televisión, siendo el 10 % del total del corpus.

Seguimos con el denominado Corpus del español, también conocido como Corpus Davies –debido a que su creador fue el profesor Mark Davies–, <https://www.corpusdelespanol.org/now/>. El corpus NOW del Corpus del español facilita el acceso a 6,6 billones palabras de periódicos y revistas basados en la web desde 2012 hasta la actualidad. Se trata de un corpus vivo en la medida en que mensualmente se añaden alrededor de 100 a 120 millones de palabras procedente de 200.000 a 250.000 artículos nuevos. Las posibilidades de búsqueda son extraordinarias: por palabras, lemas, categorías, comparaciones de frecuencias, de campos semánticos o las colocaciones (las palabras próximas a la búsqueda) hasta cuatro posiciones (Davies, 2002-2008).

Otro de las grandes corpus es el Wikicorpus <https://corpus.byu.edu/wiki/>. Es un corpus trilingüe (catalán, español, inglés) que contiene cantidades ingentes de palabras de

la Wikipedia, y que ha sido enriquecido automáticamente con información lingüística. En su versión actual alberga más de 750 millones de palabras. Ha sido extraído, segmentado y lematizado de manera automática (Atserias et alia, 2006). Además, y gracias a la biblioteca abierta FreeLing, se puede desambiguar (Reese et alia, 2010; y Medelyan et alia, 2009).

Del ámbito de las redes sociales, hemos usado algunos ejemplos del ADIBA, Atlas dialectal de Barcelona, un corpus elaborado a partir de Twitter. Nuestro uso ha sido, exclusivamente, para medir la frecuencia ciertas preposiciones en los archivos procedentes de Twitter con otros (véase la sección 3.2.3, página 142). Aunque los objetivos del corpus sean de carácter dialectológico y de contactos entre lenguas, nos ha servido como ejemplificación de una casuística más contemporánea y coloquial.

Por lo que respecta al corpus CUMBRE el uso ha sido, de nuevo, comparativo (Almela et alia, 2009). Así, hemos comparado la frecuencia de aparición de las preposiciones más habituales en este corpus con algunos de los anteriores (véase, de nuevo, la sección 3.2.3, página 142).

CAPÍTULO 3. LA GRADUALIDAD SEMÁNTICA PREPOSICIONAL. APRENDIZAJE AUTOMÁTICO, LA MEDIDA DE LA ENTROPÍA Y EL USO

Empezamos con la sección experimental de la investigación. Se compone de tres experimentos realizados a partir de los recursos y las herramientas metodológicas introducidos en el capítulo anterior. La estructura de cada uno de ellos es muy semejante: objetivos, procedencia de los datos, instrumentos de análisis y análisis de resultados. En primer lugar nos enfrentamos a la clasificación automática de un universo semántico relativamente restringido: los complementos de los verbos de movimiento. El segundo de los experimentos propone el análisis de los núcleos nominales en el contexto verbal de las preposiciones a partir de la medida entrópica de sus categorías semánticas. Esta misma herramienta nos servirá en el tercer experimento pero para contextos nominales, es decir, la X y la Z de las tripletas son nombres. Por último sugerimos el uso de hablantes –a través de un test- como estrategia para la verificación de nuestra hipótesis.

3.1 PRIMER EXPERIMENTO. La gradualidad semántica preposicional a través del aprendizaje automático (*clustering*). Un universo semántico específico: los verbos de movimiento en español

3.1.1. Objetivos y justificación del experimento

El objetivo primordial de este experimento es iniciar la verificación de la hipótesis sobre la gradualidad semántica de los valores léxico-semánticos preposicionales en español. Para ello hemos elegido un universo o área de significados específicos, el comprendido en el movimiento en el espacio, de manera real o figurada.

La preocupación, el interés y la importancia de los elementos lingüísticos que expresan conceptos y significados espaciales ha ido en aumento en las últimas dos décadas, tal y como se puede reseguir en una incesante multitud de referencias (Jackendoff, 1993; Svorou, 1994; Bloom et alia, 1996; Grabowski, 1996; Saint-Dizier, 2006b; Levinson, 2003 y 2006; Gehrke, 2006 y 2008; Aurnague, Hickmann y Vieu, 2007; Ashbury, 2008; Kelleher, Sloan y Mac Namee, 2009; Asbury et alia, 2008; Cinque y Rizzi, 2010; Demonte, 2011; Romeu, 2014; y Bosque, 2018). En esta nómina se mencionan trabajos de diversas procedencias, metodologías y lenguas, adheridas a la teoría sintáctica algunas, o a la visión psicolingüística de la gramática cognitiva otras. La expresión espacial es, en cualquiera de esas ópticas de estudio, un hecho insoslayable de la tarea semántica de las lenguas humanas; así como su relación con la expresión temporal (Haspelmath, 1997).

Asimismo, el desarrollo de los análisis computacionales ha abordado las interpretaciones semánticas de las preposiciones atendiendo a modelos predictivos de sus significados en aplicaciones distintas (Alam, 2003; Lersundi y Agirre, 2003; Costello, Kelleher y Volk, 2007; Baldwin, Kordoni y Villavicencio, 2009; y Boleda y Herbelot, 2016), o en encuentros que han tratado los aspectos más heterogéneos y variados de la categoría (Arsenijevic, Baldwin y Trawinski, 2006).

Para nuestro experimento hemos elegido 104 verbos de movimiento del español que se encuentran en la lista de (84).

(84)

Acarrear, acercarse, acodar, acompañar, acudir, alejarse, agacharse, agitar, andar, apartar, apoyarse, aproximarse, apuntar, arrastrar, arrimar, arrinconar, asentar, atraer, bajar, brincar, caer, callejear, caminar, circular, contener, conducir, continuar, correr, curvar, descender, deslizarse, desplazarse, desviar, dirigirse, doblar, doblarse, enfocar, encaminar, encarrilar, encauzar, enderezar, enviar, empujar, entrar, escalar, escapar, escorar, escurrirse, extenderse, fluir, galopar, girar, guiar, golpear, huir, hundirse, inclinarse, ir, juntar, ladear, lanzar, llevar, marcharse, montar, desmontar, montar, mover, moverse, removearse, orientar, pasar, pasear, partir, permanecer, orientar, portar, portear, rastrear, reclinar, recorrer, remolcar, retirarse, salir, saltar, sacudir, seguir, señalar, sobrevolar, subir, tirar, traer, trasladar, trajinar, transportar, traspasar, transvasar, vagabundear, viajar, volar, voltear, volver, venir, zarandear y zarpar.

En nuestra selección obviamos, por razones de inconveniencia metodológica e interpretativa, la distinción reconocida en la literatura entre verbos de desplazamiento y verbos que indican formas de desplazamiento (Cano Cambroner, 2009). Sobre la eficacia y la pertinencia sintáctica de esa separación y su reconocimiento empírico se han presentado algunas posiciones escépticas respecto de esa distinción (Morimoto, 2001; y Cuartero, 2006); y de manera complementaria algunas descripciones sintácticas sobre la verticalidad en el movimiento para algunos verbos como *subir*, *alzarse*, *ascender*, *elevarse...* o *bajar*, *descender*, *hundirse* y *caerse* (Bosque, 2015).

Es ese contexto predicativo, representado en la característica tripleta de (85), donde X es un verbo de movimiento de la lista anterior (84), P una de las preposiciones seleccionadas, *a*, *hacia*, y *hasta*, y Z cualquier clase de palabra que se identifique en la selección del corpus, donde se produce nuestro experimento.

(85) X – P – Z

Hemos elegido *a*, *hacia* y *hasta* porque algunos de sus valores son paradigmáticos de la expresión semántica locativa y espacial, tanto en sentidos reales como los figurados. Como vamos a ver, esa representatividad encaja con una de las condiciones que compone nuestra hipótesis: la flexibilidad y/o gradualidad de sus distintos usos, en la medida que conviven en diferentes contextos.

Con estas restricciones o elecciones previas la concreción del objetivo se realiza a partir de la determinación de la semántica de los nombres que son cercanos o próximos a las preposiciones a partir de las agrupaciones que ahora explicaremos.

3.1.2. Valores semánticos y clases preposicionales de *a*, *hacia* y *hasta*

En la lista de valores descriptivos que presentamos en el Capítulo 2, sección 2.3.2, página 69, las tres preposiciones del experimento participaban en un mayor número de

valores léxico-semánticos; 14 casos de la *a*, y 5 *hacia* y *hasta*. Ahora, y ya en el universo del movimiento o el desplazamiento espacial, se restringen sus posibilidades a 6, 3 y 3, respectivamente. Esta limitación, impuesta por el dominio, ha favorecido la capacidad predictiva de nuestra herramienta de cálculo de acuerdo, como vamos a analizar, con los resultados obtenidos.

En los cuadros 8, 9 y 10 se presentan las clases de preposiciones propugnadas para el experimento, las caracterizaciones descriptivo-semánticas a partir de la presentación en las asunciones semánticas (Capítulo 2, página 80 y siguientes) y ejemplos correspondientes a las tres preposiciones elegidas.

Preposición <i>a</i>	
Tipo de valor semántico y descripción	Ejemplos
<p>Funcional (F1) Introduce el objeto, el destinatario y el beneficiario de la acción verbal (los llamados CD –complemento directo u objeto-, CI -complemento indirecto-).</p> <p>Forma parte de las perífrasis verbales o de las locuciones prepositivas, adverbiales o conjuntivas.</p>	<p>Se lanzó a criticar Va a empezar Pasa a exigir Vuelve a arremeter Envían a Pedro</p>
<p>Semifuncional (SF1)</p> <p>Expresa o señala el beneficiario de la acción o el régimen verbal obligatorio.</p>	<p>Tratan de decidir Abogan por condenar</p>
<p>Semiléxico 1 (SL1)</p> <p>Expresa el medio, el instrumento, el modo o la costumbre en que se ejecuta o hace algo.</p>	<p>Fueron al rescate Andan a gritos Continúan a nado</p>
<p>Semiléxico 2 (SL2)</p> <p>Localiza en el espacio objetos, personas o conceptos respecto de un punto (sin desplazamiento, como una mera orientación), en sentido literal o figurado,</p>	<p>Parte a la reunión Acude a la convocatoria</p>

e indica el punto, la zona o el lugar concreto en el que sucede o se localiza algo.	
Léxico 1 (L1) Señala y expresa la dirección, la orientación, el destino, el término, el lugar intermedio, el curso o los límites respecto de un punto o de una acción.	Van a Madrid Se dirigen al límite
Léxico 2 (L2) Expresa el punto, el término o el límite temporal de una acción en combinación con <i>de</i> , u otras preposiciones.	Vuelven a la semana Recorren al atardecer Regresa a las tres

Cuadro 8. Descripción y ejemplos de los valores semánticos de a

Preposición <i>hacia</i>	
Tipo de valor semántico y descripción	Ejemplos
Semiléxico (SL1) Indica el receptor humano o abstracto de un movimiento real o figurado (con verbos de movimiento).	Desplazan hacia los afectados Aparta hacia la democracia
Léxico 1 (L1) Señala y expresa la dirección, la orientación, el destino, el término, el lugar intermedio, el curso o los límites respecto de un punto o de una acción.	Fluye hacia el oeste Acompañan hacia adelante
Léxico 2 (L2) Expresa el punto, el término, la dirección o el límite temporal de una acción, en combinación o no con <i>de</i> , u otras preposiciones.	Andan hacia la fase Asientan hacia siglos

Cuadro 9. Descripción y ejemplos de los valores semánticos de hacia

Preposición <i>hasta</i>	
Tipo de valor semántico y descripción	Ejemplos
Semiléxico (SL1) Introduce los complementos de infinitivo, estados abstractos, e incluyen los elementos de la acción.	Pasan hasta comprender Llegan hasta la locura
Léxico 1 (L1) Expresa término, límite o lugar donde acaba una acción.	Escalan hasta la cima Transportan hasta Amsterdam
Léxico 2 (L2) Expresa el punto, el término, la dirección o el límite temporal de una acción, en combinación o no con <i>de</i> , u otras preposiciones.	No se retiran hasta las tres Van hasta el domingo

Cuadro 10. Descripción y ejemplos de los valores semánticos de hasta

Establecemos, por tanto, la pertinencia de entre tres grupos (*hacia* y *hasta*) y seis grupos de valores semánticos (*a*). Las denominaciones que usamos son funcional, semifuncional, semiléxico 1, semiléxico 2, léxico 1 y léxico 2.

Esas denominaciones fueron surgiendo durante el análisis sintáctico-categorial, como examinamos en el estado del arte, de manera paulatina a partir de la década de los 80 del siglo XX. Se concentraron, entonces, un conjunto de prospecciones empíricas preposicionales en diferentes familias lingüísticas: anglo germánica, románica, escandinava y eslava, por ejemplo. Ese nuevo elenco de datos junto con la propuesta explicativa de las categorías funcionales abre un camino muy fructífero: la disgregación del binarismo chomskiano [+/-N] y [+/-V].



La división inicial entre categorías funcionales y categorías léxicas fue aproximándose a nuevas subdivisiones (semiléxicas y semifuncionales, fundamentalmente) a medida que se mejoraba la profundidad empírica (Van Riemsdijk, 1990; Zeller, 1999; Dechaine, 2005; y Zwarts, 2005).

De manera paralela, y ya desde los orígenes de la gramática cognitiva y su teoría de los prototipos (Berlin y Kay, 1969; y Rosch, 1973), se adoptó una visión gradualista y

no discreta de la definición de las categorías lingüísticas. En este entorno de propuestas surgió la noción de gradualidad para indicar que dentro una clase de palabras existiría una más representativa que las demás, denominado el ejemplar prototípico (Lakoff, 1987; Langacker, 1991 y 1999). Sobre cómo se ha formulado ese estilo de categorización y sus consecuencias en la interpretación del significado tenemos algunas referencias necesarias (Cifuentes, 1996; Moure, 1996; Rodríguez y Pena, 2008; e Ibarretxe-Antuñano, 2010, 2013 y 2015).

Los cuatro tipos y sus divisiones, que han sido abreviados como F1, SF1, SL1, SL2, L1 y L2 en las tres tablas anteriores, constituyen los diversos grupos del modelo clasificado manualmente y se usan para la comparación con las clases resultantes del *clustering* automático.

Recuperamos el cuadro 11, presentado en el capítulo sobre las asunciones metodológicas, para recordar que en el primer experimento se produce, gracias a las evidencias de los resultados del *clustering*, una división en la clase léxica para las tres preposiciones.

CUADRO SOBRE LA ORGANIZACIÓN FLEXIBLE DE LAS CLASES PREPOSICIONALES						
Experimentos 	Clases preposicionales 	Funcional	Semifuncional	Semiléxica	Léxica	
Primer experimento (preposiciones <i>a</i> , <i>hacia</i> y <i>hasta</i>)					<i>L1</i>	<i>L2</i>
Segundo experimento Estructuras del ámbito verbal (todas las preposiciones)						
Tercer experimento Estructuras del ámbito nominal (todas las preposiciones)						
Cuarto experimento (todas las preposiciones)						

Cuadro 11. Organización de las clases preposicionales por experimentos

3.1.3. Procedimientos: el *clustering* en el aprendizaje automático

Como vimos en la sección 2.4.3.1, el aprendizaje automático o *machine learning* se ha venido incorporando a dos ámbitos que se solapan: la inteligencia artificial y las ciencias de la computación. En la figura 1 (Gómez, Juristo, Montes y Pazos, 1997) se recoge una representación de esta intersección múltiple e interdisciplinaria que visualiza el estilo metodológico de nuestra investigación. Sin esa conexión entre disciplinas, herramientas, conceptos, protocolos y tradiciones sería del todo imposible el acceso a determinadas cuestiones que surgen en el curso de nuestra hipótesis. De procedencias distintas se ha ido convergiendo en un interés común: el estudio de los lenguajes formales, artificiales y naturales como un área con muchas confluencias.

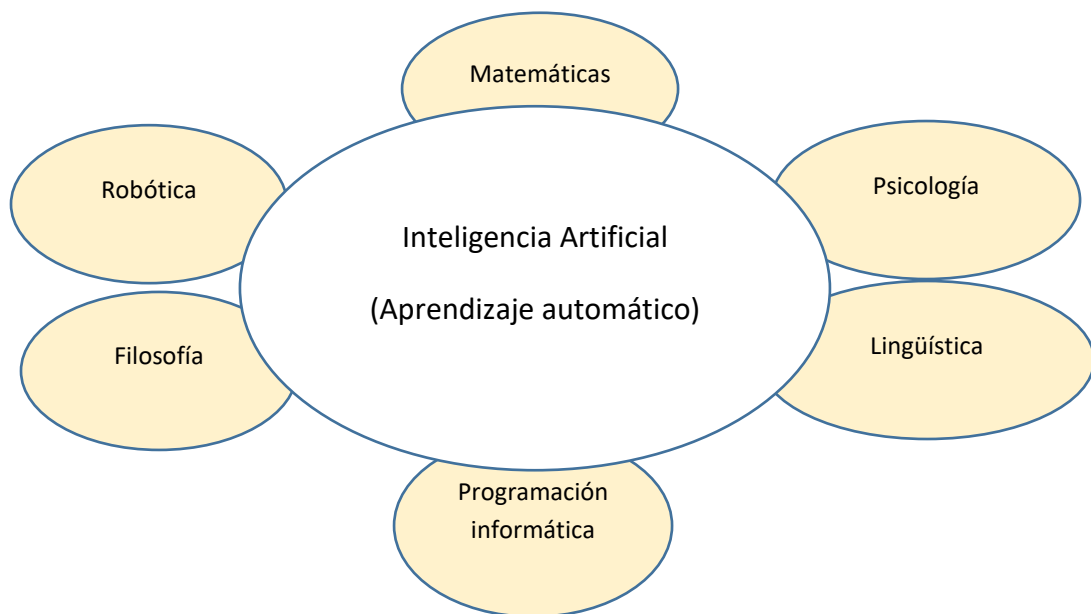


Figura 1. La ubicación del aprendizaje automático en relación a otras áreas de conocimiento (Gómez, Juristo, Montes y Pazos, 1997)

La historia de esta sucesiva colaboración entre áreas del conocimiento con objetivos comunes se ha tratado en relación a los periodos que han conformado el problema del aprendizaje (Langley, 1996; y Vapnik, 1998). Cómo se consiga aprender es, efectivamente, el reto cardinal de este espacio interdisciplinar. Se entiende, por tanto, que el aprendizaje automático es el conjunto técnicas, metodologías y algoritmos con capacidad para el aprendizaje y la mejora de su eficacia.

A lo largo del tiempo, y de manera simultánea, se han propuesto un conjunto de estrategias, metodologías, heurísticas y recursos, tal y como describe la figura 2 (Cetinsoy, Martin, Ortega y Peterson, 2015). En el eje de las ordenadas se indica la interpretabilidad como una la característica que ha ido en aumento en las potencialidades de esos procedimientos: desde el Perceptron hasta el Deep Learning.

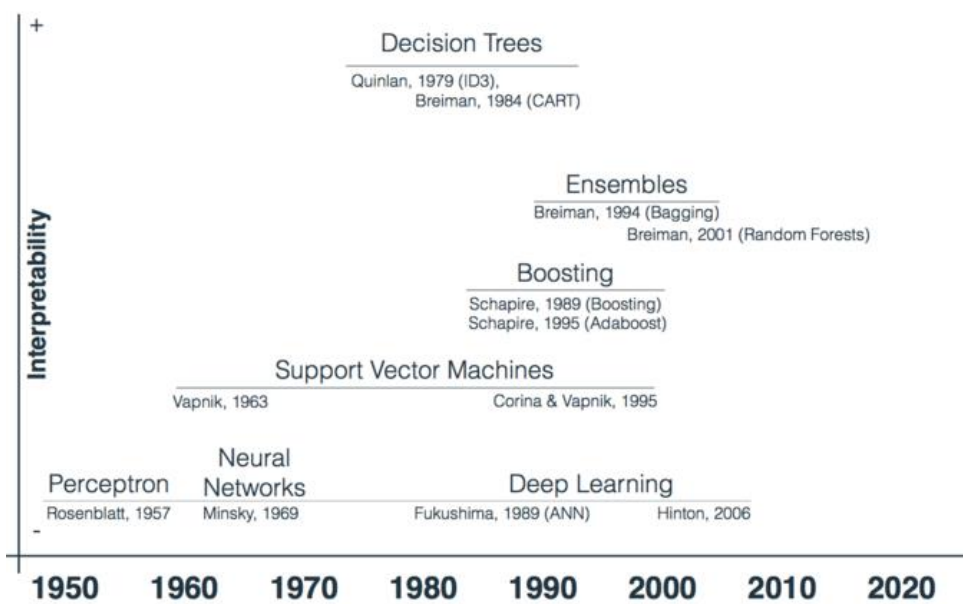


Figura 2. Cronología de los procedimientos y heurísticas sobre el aprendizaje automático

Nuestra herramienta de análisis pertenece al aprendizaje no supervisado, es decir, un tipo de algoritmo que no dispone de elementos de comparación, etiquetaje o referencia para comparar los resultados obtenidos. Digamos que no hay criterios previos para orientar, sugerir o mostrar si estamos ante una buena agrupación.

Uno de los tipos de algoritmos que opera en el aprendizaje no supervisado es el *clustering* o agrupación. De manera general diremos que estamos ante una técnica de agrupación de datos y/o objetos cuya intención es identificar similitudes de algún tipo. Se usa como un proceso para encontrar una estructura significativa, procesos subyacentes explicativos, características generativas y agrupaciones inherentes a un conjunto de ejemplos. A partir de la división de la población (expresión genérica usada en el área) – serían los conjuntos de tripletas en nuestra hipótesis- o los puntos de datos en varios grupos, se busca que sean más similares a otros puntos de datos en el mismo grupo y diferentes a los puntos de datos en otros grupos.

Existen cinco métodos de agrupación, atendiendo a los criterios de organización, selección y disposición de las semejanzas, tal y como se relacionan en (86).

(86)

- Método jerárquico
- Método basado en la densidad
- Método basado en cuadrícula
- Método basado en modelos
- Método basado en restricciones

El utilizado en nuestro experimento es de base jerárquica. Los clústeres formados en este método forman una estructura arbórea basada en la jerarquía. Los nuevos grupos surgen a partir del inicial. Se divide en dos categorías: aglomerativos o aglomerante, y divisivo (Holzinger, 2016).

La herramienta de análisis que utilizamos es la aplicación CLUTO/GLUTO (Karypis, 2003).

CLUTO es un paquete de software para agrupar conjuntos de datos de baja y gran dimensión y para analizar las características de los distintos clusters. CLUTO es muy adecuado para almacenar conjuntos de datos que surgen en diversas áreas de aplicación, incluida la recuperación de información, transacciones de compras de clientes, web, ciencia y biología.

Hemos utilizado la técnica del *Word Embeddings* para construir los vectores de CLUTO (Sahlgren, 2006; Turney y Pantel, 2010; Erk, 2012; y Mikolov et alia, 2013).

Las *Word Embeddings* son el nombre genérico que recibe un conjunto de modelos de lenguaje y técnicas de aprendizaje de características en el Procesamiento del Lenguaje Natural (PLN) donde las palabras o frases del vocabulario se asignan a vectores de números reales. Esta forma de proceder supone una correspondencia matemática desde un espacio con una dimensión por palabra a un espacio vectorial continuo con una dimensión mucho más baja. Los métodos para generar estos vínculos han sido diversos, desde las redes neuronales, hasta modelos probabilísticos (Levy y Goldberg, 2014).

En el ámbito lingüístico, las *Word Embeddings* han sido utilizadas desde la investigación en semántica distribucional (Boleda y Herbelot, 2017). El propósito fue la cuantificación y la clasificación de similitudes del significado entre componentes lingüístico a partir de los rasgos de su distribución en muestras enormes de corpus lingüísticos. Subyace una vieja idea de uno de los fundadores de la lingüística cuantitativa J. R. Firth (1957) según la cual, una palabra se puede caracterizar a partir de la compañía que tiene.

En la primera década del siglo XXI se comienza a desarrollar la técnica de las *Word Embeddings* (Bengio et alia, 2006) con la intención de reducir la alta dimensionalidad de las representaciones de palabras en contextos. El despegue del área llega después de 2010 con el avance en la calidad de los vectores y la velocidad de entrenamiento del modelo. En 2013, un equipo de Google liderado por Tomas Mikolov creó word2vec, un kit de herramientas que puede entrenar modelos de espacio vectorial mejorando, todavía más, los modelos anteriores.

El objetivo de este procedimiento, por tanto, es cuantificar y categorizar propiedades semánticas entre elementos lingüísticos a partir de los contextos donde coocurren. Estos modelos de espacio vectorial representan (incrustan) palabras en un espacio vectorial continuo en el que palabras semánticamente similares se asignan a puntos cercanos ("se incrustan entre sí").

Usamos *Word Embeddings* porque es una manera de representar las palabras que captura información semántica en el contexto o la distribución de las preposiciones.

El parámetro de algoritmo de clusterización de CLUTO que mejor agrupa, a partir de los resultados obtenidos en el experimento, se denomina *-direct*. En este método, la solución de clúster de *k-way* deseada se calcula al encontrar simultáneamente todos los *k* grupos. En general, la computación de una agrupación *k-way* directamente es más lenta que la agrupación a través de bisecciones repetidas. En términos de calidad, y solo para valores razonablemente pequeños de *k* (generalmente menos de 10-20), el enfoque directo (*-direct*) conduce a mejores conglomerados que aquellos obtenidos mediante bisecciones repetidas. Sin embargo, a medida que *k* aumenta, el enfoque de bisección repetida tiende a ser mejor que el agrupamiento directo.

3.1.4 Datos lingüísticos

Partimos de un corpus de análisis de 71.206 tripletas procedentes de dos corpus: Ancora y Wikicorpus. Su distribución por preposición es de 57.815 para la *a*, 6.389 para *hacia* y 7.102 para *hasta*.

3.1.5 Experimento

A partir de la triple elección: 103 verbos de movimiento del español, las preposiciones *a*, *hacia* y *hasta*, y tres grupos de valores léxico-semánticos (indicados en los cuadros 8, 9 y 10), procedemos a la extracción de los ficheros de los corpus de las tripletas con estructuras sintácticas como las de (87).

(87)

V + P + N (verbo, preposición y nombre) y/o V + P + V_{infinitivo} (verbo, preposición y verbo en forma infinitiva).

Nuestro análisis ha partido de las formas de los corpus. La comprobación con los lemas no ha mejorado ni ha añadido información relevante en la confirmación y la validación de nuestra hipótesis. Asimismo, hemos elaborado el experimento con las funciones sintácticas sin obtener resultados relevantes ni significativos.

El siguiente paso ha sido confeccionar manualmente un archivo de evaluación con anotación humana a partir de los grupos funcional, semifuncional, semiléxico y léxico. El anotador principal clasifica un total de 1731 tripletas de la preposición *a*, 654 de la preposición *hacia* y 779 de la preposición *hasta*. Para mejorar nuestra clasificación y disminuir el riesgo intersubjetivo hemos realizado una segunda anotación humana a partir de una muestra de 150 tripletas. El grado de coincidencia en la clasificación de ambos anotadores es del 79,2 %.

Una vez elaborados y disponibles los archivos se generan los vectores con Word2vec y CLUTO los agrupa. Desde estos vectores, y como último paso, se produce una comparación entre los archivos de anotación humana y los resultados de agrupación de CLUTO.

El programa de comparación facilita una serie de cuadros que relacionan los valores (F1, SF1, SL1, SL2, L1 y L2) con el número de los clusters, junto con un índice porcentual de pureza y de pureza inversa en la organización de los grupos.

Por *pureza* entendemos la concentración de cada columna de la tabla. Te indica como de distribuidos entre diferentes categorías están los elementos que han quedado en el mismo clúster. Si todos los elementos del clúster están en una misma categoría la pureza de aquella columna sería el 100 %. Frente a la *pureza inversa* que significa lo mismo pero observando la concentración por las filas. Te indica como de distribuidos en diferentes

clusters están los elementos de la misma categoría. Si todos los elementos de la categoría están en el mismo clúster, la pureza inversa de la fila sería, a su vez, de 100%

Hemos realizado la comparación para el caso de la *a* con 6, 5, 4 y 3 clusters respectivamente; y con 3 para *hacia* y *hasta*.

Veamos, a continuación, la valoración de los resultados a partir, primero de los porcentajes globales; y luego sobre el tipo de concentración por fila (clase preposicional) y por columna (clúster del algoritmo), tanto cuantitativamente como cualitativamente, para cada preposición. Para el análisis cualitativo del contenido de los clusters, es decir, las palabras adyacentes a las preposiciones en posición de complemento usamos las denominaciones de las categorías semánticas propuestas en las asunciones (Capítulo 2, sección 2.4.3, página 83 y siguientes), y que repetimos en (88).

(88) *Animado, Inanimado, Entidad abstracta, Evento, Temporal y Locativo.*

Por lo tanto, una vez codificados los nombres usando los vectores y agrupados por proximidad semántica podremos comparar el resultado respecto de la clasificación manual.

3.1.6. Análisis de los resultados

La contrastación empírica de nuestra hipótesis produce la confirmación en la medida en que las agrupaciones establecidas y comparadas con las propuestas por el algoritmo de CLUTO, la herramienta de aprendizaje automático propuesta, se corresponden de manera mayoritaria con los ejemplos etiquetados manualmente. Las cifras de coincidencia oscilan entre el 54,13 % y el 64 %, dependiendo del número de clusters, y de la preposición; y supera el 70 % en algunos casos si apreciamos la similitud con las celdas adyacentes o el contexto de los clusters de cada preposición, tal y como comentaremos a continuación.

En el gráfico 1 hemos recogido los porcentajes promedio de coincidencia entre la agrupación automática que hace el algoritmo de CLUTO y la clasificación del anotador humano para las tres preposiciones, *a*, *hacia* y *hasta*.

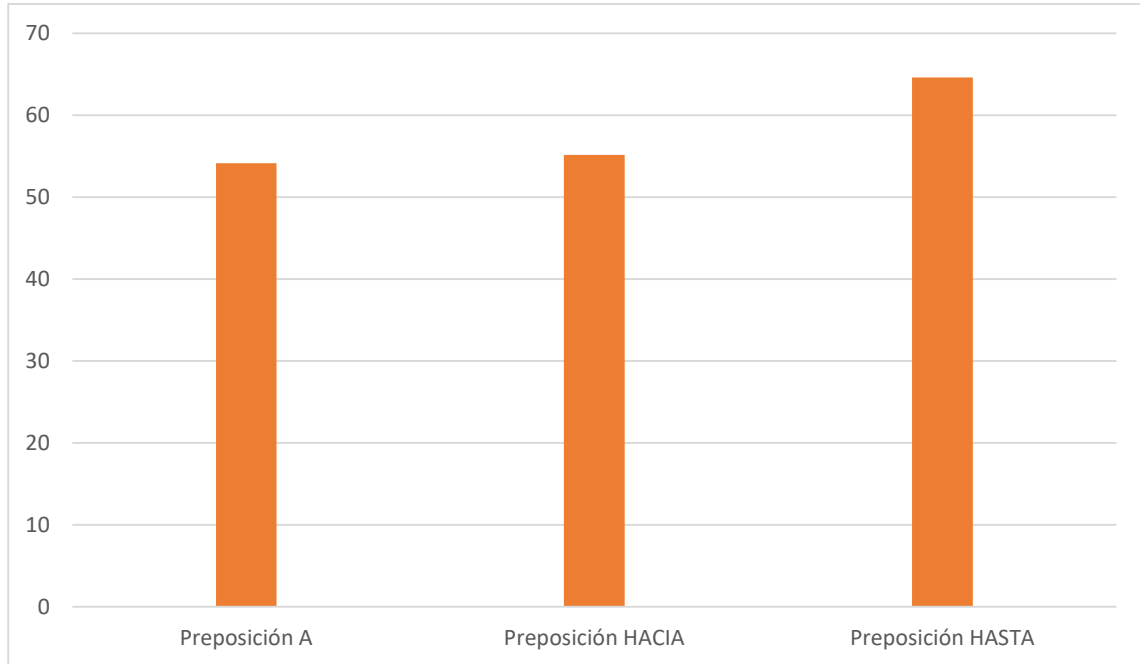


Gráfico 1. Comparación porcentual entre las tres preposiciones del experimento.

En el caso de la *a*, cuya organización semántica es más diversa, compleja y polisémica, hemos superado el 50 % en cuatro puntos (54,13 %) y, cómo vamos a ver, con una coherencia interna de los grupos relativamente estable. La preposición *hacia* sube casi un punto más y llega a 55,17 %. Ese aumento tiene que ver, también, con la progresiva reducción de su espectro interpretativo. Y, por último, *hasta* alcanza un total 64,6%. Este hecho, como destacaremos, guarda relación con la delimitación conceptual que afecta a su contexto semántico.

Seguidamente detallamos los resultados cuantitativos para cada una de las preposiciones, *a*, *hacia* y *hasta*, así como las clases preposicionales que se albergan en los clusters. En las tablas se recogen el nombre de la preposición, el parámetro de

clusterización (*direct*), el número de clusters, el número de tripletas evaluadas en la comparación, y las columnas con la numeración asignada por el programa CLUTO.

3.1.6.1 La preposición *a*

Hemos ensayado cuatro agrupaciones distintas para la *a*. El objetivo de estos exámenes es comprobar qué agrupación es significativa semánticamente para nuestra hipótesis en cada una de las clases propuestas; desde los seis valores que explicitan los significados de los verbos de movimiento hasta una progresiva concentración (de 6 a 5, de 5 a 4 y de 4 a 3); primero del espectro funcional (F1 + SF1), después del semiléxico (SL1 + SL2) y, por último, del léxico (L1 + L2).

Las fases de simplificación de las clases se representan visualmente en el cuadro 12.

6 grupos	5 grupos	4 grupos	3 grupos
Funcional	Funcional	Funcional	Funcional
Semifuncional			
Semiléxica 1	Semiléxica 1	Semiléxica	Semiléxica
Semiléxica 2	Semiléxica 2		
Léxica 1	Léxica 1	Léxica 1	Léxica
Léxica 2	Léxica 2	Léxica 2	

Cuadro 12. Simplificación de la agrupación automática para la preposición *a*

Los resultados de pureza y pureza inversa, tal y como se observa en el gráfico 2, han sido variables en esas cuatro agrupaciones. El porcentaje máximo de variación entre los distintos grupos ha oscilado entre los 11 % puntos de la pureza (azul) hasta los 17 % puntos de la pureza inversa (marrón).

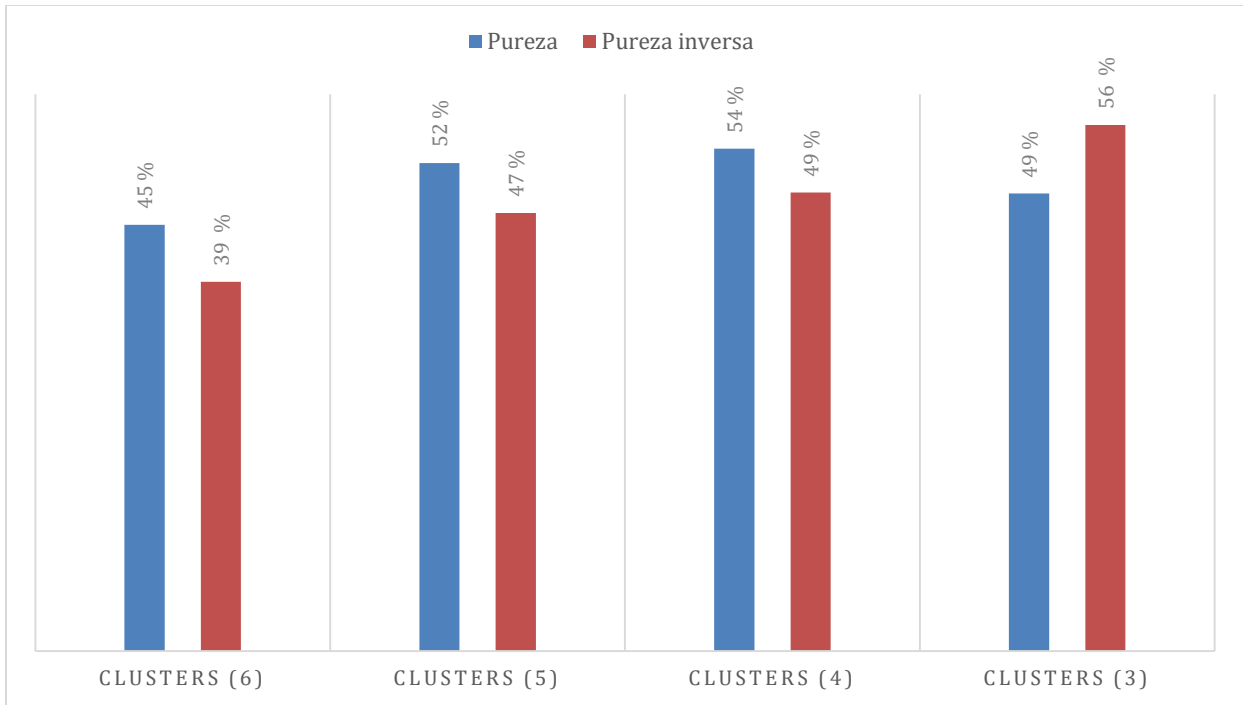


Gráfico 2. Porcentajes de pureza (azul), y pureza inversa (marrón) por el número de clúster para la preposición a.

Una vez observados los resultados, las agrupaciones más representativas de la capacidad automática de organización por similitud semántica en relación a nuestra hipótesis pertenecen a la fase 4.

A su vez, en la tabla 1 hemos reunido los resultados numéricos de la agrupación de 4 clusters, el número de tripletas y el método de agrupación (*direct*), junto con las cuatro clases preposicionales (funcional, semiléxico, léxico 1 y léxico 2) que vamos a analizar de manera pormenorizada porque son los que se adaptan mejor a nuestra hipótesis. Las otras agrupaciones, 6, 5 y 3 ofrecían distribuciones con menos relevancia sobre la predicción.

El orden de la numeración superior (0, 3, 1 y 2) es arbitrario y está propuesto por el algoritmo.

A/DIRECT/4 1731	0	3	1	2
F1	279	73	118	141
SL	100	281	75	52
L1	5	162	236	45
L2	54	18	33	59

Tabla 1. Agrupación en 4 clusters y 4 valores de la preposición a

La diagonal, sombreada en color gris, recoge un total de 855 casos (49.39 % del total). Aunque la numeración de los grupos se asigne de manera azarosa, el hecho de que coincidan el número más alto de palabras en esa línea –y de manera destacada en tres de los grupos- permite ser optimista en la predicción antes de analizar su contenido. Si, además, sumamos las dos diagonales próximas (superior e inferior) –donde se hallan ejemplos, tipos y/o clases cercanas semánticamente, alcanzamos 1343 casos (77,58 %). No obstante, esta circunstancia no sucede en el clúster 2, donde el grupo (L2 2) no corresponde con el grupo más numeroso.

Por lo que respecta al análisis de casos particulares iniciamos nuestro comentario con la funcionalidad preposicional que se recoge en los clusters (F1 0) y (F1 2), sumando un total de 420 casos, es decir, un 69,30 % del total de la fila F1.

En el primer clúster (F1 0) la totalidad de las palabras en posición de complemento de la triplete son del ámbito nominal, en su expresión categorial recta, el nombre; u otras oblicuas o derivadas como el infinitivo o adjetivos sustantivados. Mayoritariamente concurren infinitivos verbales (216 de 279): *acarrear a contradecir, acelerar a aparecer,*

ir a cantar, llevar a actuar, acudir a presentar, aparcar a ver, continuar a denunciar, detener a atravesar, enfocar a conseguir, desmontar a soportar, doblar a recibir, entrar a actuar, mover a otro, hundir a salir, ir a vivir, lanzar a arrancar, orientar a apoyar, remover a limpiar, saltar a tocar, salir a buscar o transportar a investigar. El resto de palabras (63 de 279) son palabras en función nominal que pertenecen a la categoría semántica de *Animado*. Esta es la propiedad más frecuente de los sintagmas nominales en posición argumental de complemento directo o indirecto. Algunos ejemplos son *anarquistas, adversarios, enemigos, espectador, ambos, otro, alma, dioses, antepasado, alguien, intruso, persona, necesitado, anciano, agente, jueces, nadie, lector, sujeto, contacto, forastero, afectado, director, madre, anfitriones, futbolistas, amigos, grupo, sospechoso, enemigo, humanos, ancianos, abogado, cliente, hijo, villano, pareja, chicos, novia y sujeto.*

El segundo clúster más representativo de la funcionalidad es el (F1 2). Contiene un total de 141 palabras, de las cuales 139 casos pertenecen, de nuevo, a la categoría semántica *Animado* con la forma de nombres como *director, padre, actriz, esposo, japonesa, poeta, adolescente o personaje.*

En ambos casos (F1 0) y (F1 2), la tarea contributiva de la preposición es casi nula. El algoritmo asocia un contexto de semejanzas que permite observar esa coincidencia tan alta en los dos grupos mayoritarios. Parecería que en el primer grupo la coincidencia de formas verbales crea una similitud al margen del significado. En cambio, en el caso de los nombres, es muy revelador que la totalidad de casos compartan el rasgo de la categoría semántica *Animado*. En relación con esta clase de complementos, tradicionalmente denominados dativos (Barnes, 1980; Czepluch, 1982; Belle y Langendonck, 1996; y Fernández-Alcalde, 2013), se ha presentado una tipología que incluye este rasgo (Delbecque y Lamiroy, 1999), junto con hechos semánticos sobre el potencial explicativo de la preposición (Saavedra, 2017).

En los otros dos clusters funcionales (F1 1) y (F1 3), en cambio, se concentran nombres procedentes de otras categorías semánticas como *Inanimado (aeroplano, petrolero, objeto, botones, llama), Evento (batalla, casos, conflicto), y Entidad abstracta (budismo, sociedad, belleza, control).*

La expresión de la funcionalidad se produce y se concentra por tanto, a través de dos expresiones formales: los infinitivos (clúster 0) y los nombres de categoría *Animado* (clúster 2). Por otra parte, en los dos clusters secundarios de la funcionalidad respecto al número de tripletas, (1 y 3) continúan la tendencia.

Además, en una observación de los nombres que encontramos en (1 y 2) se aprecian algunos rasgos descriptivos coincidentes. En el primero hallamos un predominio de profesiones o funciones sociales como *comerciantes, empresarios, funcionarios, militares, brigada, caballero, ministro, almirante* o *jinetes* y, en cambio, en el segundo sobresalen más relaciones personales y familiares como *madre, esposo, hijo, amigo, marido, padre, pareja, hermano, amantes* o *sobrinos*.

En la comparación con el segundo clúster de la columna, el (SL 0), certificamos la similitud y la continuidad a partir de palabras que expresan *Evento* y *Animado*, representando un 30 % del total.

El segundo grupo de valores semánticos se encuentran en la expresión de la semilexicidad (SL), a medio camino entre la marca funcional, meramente gramatical o de vacío semántico, y la progresiva adquisición de contenido léxico. En el clúster (SL 3), con 281 casos, se localizan ejemplos que pertenecen a cuatro categorías semánticas *Locativo* (86), *Inanimado* (38), *Animado* (34) y *Entidad abstracta* (11).

Además, todavía en ese clúster, encontramos un conjunto considerable de tripletas modales, considerados por la gramática descriptiva como locuciones prepositivas o frases fosilizadas o gramaticalizadas (Cifuentes, 2003). Son un total de 112 casos. Serían ejemplo como *a ojo, a pie* o *a caballo*. Este fenómeno gramatical no altera la tarea de comprobación de la similitud semántica. El hecho de que confluyan en el mismo grupo, y el algoritmo incluya en un contexto próximo de similitud a estas expresiones avala nuestra predicción porque carece de esa información formal (el hecho de que sean frases hechas). Es revelador, asimismo, que en la organización del clúster se encuentren entre los semiléxicos. De algún modo, el papel de la preposición empieza a ser valioso porque convive la manera de la acción con el lugar donde se realiza.

Se trata de ejemplos como los de (89).

(89) *Acudir a fuerza , acudir a pie, alejar a estilo andar a aire, andar a galope, andar a pedal, andar a pie, andar a ritmo, andar a velocidad,, atraer a mano, apuntar a forma, arrastrar a mano, caer a coste, conducir a riesgo, conducir a vista, correr a oscuras, correr a ritmo, desmontar a nivel, enfocar a aspecto, entrar a ojo, entrar a nivel, recorrer a galope, acelerar a pie, andar a aire, aproximar a mano, arrastrar a cuerda, bajar a pie, circular a profundidad, descender a exceso, descender a paso, entrar a ojo, extender a complejos, galopar a velocidad, fluir a pie o montar a caballo...*

Por lo que respecta a la expresión del lugar algunos ejemplos se pueden leer en (90).

(90) *Aparcar a aire, bajar a profundo, brincar a techo, caminar a largo, conducir a extensiones, contener a alojamiento, curvar a arco, arrastrar a afuera, circular a profundidad, y desplazar a alrededor...*

En la tabla 2 mostramos una distribución porcentual para mostrar cómo se distribuyen las categorías semánticas en el clúster (SL 3), el mayoritario.

Tipo de clase semántica	Número de casos en el clúster	Porcentaje
Modal	112	39,85 %
Locativo	86	30,60 %
Animado	38	13,52 %
Inanimado	34	12 %
Entidad abstracta	11	3,9 %

Tabla 2. Porcentajes de categorías semánticas del clúster semiléxico correspondiente a la preposición a

El aumento de dispersión de las categorías semánticas en esta clase se corresponde con nuestro pronóstico. Este hecho supone, además, que la aportación semántica de la preposición también crece. Así, el predominio de la expresión de modo en las tripletas (39,85 %) permite reconocer una progresión en cierta interpretación de la partícula. El segundo grupo más numeroso con un 30,60 % es la categoría locativa –en dirección ya hacia el valor predominante en la tercera clase.

El segundo clúster que apila ejemplos semiléxicos de la Tabla 1 es el (SL 0) con 100 casos. De nuevo nos hallamos, pero en menor proporción, con expresiones de matiz

modal como en *desviar a manos, conducir a paranoia, correr a antojo, atraer a desaliento, conducir a indiferencia, conducir a colapso, llevar a caos, inclinar a modestia, permanecer a manos o volar a acecho.*

En cambio, en el clúster (SL 1) se encuentran eventos y modos como *apoyar a fuerte, correr a mandato, empujar a exilio, entrar a pacto, extender a matanzas, girar a defensa, guiar a racha o volver a clandestinidad.* Disminuyen la expresión de la modalidad y aumentan los eventos. De hecho, será el clúster semiléxico (SL 2) donde se ensancha considerablemente la expresión de eventos temporales como *acudir a baile, acudir a funeral, acudir a torneo, dirigir a aventura, enviar a audiencia o permanecer a regreso.* Parecería que el algoritmo diferencia en algunos casos muy claramente entre modos y eventos. El carácter heterogéneo o de tránsito de este valor semántico sigue adelante en los dos clusters minoritarios.

En la observación del contacto del grupo semiléxico con el clúster inferior, el (L1 3), se aprecia una coincidencia del 30 % de la expresión locativa.

La primera fase del clúster léxico L1 es, en su mayoría, locativo-espacial. La concentración más alta se observa en la agrupación (L1 1) con 236 casos, un 52,67 % de la fila, y el (L1 3) con 162 casos, un 36,16 %. Los ejemplos son siempre espaciales, desde lugares concretos (edificios, instalaciones o localizaciones humanas y geográficas de todo tipo) como *acelerar a ciudad, arrastrar a centro, andar a iglesia, aproximar a ciudad, caer a afueras, caer a océano, correr a palacio, descender a ultramar, desplazar a aeropuerto, entrar a casa, escapar a desierto, mover a campamento, hasta topónimos como acudir a Barcelona, atraer a Vizcaya, continuar a Europa, dirigir a Nicaragua,* o indicaciones de límite o direcciones como *asentar a oeste, caer a norte, fluir a noreste, partir a sur o remover a centro.*

La cantidad, la variedad y el alcance son muy amplios como se aprecia en el cuadro 13.

Tipo de expresión locativo-espacial	Ejemplos
Nombres comunes de expresión genérica sobre el espacio	Ciudad, aldea, calle...
Topónimos	Estella, Grecia, Colombia, Sevilla, Venezuela, Vigo...
Límites, acceso, orientación y localizaciones genéricas	Costa, oeste, orilla, norte, centro, afuera, ribera...
Instalaciones y construcciones (recintos cerrados o semicerrados)	Cuartel, imprenta, edificio, palacio, campamento, mausoleo, apeadero...
Espacios o zonas abiertas	Desierto...

Cuadro 13. Tipos y ejemplos de expresión espacial en clusters de a (L1 1) y (L1 3)

Los otros dos clusters locativos, (L1 0) y (L1 2), son residuales y no representativos de la predicción. En el 1 aparece alguna excepción, nombres eventivos y con rasgo *Animado*; y en el 2 solo identificamos cinco nombres, y solo dos son locativos, *camarote* e *infierno*.

Por lo que respecta al contexto en la columna, el grupo superior (SL 1) contiene 19 palabras de un total de 65 palabras, es decir, una agrupación por similitud del 29,23 %.

La semántica del tiempo para la preposición *a* se recoge de manera muy distribuida no apareciendo ningún predominio en ningún clúster. Cualquiera de los cuatro clusters presenta sintagmas con indicación temporal de muchos tipos: nombres de mes, momentos, periodos de tiempo o partes del día. La concentración más alta se encuentra en los clusters (L2 0) y (L2 2), con 54 y 59 casos, representando un 68.90 % del total. Algunos ejemplos de este valor léxico serían que *acudir a ayer, caer a final, bajar a meses, correr a temporada, extender a noche, hundir a atardecer, inclinar a comienzo, permanecer a etapa, permanecer a semana, retirar a minutos, salir a mes o trasladar a mayo*.

A modo de ejemplo, veamos algunas de las palabras más frecuentes y algunas de las repeticiones según el clúster: (L2 0), 54 casos y repeticiones de *fin* (8), *momento* (7), *principio* (5) y *23 tiempo*; (L2 1), 33 casos temporales y eventos, y repeticiones de *edad* (13), *inicio* (5) y *llegada* (4); (L2 2), 59 ejemplos, y repeticiones de *amanecer* (5), *atardecer* (4), *comienzo* (6), *diario* (2), *final* (20), *mes* (6), *meses* (3), *semana* (2) y *temporada* (2), y (L2 3), 18 ejemplos, y repeticiones de *hora* (4), *intervalo* (3) y *minuto* (2).

Finalmente mostramos los siguientes gráficos con las distribuciones porcentuales de los valores semánticos de cada uno de los cuatro clusters. Son algunas de las categorías que ya han sido presentadas en las asunciones semánticas: *Animado* (presencia de entidades o partes de seres vivos), *Evento* (sucesos o acontecimientos), *Locativo* (el lugar, emplazamiento o ubicación) y *Temporal* (la duración, su límite o un periodo), más la expresión de modalidad sobre cómo se lleva a cabo la acción verbal.

Como veremos en las conclusiones parciales esta observación doble –clases y clusters- y los contenidos de sus categorías semánticas refuerzan la verificación de la hipótesis.

Gráficos sobre los porcentajes de valores semánticos en clases y agrupaciones de a
(Clústeres 0, 1, 2 y 3, dispuestos de arriba a abajo)

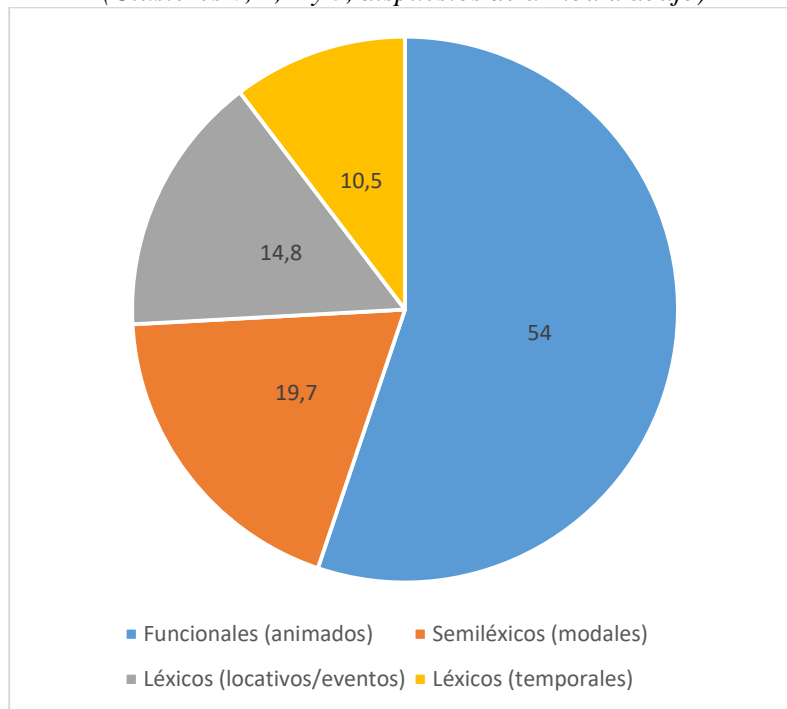


Gráfico 3. Clúster 0 de la a

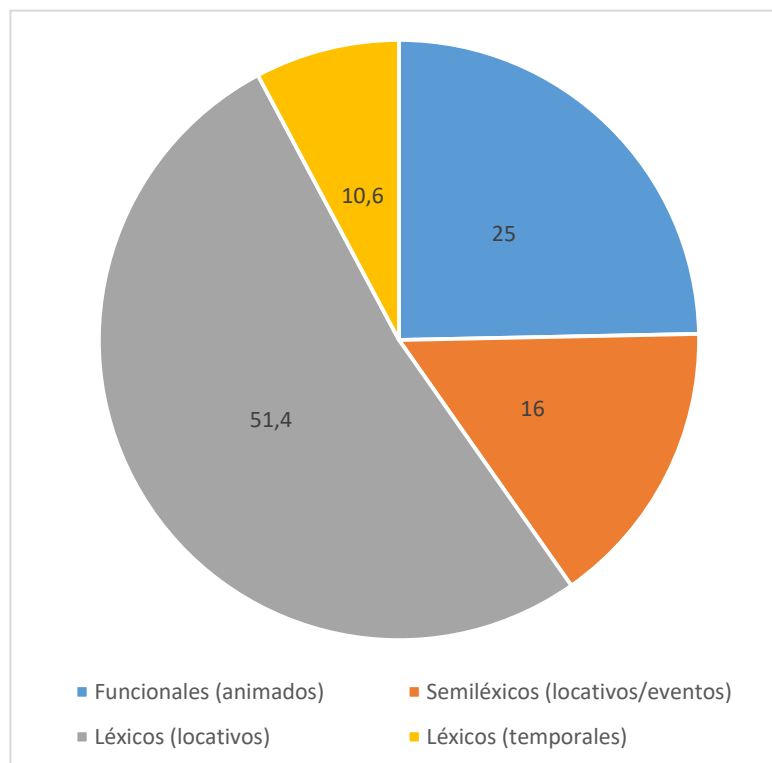


Gráfico 4. Clúster 1 de la a

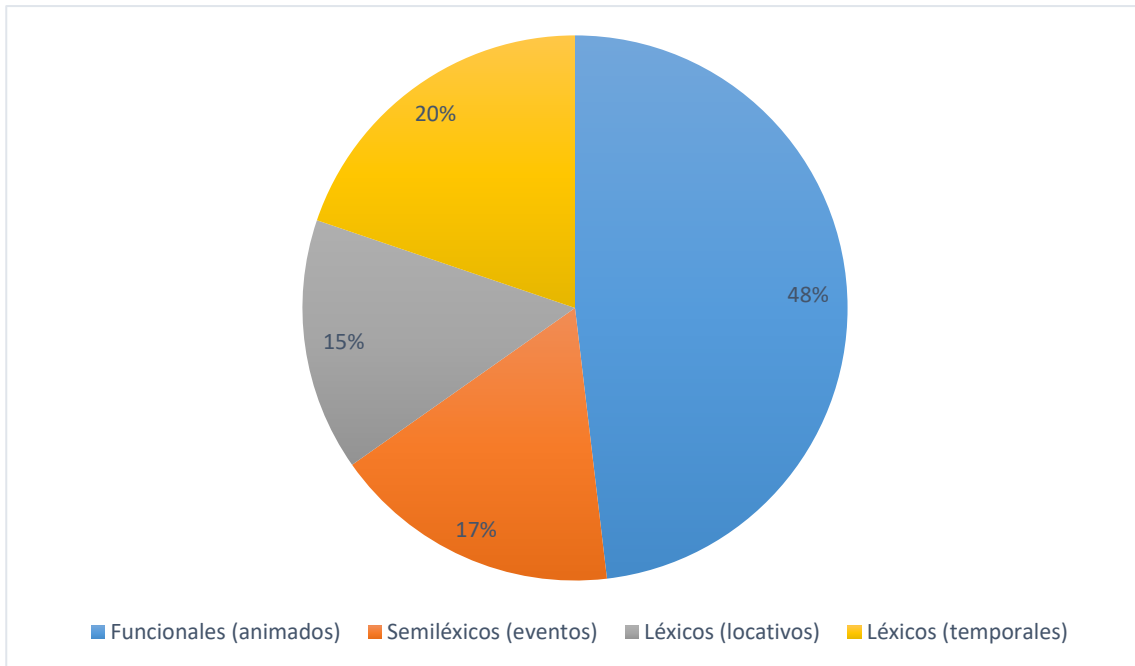


Gráfico 5. Clúster 2 de la a

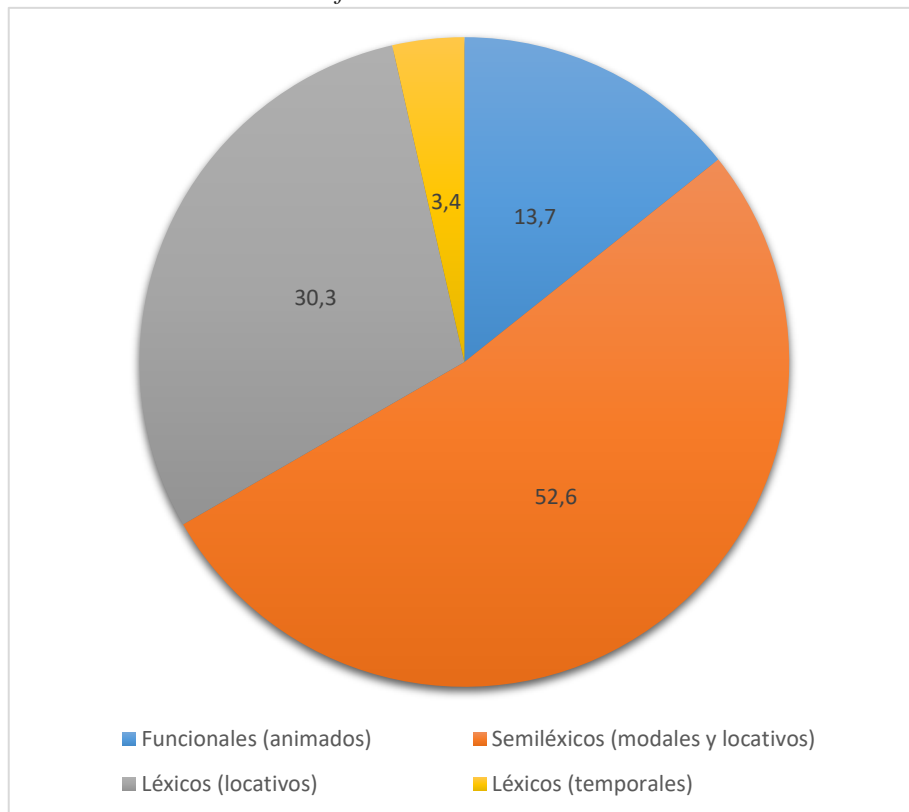


Gráfico 6. Clúster 3 de la a

3.1.6.2 La preposición *hacia*

A diferencia de la preposición *a*, la identificación descriptiva de los significados de *hacia* viene determinada por la propia historia etimológica de la partícula (Portilla, 2011). Se ha recogido el origen romance de esta preposición (*fazia*, *faza*, ‘de cara a’) y como han ido forjándose sus interpretaciones y sus sentidos entorno a la noción de orientación y dirección del sentido real o figurado de la acción, la hegemónica en su universo semántico.

Los resultados que obtenemos de la agrupación en tres clusters y en tres clases preposicionales se organizan en la tabla 3 y avalan, como vamos a comentar, su génesis histórica.

HACIA/DIRECT/ 3 654	2	0	1
SL	117	20	72
L1	169	171	51
L2	33	1	20

Tabla 3. Agrupación en 3 clusters y 3 valores de la preposición *hacia*

El primer valor preposicional, el semiléxico, concentra los sintagmas que denotan beneficiarios y términos de la acción verbal, bien sean *Animado* o *Entidad abstracta*. Estamos ante 117 en el clúster (SL 2), es decir, un 57.92 % del total de la fila. Nos encontramos con ejemplos como *acarrear hacia animales*, *acelerar hacia blanco*, *apoyar hacia gente*, *aproximar hacia contacto*, *apuntar hacia observador*, *apuntar hacia tipo*,

desviar hacia tema, desplazar hacia familia, dirigir hacia interlocutor, detener hacia Dios, girar hacia amo, pasar hacia actores, viajar hacia observador o seguir hacia conocimiento.

Este predominio de la categoría semántica *Animado* (un total de 52 palabras), lo comparte con las que indican *Entidad abstracta*, donde descubrimos hasta 42 posibilidades. Algunas muestras serían *apuntar hacia existencia, continuar hacia leyenda, encauzar hacia restauracionismo, lanzar hacia destino, orientar hacia academicismo, orientar hacia pensamiento, ir hacia socialismo, aproximar hacia objetivo o encaminar hacia dictadura.*

Esta división interna, e incluso una cierta dispersión, se corresponde con una de las aproximaciones de la hipótesis: las clases intermedias o de transición (la semifuncionalidad que vimos en los resultados de la preposición *a*, y ahora la semilexicidad) se diseminan en el clúster de un modo más heterogéneo. Este hecho, como veremos al desarrollar los experimentos siguientes, la presencia de nombres de naturaleza abstracta en las clases intermedias (semifuncionales o semiléxicas) se confirmará desde el análisis entrópico (experimentos 2 y 3) y desde el uso de los hablantes (experimento 4).

En el grupo inferior de la columna, el clúster (L1 2), aparecen indicaciones espaciales genéricas, vagas o abstractas (*abajo, adelante, lado, parte, afuera, fondo, interior, exterior, o arriba*). Esta proximidad con la abstracción del clúster semiléxico es coincidente hasta en un 28,99 % (49 palabras de 169).

La expresión locativa de la preposición, por su parte, se recoge en los clusters (L1 0) y (L1 2). Suponen un 86.95 % del total de la fila. No se aprecian diferencias semánticas substanciales entre los dos grupos. Los ejemplos que aparecen abarcan la mayoría de posibilidades de localización y ubicación: lugares definidos y concretos, espacios abiertos y cerrados, y topónimos. Así tenemos tripletas como *acelerar hacia ciudades, acelerar hacia nordeste, apuntar hacia sierra, atraer hacia Antillas, bajar hacia sur, correr hacia lago, correr hacia tierras, fluir hacia capital, ir hacia avenida, oscilar hacia centro o zarpar hacia Asia*. En cambio, los ejemplos de (L1 2) se concentran mayoritariamente en situaciones espaciales inespecíficas, direcciones y orientaciones. De esta forma tenemos *acelerar hacia abajo, agitar hacia adelante, agitar hacia lados, alejar hacia mitad,*

apuntar hacia afuera, apuntar hacia extremo, arrastrar hacia afuera, bajar hacia interior, deslizar hacia izquierda, enderezar hacia costado, orientar hacia camino y sobrevolar hacia este.

En el cuadro 14, en forma de sinopsis organizada, se muestran algunos ejemplos contenidos en los clusters locativos.

Tipos de expresión locativo espacial	Ejemplos
Topónimos	Ebro, Asia, Urales, Amazonas...
Límites, acceso, orientación y localizaciones genéricas	Litoral, norte, barlovento, salida, franja, sur, oeste, oriente...
Instalaciones y construcciones (espacios funcionales)	Canal, cementerio, puerto, barracas, iglesia, puente, aeropuerto...
Espacios o zonas abiertas y naturales	Mar, cumbre, calle, sierra, playa, aguas, isla...

Cuadro 14. Tipos y ejemplos de nombres de expresión espacial en clusters de hacia

En el segundo clúster locativo (L1 2) se reúnen 169 casos. La diferencia ahora es la presencia de localizaciones espaciales genéricas como *arriba, abajo, adelante, fondo, o direccionales como extremo o izquierda*; y locativos direccionales, figurados o abstractos, tales como *parte, abismo, posiciones, proximidades, fondo y fuera*. De manera residual, localizamos siete palabras con rasgo *Animado* como *rodilla, cabeza, laringe, ombligo, hembra, abdomen o cerebro*.

A su vez, en el (L1 1) reaparecen los topónimos hasta en 30 casos con nombres como *Ecuador, Ibiza, Finlandia, Madrid, Egipto, Trujillo, Asturias o Cuba*.

Por lo que respecta a los valores semánticos de tiempo, la mayoría de casos, 33 de 54 (61,11 % de la fila), se encuentran en el clúster (L2 2). Los ejemplos de valores temporales con un sentido abstracto tales como *acelerar hacia final, andar hacia fase, aproximar hacia mientras, apuntar hacia eternidad, apuntar hacia futuro, caminar hacia vejez, conducir hacia fin, dirigir hacia proceso, llevar hacia expectativas o venir hacia presente*.

Es interesante observar que, en la agrupación temporal, y quizás debido a ese carácter abstracto, el grupo más numeroso se encuentra en la misma columna de los semiléxicos donde encontrábamos hasta 42 palabras, un 36,52% del total, clasificables como entidades abstractas. El algoritmo de CLUTO prefiere dejarlo en la misma columna por su similitud con el clúster superior (L1 2). Este hecho podría explicar que en la fila de la temporalidad (L2) el mayoritario no sea el 1.

En cambio, en el clúster (L2 1) se produce una concentración de la palabra *finales*, en 17 casos, y solo tres palabras nuevas: *retiro*, *septiembre* y *conquista*. No se puede considerar representativa cualitativamente.

En los tres gráficos sucesivos resumimos los predominios de las categorías semánticas internas a cada clúster.

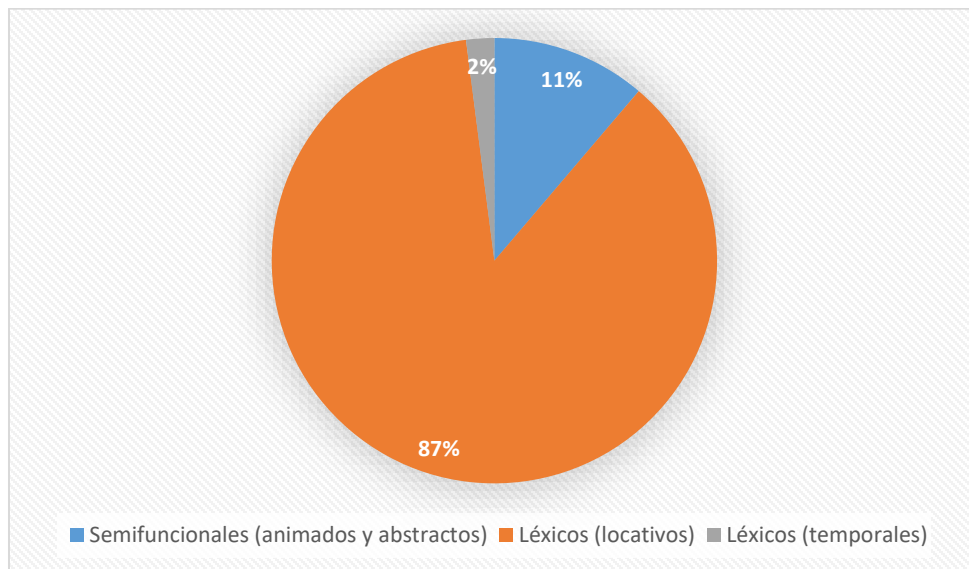


Gráfico 7. Porcentajes de valores semánticos del clúster 0 de hacia

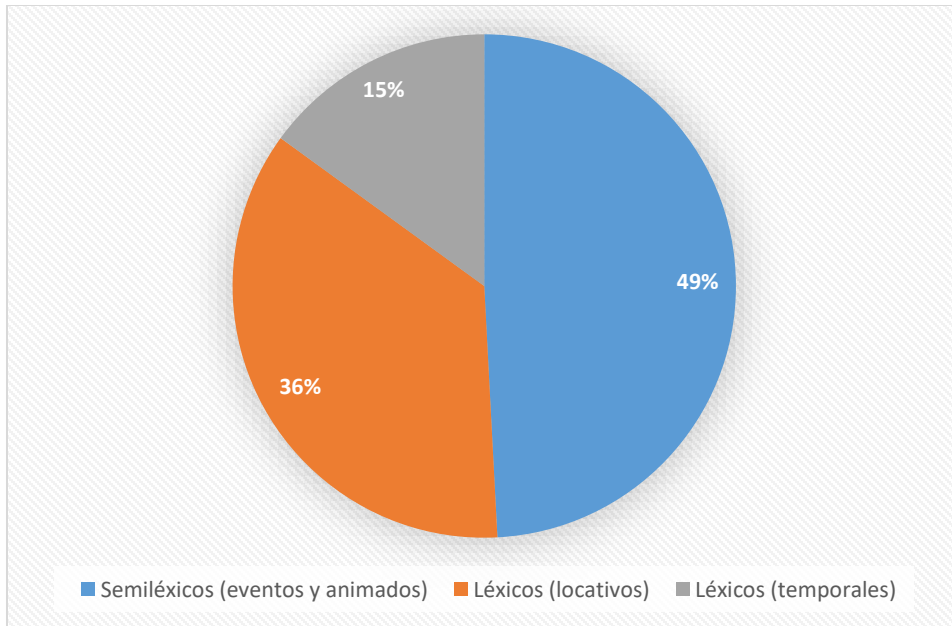


Gráfico 8. Porcentajes de valores semánticos del clúster 1 de hacia

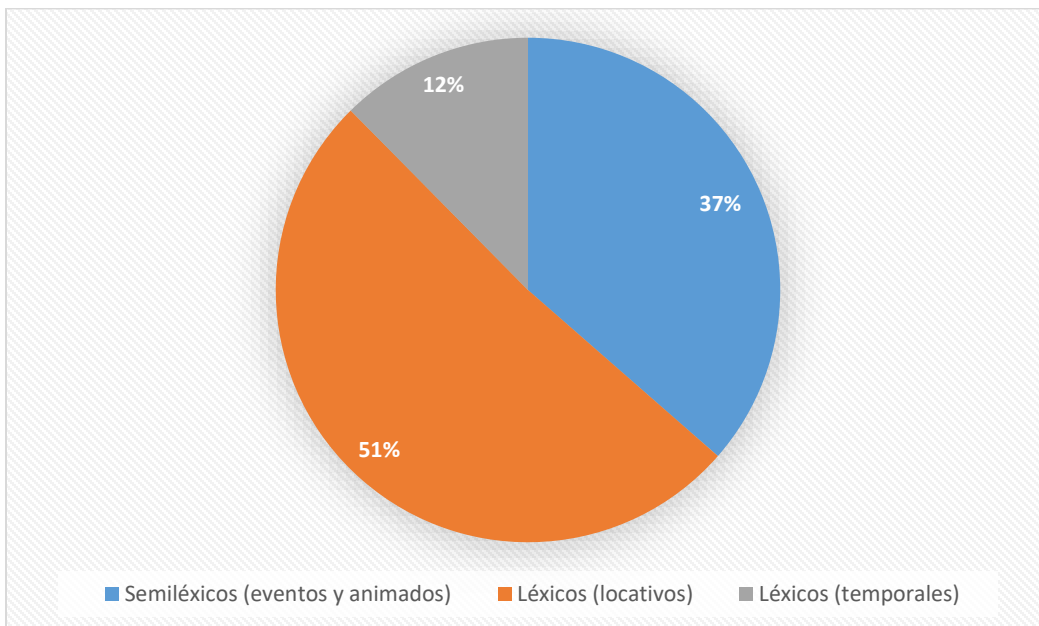


Gráfico 9. Porcentajes de valores semánticos del clúster 2 de hacia

3.1.6.3 La preposición *hasta*

Hasta es la preposición que obtiene los mejores resultados en relación al éxito de la agrupación automática y la anotación manual humana por lo que respecta a la concentración. Nos hallamos ante un 64,6 % de pureza de promedio.

Este hecho, seguramente, guarda relación con la propia naturaleza de su significado. Se trata de la expresión de un límite, en cualquiera de sus sentidos: de una acción, de un beneficiario, de un evento o suceso, de un movimiento en el espacio (valores locativos), y de un movimiento en el tiempo (valores temporales). Las descripciones sobre su historia y sus usos sincrónicos han sido profusamente comentadas (Morera, 1999 y 2016).

Pasemos al análisis de la estructura interna de los grupos a partir de los datos numéricos organizados en la tabla 4

HASTA/DIRECT/3 778	1	2	0
SL	144	22	29
L1	28	240	41
L2	129	36	109

Tabla 4. Agrupación en 3 clusters y 3 valores de la preposición *hasta*

Los datos del primer grupo, el semiléxico, son bastante reveladores, por la cantidad y por la calidad de sus evidencias. Se reúnen en el clúster (SL 1), siendo 144 casos y representando un 73,84 % de pureza de promedio. Los ejemplos son de dos tipos. El mayoritario, con 97 casos, expresa un evento complejo con un límite como *alejarse hasta*

haber, acelerar hasta alcanzar, alejar hasta terminar, asentar hasta ofrecer, caminar hasta agoten, continuar hasta situar, enfocar hasta trabajar, girar hasta conseguir, orientar hasta alcanzar, pasar hasta acabar, pasear hasta llegar, subir hasta igualar, tirar hasta llegar o volar hasta llegar. En el segundo grupo se aglutinan los nombres de categoría semántica *Animado*, siendo un total de 22 casos. Los ejemplos serían *atraer hasta espectadores, caer hasta consumidores, contener hasta animales, correr hasta personajes, entrar hasta dioses, enviar hasta adjuntos, enviar hasta condenados, extender hasta hombres, extender hasta rebeldes, ir hasta criadores, llevar hasta asesino, llevar hasta dioses, juntar hasta personas y viajar hasta receptor.* La gramática descriptiva considera este último caso como adverbial. A efectos de nuestra interpretación no resulta pertinente su conducta o adscripción categorial o sintáctica en la medida en que continúa confirmando la agrupación en virtud de su distribución semántica.

Es interesante comprobar cómo, en ambos ejemplos, la preposición está expresando el límite y el término, sea de un acontecimiento expresados por los dos verbos, o sea de los participantes en esa acción. Los dos grupos semiléxicos, (SL 0) y (SL 2), respectivamente, no aportan información valiosa en relación ni a la gradualidad ni a la organización interna.

La similitud dentro de la columna 1 la encontramos en el clúster (L2 1) donde el predominio de la expresión temporal es casi completo. El hecho de que los nombres eventivos supongan contacto en la parte inferior con el clúster de la columna (L1 1) confirma la continuidad entre ambos grupos.

En la segunda fase de significados, identificada en el clúster (L1 2), se reúnen las palabras que indican la localización en el espacio. El porcentaje de coincidencia es considerable con un total de 240 casos (77,66 %) con el rasgo de *Locativo*. Salvo *batalla, antebrazo, brazos, vientre, distancia, trenes y bolsa*. El resto expresan la ubicación, la localización y otros elementos lingüísticos del espacio. En el cuadro 15 se advierte la diversidad de recursos expresivos relacionados con el espacio.

Tipo de expresión locativo-espacial	Ejemplos
Nombres comunes de expresión genérica sobre el espacio	Lugar, poblado, provincia, localidad, municipio, capital...
Topónimos	Ponferrada, Martorell, Terranova, Duero, Arabia, Vinaroz, Motril...
Límites, acceso, orientación y localizaciones genéricas	Fondo, sur, centro, punto, interior, frontera, entrada, puerta, emplazamiento, desembocadura, margen...
Instalaciones y construcciones (espacios cerrados)	Santuario, muelle, pantano, iglesia, muro, fortaleza, albergue, casa, tumba...
Espacios o zonas abiertas	Mar, llanura, valle, fosa, plaza, manantial...

Cuadro 15. Expresión de la ubicación y la localización en el espacio del clúster léxico de hasta

En el siguiente clúster de carácter léxico, el (L1 0), se concentran, de manera casi absoluta (37 de 40), topónimos como *acudir hasta China*, *circular hasta California*, *continuar hasta Marsella*, *extender hasta Pensilvania*, *extender hasta Italia*, *ir hasta Jamaica*, *pasar hasta Vietnam*, *remolcar hasta Sevilla* o *viajar hasta Ucrania*.

Por último, la temporalidad se organiza en el clúster (L2 1), siendo el más alto con 129 casos, un 47,08 % de la fila. En él se concentran triplas que indican las fases de la secuencia temporal como *acudir hasta momento*, *agitar hasta veces*, *caer hasta final*, *continuar hasta comienzo*, *girar hasta ahora*, *permanecer hasta episodio*, o *salir hasta pronto*. El segundo clúster de triplas temporales es el (L2 0), que reúne mayoritariamente el nombre de meses, estaciones del año y algunos temporales genéricos. Los ejemplos serían *apoyar hasta enero*, *conducir hasta setiembre*, *continuar hasta periodo*, *correr hasta temporada*, *dirigir hasta principios*, *doblar hasta fecha*, *llevar hasta octubre* y *volver hasta semana*. En consonancia con esa homogeneidad, el tercer clúster léxico-temporal (L2 2), también incluye nuevas palabras sobre el desarrollo puntos, fases, momentos o periodos: *hoy*, *mediados*, *actualidad*, *siglo*, *madrugada* o *invierno*.

De nuevo nos encontramos, como con *hacia*, que la clase más numerosa aparece en la columna 1 (en lugar de la 0, la que correspondería con la diagonal). La causa sería

semejante al caso anterior. Es probable que la distribución semántica coincida en la medida en que el carácter abierto, inespecífico o abstracto de la expresión de la temporalidad se asocia con expresiones locativas también de carácter general. Este rasgo sería una posible explicación de porqué está junto a los locativos genéricos o abiertos que aparecen en el clúster superior, el (L1 1). Aquí topamos con ejemplos como *alejarse hasta entonces*, *apoyarse hasta niveles*, *asentar hasta descenso*, *desviar hasta nivel*, *continuar hasta salida*, *guiar hasta destino* y *galopar hasta meta*.

Recogemos la distribución interna de los clusters para la preposición *hasta* en los gráficos que siguen. Como en el caso anterior, en las conclusiones parciales indicaremos la oportunidad de esta información en relación a la certidumbre de la hipótesis.

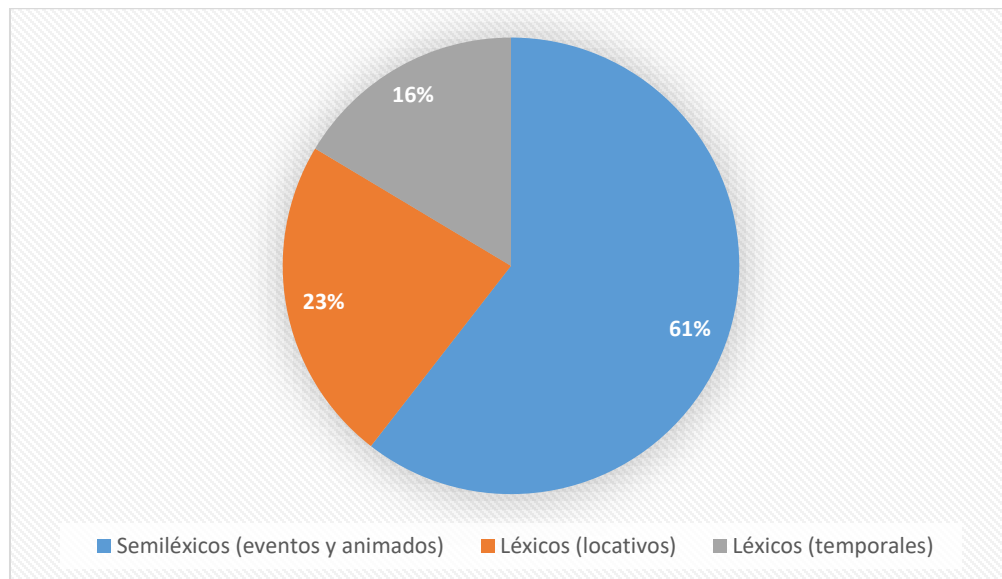


Gráfico 10. Porcentajes de valores semánticos del clúster 0 de hasta

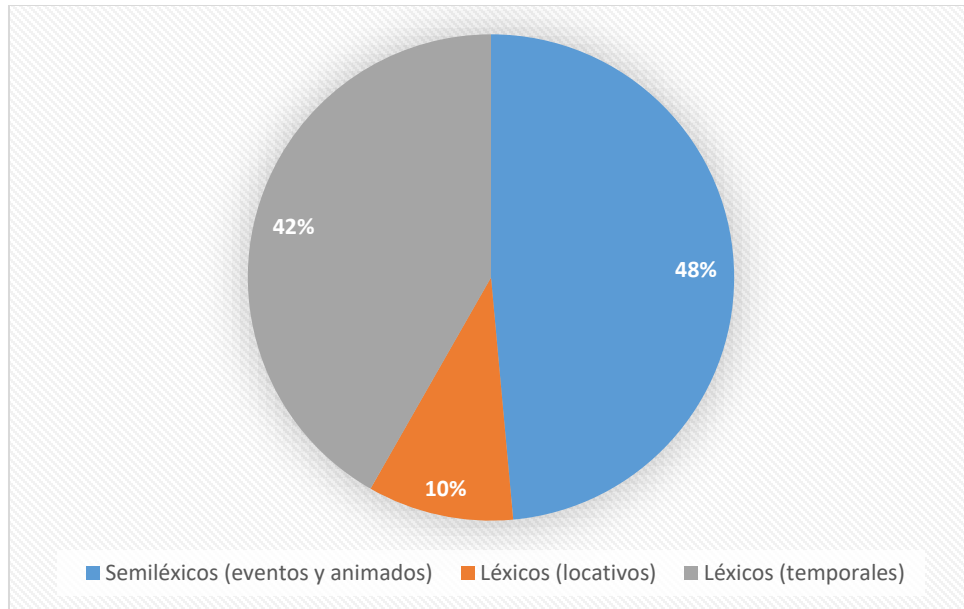


Gráfico 11. Porcentajes de valores semánticos del clúster 1 de hasta

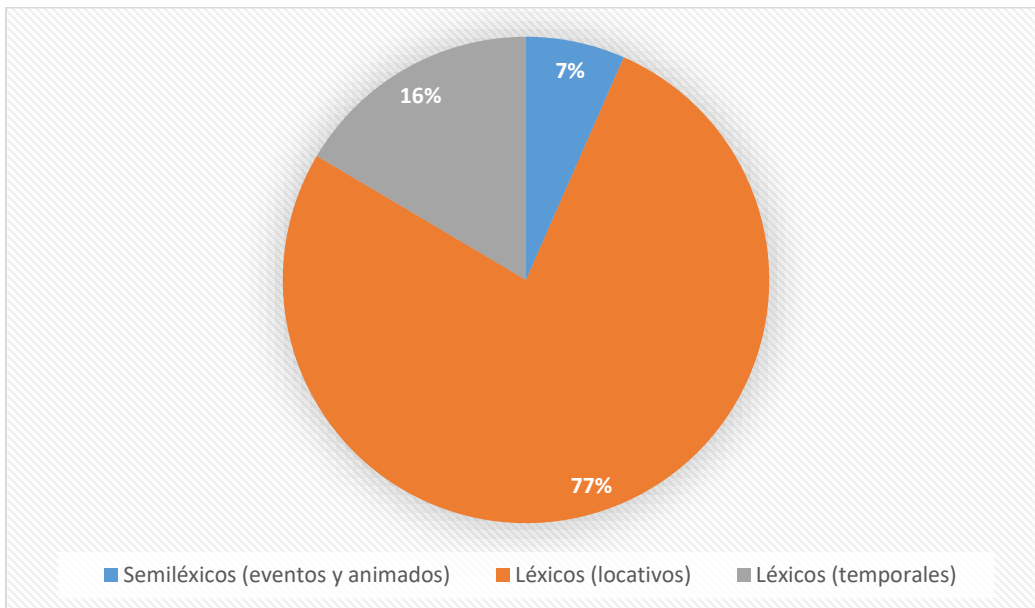


Gráfico 12. Porcentajes de valores semánticos del clúster 2 de hasta

3.1.7 Conclusiones parciales

Los resultados de la agrupación automática de las tres preposiciones recogen algunos aspectos que son parcialmente orientadores respecto de la gradualidad semántica de los valores preposicionales, y que dan apoyo y son consistentes con la hipótesis: el significado de las piezas preposicionales se produce de manera progresiva, paulatina y distribucional.

Conviene matizar, también, que esos resultados solo capturan información semántica que es relevante para nuestra hipótesis pero no la explica de manera definitiva. Existen, como vamos a ver en los siguientes experimentos, otros factores que intervienen y que han quedado fuera del alcance del experimento.

El primero de los aspectos positivos es que la concentración obtenida en las agrupaciones confirma que la agrupación realizada por el algoritmo usando exclusivamente la información semántica obtenida mediante las *Word Embeddings* coincide en gran medida –por encima del 50 %- con la clasificación humana, demostrando que la semántica explica en gran parte la función de la preposición. En una observación pormenorizada en cada una de las preposiciones, en sus filas y sus columnas, sigue siendo exitoso en 32 clusters de 34, con porcentajes variables.

El segundo elemento avala que sea la distribución contextual –la clase de palabra que se halla en la vecindad de la preposición- donde se encuentre su nivel de influencia semántica o peso en el significado (funcional, semifuncional, semiléxico o léxico). Hemos visto que analizados los 10 clusters más representativos de las clases propuestas, en 7 de sus clusters adyacentes de la columna el porcentaje de coincidencia o graduación es del 30 %.

La tercera y última consideración se revela a partir del análisis de la composición de las categorías semánticas de los clusters, recogidos en los gráficos finales de los resultados.

De este modo, la funcionalidad de la *a* reúne contextos donde su acción interpretativa es nula y los resultados de los porcentajes son sustanciales, concentrándose en el clúster 0, con un 54 % de palabras de categoría *Animado*. El clúster que aglutina las clases semiléxicas, el 3, aún un 52 % de la categoría *Locativo* más toda la expresión de la modalidad. Se empieza a observar un aumento de la diversidad de las clases que, además, se aproximará con el siguiente grupo. Este clúster, el 1, precisamente, es el que concentra un 51 % de la categoría *Locativo*. El 2, el último clúster es el más heterogéneo con un 47 % de *Animado* y un 19 % de *Temporal*.

En *hacia* se aprecia que el predominio semántico interno se produce con la categoría semántica *Locativo*, concentrado en el clúster 0, con un 77 % de palabras de ese tipo. Los dos otros dos clusters reparten su predominancia. El clúster 2 entre *Locativo* y *Animado* (con un 51 % y 37 %, respectivamente), y el clúster 1 entre *Animado* y *Evento* (49% para ambos) y *Locativo* (36%). Se produce una especie de evolución de la expresión locativa que, de manera gradual, comparte espacio con otros valores.

La preposición *hasta*, a su vez, acusa una conducta semántica muy perfilada en las distribuciones de la temporalidad. En el clúster 1 obtiene un 42 % que alterna con un 48 % de *Animado* y *Evento*, con infinitivos verbales. A su vez, en el clúster 0 continúan *Evento* y *Animado* pero llegando hasta el 61 %. Y, por último, el espacio se expresa en el clúster 2 con un 77 % de *Locativo*.

Desde el punto de vista metodológico de nuestra investigación se confirma el poder analítico de las denominadas *Word Embeddings*, vectores de valores obtenidos por similitud de aparición. No solo por la evidencia cuantitativa ya detallada sino también porque se están orillando las dificultades teóricas que se atribuyeron a las prescripciones categoriales de la preposición, en particular cómo determinar su contribución semántica en los enunciados a través de justificaciones sintácticas o estructurales. El cálculo de las *Word Embeddings* se realiza teniendo en cuenta las formas de las palabras (no los lemas) pero sin distinguir los sentidos en el caso de la polisemia. Esta contingencia ha limitado el alcance de la experimentación propuesta.

Como síntesis, podríamos decir que la distribución que se obtiene por similitud distributiva y contextual, y a través de aprendizaje automático, responde a patrones

semánticos y no al tipo sintáctico o estructural que aparece en la que participa. En los experimentos 2 y 3 abordamos, desde otra herramienta de cálculo estadístico, la entropía, como profundizar en la hipótesis de la gradualidad a partir de la caracterización de esos patrones semánticos en nuevos contextos.

3.2 SEGUNDO EXPERIMENTO. Entropía y significado preposicional en el ámbito verbal

3.2.1 Objetivos y justificación del experimento

En el segundo experimento continuamos con el objetivo central de la hipótesis: la justificación empírica de la identificación de la gradualidad semántica para las preposiciones en español.

El propósito de este experimento es medir la entropía de un conjunto de nombres (y las categorías semánticas a las que pertenecen según nuestras asunciones), insertos en tripletas que son complementos verbales de diverso tipo y naturaleza sintáctica.

La medición de esa entropía, es decir, la capacidad de predecir la aparición de una categoría semántica, va a permitir correlacionar la distribución de esas categorías con la aportación semántica preposicional de cada una de las clases propuestas: funcional, semifuncional o léxico. Como examinaremos, el resultado entrópico nos permitirá confirmar que a menor entropía en el conjunto de sintagmas nominales menor contribución de significado por parte de la preposición, y viceversa.

A diferencia del primer experimento (3.1), donde la agrupación automática se llevó a cabo en un universo semántico, el espacial, ahora no restringimos ni el alcance de la expresión semántica de los verbos, ni de las preposiciones (participan la mayoría de las propuestas en la hipótesis).

3.2.2 Las clases de la gradualidad semántica preposicional

En el desarrollo del primer experimento, y dentro de la metodología del aprendizaje automático, consideramos que en la agrupación de los valores léxico-semánticos podían existir diversos grupos, para el caso particular de la *a*, atendiendo a la calidad de los resultados. Metodológicamente resultó una decisión apropiada para la detección de las condiciones distributivas que proponemos.

Ahora, en cambio, establecemos que se dan tres clases de gradualidad semántica: la funcional, la semifuncional y la léxica.

Así, en el cuadro 16, identificamos las clases propuestas en el conjunto global de la investigación. Como ya señalamos en el primer experimento, esta fluctuación tiene un fundamento empírico, es decir, son los valores asociados a los datos lingüísticos los que empujan esta disposición.

CUADRO SOBRE LA ORGANIZACIÓN FLEXIBLE DE LAS CLASES PREPOSICIONALES						
Experimentos	Clases preposicionales	Funcional	Semifuncional	Semiléxica	Léxica	
Primer experimento (preposiciones <i>a</i> , <i>hacia</i> y <i>hasta</i>)					<i>L1</i>	<i>L2</i>
Segundo experimento Estructuras en el ámbito verbal (todas las preposiciones)						
Tercer experimento Estructuras en el ámbito nominal (todas las preposiciones)						
Cuarto experimento (todas las preposiciones)						

Cuadro 16. Organización de las clases preposicionales por experimentos.

En este experimento la clase intermedia sugerida es la semifuncional. La razón procede de la especie de complementos verbales que la componen: los de régimen verbal. Como ya indicamos en la historia gramatical –Capítulo 1, sección 1.4 (Alarcos, 1968 y 1990; Martínez, 1983 y 1986; Gutiérrez Ordóñez, 1996; Cuesta, 1997; y Tornel, 2006a y 2006b)- la relación semántica entre los verbos de régimen y sus nombres están mediatizada u orientada solo superficialmente por la preposición. Este arbitraje formal apenas acomoda el significado sin contribuir a la interpretación léxica global, a diferencia de lo que sucedía en las semiléxicas de las preposiciones *a*, *hacia* y *hasta* del primer experimento donde se percibían valores espaciales o temporales figurados, Capítulo 3, sección 3.1.2, cuadros 8, 9 y 10, páginas 99 y 100.

3.2.3. Selección y procedencia de los datos

Los criterios semánticos para la selección de los verbos abarcan una variedad considerable, tanto desde el punto de vista del aspecto o la modalidad de la acción (estado, eventos, realizaciones o logros) como desde el punto de vista del contenido léxico-semántico (psicológicos, de habla, de movimiento, de acción...).

Su naturaleza sintáctico-argumental cubre los tres tipos posibles en la predicación verbal del español, recogido en (91 a-c).

(91)

(a) Argumento o complementos directos e indirectos (sintagmas preposicionales funcionales): *enviaron al mensajero, compraron al árbitro, regalaron libros a la maestra, o mandaron las cartas a Pedro.*

(b) Argumentos de régimen verbal (sintagmas preposicionales semifuncionales): *trataron sobre sociología, abogan por el diálogo, inciden en la negociación o han acabado con la expectativa.*

(c) Adjuntos o complementos circunstanciales (sintagmas preposicionales léxicos): *actúan con rapidez, se ofrecen para la compra, beben sin medida o se distrajeron durante la ceremonia.*

Nuestra selección de verbos ha seguido la clasificación del proyecto ADESSE de la Universidad de Vigo para el español. El propósito ha sido cobijar, con una representatividad consistente de significado, la mayor parte de representaciones semánticas posibles de la extensísima categoría verbal y sus clasificaciones (Báez San José, 1990; García-Miguel y Albertuz, 2005; Albertuz, 2007; y Castellón et alia, 2015).

En el cuadro 17, procedente de (<http://adesse.uvigo.es/data/clases.php>), reseñamos la taxonomía de referencia junto a algún ejemplo de muestra.

Macroclase	Clase	Subclase	Ejemplos
Proceso mental	Mental	General	Distraer...
	Sensación	Volición	Querer...
	Percepción	General	Ver...
	Cognición	Conocimiento	Pensar, entender...
	Elección	Creencia	Saber...
	Atribución	Relación	Decidir, elegir...
Posesión		Propiedad	Estar...
		Medida	Representar...
		Apariencia	Valer...
		Denominación	Oler...
		Cambio de estado	Llamar, nombrar...
		Adquisición	Ponerse...
		Transferencia	Necesitar...
	Proceso material	Espacio	
			Dar, pagar...
			Mover...

	Cambio	Desplazamiento Localización Postura-posición Orientación Manera Unión Creación Modificación Cuidado Destrucción Meteorología Emisión Contacto Control Uso Consumo Sustitución Actividad Competición	Abalanzar... Abanderar... Levantar, extender... Orientar... Girar, sacudir... Añadir... Apuntar... Cambiar... Duchar... Eliminar... Llover... Brillar... Tocar... Agarrar, coger... Usar... Consumir... Relevar... Trabajar... Luchar... Burlar... Dormir... Comer, beber... Casar...
	Otros hechos		
	Comportamiento	Fisiología Ingestión Relaciones sociales	
Proceso verbal	Comunicación Valoración Emisión de sonido	General Petición	Decir... Mandar, pedir... Acusar... Cantar...
Proceso existencial	Existencia Fase-tiempo Vida	Fase Tiempo	Surgir... Continuar... Persistir... Morir...
Modulación	Causación Disposición Aceptación Verbos de apoyo	Inducción Permiso Obligación	Comprometer... Dejar... Obligar... Tratar... Aceptar... Dar, hacer...

Cuadro 17. Clasificación de verbos procedente de <http://adesse.uvigo.es/data/clases.php>

Con los descriptores semánticos anteriores de Adesse se extraen 3898 tripletas de estos verbos. A partir de aquí se seleccionan los nombres que aparecen en posición de complementos. Proceden de su coaparición con alguno de los 319 verbos de (92).

(92)

abalanzar, abanicar, abanderar, abandonar, abarcar, abastecer, abogar, abordar, abrazar, abrir, abusar, acarrear, acceder, aceptar, acoger, acompañar, actuar, acudir, acusar, adelantar, adentrar, advertir, aficionar, afiliarse, agarrar, aguantar, alimentar, almacenar, amar, amparar, andar, añadir, aparcar, aparear, apelar, aplicar, apostar, apoyar, aprender, apuntar, arriesgar, asociar, asustar, atacar, atender, aterrizar, atrever, atribuir, avisar, ayudar, bailar, beber, besar, borrar, brillar, burlarse, buscar, calibrar, calificar, caminar, campar, cansar, cantar, caracterizar, carecer, casar, ceder, cejar, cesar, citar, coincidir, colocar, combatir, comer, comulgar, concentrar, condenar, conducir, confrontar, consentir, conservar, consistir, contar, contestar, continuar, contrastar, convencer, conversar, convertir, convidar, convivir, convocar, corresponder, crear, creer, dañar, dar, decidir, decir, declarar, definir, degenerar, deleitar, delimitar, depender, depositar, derivar, derrotar, desarrollar, descansar, descender, desconfiar, descubrir, desembocar, destacar, desviar, diferenciar, diferir, dirigir, discurrir, digerir, discrepar, discutir, disenter, diseñar, disertar, disfrutar, disparar, disponer, distraer, dividir, dominar, dormir, duchar, elegir, eliminar, elevar, empeñarse, empujar, encargar, emigrar, encerrar, encontrar, enfrentar, engullir, entender, enterar, entrar, entregar, entrenar, enviar, escoltar, escribir, escuchar, estar, estimar, estudiar, exceder, excluir, exculpar, existir, extender, exterminar, facilitar, fascinar, formar, fundar, ganar, generar, girar, gozar, hablar, hilar, herir, huir, inclinar, incorporar, incurrir, indicar, infectar, influir, intimidar, integrar, interesar, insistir, instar, invitar, jactar, jugar, lastimar, leer, levantar, limpiar, llamar, llegar, llevar, luchar, madurar, mandar, mantener, maravillarse, marcar, medir, meditar, mencionar, mirar, molestar, morir, multar, mostrar, nacer, nadar, necesitar, nombrar, obedecer, obligar, obtener, ocultar, odiar, ofender, ofrecer, oír, oler, organizar, orientar, otorgar, padecer, pagar, parar, participar, pasar, pecar, pedir, pegar, pelear, pensar, perder, permanecer, persistir, planchar, poner, precisar, preguntar, prescindir, preocupar, preparar, prevalecer, procurar, prolongar, proponer, proporcionar, proseguir, proteger, proveer, prevenir, proyectar, pugnar, realizar, recaer, recibir, recordar, recriminar, recurrir, reflexionar, regir, relevar, renegar, renunciar, reparar, repartir, representar, requerir, responder, restituir, robar, romper, saber, sacar, sacrificar, sacudir, saltar, saludar, seleccionar, seguir, sepultar, sobresaltarse, soltar, sospechar, subir, sugerir, sumar, superar, surgir, tachar, terminar, tocar, tomar, topar, trabajar, traer, tratar, ufanarse, usar, utilizar, valer, velar, venir, ver, vivir, volcar, volver y votar.

Por lo que respecta a las tres clases preposicionales, en (93a-d) identificamos los ejemplos gramaticales arquetípicos de la funcionalidad; descritos como complementos directos de persona (a), complementos indirectos (b), las perífrasis verbales (c) o las locuciones prepositivas (d).

(93)

(a) Envía *a* María

(b) Traen medicinas *a* los afectados

(c) Echaron *a* correr

(d) Se recauda *en* beneficio *de* una ONG

En estos casos la aportación semántica descrita por algunas de las referencias tradicionales como (Fernández, 1999; Bosque y Demonte, 1999; *NGLE*, 2009; y Moliner, 2012) es muy baja o nula. La preposición es un mero índice que apenas contribuye a la interpretación del predicado. En nuestro corpus hemos excluido la presencia de locuciones que, según creemos, requieren un tratamiento diferenciado respecto de sus formas de gramaticalización y/o productividad (Kazumi, 1997).

A continuación presentamos los sintagmas preposicionales que reflejan una situación intermedia desde la perspectiva de la aportación semántica, y que denominamos semifuncionales. Para ello tomamos como punto de partida los verbos de régimen o preposicionales, es decir, aquellos verbos que exigen, mandan o condicionan una preposición, y cuya interpretación semántica está más cohesionada entre ambos términos, el núcleo verbal y el núcleo nominal (Alcoba, 2000; del Barrio, 2004; Candalija, 2006; y Jiménez, 2011). Tengamos, entonces, los ejemplos de (94a-d).

(94)

- (a) El libro trata *de* los problemas actuales
- (b) La comida consistió *en* dos primeros y dos segundos
- (c) Ha accedido *a* la cátedra sin dificultad
- (d) Se preguntaban *por* la desaparición de Laura

Siguiendo con el incremento en la aportación semántica nos hallamos ante las preposiciones *léxicas* en los casos de (95a-h).

(95)

- (a) Depositó el libro *sobre* la mesa
- (b) Iremos *a* Blanes *de* vacaciones
- (c) Quedamos *a* las seis
- (d) Vivimos *en* Sevilla *con* nuestro sobrino
- (e) Murieron *por* el bombardeo
- (f) Escribe *con* cuidado
- (g) Se enamoraron *durante* unas vacaciones
- (h) Corrieron *tras* el ladrón

En estos sintagmas, los valores de las preposiciones *sobre, a, de, en, por, con, durante* o *tras* ayudan a la especificación semántica del verbo y de todo el predicado, tal y como ya se apuntaba, desde una observación descriptiva (Cano, 1981) o desde la visión de la teoría generativista (Demonte, 1992).

Como vamos a intentar demostrar, los ejemplos entre (a) y (h) representan una gradación medible a través de la naturaleza semántica de los nombres de su contexto.

Hemos procurado, también, que la representatividad de las frecuencias esté garantizada. Las cantidades de apariciones de las siete preposiciones más frecuentes se han puesto en relación con las frecuencias totales de las preposiciones en tres corpus, como se observa en la tabla comparativa 5: el CREA de la RAE, el Corpus del español (o Corpus

Davies) y el corpus ADIBA. Este último es una base de datos procedente de ejemplos de Twitter. Su acrónimo es *Atles dialectal de Barcelona*. Hemos consultado un conjunto de archivos para establecer también esta medida de frecuencia de aparición de preposiciones en los tuits producidos en Barcelona y su área metropolitana, <http://adiba.cat/>, con el propósito de recoger usos preposicionales actuales.

Estas siete preposiciones, las más frecuentes, representan un 75 % del total de casos utilizados en el experimento (2923 de 3898).

En la columna de la derecha (5) se indican los datos de nuestro experimento, el orden en el ránking y el número ocurrencias que aparecen en nuestro trabajo. Las otras columnas (2, 3 y 4) indican el orden en cada uno de los corpus.

Preposición	CREA (RAE)	Corpus del español (Davies)	ADIBA (Atles dialectal de Barcelona)	Experimento 2
De	1	1	1	3 (643)
A	3	2	2	1 (1371)
En	2	3	3	2 (766)
Por	4	5	4	5 (164)
Con	5	4	5	4 (345)
Para	6	6	6	6 (138)
Sobre	8	7	7	7 (96)

Tabla 5. Comparación de frecuencias de aparición de las siete preposiciones más comunes en el experimento con tres corpus representativos

La observación de las coincidencias y de las diferencias en esa tabla de frecuencias está condicionada, sin lugar a dudas, por las diferencias entre las cantidades absolutas que se comparan (Alonso Ramos, 2012). Pasamos de los millones de casos que se contienen en estos corpus a los 3987 de nuestro experimento. A pesar de ello, en cinco de siete preposiciones *en*, *por*, *con*, *para* y *sobre*, o se coincide con los tres corpus o se ocasiona un desacuerdo de una sola posición (arriba o abajo). Por otra parte, nuestros ejemplos solo proceden del ámbito verbal. Este hecho explicaría la eminencia de la *a*, y el número

relativamente bajo para la *de*, que suele ser más común y frecuente en su complementación nominal (como veremos el experimento tercer, sección 3.3.5.1.1).

Por lo que respecta a la procedencia de los datos, indicaremos que la selección se ha realizado a través de una búsqueda arbitraria –sin criterio temporal ni espacial de su procedencia- de tres corpus, de dimensiones y de reconocimiento académico internacional tanto en sus volúmenes de documentos como sus posibilidades de búsqueda, de análisis y de manejo de los datos. Estos corpus han sido referenciados y citados en el Capítulo 2, sección 2.4.4.1, página 92. Hemos elegido de manera mayoritaria tripletas del Wikicorpus (2.310); en segundo lugar el Corpus Davies, 992 tripletas más; y el resto, 596 del CREA de la RAE.

3.2.4 Metodología: el cálculo de la entropía y las categorías semánticas

El concepto de entropía ha resultado controvertido y altamente polisémico en el ámbito de la ciencia y la tecnología. Desde sus orígenes, en el campo de la termodinámica durante el siglo XIX (Carnot, 1824; Clausius, 1867; y Boltzmann, 1877), ha sido una magnitud explicativa en la medición de fenómenos físicos relacionados con la distribución de las partes de un conjunto, desde la mecánica estadística hasta la cosmología. Y también ha sido utilizado en tareas comparativas entre familias lingüísticas (Montemurro y Zanette, 2011).

Nosotros utilizamos la versión que procede de la teoría de la información de Claude Shannon (1948), según la fórmula que presentamos en las asunciones metodológicas y que repetimos aquí en (96).

$$(96) H(X) = - \sum_{x \in X} p(x) \log_2 p(x)$$

Si bien es cierto que, originalmente, el objetivo ingenieril de la propuesta de Claude Shannon era obtener la máxima economía de tiempo, energía y dinero en el diseño de señales y canales técnicos, sus repercusiones posteriores como herramienta dentro de la teoría de la probabilidad han sido abundantes.

En términos generales esta medida poliédrica, en sus aplicaciones, se ha caracterizado como legitimadora de los siguientes atributos o propiedades, descritos en (97 a-e).

(97)

- (a) una medida de desorden
- (b) una medida de incertidumbre
- (c) una medida de información
- (d) un grado de ignorancia
- (e) una medida de diversidad

De estas, las que nos sirven como instrumentos explicativos son las que valoran la diversidad y el orden de un conjunto, de una población o de un sistema, sean datos para un mensaje cifrado o, como en nuestro trabajo, las categorías semánticas de un grupo de nombres.

En nuestro caso la clasificación en estas categorías semánticas para los nombres de los tres tipos de preposiciones y su distribución ofrecerá un resultado en el grado total de la entropía (diversidad y orden) de ese conjunto. Entenderemos que cuanto más baja es la entropía más ordenado es el conjunto, menos diverso y más previsible en su caracterización (saber la naturaleza semántica del nombre que vendrá después de la preposición). Y viceversa, a medida que aumenta la entropía crece el desorden y disminuye la posibilidad de saber qué sucederá.

La contribución de la preposición es baja o casi nula para la recuperación o la predicción del tipo de nombre que aparecerá en la clase funcional o gramatical. Por otra

parte, y a medida que la entropía aumenta, crecerá la diversidad y el desorden, es decir, su distribución en las categorías semánticas y la preposición jugarán un papel semántico más decisivo.

Recordemos en (98) cuáles eran las seis categorías semánticas que presentamos en las asunciones semánticas y que ya usamos en la observación de las agrupaciones del primer experimento.

(98) *Animado, Entidad abstracta, Evento, Inanimado, Temporal y Locativo.*

La herramienta clasificatoria de nuestra propuesta ha sido el TCO (Top Concept Ontology) de Wordnet.

El TCO partía de tres grandes grupos de entidades (Lyons, 1977), denominados de Primer orden, Segundo orden y Tercer orden. Se justificaba esa terna de entidades a partir de una correlación entre categorías sintácticas y valores semánticos, según la siguiente aproximación (99).

(99)

Nombres: designan personas, lugares o cosas (primer orden)

Verbos: expresan acción o estado (segundo orden)

Adjetivos: cualidades (tercer orden)

Uno de los objetivos del TCO, precisamente, consistió en detallar y mejorar esa macro taxonomía para poder establecer la red de conexiones internas de significados de los conceptos semánticos de lenguas que participan en el proyecto EuroWordnet.

Nuestro ajuste busca que las realizaciones semánticas posibles en el entorno preposicional se acomoden con cierta naturalidad. En la sección anterior vimos la configuración de rasgos de nuestras clases en correspondencia con el árbol semántico del TCO.

Veamos, ahora, algunas aclaraciones sobre cada una de las categorías. Así, la clase *Animado* incluye nombres que aluden a entidades con propiedades atribuibles a seres vivos (humanos o animales) o partes u órganos de ellos, desde las más genéricas: *individuo, hombre, mujer, niño, ternero* o *saltamontes*, hasta otras más específicas o restringidas: *consultora, teniente* o *instigador*. La *Entidad abstracta* recoge las ideas o los conceptos que no se identifican en el espacio-tiempo real, sobre los que no se puede predicar si son reales pero sí en cambio sí son verdaderos (Lyons, 1977). Hallamos nociones como *libertad, ideología, ideal* o *cristianismo*. *Evento*, por su parte, incluye todos los infinitivos que aparecen, y los nombres que denotan alguna forma de acción y suceso con un proceso temporal, tales como *partida, asalto, boda, exterminio* o *despedida* (Resnik, 2011). *Inanimado* engloba los que se refieren a entes creados con independencia de su tamaño o dimensión, tales como *vagón, maleta, jardinera* o *cazuela*. En este grupo también se incluyen los resultados de acciones como *golpe*; elementos naturales como *gasolina, pan*; o cantidades como *litro* o *metro*. En el caso de la ubicación espacial, la categoría *Locativo* se refiere a indicaciones espaciales concretas tales como *norte, sur, interior, exterior*, o lugares abiertos: *calle, avenida, playa*, topónimos: *Barcelona, Asia*, y construcciones, edificios o instalaciones: *ermita, ayuntamiento, hospital* o *gimnasio*. Por último, la categoría *Temporal* se compone de palabras que señalan, denotan e identifican el tiempo en cualquiera de sus modalidades, en un punto, en un período, en una fase o en una secuencia: *día, mañana, verano, martes, julio*, etcétera.

3.2.5 La realización del experimento

En primer lugar, buscamos, del conjunto de verbos de (83), los ejemplos de los que extraer los sintagmas preposicionales funcionales, semifuncionales y léxicos. Esa

indagación sigue los criterios sintáctico-formales establecidos en la sección 3.2.1.3, página 140, con objeto de garantizar que se dé también variedad sintáctico-argumental.

Una vez elaborados los tres conjuntos, se inicia la clasificación manual de todas las palabras que están en posición de complemento de la tripleta $X - Y - Z$, donde Y es una preposición del español y Z es su complemento.

Se obtienen 3898 tripletas. De ellas reciben una doble anotación manual un total de 719 con un acuerdo de 518, es decir, un 73 %.

Una vez clasificadas las palabras en esa posición elaboramos una tabla de Excel con las seis categorías semánticas y su frecuencia. La tabla utiliza el logaritmo de medición de la entropía sobre los datos que introducimos.

A partir de aquí se obtiene un índice de entropía para los seis categorías semánticas y para las tres clases de preposiciones, funcionales, semifuncionales y léxicas, con nombres repetidos y sin repetir. Con los datos de esas agrupaciones y los índices del valor de la entropía de los tres conjuntos procedemos a examinar los resultados.

Por último, hemos realizado una prueba de azar para ratificar que los resultados no podrían obtenerse de manera arbitraria, o ni tan siquiera aproximarse. La igualación resultante certifica que no es posible responder de manera accidental o casual a los datos de la entropía.

3.2.6 Análisis de los resultados

En las tablas 6 y 7 se reúnen los datos informativos con los resultados de la clasificación en categorías semánticas de los nombres y el valor de entropía (H) de las tres clases preposicionales propugnadas con repeticiones y sin repeticiones de nombres.

CON REPETICIONES			
CATEGORÍAS SEMÁNTICAS (6963 casos)	CLASES PREPOSICIONALES		
	FUNCIONALES	SEMIFUNCIONALES	LÉXICAS
Animado	1437	388	215
Inanimado	70	281	173
Entidad abstracta	261	1183	546
Evento	82	250	260
Locativo	164	120	627
Temporal	30	77	171
Total de casos	2044	2299	1992
Entropía total (H) =	1,47	2,031	2,37

Tabla 6. Clasificación de categorías semánticas y clases preposicionales y entropía (con repeticiones)

SIN REPETICIONES			
CATEGORÍAS SEMÁNTICAS (3898 casos)	CLASES PREPOSICIONALES		
	FUNCIONALES	SEMIFUNCIONALES	LÉXICAS
Animado	871	203	120
Inanimado	55	160	152
Entidad abstracta	146	696	374
Evento	73	168	181
Locativo	90	74	376
Temporal	23	26	110
Total de casos	1258	1327	1313
Entropía total (H) =	1,54	1,99	2,40
Diferencia entre F y SF	0,45		
Diferencia entre SF y L		0,41	
Diferencia entre F y L			0,86

Tabla 7. Clasificación de categorías semánticas y clases preposicionales y entropía (sin repeticiones)

El análisis global de los resultados obtenidos, y que se presentan en la tabla 7, pasa por tener en cuenta tres aspectos. En primer lugar, determinar si son pertinentes las

diferencias entre las repeticiones o no de los sintagmas nominales. Un segundo hecho es el valor de graduación o de progreso entre los tres grupos: funcionales, semifuncionales y léxicos. Y por último, el análisis pormenorizado de los contenidos semánticos de cada clase preposicional en relación a las seis categorías semánticas.

El total de nombres repetidos es de 2338. Este hecho podría distorsionar los resultados globales habida cuenta del porcentaje que representan, un 44,06 %. Pero no ha sido así. En términos absolutos las repeticiones apenas suponen variación de los totales. En la tabla 8 podemos ver la comparación entre repeticiones y no repeticiones, y las diferencias entre los grupos.

Tabla de comparación entre sintagmas nominales repetidos y no repetidos.	Repetidos	No repetidos	Diferencias
Funcionales	1,47	1,54	-0,07
Semifuncionales	2,03	1,99	0,03
Léxicos	2,37	2,40	0,02

Tabla 8. Comparación entre los totales de entropía con repeticiones y sin repeticiones

La repetición de los nombres no genera desviación significativa de los resultados. En dos casos, funcionales y léxicas, la no repetición provoca un aumento de la entropía. Las cifras, sin embargo, son tan bajas que no pueden valorarse como manifestaciones de ninguna tendencia cualitativa que merezca la pena investigar.

El segundo aspecto a comentar procede de las diferencias del índice de entropía en el ámbito de las no repetidas, y diana global de nuestra investigación en el experimento.

La diferencia entre las funcionales y las semifuncionales es de 0,45; y 0,40 la que separa a los semifuncionales respecto de las léxicas. En el gráfico 13, se recogen esos resultados.

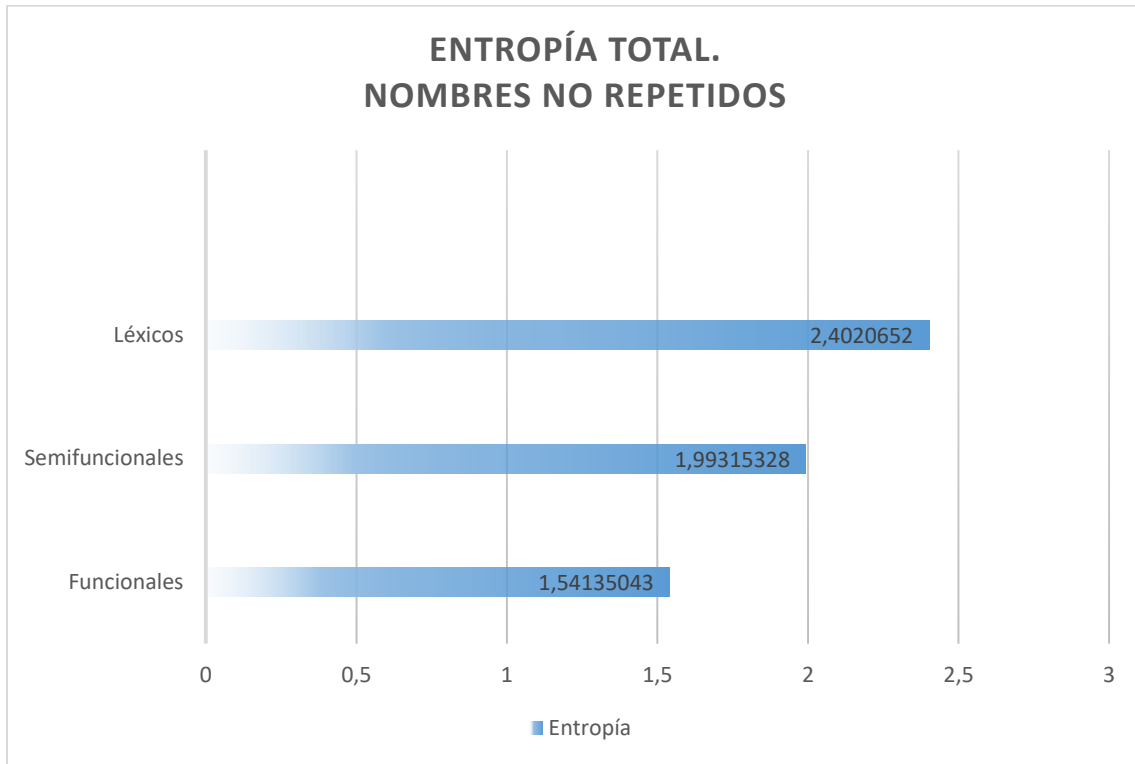


Gráfico 13. Entropía para las clases preposicionales según las categorías semánticas

La facilidad predictiva de las categorías semánticas de los nombres augura en la clase funcional que la contribución de la partícula en la tripleta es baja porque no necesitamos de su carga semántica para saber el tipo de significado que ocurrirá. Sabemos interpretar el contexto de esa clase de preposiciones porque, en definitiva, son poco decisivas en la composición global. Son los verbos y los nombres los que dirigen la arquitectura semántica de las tripletas, su grueso interpretativo. Se trata, en consecuencia, de resultados satisfactorios y coincidentes con la hipótesis sobre la ausencia casi completa de significado en el caso de las preposiciones con valores funcionales.

En el segundo grupo, sin embargo, la distribución de casos aumenta hacia los nombres de *Entidad abstracta*. Las preposiciones del régimen verbal, o semifuncionales, están orientando o dirigiendo el significado de los verbos que acompañan y de ahí el aumento entrópico en la organización de los nombres. Esa “orientación” supone una participación más distribuida del significado entre verbo y preposición; o si se quiere, un rol que aumenta su capacidad de intervención.

Para acabar, el grupo con mayor entropía coincide con las preposiciones con mayor significado: las léxicas. A medida que la preposición juega un papel semántico es más difícil saber qué sucederá en la posición de complemento, o qué requisitos seleccionales tendrá el verbo. Lo que en teoría de la información se denomina “desorden” o “incertidumbre” del mensaje se relaciona, en nuestra perspectiva, con una mayor acción expresiva para la preposición que dificulta la predicción sobre la categoría semántica. No solo no sabemos cómo serán los nombres (a qué categoría semántica pertenecerán) sino que la comprensión general del enunciado se apoyará también en la preposición.

La medida entrópica de los tres grupos preposicionales obtenidos de la clasificación de sintagmas nominales confirma, por tanto, que a mayor entropía y dificultad en la predicción del tipo semántico, mayor aportación de la preposición en la construcción del significado.

La prueba de azar prevista en la metodología se configura del siguiente modo. Se han construido tres grupos. Cada uno de ellos, (G1, G2 y G3) contiene, de manera proporcional, una tercera parte de los nombres de las tres clases, funcionales, semifuncionales y léxicos. Se ordenan alfabéticamente los verbos del experimento hasta que tienen una cifra arbitraria de cada tipo y una variación aleatoria de las tres clases, sin atender a ningún criterio posterior. A partir de aquí se vuelven a introducir los datos de clasificación de los nombres en la hoja de cálculo para verificar los nuevos resultados.

En la tabla 9 se recogen las nuevas combinaciones y los nuevos resultados de control.

Categorías semánticas	Segundo experimento Prueba de azar		
	Grupo 1	Grupo 2	Grupo 3
Animado	441	371	370
Inanimado	123	113	130
Entidad abstracta	341	424	451
Evento	123	155	143
Locativo	155	182	200
Temporal	49	50	55
Total de nombres al azar	1232	1295	1349
Entropía (H) =	2,26	2,29	2,30

Tabla 9. Resultados de la prueba de azar del experimento 2

Los índices de entropía obtenidos en los tres grupos son muy parecidos, de manera que el azar no daría cuenta en modo alguno de nuestros resultados, ni tan siquiera de manera aproximada. La irrelevancia de los resultados no permite introducir factores aleatorios en nuestra justificación.

Entre G1 y G2 la diferencia es 0,028; entre G2 y G3 0,008; y entre G1 y G3 0,036.

Pasamos a observar los detalles internos de cada clase, de un modo análogo a como lo hicimos en las agrupaciones del primer experimento.

3.2.6.1 La clase funcional

En el gráfico 14 se distribuyen porcentualmente las categorías semánticas para los sintagmas preposicionales funcionales, es decir, aquellos donde la preposición es muy débil semánticamente.

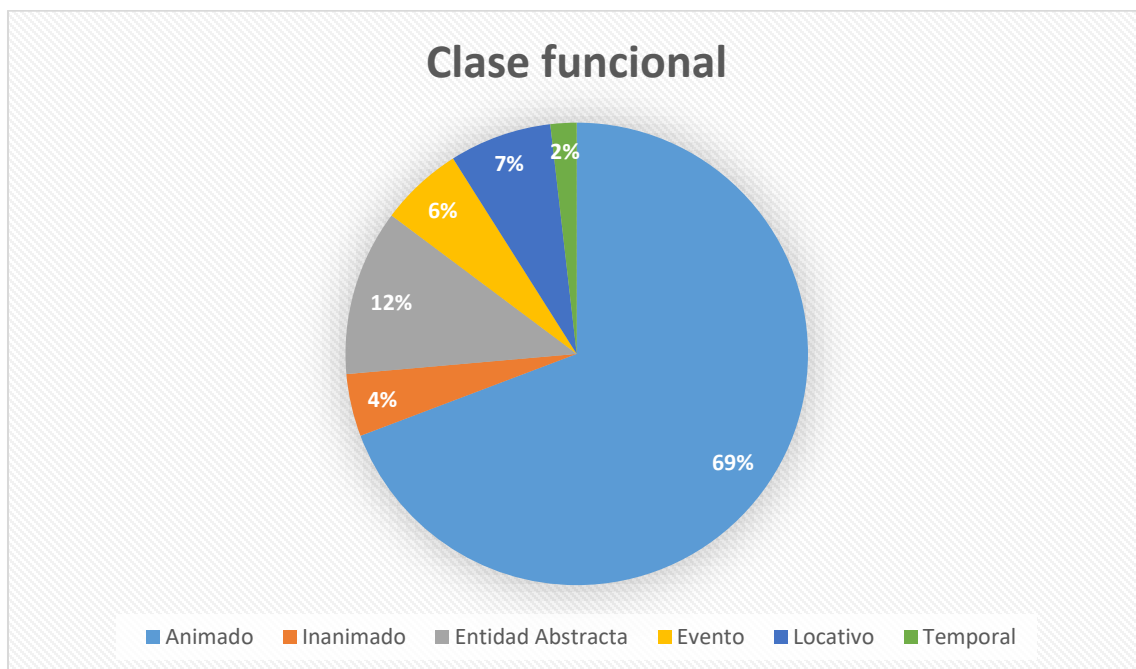


Gráfico 14. Distribución de categorías semánticas en la clase funcional

La preposición que domina de manera absoluta la funcionalidad en nuestro corpus es la *a*. La primera evidencia empírica es la concentración -871 de 1258- un 69,23 % del total de sintagmas nominales analizados, pertenecientes a la categoría semántica *Animado*. El concepto más frecuente es el vinculado a la condición humana. El rasgo “humano” engloba desde las denominaciones más genéricas: *hombre, mujer, niño, joven, viejo, adolescente*, hasta una diversidad muy extensa de condiciones, oficios, estados, ocupaciones o actividades encarnadas por los seres humanos: *vecino, madre, funcionario*,

investigador, profesional, esposo, soldado, agente, víctima, escritor, trabajador. En este grupo también se incluyen los cargos públicos, electos o no, y los colectivos u organizaciones, tales como *público, población, ciudadanía, autoridad, política, presidente, ministro, sindicato, universidad o grupo.*

Por otra parte, todos los nombres que se refieran a seres vivos o a partes u órganos de seres vivos se incluyen aquí como *avestruz, paloma, perro, árbol, cuello o nervio.*

Una síntesis de esta pluralidad expresiva se recoge en el cuadro 18.

Subdivisiones de la categoría <i>Animado</i>	Ejemplos
Nombres genéricos	Señora, varón, persona, chaval, churumbel...
Gentilicios	Cretenses, croatas, madrileño, bávaros...
Topónimos	Barranquilla, España...
Nombres propios	Marta, Nina...
Oficios y cargos	Empresario, tejedores, profesores, redactora, retratista, emperador, ingeniero, soldado, chef...
Nombres sobre relaciones personales	Sobrina, esposa, hijo, abuela...
Pronombres	Usted, quién, cualquiera...
Colectivos	Avanzadilla, colectividad, contingente, sección...
Numéricos	Treintena...
Partes de un ser animado	Cuello, hombro, nervio, muslo, ventrículo, córnea...
Condiciones (adjetivos o participios sustantivados)	Encadenados, allegados, concurrente, maligno, extraño, avaros, acusado...

Cuadro 18. Distribución de subdivisiones en la clase funcional para la categoría semántica

Animado

El segundo grupo que predomina, a una distancia remarcable, son los sintagmas nominales que representan la categoría *Entidad abstracta*, un total de 146 respecto de 1258, con un 11,60 %. Aquí nos encontramos nombres como *ley, idiotez, capacitación, argumento, aprobación, abstinencia, aquiescencia, industria o fuero.* La distancia cuantitativa respecto del primer grupo es considerable en términos relativos.

Y el tercer grupo, numéricamente soslayable, está formado por los que indican lugares o localizaciones con 90 de 1258, es decir, un 7,15 % del total. Sirvan como ejemplos casos como *derecha, izquierda, aquí, ahí*; hasta lugares genéricos tales como *muralla, oficina, calle, fuente, avenida, fábrica, tienda, puerta, templo, pantano, canal, país, ciudad, pueblo, casa o vivienda*.

Los datos anteriores están en consonancia con la previsión descriptiva de la gramática. De hecho, se asume que los complementos directos de personas, así como los complementos indirectos van introducidos por una *a*. Los resultados estadísticos confirman una predicción conocida y regular en el uso y en el conocimiento gramatical de la lengua.

Por lo que respecta a los tres grupos últimos, representan un 12 % del total (186 casos). La cifra inferior es la de temporalidad con nombres como *año, hoy, lunes, primavera o década*, y apenas un 1,76 % (23 casos). Las otras dos, *Evento* (73) e *Inanimado* (55), poseen un 5,12 % y 4,5 % respectivamente.

El hecho cuantitativo que mejor se adapta a nuestro argumento sobre la gradualidad es la alta concentración de una categoría semántica, *Animado*, indicio de la escasa relevancia interpretativa de la preposición *a*. Este hecho ha sido analizado recientemente en relación con la acumulación diacrónica de sus valores (Company, 2019).

3.2.6.2 La clase semifuncional

Si observamos el gráfico 15 comprobaremos que la distribución de categorías semánticas varía respecto del primer tipo. Esta alteración se produce con un aumento parcial o relativo de los tipos y, por tanto, un reparto de casos superior y una mayor complejidad en el análisis.

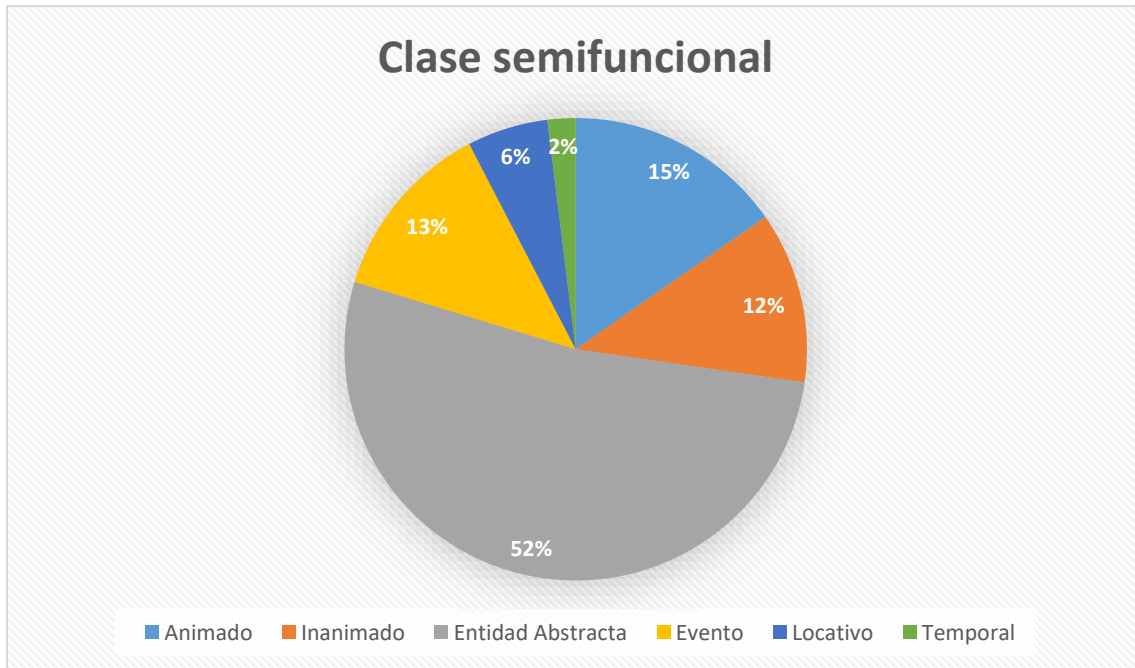


Gráfico 15. Distribución porcentual de categorías semánticas en la clase semifuncional

El segundo grupo de nombres lo forman los que aparecen con los verbos de régimen verbal. Los resultados nos indican que 695 de ellos, un 52,41 %, pertenecen a la categoría semántica *Entidad abstracta*, tales como *seguridad, fe, amor, concepto, verdad, igualdad, odio, bondad, alma, política, factorización, empecinamiento, conciencia, traición, democracia, idea, ilusión o deseo*.

De nuevo, como en el caso de los funcionales, presentamos el cuadro 19 ejemplificatorio del espectro de la categoría semántica mayoritaria.

Subdivisiones de la categoría semántica <i>Entidad abstracta</i>	Ejemplos
Nombres sobre ideas y conceptos (políticas, religiosas, morales...)	Ideales, anticlericalismo, autonomía, pragmatismo, democracia...
Nombres sobre actividades y operaciones humanas culturales, científicas...	Música, arte, cultura, zoología, agricultura, filatelia, metáfora...
Nombres sobre la cognición, la conducta y la mente	Conocimiento, ignorancia, voluntad, cordura, capacidad, memoria...
Nombres sobre emociones y actitudes	Sectarismo, intransigencia, victimismo, astucia, coraje...

Cuadro 19. Distribución de subdivisiones para la categoría semántica de Entidad abstracta

La presencia, el uso o la coocurrencia de nombres con valor abstracto se relaciona con la clase de significados incompletos o condicionados que coocurren con los verbos de régimen en español (Álvarez y Sánchez, 2009; Barrio, 2004; y Jiménez Norberto, 2011) o en otras lenguas como el inglés, en los denominados *Phrasal Verbs*. Así, los verbos preposicionales –otra denominación que reciben– como *calificar de, influir en, constar de, contar con, creer en, consistir en, tratar de, abogar por, incidir en, incitar a, andar en, dudar de, luchar por o pecar de*, por ejemplo, esperan que su caracterización semántica sea consumada con un significado externo que posibilite la determinación de todo el significado. A menudo se utiliza el concepto de “tema” para aludir a ese complemento de régimen, es decir, aquello que predica el verbo. Otra forma de presentarlo sería indicar que el significado entre el verbo y el nombre está compartido, o expresan una particular e inexcusable solidaridad semántica, como una especie de trabazón irreductible.

Es cierto igualmente que algunos de ellos poseen diversas acepciones. De hecho, algunos de estos verbos pueden aparecer con o sin preposición, circunstancia que afecta a la telicidad del predicado (Demonte, 1991 y 2011). Sería el caso de *pensar* (a) / *pensar en* (b) de (100).

(100)

(a) Pienso las preguntas del alumno

(b) Pienso *en* las preguntas del alumno

Al margen de las propiedades delimitativas sobre el aspecto y la telicidad de esa diferencia, nuestra hipótesis confirma que la mayor parte de nombres que se combinan con esta clase de verbos consigue un 52 % para un cierto tipo de expresión semántica: las entidades abstractas. La fuerza del dato orienta nuestra confirmación más allá de los matices introducidos por la preposición.

El segundo gran grupo, con cifras semejantes entre ellas, lo componen las categorías de *Animado* (203), *Evento* (168) e *Inanimado* (159), en ese orden.

Resulta muy pertinente para nuestra propuesta que la primera categoría del segundo grupo sea *Animado*, en tanto que una buena parte de los verbos de régimen se caracterizan como procesos mentales o verbales, exclusivos de la condición y la naturaleza humana, con rasgos de valoración, comunicación o sensación. Algunos de esos verbos son *advertir*, *consistir*, *tratar*, *decidir*, *advertir*, *abogar*, *acusar*, *apelar*, *ceder* o *complacer*. Del mismo modo, el desarrollo de la categoría *Evento* implica, en muchos casos, que los agentes, los actores o los participantes sean, necesariamente, *Animado*.

La temporalidad interna o tácita de los eventos se expresa en ejemplos como *celebración*, *reunión*, *encuentro*, *tribunal*, *conferencia*, *fiesta*, *actividad*, *concierto*, *boda*, *elecciones*, *cumpleaños*, *congreso* o *concurso*.

El tercer grupo, con porcentajes de 6 % y 2 %, respectivamente, incluye a *Locativo* y *Temporal*. Cualquiera de estas circunstancias expresivas matiza, de manera ulterior, el significado del núcleo verbal pero no son indispensables en la realización argumental de los verbos.

La gradualidad semántica se manifiesta con la reducción del grupo mayoritario de la categoría, desde el 65 % de *Animado*, en la clase funcional a un 52 % en la categoría *Entidad abstracta*. Son 13 puntos porcentuales de diferencia en el aumento entrópico y en la distribución. Luego existiría un segundo grupo de categorías semánticas entre un 12 % y un 15 %, y, por último, las que están por debajo del 6 %.

Crece la entropía respecto de la funcionalidad, se dispersan las categorías semánticas de los contextos y van ganando protagonismo interpretativo las preposiciones, asociadas inexorablemente con el significado del verbo.

3.2.6.3 La clase léxica

La tercera clase, la léxica, muestra su distribución entrópica en el gráfico 16 y confirma nuestra expectativa respecto de su distribución entrópica: es el más diverso y disperso en categorías semánticas.

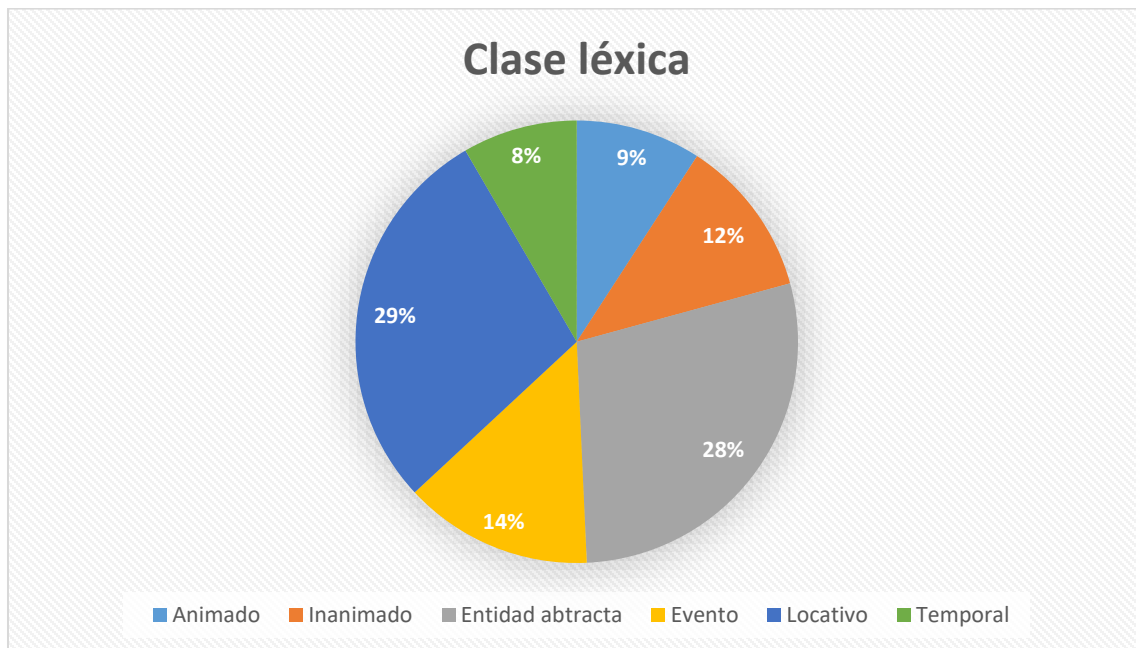


Gráfico 16. Distribución porcentual de categorías semánticas en la clase léxica

El primer hecho reseñable es el relativo equilibrio en la cantidad de sintagmas nombres correspondientes a cada una de las seis categorías. Estamos en la clase preposicional donde la entropía es mayor, es decir, ni la predicción, ni el orden, ni la certidumbre, tres propiedades de la medida instrumental del experimento se cumplen. Entre las categorías semánticas con menos nombres, la *Temporal*, con 110, y la que más, la *Locativo*, con 376, solo existe una diferencia de 266 nombres. Pensemos que ese contraste,

se ha ido reduciendo desde los dos extremos de la clase funcional (de 21 casos a 808) y de la semifuncional (de 25 casos a 685).

Asimismo, es importante comprobar cómo ha cambiado el protagonismo de los grupos semánticos en cabeza. En la clase semifuncional la posición central era para *Entidad abstracta*. Ahora, en cambio, se comparte la primacía con el *Locativo* (376).

De nuevo, como sucedía en el primer experimento para las preposiciones *a*, *hacia* y *hasta*, la expresión espacial adquiere una presencia notable en el operativo semántico de las preposiciones. En este sentido, se ha presentado una categorización lingüística de las entidades espaciales que ahonda en su riqueza expresiva (Aurnage, Hickmann y Vieu, 2007). Asimismo, los anclajes cognitivos de esa tarea lingüística también fueron sondeados (Evans y Chilton, 2010).

No cabe duda, además, que la naturaleza de la localización espacial vuelve a ser muy diversa y heterogénea, tal y como se recogen en las subdivisiones del cuadro 20.

Subdivisiones de la categoría semántica <i>Locativo</i>	Ejemplos
Nombres comunes de carácter genérico	Ciudad, aldea, municipio, lugar, sitio...
Topónimos	Nepal, Neptuno, Cuba, Segovia, Cádiz, Abisina...
Indicaciones generales y orientaciones	Espacio, allá, este, norte, salida, afueras...
Espacios abiertos (naturales o no)	Cima, lomas, llanura, zoco, carretera, cauce, paso, cenagal...
Espacios cerrados	Ducha, garita, laberinto...
Edificios, construcciones o instalaciones	Comisaría, castillo, piso...

Cuadro 20. Subdivisiones de la categoría semántica de Locativo

Por último, nos encontramos que el grupo preposicional que menos nombres incluye es el de la categoría semántica *Animado*. Parecería que el significado verbal no espera –en términos generales- ni un receptor, ni un beneficiario, ni un destino de carácter personal, ni la culminación o término, es decir, ninguno de los referentes habituales incluidos en esa categoría semántica.

El último aspecto destacado de la clase léxica es el crecimiento notable de las dos categorías hasta entonces minoritarias (*Locativo* y *Temporal*), como se aprecia en (101), en términos absolutos y en porcentaje.

(101)

Locativo: F (90 / 7,15 %) SF (74 / 5,5 %) L (376 / 28,63 %)

Temporal: F (23 / 1,8 %) SF (26 / 1,9 %) L (110 / 8,3 %)

La ubicación espacial es fundamental, como ya hemos indicado, en el despliegue semántico de los predicados verbales. El crecimiento en esta categoría semántica de entre un 400 % a un 500 % de las semifuncionales a las léxicas guarda relación con la ampliación de los campos semánticos del lugar y el tiempo en ese grupo de verbos. Sobre las semejanzas en las construcciones de tiempo y lugar se aportaron datos desde la óptica comparada de los universales lingüísticos (Traugott, 1978).

Además, el papel de la preposición es fundamental en estas tesituras porque orienta, dirige, delimita y sitúa el propósito del sentido o la acción verbal. En las lenguas con preposiciones, por tanto, su tarea es cardinal para la comprensión del significado y evitar una sobreinterpretación impropia de los requisitos de las lenguas humanas.

3.2.7 Conclusiones parciales

Desde el punto de vista de la comparación de los datos, en el gráfico 17 recogemos los resultados que permiten fortalecer la hipótesis en relación a las dos conclusiones que pretende alcanzar el experimento.

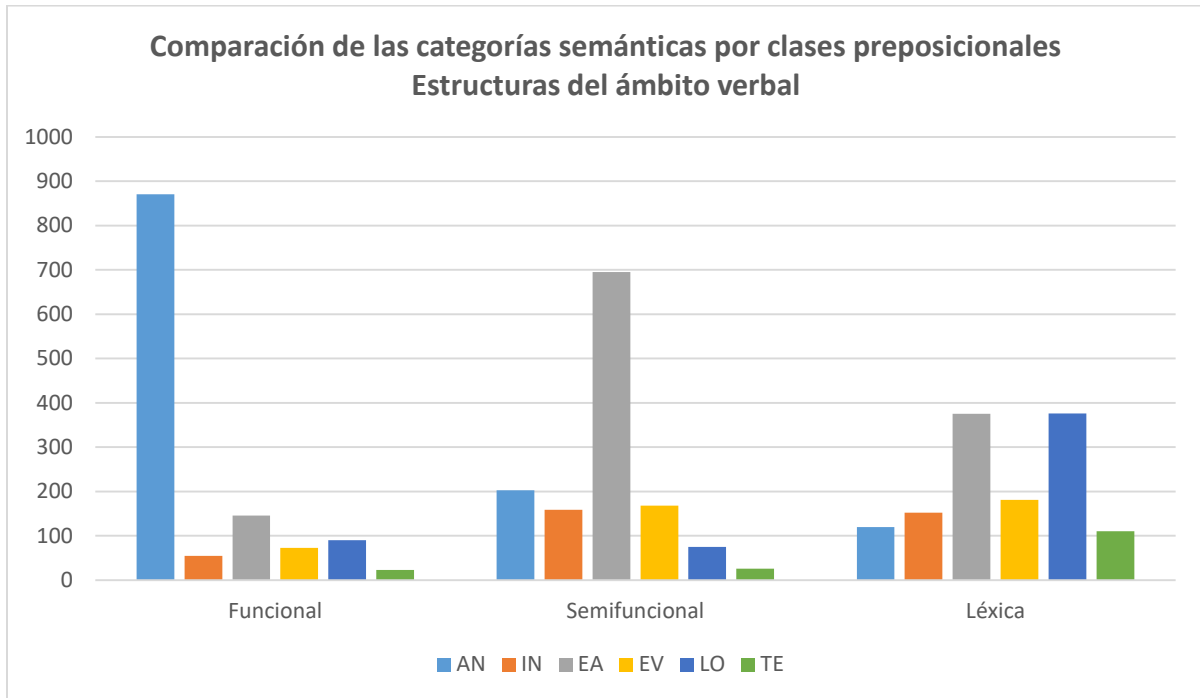


Gráfico 17. Comparación de las categorías semánticas para las tres clases preposicionales (Leyenda por orden de aparición en el gráfico: AN: Animado, IN: Inanimado, EA: Entidad abstracta, EV: Evento, LO: Locativo, TE: Temporal)

De la concentración inicial (*Animado* en funcional) se va transitando hacia una disgregación de posibilidades semánticas (*Entidad abstracta* y *Animado*) en semifuncional, hasta la dispersión de las léxicas con dos mayoritarias muy igualadas: *Locativo* y *Entidad abstracta*, y un reparto homogéneo de las otras con unas cifras relativamente altas). En la tradición francófona este crecimiento paulatino de relevancia semántica se presentaba, metafóricamente, como la emergencia de color en las preposiciones (Cadiot, 1997), es decir, de valores y matices de significado frente a la gramaticalidad.

Recordemos en (102 a-c) cuál es la cifra de separación entre las tres clases preposicionales.

(102)

(a) Entre funcionales y semifuncionales: 0,451

(b) Entre semifuncionales y léxicas: 0,40

(c) Entre funcionales y léxicas: 0,86

Estas diferencias avalan la primera conclusión parcial del experimento: existe una correlación gradual entre la cantidad de entropía y las categorías semánticas de los nombres introducidos por los tres tipos de valores preposicionales.

Otra consideración destacada a efecto de confirmar el carácter progresivo de ese cambio en el peso semántico de la preposición es que la distancia entrópica aumenta con una proporción bastante equilibrada.

El peso semántico de la preposición es inferior en la clase funcional en la medida en que la entropía de las categorías semánticas es baja puesto que el motor semántico se sitúa en el verbo y el nombre de la tripleta. A medida que se dispersan las categorías semánticas la preposición irá ganando protagonismo en el significado. La prueba es que en el segundo grupo se inicia la propagación y se corrobora cuantitativamente en el tercero.

El corolario de las dos conclusiones anteriores tendría una forma como la de (103).

(103) A medida que crece la entropía en la distribución contextual de las categorías semánticas de los nombres se enfatiza el papel o protagonismo semántico de las preposiciones en la interpretación de la tripleta, y viceversa.

Así, la *a* funcional de los sintagmas preposicionales contribuye con poco significado dado que ese valor semántico se puede inferir del conjunto o enunciado, o bien del núcleo (verbo) o bien del modificador o complemento (nombre). Una entropía baja determina, por tanto, que el conjunto de elementos –en este caso nombres– es más previsible porque se distribuye en menos categorías semánticas.

El hecho de que la mayoría (un 70% del total) de sintagmas nominales de los verbos que tienen estructuras sintácticas con preposiciones argumentales sea *Animado* confirma una predicción consolidada en la investigación sobre los complementos directos e

indirectos de persona. La irrelevancia o el carácter meramente funcional en estos casos ya fue comentado en el estado del arte (Saavedra, 2017) como una de las más aportaciones más recientes sobre las limitaciones semánticas de la construcción dativa, uno de los subtipos estructurales incluidos en esta clase.

En los otros dos grupos, en cambio, no existe predominio exclusivo de ninguna preposición. En la clase semifuncional aparecen siete preposiciones con el siguiente total de ejemplos entre paréntesis y ordenados en (104).

(104)

Contra (5) hacia (10), sobre (61), a (77), por (125), con (177), en (270) y de (602)

En este caso casi el 50 % de los casos se realizan con la preposición *de*, la mayoritaria en las estructuras de régimen verbal. La capacidad combinatoria de esta preposición está relacionada con su riqueza polisémica y su altísima frecuencia en variedad de lenguas (Burnard, 2000; y Rojo, 2011), tal y como comentaremos en el próximo capítulo (sección 3.3.5.1.1).

La multiplicidad más absoluta se aprecia en la clase léxica donde hemos encontrado ejemplos para dieciocho preposiciones, con la siguiente relación de casos, en la serie (105).

(105)

Ante (1), sin (1), según (5), bajo (6), entre (7), tras (21), desde (29), durante (32), sobre (36), mediante (39), a (40), por (44), de (49), hasta (58), hacia (108), para (119), con (218) y en (502).

En la serie anterior se observa, también, que a medida que crece la diversidad de valores léxico-semánticos atribuidos a las piezas se incrementa los ejemplos que hemos

encontrado en los corpus (véase el cuadro 4, sección 2.3.2, pág. 71). Las dos que han mostrado más productivas han sido *con* y *en*, destacadas transmisoras de la información modal y espacial en las tripletas y en los predicados de los ejemplos.

La entropía, por tanto, es una medida estable y seguro en la inferencia del significado de la preposición. La aportación preposicional va en aumento siempre que el desorden, representado en la entropía, crece. Del mismo modo, la gradualidad semántica de las preposiciones se observa a través de la medida de la organización semántica de su contexto en la posición de complemento. Además, la organización interna de cada una de las clases preposicionales –funcionales, semifuncionales y léxicas- se caracteriza a partir de rasgos y condiciones semánticas predecibles en las gramáticas normativa, descriptiva y teórica. La experimentación cuantitativa y computacional avala el reconocimiento de la descripción lingüística.

3.3 CAPÍTULO TERCERO

TERCER EXPERIMENTO.

Entropía y significado preposicional en estructuras nominales

3.3.1 Objetivos y justificación del tercer experimento

El objetivo de este experimento es confirmar la gradualidad semántica en la distribución sintáctica nominal de las preposiciones, es decir, cuando sus núcleos son nombres o categorías con rasgos nominales (como infinitivos o adjetivos nominalizados) (Girju, 2008). Metodológicamente continuamos con la medición de la entropía de las categorías semánticas de sus complementos (nombres) como variable de observación y examen para extraer conclusiones sobre nuestra hipótesis.

De nuevo, en una tripleta como (106), X y Z tendrán rasgos nominales, y P podrá ser cualquier preposición del español.

(106) X - P - Z

De la mayoría de valores léxico-semánticos que ha registrado la gramática normativa y descriptiva, que detallamos en las asunciones (Capítulo 2, sección 2.3.2, página 75) y que recordaremos en la sección siguiente, una parte de ellos se expresan exclusivamente en el ámbito de la complementación nominal, donde el sintagma preposicional es complemento de un nombre como en los ejemplos de (107 a-i).

(107)

- (a) Tarjeta de crédito
- (b) Hecho a mano
- (c) Café con leche
- (d) Viaje hacia el norte
- (e) Aclamación a gurú
- (f) Estatuto para capital
- (g) Expertos en feminismo
- (h) Examinados desde el rectorado
- (i) Cierre por deceso

Nuestro análisis consistirá en comprobar cómo se organiza el contexto nominal inmediato a la preposición a partir de las seis categorías semánticas establecidas y cómo se constituyen las posibles subdivisiones que encontremos en relación a la potencia o el peso semántico de las preposiciones.

3.3.2 Los valores léxico-semánticos en estructuras nominales

Traigamos otra vez la colección de valores léxico-semánticos que aparecerán en los sintagmas que analizamos en el ámbito nominal. En el cuadro 21 aparecen, organizados en tres columnas para nuestra clasificación, los enunciados descriptivos junto a las preposiciones que lo expresan y algún ejemplo para las estructuras nominales (procedentes de las referencias que ya indicamos en (72)).

VALORES LÉXICOS-SEMÁNTICOS DESCRIPTIVOS
Ejemplificación para estructuras nominales

Nombre de la clase preposicional y descripción de los valores léxico-semánticos	Preposiciones que lo expresan	Ejemplos
Funcional. * Introduce complementos que expresan agente, proceso o destinatario de alguna acción o proceso.	De	La decisión del profesor El cuidado de los gatos La salida del tren
Semiléxicos * Expresa el medio, el instrumento, el modo o la costumbre como se ejecuta o se hace algo.	A, con, de, en, por y según	Trabajo con sierra Clasificación por autores
* Expresa la causa o la consecuencia por la que se lleva a cabo una acción.	A, de, en, para, por y según	Huida por terror
* Expresa el precio o la materia de cualquier producto.	A, de, en y por	Una casa de tres millones Pizza de verduras
* Expresa la distribución, la proporción, el reparto o la sucesión de las cosas.	A y por	Tres juguetes por niños
* Expresa la velocidad, la cantidad límite o las medidas de las cosas.	A, de, hasta, por y sobre	Llegada sobre seis
* Expresa una relación de propiedad o posesión.	De	Esta camisa de Pablo
* Indica el encuentro, desencuentro, oposición, compañía o ausencia, o el contenido de las cosas.	De, con, sin y contra	Cita con mi primo Fusión contra corporación Vaso con agua Jardín sin flores
* Introduce el tema, el ámbito, el asunto o la materia	De, en y sobre	Lección de anatomía Agenda sobre clima

de qué trata alguna entidad o actividad.		
* Introduce y presenta los rasgos, las partes, el estado medio, los elementos integrantes de algo o de una serie.	De, con, entre, por	Una porción de pastel
* Expresa la asociación, la reciprocidad, la inclusión, la participación y la relación de varias personas o cosas.	Entre	Guerra entre árabes Regalo entre todos
* Expresa finalidad.	Para, por	Estudios para aprender
Léxicos * Localiza en el espacio objetos, personas o conceptos respecto de un punto (sin desplazamiento, como una mera orientación), en sentido literal o figurado e indica el punto, la zona o el lugar concreto en el que sucede o se localiza algo.	A, ante, bajo, entre, hacia, por, sobre, tras	Ropa al sol Paseo por ahí Puente sobre Ter Manifestación ante puerta
* Señala el origen, la procedencia, el punto de partida o el inicio de un movimiento o una observación.	De, desde	Excursión desde facultad Salida de Madrid
* Señala y expresa la dirección, la orientación, el destino, el término, el lugar intermedio, el curso o los límites respecto de un punto o de un desplazamiento.	A, contra, entre, hacia, hasta, para, por, sobre y tras	Dirección a París Camino hasta ermita Autopista hacia Lyon Cama contra ventana
* Indica la distancia exacta respecto de un punto o la localización en una escala.	A, de	Una distancia de 10 metros
* Introduce, indica y sitúa la hora concreta, el momento, la	A, de, entre	De tres a cuatro

situación, o un suceso en distintas fases del desarrollo temporal o de un acontecimiento de un modo concreto o aproximado.		Entre la una y las dos
* Expresa el punto, el término o el límite temporal en combinación con <i>de</i> , u otras preposiciones.	De, a, hasta	De martes a jueves De hoy hasta mañana De seis a siete
* Expresa, señala o localiza la aproximación, la frecuencia, la duración, la posteridad y la periodicidad (en meses, años, estaciones) con que se lleva a cabo una acción o acontecimientos, o cuando empieza.	A, bajo, desde, en, hacia, para, por, sobre y tras	Tres veces al mes Construcción en años Pruebas hacia marzo

Cuadro 21. Valores léxico-semánticos de las preposiciones en estructuras nominales

Tanto en el primer experimento como en el segundo (el dedicado a las estructuras verbales), los valores léxico-semánticos descriptivos han emergido en contextos que condicionaban la aparición de unos u otros, bien los verbos de movimiento para el primero, o bien la selección de 340 verbos. Digamos que los ejemplos estaban motivados o sesgados a partir de decisiones previas, bien la restricción del universo semántico, o bien el núcleo de la preposición. Ahora, en cambio, la obtención de los ejemplos y por tanto los valores que se observaran se realiza de tales valores sin atender a requisitos o restricciones sintácticas o semánticas entre los núcleos de las preposiciones.

En la agrupación del cuarto experimento recuperamos, como en el primer experimento con *clustering*, las denominaciones semifuncional y semiléxica para referirnos a dos de las clases de nuestra clasificación. Como vemos, se está produciendo una fluctuación en los grados semánticos, tal y como se recoge en el cuadro 22, donde se mancha de gris las agrupaciones que hemos supuesto en cada experimento.

Experimentos ↓	Clases preposicionales →	Funcional	Semifuncional	Semiléxica	Léxica
Primer experimento (preposiciones núcleos verbales con <i>a</i> , <i>hacia</i> y <i>hasta</i>)					
Segundo experimento (Estructuras en el ámbito verbal con todas las preposiciones)					
Tercer experimento (Estructuras en el ámbito nominal todas las preposiciones)					
Cuarto experimento (todas las preposiciones)					

Cuadro 22. Organización de las clases preposicionales por experimentos

Este hecho está motivado por la propia naturaleza de los datos, es decir, la clasificación de las categorías semánticas varía según los contextos sintáctico-argumentales, según el espectro semántico seleccionado (la localización espacial, real o figurada para el experimento 1), o según el núcleo categorial que controla la preposición (sean nombres o verbos para el experimento 2).

En este caso, la funcionalidad en las construcciones nominales proviene mayoritariamente de los ejemplos procedentes de nombres deverbales y no se observa semifuncionalidad en las clasificaciones. La transición entre clases, como vamos a ver, se aprecia mejor entre la semiléxica y la léxica. De hecho, la noción de semifuncionalidad, como tratamos en la sección 3.1.2, página 101, tuvo una motivación empírica que arrancaba en ejemplos y fenómenos de lenguas como el holandés. La fortuna de esa subdivisión ha continuado hasta nuestros días pero para categorías vinculadas a la categoría verbal donde sí parecen existir datos que avalan evidencias empíricas (Song, 2019).

3.3.3 Procedencia de los datos y procedimientos: entropía y categorías semánticas

Una vez más el corpus inicial que utilizamos es el Wikicorpus. Las características descriptivas del recurso se presentaron en la sección 2.4.4.1, página 94. A partir de un fichero inicial de 8.036 sintagmas nominales y preposicionales con repeticiones,

realizamos la elección según los valores léxico-semánticos de las tres preposiciones. Para disponer de una muestra más representativa y más diversa de nombres hemos usado, también, una lista de nombres deverbales de 4.017 extraídos del Diccionario Freeling que han complementado los datos finales (Padró, 2011 y 2012). Una fuente suplementaria ha sido el Corpus del español (Corpus Davies) extrayendo ocurrencias de estos nombres deverbales. La distribución total ha sido de 3.150 de Wikicorpus y 753 del Corpus Davies sin repetir.

Las categorías semánticas vuelven a ser las mismas que en las estructuras verbales, recordadas en (106) para facilitar el seguimiento de nuestra exposición.

(108) *Animado, Entidad abstracta, Inanimado, Evento, Tiempo y Locativo.*

De nuevo es la medición de la entropía la que nos permite seguir avanzando en la verificación de nuestra hipótesis. La fórmula que mide la magnitud entrópica (H) se reproduce en (109).

$$(109) H(X) = -\sum_{x \in X} p(x) \log_2 p(x)$$

3.3.4 La realización del experimento

El procedimiento que hemos seguido es la creación de un fichero principal donde se acumulan los datos procedentes de los corpus. A continuación seleccionamos los sintagmas preposicionales que representan los valores léxico-semánticos del cuadro 21 (sección 3.3.3, pág. 170) y procedemos, por último, a clasificar los nombres en posición de complemento en alguna de las seis categorías semánticas propuestas, tal y como hiciéramos en el segundo experimento.

Tras el recuento de las categorías introducimos los datos en la tabla Excel que actúa siguiendo el cálculo de la fórmula sobre la entropía.

Como en los dos experimentos anteriores, hemos realizado una evaluación manual paralela de control de una parte del corpus. Se han clasificado 968 nombres con una coincidencia de 779 entre ambos anotadores, es decir, un 80,47 %.

3.3.5 Análisis de los resultados

En la tabla 10 recogemos los datos resultantes del cálculo sobre la entropía (H) para los tres grupos preposicionales dentro de estructuras nominales –funcionales (F), semiléxicas (SL) y léxicas (L)- y las categorías semánticas sugeridas de un total de 3903 tripletas.

CATEGORÍAS SEMÁNTICAS	CLASES PREPOSICIONALES (En estructuras del ámbito nominal)		
	FUNCIONALES (F)	SEMILÉXICAS (SL)	LÉXICAS (L)
Animado	741	358	192
Inanimado	214	367	197
Entidad abstracta	191	329	198
Evento	60	99	195
Locativo	121	63	356
Temporal	24	45	153
Total de tripletas (3903)	1351	1261	1291
Entropía total (H) =	1,90	2,21	2,52
Diferencia entre F - SL	0,30		
Diferencia entre SL - L		0,31	
Diferencia entre F - L			0,61

Tabla 10. Clasificación de categorías semánticas y clases preposicionales y entropía en el ámbito nominal

Las cifras de los resultados confirman la hipótesis sobre la gradualidad semántica en tres grupos para un corpus de sintagmas preposicionales como complementos del nombre. Así, el grupo donde la preposición aporta menos al significado tiene una entropía

más baja (funcional). A medida que crece la contribución, la entropía aumenta (semiléxico y léxico, sucesivamente).

Esa gradualidad –de menor a mayor, significado y entropía- se produce, además, de un modo más equilibrado que en las estructuras verbales (en las comparaciones de las conclusiones finales se apreciarán estos detalles). La diferencia entre los dos primeros grupos –funcional y semiléxico- 0,3059778-, y los dos segundos grupos –semiléxicos y léxicos -0,3101957- es de apenas 0,0042179. Como en los ejemplos de estructuras del ámbito verbal se observa una cierta simetría en la variación entre clases preposicionales.

La correlación entre el aumento de la entropía y la contribución preposicional en la interpretación de la tripleta se apoya, además, en consideraciones sobre la distribución de las categorías semánticas de los nombres. Cuanto más entropía más variación se da en los nombres; y viceversa, es decir, una entropía baja significa que la preposición aporta menos significado y la categoría semántica del nombre está más determinada y condicionada en el contexto.

A continuación comprobaremos que ha sucedido en el interior de estas clases, cómo son los predomios semánticos de unas subdivisiones u otras, y cómo se han comportado unas preposiciones frente a otras.

3.3.5.1 La clase funcional

Las dos preposiciones que concentran los valores funcionales en estructuras nominales son la *a* y la *de*, con 123 casos y 1229 respectivamente.

El motivo de esa concentración radica en la naturaleza gramatical, vacía y funcional del quehacer semántico de sendas preposiciones. Veamos, en la distribución y el contenido de las categorías, qué evidencias prueban ese carácter residual. Para ello, mostramos en el gráfico 18, el reparto porcentual de las seis categorías semánticas de los nombres.

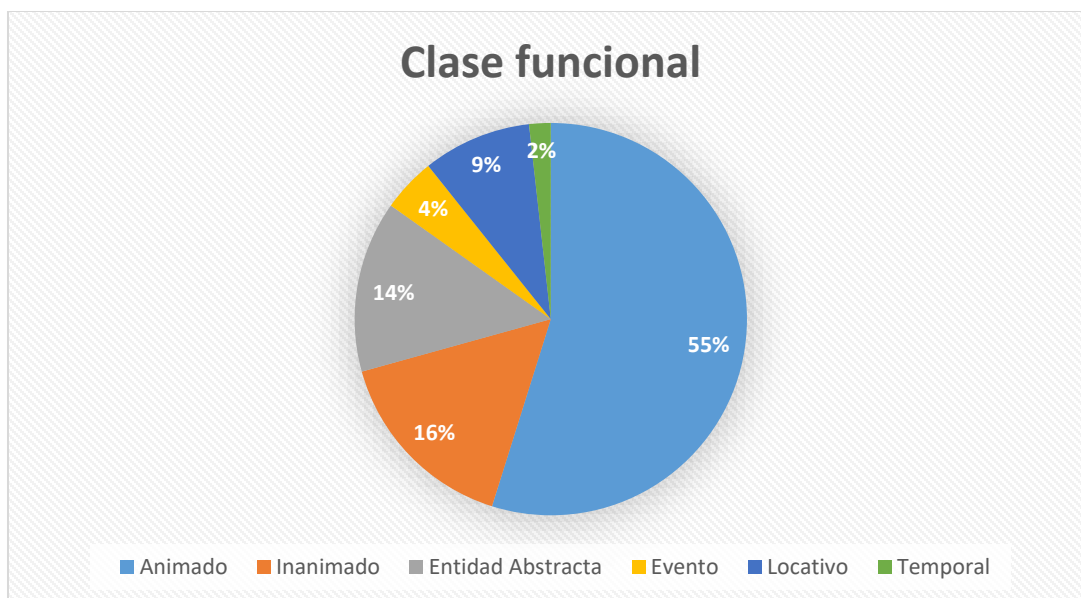


Gráfico 18. Distribución porcentual de las categorías semánticas. Clase funcional.

El predominio de la categoría *Animado* supera la mitad del total por cinco puntos porcentuales (55 %). En el seno de este grupo nos encontramos una diversidad profunda de posibilidades expresivas del rasgo semántico. En el cuadro 23 presentamos hasta 10 subdivisiones posibles de esa riqueza.

Subdivisiones de la categoría <i>Animado</i> de la clase funcional	Ejemplos
Nombres genéricos	Señora, varón, persona, chaval, churumbel...
Gentilicios	Cretenses, croatas, madrileño, bávaros...
Nombres propios	Marta, Nina...
Oficios y cargos	Empresario, tejedores, profesores, redactora, retratista, emperador, ingeniero, soldado, chef...
Nombres sobre relaciones personales	Sobrino, esposa, hijo, abuela...
Pronombres	Usted, quién, cualquiera...
Colectivos	Avanzadilla, colectividad, contingente, sección...
Cuantificadores	Treintena, decena...
Partes de un ser animado	Cuello, hombro, nervio, muslo, ventrículo, córnea...
Condiciones (adjetivos o participios; propiedades vinculadas a la clase)	Encadenados, allegados, concurrente, maligno, extraño, avaros, acusado...

Cuadro 23. Subdivisiones de la categoría *Animado* de la clase funcional

Desde el punto de vista sintáctico de muchos ejemplos de esta clase podríamos realizar versiones estructurales con sintagmas verbales. De hecho, y cómo vamos a comentar en breve, la mayoría de nombres-núcleos de esta clase son “deverbales”, es decir, una posible derivación morfológica de una categoría verbal (Peris, Taulé y Rodríguez, 2009). Estos procesos de nominalización ya fueron tratados para la subclase nominal de los eventivos (Grimshaw, 1990).

Así, en los ejemplos del cuadro 24, se manifiestan las dos posibilidades gramaticales de la expresión de la acción, la nominal y la verbal.

Tripletas del experimento	Equivalencia verbal
Abertura de <u>condensador</u>	El condensador se abre
Interrogatorio de <u>vendedores</u>	Los vendedores interrogan Los vendedores son interrogados
Límite de <u>censores</u>	Los censores limitan
Homenaje a <u>escritor</u>	El escritor es homenajeado
Aclamación a <u>gurú</u>	El gurú es aclamado

Cuadro 24. Ilustración de equivalencias nominal-verbal en la clase funcional con de-a

En cualquiera de los ejemplos el nombre en posición de complemento de nuestras tripletas pasa a ser el sujeto sintáctico de sus equivalentes verbales. Las versiones verbales se expresan en voz activa y pasiva. Aunque se trate de enunciados sintácticamente diferentes no se altera el sentido de la equivalencia que señalamos.

La alternancia –nominal versus verbal- podría insinuar dos aspectos relevantes. El primero es el vínculo empírico entre el predominio de la categoría semántica *Animado* y la posición de sujeto sintáctico de las oraciones. El segundo aspecto, y quizás más apreciable en nuestra predicción, es que la posibilidad gramatical de desaparición de las preposiciones *a* y *de* –en ese proceso de nominalización- no resta ninguna eficacia semántica a la versión verbal, es decir, la interpretación de los enunciados es idéntica. En el caso de la nominalización, una eliminación de la preposición no repercutiría en la interpretación del sintagma ya que el hecho de que en su versión verbal la preposición no se realice parece

indicar que la contribución de la preposición a la interpretación semántica es prescindible o no es importante.

La literatura sobre los nombres deverbales, también denominados nombres argumentales, procede de diversas metodologías e intenciones explicativas.

La importancia y el valor gramatical y semántico de este tipo nominal ha sido comentada, tanto desde la perspectiva de la morfología derivativa de estos nombres, (Santiago y Bustos, 1999), como desde la historia de los sufijos en español (Pharies, 2002) o incluso con herramientas de consulta on line (Faitelson-Weiser, <http://www.sufijos.lli.ulaval.ca/presentacion/>).

Por otra parte, las aproximaciones computacionales y sus utilidades para tareas en el procesamiento del lenguaje natural orientan sus propuestas para profundizar en la distribución semántico-argumental (Peris, 2010; y Taulé et alia, 2012).

Otros tipos nominales, como los eventivos, los abstractos, los concretos o los temporales han recibido también atención (García Mesager, 2007; Fábregas, 2010; y Resnik, 2011).

La segunda categoría semántica es la de *Inanimado* con 214 casos y un porcentaje del 16 %. Esta posición está en consonancia con la posibilidad sintáctica de que este tipo de nombres también sean candidatos gramaticales a la función sintáctica de sujeto. Aunque no incluyan el rasgo [+Humano] no tienen ninguna restricción añadida, tal y como se aprecia en los ejemplos de (110).

(110) Jalonamiento *de* balizas

Liquidación *de* prendas

Neutralización *de* toxina

Pago *de* recibos

Tachadura *de* tinta

En el seno de esta categoría semántica disponemos de un conjunto variado de casuística semántica. En el TCO (Top Concept Ontology) de EuroWordnet, presentado en la exposición descriptiva del primer experimento (Cap. 2, sección 2.4.2), se ofrecen una serie de rasgos identificativos dentro de las entidades de primer orden (subgrupo de Función) que usamos en el cuadro 25 para mostrar la diversidad que alcanza esta categoría.

Rasgos de Función (procedente del TCO)	Ejemplos en la clase funcional
Artefactos	Autocar, automóvil, cazas...
Comestibles	Queso, agua, alcohol, fruto...
Recipiente	Tubería...
Mobiliario	Barrera...
Vestidos	Jersey, ropa, tejido...
Instrumento	Micro, condensador, pluma, impresora, órgano...
Representación (lingüística) (económica)	Novela Reportaje Moneda
Substancia (gas)	Butano

Cuadro 25. Ejemplificación de rasgos del Top Concept Ontology para la categoría semántica Inanimado

La tercera categoría semántica que supera el 10% es la de *Entidad abstracta*. Nos hallamos todavía en el ámbito de lo que en el TCO (Top Concept Ontology) se consideran entidades de primer orden (tanto *Animado* como *Inanimado* están dentro de ella). Son ejemplos como *derechos, energía, ideales, autonomía, Dios, promesa, creencia o música*. Lo que en nuestra propuesta atribuimos a la abstracción, entendido como una entidad o

concepto que no puede ser identificada o representada en el continuo espacio-temporal, se cataloga en el TCO como representaciones con imágenes, lingüísticas o económicas.

Las otras tres categorías (*Temporal*, *Evento* y *Locativo*) no representan, en conjunto, más que el 15 % del total de los ejemplos. Toda la información léxico-semántica asociada a esas categorías requiere, sin duda, de la participación de preposiciones más activas y orientadoras del significado, como sucederá en los grupos que vienen.

3.3.5.1.1 Frecuencia, polisemia y significado. El caso de la preposición *de*

Dado que la preposición más común, en términos cuantitativos, y más polisémica, de esta parte del experimento y del español en general es la preposición *de* le dedicaremos una sección específica para vincular esa eminencia con los resultados del experimento.

Las evidencias actualizadas de esa frecuencia se pueden localizar en dos corpus de referencia para el español, usados por nosotros también, en el segundo experimento: el CREA (<http://web.frl.es/CREA/view/inicioExterno.view>) y el Corpus del español (Davies) en <https://www.corpusdelespanol.org/now/>. En esos almacenes de datos encontraremos los datos actualizados (pensemos que el segundo corpus está en proceso de incorporación constante de nuevos documentos). Si queremos acudir a un corpus publicado, disponemos de datos de frecuencia basados en el Corpus CUMBRE organizados en subgrupos (Almela et alia, 2005).

Una de las fuentes fundamentales para estudiar el valor de la frecuencia como indicador del papel semántico de una palabra o clase de palabras fue la conocida ley de Zipf (1949). Buena parte de los estudios cuantitativos de las lenguas han usado esa correlación como punto de partida para observar, estudiar, medir y/o localizar el significado (Ha et alia, 2002). Disponemos de una panorámica completa sobre la construcción de esas correlaciones entre frecuencia y significado (Rodrigues, 2009) que recoge referencias destacadas en el ámbito (Samuelson, 1996; Montemurro, 2002 y 2004; Ha et alia, 2002; Ferrer y Solé, 2002; y Ferrer y Reina, 2002). Las relaciones entre esta preposición y la categoría adverbial ya fueron tratadas anteriormente (Martínez y Jorgensen, 2013).

Investigaciones más recientes buscan, precisamente, regularidades matemáticas y estadísticas para la localización del significado (Ferrer y Vitevich, 2018). También encontramos aportaciones desde una perspectiva gramatical que distinguen entre frecuencia de uso y de inventario, y su rentabilidad en las búsquedas para el análisis gramatical (Rojo, 2011).

En el caso de la preposición *de* del español nos encontramos que es la que expresa y participa en más valores léxico-semánticos de los recogidos en el esquema 1, en 13 de 19 casos. Es cierto que seguida, muy de cerca, de los 11 y 10 valores de la *a* y la *por*, respectivamente. Sin embargo, la localización de ejemplos en los corpus de referencia para la diversidad de *a* y de *por* ha sido bastante más difícil. En el resto de preposiciones solo tres, *en*, *entre* y *sobre*, alcanzan los 5 valores.

Para captar el perímetro semántico de la preposición *de*, recojo en el esquema 1, la colección de valores léxico-semánticos específicos de esta preposición.

Clase preposicional	Descripción del valor léxico-semántico	Ejemplos
Funcional	Introduce el complemento que expresa agente, destinatario o proceso	<i>Lectura de acta</i> <i>Abdicación de rey</i>

Clase preposicional	Descripción del valor	Ejemplos
Semiléxica (modales y nocionales)	Medio, instrumento, y modo	<i>Ceremonia de pie</i> <i>Hombre de carácter</i>
	Causa	<i>Ataque de vergüenza</i>
	Precio, materia de un producto	<i>Silla de metal</i> <i>Carne de ternera</i> <i>Patatas de un euro</i>
	Medida de las cosas	<i>Tres de largo</i>
	Posesión	<i>Libro de Miquel</i>
	Contenido	<i>Lata de gasolina</i>
	Parte de un todo	<i>Poco de verdura</i>
	Tema	<i>Tesis de historia</i>

PREPOSICIÓN DE
Correspondencia
entre su adscripción a las
clases de preposiciones
y sus valores léxico-semánticos

Clase preposicional	Descripción del valor	Ejemplos
Léxica (locativa y temporal)	Origen, procedencia, destino y punto de partida	<i>Llegada de Bilbao</i>
	Distancia	<i>Viaje de kilómetros</i>
	Inicio de un periodo de tiempo	<i>Cierre de tres</i>
	Momento, situación, etapa o edad donde sucede un evento	<i>Dedicación de día</i>

Esquema 1. Correspondencia entre su adscripción a las clases de preposiciones y sus valores léxico-semánticos

La aparición de ejemplos con *de* representa un 51,85 % del total (2024 de 3903 ejemplos). Su frecuencia de aparición se reparte en los tres grupos funcional (1229 de 1351, un 90,96 %), semiléxicos (780 de 1259, un 61,95 %), y léxicos (15 de 1287, un 1,16 %).

La proporción de casos indica que en su mayoría los casos de la preposición *de* son los menos densos semánticamente. Sin embargo, el aluvión de posibilidades expresivas es considerable. La correlación de la ley de Zipf sobre la frecuencia de la *de* se apoya, además, en la polisemia de la preposición. La importancia de la variedad de significados desde la perspectiva computacional fue tratada desde diferentes ópticas (Ravin y Leacock, 2000); aunque otras caracterizaciones prefieren justificar esa amplitud de espectro semántico como un contexto polifuncional no solo de la *de* sino de todas las preposiciones (Company, 2019).

De hecho, se ha estudiado un ejemplo de ese vínculo: frecuencia y polisemia a partir de un texto concreto obteniéndose evidencias estadísticas significativas (Petrík, 2011); también a partir de preposiciones concretas del inglés (Alam, 2008), del holandés (Colombo y Flores d'Arcais, 1984), o en la indización automática para la identificación de contenidos en un texto (Urbizagástegui y Restrepo, 2011).

Como vimos, el protagonismo de la preposición *de* ha sido extensamente documentado en la historiografía gramatical (Gutiérrez Ordóñez, 1978). Las referencias bibliográficas sobre su carácter polisémico se han sucedido en la larga tradición descriptiva. Algunas de esas aportaciones ya fueron citadas en nuestro estado del arte, Capítulo 1, sección 1.4, página 20 y siguientes. Las más comunes proceden de los enfoques prototípicos (Funes, 2011) o de la diacronía de sus significados desde el siglo XIII hasta el siglo XXI (Granvik, 2012). Del mismo modo, su tarea semántica en ciertos fenómenos sintácticos como el “dequeísmo” ha sido objeto de interés y exégesis (Almeida, 2007).

3.3.5.2 La clase semiléxica

Como podemos apreciar en el gráfico 19, que representa la distribución de categorías semánticas en esta clase, ya se aprecia una de las características definitorias de la hipótesis gradual sobre el significado preposicional, el aumento de la dispersión semántica en el contexto.

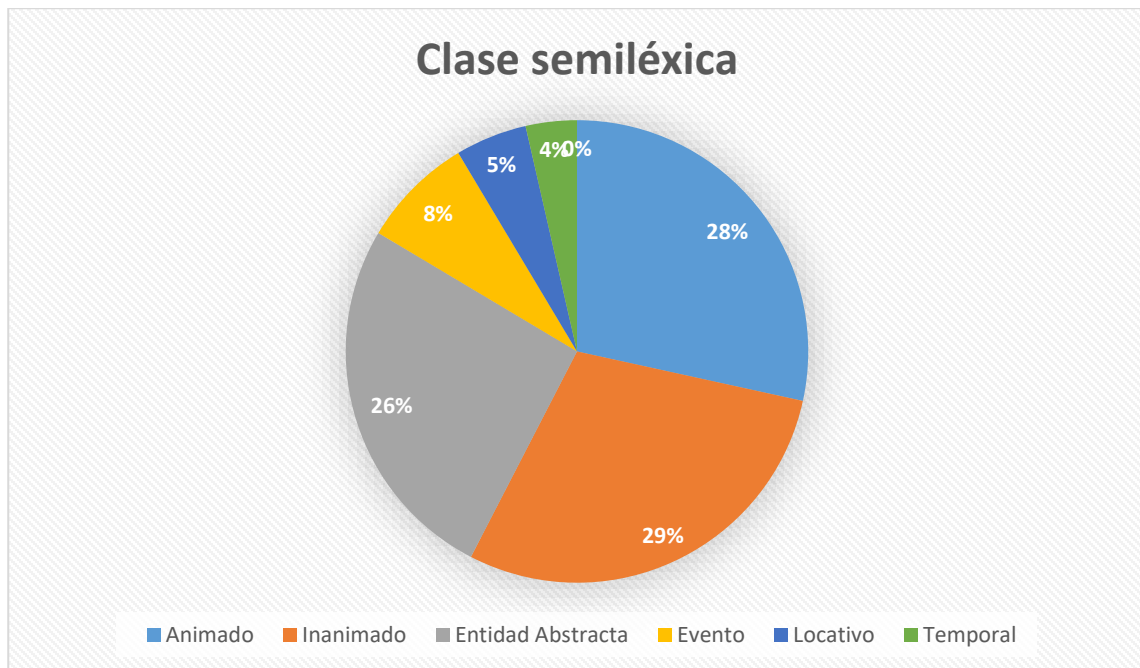


Gráfico 19. Distribución porcentual de las categorías semánticas en la clase semiléxica

Un hecho previo, que es coadyuvante de ese proceso de dispersión a partir del aumento de la entropía, es la presencia de 7 preposiciones (*de, con, sin, sobre, entre, por y para*) en lugar de las dos de la clase funcional (*a, de*).

El total de ejemplos por cada una de las piezas se recoge en la tabla 11.

Tabla de ocurrencias de las preposiciones en la clase semiléxica Ordenadas de mayor a menor y con porcentajes (1261)	
Preposición	Número de ocurrencias
De	782 (61,95 %)
Con	176 (13,97 %)
Sin	88 (6,98)
Sobre	82 (6,51 %)
Para	74 (5,87 %)
Por	45 (3,57 %)
Entre	16 (1,27 %)

Tabla 11. Ocurrencias de las preposiciones de la clase semiléxica

El reparto entre preposiciones y categorías semánticas no permite establecer correlaciones vinculadas con la gradualidad. Sí, en cambio, respecto de los propios valores léxico-semánticos de cada una de ellas. Uno de los ejemplos más representativos es el caso de *entre* que en sus 16 ejemplos selecciona contextos con nombres de la categoría *Animado* como *hermanos, maestros, acompañantes, responsables* o *cónsules*.

Otras dos preposiciones donde se produce una correspondencia entre su aparición y el predominio de una categoría semántica son *por* y *para*. En el primer caso se recogen las expresiones de causa con un predominio semántico de las *Entidades abstractas* (22 casos), seguidos de la categoría *Inanimado* (14). Por lo que respecta a *para*, baluarte de la expresión de finalidad, se acompaña de dos categorías mayoritariamente: *Evento*, con 31 ejemplos, e *Inanimado* con 33. Parece razonable porque en esos tipos de nombres se

concentran los objetos y los límites de la finalidad del significado: sean objetos u entidades materiales como *capazo, bote, camiones, armario, dispositivo, estuche o mástil*, acciones en forma de infinitivo verbal como *purificar, subyugar, aprender, tensar, averiguar, o contratar*; o nombres eventivos como *labranza, batalla o entierro*.

En el examen del crecimiento de las categorías semánticas se observa que pasamos de una sola categoría predominante, *Animado*, a tres, *Animado, Inanimado y Entidad abstracta*, alcanzándose un 81 % de los miembros de la clase preposicional.

Las dos categorías con mayor número de casos son *Animado e Inanimado* con 358 y 367 ejemplos respectivamente, siendo porcentualmente casi idénticas (28 %). Representan más de la mitad del grupo. Las propiedades descriptivas de estos nombres se pueden subdividir de modo semejante a como hicimos en la clase funcional. Añadimos algunos tipos nuevos para completar el análisis del espectro, como se aprecia en el cuadro 26.

Subdivisiones de la categoría <i>Animado</i> de la clase semiléxica	Ejemplos
Nombres genéricos	Iguales, vecinos, principiante...
Nombres colectivos	Población, raza, banda, peña, congregación, vecindario...
Oficios, profesiones, cargos e intereses	Bailaora, buzo, soldado, gruista, espías, músicos, ujier, jornalero...
Nombres sobre relaciones personales	Vecino, amante, heredero, pareja...
Animales y especies de animales	Saurios, insectos, mejillones, langostinos...

Cuadro 26. Subdivisiones de la categoría semántica Animado en la clase semiléxica

La tercera categoría semántica, *Entidad abstracta*, está muy cerca de las anteriores con 327 casos (26 %). Se reúnen aquellos nombres que aluden y refieren a entidades del universo que no pueden ser identificadas física o materialmente. Nuevamente, la calidad de los matices semánticos de esta categoría nos permite afinar la taxonomía interna como la del cuadro 27.

Subdivisiones de la categoría <i>Entidad abstracta</i> de la clase semiléxica	Ejemplos
Conceptos y nociones de carácter filosófico.	Fundamento, paz...
Nombres que implican un transcurso.	Desarrollo, proceso...
Rasgos, disposiciones o posibilidades mentales, cognitivas y emocionales.	Actitud, felicidad, mando, voluntad, comportamiento...
Hechos culturales de la condición humana.	Historia, política, música, negocio, enseñanza, lingüística...

Cuadro 27. Subdivisiones de la categoría semántica Entidad abstracta en la clase semiléxica

El segundo bloque de categorías es minoritario (17 %), y se compone de los nombres clasificados en las categorías *Locativo*, *Temporal* y *Evento*. Sin embargo, y como vemos en la comparación de la tabla 12, se ha producido un crecimiento en dos de las tres entidades respecto de la clase anterior, la funcional.

Categoría semántica	Clase preposicional	
	Funcional	Semiléxica
Evento	60	99
Locativo	121	62
Temporal	24	45

Tabla 12. Comparación de las categorías semánticas minoritarias de las clases funcional y semiléxica

En la clase semiléxica crecen significativamente los casos *Temporal* y *Evento*; con un aumento de casi el 60,60 % en el primer caso, y del 53,33 % en el segundo.

En el caso de la categoría semántica *Evento* la explicación se sustenta en la aparición de la preposición *para* en este grupo debido a la irrupción de tripletas con formas verbales de infinitivo.

El aumento de la entropía respecto de la clase funcional obedece, por tanto, a la diseminación de los valores léxico-semánticos en un conjunto más extenso de preposiciones. Además, el protagonismo de las preposiciones se muestra en el incremento considerable de las categorías semánticas, y en la riqueza, la subdivisión y la matización de sus componentes.

3.3.5.3 La clase léxica

El primer aspecto destacable en el tercer grupo de preposiciones es la extensión de los valores léxico-semánticos de 7 piezas a 16. Esta presencia de casi la totalidad de ejemplos de la categoría, salvo *según, con, sin*, tiene que ver con la naturaleza semántica de la clase. El ascenso y la participación de las preposiciones en la construcción del significado de las tripletas se apreciará en el abanico de posibilidades que ofrece.

En la tabla 13 indicamos el número de nombres para cada una de las preposiciones.

Casos por preposición (ordenadas de mayor a menor). Clase léxica	
En	592
a	122
Hacia	121
Ante	95
Contra	61

Hasta	55
Durante	53
Mediante	45
Desde	42
Por	28
Sobre	27
Tras	20
De	15
Bajo	9
Para	6
TOTAL	1291

Tabla 13. Relación de ocurrencias por preposición en la clase léxica

El predominio de la preposición *en* guarda relación, como indicaremos a continuación, con la importancia de la expresión locativo-espacial en esta clase preposicional. De hecho, la categoría semántica más numerosa es *Locativo* (356 casos). No en vano, es una de las grandes funciones semánticas de las preposiciones: la tarea topológica, tal y como referimos en el estado del arte y en los objetivos del primer experimento (Capítulo 3, sección 3.1.2, página 95). Y así, las otras más numerosas en casos, *a* y *hacia*, también presentan en sus contextos nombres relacionados con este tipo de significado.

Pasemos al análisis de la distribución por categorías semánticas de los nombres que la componen, a partir del gráfico 20.

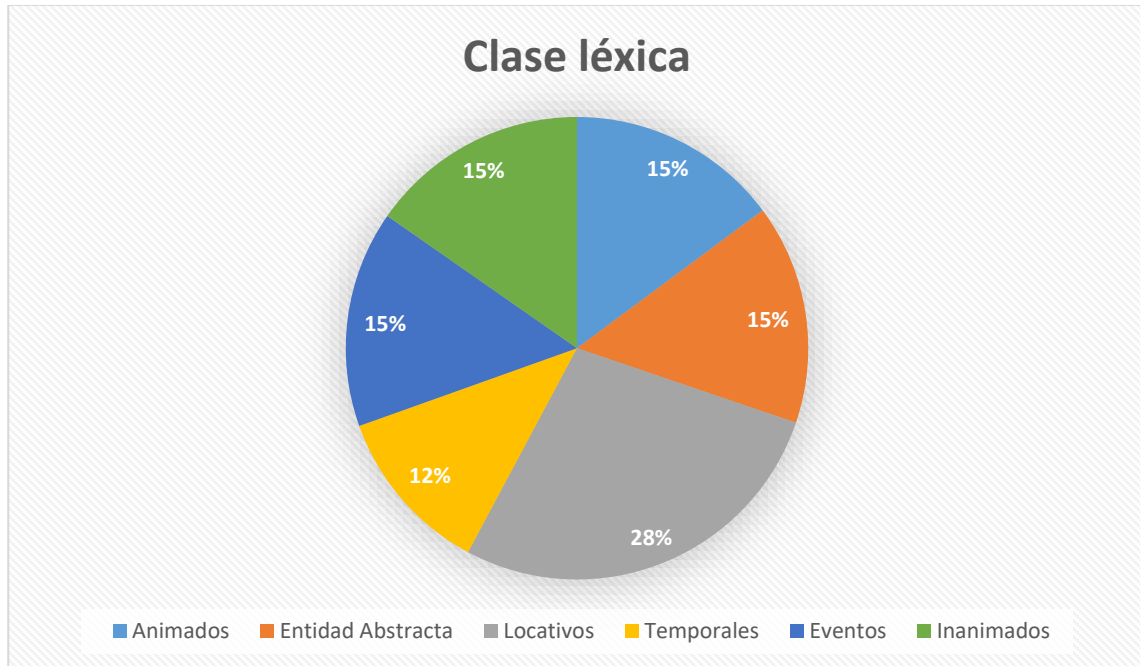


Gráfico 20. Distribución porcentual de las categorías semánticas de la clase léxica

Tres fenómenos resultan valiosos en la confirmación de nuestra hipótesis sobre la gradualidad semántica: (i) la concentración relativa en una categoría semántica paradigmática de la categoría –en la *Locativo*–, (ii) un reparto cuantitativo semejante en la mayoría de categorías (4 de 7), y (iii) la proporcionalidad del crecimiento entre categorías semánticas respecto de las otras clases, la funcional y la semiléxica.

Efectivamente es la expresión del espacio, el lugar o la localización la que aglutina el mayor número de casos (355) con un 28 %. Son cuatro preposiciones las que participan activamente en esta configuración: *en*, *a*, *hacia* y *sobre*. Así, las subdivisiones alcanzan muchos aspectos sobre la ubicación, tal y como se coleccionan en el cuadro 28.

Subdivisiones de la categoría semántica <i>Locativo de la clase léxica</i>	Ejemplos
Nombres generales	Superficie, entorno, entrada, elevación, enclave...
Espacios determinados por razones diversas	Comarca, provincia, país, terruño...
Construcciones, edificaciones e instalaciones, o partes interiores de ellas	Cabildo, chavelas, cortes, talleres, catedral, castillo, aeródromo, guardería, terraza, salón...
Espacios naturales	Parque, barranco
Espacios urbanizados	Avenida, calle, ciudad, plazuela, glorieta, rotonda, boulevard...
Topónimos	África, Cuba, París, Irán, Chicago...

Cuadro 28. Subdivisiones de la categoría semántica Locativo de la clase léxica

Conviene comentar también aquellos ejemplos donde el complemento del nombre es un objeto que, por la naturaleza semántica de todo el sintagma es interpretado como un lugar, entendido como un recipiente, receptáculo o espacio específico. Serían los casos de *aguarrás en botellín, alojamiento en nave, anillo en joyero o avispa en parra*. En estos casos se manifiesta, una vez más, el sentido composicional en la interpretación global de las tripletas (como ya subrayamos en nuestras asunciones semánticas), y el papel decisivo de la preposición en la comprensión del sentido de los sintagmas.

En segundo lugar se aprecia una relativa coincidencia numérica en los casos entre *Entidad abstracta, Inanimado y Animado*. Constituyen un 45 % de la clase con un reparto de 198, 197 y 191 casos respectivamente. Son las mismas categorías que resultaron mayoritarias en la clase semiléxica. La diferencia, sin embargo, reside en el descenso de casos y la proximidad numérica que se ha producido ahora en todas las categorías, es decir, un aumento de la dispersión o un reparto de los casos.

El último aspecto destacable es la proporcionalidad en la evolución en la mayoría de las categorías de la clase. Así, en la tabla 14, comparamos las categorías de la clase léxica, indicando la diferencia (Dif.) en número de una a otra, y de mayor a menor.

Diferencia numérica entre categorías semánticas en la clase léxica (ordenadas de más a menos, e indicando el porcentaje)						
Léxicas	LO	EA	IN	AN	EV	TE
	355	198	197	191	195	151
		Dif. 157	Dif. 1	Dif. 6	Dif. 4	Dif. 34

Tabla 14. Diferencias numéricas entre las clases semánticas de la clase léxica

La diferencia entre las dos más separadas, la *Locativo* y la *Entidad abstracta*, es de 157 ejemplos. Pero lo significativo es que la variación entre las demás categorías semánticas oscila solo entre 1 y 34.

Si realizamos la comparación de esos resultados con las clases anteriores comprobaremos que esa proporcionalidad era inferior en las clases funcional y semiléxica. Lo podemos ver en la tabla 13.

Diferencias numéricas de categorías semánticas entre los tres clases preposicionales (ordenadas de más a menos, e indicando la diferencia)						
Funcionales	AN	IN	EA	LO	EV	TE
	741	214	191	121	60	24
		Dif. 427	Dif. 23	Dif. 70	Dif. 60	Dif. 44
Semiléxicas	IN	AN	EA	EV	LO	TE
	367	358	327	99	63	45
		Dif. 9	Dif. 31	Dif. 229	Dif. 36	Dif. 8
Léxicas	LO	EA	IN	AN	EV	TE
	355	198	197	191	195	151
		Dif. 164	Dif. 1	Dif. 6	Dif. 4	Dif. 34

Tabla 15. Comparación de las diferencias por ocurrencias entre categorías semánticas de las tres clases preposicionales

En la clase funcional el gran salto se produce entre la primera y la segunda (*Inanimado* y *Animado*) con 427 ejemplos de diferencia; y en el interior de la clase se oscila

entre 23 y 70, el doble respecto de la léxica. En cambio, en la semiléxica la variación más fuerte se encuentra entre las tres primeras (*Animado*, *Inanimado* y *Entidad abstracta*) y la cuarta (*Evento*), 229 ejemplos. Este hecho explica cómo ha crecido la entropía en la clase semiléxica y cómo la distribución se reparte ahora en tres categorías semánticas (*Animado*, *Inanimado* y *Entidad abstracta*) apuntando la dispersión del significado.

Asimismo, en la clase léxica, la categoría que aumenta de modo más destacado es la *Temporal*, respecto de lo que sucedía en la clase semiléxica. Pasa de 45 casos a 151. Los ejemplos responden a la diversidad de nombres referidos a la dimensión física donde se incardina la existencia, las actividades y las acciones humanas o de otros seres vivos. A modo de ejemplificación, disponemos el cuadro 29 con algunos casos de las subdivisiones que han aparecido en nuestro corpus.

Subdivisiones de la categoría semántica <i>Temporal</i> de la clase léxica	Ejemplos
Nombres genéricos	Tiempo, ciclo, fase, día, noche, madrugada, lapso, actualidad...
Periodos	Día, semana, mes, año, trimestre, semestre, siglo, era...
Fechas y números de horas	2006, dos, tres...
Días de meses, días y estaciones	Lunes, verano, solsticio...
Fases de la vida	Vejez, infancia, niñez, adolescencia...

Cuadro 29. Ejemplos de las subdivisiones de la categoría semántica *Temporal* en la clase léxica

Los resultados de la organización entrópica de la clase muestran, junto a las propiedades de crecimiento repartido y proporcional, una presencia destacada de los significados más estructurados en el universo semántico: el espacio y el tiempo. De hecho, la concentración en los nombres reservados a las circunstancias de tiempo y lugar son las que crecen más. Son 506 casos, un 39,31 % del total.

Se trata, por tanto, de la clase donde la predicción sobre qué categoría semántica de nombres acompaña a la preposición es más difícil. Como venimos señalando esta característica contribuye a considerar el peso semántico de la preposición como imprescindible en la interpretación de los sintagmas. Hemos visto, además, que la

ampliación del uso de más piezas preposicionales en más del doble (de 7 a 16) supone una diversificación de los valores léxico-semánticos disponibles.

3.3.6 Conclusiones parciales

Los resultados obtenidos en la medida de la entropía de los 3903 sintagmas preposicionales complementos de un nombre y clasificados en tres clases confirman nuestra hipótesis sobre la gradualidad semántica en la relación básica que venimos considerando: entre la predictibilidad de la categoría semántica de los nombres y el aumento de la contribución de la preposición en el significado total del sintagma.

Veamos, en el análisis comparado del gráfico 21, qué cifras son destacables respecto de la estructura interna (u orden) de las tres clases y el progreso a partir de determinadas categorías semánticas.

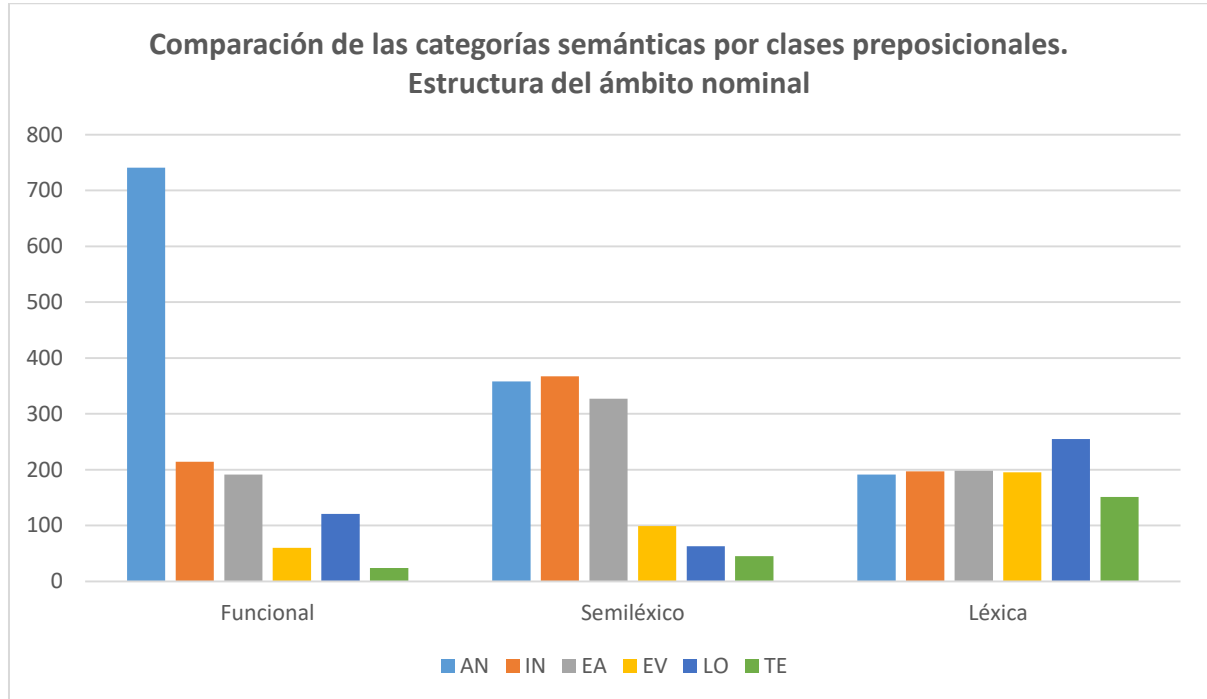


Gráfico 21. Comparación de las categorías semánticas por clases preposicionales

La funcionalidad y la transparencia semántica junto a valores meramente gramaticales son las propiedades de la clase preposicional en la que es más fácil predecir qué clase semántica aparecerá. Hasta en un 55 % del total nos encontramos con nombres de la categoría *Animado*. Este resultado es consistente con aspectos sintácticos y morfológicos sobre el tipo de nombres y su configuración, y sus funciones sintácticas en la oración. Además, solo dos preposiciones, la *a* y la *de* son las protagonistas de esta tarea meramente marcativa, en consonancia con su frecuencia reconocida en grandes corpus.

En la semilexicalidad, en cambio, se inicia el crecimiento de la entropía de la clase y, en consecuencia, es más difícil averiguar qué universo semántico se identifica en el contexto preposicional. Efectivamente se pasa de 2 a 7 preposiciones y es otro indicador inequívoco de que la densidad asciende en el grupo porque la relevancia de la preposición está justificada cuantitativamente.

Por último, en la clase léxica, y siguiendo la dispersión semántica del entorno preposicional, suben todas las categorías (hasta cinco colores del gráfico) por encima de 150 casos y cuatro de ellas de un modo parejo. Las preposiciones ejercen una actividad semántica en la composición de la tripleta, y surge la expresión de las circunstancias que suelen relacionarse con la acción verbal, como el espacio y el tiempo.

La conclusión parcial se puede volver a formular, por tanto, en los términos del corolario que propusimos en el segundo experimento. Lo repetimos aquí en (111).

(111) A medida que crece la entropía en la distribución contextual de las categorías semánticas se enfatiza el papel o protagonismo semántico de las preposiciones en la interpretación de la tripleta, y viceversa.

Comprobamos, por consiguiente, que a mayor variación semántica más significado preposicional. La gradualidad preposicional, observada en la distribución contextual, se ha medido a través de la entropía. Empíricamente queda demostrada esa variación gradual y su incremento.

Como en el segundo experimento hemos sometido los datos a pruebas de azar. El método de organización de las tripletas varía respecto del segundo experimento porque no lo hemos hecho correlativo, añadiendo otro factor de azar.

El grupo 1 contiene el tercer tercio, alfabético y correlativo, de la clase funcional, más el primer tercio de la clase semiléxica y el segundo tercio de la clase léxica. En el Grupo 2 tenemos el primer tercio de la clase funcional, más el segundo tercio de la clase semiléxica y, por último el tercer tercio de la léxica. Y el último, el Grupo 3, tiene el segundo tercio de la funcional, el tercero de la semiléxica y el primero de la léxica.

La tabla 16 presenta los resultados obtenidos en la prueba.

Clases semánticas	Tercer experimento. Prueba de azar. Ámbito nominal (3903)		
	Grupo 1 (G1)	Grupo 2 (G2)	Grupo 3 (G3)
Animado	442	416	433
Inanimado	244	259	299
Entidad abstracta	191	248	281
Evento	97	130	116
Locativo	181	148	171
Temporal	83	67	77
Total de nombres al azar	1249	1278	1376
Entropía (H) =	2,36	2,37	2,37

Tabla 16. Resultados de entropía con criterios aleatorios de organización de las clasificaciones (ámbito nominal)

Nuevamente, y como sucedía entre las estructuras verbales, las diferencias entre los tres grupos son casi inexistentes y confirman que los resultados entrópicos no pueden conseguirse por combinación aleatoria de las clasificaciones.

Entre G1 y G2 la diferencia es de 0,008, mientras que entre G2 y G3 aumenta hasta 0,004. En ambos casos es una cifra irrelevante y confirmadora del fracaso del azar en la obtención de los resultados equivalentes.

Estas conclusiones parciales adquieren matices y aspectos complementarios si comparamos los nombres dependientes de un verbo, segundo experimento, con el actual.

3.3.5.7 Comparación de resultados del ámbito verbal (segundo experimento) y del ámbito nominal (tercer experimento)

En términos absolutos los resultados de la entropía, tal y como se aprecian en la tabla 17, son más bajos en las estructuras del ámbito verbal para las tres clases preposicionales.

Comparación de los resultados de entropía entre clases preposicionales y ámbitos sintácticos			
Clase preposicional Ámbitos	Funcional	Semifuncional / Semiléxica	Léxica
Entropía (H) Verbal	1,54	1,99	2,40
Entropía (H) Nominal	1,90	2,21	2,52

Tabla 17. Comparación de resultados entrópicos en clases preposicionales en entorno verbal y nominal

La primera conclusión conjunta de ambas partes del experimento es que la predicción sobre el significado de la preposición está relacionada con cómo se organizan las categorías semánticas de su contexto (sus predominios, hegemonías y contenidos) con independencia de las dependencias sintácticas y las restricciones seleccionales y/o argumentales que se establecen entre verbos y preposiciones, y nombres y preposiciones.

Así, en la tripleta que venimos manejando como modelo, reproducida en (112), las condiciones categoriales entre sus elementos no son determinantes.

(112) X - P - Z

Las relaciones formales y estructurales entre X y P, cuando X sea un nombre o un verbo, no condicionan nuestras predicciones sobre el grado, el peso o la influencia

semántica de la preposición en todo el conjunto. Este hecho, además, se relaciona con la sugerencia que indicamos en el estado del arte (secciones 1.5, 1.6 y 1.7 del Capítulo 1) sobre la necesidad de abordar la categoría preposicional, no ya desde su idiosincrasia sintáctico-estructural sino desde cuál sea su papel en la distribución, la obtención y la composición del significado. Lo que está siendo el centro de nuestra tarea y el eje de nuestra hipótesis.

En el ámbito verbal la entropía es más baja en las tres clases preposicionales. Es probable que este hecho esté relacionado con la naturaleza semántica del núcleo de la preposición: el verbo. Las restricciones seleccionales sobre cómo se articule el significado de los enunciados son más estrechas y matizables con la categoría verbal, frente a lo que sucede en el ámbito nominal donde la predicción combinatoria es más difícil y, por tanto, la entropía crece.

Si comparamos las diferencias entre las seis categorías semánticas de las dos estructuras, verbal y nominal, comprobamos que a medida que crece el significado de la clase de +F(uncional) a +L(éxica) la entropía (H) aumenta. La gradación, de menos a más, es la que indicamos en (113).

(113) $H [F(\text{Verbal})] < H [F(\text{Nominal})] < H [SF(\text{Verbal})] < H [SL(\text{Nominal})] < H [L(\text{Verbal})] < H [L(\text{Nominal})]$

Parecería que a medida que se produce una apertura en las restricciones estructurales (de Verbal a Nominal) la entropía de los grupos crece a la vez que las aportaciones preposicionales al significado global de la tripleta. Del mismo modo, además, la distancia entrópica entre las categorías semánticas de cada clase preposicional va disminuyendo, como se aprecia en la tabla 18.

Ámbito sintáctico	Clase preposicional		
	Funcional	Semifuncional / Semiléxica	Léxica
Nominal – Verbal	0,36	0,22	0,12

Tabla 18. Diferencia de entropía entre clases preposicionales y estructuras sintácticas por ámbitos

Esa disminución, cruzando resultados de estructuras y clases preposicionales, confirma que esos tres espectros de densidad o valor semántico son cada vez más parecidos con independencia de quién controle o dirija el significado de la preposición, sea un verbo o un nombre.

Pasemos ahora a la comparación global –entre las estructuras verbales y las nominales- de las seis categorías semánticas de los tres tipos de preposiciones. Para ello, reproducimos de manera conjunta los gráficos 22 y 23 de ambas partes, con la numeración correspondiente en el momento de su aparición, para facilitar el comentario y el análisis de algunas de esas cuestiones que son pertinentes en nuestro examen.

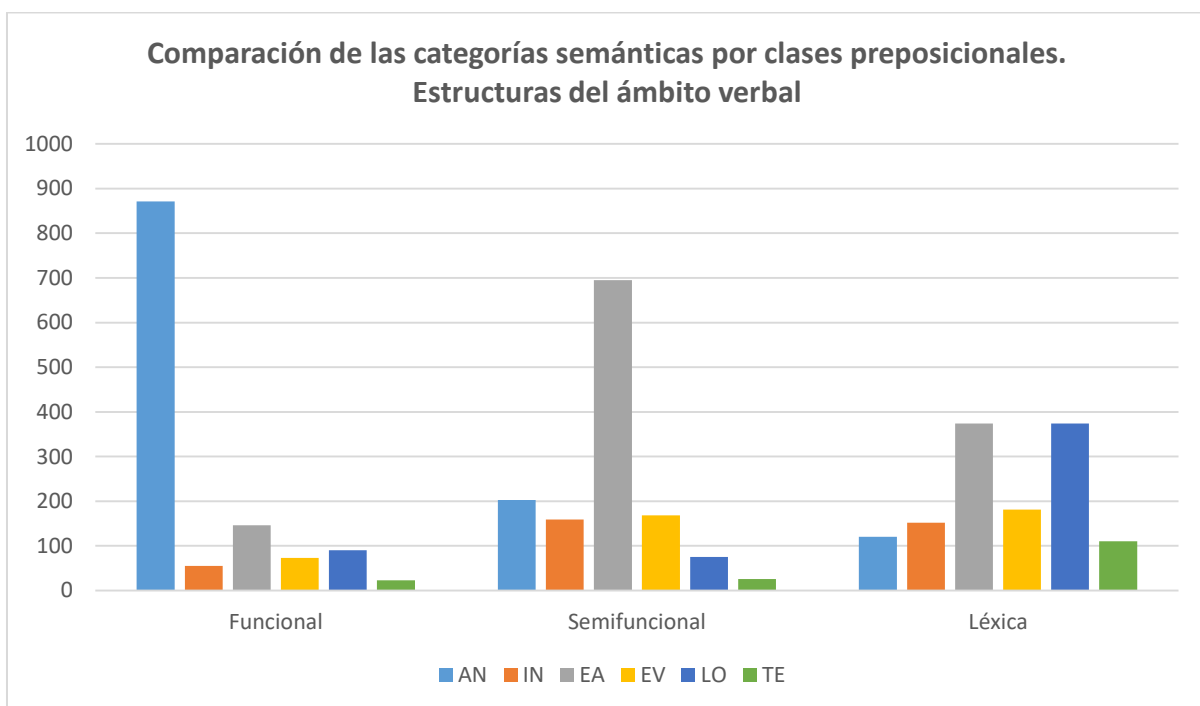


Gráfico 22. Comparación de las categorías semánticas para las clases preposicionales. Estructuras del ámbito verbal

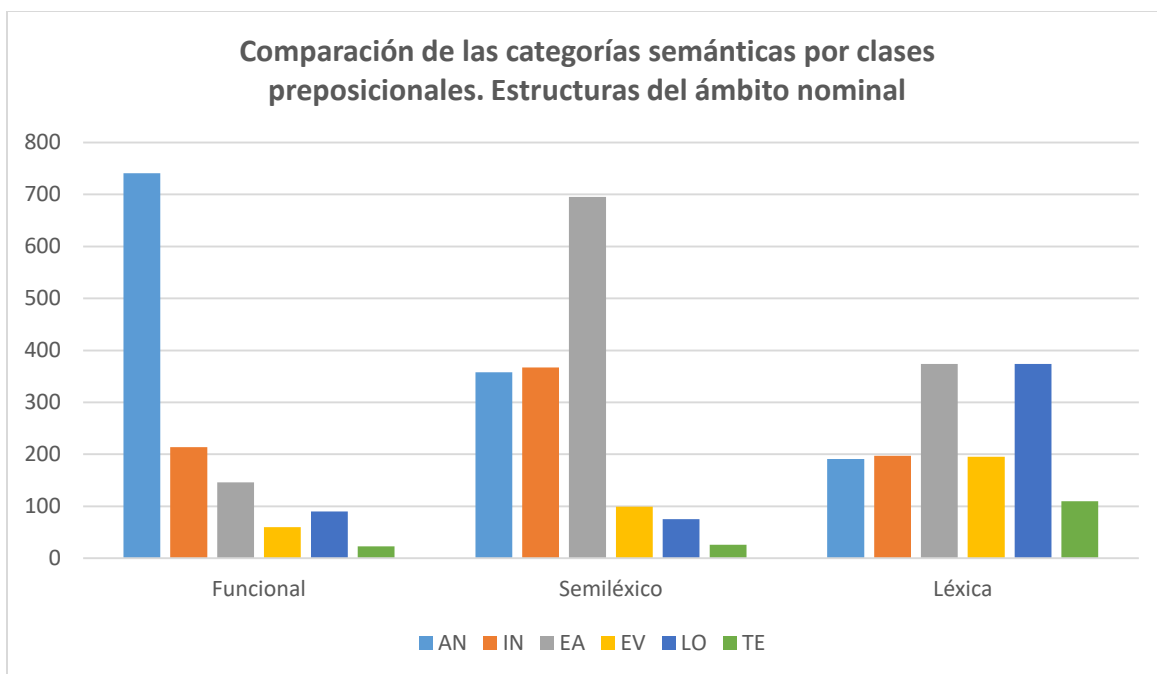


Gráfico 23. Comparación de las categorías semánticas por clases preposicionales. Estructuras del ámbito nominal

(Leyenda: AN: Animado, EA: Entidad abstracta, LO: Locativo, TE: Temporal, EV: Evento e IN: Inanimado)

Una mirada impresionista, conducida por la dispersión de los colores en los dos gráficos 22 y 23, permite afirmar que el proceso de expansión de las categorías es semejante en ambos contextos sintáctico-estructurales.

Pasamos del predominio mayoritario de una categoría semántica en sendos entornos para la clase funcional, *Animado*, a una progresiva aparición de equilibrios en el reparto de categorías semánticas.

En la clase semiléxica surgen, en las estructuras verbales, las *Entidades abstractas* como consecuencia de ciertas condiciones semánticas aplicadas por los verbos de régimen frente al crecimiento más repartido del ámbito nominal donde crecen simultáneamente *Animado*, *Entidad abstracta* e *Inanimado*.

A su vez, entre las léxicas germinan con vigor las categorías que expresan, matizan y se aproximan a dos núcleos universales de la circunstancialidad, es decir, el espacio y el tiempo en ambas estructuras. En el contexto verbal empujan dos categorías semánticas, *Entidad abstracta*, *Locativo* y, mientras que en las estructuras nominales hemos reseñado un crecimiento más unitario de *Animado*, *Inanimado*, *Entidad abstracta* y *Evento*. Esta uniformidad (apenas cuatro ocurrencias de diferencia) en cuatro grupos de seis reafirma el papel sustancial de la preposición en esos enunciados.

3.4 CUARTO EXPERIMENTO.

Los grados semánticos a partir del uso de las preposiciones

3.4.1 Objetivos y justificación del experimento

En el camino de confirmación de nuestra hipótesis vamos, en esta última prospección empírica, a buscar evidencias y mediciones para la gradualidad semántica de los significados preposicionales en los resultados obtenidos a partir del uso, y gracias a la comparación del acuerdo entre las propuestas de los datos del corpus –el modelo- y de los hablantes o usuarios.

A partir de la variación de los resultados, a causa de la inseguridad y la ambigüedad en el uso de los hablantes, intentaremos mostrar que el significado de las preposiciones se organiza de manera paulatina y expresa la diversidad polisémica léxico-semántica en cualquier tipo de estructura donde participa o aparece.

Este último experimento consiste en la ejecución de un test por parte de hablantes de la lengua, que deberán proponer la opción que consideran más adecuada. Con la verificación externa de hablantes añadimos un apoyo suplementario a la metodología probatoria de la investigación, sumándose al aprendizaje automático (Cap. 3, sección 3.1) y la medición de la entropía de las categorías semánticas de los contextos preposicionales en estructuras verbales y nominales (Cap. 3, secciones 3.2 y 3.3, respectivamente).

Nuestro objetivo es legitimar que cuanto más acierto o acuerdo se observa en el empleo de las preposiciones más fácil es la predicción del peso de su significado, o menor es su contribución al significado global del enunciado. La identificación de la semántica de la preposición sigue vinculada, como en toda nuestra investigación, a la naturaleza de su contexto. En este experimento la estabilidad o acuerdo en el uso será una nueva variable

analítica de la gradualidad preposicional. Asimismo, examinaremos el grado de dispersión en las respuestas como una medida que se añade a la incertidumbre en la elección de una preposición u otra. De modo que cuantas más posibilidades proponen (dispersión alta) más significado tienen, tal y como mostraremos. Esta dispersión, además, será interpretada a la luz de la entropía que ya usamos en el segundo y el tercer experimento, es decir, comprobaremos si existe alguna correlación entre la clase de preposición y el desorden entrópico de las opciones elegidas por los participantes del test.

A continuación describiremos cuestiones metodológicas previas, así como las partes del test en cuestión. Al final de la sección pasaremos a analizar los resultados representados gráficamente.

3.4.2 El uso de las preposiciones: de la ambigüedad al significado.

El uso de las preposiciones entraña ciertas dificultades en aquellos casos donde se produce una ambigüedad en la interpretación del enunciado, sea por proximidad semántica del contexto o sea por una interpretación diferente de los hablantes o usuarios de la lengua. Así, por ejemplo, en (114) y (115), extraídos del test que proponemos, podemos encontrar ambas circunstancias.

(114) Cree **_en_** su inocencia y está dispuesto a todo **_en / durante_** el procedimiento judicial

(115) Tres cartas fueron enviadas **_por / al_** el secretario judicial sin ningún remitente

La doble posibilidad de (114), entre *en* y *durante* manifiesta esta cierta capacidad interpretativa de la que disponen las piezas preposicionales. Esa elección no supone, en este caso, ningún error normativo sino dos lecturas que van más allá del estricto valor

léxico-semántico de la pieza. En (115), por otro lado, parece que el decantamiento hacia una u otra preposición depende del papel temático que se le atribuya al sintagma nominal *secretario judicial*.

Ha sido en las áreas que se ocupan de la adquisición lingüística y del aprendizaje de lenguas extranjeras donde es habitual destacar las dificultades en el uso como objetivos de análisis de estas subdisciplinas lingüísticas, apelando al carácter específico e idiosincrático de las preposiciones en la expresión propia de cada lengua natural (Pérez Pereira, 2004; Fernández López, 1995; Perea, 2007; y Alonso y Fernández, 2011).

En la demostración de nuestra hipótesis las cuestiones sugeridas y los propósitos de esas perspectivas, la adquisición y la didáctica del ELE, solo nos resultan necesarios de una manera complementaria y accesoria, es decir, como un incremento empírico de los argumentos ya expuestos en los experimentos anteriores (Fernández Vallejo, 1994; Fernández López, 1995; Hernando Cuadrado, 2002; Campillos, 2014; y Rakotovoavy, 2018). No en vano, los participantes del test que conforma el experimento tienen como lengua materna o de uso habitual el español.

De este modo, las variaciones, los desacuerdos, la diversidad o las inadecuaciones en el uso de unas preposiciones conducen nuestra justificación desde este ángulo hacia una de las nociones clave de la propuesta: el peso semántico de las preposiciones y la manifestación empírica de su gradualidad.

Nuestras consideraciones en la prospección de los datos y los resultados obtenidos en la realización del test no se ocuparán, tampoco, de cuestiones ni valoraciones normativas, estilísticas o de corrección. En este sentido a las diversas posibilidades que proponen los participantes las consideramos como de acuerdo o no con el modelo sin entrar en juicios o explicaciones metodológicamente ajenas a nuestra heurística de trabajo. De todos modos, conviene recordar las referencias descriptivas, que ya mencionamos en las asunciones descriptivas, Capítulo 2, sección 2.2, página 70 (López 1970; Trujillo, 1971; Luque, 1973; Morera, 1984; García Yebra, 1988; Osuna, 1991; Nández, 1995; Cifuentes, 1996a; Fernández, 1999; Bosque & Demonte, 1999; *NGLE*, 2009; y Moliner, 2012). Junto

a ellas, también se pueden citar algunos manuales de construcción, redacción y estilo que se ocupan de cuestiones sobre el uso (Paredes et alia, 2012 y 2013; y RAE, 2013).

3.4.3 Metodología: diseño y características del test de uso preposicional

Se ha elaborado un test dividido en cuatro partes de manera aleatoria, en cantidad de ocurrencias (el número de veces que se debe completar un espacio en blanco con una preposición) y en las clases de preposiciones de nuestra hipótesis (funcionales, semifuncionales, semiléxicas y léxicas). El objetivo de esa división era facilitar la tarea a los participantes, en tiempo y en calidad de atención. Con un único test hubiesen sido 40 enunciados con 90 espacios en blanco (!).

Hemos elegido la casi totalidad de las preposiciones (13 de 19), a excepción de algunas cuyo significado es muy previsible, como *desde*, *durante*, *mediante*, *según* o *ante*. Esta facilidad en la adecuación y en la coincidencia aportaría poca información a nuestro propósito y desviaría, probablemente, los resultados.

El objetivo central en el diseño del test ha sido favorecer la diversidad y la representatividad de los valores léxico-semánticos y de las estructuras sintáctico-argumentales de las preposiciones en español.

En la tabla 19 se recogen los datos descriptivos del test que, como decimos, pretenden abarcar el perímetro semántico mayoritario de las preposiciones.

Distribución de ocurrencias, valores léxico semánticos y clases preposicionales			
Lista de preposiciones que aparecen en los test	Ocurrencias totales de cada una de las preposiciones	Valores léxico-semánticos asociados (repetidos)	Número de valores asociados a las clases F, SF-SL y L
A	21	14	14 F / 1 SF / 6 L
BAJO	1	3	1 L
CON	6	4	3 SF / 2 SL / 1 L
DE	18	14	3 F / 4 SF / 9 SL / 2 L
EN	21	8	1 F / 4 SF / 2 SL / 14 L
ENTRE	1	6	1 SL
HACIA	2	4	1 SF / 1 L
HASTA	2	3	1 SL / 1 L
PARA	6	6	1 SF / 4 SL / 1 L
POR	8	13	3 F / 2 SF / 2 SL / 1L
SIN	2	1	2 SL
SOBRE	1	6	1 SF
TRAS	1	3	1 L
TOTALES	<i>PREPOSICIONES</i> 13 <i>OCURRENCIAS</i> 90 (Por partes: P1 22, P2 18, P3 25 y P4 25)		<i>Funcionales</i> 21 <i>Semifuncionales</i> 17 <i>Semiléxicos</i> 21 <i>Léxicos</i> 31

Tabla 19. Distribución de ocurrencias, valores y clases preposicionales en el test

La distribución de las clases preposicionales en cada una de las partes no ha seguido una planificación determinada, apoyada en una estrategia o previsión metodológica. La única restricción que hemos aplicado es el predominio de la representatividad en las preposiciones más comunes y frecuentes (*a, de, en y por*) y sus valores léxico-semánticos más diversos y cercanos. Hemos considerado que cuanto más arbitrario resultase el diseño del test más contribuiría a la validación y la significación de las conclusiones en la medida en que se asemeja más a la realidad. Esa arbitrariedad guarda relación con el orden de las partes, con la distribución de los valores elegidos, con la repetición de preposiciones, con

los contenidos de los enunciados de los ejemplos, y, por supuesto, con las características personales de los participantes.

3.4.3.1 Procedencia de los datos, procedimiento de realización y rasgos de los participantes

Los datos proceden del corpus Ancora, ya utilizado en el primer experimento, <http://clic.ub.edu/corpus/es/ancora>, adaptándose algunos enunciados para facilitar el acceso. El motivo de su elección tiene que ver con el tipo de documentos que lo componen: los textos periodísticos. El registro de la lengua en los medios de comunicación escrita acostumbra a seguir un cierto estándar que, debemos suponer, unifica las posibles vacilaciones o dudas sobre el uso preposicional. Se trata de un contexto semántico no restringido, no pertenecen a un dominio concreto, sino que incluyen registro escrito con un carácter muy heterogéneo. Por otro parte, y por su disponibilidad y su volumen, es fácil realizar un corpus de prensa ya que cada día se generan hechos lingüísticos de este registro.

Se han obtenido 626 respuestas en las cuatro partes del test (142, 187, 152 y 145, respectivamente). Esas respuestas las han realizado un total de 366 participantes diferentes. Muchos de ellos, han podido responder 1, 2, 3 o 4 de los test. Esta diversidad en la participación, así como la relativa libertad de opción, pretendía favorecer la calidad de los resultados en la medida en que requieren un tiempo de atención. Al tratarse de una colaboración voluntaria consideramos que era preferible favorecer las condiciones materiales de participación.

El proceso de realización del test ha sido el siguiente.

Se han enviado cuatro enlaces de *google formularios* para el acceso a los participantes que han procedido a responderlos sin ninguna ayuda ni indicación. Las instrucciones recibidas en la cabecera de los formularios advertían que las respuestas debían ser únicas. También se señalaba que se trataba de una indagación sobre el uso de

las preposiciones sin entrar en más detalles que dirigiesen su atención o la intención de nuestra hipótesis. No había ningún control de duración del tiempo.

Durante tres meses, entre el 18 de febrero de 2019 y el 20 de mayo de 2019, se han recogido las respuestas que se han guardado en los formularios. Una vez cerrada la posibilidad de participar se descargan las hojas con las respuestas y se preparan los datos para el vaciado, la observación, el comentario y el análisis final.

Respecto al perfil sociolingüístico de los participantes se recogen, en el cuadro 30, las variables descriptivas más representativas.

Variables sociolingüísticas de los perfiles de los participantes en el test	
Sexo	Predominio de mujeres respecto de hombre (60 % frente al 40 %)
Edad	A partir de 16 años
Formación, estudios y/o actividad profesional	Educación secundaria (obligatoria y postobligatoria) y universitaria de todos los ámbitos (desde ingenierías hasta filologías, pasando por ciencias físicas, por ejemplo)
Dedicación y actividad profesional	Estudiantes, docentes, profesiones liberales, administrativas y trabajadores manuales
Lugares de residencia	Cataluña y algunas ciudades de España (Madrid, Sevilla y Huelva)
Lenguas maternas y de uso habitual	Español, catalán y otras posibles (sin contabilizar pero sin ser predominantes)

Cuadro 30. Descripción de las variables sociolingüísticas de los participantes del test

Algunas de las referencias que se ocupan de los efectos, la pertinencia y la influencia de las variables anteriores en relación a cómo estudiar los hechos lingüísticos han sido objeto de la sociolingüística en los últimos cuarenta años (Hudson, 1980; Fasold, 1990; Coulmas, 1996; Moreno Fernández, 1998; y Bayley et alia, 2013).

3.4.3.2 El test: información sobre las respuestas

A continuación encontramos las cuatro partes del test con los enunciados y los espacios en blanco completados con la preposición esperada (la que procede del corpus de referencia que ya presentamos en la sección 2.4.4.1, página 93, y que volvemos a recordar en el apartado siguiente). En la parte final de los enunciados se halla una abreviatura de la pertenencia a una de las cuatro clases de preposiciones de nuestra hipótesis, F(uncional), SemiF(uncional), SemiL(éxica), y L(éxica). Esa información responde a las clasificaciones de los valores léxico-semánticos que presentamos en las asunciones (Cap. 2, sección 2.1). En último lugar, también se añade el tipo de estructura o núcleo del que depende el sintagma preposicional, sea nominal o verbal.

Cada parte del test presenta un recuadro inicial con una información para facilitar el seguimiento, lo que podríamos denominar el “modelo”. Se trata del número de preposiciones usadas en cada test, el número de ocurrencias total y el número de ocurrencias para cada preposición.

Parte 1. 8 preposiciones. 22 ocurrencias distribuidas en A 4, BAJO 1, DE 3, CON 3, EN 6, ENTRE 1, POR 2, SIN 1 y SOBRE 1.
--

1. Han requerido a los asistentes de la ceremonia. **F / L. 1 verbal y 1 nominal.**
2. Aunque no se complacen en atacar a los visitantes, su actitud es claramente impropia de su posición. **SF / F. 1 verbal y 1 nominal.**

3. El fiscal no cesa **_de_** insistir en la idea presentada **_en_** las cuestiones preliminares. La sala ha respondido con un murmullo de contrariedad cuando ha comparado la situación **_con_** Corea del Norte. **SF / L / SF. 2 nominales y 1 verbal.**
4. Se han dirigido **_a_** la salida **_sin_** atender las advertencias del regidor del plató. **L / SL. 1 verbal y 1 nominal.**
5. Lo realizaron **_entre_** todos los presentes que estuvieron regidos **_por_** una serie de especialistas **_en_** la materia. **SL / F / SF. 1 verbal y 2 nominales.**
6. Los turistas han dormido **_en_** el hotel que les recomendaron **_a_** la salida del aeropuerto. **L / L. 2 verbales.**
7. Contaban **_con_** su confianza aunque ya habían tenido alguna discrepancia **_en_** los últimos meses. **SF / L. 2 verbales.**
8. Los guardias velan **_por_** la seguridad de los cardenales **_bajo_** la atenta supervisión del capitán suizo. **SF / L. 2 verbales.**
9. Se divierten **_con_** la reproducción de los videos sobre caídas y ocurrencias. Las carcajadas se oyen **_en_** la habitación **_de_** abajo. **SF / L / SL. 2 verbales y 1 nominal.**
10. No han meditado **_sobre_** las consecuencias que tendrá la suspensión del presupuesto. **SF. 1 verbal.**

Parte 2. 8 preposiciones. 18 ocurrencias distribuidas en A 4, CON 1, DE 3, EN 3, HACIA 1, HASTA 1, PARA 3 y POR 2.

1. Han abastecido **_de_** agua y medicinas imprescindibles a los primeros refugiados. **SF. 1 verbal.**
2. Creen que recogerán **_a_** los pasajeros en la terminal B. **F. 1 verbal.**
3. Tres cartas fueron enviadas **_por_** el secretario judicial sin ningún remitente. **F. 1 verbal.**
4. Nos inclinamos **_hacia_** esa nueva opinión porque ha presentado argumentos más convincentes que los anteriores. **SF. 1 verbal.**
5. Los soldados fueron reprendidos **_en_** la ceremonia **_por_** su conducta grosera y ostensiva. **L. / SL. 1 verbal y 1 nominal.**
6. Este memorial se ha escrito **_para_** nuestro querido profesor. **SF. 1 verbal.**

7. Se han apoyado **_en_** una teoría antigua **_para_** su estudio **_de_** la célula más importante del hígado. El público ha preguntado con mucho entusiasmo durante el turno de intervenciones. **SF. / SL. / SL. 1 verbal y 2 nominales.**

8. El precio se ha elevado **_hasta_** cotas inimaginables **_en_** el mercado energético. **SL. / L. 1 verbal y 1 nominal.**

9. No tienen credibilidad **_para_** convencer **_a_** los votantes ya que tienen un pasado probado **_de_** corrupción. **SL. / F. / SL. 1 verbal y 2 nominales.**

10. Viajaron **_a_** Melilla **_con_** la intención de solucionar el problema de la herencia materna. **_Al_** comprobar la actitud de sus primos supieron que sería muy fácil. **L. / SL. / L. 2 verbales y 1 nominal.**

Parte 3. 6 preposiciones. 25 ocurrencias distribuidas en A 9, DE 5, EN 7, HASTA 1, POR 2 y TRAS 1
--

1. En el intermedio de ese programa de televisión se acostumbra **_a_** servir un refrigerio a base de patatas y aceitunas. **A_1** final se avisa al público por donde deben salir. **F. / L. 2 verbales.**

2. Acabaron **_por_** interesarse en la cuestión que planteó el diputado **_en_** su intervención. **SF. / L. 2 verbales.**

3. He mirado dos veces pero creo que está **_en_** la mesa. **L. 1 verbal.**

4. El economista reconoció que estaba preocupado **_por_** el elevado valor de la libra que incitaba **_a_** numerosas empresas a abandonar el país debido a la pérdida de competitividad de sus negocios. Y declaró: "**_Hasta_** que no se invierta esa tendencia monetaria no habrá nada que hacer". **SL. / F. / L. 2 nominales y 1 verbal.**

5. Hemos ido **_de_** Lugo **_a_** Santiago **_en_** tres horas. **L. / L. / L. 1 verbal y 2 nominales.**

6. Una serie **_de_** hombres y mujeres **_de_** apariencia burguesa, demacrados y parapetados **_tras_** sus gafas de sol, consumen las últimas gotas **_de_** alcohol de una juerga nocturna junto **_a_** una especie de piscina. **F. / SL. / L. / SL. / F. 4 nominales y 1 verbal.**

7. No abandonará **_a_** su sobrino a pesar de las condiciones económicas que le exige el Departamento de Hacienda del condado. Cree **_en_** su inocencia y está dispuesto a todo **_en_** el procedimiento judicial. **F. / SF. / L. 2 verbales y 1 nominal.**

8. Los atletas se han cansado de esperar la llegada **_de_** equipo de masajistas y han empezado **_a_** calentar y estirar en la pista habilitada para ello. **F. / F. 2 verbales.**

9. Los más destacados lutieres (constructores de violines) del mundo se dan cita **_en_** Valencia para rendir tributo **_a_** los instrumentos de cuerda. **L. / SF. 1 verbal y 1 nominal.**

10. Practicaron un examen **_en_** profundidad **_a_** los afectados. **SL. / F. 1 nominal y 1 verbal.**

Parte 4. 8 preposiciones. 25 ocurrencias distribuidas en A 5, DE 7, EN 6, HACIA 1, PARA 3, POR 2 y SIN 1

1. Te entregaron el libro **_de_** física **_para_** revisar. **SL / SL. 2 nominales**

2. Chávez se despidió del mundanal ruido el pasado jueves **_en_** su más breve discurso **_en_** televisión, tildado de "institucional" por los observadores, en el que aceptó el aplazamiento de los comicios y valoró la " pedagógica " medida como un triunfo de la democracia venezolana. **L. / SL. 1 verbal y 1 nominal.**

3. Dirígete **_hacia_** el norte un kilómetro y luego vuelves **_a_** preguntar **_en_** la rotonda. **L. / F. / L. 1 nominales y 2 verbal.**

4. Todavía le quedaron fuerzas **_para_** recabar fondos **_en_** la gala de entrega de premios. **SL. / L. 2 nominales.**

5. Abusan **_de_** la confianza **_de_** sus criados. **SF. / SL. 1 nominal y 1 verbal.**

6. Han acompañado **_a_** los alumnos **_a_** la vacunación **_sin_** petición de la directora. **F. / L. / SL. 3 verbales.**

7. Quiénes abanderaron **_a_** los comuneros **_en_** las jornadas **_de_** lucha ahora estaban buscando pactos **_con_** las autoridades que intentaron imponer aquellas condiciones de trabajo. **F. / L. / SL. / SL. 2 verbales y 2 nominales.**

8. **_Para_** mañana es imposible que tengamos hechos los ejercicios. **L. 1 verbal.**

9. Los encausados carecen **_de_** defensa pública. De manera que el juez ha propuesto un aplazamiento **_a_** la espera **_de_** que se solucione la cuestión. **SF. / F. / F. 2 verbales y 1 nominal.**

10. Alcanzaron un acuerdo **_de_** infraestructuras que fue ratificado **_por_** la cámara el jueves **_por_** la tarde. **SL. / F. / L. 1 verbal y 2 nominal.**

3.4.4 Descripción de los resultados del test

En esta sección describimos, a partir de la información representada en las tablas 20, 21, 22 y 23, la información sobre las 626 respuestas obtenidas.

En las cabeceras se indican, por orden de columna, (i) el número de la pregunta con su respuesta adecuada, (ii) las series de preposiciones respondidas con sus ocurrencias (la primera cifra en negrita son las acertadas y la que establece el porcentaje de acierto), (iii) la clase de preposición a la que pertenece según nuestra hipótesis (cuatro colores de identificación), (iv) el porcentaje de acuerdo, (v) el nivel de dispersión de la respuesta (de 1 a 19), y (vi) la medida de la entropía de la dispersión.

El nivel de dispersión de la respuesta es, tal y como dijimos en los objetivos del experimento, la cantidad de preposiciones de más que los participantes han propuesto en los espacios en blanco. Es un valor de 1 a 19, el total de preposiciones del repertorio disponible (teniendo en cuenta que los participantes no sabían que solo hemos usado 13 de ellas). Este aspecto o variable, que también vamos a usar como medida de nuestra predicción, se relaciona, como veremos en el análisis, con la interpretación del enunciado y con la naturaleza de la clase preposicional en cuestión.

Por último, en la columna seis (vi), el número de ocurrencias en cada una de las opciones que se proponen se mide con la fórmula de la entropía que hemos utilizado a lo largo de la investigación (experimentos 3.2 y 3.3), y que reproducimos aquí en (116)

(116)

Parte 1. Resultados de las respuestas identificados por número, respuestas y porcentaje de acuerdo

Total de respuestas recibidas, 142

F(3) SF(8) SL (3) L(8)

Número de pregunta y preposición más adecuada (i)	Respuestas ordenadas de más a menos (ii)	Clase de preposición según nuestra hipótesis (iii)	Porcentaje de acuerdo (iv)	Dispersión de la respuesta (número de opciones distintas) (v)	Medida de la entropía de la dispersión (vi)
1.1 A	123/ 16 DE / 1 HASTA / 1 SIN / 1 HACIA	FUNCIONAL	86,61 %	4	0,68
1.2 DE	61 / 38 PARA / 22 EN / 20 A / 1 HASTA	LÉXICA	42,25 %	4	1,94
2.1 EN	94 / 23 CON / 13 EN / 6 PARA / 3 A / 2 SOBRE / 1 ANTE / 1 DURANTE	SEMIFUNCIONAL	66,19 %	7	1,53
2.2 A	137/ 2 CON / 1 CONTRA / 1 HACIA / 1 DE	FUNCIONAL	96,47 %	4	0,28
3.1 DE	84 / 40 EN / 12 A / 4 PARA / 1 CON / 1 HACIA	SEMIFUNCIONAL	57,74 %	5	1,43
3.2 EN	71 / 22 DE / 15 PARA / 13 SOBRE / 12 A / 9 POR	LÉXICA	50 %	5	2,12
3.3 CON	86 / 37 DE / 16 EN / 1 ANTE / 1 A / 1 SOBRE	SEMIFUNCIONAL	60,56 %	5	1,62

4.1 A	96 / 44 HACIA / 2 HASTA / 1 ANTE / 1 CON	LÉXICA	67,60 %	4	1,24
4.2 PARA	91 / 23 TRAS / 16 A / 10 SIN / 2 CON	SEMILÉXICA	64,08 %	4	1,70
5.1 ENTRE	49 / 36 CON / 27 PARA / 16 ANTE / 13 EN / 1 POR	SEMILÉXICA	34,50 %	5	2,29
5.2 POR	101 / 19 EN / 16 A / 4 CON / 1 SEGÚN / 1 SOBRE	FUNCIONAL	71,12 %	5	1,56
5.3 EN	93 / 35 DE / 4 CON / 5 SOBRE / 3 POR / 2 PARA	SEMIFUNCIONAL	65,49 %	5	1,48
6.1 EN	133 / 8 DURANTE / 1 SOBRE	LÉXICA	93,66 %	2	0,45
6.2 A	89 / 34 EN / 10 TRAS / 4 DE / 4 DESDE / 1 PARA	LÉXICA	62,67 %	5	1,54
7.1 CON	134 / 4 DE / 2 EN / 1 BAJO / 1 DESDE	SEMIFUNCIONAL	94,36 %	4	0,56
7.2 EN	106 / 25 DURANTE / 8 CON / 2 TRAS 1 / ENTRE /	LÉXICA	74,64 %	4	1,17
8.1 POR	122 / 7 PARA / 8 A / 3 EN / 1 SEGÚN / 1 HACIA	SEMIFUNCIONAL	85,91 %	5	0,81

8.2 BAJO	60 / 34 EN / 21 ANTE / 10 HASTA / 8 POR / 2 DE / 2 A / 2 HACIA / 2 MEDIANTE / 1 DESDE	LÉXICA	42,25 %	9	2,28
9.1 CON	112/ 21 EN / 3 POR / 3 DURANTE / 2 TRAS / 1 DE	SEMIFUNCIONAL	78,87 %	5	1,04
9.2 DESDE	84 DESDE / 39 EN / 10 HASTA / 4 POR / 4 DE	LÉXICA	59,15 %	5	1,60
9.3 DE	117 / 12 EN / 8 HASTA / 3 POR / 1 PARA / 1 HACIA	SEMILÉXICA	82,39 %	5	1,01
10 SOBRE	93 / 30 EN / 5 A / 5 ANTE / 4 CON / 3 SEGÚN / 2 POR	SEMIFUNCIONAL	65,49 %	6	1,47

Tabla 20. Parte 1. Resultados de las respuestas identificadas por número, respuestas y porcentaje de acuerdo

<p align="center">Parte 2. Resultados de las respuestas identificadas por número, respuestas y porcentaje de acuerdo</p> <p align="center">Total de respuestas recibidas, 187</p> <p align="center">F(3) SF(4) SL (7) L(4)</p>					
Número de pregunta y respuesta más adecuada	Respuestas ordenadas de más a menos	Clase de preposición según nuestra hipótesis	Porcentaje de acuerdo %	Dispersión de la respuesta (número de opciones distintas)	Entropía de la dispersión del uso
1 DE	122/ 55 CON / 5 EN / 3 ENTRE / 1 ANTE / 1 TRAS	SEMIFUNCIONAL	65,24 %	5	1,11

2 A	183 / 1 PARA / 1 ENTRE / 1 POR / 1 CON	FUNCIONAL	97,86 %	4	0,19
3 POR	89 / 58 A / 23 PARA / 16 HACIA / 1 ANTE	FUNCIONAL	47,05 %	4	1,74
4 HACIA	65 / 47 POR / 32 ANTE / 20 A / 17 EN / 2 SOBRE / 2 PARA / 2 BAJO	SEMIFUNCIONAL	34,75 %	6	2,19
5.1 EN	89 / 32 A / 21 DE / 18 POR / 13 DURANTE / 8 TRAS / 3 CON / 2 CONTRA / 1 SOBRE	LÉXICA	47,59 %	8	2,03
5.2 POR	121 / 52 DE / 8 CON / 3 EN / 2 TRAS / 1 SEGÚN /	SEMILÉXICA	64,70 %	5	1,52
6 PARA	107 / 43 POR / 19 A / 5 CON / 5 SOBRE / 4 DE / 4 HACIA	SEMIFUNCIONAL	57,21%	6	1,68
7.1 EN	140 / 25 A / 9 SOBRE / 7 DE / 5 CON / 1 TRAS	SEMIFUNCIONAL	72,19 %	5	1,29
7.2 PARA	58 / 49 DE / 32 EN / 20 SOBRE / 13 POR / 8 A / 4 CON / 3 BAJO	SEMILÉXICA	31,01 %	7	2,32
7.3 DE	91 / 67 SOBRE / 12 EN / 5 HACIA / 5 POR / 4 CON / 2 SEGÚN / 1 MEDIANTE	SEMILÉXICA	43,31 %	7	1,84

8.1 HASTA	97 / 40 HASTA / 21 CON / 16 EN / 7 POR / 2 HACIA / 2 ANTE / 1 HASTA / 1 BAJO	SEMILÉXICA	51,87 %	8	2,06
8.2 EN	130 EN / 19 POR / 15 DE / 10 SOBRE / 3 ANTE / 3 TRAS / 2 CON / 2 HASTA / 2 PARA / 1 SIN	LÉXICA	69,51 %	10	1,44
9.1 PARA	132 / 21 DE / 15 A / 14 EN / 3 POR / 1 SOBRE / 1 TRAS	SEMILÉXICA	70,58 %	6	1,41
9.2 A	175 / 9 PARA / 1 BAJO / 1 CON / 1 EN	FUNCIONAL	93,58 %	4	0,42
9.3 DE	95 / 47 POR / 19 EN / 9 SOBRE / 7 MEDIANTE / 6 ENTRE / 3 SIN / 1 ANTE	SEMILÉXICA	50,80 %	7	1,96
10.1 A	152 / 19 HASTA / 9 DESDE / 5 HACIA / 2 DE	LÉXICA	81,28 %	4	1,07
10.2 CON	166 / 13 POR / 4 PARA / 2 TRAS / 2 EN / 1 DESDE	SEMILÉXICA	88,77 %	5	0,73
10.3 A	49 / 46 TRAS / 39 PARA / 26 EN / 15 SIN / 9 CON / 2 POR / 1 MEDIANTE	LÉXICA	26,20 %	7	2,29

Tabla 21. Parte 2. Resultados de las respuestas identificadas por número, respuestas y porcentaje de acuerdo

Parte 3. Resultados de las respuestas identificados por número, respuestas y porcentaje de acuerdo
Total de respuestas recibidas, 152
F(8) SF(3) SL(4) L(10)

Número de pregunta y respuesta más adecuada	Respuestas ordenadas de más a menos	Clase de preposición según nuestra hipótesis	Porcentaje de acuerdo %	Dispersión de la respuesta (número de opciones distintas)	Entropía de la dispersión en el uso
1.1 A	141 / 9 EN / 2 PARA / 1 HACIA	FUNCIONAL	92,76 %	3	0,48
2.1 A	52 / 38 EN / 32 HACIA / 23 TRAS / 2 POR / 2 DURANTE / 1 SOBRE / 1 PARA / 1 DESDE	LÉXICA	34,21 %	6	2,25
2.2 POR	106 / 40 DE / 3 EN / 2 SIN / 1 CON	SEMIFUNCIONAL	69,73 %	4	1,16
2.3 EN	95 / 24 DURANTE / 12 TRAS / 11 DE / 4 CON / 4 PARA / 2 ANTE	LÉXICA	63,33 %	6	1,65
3 EN	76 / 62 SOBRE / 9 POR / 4 DEBAJO / 1 A	LÉXICA	52,41 %	4	1,52
4.1 POR	134 / 10 DE / 3 EN / 2 ANTE / 2 CON / 1 SOBRE	SEMILÉXICA	88,15 %	5	0,70
4.2 A	141 / 5 EN / 3 SOBRE / 2 DE / 1 DURANTE	FUNCIONAL	92,76 %	4	0,50
4.3 HASTA	104 / 16 PARA / 9 A / 8 MIENTRAS	LÉXICA	69,33 %	10	1,39

	/ 6 POR / 5 SIN / 1 DE / 1 MEDIANTE / 1 VERSUS / 1 DURANTE / 1 SIN				
5.1 DE	88 / 43 A / 16 DESDE / 3 POR / 1 HACIA / 1 EN	LÉXICA	58,66 %	5	1,59
5.2 A	97 / 32 POR / 12 HASTA / 8 DESDE / 2 DE / 1 CON	LÉXICA	64,66 %	5	1,66
5.3 EN	105 / 35 DURANTE / 8 A / 1 DESDE / 1 HASTA / 1 POR / 1 TRAS	LÉXICA	72,41 %	6	1,34
6.1 DE	145 / 4 PARA / 2 ENTRE / 1 HACIA	FUNCIONAL	98,66 %	3	0,38
6.2 DE	89 / 53 CON / 7 EN / 1 / 1 HACIA / 1 BAJO	SEMILÉXICA	58,55 %	4	1,49
6.3 TRAS	81 / 34 CON / 17 EN / 8 BAJO / 5 POR / 4 DE / 2 A	LÉXICA	53,28 %	6	1,81
6.4 DE	139 / 9 A / 3 BAJO / 1 CON	SEMILÉXICA	91,44 %	3	0,52
6.5 A	144 / 3 CON/ 3 DE / 2 EN	FUNCIONAL	94,73 %	3	0,43
7.1 A	149 / 3 CON	FUNCIONAL	98,02 %	1	0,22
7.2 EN	146 / 6 SOBRE	SEMIFUNCIONAL	96,05 %	1	0,46
7.3 EN	50 / 32 DURANTE / 25 ANTE / 17 POR / 6 MEDIANTE / 6 CONTRA / 5 PARA / 4 TRAS / 2	LÉXICA	32,89 %	12	2,15

	HASTA / 1 CON / 1 SOBRE / 1 SEGÚN / 1 A / 1 DESDE				
8.1 DE	143 / 9 POR	FUNCIONAL	94,07 %	1	0,45
8.2 A	145 / 5 TRAS / 2 POR	FUNCIONAL	95,39 %	2	0,37
9.1 EN	141 / 8 PARA / 1 A / 1 DE / 1 DESDE	LÉXICA	92,76 %	4	0,45
9.2 A	133 / 10 CON / 7 DE / 2 EN	SEMIFUNCIONAL	87,5%	3	0,75
10.1 EN	109 / 32 DE / 6 CON / 5 A	SEMILÉXICA	71,71 %	3	1,37
10.2 A	96 / 21 DE / 13 PARA / 9 POR / 7 SOBRE / 6 ENTRE	FUNCIONAL	63,15 %	5	1,56

Tabla 22. Parte 3. Resultados de las respuestas identificadas por número, respuestas y porcentaje de acuerdo

<p>Parte 4. Resultados de las respuestas identificados por número, respuestas y porcentaje de acuerdo Total de respuestas recibidas, 145 F(6) SF(2) SL(8) L(9)</p>					
Número de pregunta y respuesta más adecuada	Respuestas ordenadas de más a menos	Clase de preposición según nuestra hipótesis	Porcentaje de acuerdo %	Dispersión de la respuesta (número de opciones distintas)	Entropía de la dispersión en el uso
1.1 DE	128 / 17 SOBRE	SEMILÉXICA	89,51 %	1	0,70
1.2 PARA	110 / 22 SIN / 5 A / 5 POR	SEMILÉXICA	75,86 %	3	1,23

2.1 CON	70 / 52 EN / 17 TRAS / 2 MEDIANTE / 2 DE / 1 HASTA / 1 ANTE	LÉXICA	48,95 %	6	1,62
2.2 EN	82 / 35 DE / 26 POR / 1 SOBRE / 1 PARA	LÉXICA	57,34 %	4	1,63
3.1 HACIA	125 / 10 POR / 5 A / 1 DESDE / 1 PARA / 1 EN / 1 CON / 1 HASTA	LÉXICA	87,41 %	7	0,88
3.2 A	125 / 13 PARA / 2 SIN / 1 HACIA / 1 HASTA	FUNCIONAL	86,20 %	3	0,73
3.3 EN	78 / 58 POR / 5 SOBRE / 2 A / 1 PARA / 1 HASTA	LÉXICA	54,54 %	5	1,43
4.1 PARA	128 / 12 TRAS / 3 EN / 2 DE	SEMILÉXICA	89,51 %	3	0,66
4.2 EN	70 / 44 PARA / 20 DE / 6 DURANTE / 2 POR / 1 HASTA / 1 SOBRE / 1 TRAS	LÉXICA	48,95 %	7	1,79
5.1 DE	129 / 10 CON / 3 ANTE / 1 BAJO / 1 SOBRE / 1 CONTRA	SEMIFUNCIONAL	88,96 %	5	0,64
5.2 DE	120 / 16 CON / 4 HACIA / 2 SOBRE / 1 A / 1 EN / 1 ENTRE	SEMILÉXICA	82,75 %	5	0,84
6.1 A	143 / 2 POR	FUNCIONAL	98,62 %	1	0,29

6.2 A	84 / 33 PARA / 18 EN / 3 DE / 2 HASTA / 2 DURANTE / 2 HACIA / 1 POR	LÉXICA	58,74 %	7	1,57
6.3 SIN	73 / 61 POR / 5 DE / 3 CON / 1 CONTRA / 1 EN / 1 SEGÚN	SEMILÉXICO	50,34 %	6	1,48
7.1 A	136 / 5 EN / 1 ANTE / 1 CABE / 1 POR / 1 DE	FUNCIONAL	92,41 %	5	0,45
7.2 EN	85 / 32 DURANTE / 15 DE / 4 PARA / 4 POR / 2 A / 1 HACIA / 1 ENTRE / 1 TRAS	LÉXICA	59,44 %	8	1,70
7.3 DE	113 / 17 EN / 5 POR / 4 PARA / 4 CON / 1 A / 1 SOBRE	SEMILÉXICA	77,93 %	6	1,14
7.4 CON	100 / 25 PARA / 7 DE / 6 EN / 2 ENTRE / 2 CONTRA / 1 HACIA / 1 A / 1 SOBRE	SEMILÉXICA	69,93 %	8	1,45
8 PARA	103 / 25 HASTA / 13 POR / 2 DURANTE / 1 DESDE / 1 DE	LÉXICA	72,02 %	5	1,30
9.1 DE	122 / 20 EN / 1 BAJO / 1 CON / 1 POR	SEMIFUNCIONAL	84,13 %	4	0,78
9.2 A	111 / 11 DE / 10 EN / 7 PARA / 3 HASTA / 2	FUNCIONAL	76,55 %	6	1,28

	ANTE / 1 TRAS				
9.3 DE	122 / 11 PARA / 9 HASTA / 1 POR	FUNCIONAL	84,13 %	3	0,88
10.1 DE	67 / 39 SOBRE / 28 EN / 7 PARA / 3 TRAS / 1 CON	SEMILÉXICA	45,45 %	5	1,62
10.2 POR	84 / 50 EN / 17 PARA / 2 ANTE / 1 DESDE / 1 A	FUNCIONAL	57,93 %	5	1,54
10.3 POR	123 / 16 A / 3 PARA / 2 DURANTE / 1 HASTA	LÉXICA	86,01 %	4	0,78

Tabla 23. Parte 4. Resultados de las respuestas identificadas por número, respuestas y porcentaje de acuerdo

Por lo que respecta a la representación de las respuestas hemos elaborado cuatro gráficos 24, 25, 26 y 27, que visualizan los datos de manera transversal, siguiendo las clases preposicionales: la funcional, la semifuncional-semiléxica y la léxica. Como se aprecia el número de respuestas por clase y por test no es simétrico. Sin embargo, esta variación no afectará, desde el punto de vista cuantitativo ni cualitativo, el análisis de los datos.

En el primero de ellos, el gráfico 24, se indica el porcentaje de acuerdo de las ocurrencias de las preposiciones funcionales, 20 en total, distribuidas en 3, 3, 8 y 6, respectivamente. El resultado se obtiene de sumar los porcentajes y realizar una media aritmética de los porcentajes parciales.

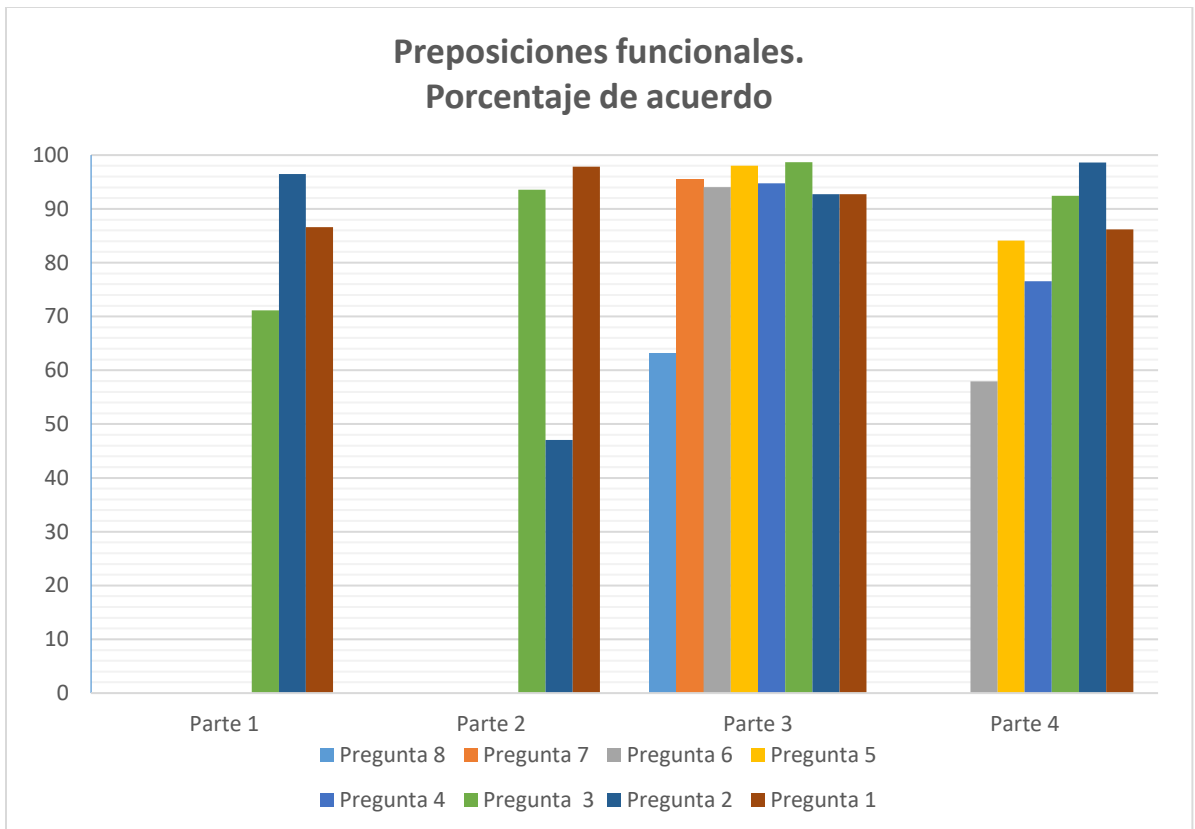


Gráfico 24. Porcentaje de acuerdo en las preposiciones funcionales

En el gráfico 25 se representan los porcentajes para las respuestas de preposiciones consideradas semifuncionales. Aquí la distribución de respuestas es más variable todavía y encontramos 8, 4, 3 y 2 posibilidades, con un total de 17 casos. En esta ocasión es la parte 1 la que concentra más opciones de esta clase. Ya ha empezado el descenso de los aciertos, como vamos a comentar.

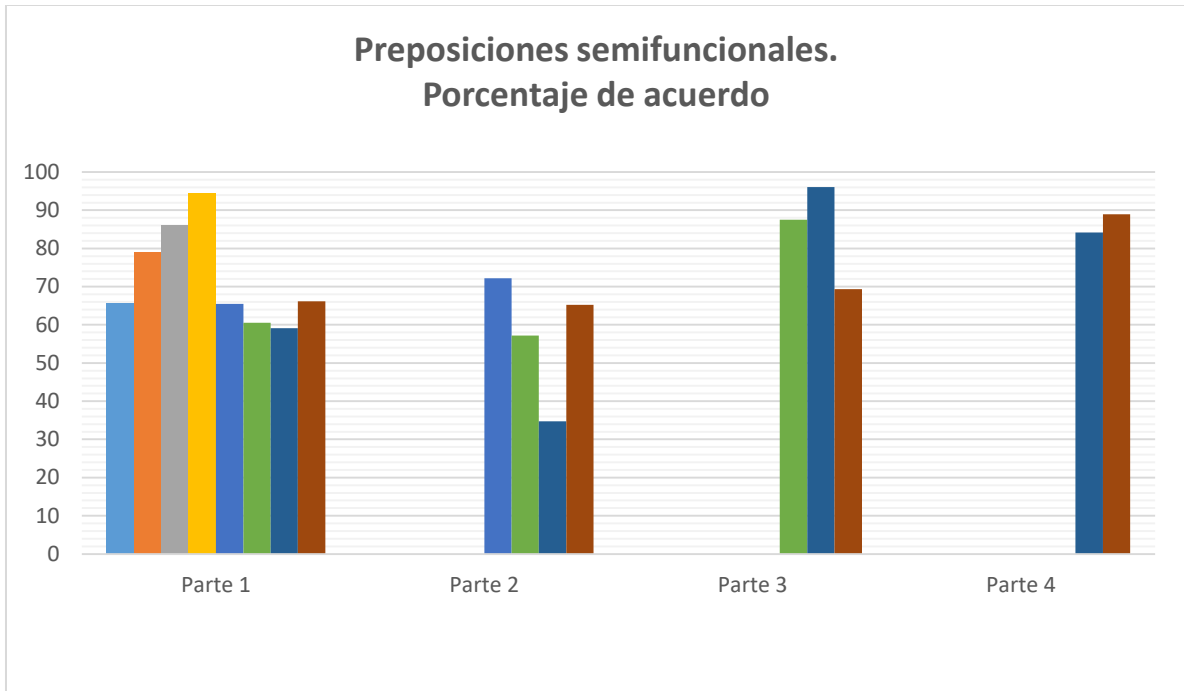


Gráfico 25. Porcentaje de acuerdo en las preposiciones semifuncionales

En tercer lugar presentamos el gráfico 26, el de las preposiciones semiléxicas con dos partes del test, la 2 y la 4 cargadas de esta opción, 7 preguntas en cada uno, frente a los otros dos, solas con tres. El descenso en el acuerdo de las respuestas se asemeja, cuantitativamente, a las anteriores.

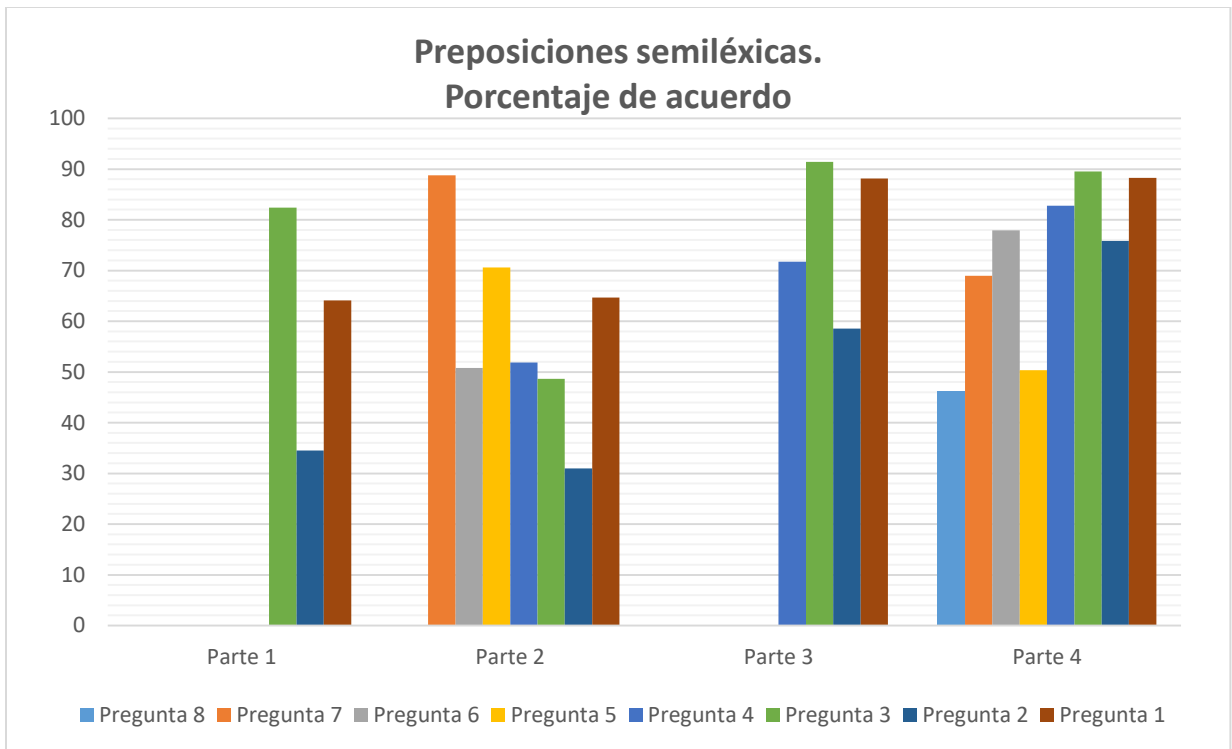


Gráfico 26. Porcentaje de adecuación y acierto en las preposiciones semiléxicas

Y por último, en el gráfico 27 encontramos los datos de las preposiciones léxicas, 31 ocurrencias distribuidas en 8, 4, 10 y 9 por parte. Se observa que la disminución en el acuerdo de las respuestas es muy notable por los motivos que tratamos inmediatamente.

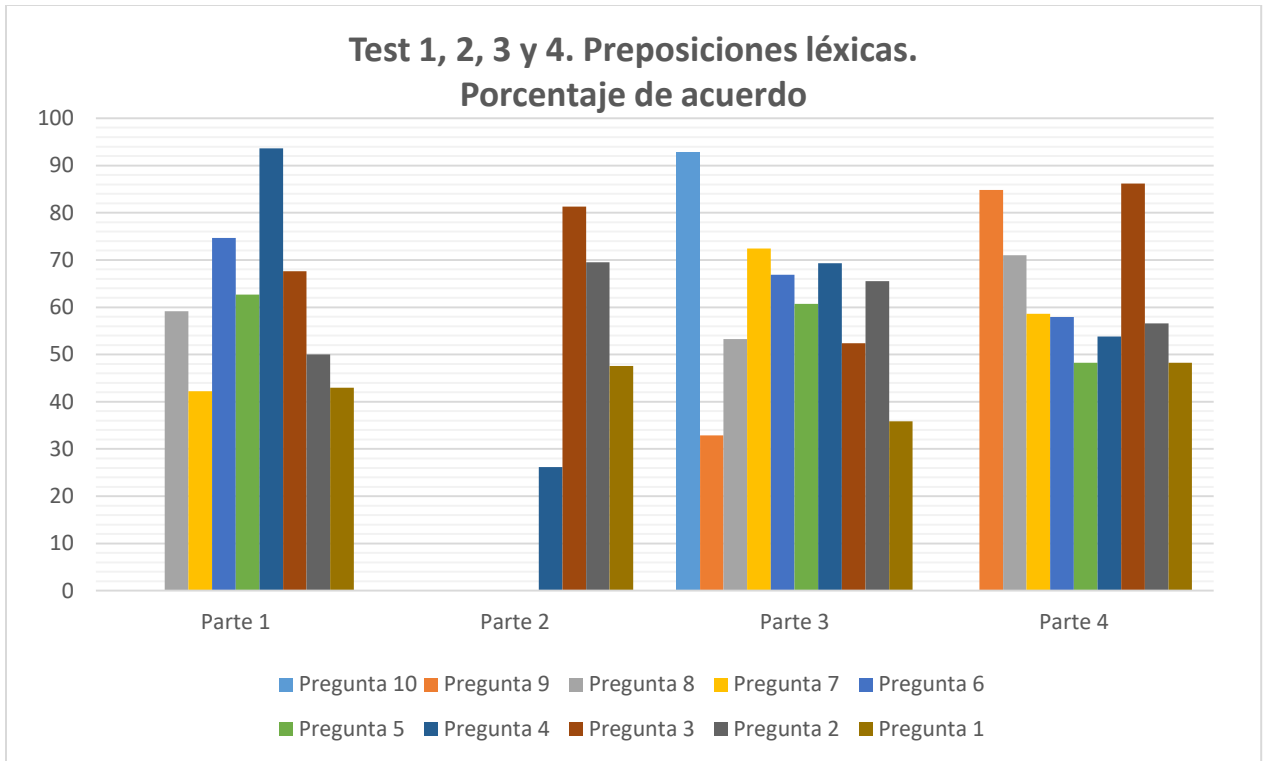




Gráfico 27. Porcentaje de acuerdo en las preposiciones léxicas

3.4.5 Análisis de los resultados: acuerdo, dispersión y entropía

Tras la presentación de los datos, las respuestas y el detalle porcentual de la totalidad de las preguntas procedemos a verificar la fortaleza empírica de la hipótesis gradual para el significado preposicional. En este experimento la correlación se establece de manera doble, como queda representado en el cuadro siguiente.

Correlaciones observadas en el análisis de los resultados del test	
Grado porcentual de acuerdo entre hablantes y modelo 	Tipos clases preposicionales (según su contribución semántica)
Dispersión de las respuestas (medida en términos absolutos y con entropía) 	

Cuadro 31. Correlaciones observadas entre los resultados del test y la hipótesis

Tal y como sucedía con el valor de la entropía en el segundo y el tercer experimento, nos va a permitir medir cuánto desorden existe en la dispersión de las respuestas y, cómo vamos a indicar, ese desconcierto guarda relación con la cantidad de significado de la pieza preposicional.

Si empezamos por el análisis del acuerdo entre los hablantes y el modelo podemos observar el gráfico 28 donde se comparan las cuatro partes del test y los resultados de las cuatro clases. Como decimos las funcionales obtienen las cifras más altas de acuerdo. A su vez comprobamos que, dependiendo de la parte, las distancias varían y, en la parte 4, las semifuncionales superan a las funcionales; o en la parte 2 semiléxicas y léxicas crecen en lugar de disminuir. Estas diferencias, sin embargo, no afectan a la validez global de los resultados porque lo que cuentan son las cifras totales que permiten reforzar nuestra hipótesis. En consecuencia, no tiene sentido buscar información valiosa, demostrativa o interpretable en los análisis individuales de las cuatro partes por la propia naturaleza unitaria de su diseño y confección.

Además, las distribuciones de preposiciones, ocurrencias y clases distintas en cuatro partes buscaban, precisamente, corregir las posibles desviaciones. En los comentarios por clases preposicionales nos detendremos en el examen de algunas circunstancias gramaticales que justifiquen algunas variaciones puntuales por razones de

carácter léxico o sintáctico, compatibles con el razonamiento semántico que sostenemos en nuestra investigación.

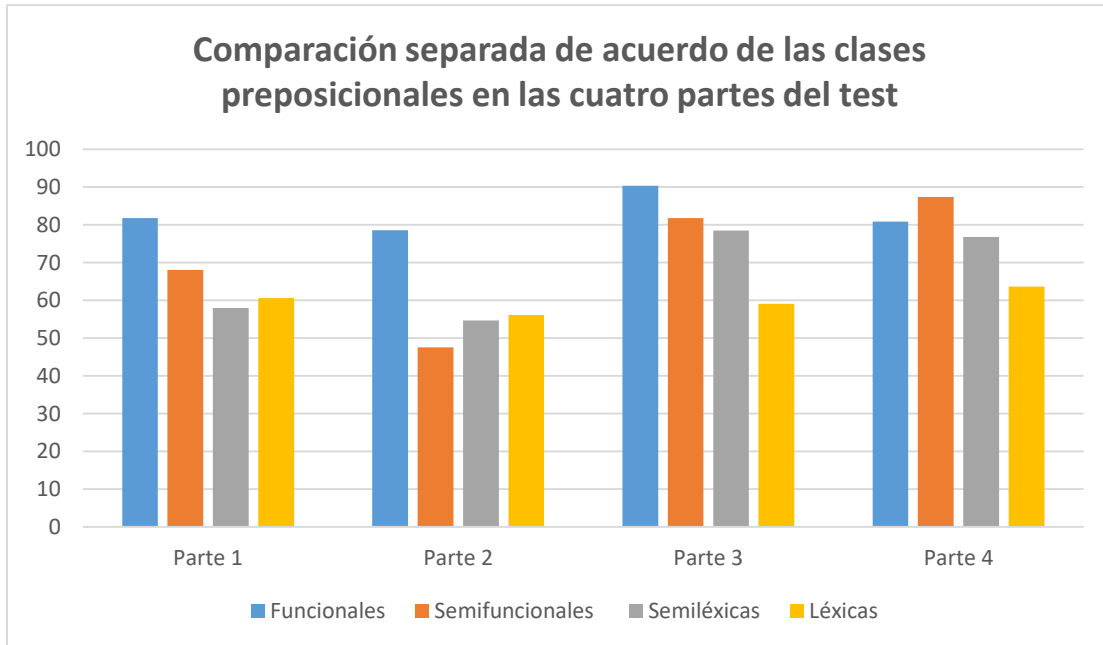


Gráfico 28. Comparación entre las cuatro clases preposicionales de las cuatro partes

Una vez tratados de manera unitaria –sumados los resultados parciales- en el gráfico 29 se aprecia, con claridad, la hipótesis gradualista del significado preposicional.

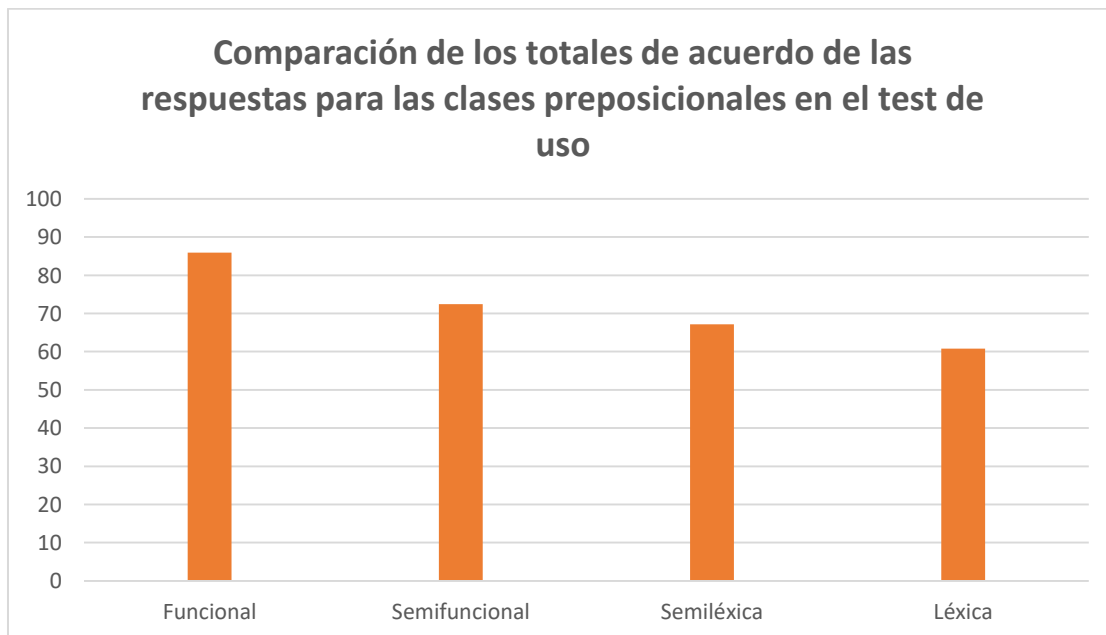


Gráfico 29. Comparación del grado de acuerdo de las clases preposicionales

La diferencia porcentual entre las cuatro series se señala en la serie de (117) a (119).

(117) De F (85,90 %) a SF (72,41 %): 13,49 puntos

(118) De SF (72,41 %) a SL (67,13 %) 5,28 puntos

(119) De SL (67,13 %) a L (60,80 %) 6,33 puntos

El descenso más acentuado entre funcionales y semifuncionales (más del doble respecto de la siguiente variación) se explicaría, como sucedió en el experimento tercero sobre la entropía, por la clara ausencia de significado de las preposiciones *a*, *de* y *por*. La facilidad para su uso reafirma el carácter casi nulo o transparente de su contribución al contexto, mientras que en las siguientes ya se crean expectativas semánticas en sus enunciados aumentando, en consecuencia, la incertidumbre del hablante en el momento de elegir una opción.

La segunda variable que hemos empleado para la predicción de la gradualidad es el nivel de dispersión en la elección de las respuestas, es decir, cuántas opciones de más han propuesto los hablantes.

En las dos últimas columnas de las tablas 20, 21, 22 y 23 se recogieron, primero una cifra que indica el número de otras preposiciones que se han propuesto como respuestas; y después el valor de la entropía de las cantidades de ocurrencias de cada una de esas preposiciones. Seguidamente resumimos en el gráfico 30 el promedio de esa variable en las cuatro partes para cada una de las clases.

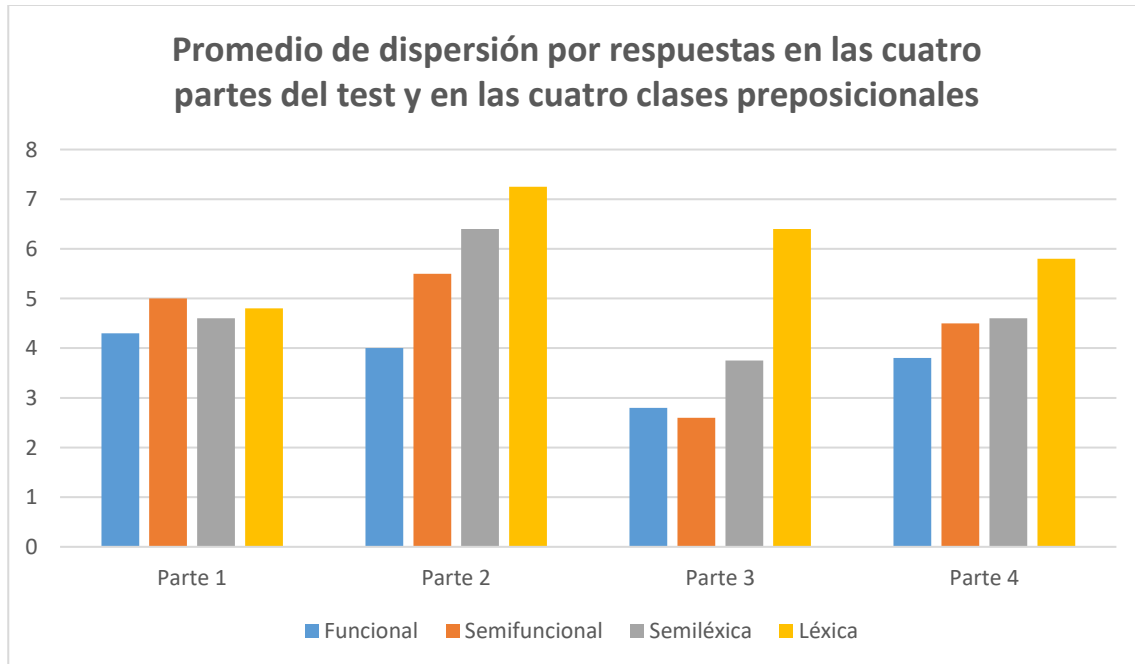


Gráfico 30. Comparación de la dispersión de las clases preposicionales

La dispersión indica el número de preposiciones distintas en el test que sugieren los participantes. En las tablas anteriores hemos aportado la medición de la dispersión de dos maneras: con el término absoluto (2, 3, 4, 5...) y con la entropía, según la fórmula que venimos empleando en nuestra investigación. En el gráfico 31 se comprueba que las preposiciones funcionales aceptan menos opciones, frente a las clases intermedias, semifuncionales y semiléxicas, que crecen. Las léxicas, son, en correlación inversa con el acuerdo en el uso, las que más se dispersan.

A diferencia de lo que sucede con el uso, la dispersión más baja, o concentración en pocas preposiciones, se da en las preposiciones funcionales. Son las que aportan menos carga semántica y, en consecuencia, a los hablantes sugieren menos dudas.

El total de la dispersión de cada clase se reúne en el gráfico 31. En él se aprecia con mayor claridad la evolución inversa entre acuerdo o certidumbre en el uso y concentración, aunque cómo vamos a comentar en el análisis por clases, en las clases intermedias (semifuncionales y semiléxicas) la diferencia de dispersión resulta poco significativa.

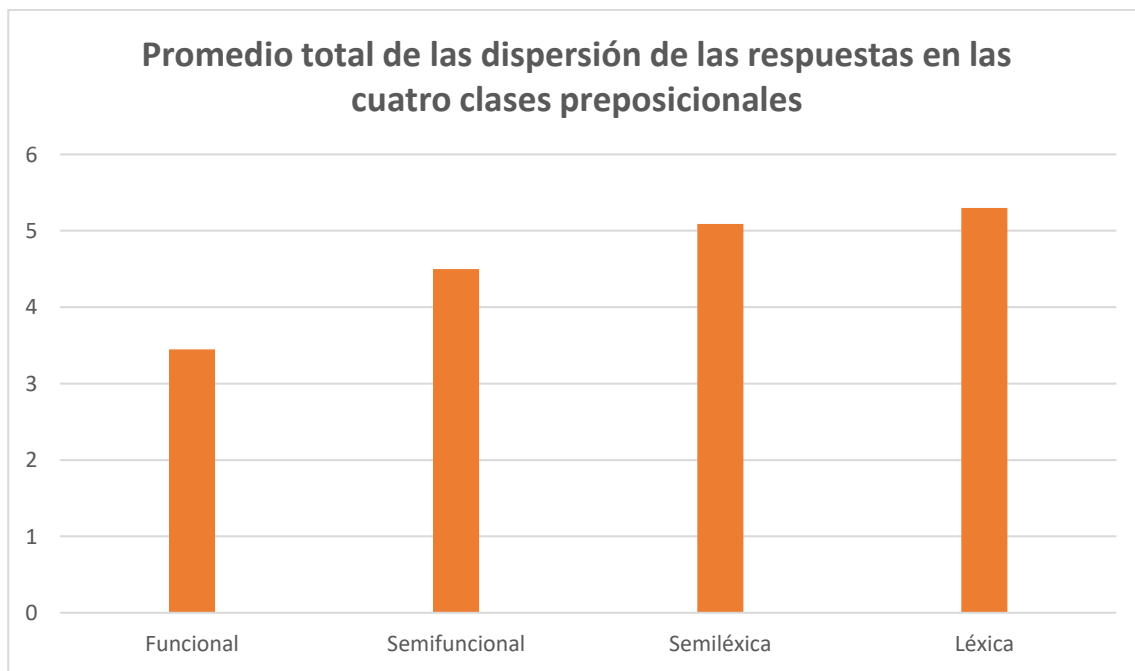


Gráfico 31. Comparación de la dispersión global por clases preposicionales

Otro aspecto fundamental en la medición de la dispersión es cómo de entrópica es la distribución de sus respuestas. Los resultados sobre la entropía de la dispersión son reveladores del fenómeno de la gradualidad porque expresan el aumento de la incertidumbre en relación con el significado. Las preposiciones funcionales tienen un valor más bajo de entropía, es decir, se eligen menos alternativas (medida de dispersión) y en menor cantidad (medida de entropía).

Recordemos como se han obtenido las cifras sobre la entropía, por ejemplo, en las preposiciones funcionales de la primera parte del test. En la tabla 24 se ejemplifica el estilo de tabla que usa, idéntica a las empleadas en el cálculo de los experimentos dos y tres.

Ejemplo de tabla de cálculo de la entropía (preposiciones funcionales de la primera parte del test)			
Número de ocurrencias por preposición	Espacio en blanco 1 del test	Espacio en blanco 2 del test	Espacio en blanco 3 del test
.	122	139	88
.	18	2	21
.	1	1	16
.	1	-	9
.	-	-	4
.	-	-	1
.	-	-	1
TOTAL	0,74	0,34	1,62

Tabla 24. Ejemplo de tabla de cálculo de la entropía (preposiciones funcionales de la primera parte del test)

Así, en el apéndice se encontrará la tabla A, página 315, con los resultados completos para todas las partes y todas las clases de preposiciones. Se introducen las cifras de ocurrencias de cada una de las preposiciones elegidas por los hablantes para cada espacio en blanco del test en la hoja Excel y se calcula el índice. Esa tabla también incluye la cifra de dispersión y el porcentaje de acuerdo.

Una vez introducidos y realizados todos los cálculos se busca la media aritmética y se consiguen los resultados globales de la tabla 25 que hemos detallado también por partes.

Tabla de la medida de la entropía de la dispersión para las clases preposicionales en el test				
Clases de preposición	Funcionales	Semifuncionales	Semiléxicas	Léxica
Parte 1	0,84	1,21	1,61	1,5
Parte 2	0,78	1,56	1,59	1,72
Parte 3	0,54	0,79	0,97	1,58
Parte 4	0,86	0,71	1,14	1,39
Entropía (H)	0,72	1,16	1,32	1,52

Tabla 25. Entropía de la dispersión en el uso de las clases preposicionales

Una recolección pormenorizada de los datos se refleja en el gráfico 32 donde se aprecia en columnas como se produce el crecimiento entrópico de las clases preposicionales por partes, aunque se aprecian algunas irregularidades en el progreso entrópico.

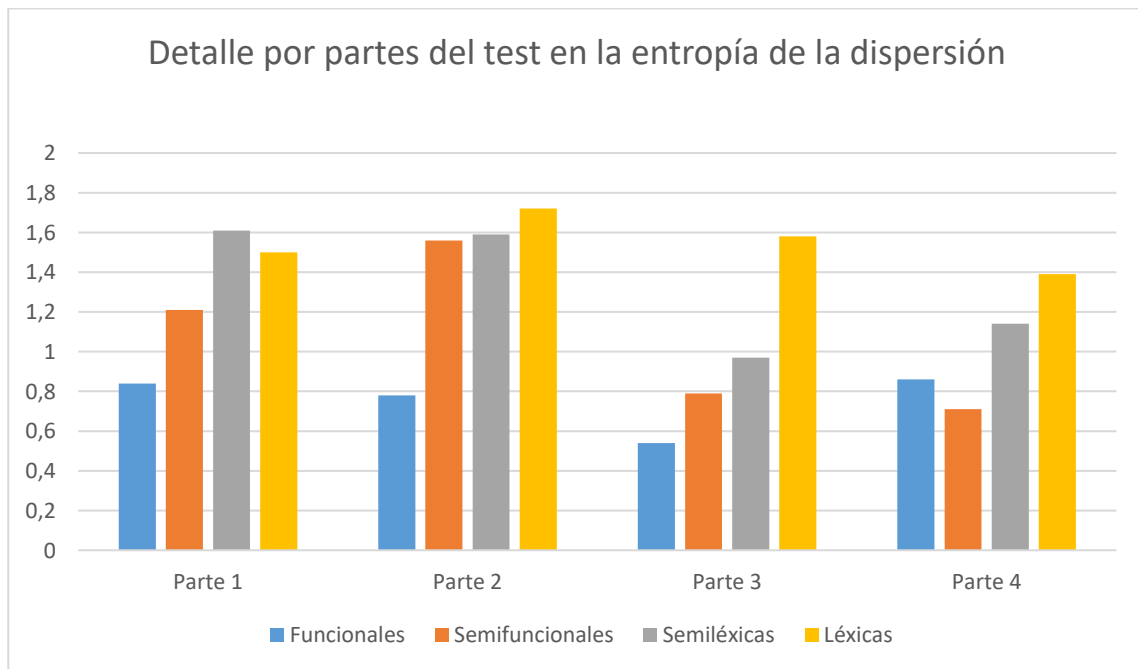


Gráfico 32. Resultados de entropía de la dispersión de las cuatro partes del test

Las fluctuaciones inevitables como resultado de la diferencia de número de ocurrencias en cada clase van a desaparecer una vez que tratamos los resultados globalment. Así, y ee manera unitaria, el gráfico 33 nos muestra, en cambio, el crecimiento de la entropía desde las funcionales hasta las léxicas, vinculado a la hipótesis de toda nuestra experimentación: a menor entropía menor significado preposicional.

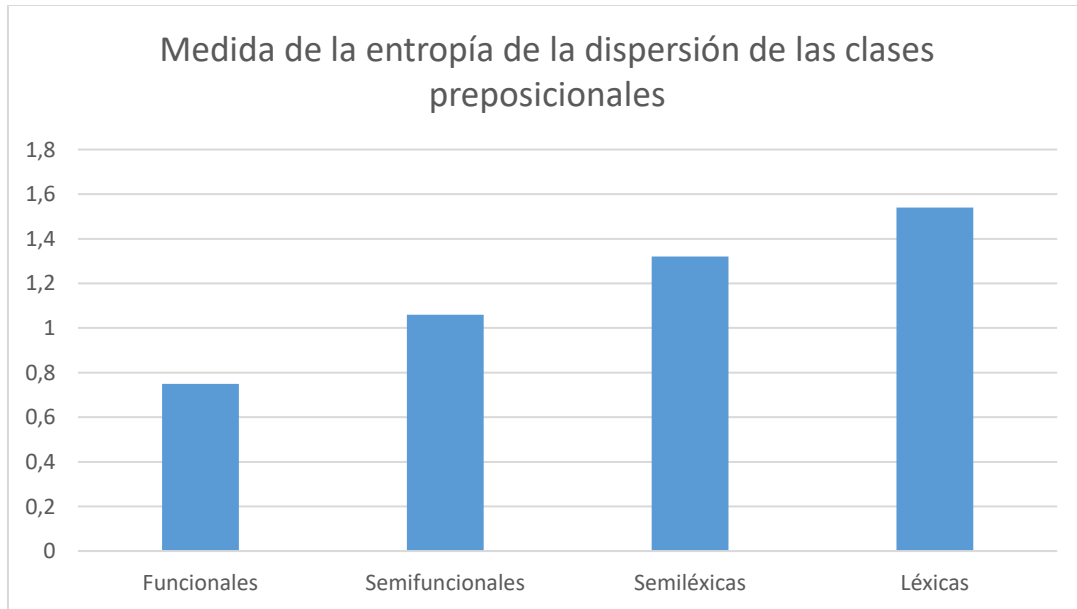


Gráfico 33. Resultados de entropía de la dispersión del test completo

Como se aprecia en el gráfico, la variación entre las cuatro clases es muy parecida. Así tenemos que entre funcionales y semifuncionales se da un 0,31 punto de diferencia; 0,26 entre semifuncionales y semiléxicas, y solo un 0,22 entre semiléxicas y léxicas. Esta progresiva disminución es una constante en las fases de gradualidad de los anteriores experimentos. A medida que crece el significado las diferencias entre clases preposicionales se van atenuando.

La síntesis del análisis global es que el acuerdo entre el modelo y los usuarios aumenta a medida que desaparece el valor semántico de la preposición; frente a la incremento de la dispersión y su entropía con la aparición de significado en la interpretación del enunciado. Hemos comprobado que la entropía nos indica que las preposiciones funcionales son preposiciones poco dispersas y ordenadas, es decir, se da la concentración en pocos casos y con pocas ocurrencias.

Pasemos a examinar algunos aspectos relevantes en cada uno de las clases con el propósito de adentrarnos en algunas causas de los resultados obtenidos.

3.4.5.1 La clase funcional

El resultado global de las preposiciones funcionales es de 85,09 % de acuerdo. La solidez de la cifra permite comprobar que el acuerdo con el modelo de esta clase está directamente relacionada con la inmediatez en la decisión semántica del usuario, es decir, saben qué deben responder porque el significado preposicional no es relevante; es meramente gramatical. No hay debate interpretativo.

Solo en esta clase, y para dos preguntas, 6.1 de la parte 3 y 6.1 de la parte 4, se produce los acuerdos más altos: el 98,66 % y del 98,62 % de todo el test.

Asimismo, de las 20 ocurrencias funcionales, 12 superan el 90 % y 3 el 80 %; y 5 casos se encuentran entre el 57,93 % y el 76,55 % siempre a relativa distancia del 50 %.

Desde el punto de vista descriptivo, la clase de ejemplos y las preposiciones (*a*, *de* y *por*) constituyen las estructuras gramaticales arquetípicas de la funcionalidad, tal y como se recoge en (120 a-d).

(120)

(a) Recogerán *a* los pasajeros

(b) Acostumbra *a* servir un refrigerio

(c) La llegada *del* equipo

(d) Fue ratificado *por* la cámara

Desde el complemento directo personal de (a), hasta la perífrasis de (b), pasando por el complemento nominal en (c) y el complemento agente de la voz pasiva en (d), la funcionalidad está representada extensamente (Funes, 2011b).

A pesar de la fuerza cuantitativa de los resultados, nos encontramos con algunos datos en los que conviene detenerse. Así, en la parte 2, pregunta 3, la preposición *por* solo obtiene un 47,05 % de acuerdo. Recordemos el ejemplo en (119).

(121) Tres cartas fueron enviadas *por* el secretario judicial sin ningún remitente

Las otras opciones que proponen, con su número de ocurrencias, son 31.01 % (58) para la *a*, 12.29 % (23) para la *para*, 8.55% (16) para la *hacia*, y 1.14% (1) para *ante*. En la construcción de un verbo en forma pasiva podría esperarse sistemáticamente la preposición *por*, pero son gramaticales, correctas y adecuadas las opciones con *a*, *para*, *hacia* y *ante*. Todas admiten, en este contexto, el rasgo +Animado e indican un receptor o destinatario. Aquí la idiosincrasia del verbo *enviar* permite este abanico de posibles interpretaciones. Si hubiese sido un verbo más restringido o delimitado como *construir*, la situación, probablemente, hubiese sido distinta. Así, en (122a) no se admitiría tanta dispersión; y (122b) sería agramatical directamente.

(122)

(a) Tres apartamentos fueron construidos por los obreros

(b) * Tres apartamentos fueron construidos a los obreros

Otro hecho destacable aparece en la parte 1, pregunta 5.2, donde nos encontramos con la preposición *por* y un 71,12 (92) % de acuerdo. Recuperamos el ejemplo en (123).

(123) Lo realizaron entre todos los presentes que estuvieron regidos *por* una serie de especialistas en la materia.

Las opciones de la dispersión son 21 *en*, 16 *a*, 9 *con* y 4 *según*. La atribución de un significado con otras preposiciones solo es posible si se entiende el nombre *serie* como el conjunto de capítulos de una emisión audiovisual o un sentido semejante.

Otro aspecto que digno de mención es lo que sucede en la parte 4, en sus preguntas 9.2 y 9.3, donde el posible descenso de aciertos tiene que ver con la conexión o correlación entre los significados de las dos preposiciones. Es el ejemplo de (124).

(124) De manera que el juez ha propuesto un aplazamiento *a* la espera *de* que se solucione la cuestión

La opción vencedora, que lo es con un 76,55 %, ha presentado como segunda opción 11 respuestas con *de*, que solo será gramaticalmente posible si de igual forma se cambia la segunda preposición, y se propone *para*, como se lee en el enunciado (125).

(125) De manera que el juez ha propuesto un aplazamiento *de* la espera *para* que se solucione la cuestión

Estas variaciones, justificadas gramaticalmente por la interacción entre los elementos del enunciado, no pueden considerarse excepciones en nuestra investigación porque no guardan relación con la expresión de la gradualidad semántica, sino con las diferentes interpretaciones o lecturas del enunciado.

Por lo que respecta a la dispersión de la clase, en las 18 ocurrencias funcionales se oscila entre 1 y 6, y alcanza un promedio de 3,3, el más bajo de las cuatro clases.

Se reafirma, por tanto, que es la clase cuantitativamente más sólida en una de las correlaciones que establecemos en la hipótesis: la facilidad de elección o acuerdo se debe a la práctica ausencia de interpretación semántica de la preposición. Ni la posible ambigüedad en la interpretación de la secuencia ha propiciado la distorsión de estos resultados.

La entropía es la más baja (0,75) y demuestra que los usuarios optan por propuestas casi automatizadas en la medida en que su aporte a la comprensión global es residual.

3.4.5.2 Las clases semifuncional y semiléxica

Los datos de las clases semifuncional y semiléxica corroboran el tránsito hacia las variaciones sobre la elección de una preposición u otra. Porcentualmente, la primera clase se encuentra en un 72,41 %, ha disminuido el acuerdo respecto de la funcional; y la segunda baja hasta un 67,13 %. A medida que crece la aportación en el significado de la secuencia el hablante parece que se muestra más inseguro porque debe interpretar un significado que no es extraíble exclusivamente del contexto.

Algunos ejemplos se observan, para las semifuncionales, en (126 a-c).

(126)

(a) Aunque no se complacen *en* atacar a los visitantes, su actitud es claramente impropia de su posición

(b) Se dan cita en Valencia para rendir tributo *a* los instrumentos de cuerda

(c) Contaban *con* su confianza

Son, en su mayoría, preposiciones en construcciones controladas por el verbo, o de régimen verbal. Como hemos indicado en el experimento 2, al encontrarnos con estas estructuras o tipos de complementos, el papel de las preposiciones es más activo que en las funcionales puesto que el significado del verbo necesita de un término (el núcleo del sintagma) y una orientación que corre a cargo de la preposición. Desde el punto de vista composicional la preposición activa su rol semántico en estas situaciones.

La diversidad de respuestas en las semifuncionales, ese casi 20 % de diferencia respecto a la clase anterior, tiene que ver con los matices, las posibilidades o las variaciones que atribuyen los hablantes y esas opciones, como estamos indicando, provienen de la nueva complejidad léxica de la preposición.

Los porcentajes de acierto fluctúan entre un 34,75 % (parte 2, pregunta 4) y un 96,05 % (parte 3, pregunta 7.2). Veamos los ejemplos en (127) y (128), y examinemos las causas de las diferencias porcentuales respecto de la previsión.

(127) Nos inclinamos *hacia* esa nueva opinión

(128) Una serie *de* hombres y mujeres de apariencia burguesa

En (127) se han propuesto 47 ocurrencias para la *por*, es decir, a solo 15 ocurrencias de la considerada más adecuada (*hacia*). En la realización del valor léxico-semántico de *meta* convergen las dos preposiciones: *hacia* y *por*. En el primer caso parece que algunos usuarios han optado por el sentido espacial y de orientación del verbo *inclinarse* y prefieren la preposición *hacia*; frente a los que consideran más decisivo el valor del *meta*, procedente, quizás, del *por* en las construcciones en voz pasiva.

En cambio, en (128) la explicación guarda relación con la compleja y polisémica naturaleza de la preposición *de*. Como vimos en el experimento 3, sección 3.3.5.1.1, el abanico léxico-semántico de esta preposición es extenso, aunque la mayoría de esos valores se concentran en las clases funcional y semifuncional. En el ejemplo (128), que indica “parte de” o “contenido de”, no parece que exista ninguna duda semánticamente viable. De hecho las otras dos preposiciones elegidas han sido *para* y *entre*, que solo serían posible gracias a otra acepción del nombre *serie*.

Inspeccionemos, ahora, la comparación entre la semifuncionalidad y la semilexicalidad. Como indicamos, la diferencia en puntos porcentuales entre ellas es de 4,33 %, la más baja de los tres grados de comparación (F-SF, SF-SL, SL-). La elasticidad semántica de estas dos clases ya ha sido referida en los anteriores experimentos.

Si observamos algún ejemplo de la clase semiléxica, los recogidos entre (129) y (132), comprobamos que los matices interpretativos pertenecen al ámbito de la modalidad, es decir, la especificación, el modo, la causa o la finalidad de las acciones.

(129) Todavía le quedaban fuerzas *para* recabar fondos

(130) Viajaron *con* la intención de solucionar el problema

(131) La habitación *de* abajo

(132) Los soldados fueron reprendidos *por* su conducta grosera y ostensiva

Algunas de las respuestas obtenidas ponen en evidencia como la expresión semántica de esos valores representa una mayor dificultad de elección para el usuario. Por ejemplo, en la respuesta 5.1 de la parte 1, nos encontramos con uno de los resultados más bajo, un 34,50 % para la preposición *entre*. Precisamente, sobre la bivalencia de esta preposición se han presentado algunas evidencias (Cabezas, 2014a y 2014b). La dispersión de las respuestas es la siguiente: 36 *con*, 27 *para*, 16 *ante*, 13 *en*, 1 *por*, 1 *según*, 1 *contra* y 1 *mediante*. El verbo *realizar* admite un conjunto de posibilidades tan genérico que se refleja en la multiplicidad de decisiones que adoptan los hablantes en sus respuestas. A nuestros efectos, lo que resulta decisivo, es que cualquiera de las tres opciones mayoritarias denota un peso semántico atribuible solo a la preposición.

En relación con la variable de la dispersión, la comparación de ambas clases, semifuncionales y semiléxicas, no permite apreciar apenas diferencias en la decisión que toman los hablantes. Tal y como se observa en los promedios de dispersión de las cuatro partes para las dos clases, semifuncional (133) y semiléxica (134), están casi igualados.

(133) P1 (5), P2 (5,5), P3 (2,6), P4 (4,5) = 4,4

(134) P1 (4,6), P2 (6,4), P3 (3,75), P4 (4,6) = 4,8

Nos hallamos, como en los anteriores experimentos, en una zona de vacilación, oscilación e incertidumbre dentro de la gradualidad semántica de la categoría. La distinción semifuncional y semiléxico se sostiene, a partir de los resultados examinados, desde el punto de vista de los usos globales aunque no resulta relevante la comparación de la dispersión como marcadora de la ambigüedad o de la polisemia en la interpretación de los enunciados.

Por lo que respecta a la entropía los niveles se aproximan entre sí en estas dos clases, 1,06 y 1,32 respectivamente. A medida que aumenta el valor semántico se amplifica la dificultad en la predicción (0,31 entre F y SF) y cuando se alcanza la clase léxica esa distancia disminuye (0,22 entre SL y L).

3.4.5.3 La clase léxica

Los resultados globales de acuerdo de la clase léxica son de 58,99 % en nuestro modelo (16 respuestas de 31 están por debajo del 60 %). Es una cifra claramente baja teniendo en cuenta que, como vamos a ver, los ejemplos no contienen, según creemos, situaciones demasiado conflictivas, ni desde la perspectiva de la doble interpretación ni sobre cuestiones gramaticales o sintácticas conexas.

Sin embargo, nos encontramos con algunas excepciones como la pregunta 6.1 (93,66 de acuerdo %), de la parte 1, la 10.1 (81,28 % de acuerdo) de la parte 2, y la 3.1 (86,20 % de acuerdo) de la parte 4, que reproducimos en ese orden entre (135) y (137).

(135) Los turistas han dormido *en* el hotel que les recomendaron a la salida del aeropuerto

(136) Viajaron *a* Melilla con la intención de solucionar el problema de la herencia materna. Al comprobar la actitud de sus primos supieron que sería muy fácil

(137) Dirígete *hacia* el norte un kilómetro y luego vuelves a preguntar en la rotonda

En nuestra predicción, el aumento de significado favorece la indecisión o la incertidumbre salvo cuando las condiciones contextuales o distributivas de interpretación son incontestablemente claras. De algún modo, diríamos que en ambos extremos de la gradación semántica del significado preposicional la variación de los hablantes disminuye a partir de razones empíricas; cuanto más informativo es el contexto más fácil es el acuerdo entre el modelo del test y las respuestas.

De este modo, en (135), la selección semántica del verbo *dormir* exige un lugar o artefacto. Esa información es ajena a la interpretación que pueda añadir la preposición, de ahí el elevado acuerdo.

El razonamiento es parecido para (136) y (137). Los verbos *viajar* y *dirigir*, ambos de movimiento, presuponen en el conocimiento semántico de los hablantes, lugares, direcciones o destinos, tal y como indican el nombre de la ciudad de *Melilla* o la orientación espacial *norte*.

Esta fuerza del peso semántico como apoyo en la decisión se relaciona con investigaciones que trataban sobre la gradación desde las preposiciones más léxicas a categorías, también cerradas, como adverbios y conjunciones (Déchaine & Tremblay, 1996; y Cinque, 1999). Esa conexión ya fue descrita y explicada a fondo para el español (Horno, 2002; y Pavón, 2003). De hecho, en respuestas recibidas, pregunta 3 y 4 de la parte 3, aparecen el adverbio *debajo* y la conjunción *mientras*.

La mayoría de ejemplos proceden de los universos expresivos del lugar y del tiempo, tal y como vemos entre (138) y (141).

(138) Las carcajadas se oyen *en* la habitación de abajo

(139) Al comprobar la actitud de sus primos supieron que sería muy fácil

(140) Hemos ido *de* Lugo *a* Santiago *en* tres horas

(141) *Para* mañana es imposible que tengamos los ejercicios

El porcentaje de acuerdo cae, precisamente, porque los valores léxico-semánticos y las piezas preposicionales que las realizan son muy diversos y a menudo complementarios y compatibles. La incertidumbre en el uso se ensancha porque, a menudo, las posibilidades no están desambiguadas en el contexto (a diferencia de lo que sucedía en los ejemplos (135), (136) y (137)).

Ese aumento de la variación hace bajar el porcentaje de los resultados del acuerdo entre el modelo y las propuestas de los hablantes. En cambio, la dispersión en las opciones

ha crecido respecto de la clase semiléxica hasta 0,60 puntos. Esta variable, en algunos ejemplos, alcanza niveles muy altos como en (142), (143) y (144) con 9, 9 y 10 puntos de dispersión respectivamente.

(142) El precio se ha elevado hasta cotas inimaginables *en* el mercado energético

(143) *Hasta* que no se invierta esa tendencia monetaria no habrá nada que hacer

(144) Cree en su inocencia y está dispuesto a todo *en* el procedimiento judicial

Una explicación de las variaciones propuestas procede de la clase de significado léxico. De origen locativo y temporal, sin embargo, presentan un sentido metafórico o general que sería la causa de esa variación. En el último ejemplo (144) no encontramos 32 *durante*, 25 *ante*, 17 *por*, 6 *mediante*, 5 *para*, 4 *tras*, 2 *hasta*, 1 *con*, 1 *sobre* y 1 *según*.

Conviene, a su vez, indicar que esa dispersión se produce con unas cifras de ocurrencias muy altas en cualquiera de las partes del test. No es solo que elijan muchas otras posibilidades como hemos visto en los ejemplos anteriores sino que lo hacen en abundancia. A modo de ejemplo, tenemos algunas series de ocurrencias recuperadas de las tablas, y que reproducimos en la tabla 24, que reflejan la consistencia de los valores semánticos para los hablantes.

Parte del test	Pregunta	Ocurrencias y número de preposiciones
Parte 1	Pregunta 1.2	61 de, 38 para, 22 en y 20 a.
Parte 1	Pregunta 3.2	71 en, 22 de, 15 para y 13 sobre
Parte 2	Pregunta 5.1	89 en, 32 a, 21 de, 18 por y 13 durante
Parte 3	Pregunta 5.1	88 de, 43 a, 16 desde y 3 por
Parte 4	Pregunta 2.1	70 con, 52 en, 17 tras y 2 mediante
Parte 4	Pregunta 6.2	84 a, 33 para, 18 en y 3 de

Tabla 26. Algunos ejemplos representativos de dispersión de la clase léxica en el test

Esta diversidad de preposiciones responde a enunciados gramaticales que son interpretados gracias al peso y la aportación semántica de esta clase de preposiciones. De este modo, la entropía ha crecido casi un 100 % desde las funcionales. Ahora nos encontramos en el 1.54. Si bien es cierto que ese la progresión de ese crecimiento va atenuándose a medida que las preposiciones se cargan de valor semántico. Así, la diferencia respecto de las semiléxicas es solo de 0,22.

3.4.6 Consideraciones estadísticas sobre los resultados del test

Algunas pruebas son necesarias para garantizar la significación estadística del test, es decir, la verificación matemática de una muestra que descarte el azar y pueda establecer una relación causal entre nuestra hipótesis y los resultados utilizados para su confirmación. Los resultados, o en su caso los efectos, serán estadísticamente significativos cuando sea improbable que se obtengan por azar. Es una manera de identificar las razones que justifican el experimento (Freedman et alia, 1978; Womacott, 1997; Álvarez, 2000; Montgomery, 2002; Johnson & Kuby, 2004; y Ross, 2007).

En primer lugar vamos a garantizar que la medida de la muestra elegida es representativa (representatividad cuantitativa de las submuestras). Tenemos un total de 13.915 respuestas (ocurrencias), repartidas en cuatro clases de preposiciones –según nuestra hipótesis- de 3.073 (funcionales), 2.630 (semifuncionales), 3.503 (semiléxicas) y 4.709 (léxicas).

Para calcular la muestra necesaria para que sea representativa a partir de una población conocida (muestra finita), la fórmula estándar que se utiliza es la de (145)

$$(145) \quad n = \frac{N \cdot Z^2 \cdot pq}{E^2(N-1) + Z^2 \cdot pq}$$

En este caso, **n** es la muestra buscada, **N** es el universo o población, **Z** el parámetro estadístico (1,96, el elegido para un 95 % de significancia), **p** y **q** representan la varianza de la proporción (las probabilidades de que sí suceda y no suceda respectivamente) y **E** es el error máximo permitido que estimamos en el 3 % (aunque también lo calcularemos con el 2 %).

Las muestras representativas serían las de la tabla 27 para tres porcentajes de error y con un 95 % de confianza.

Clase preposicional	Tamaño de las submuestras del test	2 % de error (95 % de nivel de confianza)	3 % de error (95 % de nivel de confianza)	3 % de error (99 % de nivel de confianza)
Funcionales	3073 >	2047	991	1632
Semifuncionales	2630 >			
Semiléxicas	3503 >			
Léxicas	4709 >			

Tabla 27. Tamaño de las muestras del test para cada clase preposicional en relación a tres estándares de error y confianza estadística

Como se observa, en cualesquiera de los casos, las muestras que hemos analizado son mayores que las recomendadas por el cálculo establecido con la fórmula de (145).

Por lo que respecta a la representatividad cualitativa, y tal y como indicamos en el apartado 3.3.3.2, en la página 207, hemos elegido una tipología variada de hablantes del español, respecto de la edad, la formación y los lugares de residencia (véase el cuadro 30). Dado que el objetivo del test no es medir aspectos de corrección sino de concordancia y acuerdo a partir de nuestra hipótesis no hemos considerado oportuno introducir otras variables como la territorialidad (países donde se habla español distintos a España) u otros niveles de formación (que no hubiesen obtenido el nivel de estudios primarios). Lo mismo sucedería con la cuestión de la edad. Las franjas, inferior y superior, respectivamente, hubiesen provocado distorsiones relacionadas con las fases de adquisición lingüística –

entre 0 y 12 años- y con posibles problemas de degeneración relacionados con el envejecimiento.

De todos modos, la totalidad de nuestros participantes (626) está dentro de los grupos mayoritarios tanto por lo que respecta a la edad como por lo que respecta a la formación. Como se observa en las tablas 28 y 29, con datos procedentes del Instituto Nacional de Estadística, con nuestros perfiles del test cubrimos el 77 % respecto de los promedios de edad, y el 80 % en relación con la formación del conjunto de la población española (son las filas sombreadas). Además, nuestros datos reales disponibles son para el 50 % de participantes están repartidos de un modo bastante homogéneo en ambas variables, la formación y la edad. El resto de datos no está disponible porque no obtuvimos respuestas de esa información por parte de los participantes.

PORCENTAJES DE FORMACIÓN EN ESPAÑA. Fuente: INE https://www.ine.es/dyngs/IOE/es/operacion.htm?id=1259931075839		Cobertura con nuestras muestras (80 %)	Datos reales disponibles del 50 % (313)
Analfabetos	1,5 %	9	0
Primaria incompleta	5,2 %	32	0
Primaria	12,7 %	86	0
ESO (1er. Ciclo)	29 %	183	105
ESO (2do. Ciclo)	21,6 %	140	80
Superior	30 %	186	123

Tabla 28. Porcentajes de formación académica en España y correspondencia respecto del total en el test y con el 50 % de los datos reales de participantes del test

PORCENTAJES DE FRANJAS DE EDAD. Fuente: INE https://www.ine.es/jaxi/Tabla.htm?path=/t20/e245/p08/l0/&file=02002.px		Cobertura con nuestras muestras cifras (77 %)	Datos reales disponibles del 50 % (313)
0 – 14	14 %	86	100
15 – 29	15,19 %	89	85
30 – 44	22,53 %	140	25
45 – 59	22,59 %	141	113
60 – 74	15,44 %	91	10
74 +	9,41 %	89	

Tabla 29. Porcentajes de franjas de edad en España y correspondencia respecto del total en el test y con el 50 % de los datos reales de participantes del test

A continuación presentaremos un estadístico de contraste de nuestra hipótesis, el χ^2 cuadrado. Este estadístico es uno de los más reconocidos en el análisis de variables cualitativas. Mide la discordancia entre una distribución observada y otra teórica, mostrando en qué medida las diferencias existentes entre ambas, caso de haberlas, se deben al azar en el contraste de hipótesis. Asimismo, es empleada para validar la independencia entre dos variables entre sí, a través de los datos que se suelen recoger en una tabla de contingencia. Cuanto mayor sea el valor de χ^2 menos verosímil es que la denominada hipótesis nula (H_0) –en nuestro caso sería asumir que no existen diferencias de resultados en las cuatro clases de preposiciones- sea correcta (Nikulín, 1973; y Corder y Foreman, 2009).

La fórmula habitual para el cálculo del estadístico χ^2 es la consignada en (146).

(146)

$$\chi^2 = \sum \frac{(fo - ft)^2}{ft}$$

En esta fórmula **fo** serían los datos sobre las frecuencias observadas, y **ft** las frecuencias teóricas. En la tabla 30 recogen los datos del test para nuestros resultados.

CONCORDANCIA DE LA FRECUENCIA PARA EL CÁLCULO DE LA χ^2					
	Funcionales	Semifuncional	Semiléxicas	Léxicas	Total
Concordancia (Frecuencia observada)	2623	2028	2347	2825	9823
Discordancia	450	602	1156	1884	4092
Total	3073	2630	3503	4709	13915

Tabla 30. Datos estadísticos para el cálculo del valor χ^2

A partir de la tabla 30 de contingencia en el apéndice, página 325, tabla B, se obtiene la frecuencia teórica necesaria para aplicar la fórmula anterior (147).

Una vez realizados los cálculos obtenemos una χ^2 de 39. Los grados de libertad (gl) para nuestros datos son 3 (resultado de aplicar el siguiente producto: $(r - 1) (k - 1)$, donde r son las filas (2) y k son las columnas (4)).

Para un nivel de significancia estadística α del 0,05 % -porcentaje convencional en esta clase de pruebas- en la tabla de valores críticos del χ^2 asociado a un grado de libertad 3 se supone un valor de 7,81. Dado que el resultado de χ^2 es mayor $39 > 7,81$ se rechaza la H_0 . De manera que nuestra hipótesis alternativa (H_1) tiene un nivel de significancia estadística aceptable, grosso modo, la formulación de la (H_1) sería: *existirían cuatro tipos o grupos de preposiciones según su naturaleza semántica que provocan una serie de variaciones en las respuestas de los usuarios.*

Por último, calcularemos el coeficiente o índice kappa (κ) de Cohen (Cohen, 1960). Se trata de un medidor (κ) que se corresponde con la proporción de concordancias observadas ($P(o)$) sobre el total de observaciones, habiendo excluido las concordancias atribuibles al azar ($P(a)$). El coeficiente kappa (κ) toma valores entre -1 y +1; mientras más cercano a +1, mayor es el grado de concordancia entre observadores. Por el contrario, un valor de $\kappa = 0$ refleja que la concordancia observada es precisamente la que se espera a causa exclusivamente del azar. La interpretación del coeficiente kappa (κ) se realiza correlacionando su valor con una escala cualitativa que incluye distintos niveles de fuerza de concordancia: pobre, leve, aceptable, moderada, considerable y excelente.

El procedimiento habitual de cálculo del coeficiente kappa (κ) es el de (147).

(147)

$$\kappa = \frac{P(o) - P(a)}{1 - P(a)}$$

El uso que hacemos de este coeficiente es doble. En primer lugar medimos una selección de 11 participantes -elegidos como representación de las cuatro franjas de edad y las tres de formación que hemos comentado anteriormente- y los comparamos entre ellos.

Realizamos un total de 21 comparaciones de estos 11 participantes; 8 de ellos se comparan en 4 ocasiones con otros de manera aleatoria, y 3 en 3 ocasiones. La tabla de

contingencia (Everitt, 1992) y los resultados, caso a caso, se encuentran en la tabla D del apéndice, página 321.

En la tabla 31 mostramos los resultados comparativos para las clases preposicionales del total de las 21 comparaciones, una vez implementados los cálculos con Excel, y a partir de la fórmula de (148).

Clase preposicional	Valor del coeficiente kappa de Cohen
Clase funcional	0,64
Clase semifuncional	0,80
Clase semiléxica	0,38
Clase léxica	0,54
PROMEDIO	0,59

Tabla 31. Coeficientes kappa (k) de Cohen para la comparación de resultados entre 11 participantes distintos

El segundo cálculo con el índice kappa (κ) procede de la comparación de esos mismos 11 participantes en respuestas producidas en dos momentos distintos (mayo de 2019 y octubre de 2019), para observar la coherencia de esos hablantes. Los resultados por preguntas se pueden consultar en el cuadro C del apéndice, página 319. En la tabla 32, a su vez, ofrecemos los datos obtenidos por el mismo procedimiento.

Clase preposicional	Valor del coeficiente kappa de Cohen
Clase funcional	0,79
Clase semifuncional	0,95
Clase semiléxica	0,46
Clase léxica	0,42
PROMEDIO	0,65

Tabla 32. Coeficientes kappa (k) de Cohen para la comparación de resultados de los 11 participantes en dos momentos distintos

La kappa más alta es la semifuncional, seguida de la funcional. El acuerdo aumenta en la medida en que disminuye el significado, y viceversa: en las léxicas es más baja. El cálculo nos asegura que estas correlaciones no se producen por azar. Los detalles de los resultados se pueden consultar en la tabla E del apéndice, página 336.

El promedio de ambas comparaciones (tablas 31 y 32), entre participantes y de ellos mismos, es de 0.62. Según algunos autores (Landis y Koch, 1977; Everitt, 1992; y Hernández-Nieto, 2002) este resultado estaría considerado aceptable o considerable.

Como conclusión de estas consideraciones estadísticas se confirma que las tres verificaciones: la representatividad de la muestra, el estadístico de contraste y los coeficientes kappa (κ), avalan la consistencia estadístico-matemática de nuestra hipótesis para el experimento cuarto; tanto en la elección de la población como el papel que el azar ha podido suponer en la obtención de los resultados.

3.4.7 Conclusiones parciales

La hipótesis gradual, asociada a los resultados sobre el uso preposicional, ha puesto de manifiesto que apenas existe variación en la elección de preposiciones cuando éstas apenas aportan significado. La inestabilidad en el uso aumenta de manera simultánea a la contribución semántica de las piezas.

En este experimento se presenta el contexto de la preposición y son 366 hablantes distintos a los dos anotadores que habían intervenido en los experimentos anteriores, los que, con sus respuestas avalan la predicción sobre las clases funcional, semifuncional, semiléxica y léxica en la preposición española. Esta confirmación se debe a que a medida que crece la incertidumbre, aparecen fases distintas y análogas a las que se habían establecido gracias al aprendizaje automático (experimento 1) y la medición de la entropía de las categorías semánticas de los nombres en el contexto preposicional (experimentos 2 y 3).

Desde el punto de vista metodológico, el hecho de que los datos empíricos analizados en el test contengan esos contextos semánticos contribuye a mejorar la calidad de nuestra hipótesis porque esas decisiones las toman otros usuarios distintos a los anotadores y se alcanzan conclusiones muy semejantes.

Incluso en partes muy diferentes del test, con variables de frecuencia en las preposiciones, de los valores léxico-semánticos, de las clases preposicionales y de los contextos verbales y nominales, los resultados siguen estableciendo correlaciones convincentes.

Propongamos, en el corolario (148), y como hemos hecho en los anteriores experimentos, un enunciado que completa la descripción de nuestra hipótesis a partir del análisis de los datos obtenidos.

(148) *La gradualidad semántica se evidencia en el uso de las preposiciones a partir de la correlación entre el acuerdo sobre el uso y la cantidad de significado. A mayor concordancia menor significado preposicional, y viceversa; a mayor discrepancia entre el modelo y los participantes, mayor significado preposicional.*

Otro hecho que ha quedado probado es que las clases preposicionales intermedias, la semifuncional y la semiléxica, fluctúan menos entre ellas. Esa distinción, que hemos visto variar en los tres experimentos anteriores, consigue su validación doblemente.

De un lado, se han aportado ejemplos más cercanos a la funcionalidad (semifuncionales) y otros a la lexicidad (semiléxicos); y del otro, las relaciones entre ellas son de contacto, de cercanía y de similitud.

El uso de la medida de la entropía al tratar sobre la dispersión de las diversas respuestas ha enriquecido la visión que ofrecía el experimento sobre los contextos nominales de la preposición. En este caso, sin embargo, la asociación se establece entre la cantidad de repeticiones de las ocurrencias y su significado. Podríamos reflejar este hecho en el subcorolario (149).

(149) *El aumento de la entropía de la dispersión en el uso preposicional guarda relación con su aporte o carga semántica.*

Igualmente, al presentar los resultados de la entropía de la dispersión de preposiciones hemos comprobado que la fluctuación de los cambios entre fases o clases se va ralentizando a medida que aparece la interpretación semántica de estas piezas.

Por último, hemos presentado tres condiciones estadísticas para velar por el buen diseño estadístico de la prueba: la representatividad de la muestra, el efecto del azar sobre los resultados y una medida de contraste entre participantes.

CAPÍTULO 4.

CONCLUSIONES Y PERSPECTIVAS FUTURAS

La categoría preposicional ha representado, en los mismos orígenes de la especulación gramatical, un cierto enigma en relación a cómo determinar la naturaleza de su significado, dónde localizarlo, y con qué herramientas –metodológicas y/o teóricas- proceder a su análisis.

Nuestra investigación ha pretendido, desde el reconocimiento de esa dificultad –en forma de controversias, de limitaciones o de restricciones- la incorporación de la noción de gradualidad como concepto explicativo a partir de la aportación de evidencias empíricas. Hemos procurado entender esas clases como adaptativas a un contexto de un modo plástico y fluctuante.

Efectivamente, el objetivo central ha sido contribuir con datos empíricos a la justificación de la gradualidad semántica de la categoría preposicional en español. Hemos convenido que la expresión de esos grados se produce desde una posición o límite funcional, gramatical o vacío –en su aporte de significado- hasta otra posición léxica, autónoma y densa (con fases o clases intermedias) como se representa en la doble línea de (150).

(150)

Funcional ----- Semifuncional ----- Semiléxica ----- Léxica

Menos significado

Más significado

Significado pleno

Para ello, hemos reorganizado los valores léxico-semánticos de las piezas preposicionales -recogidos, presentados y descritos a lo largo de la historia gramatical de la lengua- en función de los grados de expresión semántica de las clases preposicionales.

Hemos asumido, del mismo modo, que el significado preposicional es relacional, contextual y distribuido en el entorno sintagmático de las estructuras donde se insertan. De la mano de las asunciones semánticas del neodistribucionalismo (Boleda y Herbelot, 2017) hemos procedido a la realización de cuatro experimentos que podrían arrojar luz sobre los datos que sustentarían esas clases preposicionales: funcionales, semifuncionales, semiléxicas y léxicas.

La localización de la expresión de esos valores léxico-semánticos se lleva a cabo en los contextos y las distribuciones en las que participan las preposiciones.

De este modo, la formulación de la hipótesis rectora de nuestra investigación se recupera en (151).

(152) Dado un conjunto de valores semánticos descriptivos asociados a las preposiciones y procedentes de la gramática del español, predecimos que la identificación y la distribución de estos valores ocurre gradualmente desde la clase más funcional (con nulo o poco contenido, carga, peso o efecto semántico, y función meramente gramatical) hasta la clase léxica (con un peso semántico relevante e interpretable en el contexto); con la posibilidad de clases intermedias, semifuncionales o semiléxicas, según la estructuras y/o la construcción en la que participen.

Metodológicamente nos hemos situado en la perspectiva de la lingüística cuantitativa, computacional y de corpus, tanto en las herramientas de observación, contrastación, medición y análisis, como en la procedencia de los datos que nos sirven de fuente de nuestras predicciones. Los recursos y las herramientas estadísticas han servido como procedimientos verificadores de los cálculos en relación a la hipótesis.

Recuperemos, de manera sintética, cuáles han sido los resultados y de qué manera han conformado la posible certidumbre de nuestra investigación.

En el primer experimento, dentro del aprendizaje automático no supervisado, hemos seleccionado un universo semántico: los verbos de movimiento en español; y tres preposiciones: *a*, *hacia* y *hasta*, a partir de 71.206 tripletas. Tras su agrupación, basada en

Word Embeddings, hemos obtenido resultados predictivos válidos en más de un 50 % para las tres clases preposicionales supuestas. De manera pormenorizada esa coincidencia entre la hipótesis sobre las clases preposicionales y la agrupación del mecanismo –*Word Embeddings*- para las propiedades semánticas de los contextos se recoge en la tabla 33 para la preposición *a*.

A/DIRECT/4 1731	0	3	1	2
F1	279	73	118	141
SL	100	281	75	52
L1	5	162	236	45
L2	54	18	33	59

Tabla 33. Datos de agrupación por clustering 0, 1, 2 y 3 para la preposición *a*

Los grados de concentración por columna (la pureza) y por fila (pureza inversa) se muestran porcentualmente en los gráficos siguientes. Se indican, en el gráfico 34, los promedios para la preposición *a*, para todas sus agrupaciones.

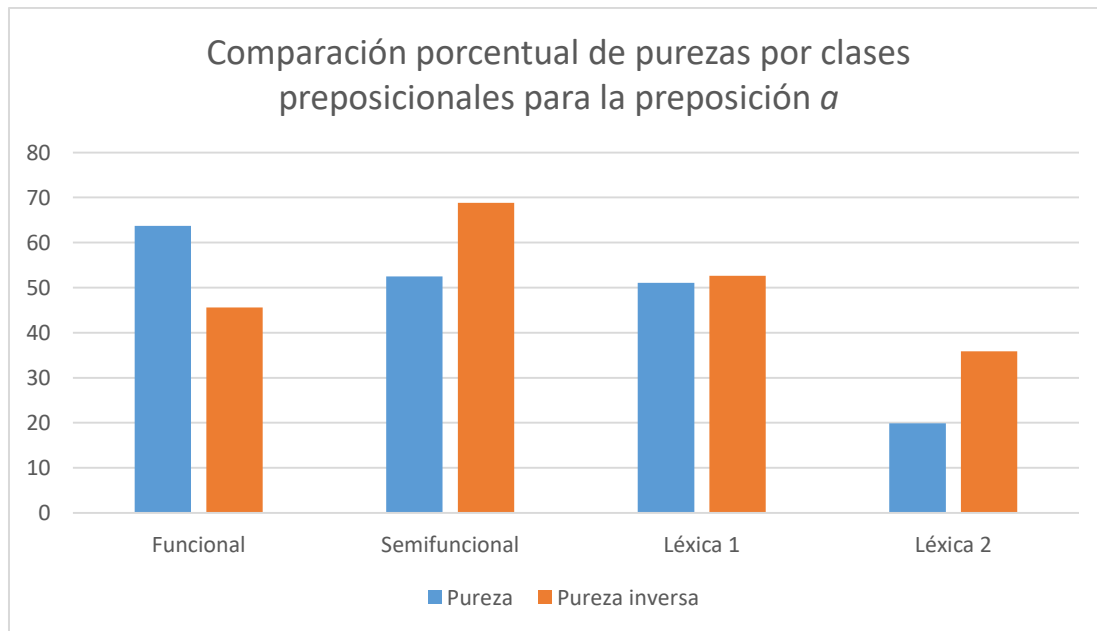


Gráfico 34. Comparación de purzas por clases preposicionales para la *a*

En cinco de los nueve resultados porcentuales se superan el 50 %. La concentración por columnas es superior al 50 % en tres de cuatro casos. Y en una de las clases previstas, la semifuncional, tanto en la pureza como en la pureza inversa se supera el 50 %, llegando en la inversa al 68 %.

Estas cifras, lejos de ser altamente representativas, indican, sin embargo una tendencia y una orientación de uno de los principios de nuestra hipótesis. Se está produciendo una agrupación automática por similitud semántica en el contexto de las preposiciones y esa organización coincide en un grado no despreciable con los resultados establecidos por los anotadores manuales. El algoritmo asocia palabras parecidas en contextos análogos y esa proximidad está en consonancia con la intuición semántica que poseen los seres humanos.

En el caso de las otras dos preposiciones, *hacia* y *hasta*, los resultados se presentan en la tabla 34. Otra vez las diagonales sombreadas en gris son las que informan de los predominios que se producen en las agrupaciones.

HACIA/DIRECT/3 654	2	0	1
SL	117	20	72
L1	169	171	51
L2	33	1	20
HASTA/DIRECT/ 3 778	1	2	0
SL	144	22	29
L1	28	240	41
L2	129	36	109

Tabla 34. Datos de agrupación automática para *hacia* y *hasta*

En las dos diagonales señaladas, salvo en (L2 1) de *hacia*, se concentran las cifras más altas y, además, se produce una cierta concordancia de la previsión de los anotadores manuales con lo que halla la máquina. Así, el análisis interno de los grupos (clusters) ha permitido comprobar que los resultados automáticos son coincidentes en las previsiones de la gramática descriptiva.

Tal y como se muestra en los gráficos 35 y 36, los porcentajes de similitud superan el 50 % para ambas preposiciones (en 5 de 8 casos para ambas preposiciones).

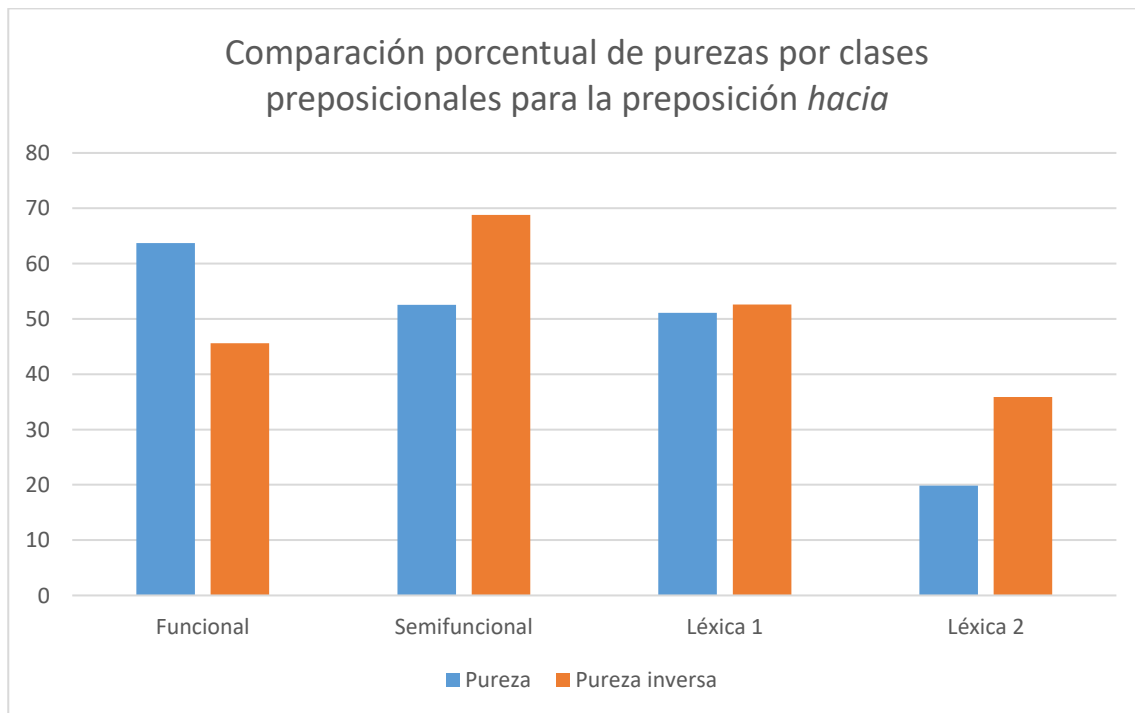


Gráfico 35. Comparación porcentual de purezas por clases para *hacia*

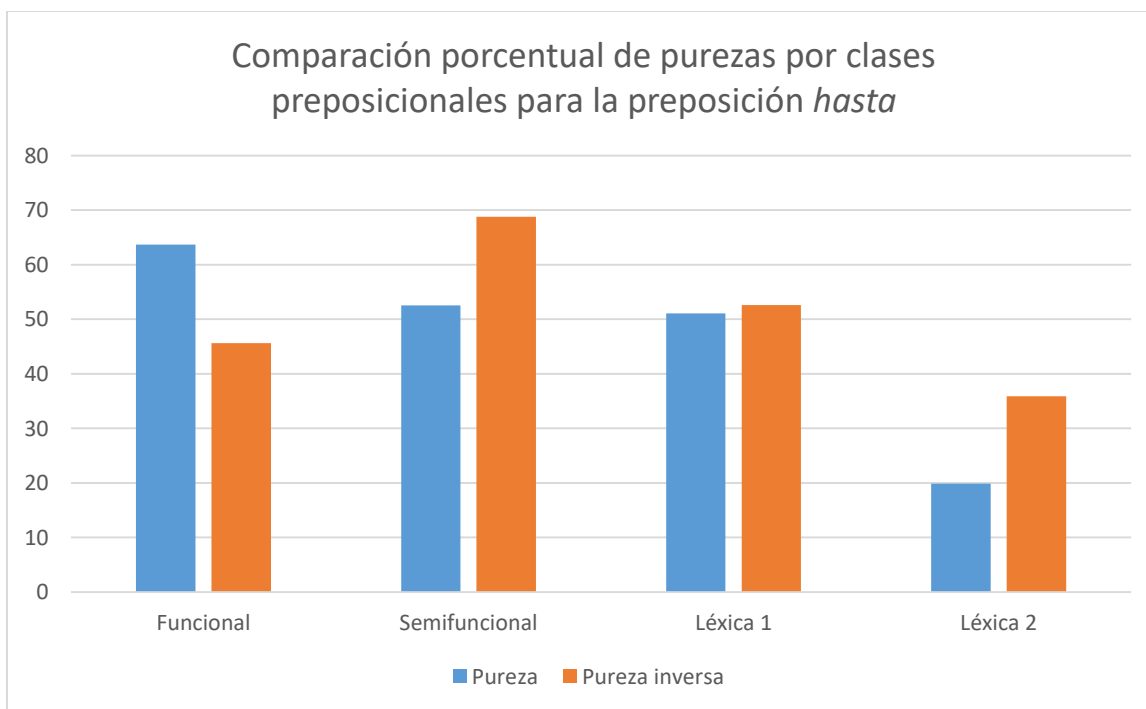


Gráfico 36. Comparación porcentual de purezas por clases para hasta

En el segundo y en el tercer experimento nuestra medida de análisis ha sido la entropía, procedente de la teoría de la información (Shannon, 1948). Primero hemos analizado los nombres de un total de 3898 tripletas preposicionales que dependen de una variedad exhaustiva –sintáctica y semánticamente- de verbos; y después 3903 tripletas que dependen de nombres.

Las seis categorías semánticas utilizadas fueron *Animado*, *Inanimado*, *Evento*, *Entidad abstracta*, *Temporal* o *Locativo*. Una vez clasificados los nombres en posición de complemento preposicional obtuvimos los resultados por clases preposicionales y por categorías semánticas fueron los que recuperamos en la doble tabla 35.

CATEGORÍAS SEMÁNTICAS	CLASES PREPOSICIONALES (Experimento 2. Ámbito verbal)		
	FUNCIONALES	SEMIFUNCIONALES	LÉXICAS
Animado	871	203	120
Inanimado	55	160	152
Entidad abstracta	146	696	374
Evento	73	168	181
Locativo	90	74	376
Temporal	23	26	110
Total de casos (3897)	1258	1327	1313
Entropía total (H) =	1,54	1,99	2,40
Diferencia entre F y SF	0,45		
Diferencia entre SF y L		0,41	
Diferencia entre F y L			0,86
CATEGORÍAS SEMÁNTICAS	CLASES PREPOSICIONALES (Experimento 3. Ámbito nominal)		
	FUNCIONALES	SEMILÉXICAS	LÉXICAS
Animado	741	358	192
Inanimado	214	367	197
Entidad abstracta	191	329	198
Evento	60	99	195
Locativo	121	63	356
Temporal	24	45	153
Total de casos (3903)	1351	1261	1291
Entropía total (H) =	1,90	2,21	2,52
Diferencia entre F - SL	0,3		
Diferencia entre SL - L		0,31	
Diferencia entre F - L			0,61

Tabla 35. Resultados de la entropía de las clases preposicionales (ámbito verbal y nominal)

Los resultados sobre la entropía han confirmado los dos extremos de nuestra hipótesis: que las preposiciones se organizan en clases según su peso, deriva o acción semántica (funcionales, semifuncionales, semiléxicas y léxicas) y que esa organización guarda relación con el aporte de significado de la preposición en el sintagma o tripleta.

Una entropía más baja ha mostrado, en ambos experimentos, que la mayoría de nombres se concentraba en un tipo (*Animado*, fundamentalmente). A medida que crece la entropía se reparten la cantidad de posibilidades semánticas: *Entidad abstracta*, *Animado* e *Inanimado* respectivamente). En el grupo de las preposiciones léxicas la dispersión o el reparto de nombres aumentan casi de manera equitativa, sobre todo en las tripletas del ámbito nominal.

En el gráfico 37 se aprecian los resultados de los dos experimentos. La progresión de la entropía por clases y en sendos ámbitos confirma el pronóstico de nuestra hipótesis.

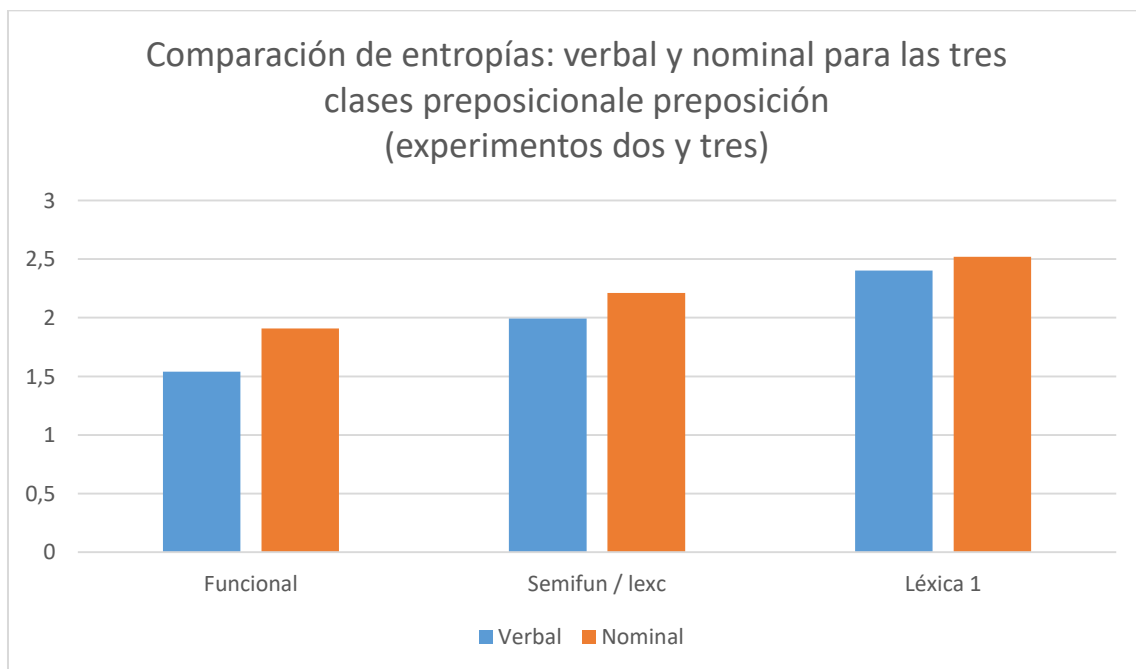


Gráfico 37. Comparación de entropías en los ámbitos nominal y verbal

La relación entre entropía y desorden queda relativamente clara: los resultados más altos tienden a indicar que la preposición es más significativa e impredecible. La progresión es visible en el gráfico, por clases preposicionales y por ámbitos (nominal y verbal).

Por último hemos acudido, en el cuarto experimento, al uso preposicional a partir de un test elaborado con una colección de 90 ejemplos de valores y preposiciones que representan las cuatro clases propuestas.

En esta ocasión las conclusiones avalan también la predicción: el grado de acuerdo en el uso y la entropía de la dispersión en las respuestas (en cantidad y en variedad) confirman que se produce la gradualidad preposicional. En el gráfico 38 se observa como los participantes en el test coinciden más con el modelo propuesto cuanto menos significado tiene la preposición.

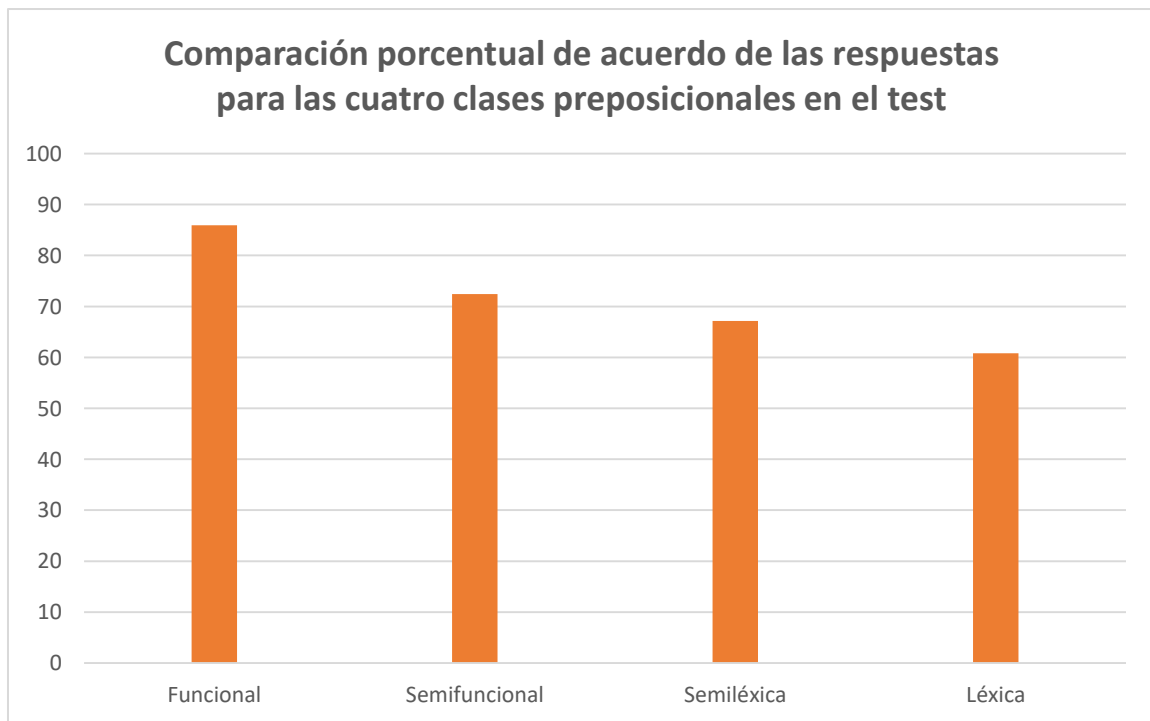


Gráfico 38. Comparación porcentual de acuerdo entre el test y los participantes

Cuanto menos significado tiene la preposición más acuerdo existe entre el modelo del corpus y la elección de los usuarios. El hablante puede deducir del contexto su significado. Intuitivamente, diríamos que la incertidumbre, noción también utilizada en los experimentos con entropía, va contra la determinación del significado.

Por lo que respecta a la entropía de la dispersión los resultados también son aceptablemente positivos. No obstante, la correlación es inversa a lo que sucedía con el acuerdo entre el modelo del test y los participantes. Una mayor dispersión y una mayor entropía tiene que ver con el significado de la preposición, tal y como se recoge en el gráfico 39.

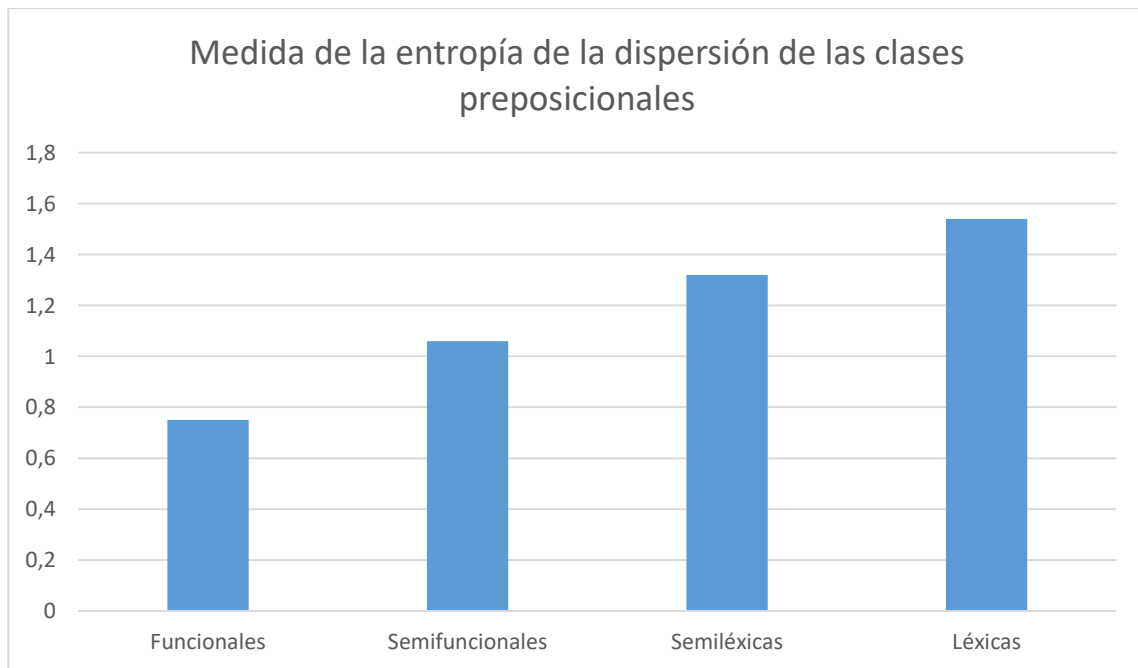


Gráfico 39. Medida de la entropía de la dispersión en las cuatro clases preposicionales

Parecería claro que los hablantes no dudan y aciertan cuando la preposición no ofrece aportación semántica. La entropía es, por tanto, muy baja. Progresivamente, en las preposiciones semifuncionales y semiléxicas los usuarios dudan y proponen otras alternativas que elevan las cifras de entropía.

En síntesis, y como conclusión final se comprueba que los resultados de esa triple perspectiva experimental representan tres formas de aproximarse a nuestra hipótesis: el significado preposicional es gradual, contextual y reconocible desde heurísticas diferentes dentro de una metodología empírica, cuantitativa y computacional.

Por un lado, la agrupación automática nos permite mostrar que es en la similitud contextual que se identifica de manera automática su coincidencia con las predicciones humanas. Las variaciones de grado de las preposiciones empiezan a perfilarse con unos porcentajes de coincidencia que ya son indicativos.

Por otra parte, y cuando ampliamos el foco del espectro de observación y usamos otra variable de medida: la entropía, la confirmación acrecienta su valor en la medida en que la correlación entre entropía y significado se establece de nuevo en el contexto pero con la intervención de criterios humanos en la discriminación.

Emerge una organización entrópica de las categorías semánticas de los nombres vinculada a las clases de preposiciones. De modo que una entropía baja supone que las preposiciones son funcionales y la categoría semántica es la del rasgo *Animado*. A medida que crece el desorden y, por tanto, el aumento de la entropía, surgen las preposiciones semifuncionales y semiléxicas. Y, en último lugar, la entropía de los nombres de las preposiciones léxicas es más alta debido a la relevancia de estas piezas en la composición global del significado.

En el último experimento de la investigación, en el test de uso, hemos comprobado como en una muestra representativa de usuarios del español se produce mayor variación cuanto más significado tienen las preposiciones, como se recoge en (152)

(152) Mayor acuerdo de uso = menor significado preposicional

Menor acuerdo de uso = mayor significado preposicional

También hemos comprobado que esa incertidumbre interpretativa –relacionada con el acuerdo- produce un efecto de dispersión en las ocurrencias que se usan para cada caso

que, una vez medida con la variable entrópica, produce una segunda correlación. La entropía de la dispersión en el uso y el significado de las preposiciones crecen en paralelo. De manera que podríamos añadir a (152) un tercer enunciado a la igualdad, la de (153)

(153) Mayor acuerdo de uso = menor significado preposicional = menor entropía de dispersión

Menor acuerdo de uso = mayor significado preposicional = mayor entropía de dispersión

Las clases preposicionales, por tanto, se han justificado en el contexto de corpus lingüísticos y a partir de tres clases de heurísticas bien diferenciadas: el aprendizaje automático, la medición entrópica de la diversidad semántica y el acuerdo en el uso preposicional. Esa diversidad en la naturaleza empírica de los procedimientos mejora, según creemos, la consistencia de nuestra investigación.

Además, desde la perspectiva de la robustez del análisis estadístico de los resultados, hemos realizado doble anotación en el experimento 1 y 2; pruebas de azar en la clasificación de los nombres del experimento 2 y 3, y tres controles estadísticos del test sobre el uso preposicional: la representatividad de la muestra del test, la nulidad del efecto del azar en los resultados (valores Z y p) y la aplicación de un coeficiente de coincidencia de anotadores (kappa de Cohen).

Por último, consideramos que existen dos perspectivas para la continuidad explicativa de nuestra hipótesis.

La primera de ellas es de carácter conceptual. Las analogías y las observaciones de diferencias y de similitudes con categorías lingüísticas análogas, en su función y distribución, como los adverbios y las conjunciones, o con partículas y prefijos morfológicos podrían usarse como un nuevo objeto de investigación en la corroboración de la hipótesis de la gradualidad a partir de comparaciones y mediciones nuevas (Biskup, 2000 y 2019; y Gehrke, 2005).

Por otra parte, la segunda perspectiva procede de la naturaleza empírica de los datos. Tanto la afasiología como la comparación tipológica e interlingüística son dos áreas en las que algunas predicciones utilizadas aquí ya han sido exploradas con relativo éxito.

A lo largo de nuestra investigación hemos indicado que la presencia de tipologías lingüísticas diversas respecto de las adposiciones: preposiciones, posposiciones o circunposiciones ha sido una de las dianas que ha hecho cambiar y progresar las predicciones sobre cómo abordar el significado dentro de las teorías gramaticales (Clark, 1978; y Hasplemath, 2005) más allá de las restricciones sintáctico-categoriales o de condiciones a priori sobre la explicación teórica.

Hemos comprobado que en lenguas muy disímiles, tanto en la variación categorial como en la expresión de las piezas (por su número y su forma), se describen fenómenos parecidos a la funcionalidad y la lexicidad, en algunas lenguas eslavas como el ruso, checo y polaco (Comrie y Greville, 1993), en la lengua urálica, el finlandés (Hakulinen, 1978), en el igbo y sus dialectos de la familia níger-congolesa (Sopuruchi y Ibenene, 2018) o el chino mandarín (Liang, 1971; y Ross, 1991). En todas ellas, las formas lingüísticas semejantes a las preposiciones del español admiten fases y grados, tanto de significado como de construcción sintáctico-morfológica.

Finalmente, en la disciplina de la afasiología la separación entre palabras funcionales y palabras léxicas fue un recurso descriptivo muy temprano en la historia del área (Zurif y Caramazza, 1976; Berndt y Caramazza, 1980; Friederici, 1981; Caplan, 1987; Miceli et alia, 1989; y Grodzinsky, 1990). De manera general diríamos que los pacientes con afasias tienden a obviar, a equivocarse o a perder antes las preposiciones funcionales, es decir, aquellas que tienen menos presencia o valor en el establecimiento de las relaciones semánticas mientras mantienen las léxicas en más ocasiones (Tesak y Hummer, 1994; Montes de Oca, 1997; Froud, 2001; Kemmerer, 2003 y 2005; Franco, Zampieri y Meneghello, 2013; Martínez Sánchez, 2008; y Martínez-Ferreiro et alia, 2019). A su vez, desde la perspectiva de los trastornos infantiles del lenguaje se investiga en las dificultades para recuperar, precisamente, las preposiciones más polisémicas (Auza y Morgan, 2013).

La búsqueda y el escrutinio desde estas perspectivas apuntadas podrían mejorar y ampliar, sin duda, el alcance conceptual y empírico de nuestra hipótesis.

La gradualidad semántica es una opción que se manifiesta desde diferentes ópticas y en diferentes fenómenos en las lenguas humanas y que, en el caso de la preposición, es particularmente interesante en la medida que explica algunos hechos lingüísticos que habían quedado relegados, subordinados o reducidos en ciertas teorías lingüísticas.

REFERENCIAS Y BIBLIOGRAFÍA

Abney, Steven (1987): *The English Noun Phrase in its Sentential Aspect*, tesis doctoral, The MIT Press, Cambridge.

Abney, Steven (1996): *Statistical Methods in Linguistics*, en Klavans, Judith L. y Resnik, Philip (1996): *The balancing act: combining symbolic and statistical approaches to languages*, The MIT Press, Cambridge, págs. 1-26.

Abney, Steven (2007): *Semi supervised Learning for Computational Linguistics*, Chapman Hall / CRC Computer Science y Data Analysis, Boca Ratón.

Abney, Steven y Bird, Steven (2011): *Towards a Data Model for the Universal Corpus*, en *Proceeding of the 4th Workshop on Building and Using Comparable Corpora*, Portland, págs. 120-127.

Aboh, Enoch O. y Essegbey, James (eds.) (20333310): *Topics in Kwa Syntax*, Springer, Dordrecht.

Abril, Gonzalo (1997): *Teoría general de la información*, Cátedra, Madrid.

Abur-Mostafa, Yaser S. (2012): *Learning from data*, AMLBooks, California.

Agud, Ana (1980): *Historia y teoría de los casos*, Gredos, Madrid.

Alam, Yukiko Sasaki (2003): *For computational treatment of the polysemy of prepositional uses of Over*, en *Proceedings of the ACL-SIGSEM Workshop on the Linguistic Dimensions of Prepositions and their Use in Computational Linguistics Formalisms and Applications*, Toulouse, págs. 66–76.

Alarcos Llorach, Emilio (1951): *Gramática estructural*, Gredos, Madrid.

Alarcos Llorach, Emilio (1968): *Verbo transitivo, verbo intransitivo y estructura del predicado*, en *Estudios de gramática funcional*, Gredos, Madrid, págs. 7-17.

Alarcos Llorach, Emilio (1972): *Grupos nominales con de en español*, en *Studia hispánica in honorem R. Lapesa*, Gredos, Madrid, págs. 85-91.

Alarcos Llorach, Emilio (1990): *La noción de suplemento en Homenaje al Profesor Francisco Marsá*, Vol I, Jornadas de Filología, Universidad de Barcelona, Barcelona, págs. 209-223.

Alarcos Llorach, Emilio (1995): *Gramática de la lengua española*, Espasa y Calpe, Madrid.

Albertuz Carneiro, Francisco (2007): *Sintaxis, semántica y clases de verbos: clasificación verbal en el proyecto ADESSE*, en Cano López, Pablo (coord.), *Actas del VI Congreso de*

Lingüística General, 3-7 de mayo de 2004, Vol. II, Tomo 2, *Las lenguas y su estructura* (IIb), Santiago de Compostela, págs. 2015-2030.

Alcina, José y Blecua, José Manuel (1975): *Gramática española*, Ariel, Barcelona.

Aijón Oliva, Miguel Ángel (2009): *Complementos más bien indirectos: el régimen verbal como hecho de variación*, Actas del VIII Congreso de Lingüística General, coord. por Antonio Moreno, UAM, Madrid, en <http://elvira.llf.uam.es/clg8/actas/pdf/paperCLG2.pdf>.

Alcoba, Santiago (2000): *Verbos de régimen. Funciones argumentales y función semántica del predicado*, en Fernández González, Jesús; y C. Fernández, María del Carmen (eds.) (2000): *Lingüística para el siglo XXI*, Universidad de Salamanca, Salamanca, págs. 143-155.

Alcoba, Santiago y del Teso, Enrique (2000): *Funciones semánticas argumentales del complemento preposicional de los verbos de régimen*, en *Cien años de investigación semántica, de Michel Breal a la actualidad. Actas del Congreso Internacional de Semántica*, Coordinador por José Manuel Oliver Frade, Cristóbal José Corrales Zumbado, María Laura Izquierdo Guzmán, Dolores García Padrón, Dolores Corbella Díaz, José Secundino Gómez Soliño, Marcos Martínez Hernández, Francisco José Cortés Rodríguez, Vol. II, La Laguna, págs. 929-944.

Almeida, Manuel (2007): *Tres tesis sobre el dequeísmo*, «Revista de Filología», págs. 49-57.

Alonso, César (2008): *Funciones matemáticas que determinan la rección de la preposición “de” o de la preposición “a” en los verbos pronominales*, «RESLA», 21, págs. 107-128.

Alonso Ramos, Margarita (2012): *Explorando la frecuencia léxica para el Diccionario de colocaciones del español*, en *Cum Corde et In Nova Grammatica. Homenaje a Guillermo Rojo*, págs. 19-40. Coordinador Tomás Eduardo Jiménez Juliá (coord.), Belén López Meirama (coord.), Victoria Vázquez Rozas (coord.), y Alexandre Veiga Rodríguez, Universidad de Santiago de Compostela, Santiago de Compostela.

Alonso, Amado (1963): *Introducción a los estudios gramaticales de Andrés Bello*, Universidad Austral de Chile, Valdivia.

Alpaydin, Ethem (2004): *Introduction to Machine Learning*, The MIT Press, Cambridge.

Alsina, Alex (1996): *The role of argument structure in grammar*, CSLI Publications, Stanford.

Alturo, Nuria (2004): *Catalan Adpositions of ‘Movement’ as Aspectual Particles: Cap (a) and Fins (a)*, en Cuyckens, Hubert (ed.) (2004): *Adpositions of Movement*, «Belgian Journal of Linguistics», 18, págs. 197–223.

Álvarez Soriano, Paula y Sánchez, Esther (2009): *La naturaleza aspectual de las preposiciones: los verbos de régimen*, Encuentro Internacional de Jóvenes Lingüistas, UAB, Barcelona.

Álvarez, Sixto Jesús (2000): *Estadística aplicada*, CLAG, Madrid.

Álvez, Javier; Atserias, Jordi; Carrera, Jordi; Climent, Salvador; Oliver, Antoni; y Rigau, German (2015): *Consistent annotation of EuroWordNet with the Top Concept Ontology*, en *Proceedings of the International Conference on Language Resources and Evaluation, European Language Resources Association, LREC 2008*, Marrakech.

Amícola, José (1975): *Observaciones sobre la alternancia para/por en castellano*, «Filología y Didáctica Hispánica», págs. 3-19.

Amiot, Dany y De Mulder, Walter (2002): *De l'adverbe au préfixe en passant par la préposition: Un phénomène de grammaticalisation?* «Linguisticae Investigationes», 25.2, págs. 247-273.

Amiot, Dany (2004): *Préfixes ou prépositions? Le cas de sur-, sans-, contre- et les autres*, «Lexique», 16, págs. 67-87.

Aparicio, Juan José (2015): *Representación computacional de las perífrasis de fase: de la cognición a la computación*, tesis doctoral, Universidad de Barcelona, Barcelona.

Arad, Maya (1995): *On the projection to ditransitive verbs*, en «UCLWPL», 7, University College London, Londres.

Arad, Maya (1999): *Transitivity as a universal: the status of v*, GLOW Newletters, págs. 14-15.

Arens, Hans (1969): *La Lingüística: sus textos y su evolución desde la antigüedad hasta nuestros días*, Vol. I, Gredos, Madrid.

Arnauld, Antoine y Claude Lancelot, Claude (1660): *Grammaire générale et raisonnée, ou La Grammaire de Port-Royal*, edición de Brekle, Herbert E. (1966), Stuttgart-Bad Cannstatt, Frommann.

Arsenijevic, Boban; Baldwin, Timothy; y Trawinski, Beata (eds.) (2006): *Prepositions '06 Proceedings of the Third ACL-SIGSEM Workshop on Prepositions*, Trento.

Asbury, Anna; Dotlačil, Jakub; Gehrke, Berit; y Nouwen, Rick (eds.) (2008): *Syntax and Semantics of Spatial P*, John Benjamins, Utrecht Institut of Linguistics.

Asbury, Anna; Gehrke, Berit; y Hegedús, Veronika (2006): *One size fits all: prefixes, particles, adpositions and cases as members of the category P* en Keskin, Cem (ed.) (2006): *UIL OTS Yearbook*, LOT, Utrecht, págs. 1-17.

Aurnague, Michel; Hickmann, Maya; y Vieu, Laure (eds.) (2007): *The categorization of spatial entities in language and cognition*, John Benjamins, Amsterdam.

Aurnague, Michel y Col, Gilles (2017): *Claude Vandeloise on proximity or the missing piece of a final triptych*, «Corela», HS 23, págs. 1-15.

Auza B., Alejandra y Morgan, Gareth (2013): *Uso de preposiciones en el recuento de una historia. Comparación de niños hispanohablantes con y sin trastorno del lenguaje*, «Infancia y aprendizaje», 36, págs. 35-49.

Báez Montero, Inmaculada C. (1996): *Las frases preposicionales con predicativo en español medieval*, en Casado Velarde et al. (1996): *Scripta Philologica in memoriam Manuel Taboada*, Universidad de Coruña, Coruña, págs. 299-306.

Báez San José, Valerio (1990): *Diccionario informatizado de construcciones oracionales y el proyecto Esquemas sintáctico-semánticos del español*, en G. Wotjak y A. Veiga (comps.) *La descripción del verbo español*, «Verba», anejo 32, págs. 33-70.

Baker, Mark C. (1988): *Incorporation*, University Chicago Press, Chicago.

Baker, Mark C. (1997): *On Particles in Universal Grammar. A review of Den Dikken (1995)*, «Natural Languages and Linguistic Theory», 15, págs. 641-666.

Baker, Mark C. (1995): *The polysynthetic parameter*, Oxford University Press, Oxford.

Baker, Mark C. (2003): *Lexical Categories. Verbs, nouns and adjectives*, Cambridge University Press, Cambridge.

Bakken, Anne Ingebord (2014): *Las preposiciones por y para: estudio sobre los complementos de finalidad con por*, Norges Arktiske Universitet, Narvik.

Bally, Charles (1941): *Les prépositions dans leur rapport avec les verbes transitifs*, en «Cahiers Ferdinand de Saussure», 1, págs. 16-8.

Baldwin, Timothy; Kordoni, Valia; y Villavicencio, Aline (2009): *Prepositions in Applications: A Survey and Introduction to the Special Issue*, «Computational Linguistics», 35.2, págs. 119-149.

Barnbrook, Geoff (1996): *Language and Computers: A Practical Introduction to the Computer Analysis of Language*, Edinburgh University Press, Edimburgo.

Barnes, Betsy (1980): *The Notion of 'Dative' in Linguistic Theory and the Grammar of French*, «Linguisticae Investigationes», IV, págs. 245-292.

Barrajón, Elisa (2006): *Análisis contrastivo locativo-nocional de la CRV en el español hablado en Alicante: usos espaciales y abstractos de salir y entrar*, Universidad de Alicante, Alicante.

Barrajón, Elisa (2014): *La función semántica del complemento preposicional nocional y local*, «Revista de Investigación Lingüística», 17, Universidad de Murcia, Murcia, págs. 11-30.

Barrajón, Elisa; Cifuentes, José Luis; y Rodríguez, Susana (eds.) (2015): *Verb Classes and Aspect*, IVITRA, Benjamins, Dordrecht.

Baroni, Marco y Lenci, Alessandro (2010) *Distributional memory: A general framework for corpus-based semantics*, «Computational Linguistics», 36.4, págs. 673–721.

Baroni, Marco; Boleda, Gemma; Bruni, Marco; y Nam-Khanh Tram (2012): *Distributional semantics in Technicolor*, en *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics*, Long Papers, Vol. I, págs. 136-145.

Barrenechea, Ana María y Orecchia, Teresa (1970): *La duplicación de OODD y OOOI en el español hablado de Buenos Aires*, «Romance Philology», XXIV, págs. 58-83.

Barrero, Tomás (2013): *Acción y construcción lógica*, «Crítica. Revista Hispanoamericana de Filosofía», Vol. LVX, núm. 133, págs. 3-26.

Barretta de, V. D. (1979): *Un intento de definir las preposiciones desde un punto de vista semántico*, «Letras», 36, págs. 61-94.

Barrio, Florencio del (2004): *El régimen de los verbos en español*, tesis doctoral, Universidad de Valladolid, Valladolid.

Barton, Edward; Berwick, Robert C.; y Ristad, Eric Sven (1986): *Computational Complexity and NLP*, The MIT Press, Cambridge.

Bartra, Anna y Suñer, Avel.lina (1992): *Functional Projection Meet Adverbs*, «Catalan Working Papers in Linguistics», 2, págs. 45-86.

Bayer, Josef y Bader, Markus (2007): *On the Syntax of Prepositional Phrases*, en Späth, A. (ed.) *Interface and Interface Conditions*, Walter de Gruyter, Berlín, págs. 157-179. http://www.ub.uni_konstanz.de/kops/volltexte/2008/5795/

Bayley, Robert; Cameron, Richard, Ceil, Lucas (Redactor) (2013): *The Oxford Handbook of Sociolinguistics*, Oxford University Press, Oxford.

Batet Santromà, Montserrat (2010): *Ontology based semantic clustering*, tesis doctoral, Universitat Rovira i Virgili, Tarragona.

Belle, William Van y Langendonck, Willy (eds.) (1996): *The Dative: Descriptive Studies*, John Benjamins, Amsterdam.

Belletti, Andrea y Rizzi, Luigi (eds.) (1996): *Parameters and functional heads*, Oxford University Press, Oxford.

Bello, Andrés (2002): *Gramática de la lengua castellana destinada al uso de los americanos*, Biblioteca Virtual Miguel de Cervantes, Alicante.

Ben-Naim, Arieh (2011): *Entropy desmystified. The Second Law Reduced to Plain Common Sense*, World Scientific Publishing, Londres.

Bengio, Yoshua; Schwen, Holger; Senécal, Jean-Sébastien; Morin, Frédéric; Gauvain, Jean-Luc (2006): *A Neural Probabilistic Languages Model. Studies in Fuzziness and Soft Computing*, págs. 137-186, Springer, Berlín.

Bennet, David C. (1975): *Spatial and Temporal uses of English Prepositions: an essay in stratificational semantics*, Longman, Londres.

Benveniste, Émile, (1958): *Catégories de pensée et catégories de langue*, PUF, París.

Benware, Wilbur (1993): *Representing prepositions: New High German um**, «Linguistics», 31, págs. 135-157.

Berenguer Sánchez, José Antonio (1992): *Distintos conceptos de partícula en la descripción lingüística*, «Revista Española de Lingüística», 22.1, págs. 55-76.

Bezos López, Javier (2005): *Manual de español urgente*, Fundéu BBVA, Debate, Barcelona.

Bhattacharya, Tanmoy (1999) *Numeral/ Quantifier-Classifier as a complex head en Semi-Lexical Categories*, en Corver, Norbert y Van Riemsdijk, Henk (eds.) (2001): *Semi-lexical categories*, Walter de Gruyter, Berlín, págs. 191-222.

Biskup, Petr (2000): *P(refixe)s and P(reposition)s*, ms. Universidad de Leipzig.

Biskup, Petr (2019): *Prepositions, Case and Verbal Prefixes: The case of Slavic*, John Benjamins, Amsterdam.

Biyun, Li (2018): *El complemento de régimen en la lengua española y la china*, «Monográficos Sinoele», 17, págs. 559-569.

Blake, Barry J., (1994): *Case*, Cambridge University Press, Cambridge.

Blansitt, Edward L. (1988) *Datives and allatives* en Hammond, Michael, Moravcsik, Edith A., Wirth, Jessica R. (eds.) (1988) *Studies in Syntactic Typology*, John Benjamins, Amsterdam, págs. 173-192.

Blume, Kerstin (1998): *A contrastive analysis of interaction verbs with datives complements*, «Linguistics», 36.2, págs. 253-280.

Bloom, Paul; Peterson, Mary A; Nadel, Lynn y Garrett, Merrill (1996): *Language and Space*, The MIT Press, Massachusetts.

Bloomfield, Leonard (1933): *Language*, Indiana University Linguistics Club, Bloomington.

- Boer, Minne-Gerben De (1988): *La preposizione di e la nozione di preposizione vuota*, «Lingua e Stile», XXIII, 1, págs. 51-65.
- Boleda, Gemma y Erk, Katrin (2015): *Distributional Semantic Features as Semantic Primitives –or not*, Association for the Advancement of Artificial Intelligence, págs. 2-4.
- Boleda, Gemma y Herbelot, Aurélie (2017): *Formal Distributional Semantics: Introduction to Special Issue*, «Computational Linguistics», págs. 619-635.
- Bolinger, Dwight (1977): *Meaning and form*, Longman, Londres.
- Bolshakov, Igor y Gelbukh, Alexander (2004): *Computational Linguistics: Models, Resources, Applications*, México, IPN-UNAM-FCE.
- Boltzmann, Ludwig (1887): *Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respective den Sätzen über das Wärmegleichgewicht*, Kaiserliche Akademie der Wissenschaften (Vienna), Mathematisch-naturwissenschaftliche Classe, Zweite Abtheilung, Sitzungsberichte, 77, págs. 373-435.
- Borer, Hagit (1984): *Parametric syntax*, Foris Publications, Dordrecht.
- Borer, Hagit (1993): *The projection of arguments*, en Benedicto, Elena y Runner, Jeffrey (eds.) (1993): *Functional Projections*, Universidad de Massachusetts, Amherst, págs. 19-49.
- Borer, Hagit (ed.) (2005): *Structuring Sense. Name Only*, Vol. I, Oxford University Press, Oxford.
- Borillo, Andrée (1993): *Prépositions de lieu et anaphore*, «Langages», 27.110, págs. 27-46.
- Bošković, Željko (2006): *Case Checking versus Case Assignment and the Case of Adverbial NPs*, Linguistic Inquiry, 37.3, págs. 522-533.
- Bosque, Ignacio (1980): *La preposición sin*, «Lingüística Hispánica», 3, págs. 71-85.
- Bosque, Ignacio (1990): *Las categorías gramaticales*, Síntesis, Madrid.
- Bosque, Ignacio (1993): *Preposición tras preposición. Contribuciones al homenaje al Profesor Ramón Trujillo*, Universidad de la Laguna, La Laguna, págs. 133-155.
- Bosque, Ignacio (1998): *Sobre los complementos de medida*, en Delbecque, Nicole y Paepe, Christian de (eds), *Estudios en Honor del Profesor Josse de Kock*, Lovaina, Leuven University Press, 1998, págs. 57-73 .
- Bosque, Ignacio y Demonte, Violeta (eds.) (1999): *Gramática descriptiva de la lengua Española*, Espasa, Madrid.

Bosque, Ignacio (2015): *Inner and outer prepositions with Spanish verbs of vertical movement*, en Barrajon, Elisa; Cifuentes Honrubia, José Luis; y Rodríguez Rosique, Susana (2015): *Verb Classes and Aspect*, IVITRA, Benjamins, Dordrecht, págs. 77–97.

Bosque, Ignacio (2018): *On source prepositions and their relationship with paths*, en António Leal (ed.), *Verbs, Movement and Prepositions*, Centro de Linguística de la Universidad de Porto, Porto, págs.11-44.

Botwinik-Rotem, Irena (2004): *The Category P. Features, projections, interpretation*, tesis doctoral, Tel-Aviv University, Tel-Aviv.

Botwinik-Rotem, Irena (2008): *Why are they different? An exploration of Hebrew locative PPs*, en Asbury, Anna; Dotlačil, Jakub; Gehrke, Berit y Nouwen, Rick (eds.) (2008): *Syntax and Semantics of Spatial P*, John Benjamins, Utrecht Institut of Linguistics, págs. 331-364.

Branchereau, Laurence y Nespoulos, Jean-Luc (1989): *Syntactic parsing and the availability of prepositions in agrammatics patients*, «Aphasiology», 3.4, págs. 411-422.

Brée, David S. Pratt-Hartmann, Ian E. (2002): *Temporal semantics of prepositions in context*, en Feigenbaum, Susanne; y Kurzon, Dennis (2002): *Prepositions in their Syntactic, Semantic and Pragmatic Context*, John Benjamins, Amsterdam.

Brea, Mercedes (1985): *Las preposiciones, del latín a las lenguas romances*, «Verba», 12, págs. 147-182.

Bresnan, Joan (1982a): *The Mental Representation of Grammatical Relations*, The MIT Press, Cambridge.

Bresnan, Joan (2017): *The Garden and the Bush*, «Computational Linguistics», 42.4, págs. 599-617.

Brody, Michael (1998): *The Minimalist Program and a Perfect Syntax: A Critical Note of Noam Chomsky's The Minimalist Program*, «Mind and Language», 13.2, págs. 205-214.

Brondal, Viggo (1940): *Théorie des prépositions. Introduction à une sémantique rationnelle*, Ejnar Munksgaard, Copenhague.

Brucart, Josep Maria (1993): *Gramática generativa y gramática del español*, «Verba», 20, págs. 93-112.

Brugman, Claudia (1981): *The story of over*, Universidad de California, California.

Brugman, Claudia (2001): *Light Verbs and Polisemy*, «Language Science», 23, págs. 551-578.

Bubenik, Vit (2006): *Cases and Prepositions in Iranian*, en Hewson, John y Bubenik, Vit (2006): *From Case to Adposition: The development of configurational syntax in Indo-European languages*, «Current Issues in Linguistic Theory», 280, págs. 131-159.

Butt, Miriam y Geuder, Wilhem (2001): *On semilexical status of Light Verbs*, en Corver, Norbert et alia (2001): *Semi-lexical categories: the function of content words and the content of function words*, Mouton de Gruyter, Berlín, págs. 323-370.

Cabezas, Emilio (2014a): *Entre la locación y la colección: sobre las propiedades léxicas de una preposición bivalente*, DICENDA, 32, págs. 7-31.

Cabezas, Emilio (2014b): *Sobre usos de la preposición entre. El sentido realizativo*, Círculo de Lingüística aplicada a la Comunicación, 66, págs. 3-21.

Calero Vaquera, María Luisa (1986): *Historia de la gramática española*, Madrid, Gredos.

Campbell, Lyle (2002): *The History of Linguistics* en Aronoff, Mark y Miller, Jeanie Rees (eds.) *The Handbook of Linguistics*, Blackwell Publishing, págs. 1-17.

Campos, Héctor (1990): *Preposition Stranding in Spanish*, ms., Georgetown University. Washington.

Campillos Llanos, Leonardo (2014): *Las preposiciones en el habla no nativa de nivel intermedio: análisis de la interlengua basado en corpus*, «Revista Nebrija de Lingüística Aplicada», 16, págs. 5-27.

Canavos, George C. (1988): *Probabilidad y estadística. Aplicaciones y métodos*, McGraw-Hill, México.

Cano, Rafael (1981): *Estructuras sintácticas transitivas*, Madrid, Gredos.

Cano, Rafael (1982): *Sujeto con preposición en español y cuestiones conexas*, «Revista de Filología Española», 62, págs. 211-258.

Cano Cambronero, María Ángeles (2009): *La entrada léxica de los verbos de movimiento: los contenidos sintácticamente relevantes*, «ELUA», 23, págs. 51-72.

Carbonero Cano, Pedro (1975): *Funcionamiento lingüístico de los elementos de relación*, Universidad de Sevilla, Sevilla.

Candalija Reina, José Antonio (2006): *Preposiciones vacías del español: complemento de régimen preposicional*, tesis doctoral, Universidad de Alicante, Alicante.

Candalija Reina, José Antonio (2013): *Transitividad preposicional: el doble complemento preposicional en español*, «Revista de Investigación Lingüística», Universidad de Murcia, Murcia, págs. 177-207.

Caplan, David (1993): *Language. Structure, processing and disorders*. Cambridge University Press, Cambridge.

Caravedo, Rocío (1999): *Lingüística del Corpus. Cuestiones teórico-metodológicas aplicadas al español*, Ediciones Universidad de Salamanca, Salamanca.

Carnot, Nicolas Leonard Sadi (1824): *Réflexions sur la puissance motrice du feu et sur les machines propres a développer cette puissance*, Bachelier, París.

Carroll, Sean (2009): *From Eternity to Here*, Plume, Nueva York.

Casadei, Federica (2001): *Le locuzioni preposizionali. Struttura lessicale e gradi di lessicalizzazione*, «Lingua e Stile», 36.1, págs. 43-79.

Casado Velarde, Manuel et alia (1996): *Scripta Philologica in memoriam Manuel Taboada*, Universidad de Coruña, Coruña, págs. 299-306.

Castellón, Irene; Coll-Florit, Marta; Gil-Vallejo, Lara; y Turno, Jordi (2015): *Hacia una clasificación verbal automática para el español: estudio sobre la relevancia de los diferentes tipos y configuraciones de información sintáctico-semántica*, «Linguamática», 7, págs. 41-52.

Cervoni, Jean (1991): *La préposition. Étude sémantique et pragmatique*, Duculot, París.

Cetinsoy, Atakan; Martín, Francisco. J; Ortega, José Antonio; y Petersen, Paul (2015): *The Past, Present and Future of Machine Learning APIs*, JMLR, en *Workshop and Conference Proceedings*, 50, págs. 43–49.

Cienki, Alan J. (1989): *Spatial Cognition and the Semantics of Prepositions in English, Polish, and Russian*, Otto Sagner, Munich.

Chapado Chorro, M^a Olga (1996): *Estudio tipológico de las adposiciones*, tesis doctoral, UAM, Madrid.

Chapelle, Olivier; Schölkopf, Bernhard; y Zien, Alexander (eds.) (2006): *Semi-supervised Learning*, The MIT Press, Cambridge.

Charniak, Eugene (1993): *Statistical Language Learning*, The MIT Press, Cambridge.

Cheng, Winnie (2012): *Exploring Corpus Linguistics. Language in Action*, Routledge, Londres.

Chevalier, Jean-Claude (1968): *Histoire de la syntaxe. Naissance de la notion de complément dans la grammaire française (1530-1750)*, Droz, Ginebra.

Chilton, Paul (2005): *Vectors, viewpoint and viewpoint shift*, «Annual Review of Cognitive Linguistics», 3, págs. 78-116.

Chomsky, Noam (1965): *Aspects of Syntactic Theory*, The MIT Press, Cambridge, Massachussets.

Chomsky, Noam (1970): *Remarks on nominalization*, en Jacobs, Roderick A. y Rosebaum, Peter (eds.): *Reading in English Transformational Grammar*, Waltham, Ginn, págs. 184-221.

Chomsky, Noam (1981): *Lectures on Government and Binding*, Foris Publications, Dordrecht.

Chomsky, Noam (1995): *The Minimalist Program*, The MIT Press, Massachussets.

Church, Kenneth W. y Mercer, Robert L. (1993): *Introduction to Special Issues on Computational Linguistics Using Large Corpora*, «Computational Linguistics», 19.1, págs. 1-24.

Cifuentes Honrubia, José Luis (1994): *Gramática cognitiva: fundamentos críticos*, Eudema, Madrid.

Cifuentes Honrubia, José Luis (1996a): *Usos prepositivos en español*, Universidad de Murcia, Murcia.

Cifuentes Honrubia, José Luis y Llopis Ganga, Jesús, (1996b): *Complemento indirecto y complemento de lugar: estructuras locales de base personal en español*, Universidad de Alicante, Alicante.

Cifuentes Honrubia, José Luis (2003): *Locuciones prepositivas. Sobre la gramaticalización preposicional en español*, Publicaciones de la Universidad de Alicante, Alicante.

Cifuentes Honrubia, José Luis y Rodríguez Rosique, Susana (2015): *Verb Classes and Aspect*, «IVITRA», 9, págs. 77–97.

Cienki, Alan J. (1989): *Spatial Cognition and the semantics of prepositions in English, Polish, and Russian*, Sagner, Munich.

Cinque, Guglielmo (1999): *Adverbs and Functional Heads. A Cross-linguistic perspective*, Oxford University Press, Oxford.

Cinque, Guglielmo y Rizzi, Luigi (eds.) (2010): *Mapping Spatial PPs. The Cartography of Syntactic Structures*, Oxford University Press, Oxford.

Clark, Alexander; Fox, Chris; y Lappin, Shalom (2013): *The Handbook of Computational Linguistics and Natural Language Processing*, Wiley-Blackwell, Oxford.

Clark, Eve V. (1978): *Locationals: existentials, locative and possessive constructions*, en Greenberg, Joseph (1978): *Universals of Human Language*, Vol. IV, págs. 85-126, Stanford University Press, Stanford.

Clausius, Rudolf (1867): *The Mechanical Theory of Heat*, Taylor and Francis, Londres.

Codoñer, Carmen (1983): *Nebrija y la introducción del Renacimiento en España*, en *Actas de la III Academia Literaria Renacentista*, Universidad de Salamanca, 9, 10 y 11 de diciembre de 1981, Salamanca, pág. 105-143.

Codoñer, Carmen y González, Juan Antonio (1994): *Antonio de Nebrija: Edad Media y Renacimiento*, Ediciones de Salamanca, Salamanca.

Coene, Martine y D'Hulst, Yves (2003): *From NP to DP. The Syntax and semantic of noun phrases. Vol I*, John Benjamins, Amsterdam.

Cohen, Jacob (1960): *A coefficient of agreement for nominal scales*, «Educational and Psychological Measurement», 20.1, págs. 37–46.

Coll-Florit, Mercè (2009): *La modalitat de l'acció. Anàlisi empírica, reformulació teòrica i representació computacional*, tesis doctoral, Universidad de Barcelona, Barcelona.

Colombo, Lucia y Flores d'Arcais, Giovanni B. (1984): *The meaning of Dutch prepositions: a psycholinguistic study of polysemy*, «Linguistics», 22, págs. 51-98.

Cook, Walter Anthony (1989): *Case Grammar Theory*, Georgetown University Press, Washington.

Company Company, Concepción (2019): *Grammatical Words and Spreading of Contexts: Evidence from the Spanish Preposition a*, «Languages», 4.10, págs. 1-16. doi:10.3390/languages4010.

Comrie, Bernard y Corbett, Greville G. (eds.) (1993): *The Slavonic Languages*, Routledge, Londres.

Confortiová, Helena (1987): *On the Semantic Analysis of Prepositions from the Quantitative Point of View*, en Hajičová, Eva; Těšitelová, Marie y Horecký, Ján (eds.) (1987): *Prague Studies in Mathematical Linguistics*, Vol. IX, John Benjamins, Amsterdam.

Corbett, Greville; Fraser, Norman; y McGlashan, Scott (eds.) (1993): *Heads in grammatical theory*, Cambridge University Press, Cambridge.

Corder, Gregory W., Foreman, Dale I. (2009): *Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach*, Blackwell, Londres.

Cormack, Annabel y Breheny, Richard (1994): *Projections for functional categories*, «Working Papers of London», Universidad de College of London, Londres, págs. 35-61.

Corver, Norbert y Van Riemsdijk, Henk (2001): *Semi-lexical categories*, Walter de Gruyter, Berlín.

Coseriu, Eugenio (1972): *Sobre las categorías verbales*, en *Gramática, semántica y universales*, Gredos, Madrid, págs. 50-79.

Coulmas, Florian (1996): *The Handbook of Sociolinguistics*, Oxford, Blackwell.

Coventry, Kevin y Ludwig, Andreas (1991): *Semantic of Prepositions. A literature survey and proposal framework for future treatment*, Centre for Cognitive Science Research Paper EUCCS/RP-45, Universidad de Edimburgo.

Coventry, Kenny R.; Carmichael, Richard; y Garrod, Simon C. (1994): *Spatial Prepositions, Object-Specific Function and Task Requirements*, *Journal of Semantics*, 11, págs. 289-309.

Crego, María Victorina (1997): *Las preposiciones como marcadores sintáctico-semánticos en Molina Redondo, José Andrés de y Luque Durán, Juan de Dios (1997): Estudios de lingüística general*, Método Ediciones, Granada, págs. 57-67,

Cresswell, Michael (1985). *Prepositions and points of view*, «*Linguistics and Philosophy*», 2:1-41.

Croft, William y Cruse, D. Alan (2004): *Cognitive Linguistics*, Cambridge University Press, Cambridge.

Cuadros, Roberto (2005): *La complementación verbal. Viejos y nuevos enfoques*, «*Language Design*», 5, págs. 105-136.

Cuartero Otal, Juan (2006): *¿Cuántas clases de verbos de desplazamiento se distinguen en español?*, «*Rilce*», 22.1, págs. 13-36.

Cuervo, Rufino (2002): *Diccionario de construcción y régimen de la lengua castellana*, Herder, Barcelona.

Cuesta, Paloma (1997): *El Suplemento: Problemas de caracterización y tipología*, BRAE, T. LXXVII. C. CCLXX, págs. 57-98.

Cuyckens, Hubert (1984): *The deictic vs. non-deictic ambiguity in the spatial uses of prepositions*, «*Linguisticae Investigationes*», 8.1, págs. 187-194.

Cuyckens, Hubert (1995): *Family resemblance in the Dutch spatial prepositions*, «*Cognitive Linguistics*», 6.2/3, págs. 183-207.

Cuyckens, Hubert (1997): *Prepositions in Cognitive Lexical Semantics*, en Dagmar Haumann, Herausgegeben von y Schierholz, Stefan J. (eds.) (1997): *Lexicalische und grammatische Eigenschaften präpositionaler Elemente*, Niemeyer, Tübingen, págs. 63-82.

Cuyckens, Hubert y Radden, Günter (2002): *Perspectives on Prepositions*, Niemeyer, Tübingen.

Czepluch, Hartmut (1982): *Case theory and the Dative Constructions*, «*Linguistic Review*», 2, págs. 1-38.

Davies, Mark (2002-2008): *Corpus del español*, Brigham Young University, Provo [en línea: www.corpusdelespanol.org].

Davidson, Donald (1967): *The logical form of actions sentences*, en Rescher, Nicholas (ed.), *The Logic of Decision and Action*, University Pittsburgh Press, Pittsburgh, págs. 81-95.

Deacon, Robert Joel (2014): *Adpositions and Lexical Categories Distributed Morphology's insight into the Preposition*, «Florida Linguistics Papers», Vol. I, págs. 1-12.

Déchaine, Rose-Marie (1989): *Verbs and Prepositions*, ms. Amherst, Universidad de Massachusets.

Déchaine, Rose-Marie (1993): *Predicates across categories*, tesis doctoral, Universidad de Massachusets, Amherst.

Déchaine, Rose-Marie (2005): *Grammar at the Borderline: A Case Study of P as a Lexical Category*, en Alderete, John et alia (ed.) (2005): *Proceedings of the 24th West Coast Conference on Formal Linguistics*, Somerville, MA.

Déchaine, Rose-Marie y Tremblay, Mireille (1996): *Adverbial PPs and prepositional adverbs in French and English*, en *CLA Annual Conference Proceedings*, «Working Papers in Linguistics», Universidad de Calgary, Calgary.

Delbecque, Nicole y Lamiroy, Béatrice (1996): *Towards a typology of the Spanish dative*, en Belle, William Van y Langendonck, Willy (eds.) (1996): *The Dative: Descriptive Studies*, Vol. I, John Benjamins, Amsterdam, págs. 73-118.

Delbecque, Nicole y Lamiroy, Béatrice (1998): *The possessive dative in Romance and Germanic languages*, en Van Belle, William y Van Langendonck, Willy (eds.) (1998): *Case and Grammatical Relations Across Languages*, Vol. II, John Benjamins, Amsterdam, págs. 29-74.

Dehé, Nicole; Jackendoff, Ray; McIntyre, Andrew y Urban, Silke (eds.) (2002): *Verb-Particle Explorations*, Walter de Gruyter, Berlín.

Demonte, Violeta (1991): *Temporal and aspectual constraints on predicative Aps*, en Campos, Héctor y Martínez-Gil, Fernando (eds.) (1991): *Current Studies in Spanish Linguistics*, Georgetown University Press, Georgetown, págs. 165-200.

Demonte, Violeta (1992): *Linking and case: The case of prepositional verbs*, en Laeufer, Christiane y Morgan, Terrell A. (ed.) (1992): *Theoretical Analyses in Romance Linguistics: Selected papers from the Linguistic Symposium on Romance Languages XIX*, Ohio State University, 21–23 de abril, John Benjamins, Amsterdam.

Demonte, Violeta (2002): *Microvariación sintáctica en español. Rasgos, categorías y virus*, en Álvarez Santullano, P.; Rosas, Minerva; y Contreras, Manuel. (eds.) (2002):

Texto, lingüística y cultura, Actas del XIV Congreso de la sociedad chilena de lingüística, Editorial Universidad de los Lagos, Osorno, págs. 9-40.

Demonte, Violeta (2011): *Los eventos de movimiento en español. Construcción léxico-sintáctica y microparámetros preposicionales*, en Cuartero Otal, Juan; García, Luis Fernández; y Sinner, Carsten (eds.) (2011): *Estudios sobre perífrasis y aspecto*, München Peniopo, Munich, págs. 16-42.

Den Dikken, Marcel (1995): *Particles*, Oxford University Press, Nueva York.

Den Dikken, Marcel y Tortora, Christina M. (eds.) (2005): *The Function of Function Words and Functional Categories*, John Benjamins, Amsterdam.

DRAE (2014): *Diccionario de la Real Academia Española de la lengua*, Espasa y Calpe, Barcelona.

Echarte, María José (1996) *Acusativo y dativo: dinámica sincrónica del latín al castellano*, «Revista Española de Lingüística», 23.1, Madrid, págs. 86-107.

Emonds, Joseph E. (1985): *A Unified Theory of Syntactic Categories*, Foris Publications, Dordrecht.

Emenanjo, N. E. (2010): *The expression of prepositional notion in Igbo*, en Uchekukwu, Chinedu y Mbah, B. M. (eds.), *The Preposition in Igbo*, Edumail Publications, Onitsha, págs. 5-13.

Erk, Katrin (2012): *Vector Space Models of Word Meaning and Phrase Meaning: A Survey*, «Language and Linguistics Compass», 6-10, págs. 635-653.

Escalígero, César (2004): *De Causis linguae latinae*, Edición de Institución Cultural “El Brocense”, Cáceres. Basada en la edición de Lyon de 1540.

Escandell Vidal, María Victoria (2000): *Categorías funcionales y semántica procedimental*, en Martínez, M. et alia (eds.) (2000): *Cien años de investigación semántica: De Michel Bréal a la actualidad*, Ed. Clásicas, Madrid, págs. 363-378.

Escandell Vidal, María Victoria (2004): *Fundamentos de Semántica composicional*, Ariel, Barcelona.

Espinal, María Teresa (coord.) (2014): *Semántica*, Akal, Madrid.

Evans, Vyvyan (2007): *A Glossary of Cognitive Linguistics*, Edinburgh University Press, Edimburgo.

Evans, Vyvyan y Tyler, Andrea (2003): *The Semantics of English Prepositions*, Cambridge University Press, Cambridge.

Evans, Vyvyan y Chilton, Paul (eds.) (2010): *Language, Cognition and Space*, Equinox Publishing, Londres.

Everitt, Brian S. (1992): *The Analysis of Contingency Tables*, Chapman & Hall-CRC Press Books, Londres.

Fábregas, Antonio (2007): *An Exhaustive Lexicalisation Account of Directional Complements*, Tromsø Working Papers on Language & Linguistics, Nordlyd, 34.2, número especial editado por Bašić, Monika; Pantcheva, Marina; Son, Minjeong; y Svenonius, Peter (eds.): 2007: *Space, Motion, and Result*, CASTL, Tromsø, págs. 165–199. <http://www.ub.uit.no/baser/nordlyd/>

Fábregas, Antonio (2010): *On Spanish Prepositional Prefixes and Cartography of Prepositions*, «Catalan Journal of Linguistics», 9, págs. 55-77.

Fábregas, Antonio (2015) : *Una nota sobre locativo y acusativo*, «Archivium», LXV, págs. 57-74.

Fagard, Benjamin; y Mardale, Alexandru (2019): *Les prépositions complexes du roumain – approche sur corpus*, en Fagard, Benjamin; Pinto De Lima, José ; y Stosic, Dejan (eds.) (2019): *Les prépositions complexes dans les langues romanes*, «Revue Romane», 54.1, págs. 174–202.

Fasold, Ralph (1990): *The Sociolinguistic of Language*, Blackwell, Oxford.

Feigenbaum, Susanne y Kurzon, Dennis (2002): *Prepositions in their syntactic, semantic and pragmatic context*, John Benjamins, Amsterdam.

Fernández López, María del Carmen (1995): *Errores e interlingua en el aprendizaje del español como lengua extranjera*, «Didáctica, 6, págs. 203-216.

Fernández López, María del Carmen (1999): *Las preposiciones en español. Valores y usos. Construcciones preposicionales*, Ediciones Colegio de España, Madrid.

Fernández Pérez, Milagros (1996): (ed.): *Avances en lingüística aplicada*, Universidad de Santiago de Compostela, Santiago de Compostela.

Fernández Vallejo, Jesús (1994): *Adquisición y uso de las preposiciones españolas para un francófono*, REALE, 2, págs. 47-63.

Fernández-Alcalde, Héctor (2013): *Datives, prepositions and argument structure in Spanish*, «Romance Languages and Linguistic Theory», págs. 125–142.

Ferrer-i-Cancho, Ramon y Reina, Francesc (2002): *Quantifying the Semantic Contribution of Particles*, «Journal of Quantitative of Linguistics», 9.1, págs. 35-47.

- Ferrer-i-Cancho, Ramon y Solé, Ricard (2001): *Two Regimes in the Frequency of Words and the Origin of Complex Lexicons*, «Journal of Quantitative Linguistics», 8, págs. 165-173.
- Ferrer-i-Cancho, Ramon y Vitevitch, Michael S. (2018). *The origins of Zipf's meaning-frequency law*, «Journal of the American Society for Information Science Technology», 69.11, págs. 1369-1379.
- Fillmore, Charles J. (1968): *The Case for Case*, en Bach, E. y Harms, R. T. (eds.) (1968): *Universals in Linguistic Theory*, Rinehart and Winston, Nueva York, págs. 1-88.
- Firth, John R. (1957): «A synopsis of linguistic theory 1930-1955» en *Studies in Linguistic Analysis* (Oxford: Philological Society), págs. 1-32. Reimpreso en Palmer, F. R. (ed.) (1968): *Selected Papers of J. R. Firth (1952-1959)*, Longman, Londres.
- Folgar de la Calle, José María (1983): *El análisis de las preposiciones en español: los sistemas de Cuervo y López*, «Verba», 10, págs. 183-201.
- Folli, Raffaella y Ranchand, Gillian (2005): *Prepositions and results in Italian and English: an analysis from event decomposition*, en Verkuyl, Henk; de Swart, Henriette; y van Hout, Angelica (eds.): *Perspectives on Aspect*, Springer, Berlín, págs. 81-105.
- Franco, Ludovico (2012): *Prepositions inside (and at the edge) of words: a view from agrammatism*, «Language Sciences», 40, págs. 95-122.
- Franco, Ludovico; Elisa Zampieri, Bencini, Giulia; y Meneghello, Francesca (2013): *On the production of axial prepositions: linking Figure and Ground in Broca's aphasia. A case study*, «Language and Information Society», 19, págs. 59-100.
- Franco, Ludovico (2018): *Improper prepositions in Old and Modern Italian*, <https://www.researchgate.net/publication/329782311> *Improper prepositions in Old and Modern Italian*.
- Freedman, David; Pisani, Robert; Purves, Roger; y Adhikari, Ani (1978): *Estadística*, Antoni Bosch Editor, Barcelona (traducción de 1991).
- Frege, Gottlob Fried (1892): *Über Sinn und Bedeutung*, en *Zeitschrift für Philosophie und philosophische*, KritikNF 100, págs. 25–50. En español en Gottlob Frege (1996): *Escritos Filosóficos*, Crítica, Barcelona (traducción de Jesús Monterín).
- Friederici, Angela D. (1981): *Production and comprehension of prepositions in aphasia*, «Neuropsychologia», 19.2, págs. 191-9.
- Froud, Karen (2001): *Prepositions and the lexical/functional divide: Aphasia evidence*, «Lingua», 111, 1-28.

Funes, María Soledad (2010): *La alternancia de las preposiciones por y de en las construcciones causales*, «Entrehojas: Revista de Estudios Hispánicos», 1.1, disponible en <http://ir.lib.uwo.ca/entrehojas/vol1/iss1/5>.

Funes, María Soledad (2011a): *El significado de la preposición 'de': un enfoque cognitivo-prototípico*, «Boletín de Filología», Tomo XLVII, Número 1, págs. 111-135.

Funes, María Soledad (2011b): *La alternancia de las preposiciones Por y De como introductorias del complemento agente*, ELUA, 25, págs. 135-161.

Fukui, Nakoi y Speas, Peggy (1986): *Specifiers and Projection*, «MIT Working Papers in Linguistics», 8, págs. 128-172.

Gaätone, David (2001): *Les prépositions: une classe aux contours flous*, «Travaux de linguistique», 42-43, págs. 23-31.

Gaätone, David (2004): *Les prepositions forment-elles una classe?* En Leclère, Christian ; Laporte, Éric ; Piot, Mireille ; y Silberztein, Max (eds.) (2004) : *Syntaxe et Lexique-Grammaire / Syntax, Lexis & Lexicon-Grammar: Papers in honour of Maurice Gross*, págs. 211–221.

Galán Sánchez, Pedro Juan (2004): *Influencias del De causis linguae Latinae de Escalígero en el Minerva de El Brocense*, «Minerva. Revista de Filología Clásica», 17, págs. 139-164.

Gallego, Ángel J. (2010a): *On the prepositional nature on non-finite verbs*, «Catalan Journal of Linguistics», 9, págs. 79-102.

Gallego, Ángel J. (2010b): *El complemento de régimen verbal*, «Lingüística Española Actual», XXXII, págs. 223-258.

Gallego, Ángel J. (2011): *“Cuando” ¿preposición o adverbio relativo?* En Escandell Vidal, María Victoria; Manuel Leonetti Jungl, María Cristina Sánchez López (coord.): *60 problemas de gramática: dedicados a Ignacio Bosque*, Akal, Madrid, págs. 9-17.

Gallego, Ángel J. (2016): *Three types of prepositions in spanish se sentences. consequences for cross-dialectal studies*, Dialectología, 17, págs. 51-70.

Garachana, Mar (2019): *The History of the Spanish Preposition Mediante. Beyond the Theory of Grammaticalization*, «Languages», 4.26, págs. 1-24.

García Martín, José María (1991): *Sobre la localización originaria del uso del OD Preposicional peculiar del castellano: el testimonio de los documentos notariales prealfonsíes*, en *Actas del III Congreso de Lingüística General*, Universidad de Salamanca, Salamanca.

García-Miguel, José María (1991): *La duplicación de complemento directo e indirecto como concordancia*, «Verba», 18, págs. 375-410.

García-Miguel, José María (1995a): *Transitividad y complementación preposicional en español*, «Verba», anexo 40, Universidad de Santiago de Compostela, Santiago de Compostela.

García-Miguel, José María (1995b): *Régimen y caso en las primeras gramáticas del español*, en *Scripta Philologica en Homenaje a Manuel Taboada*, Universidad de la Coruña, La Coruña, págs. 109-121.

García-Miguel, José María y Albertuz, Francisco (2005): *Verbs, semantic classes and semantic roles in ADESSE project*, en *Proceedings of the Interdisciplinary Workshop on the Identification and Representation of Verb Features and Verb Classes*, Saarbrücken págs. 50-55.

García Meseguer, Álvaro (2007): *Nombres concretos y abstractes: una propuesta de definición basada en pruebas sintácticas*, «ELUA», 21, págs. 1-27.

García Yebra, Valentín (1988): *Claudicación en el uso de preposiciones*, Gredos, Madrid.

Gazdar, Gerald y Mellish, Chris (1989): *Natural Language Processing in PROLOG, An Introduction to Computational Linguistics*, Addison-Wesley, Wokingham.

Gawron, Jean (1986): *Situations and prepositions*, «Linguistics and Philosophy», 9.3, págs. 327-382.

Geeraerts, Dirk (1989): *Prospects and problems of prototype theory*, «Linguistics», 27.4, págs. 587-612. Recogido en Geeraerts, Dirk (ed.) (2006): *Cognitive Linguistics: Basic Reading*, Mouton de Gruyter, Berlín, págs. 141-165.

Geeraerts, Dirk (2006) (ed.): *Cognitive Linguistics: Basic Reading*, Mouton de Gruyter, Berlín.

Geeraerts, Dirk y Cuyckens, Hubert (eds.) (2012): *The Oxford Handbook of Cognitive Linguistics*, Oxford University Press, Oxford.

Gelderen, Elly Van (1993): *The rise of Functional Categories*, John Benjamin, Amsterdam.

Gehrke, Berit (2005): *The prepositional aspect of Slavic prefixes and the goal-source asymmetry*, ESSLLI workshop on Formal Semantics and Cross-Linguistic, Edimburgo.

Gehrke, Berit (2006): *Where is the path locative prepositions*, Sinn und Bedeutung, Barcelona.

Gehrke, Berit (2008): *Ps in Motion: On the Semantics and Syntax of P Elements and Motion Events*, LOT, Utrecht.

Gili Gaya, Samuel (1943-1961) (1980): *Curso Superior de Sintaxis Española*, Bibliograf, Barcelona.

Giorgi, Alessandra y Longobardi, Giuseppe (1991a): *The syntax of Noun Phrases. Configuration, parameters and empty categories*, Cambridge University Press, Cambridge.

Giorgi, Alessandra (1991b): *Prepositions, binding and Θ -marking*, en Koster, Jan y Reuland, Eric (eds.) (1991): *Long-distance anaphora*, Cambridge University Press, Cambridge, págs. 185-208.

Girju, Roxana (2008): *The Syntax and Semantics of Prepositions in the Task of Automatic Interpretation of Nominal Phrases and Compounds: A Cross-Linguistic Study*, «Computational Linguistics», 35.2, págs. 185-228.

Girón, José Luis y Sáez, Daniel (2014): *Procesos de gramaticalización en la historia del español*, Iberoamericana Vervuert, Madrid.

Gómez Asensio, José J. (1981): *Gramática y categorías verbales en la tradición española (1771-1847)*, Ediciones de la Universidad de Salamanca, Salamanca.

Gómez Asensio, José J. (1985): *Subclases de palabras en la tradición española (1771-1847)*, Ediciones de la Universidad de Salamanca, Salamanca.

González Hernández, Ana Teresa (2007): *Lingüística contrastiva y traducción: algunas consideraciones sobre la traducción de las preposiciones en francés y en español*, en Bonnet, Dominique; Chaves García, María José; y Duchêne, Nadia (coords.), *Littérature, langages et arts: rencontres et création*, Universidad de Huelva, Huelva [CD-ROM], s.p.

Grabowski J, Weiß P. (1996): *The prepositional inventory of languages: A factor that affects comprehension of spatial prepositions*, «Language Sciences»18.1-2, págs. 19-35.

Granvik, Anton (2012): *Estudio histórico-comparativo de los usos y la semántica de la preposición de en español*, tesis doctoral, Societé Néophilologique, Helsinki.

Granvik, Anton (2014): *Hablando de, sobre, y acerca de la gramaticalización y la lexicalización: panorama diacrónico de las relaciones entre preposiciones y locuciones prepositivas dentro del campo semántico TEMA/ASUNTO*, dentro de Girón, José Luis y Sáez, Daniel (2014): *Procesos de gramaticalización en la historia del español*, Iberoamericana Vervuert, Madrid, págs. 77-118.

Grimshaw, Jane (1990): *Argument Structure*, The MIT Press, Cambridge, Massachussets.

Grimshaw, Jane (1991): *Extended Projection*. Ms, Brandeis University, Waltham, MA.

Gries, Stephan Th. (1999): *A cognitive and functional approach*, «Cognitive Linguistics», 10.2, págs. 105-145.

Gries, Stephan Th. y Steffanowitsch, Anatol (2006): *Corpora in cognitive linguistics*, Mouton de Gruyter, Berlín.

Grishman, Ralph (1986): *Introducción to Computational Linguistics*, Cambridge University Press, Cambridge.

Gross, Maurice (1981): *Les prépositions composées*, en Schwarze, Christoph (ed.) (1981): *Analyse des prépositions*, Niemeyer, Tübingen, págs. 2-25.

Grodzinsky, Yosef (1988): *Syntactic Representations in Agrammatic Aphasia: The Case of Prepositions*, «Language and Speech», 31.2, págs. 115-131.

Grodzinsky, Yosef (1989): *Comparative Aphasiology*, «Aphasiology», 33, págs. 111-142.

Grodzinsky, Yosef (1990): *Theoretical Perspective on Language Deficits*, Cambridge University Press, Cambridge, págs. 9-39.

Guéron, Jacqueline (1990): *On the Syntax and Semantics of PP Extraposition*, «Linguistic Inquiry», 11.4, págs. 637-678.

Guardón, María del Carmen (2003): *El concepto de trayectoria y las preposiciones in y on*, «EPOS», 19, págs. 179-192.

Guimier, Claude (1981): *Prepositions: an analytical bibliography*, John Benjamins, Amsterdam.

Gutiérrez Araus, María Luz (1978): *Estructuras sintácticas del español actual*, Madrid, SGEL.

Gutiérrez Araus, María Luz (1986): *El complemento verbal de régimen preposicional en la gramática española*, «Boletín de la Academia Puertorriqueña de la Lengua Española», 14, págs. 43-60.

Gutiérrez Araus, María Luz (1987): *Sobre la transitividad preposicional en español*, «Verba», 14, págs. 367-381.

Gutiérrez Araus, María Luz (1991): *Valores atípicos de las preposiciones*. En Luna, Elisabeth (coord.) (1991): *Scripta Philologia in Honorem J. M. Lope Blanch*, UNAM, México, págs. 59-78

Gutiérrez Ordóñez, Salvador (1978): *Grupos sintagmáticos N de N: sintaxis y semántica*, «Archivum», págs. 381-396, también en *Estudios ofrecidos a Emilio Alarcos Llorach*, Vol. III, Universidad de Oviedo, Oviedo, págs. 133-159.

Gutiérrez Ordóñez, Salvador (1994): *Problemas en torno a las categorías funcionales*, en Hernández Paricio, Francisco (ed.) (1994): *Perspectivas sobre la oración*, Zaragoza, págs. 71-99.

Gutiérrez Ordóñez, Salvador (1996): *¿Hablamos del suplemento?* en *Scripta Philologica in Memoriam Manuel Taboada Cid*, Universidad de la Coruña, La Coruña.

- Ha, Le Quan; Sicilia-Garcia, Elvira I.; Ming, Ji; y Smith, Francis J (2002): *Extension of Zipf's law to words and phrases*, en COLING, *Proceeding of the 19th International Conference on Computational Linguistics*, Taipei, págs. 1-6.
- Habert, Benoit; Nazarenko, Adeline; Salem, André (1997): *Les linguistiques de corpus*, Armand Colin, París.
- Hale, Ken y Keyser, Samuel J. (1991): *On the syntax of argument structure*, Center for Cognitive Science, The MIT Press, Cambridge.
- Hale, Kenneth; Keyser, Samuel J. (1993): *On the argument structure and the lexical expression of syntactic relations*, en Hale, Kenneth y Keyser, Samuel (eds.) (1993): *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger*, The MIT Press, Cambridge, págs. 53-109.
- Hale, Ken y Keyser, Samuel J. (ed.) (1993): *The view from Building 20th*, Cambridge, The MIT Press, Massachusetts.
- Hale, Ken y Keyser, Samuel J. (1993): *On argument structure and the lexical expression of syntactic relations*, en *The view from Building 20th*, Cambridge, The MIT Press, Cambridge, Massachusetts.
- Hale, Ken y Keyser, Samuel J. (2002): *Prolegomenon to a Theory of Argument Structure*, The MIT Press, Cambridge, Massachusetts.
- Halliday, Michael A. K (1961): *Categories of the theory of grammar*, «Word», 17.3, págs. 241-292.
- Hagège, Claude (2010): *Adpositions*, Oxford University Press, Oxford.
- Hakulinen, Lauri (1979): *Suomen kielen rakkene ja kehitys*, Otava, Keuruu.
- Hartrumpf, Sven; Helbig, Hemann; y Osswald, Kainer (2006): *Semantic Interpretation of Prepositions for NLP Applications*, en Arsenijevic, Baldwin, Trawinski (eds.) (2006): *Prepositions '06 Proceedings of the Third ACL-SIGSEM Workshop on Prepositions*, Trento, págs. 29-36.
- Harris, Zellig (1954): *Distributional structure*, «Word», 10.23, págs. 146–162.
- Hastie, Trevor; Tibshirani, Robert; y Friedman, Jerome (2009): *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer, Nueva York.
- Haspelmath, Martin (1997): *From Space to Time. Temporal Adverbial in the World's Languages*, Lincom Europa, Munich.
- Hausser, Roland (2014): *Foundations of Computational Linguistics*, Springer, Berlín.

Haspelmath, Martin; Dryer, Matthew S; Gil, David; y Comrie, Bernard (2005): *WALSin. The World Atlas of Languages Structures*, OUP Oxford, Oxford.

Hernández-Nieto, Rafael (2002): *Instrumentos de recolección de datos en Ciencias Sociales y en Ciencias Biomédicas*, Universidad de los Andes, Mérida (Venezuela).

Herskovits, Annette (1981): On the Spatial Uses of Prepositions in English, «Linguisticae Investigationes», 5.2, págs. 303-327.

Herskovits, Annette (1986): *Language and Spatial Cognition: An interdisciplinary Study of Prepositions in English*, Cambridge University Press, Cambridge.

Hernández Cabrera, Tomás (1998): *Las preposiciones latinas en época clásica. Estudio funcional*, tesis doctoral, Universidad de La Laguna, La Laguna.

Hernández, Patricia C. (2013): *La locución prepositiva dentro de cómo marcador de bordes. Una descripción según el enfoque cognitivo-prototípico*, *Lingüística*, 29.1, págs. 81-114.

Hernández-Pastor, Diana y Perinián-Pascual, Carlos (2016): *Developing a knowledge base for preposition sense disambiguation: A view from Role and Reference Grammar and FunGranKB*, «Onomázein», 33, págs. 251-288.

Hernando Cuadrado, Luis Alberto (1998): *Sobre la expresión de la concesividad en español*, «Revista de Filología Románica», 15, págs. 123-133.

Hernando Cuadrado, Luis Alberto (2002): *La preposición en español*, «Didáctica, lengua y literatura», 14, págs. 145-159.

Hilferty, Joseph (1993): *Semántica lingüística y cognición*, «Verba», 20, págs.29-44.

Hill, Leslie Alexander (1968): *Prepositions and adverbial particles*, Oxford University Press, Oxford.

Hjelmslev, Louis (1935): *La catégorie du cas. Étude de grammaire générale*, Universitetsforlaget, Aarhus.

Hjemslev, Louis (1939): *La noción de rección en Ensayos lingüísticos*, Madrid, Gredos (1972), págs. 189-202.

Hockett, Charles (1958): *A Course in Modern Linguistics*, The MacMillan Company, Nueva York.

Hoekstra, Teun (1984): *Transitivity: Grammatical Relations in Government-Binding Theory*, Foris Publications, Dordrecht.

Hoekstra, Teun (1995): *The Function of Functional Category*, «GLOT International», 1.2, págs. 3-6.

Holik, Federico (2016): *Teoría de la información de Claude E. Shannon*, en *Diccionario Interdisciplinar Austral*, editado por Claudia E. Vanney, Ignacio Silva y Juan F. Franck. URL=http://dia.austral.edu.ar/Teoría_de_la_información_de_Claude_E._Shannon

Holmberg, Anders y Nikkane, Urpo (eds.): (1993): *Case and Other Functional Categories in Finnish Syntax*, Mouton de Gruyter, Berlín.

Holzinger, Andreas (2016) (Ed.): *Machine Learning for Health Informatics*, LNAI 9605, Springer International, págs. 1–24, DOI: 10.1007/978-3-319-50478-0 1.

Hornstein, Norbert y Weinberg, Amy (1981): *Case Theory and Preposition Stranding*, «Linguistic Inquiry», 12.1, 55-91.

Hopper, Paul J. y Closs Traugott, Elisabeth (1993): *Grammaticalization*, Cambridge University Press, Cambridge.

Hornos Chéliz, María del Carmen (2002): *Lo que la preposición esconde. Estudio sobre la argumentalidad preposicional en el predicado verbal*, Prensas Universitarias de Zaragoza, Zaragoza.

Hornos Chéliz, María del Carmen (2010): *La gramática del sentido. Léxico y sintaxis en la encrucijada*, Prensas Universitarias de Zaragoza, Zaragoza.

Hottenroth, Priska-Monika (1991): *Prepositions and object concept: A contribution to cognitive semantics*, en Zelinsky-Wibblet, Cornelia (ed.) (1993): *The semantics of Prepositions*, Mouton de Gruyter, Berlín, págs. 170-219.

Hout, Angeliek Van (1989): *PP's and Theta Theory*, en *Proceedings of Leiden Conference for Junior Linguists*, Leiden, págs. 49-62.

Hovy, Dirk; Tratz, Stephen; y Hovy, Eduard (2010): *What's in a Preposition? Dimensions of Sense Disambiguation for an Interesting Word Class*, Coling Poster Volume, Beijing, págs. 454-462.

Hovy, Dirk; Vaswani, Ashish; Tratz, Stephen y Chiang, David (2011): *Models and Training for Unsupervised Preposition Sense Disambiguation*, en *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, Portland, págs. 323–328.

Hüllen, Werner (1990): *Understanding the Historiography of Linguistics. Problems and projects*, Nodus Publikationen, Münster.

Hunston, Susan (2002): *Corpora in Applied Linguistics*, Cambridge Applied Linguistics, Cambridge.

Ibarretxe-Antuñano, Iraide y Valenzuela, Javier (2010): *Language as a Complex Dynamic System: A View from Cognitive Linguistics*, págs. 3-38, en Bel-Enguix y Jiménez-López,

Dolores (2010): *Language as a Complex System Interdisciplinary Approaches*, Cambridge Scholars Publishing, Newcastle upon Tyne.

Ibarretxe-Antuñano, Iraide (2013): *La lingüística cognitiva y su lugar en la historia de la lingüística*, «RESLA», 16, págs. 245-266.

Ibarretxe-Antuñano, Iraide y Hijazo-Gascón, Alberto (eds.) (2015): *New Horizons in the Study of Motion: Bringing together applied and theoretical perspectives*, Cambridge Scholars, Newcastle Upon Tyne.

Ibarretxe-Antuñano, Iraide; Valenzuela, Javier, y Hilferty, Joseph (2016): *La semántica cognitiva*, en *Lingüística Cognitiva*, «Anthropos», Barcelona, págs. 34-60.

Ifeganyi Mmadike, Benjamin (2016): *A case for prepositions in Igbo*, Linguistic Online, <http://www.phil.muni.cz/linguistica/art/mmadike/mma-001.pdf> ISSN 1801-5336

Inchaurrealde, Carlos y Vázquez, Ignacio (2000): *Una introducción cognitiva al lenguaje y la lingüística*, Mira Editores, Zaragoza.

Infante Arriala, Pascuala (2005): *Esquemas locativos-espaciales de las preposiciones de trayectoria along, across y through*, tesis doctoral, Universidad de Chile, Santiago de Chile.

Isaza Calderón, Baltasar (1967): *La doctrina gramatical de Bello*, Boletín de la RAE, Madrid.

Jackendoff, Ray (1973): *The base rules for PP* a Anderson, S. R. y Kiparsky, Paul (eds.): *A Festschrift for Morris Halle*, Holt, Nueva York, págs. 345-356.

Jackendoff, Ray (1990): *Semantic Structures*, The MIT Press, Massachusetts.

Jackendoff, Ray (1993): *X-Bar Semantics*, en Pustejovsky, James (ed.), *Semantics and Lexicon*, Kluwer, Dordrecht, págs. 15-26.

Jackendoff, Ray (1996): *The Architecture of the Linguistic-Spatial Interface*, en Peter Bloom, Peter; Peterson, M.; Nadel, L; y M. Garrett, M. (eds.), *Language and Space*, págs. 1-30, The MIT Press, Cambridge.

Jacobsson, Bengt (1977): *Adverbs, prepositions and conjunctions in English: a study in gradience*, «Studia Lingüística», 21.1, págs. 38-64.

Jakobson, Roman (1936): *Contribución general a la teoría general de los casos*, en *Ensayos de lingüística General*, Ariel, Barcelona, págs. 235-302 (traducción de Josep M. Pujol y Jem Cabanes).

Jespersen, Otto (1924): *The Philosophy of Grammar*, 1924, Allen and Unwin, Londres.

Jiménez Norberto, Laura (2011): *El estatus argumental del complemento de régimen prepositivo en español*, «Signos lingüísticos», Vol. VII, 14, págs. 81-114.

Jiménez, Tomás y Lübke, Bárbara (2013): *Los contextos preposicionales en español y alemán. Análisis contrastivo*, «Verba», 40, págs. 203-252.

Jiménez Juliá, Tomás (2002): *Notas sobre combinatòria de preposiciones en español*, en R. Lorenzo et al. (eds.) *Homenaxe a Fernando R. Tato Plaza*, Universidad de Santiago de Compostela, Santiago de Compostela, págs. 457-472.

Johnson, Robert y Kudy, Patricia (2004): *Estadística elemental*, Thomson, México.

Johansson Falck, Marlene (2014): *Temporal prepositions explained: Cross-linguistic analysis of English and Swedish unit of time landmarks*, «Cognitive Linguistic Studies», 1.2, págs. 271-288.

Jolly, Julia (1993): *Preposition Assignment in English*, en Van Valin, Robert (1993): *Advances in Role and Reference Grammar*, John Benjamins, Amsterdam, págs. 275-310.

Jorgensen, Fredrik y Lonning, Jan Tore (2008): *A Minimal Recursion Semantic Analysis of Locatives*, «Computational Linguistics», 35.2, págs. 229-270.

Jurafsky, Daniel; Martin, James H.; Norvig, Peter; y Russell, Stuart (2000): *Speech and language processing: an introduction to natural language processing computational linguistics and speech recognition*, Prentice Hall, Londres.

Kamp, Hans y Partee, Barbara (1995): *Prototype theory and compositionality*, «Cognition», 57.2, págs. 129-191.

Karypis, George (2003): *CLUTO. A clustering toolkit*, Universidad de Minnesota, Technical Report, 02-017. En <http://glaros.dtc.umn.edu/gkhome/cluto/cluto/overview>

Kayne, Richard (eds.) (1984): *Connectedness and Binary Branching*, Foris Publications, Dordrecht.

Kayne, Richard (1999): *Prepositional complementizers as attractors*, «Probus», 11, págs. 39-73.

Kayne, Richard (2000): *Note on prepositions, complementizers, and Word Order Universals*, en *Parameters and Universals*, Oxford University Press, Oxford, págs. 314-326.

Kayne, Richard (2007): *Prepositions as probes*, en Kayne, Richard (2007): *Movement and Silence*, Oxford University Press, Oxford, págs. 192-212.

Kazumi, Koike (1997): *Valores funcionales de las locuciones preposicionales en español*, «Onomazein», 2, págs. 151-179.

Kemmerer, David (2005): *The Spatial and Temporal Meanings of English Prepositions can be 38 independently impaired*, «Neuropsychologia», 43, págs. 797-806.

Kemmerer, David y Tranel, Daniel (2003): *A Double Dissociation Between the Meanings of Action Verbs and Locative Prepositions*, «Neurocase», 9.5, págs. 21-435.

- Kennedy, George (1998): *An introduction to corpus linguistics*, Longman, Nueva York.
- Kelleher, John D. y Costello, Fintan J. (2009): *Applying Computational Models of Spatial Prepositions to Visually Situated Dialog*, «Computational Linguistics», 35.2, págs. 272-306
- Kelleher, John; Sloan, Colm y Mac Namee, Brian (2009): *An investigation into the semantics of English topological prepositions*, «Cognitive Processing», 10.2, págs. 233-236.
- Kézié, Koyenzi (1996): *Les locatifs relationnels en Kabiýè: leurs propriétés sémantiques et morphosyntaxiques*, «Journal of West African Languages», 26.1, págs. 103-119.
- Klavans, Judith L. y Resnik, Philip (1996): *The balancing act: combining symbolic and statistical approaches to languages*, The MIT Press, Cambridge.
- Kleiber, Georges (1995): *La semántica de los prototipos. Categoría y sentido léxico*, Visor libros, Madrid.
- Klippe, Elisabeth (1997): *Prepositions and Variation*, en Di Sciullo, Anna-Maria (1997), (ed.) (1997): *Projections and Interface Conditions. Essays on Modularity*, Nueva York, Oxford University Press, págs. 73-108.
- Koerner, E. F. Konrad (1978): *Western Histories of linguistic thought*, John Benjamins, Amsterdam.
- Koehler, Reinhard y Rieger, Burghard B. (1993): *Contributions to Quantitative Linguistics*, Kluwer, Dordrecht.
- Koopman, Hilda (2000): *Verbal complexes*, The MIT Press, Massachusetts.
- Kordoni, Valia (2006): *PPs as verbal arguments: From a computational semantics perspective*, en *The ACL-SIGSEM Workshop on Prepositions*, The European Association for Computational Linguistics (EACL), Trento.
- Kortmann, Bernd y König, Ekkehard (1992): *Categorial Reanalysis: the case of deverbal prepositions*, «Linguistics», 30, págs. 671-697.
- Kovatchev, Venelin; Salamó, Maria; y Martí, M. Antònia (2016): *Comparing Distributional Semantics Models for identifying groups of semantically related words*, «Procesamiento del Lenguaje Natural», 57, págs. 109-116.
- Kurzon, Dennis y Adler, Silvia (2008): *Adpositions. Pragmatic, semantic and syntactic perspectives*, John Benjamins, Amsterdam.
- Laguna, José (2004): *Gramática de las preposiciones*, ASELE, Actas XV, Centro Virtual Cervantes, págs. 1-8.
- Laka, Brenda (1995): *Sobre el uso del acusativo preposicional en Español en Pensado*, Carmen (1995) (ed.): *El Complemento directo preposicional*, Visor, Madrid págs. 61-92.

Lakoff, George (1987): *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*, The Chicago University Press, Chicago.

Landis, J. Richard; y Koch, Gary G. (1977): *The Measurement of Observer Agreement for Categorical Data*, «Biometrics», 33.1, págs. 159-174.

Langacker, Ronald W. (1987): *Foundations of Cognitive Grammar. Theoretical Prerequisites*, Vol. I, Stanford University Press, Stanford.

Langacker, Ronald W. (1991): *Foundations of Cognitive Grammar. Descriptive Applications*, Vol. II, Stanford University Press, Stanford.

Langacker, Ronald W. (1999): *Grammar and Conceptualization*, Mouton de Gruyter, Berlín.

Langacker, Ronald W. (2008): *Cognitive Grammar: A basic introduction*, Oxford University Press, Oxford.

Langley, Pat (1996): *Elements of Machine Learning*, Morgan Kaufmann, San Francisco.

Lapesa, Rafael (1964): *Los casos latinos: restos sintácticos y sustitutos en español*, «BRAE», Tomo XLIV, Cuaderno CLXXI, págs. 57-105.

Larson, Richard (1988a): *On the double object construction*, «Linguistic Inquiry», 19.3, 1988, págs. 335-391.

Larson, Richard (1988b): *Missing Prepositions and the Analysis of English Free Relative Clauses*, 1985, «Linguistic Inquiry», págs. 239-266.

Larson, Richard (1990): *Double Objects Revisited: Reply to Jackendoff*, «Linguistic Inquiry», 21.4, págs. 589-632.

Lastra, Yolanda (1992): *Estudios antiguos y modernos sobre el otomí*, Colegio de México, en Barriga, Rebeca y García, Josefina (eds.) (1992): *Reflexiones lingüísticas y literarias*, Colegio de México, págs. 43-68.

Lassen, Tine (2006): *An Ontology-Based View on Prepositional Senses*, en Arsenijevic, Boban; Baldwin, Timothy; y Trawinski, Beata (eds.) (2006): *Prepositions '06 Proceedings of the Third ACL-SIGSEM Workshop on Prepositions*, Trento, págs. 45-50.

Laur, Dany (1993): *La relation entre le verbe et la préposition dans la sémantique du déplacement*, «Langages», 110, págs. 47-67.

Lázaro Mota, Fernando A. (1985): *Algunas notas sobre la preposición*, en *Philologica Hispaniensa in honorem M. Alvar*, Vol. II, Gredos, Madrid, págs. 378-379.

Leal, António (ed.) (2018): *Verbs, Movement and Prepositions*, Centro de Linguística da Universidad de Porto, Porto.

Lersundi, Mikel y Agirre, Eneko (2003): *Semantic interpretations of postpositions and prepositions: a multilingual inventory for Spanish, Basque and English*, ACL-SIGSEM workshop: The Linguistic Dimensions of Prepositions and their Use in Computational Linguistics Formalisms and Applications, Institut de Recherche en informatique de Toulouse, Toulouse, págs. 56-65.

Lee, David (2001): *Cognitive Linguistics: An Introduction*, Oxford University Press, Oxford.

Leech, Geoffrey (1991): *The state of the art in corpus linguistics*, en Aijmer, K. y Altenberg, B. (Eds.): (1991): *English Corpus Linguistics. Studies in honor of Jan Svartvik*, Longman, Londres, págs. 8-29.

Leeman, Danielle (2008): *Prépositions du Français: état des lieux*, «Langue française», 157, págs. 5-19.

Lehrer, Adrienne (1974): *Semantic fields and lexical structure*, Elsevier, Amsterdam.

Lehmann, Christian (2002): *Thoughts on Grammaticalization*, Universidad de Erfurt, Erfurt.

Lenarduzzi, René (1996): *Las preposiciones POR y PARA en los complementos verbales de espacio y tiempo*, Bulzoni, Roma.

Lenci, Alessandro (2008): *Distributional semantics in linguistic and cognitive research*, «Rivista di Linguistica», 20.1, págs. 1–31.

Lenz, Rodolfo (1920): *La oración y sus partes. Estudios de gramática general y castellana*, Revista de Filología Española, Madrid.

Levinson, Stephen C. (2003): *Space in language and cognition*, Cambridge University Press, Cambridge.

Levinson, Stephen C. y Meira, Sérgio (2003). *Natural concepts in the spatial topological domain –adpositional meanings*, «Language», 79.3, págs. 485-516.

Levinson, Stephen C. (ed.) (2006): *Grammars of Space: Explorations in Cognitive Diversity*, Cambridge University Press, Cambridge.

Levy, Omar y Goldberg, Yoav (2014): *Linguistic regularities in Sparse and Explicit Word Representations*, en *Proceedings of the Eighteenth Conference on Computational Language Learning*, Baltimore, Maryland, Association for Computational Linguistics, págs. 171–180.

Liang, James C. P. (1971): *Prepositions, co-verbs or verbs? : a commentary on Chinese grammar past and present*, tesis doctoral, Universidad de Pensilvania, Pensilvania.

Linstromberg, Seth (1997): *English Prepositions Explained*, John Benjamins, Amsterdam.

Littlefield, Heahter (2005): *Lexical and Functional Prepositions in Acquisition: Evidence for a Hybrid Category*, Boston University Conference on Language Development 29, Online Proceedings Supplement.

Littlefield, Heahter (2006): *Syntax and Acquisition in the Prepositional Domain: Evidence from English for fine-grained syntactic categories*, tesis doctoral, Boston University, Boston.

Litkowski, Ken y Hargraves, Orin (2005): *The Preposition Project*, en *ACL-SIGSEM Workshop on the Linguistic Dimensions of Prepositions and their Use in Computational Linguistics Formalisms and Applications*, Colchester, págs. 171–179.

Litkowski, Ken y Hargraves, Orin (2007): *SemEval-2007 Task 06: Word-Sense Disambiguation of Prepositions*, en *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, Praga, págs 24–29.

Litkowski, Ken y Hargraves, Orin (2006): *Coverage and Inheritance in The Preposition Project*, en Arsenijevic, Boban; Baldwin, Thimoty; y Trawinski, Beata (eds.) (2006): *Prepositions '06 Proceedings of the Third ACL-SIGSEM Workshop on Prepositions*, Trento, págs. 37-44.

Litkowski, Ken (2017): *The Preposition Corpus in the Sketch Engine*, Technical Repor, Damasco.

Levy, Omer y Goldberg, Yoav (2014): *Linguistic Regularities in Sparse and Explicit Word Representations*, en *Proceedings of the Eithteenth Conference on Computational Language Learning*, Association for Computational Linguistics, Baltimore, págs. 171-180.

Levy, Omer; Goldberg, Yoav y Dagan, Ido (2015): *Improving Distributional Similarity with Lessons Learned from Word Embedding*, «Computational Linguistics», 3, págs. 211-225.

Lliteras, Margarita (1992): *Propuesta de explicación histórica de la oposición para/hacia*, En *Actas del II Congreso Internacional de Historia de la lengua española*, Universidad de Sevilla, Sevilla, págs. 605-612.

Llopis, Jesús (1996): *Gramática y cognición: el dativo de dirección*, «ELUA», 11, págs. 199-231.

Llopis-García, Reyes (2011): *La gramática cognitiva: nuevas avenidas para la enseñanza de lenguas extranjeras*, «Verba Hispanica», 19, Universidad de Ljubljana, págs. 111-127.

Llopis-García, Reyes (2015): *Las preposiciones y la metáfora del espacio: aportaciones y potencial de la lingüística cognitiva para su enseñanza*, «Journal of Spanish Language Teaching», 2, págs. 51-68.

Löbel, Elisabeth (1999): *Classifiers and the Notion of “Semi-lexical Head”*, en *Workshop on Semi-lexical Heads*, Universidad de Tilburg, Tilburg.

López, María Luisa (1970): *Problemas y métodos en el análisis de las preposiciones*, Madrid, Gredos.

López Martínez, M. Sol (1993): *O complemento directo con preposición a en galego*, «Verba», anejo 36, Santiago de Compostela.

López Rua, Paula (2003): *Birds, Colours and Prepositions. The theory of Categorizations and its Applications in Linguistics*, LINCOM, Múnich.

Luján, Marta (1999): *Minimalist Bello: Basic Categories in Bello's Grammar* en Gutiérrez-Rexach, Javier y Martínez-Gil, Fernando (eds.) (1999): *Advances in Hispanic Linguistics*, Cascadilla Press, Somerville MA, págs. 428-446.

Luo, Ying (2016): *Verbos de desplazamiento en español y en chino: un análisis subléxico de su significado y sus extensiones semánticas*, tesis doctoral, UAM, Madrid.

Luque, Juan D. (1973): *Las preposiciones. Valores generales. Vol. I. Valores idiomáticos. Vol. II*, SGEL, Madrid.

Luque, Bartolo y Parrondo, Juan M. R. (2016): *Las leyes del azar*, Bonallettera Alcompas, Barcelona.

Lyons, John (1977): *Semantics*, Cambridge University Press, Cambridge.

Maestre Maestre, José María; Barea, Joaquín Pascual, Brea, Luis Charlo(eds.) (2008): *Humanismo y pervivencia del mundo clásico: homenaje al profesor Antonio Prieto*, CSIC Press, Madrid.

Manning, Christopher D. y Schütze, Hinrich (1999): *Foundations on statistical natural languages processing*, The MIT Press, Londres.

Mardale, Alexandre (2011): *Prepositions as a semilexical category*, «Bucharest Working Papers in Linguistics», 13.2, Universidad de Bucharest, Bucharest, págs.35-50.

Markovskaya, Evguenia (2006): *Goal-Source Asymmetry and Russian Spatial Prefixes* en «Tromsø Working Papers in Linguistics», 33.2. Número especial sobre *Adpositions*, editado por Svenonius, Peter, págs. 200–219. CASTL, Tromsø. <http://www.ub.uit.no/munin/nordlyd/>

Marque-Pucheu, Christiane (2008): *La couleur des prépositions à et de*, «Langue française», 157, págs. 74-105.

Martí, Salvador y Miñana, Vicent (1988): *Les preposicions, més predicats que mai*, «Caplletra», 3, págs. 101-108.

Martínez, José Antonio (1992): *Tres hipótesis sobre el origen histórico de la 'partícula' hasta*. *Actas del II Congreso Internacional de Historia de la Lengua Española*. Tomo I, Madrid: Pabellón de España, 613-30. Consultado en:

<http://www.cervantesvirtual.com/obra/tres-hipotesis-sobre-el-origen-historico-de-la-particula-hasta-0/>.

Martínez, Hortensia (1983): *El suplemento en español partiendo de los presupuestos de la lingüística funcional*, tesis doctoral, Universidad de Oviedo, Oviedo.

Martínez, Hortensia (1986): *El suplemento en español*, Gredos, Madrid.

Martínez, Hortensia (1995): *Combinatoria y semántica de los verbos suplementarios*, «Revista Española de Lingüística», 25.2, págs. 397-409.

Martínez, Hortensia (2002): *Viejos y nuevos valores de las preposiciones españolas*, «Verba», 39, págs. 7-34.

Martínez-Ferreiro, Silvia; Bastiaanse, Roelien; y Boye, Kasper (2019): *Functional and usage-based approaches to aphasia: the grammatical-lexical distinction and the role of frequency*, «Aphasiology», DOI: 10.1080/02687038.2019.1615335.

Martínez Linares, María Antonia (1999): *Papeles semánticos, preposiciones y objetos preposicionales*, «ELUA», 13, págs. 155-192.

Martínez Sánchez, José Manuel (2008): *Neurolingüística: Patologías y Trastornos del Lenguaje*, «Revista Digital Universitaria», 9.12, págs. 1-18

Martínez, Juan Antonio y Jorgensen, Annette Myre (2013): *La adverbialización con la preposición de: patrones y frecuencias*, «Dilogía», 7, págs. 4-22.

Mateu, Jaume (1999): *On the irrelevance of the conceptual content to the syntax-semantics in interface*, en *GLOW Workshop on Sources of Universals*, Universidad de Potsdam, Potsdam.

Mateu, Jaume (2002): *Argument structure: relational construal at the syntax-semantics interface*, tesis doctoral, Universitat Autònoma de Barcelona, Barcelona.

Medelyan, Olena; Milne, David; Legg, Catherine; y Witten, Ian H. (2009). *Mining meaning from Wikipedia*, «International Journal of Human-Computer Interactions», 67.9, 716-754.

Mendizábal de la Cruz, Nieves (2012): *Valores semánticos de la preposición en: el caso del español de México en un corpus de entrevistas orales*, «Anmal Electrónica», 33, págs. 123-146.

Mendikoetxea, Amaya (2007): *En busca de los primitivos léxicos y su realización sintáctica: del léxico a la sintaxis y viceversa*, en Teresa Cabré (ed.) (2007): *Lingüística Teòrica: Anàlisi i Perspectives*, Servei de Publicacions de la UAB, Barcelona.

Mendoza Abreu, Josefa María (1992): *El uso de sintagmas prepositivos para la expresión de la causa en castellano medieval*, Actas del II Congreso, Vol. I, Sevilla, págs. 675-688.

- Menéndez Pidal, Marcelino (1904): *Manual de gramática histórica*, Editor Vicente Suárez, Madrid.
- Meya Llopart, Montserrat y Huber, Wolfgang (1986): *Lingüística computacional*, Teide, Barcelona.
- Minguell, Antonia Esther (2010): *La preposición: categoría léxica y funcional. Aportes para la enseñanza del español*, «Signos ELE», págs. 1-20.
- Mikelenic, Bojana (2017): *El complemento de régimen preposicional en español y sus traducciones al croata: observaciones preliminares*, «Verba hispánica», 25, págs. 37-53.
- Mikolov, Tomas; Yih; Wen-tau; y Zweig, Geoffrey (2013): *Linguistic regularities in continuous space word representations*, en *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics Human Language Technologies*, págs. 746–751.
- Millán Chivite, Alberto (1992): *El suplemento propio e indirecto en el Mío Cid. El suplemento propio e indirecto en el Mío Cid*, en *Actas del II Congreso Internacional de Historia de la lengua española*, Universidad de Sevilla, Sevilla, págs. 689-700
- Miyoshi, Jun-Nosuke (2015): *La partícula española hasta y su función pragmática*, «Lingüística», 31.1, págs. 7-23.
- Moisl, Hermann (2015): *Cluster Analysis for Corpus Linguistics*, De Gruyter Mouton, Munich.
- Moliner, María (2012): *Uso de las preposiciones*, Gredos, Madrid.
- Montemurro, Marcelo A. (2001): *Beyond the Zipf-Mandelbrot Law in Quantitative Linguistics. Physica A Statistical Mechanics and its Applications*, «Physica», Vol. 300, págs. 567-578.
- Montemurro, Marcelo A. (2004): *A Generalization of the Zipf-Mandelbrot Law in Linguistics*, en Gell-Mann, Murray y Tsallis, Constantino (eds.) (2004): *Nonextensive entropy. Interdisciplinary applications*, Oxford University Press, Oxford, págs. 347-356.
- Montemurro, Marcelo A. y Zanette, Damián H. (2011): *Universal Entropy of Word Ordering Across Linguistic Families*, «PlosOne», 6.5, págs. 1-9.
- Montes de Oca, Domingo Román (1997): *Uso de las preposiciones de dirección en pacientes afásicos agramáticos*, «Boletín de Filología», Tomo XXXVI, págs. 211-264.
- Montgomery, Douglas C. (2002): *Diseño y análisis de experimentos*, Limusa Wilwy, México.
- Moreno, Concha y Tuts, Martina (1998): *Las preposiciones: valor y función*, SGEL, Madrid.

- Moreno Fernández, Francisco (1998): *Principios de sociolingüística y sociología del lenguaje*, Ariel, Barcelona.
- Moreno Sandoval, Antonio (1998): *Lingüística computacional. Introducción a los modelos sintácticos, estadísticos y biológicos*, Síntesis, Madrid.
- Morera, Marcial (1984): *La semántica preposicional*, Anuario de la Universidad de La Laguna, La Laguna, págs. 113-144.
- Morera, Marcial (1988): *Estructura semántica del sistema preposicional del español moderno y sus campos de uso*, Cabildo Insular de Fuerteventura, Puerto del Rosario.
- Morera, Marcial (1990): *La preposición española contra: su evolución semántica*, «Verba», 17, Santiago de Compostela, págs. 287-313.
- Morera, Marcial (1994): *La filiación categorial de la preposición*, «Verba», 21, Santiago de Compostela, págs. 241-256.
- Morera, Marcial (1999a): *El arabismo español hasta: su evolución formal y semántica*, «Verba», Santiago de Compostela, págs. 81-95.
- Morera, Marcial (1999b): *La preposicionalización del nombre*, «Anuario de Letras: Lingüística y Filología», 37, págs. 33-52.
- Morera, Marcial (2016): *Preposiciones*, en Gutiérrez-Rexach, Javier (ed.) *Enciclopedia de lingüística hispánica*, Vol. I, Routledge, Londres, págs. 835-846.
- Morimoto, Yuko (2001): *Los verbos de movimiento*, Visor, Madrid.
- Moure, Teresa (1994): *La teoría de prototipos y su aplicación en gramática (I y II)*, «Contextos», 23-24, págs. 167-219.
- Moure, Teresa (1996): *La alternativa no-discreta en lingüística. Una perspectiva histórica y metodológica*, Universidad de Santiago de Compostela, Servicio de Publicaciones, Santiago de Compostela.
- Mourelle Lema, Manuel (1968): *Teoría lingüística en la España del siglo XIX*, Prensa Española, Madrid.
- Moya, Myriam (2007): *Postulados y retos de la lingüística cognitiva*, «Cuadernos de Lingüística Hispánica», 9, págs. 97-108.
- Muñoz, María Dolores (2017): *Pasado y presente de los diccionarios sintácticos con especial referencia al Diccionario de construcción y régimen de la lengua castellana*, en «Estudios de lingüística del Español», 38, págs. 155-172.
- Muysken, Pieter (1981): *Parametrizing the notion Head*, «Journal of Linguistics Research», 2.3, págs. 57-76.

Muysken, Pieter (1988): *Are Creoles a special type of language?* En *Linguistics: The Cambridge survey*, Vol. II. Editado por Newmeyer, Frederick Cambridge University Press, Cambridge, págs. 285–301.

Muysken, Pieter (2008): *Functional Categories*, Cambridge University Press, Cambridge.

Muysken, Pieter y Van Riemsdijk, Henk (eds.) (1985): *Features and Projections*, Foris Publications, Dordrecht.

Nañez, Emilio (1991): *Uso de las preposiciones*, SGEL, Madrid.

Napoli, Donna Jo y Nevis, Joel (1987): *Inflected prepositions in Italian*, «Phonology Yearbook», 4, págs. 195-209.

Napoli, Donna Jo (1989): *Predication Theory*, Cambridge University Press, Cambridge.

Nazar, Rogelio y Renau, Irene (2012): *Agrupación semántica de sustantivos basada en similitud distribucional. Implicaciones lexicográficas*, V Congreso Internacional de Lexicografía Hispánica, publicado en Anexo de la Revista de Lexicografía, Vol. II, págs. 272-295.

Nebrija, Antonio de (1492): *Gramática de la lengua castellana*, edición de Quilis, Antonio (1984), Editora Nacional, Madrid.

Neeleman, Ad (1997a): *PP's complements*, «Natural Language and Linguistic Theory», 15, págs. 89-137.

Neeleman, Ad (1997b): *Flexible Syntax. A Theory of Case and Arguments*, Utrecht Institute of Linguistics, Utrecht.

Nieuwenhuijsen, Dorien: (2013) *En busca de la persona agente: la explotación de un corpus diacrónico digital para el estudio de las preposiciones de y por en construcciones pasivas*, «Scriptum Digital», 2, págs. 43-58.

Nikkane, Urpo (1993): *On assigning Semantic Case in Finnish*, en Holmberg, Anders y Nikkane, Urpo (eds.): *Case and Other Functional Categories in Finnish Syntax*, Mouton de Gruyter, Berlín, págs. 75-88.

Nikulin, Mikhail S. (1973): *Chi-square test for continuous distributions with scale and shift parameters*, «Theory of Probability and its Applications», 18.3, págs. 559–568

O'Hara, Tom y Wiebe, Janyce (2008): *Exploiting Semantic Role Resources for Preposition Disambiguation*, «Computational Linguistics», 35.1, págs. 151-184.

O'Keffe, Anne y McCarthy, Michael (2019): *The Routledge Handbook of Corpus Linguistic*, Routledge, Oxon.

Octavio de Toledo, Álvaro S. (2016): *Los relacionantes locativos en la historia del español*, Mouton De Gruyter, Berlín.

- Ojeda, Ascensión (2013): *Estudio de las frases prepositivas en el habla culta de Sevilla*, tesis doctoral, Universidad de Sevilla, Sevilla.
- Osherson, Daniel N. y Smith, Edward E. (1981): *On the adequacy of prototype theory as a theory of concepts*, «Cognition», 9.1, págs. 35-68.
- Osuna, Francisco (1991): *Función semántica y función sintáctica de las preposiciones*, Ágora, Málaga.
- Ouhalla, Jamal (1991): *Functional Categories and Parametric Variation*, Routledge, Londres.
- Ouyang, Jinghui y Jiang, Jingyang (2019): *What Makes a Good Quantitative Linguistics Textbook?*, «Journal of Quantitative Linguistics», DOI: <https://doi.org/10.1080/09296174.2019.1631118>
- Padró, Lluís y Stanilovsky, Evgeny (2012): *Freeling 3.0: Toward Wider Multilinguality, Proceeding of the Eight International Conference Language Resources and Evaluation (LREC'12)*, ELRA, Estambul, págs. 2473-2479.
- Padró, Lluís (2011): *Analizadores Multilingües en FreeLing*, Departamento de Lenguajes y Sistemas Informáticos, Centro de Investigación TALP, Universitat Politècnica de Catalunya, Barcelona, págs. 1-8.
- Pardo, Leandro (1993): *Teoría de la información estadística*, «Estadística Española», 35.133, págs. 195-268.
- Paredes García, Florentino; Álvaro, Salvador; Paredes Zurdo, Luna; y Núñez, Zaida: *El libro del español correcto. Claves para hablar y escribir bien en español* (2012): Espasa-Instituto Cervantes, Barcelona.
- Paredes García, Florentino; Álvaro, Salvador; y Paredes Zurdo, Luna (2013): *Las 500 dudas más frecuentes del español*, Espasa Libros, S.L.U. Instituto Cervantes, Barcelona.
- Parodi, Giovanni (2008): *Lingüística de corpus: una introducción al ámbito*, «Revista de Lingüística Teórica y Aplicada», 46.1, Chile, págs. 93-119.
- Parret, Herman (ed.) (1976): *History of Linguistic Thought and Contemporary Linguistics*, Walter de Gruyter, Berlín.
- Perea, Francisco Javier (2007): *Análisis de Errores y enseñanza de las preposiciones de la lengua española como LE*, «MarcoELE», 4, págs. 1-13.
- Peris, Aina; Taulé, Mariona; y Rodríguez, Horacio (2009): *Hacia un sistema de clasificación automático de sustantivos deverbales*, «Procesamiento del Lenguaje Natural», 43, págs. 23-31.
- Petrík, Daniel (2011): *Aplicación de las leyes de Zipf en los textos españoles*, «Romanica Olomucensia», 27.2, págs. 137-146.

Obediente, Enrique y D'Introno, Francesco (2001): *Andrés Bello: Sus antecedentes en la filosofía analítica y su proyección en la lingüística moderna*, en *History of Linguistics in Spain/Historia de la lingüística en España*, Volumen II. Ed. Koerner, E. F. Konrad y Niederehe, Hans-Josef, John Benjamins, 2001, Amsterdam, págs. 406-22.

Pavón, María Victoria (1995): *Clases de partículas y estructura de constituyentes*, tesis doctoral, Universidad Complutense de Madrid, Madrid.

Pavón, María Victoria (2003): *Sintaxis de las partículas*, Visor, Madrid.

Pensado, Carmen (1996): *El complemento directo preposicional. Estado de la cuestión y bibliografía comentada*, Visor, Madrid.

Peñalver, Manuel (1992): *Las ideas gramaticales de Vicente Salvá*, en las *Actas del II Congreso Internacional de Historia de la Lengua Española*, Tomo I, Sevilla, 1992, págs. 739-759.

Plann, Susan (1988): *Prepositions, Postpositions and Substantives*, «Hispania», 71, págs. 920-926.

Pollard, Carl y Sag, Ivan A. (1994) *Head-Driven Phrase Structure Grammar*, University of Chicago Press and Stanford, CSLI Publications, Chicago.

Portilla, Mario (2011): *El origen de las preposiciones en español*, «Filología y Lingüística», 37.1, págs. 229-244.

Porto Dapena, José Alvaro (1987): *Contribución a una teoría de las preposiciones: factores que determinan la elección éstas en el discurso*, «Thesaurus», 41, págs. 623-646.

Pottier, Bernard (1962): *Systematique des éléments de relation*, Klincksieck, París.

Pottier, Bernard (1968): *L'emploi de la préposition a devant l'objet en espagnol...* «Bulletin Société Linguistique de Paris», 63.1, págs. 68-72.

Pratt-Hartmann, Ian E. y Francez, Nissim (2001): *Temporal Prepositions and Temporal Generalized Quantifiers*, «Linguistic and Philosophy», 24, págs. 187-222.

Pratt-Hartmann, Ian E. (2005): *Temporal Prepositions and their Logic*, «Artificial Intelligence», 166.1-2, págs. 1-36.

Pustejovsky, James y Stubbs, Amber (2013): *Natural Language Annotation for machine learning*, O'Reilly, Sebastopol.

Ramat, Paolo (1999): *Linguistic categories and linguists categorizations*, «Linguistics», 37.1, págs. 157-180.

Rakotovoavy, Bakolinirina (2018): *Los errores en el uso de las preposiciones por parte de los profesores malgaches de ELE*, tesis doctoral, Universidad Complutense de Madrid, Madrid.

RAE (1771): *Gramática de la Real Academia Española*, edición facsímil y apéndice documental de Ramón Sarmiento (1984), Editora Nacional, Madrid.

Real Academia Española y Asociación de Academias de la Lengua Española (2013): *Diccionario panhispánico de dudas*, Santillana Ediciones Generales, Barcelona.

Real Academia Española y Asociación de Academias de la Lengua Española (2009): *Nueva gramática de la lengua española*, Espasa Libros, Madrid.

Real Academia Española y Asociación de Academias de la Lengua Española (2013): *El buen uso del español*, Espasa Libros, S.L.U, Barcelona.

Rauh, Gisa (1991): *Approaches to prepositions*, Günter Narr, Tübingen.

Rauh, Gisa (1993): *On the grammar of lexical and non-lexical prepositions in English* en Zelinsky-Wibblet, Cornelia (ed.) (1993): *The semantics of Prepositions*, Mouton de Gruyter, Berlín, págs. 99-150.

Rauh, Gisa (1994): *Prépositions et Roles: point de vue syntaxique et sémantique*, «Languages», 114-115, págs. 67-75.

Rauh, Gisa (2002): *Word-classes as prototypical categories*, en Scholz; S.; Klages; M.; Hantson, E.; y Römer, U. (eds.) (2002): *Language: Context and Cognition. Papers in Honour of Wolf-Dietrich Bald's 60th Birthday*, Langenscheidt-Longman, Munich, págs. 259-270.

Ravin, Yael y Leacock, Claudia (2000): *Polysemy. Theoretical and Computational Approaches*, Oxford University Press, Oxford.

Real, Cristina (2013): *Lexicalization by phase: the role of prepositions in argument structure and its cross-linguistic variation*, tesis doctoral, Universitat Autònoma de Barcelona, Barcelona.

Reese, Samuel; Boleda, Gemma; Cuadros, Montse; Padró, Lluís; y Rigau, German (2010): *Wikicorpus: A Word-Sense Disambiguated Multilingual Wikipedia Corpus*, en *Proceedings of 7th Language Resources and Evaluation Conference (LREC'10)*, La Valleta.

Resnik, Gabriela (2011): *Los nombres eventivos no deverbales en español*, tesis doctoral, Universidad Pompeu Fabra, Barcelona.

Rezai, Vali (2012): *A Functional Analysis of Prepositional Phrases in Persian*, «International Journal of Linguistics», 4.4, págs. 309-325.

Reuland, Eric; Bhattachanya, Tammoy; y Spathas, Giorgos (2007): *Argument Structure*, John Benjamins, Amsterdam.

Rico, Francisco (1978): *Nebrija frente a los bárbaros: El canon de gramáticos nefastos en las polémicas del humanismo*, Universidad de Salamanca, Salamanca.

Riemsdijk, Henk Van (1978): *A Case Study in Syntactic Markedness. The Binding Nature of Prepositional Phrases*, 1982, Foris Publications, Dordrecht.

Riemsdijk, Henk Van y Williams, Edwin, (1985): *Introduction to the Theory of Grammar*, The MIT Press, Cambridge.

Riemsdijk, Henk Van (1990): *Functional Prepositions*, en Pinkster, H. y Genee, I. (eds.) *Unity in Diversity. Papers to Simon C. Dik on his 50th birthday*, Foris Publications, Dordrecht, págs. 229-242.

Riemsdijk, Henk Van (1996): *The extension of projections*, «Tsuru University Working Papers in Linguistics», págs. 42-80.

Riemsdijk, Henk Van (1998): *Categorial feature magnetism: The endocentricity and distribution of projections*, «Journal of Comparative Germanic Linguistics», 2, págs. 1-48.

Riemsdijk, Henk Van (2007): *Case in Spatial Apositional Phrases: The Dative-Accusative Alternation in German*, en Alboiu, Gabriela; Avram, Andrei; Avram, Larisa; y Isac, Dana (eds.) (2007): *Festschrift for Alexandra Cornilescu*, Bucharest University Press, Bucarest, págs. 1-19.

Rice, Sally (1999): *Aspects of prepositions and prepositional aspect*, en Stadler, Leon de y Eyrich, Christoph (eds.) (1999): *Iusses in Cognitive Linguistics*, Mouton, Berlín, págs. 225-247.

Rigau, Gemma (1990): *The semantic nature of some Romance prepositions* en Mascaró, Joan y Nespó, M. (eds.) (1990): *Grammar in progress*, Foris Publications, Dordrecht, págs. 363-374.

Rizzi, Luigi (2004): *The Structure of CP and IP: the Cartography of Syntactic Structures*, Oxford University Press, Vol. II, Oxford.

Robins, Robert H. (1966): *The Development of the Word Class System*, «Foundations of Language», II, págs. 3-19.

Robins, Robert H. (1967): *A short history of linguistics*, Longmans, Green and Co, Londres.

Rodrigues, José María (2007): *Las preposiciones: ¿elementos de enlace o especificadores semánticos?*, en *Actas del XVI Congreso AIH*, págs. 1-13.

Rodrigues, José María (2008): *Las preposiciones en la gramática española: estudio diacrónico sobre los "especificadores semánticos"*, «Anais do Congresso Brasileiro de Hispanistas», V. C749a, págs. 2785-2797.

Rodrigues, José María (2009): *La lingüística del corpus y el análisis gramatical del español: propuesta de tratamiento de las preposiciones del español como especificadores semánticos*, tesis doctoral, Universidad de Valladolid, Valladolid.

Rodríguez Adrados, Francisco (1986): *Reflexiones sobre los sistemas de preposiciones del griego antiguo a partir del DGE*, «Revista Española de Lingüística», 16.1, págs. 71-82.

Rodríguez Espiñeira, María José y Pena Seijas, Jesús (2008): *Categorización lingüística y límites intercategoriales*, «Verba», anejo 61, Universidad de Santiago de Compostela, Santiago de Compostela.

Rojo, Guillermo (2011): *Sobre la frecuencia de verbos y esquemas sintácticos*, en Bustos Tovar, José Jesús (coord.) (2011): *Homenaje a Antonio Narbona*, Vol 2, Universidad de Sevilla, Sevilla, págs. 902-907.

Romeu, Juan (2012a): *Verbal Prefixes are not Ps*, P-Workshop, Stuttgart.

Romeu, Juan (2012b): *Cartography and polysemy of spatial Ps. The case of in in English*, Ruhr-Universität Bochum, Bochum.

Romeu Fernández, Juan (2014): *Cartografía mínima de las construcciones espaciales*, tesis doctoral, UCM, Madrid.

Romeu Fernández, Juan (2017): *Grammatical categories at the Lexicon-Syntax-Semantics interface: The case of prepositions*, en Fernández-Soriano, Olga; Castroviejo, Elena; y Pérez-Jiménez, Isabel (ed.) (2017): *Boundaries, Phases and Interfaces: Case studies in honor of Violeta Demonte*, John Benjamins, Amsterdam, págs. 326–342.

Romo, Francisco (2016): *Un estudio cognitivista de las preposiciones espaciales del español y su aplicación a la enseñanza E/LE*, tesis doctoral, UAB, Barcelona.

Ross, Claudia (1991): *Coverbs and category distinctions in Mandarin Chinese*, «Journal of Chinese Linguistics», 19.1, págs. 79-115.

Ross, Sheldon M. (2007): *Introducción a la estadística*, Reverté, Barcelona.

Rudzka-Ostyn, Brygida (1987): *Topics in cognitive linguistics*, John Benjamins, Amsterdam.

Ruiz de Mendoza, Francisco José y Pérez, Lorena (2000): *Primitivos semánticos y modelos cognitivos en la organización del conocimiento*, «Scire», 6.2, págs. 79-97.

Ruiz de Mendoza, Francisco José (2001): *Lingüística cognitiva: semántica, pragmática y construcciones*, «CLAC», 8, Universidad de la Rioja, La Rioja.

Saeed, John I. (2016): *Semantics*, Blackwell, Londres.

Sahlgren, Magnus (2006): *The Word-Space Model: using distributional analysis to represent syntagmatic and paradigmatic relations between words in high-dimensional vector spaces*, tesis doctoral, Estocolmo.

Sahlgren, Magnus (2008): *The distributional hypothesis*, «Rivista di Linguistica», 20.1, págs. 33-53.

- Saint-Dizier, Patrick (2006a): *Syntax and Semantics of prepositions*, Springer, Dordrecht.
- Saint-Dizier, Patrick (ed.) (2006b): *Computational Linguistics Dimensions of Syntax and Semantics of Prepositions*, Kluwer Academic, Dordrecht.
- Saint-Dizier, Patrick (2006c): *PrepNet: A multilingual lexical description of prepositions*, en *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006)*, págs. 877–885, Génova.
- Saint-Dizier, Patrick y Vázquez, Glòria (2001): *A Compositional framework to Prepositions*, IWCS4, Tilburg.
- Sampson, Geoffrey y McCarthy, Diana (eds.) (2004): *Corpus Linguistics: Reading in a Widening Discipline*, Continuum, Londres.
- Samuelson, Christer (1996): *Relating Turing's Formula and Zipf's Law*, en *Proceedings of the 4th Workshop on Very Large Corpus*, Copenhagen.
- San Vicente, Felipe (Coord. y dir.) (2015): *Gramática de referencia de español para itálofonos*, Ediciones de la Universidad de Salamanca, Salamanca.
- Sánchez de las Brozas, Francisco (1995): *Minerva o de causis linguae latinae*, Institución Cultural “El Brocense”, Cáceres.
- Sancho Cremades, Pelegrí (1995): *La categoría preposicional*, Universidad de Valencia, Valencia.
- Sanromán Vilas, Begoña y Carrasco, Ángeles (2019): *Introducción: categorías semiléxicas. Haciendo balance entre el contenido léxico y la función gramatical*, «RSEL», 49, pág. 7-22.
- Sapir, Edward (1921): *Language: An Introduction to the Study of Speech*, Hartcourt, Nueva York.
- Sarmiento, Ramón (1978): *La gramática de la Academia: historia de una metodología*, «BRAE», Madrid, 58, págs. 435-46.
- Schwarze, Christoph (ed.) (1981): *Analyse des prépositions*, Niemeyer, Tübingen.
- Seco, Rafael (1930): *Manual de gramática española*, Aguilar, Madrid.
- Serradilla Castaño, Ana María (1995): *Sobre las primeras apariciones de construcciones preposicionales ante que completivo en español medieval*, «EPOS. Revista de Filología», págs. 1-17.
- Serradilla Castaño, Ana María (1997a): *El complemento de régimen verbal: criterios para su identificación*, «CAUCE. Revista de Filología y su Didáctica», 20-21, págs. 1017-1051.

Serradilla Castaño, Ana María (1997b): *El régimen de los verbos de entendimiento y lengua en español medieval*, Universidad Autónoma de Madrid, Madrid.

Serradilla Castaño Ana María (2001): *Frecuencia de preposiciones en el régimen de construcciones transitivas: hacia una periodización lingüística de la Edad Media*, en «Cahiers de linguistique et de civilisation hispaniques médiévales», 24, págs. 143-165.

Sethna, James P. (2006): *Statistical Mechanics: Entropy, Order Parameters and Complexity*, Oxford University Press, Oxford.

Shannon, Claude E. (1948): *A Mathematical Theory of Communication*, «The Bell System Technical Journal», 27, págs. 379–423 y 623–656, julio y octubre.

Sirvent, Ángeles (ed.) (2005): *Espacio y texto en la cultura francesa*, Universidad de Alicante, Alicante.

Slager, Emile (1997): *Pequeño diccionario de construcciones preposicionales*, Visor, Madrid.

Slager, Emile (2004): *Diccionario de uso de las preposiciones españolas*, Espasa & Calpe, Madrid.

Slager, Emile (2010): *Las preposiciones en español*, Editorial Castalia, Madrid.

Sopuruchi, Christian y Ibeneme, Joy Charles (2018): *A Comparative Analysis of Lexical Prepositions in the Ibibio and Igbo Languages*, «Macrolinguistics», 6.9, págs. 36-45. doi: 10.26478/ja2018.6.9.3.

Srikumar, Vivek y Roth, Dan (2013a): *Modeling Semantic Relations Expressed by Prepositions*, Transactions of the Association for Computational Linguistics, págs. 231-242.

Srikumar, Vivek y Roth, Dan (2013b): *An inventory of Preposition Relations*, <https://arxiv.org/abs/1305.5785v1>

Solías, Teresa; Pisabarro Marrón, Alma M.; y Cardeñoso-Payo, Valentín (2002): *Análisis del Castellano Guiado por una Jerarquía Preposicional*, «Revista de la SEPLN», 29.1, págs. 113-120.

Song, Chenchen (2019): *On the Formal Flexibility of Syntactic Categories*, tesis doctoral, Universidad de Cambridge, Cambridge.

Stadler, Leon de y Eyrich, Christoph (eds.) (1999): *Issues in Cognitive Linguistics*, Mouton, Berlín.

Stewart, John (2015): *Learning Functional Prepositions*, tesis doctoral, CUNY Academic Works, Nueva York.

Strzalkowski, Tomek (1999): *Natural Language Information Retrieval*, Kluwer Academic Publishers, Dordrecht.

Suárez, Armando; Palomar, Manuel; y Rigau, German (2005): *Reentrenamiento: Aprendizaje Semisupervisado de los Sentidos de las Palabras*, «Procesamiento del Lenguaje Natural», 35, págs. 49-66.

Surányi, Balázs (2009): *Verbal particles inside and outside vP*, «Acta Linguistica Hungarica», 56.2, págs. 201-249. En DOI: 10.1556/ALing.56.2009.2-3.3.

Svenonius, Peter (2006): *The Extended Projection of P*. University of Tromsø. Fuente http://www.diffusion.ens.fr/databis/diffusion/bonus/2006_09_19_svenonius.pdf.

Svenonius, Peter (2007): *Adpositions, Particles, and the Arguments They Introduce*, en Reuland, Eric; Bhattacharya, Tanmoy; y Spathas, Giorgos (eds.) (2007): *Argument Structure*, John Benjamins, Ámsterdam, págs. 71–110.

Svenonius, Peter (2010): *Spatial P in English*, en Cinque, Guglielmo y Rizzi, Luigi (Eds.): (2010): *Mapping Spatial PPs: The Cartography of Syntactic Structures*, Vol. VI, Oxford University Press, Oxford, págs. 127–160.

Svenonius, Peter (2012): *Structural decomposition of spatial adpositions*. En *The Meaning of P*, Universität Bochum, http://www.linguistics.rub.de/TheMeaningofP2012/Svenonius_Handout.pdf

Svenonius, Peter (2016): *Emergent Extended Projections*, CASTL, Universidad de Troms, The Arctic University of Norway, Troms.

Svorou, Soteria (1994): *The grammar of space*, John Benjamins, Amsterdam.

Talmy, Leonard (1983): *How language structures space*; en Pick, Herbert L. y Acredolo, Linda P. (eds.), *Spatial orientation: Theory, research, and application*, Plenum Press, Nueva York, págs. 225-282.

Talmy, Leonard (2000): *Toward a Cognitive Semantics*, The MIT Press, Cambridge, Massachussets.

Talmy, Leonard (2010): *How spoken in and signed language structure space differently*. En Vyvyan Evans y Paul Chilton (eds.), *Language, cognition, and space: The state of the art and new directions*, Equinox, Londres.

Tammenga-Helmantel, Marjon (1998): *Simplex Adpositions and Vector Theory*, «The Linguistic Review», 15, págs. 361-388.

Taulé, Mariona; Martí, M. Antònia; y Recasens, Marta (2008): *Ancora: Multilevel Annotated Corpora for Catalan and Spanish*, en *Proceedings of 6th International Conference on Language Resources and Evaluation*, Marrakesh, págs. 96-101.

- Taulé, Mariona; Martí, M. Antònia; Peris, Aina; Rodríguez, Horacio; Moreno, Lidia y Moreda, Paloma: (2012): *IARG-Ancora: Anotación de los corpus AnCora con argumentos implícitos*, «Procesamiento del Lenguaje Natural», 49, Madrid, págs. 181-184.
- Taylor, John (1986): *Contrasting Prepositional Categories: English and Italian*, en Rudzka-Ostyn, B. (1987): *Topics in cognitive linguistics*, John Benjamins, Amsterdam, págs. 299-326.
- Taylor, John (1995): *Linguistic Categories: Prototypes in Linguistic Theory*, Clarendon Press, Oxford.
- Taylor, John (2004): *Linguistic Categorization*, Oxford University Press, Oxford.
- Terzi, Arhonto (2007): *Locative Prepositions and Place*, en Cinque, Guglielmo y Rizzi, Luigi (eds.) (2010): *Mapping Spatial PPs. The Cartography of Syntactic Structures*, Oxford University Press, Oxford, págs. 32-62.
- Tesak, Jürgen y Hummer, Peter (1994): *A note on Prepositions in Agrammatism*, «Brain and Language», 46, págs. 463-468.
- Thomas, Jenny y Short, Mick (eds.) (1996): *Using Corpora for Language Research*, Longman, Londres.
- Tognini- Bonelli, Elena (2001): *Corpus linguistics at work*, Benjamins, Amsterdam.
- Toledo y Huerta, de, Álvaro S. Octavio (2016): *Los relacionantes locativos en la historia del español*, Walter de Gruyter, Berlín.
- Tordera Yllescas, Juan Carlos (2011): *Lingüística computacional. Tecnologías del habla*, Quaderns de Filologia, Anejo 74, Universidad de Valencia, Valencia.
- Tordera Yllescas, Juan Carlos (2012): *El abecé de la Lingüística computacional*, Arco Libros, Madrid.
- Tornel Sala, José Luis (2006a): *El suplemento y el complemento adverbial: precisiones categoriales*, «Philologia Hispalense», 20, págs. 42-52.
- Tornel Sala, José Luis (2006b): *La obligatoriedad del suplemento*, «AnMal», 20, págs. 1-24.
- Trabelsi, Hamida (2014): *Étude diachronique des prépositions latines ad et de*, «Lingvisticæ Investigationes», 37.1, págs. 23-68.
- Tremblay, Mireille (1998): *Lexical and non-Lexical Prepositions in French*, en Di Sciullo, A. M. (ed.) (1998): *Configurations*, Cascadilla Press, Somerville, págs.77-98.
- Tremblay, Mireille (1999): *Du estatus des prépositions dans la grammaire*, «Revue Québécoise de Linguistique», 27.2, págs. 167-183.

Trovato, Giuseppe (2013): *El régimen preposicional en español e italiano: breve estudio contrastivo y traductológico*, «Ogigia», 13, págs. 19-34.

Trujillo, Arturo (1995): *Towards a Cross-Linguistically Valid Classification of Spatial prepositions*, «Machine Translation», 10.1-2, págs. 93-141.

Trujillo, Ramón (1971): *Notas para un estudio de las preposiciones españolas*, «Thesaurus», 26, págs. 234-279.

Tseng, Jesse (1999): *A Semantic Approach to Lexical and Functional Approach*, Workshop on Semilexical Heads, Universidad de Tilburg, Tilburg.

Tseng, Jesse (2005): *The representation and Selection of Prepositions*, tesis doctoral, Universidad de Edimburgo, Edimburgo.

Tungseth, Mai Ellin (2008): *Verbal Prepositions and Argument Structure in Norwegian*, John Benjamins, Amsterdam.

Turney, Peter D. (2006): *Similarity of Semantics Relations*, «Computational Linguistics», 32.3, págs. 379-416.

Turney, Peter D. y Pantel, Patrick (2010): *From frequency to meaning: Vector space models of semantics*, «Journal of Artificial Intelligence Research», 37.1, págs. 141–188.

Ulinic, Daniel (2018): *Why Neural Networks and Deep Learning Are the Future in Machine Learning*, Escuela de Ingeniería Informática, UAB, Barcelona.

Ungerer, Friedrich y Schmid, Hans-Jörg (1996): *Introduction to Cognitive Linguistics*, Longman, Londres.

Urbizagástegui, Rubén y Restrepo, Cristina (2011): *La ley de Zipf y el punto de transición de Goffman en la indización automática*, «Investigación Bibliotecológica», 25.54, págs. 71-92.

Urrutia, Hernán (1984): *Conocimiento, lenguaje y gramática en la obra de Andrés Bello (1781-1865)*, «Historiografía Lingüística», 11, Amsterdam, págs. 263-286.

Ursini, Francesco-Alessio y Akagi, Nobuaki (2013): *Another look at Modification in Spatial Prepositions*, «International Theoretical Linguistics», 5.2, págs. 38-84.

Vallez, Mari y Pedraza-Jiménez, Rafael (2007): *El Procesamiento del Lenguaje Natural en la Recuperación de Información Textual y áreas afines*, «Hipertext», 5. [Http://www.hiertext.net](http://www.hiertext.net)

VVAA (2014): *Computational Linguistics*, <https://plato.stanford.edu/entries/computational-linguistics/>

Vandeloise, Claude (1986) : *L'espace en français: sémantique des prépositions spatiales*, Editions du Seuil (Travaux Linguistiques), París.

Vandeloise, Claude (1991): *Spatial Prepositions: A case Study for French*, University Chicago Press, Chicago.

Vandeloise, Claude (1993): *La couleur des prépositions*, «Langages», 110, págs. 281-308.

Vandeloise, Claude (1994): *Methodology and analyses of the preposition*, en «Cognitive Linguistics», 5, págs. 157-184.

Vapnik, Vladimir (1998): *The Nature of Statistical Learning Theory*, Springer, Nueva York.

Varas San Vicente, Miguel Ángel (2008): *Categorías funcionales, léxicas y rasgos. Un intento de simplificación*, «DICENDA. Cuadernos de Filología Hispánica», 26, págs. 291-322.

Vate, Marleen van de (2008): *Prepositions in Krio*, en Nordlyd, «Tromsø Working Papers in Linguistics», 33.2, número especial sobre *Adpositions*, editado por Svenonius, Peter, págs. 234–252. CASTL, Tromsø. <http://www.ub.uit.no/munin/nordlyd>.

Vázquez, Glòria; Alonso, Laura; Capilla, Joan A.; Castellón, Irene; y Fernández, Ana (2006): *SenSem: sentidos verbales, semántica oracional y anotación de corpus*, «Procesamiento del Lenguaje Natural», 37, págs. 113-119.

Vendler, Zeno (1967): *Linguistics in Philosophy*, Cornell University Press, Ithaca y Londres.

Vez Jeremías, José Manuel (1982): *El modelo casual de Charles J. Fillmore: bases para una orientación semántica en la enseñanza del inglés*, «Cauce», 5, págs. 235-249.

Waluch-de la Torre, Edyta (2007): *Análisis comparado de las preposiciones espaciales del español, el portugués y el polaco*, tesis doctoral, Universidad de Varsovia y Universidad de Granada, Editorial de la Universidad de Granada, Granada.

Wermater, Stefan; Riloff, Ellen; y Scheler, Gabriele (eds.) (1996): *Connectivist, statistical and Symbolic approaches to learning for Natural Language Processing*, Springer, Berlín.

Wonnacott, Thomas H. y Wonnacott, Ronald J. (1997): *Introducción a la Estadística*, Limusa, México.

Woolford, Ellen (2006): *Lexical Case, Inherent Case, and Argument Structure*, «Linguistic Inquiry», 37.1, págs. 111-130.

Wunderlich, Dieter (1991): *How do prepositional phrases fit into compositional syntax and semantics*, «Linguistics», 29, págs. 591-621.

Wunderlich, Dieter (1993): *On German um: semantic and conceptual aspects*, «Linguistics», 31.1, págs. 111-133.

Wunderlich, Dieter (1996): *Lexical Categories*, «Theoretical Linguistics», 22, págs. 1-47.

Yadroff, Michael (1999): *Formal Properties of Functional Categories: The Minimalist Syntax of Russian Nominal and Prepositional Expressions*, tesis doctoral, Indiana University, Bloomington.

Yadroff, Michael y Franks, Steven (1999): *The Origin of Prepositions*, «FDSL», 3, Leipzig, pág. 1-18.

Zajara-Nekrasova, E. (1993): *Casos y preposiciones en el idioma ruso*, Rubiños, Madrid.

Zamboni, Alberto (1992): *Postille alla discussione sull'accusativo preposizionale*, en Lorenzo, Ramón (ed.) (1992): *Actas do XIX Congreso Internacional de Lingüística e Filoloxía Románicas*, Universidad de Santiago de Compostela, Santiago de Compostela, págs. 777-802.

Zegrean, Iulia (2007): *From locative to De directional phrases in Romanian*, *Diacronia*, Universidad de Bucarest, Bucarest, págs. 160-170.

Zeller, Jochen (1999): *Lexical particles and semi-lexical postpositions*, en Corver, Norbert y Riemsdijk, Henk van (eds.) (2001): *Semi-Lexical Categories*, Mouton de Gruyter, Berlín, págs. 505-549.

Zeller, Jochen (2001): *Particle Verbs and Local Domains*, John Benjamins, Amsterdam.

Zeller, Jochen y Ngoboka, Jean Paul (2018). *Agreement with locatives in Kinyarwanda: a comparative analysis*, «*Journal of African Languages and Linguistics*», 39.1, págs. 65-106.

Zelinsky-Wibblet, Cornelia (ed.) (1993): *The semantics of Prepositions*, Mouton de Gruyter, Berlín.

Zipf, George Kingsley (1949): *Human Behaviour and the Principle of Least-Effort*, Addison-Wesley, Cambridge.

Zwarts, Joost (1992): *X'-Syntax – X'-Semantics. On the interpretation of Functional and Lexical Heads*, Universidad de Utrecht, Utrecht.

Zwarts, Joost (1997a): *Vectors as relative positions: A compositional semantics of modified PP*, «*Journal of Semantics*», 14, págs. 57-86.

Zwarts, Joost (1997b): *Lexical and Functional Properties of Prepositions*, en Haumann, Dagmar y Schierholz, Stefan J. (eds.) (1997): *Lexicalische und grammatische Eigenschaften präpositionaler Elemente*, Max Niemeyer, Tübingen.

Zwarts, Joost (1999): *Vector Space Semantics: A Model-Theoretic Analysis of Locative Prepositions*, «*Journal of Logic, Language and Information*», 9.2, págs. 161-211.

Zwarts, Joost (2005): *Prepositional Aspect and the Algebra of Paths*, «*Linguistics and Philosophy*», 28, págs. 739-779.

Zwarts, Joost y Peter Gärdenfors (2016): *Locative and directional prepositions in conceptual spaces: The role of polar convexity*, «*Journal of Language, Logic and Information*», 25.1, págs. 109-138.

Zwarts, Joost (2017): *Spatial Semantics: Modelling the meaning of prepositions*, «*Language and Linguistics Compass*», 12.5, págs. 1-20.

Zwicky, Arnold M. (1987): *French prepositions. No peeking*, «*Phonology Yearbook*», 4, págs. 211-27.

APÉNDICE

A. Tabla de datos del test del cuarto experimento

TABLA DE RESULTADOS (POR CLASES PREPOSICIONALES) ACUERDO PORCENTUAL, DISPERSIÓN Y ENTROPÍA DE LA DISPERSIÓN DEL USO PREPOSICIONAL														
D = dispersión A= Acuerdo en porcentaje H= Entropía														
CLASE PREPOSICIONAL Número de respuestas	NÚMERO Y POSICIÓN DE LAS PREPOSICIONES PROPUESTAS										D	A %	H	
F(1) 142	123	16	1	1	1							4	86,61	0,68
F(2)	137	2	1	1	1							4	96,47	0,28
F(3)	91	21	16	9	4	1						5	71,12	1,56
F(4) 187	183	1	1	1	1							4	97,86	0,19
F(5)	89	58	23	16	1							4	47,05	1,74
F(6)	175	9	1	1	1							4	93,58	0,42
F(7) 152	140	9	2	1								3	92,76	0,48
F(8)	141	5	3	2	1							4	92,76	0,50
F(9)	145	4	2	1								3	98,66	0,38
F(10)	144	3	3	2								3	94,73	0,43
F(11)	149	3										1	98,02	0,22
F(12)	143	9										1	94,07	0,45
F(13)	145	5	2									2	95,39	0,37
F(14)	96	21	13	9	7	6						5	63,15	1,56
F(15) 145	125	16	3	1								3	86,20	0,73
F(16)	143	2										1	98,62	0,29
F(17)	136	6	1	1	1							4	92,41	0,45
F(18)	111	16	12	3	2	1						5	76,55	1,28
F(19)	124	11	9	1								3	84,13	0,88
F(20)	84	45	13	1	1	1						5	57,93	1,54
SF(1) 142	94	26	13	6	2	1						5	66,19	1,53
SF(2)	84	40	12	4	1	1						5	59,15	1,50
SF(3)	86	37	16	1	1	1						5	60,56	1,44
SF(4)	93	35	5	4	3	2						5	65,49	1,41
SF(5)	134	4	2	1	1							4	94,36	0,47
SF(6)	122	8	7	3	1	1						5	85,91	0,85

SF(7)	112	21	3	3	2	1						5	78,87	1,04
SF(8)	93	30	5	5	4	3	2					6	65,49	1,47
SF(9) 187	127	55	3	1	1							4	65,24	1,11
SF(10)	67	47	32	20	17	2	2					6	34,75	2,19
SF(11)	107	43	19	5	5	4	4					6	57,21	1,68
SF(12)	140	25	9	7	5	1						5	72,19	1,29
SF(13) 152	106	40	3	2	1							4	69,73	1,16
SF(14)	146	6										1	96,05	0,46
SF(15)	133	10	7	2								3	87,8	0,75
SF(16) 145	129	10	3	1	1	1						5	88,96	0,64
SF(17)	122	20	1	1	1							4	84,13	0,78
SL(1) 142	91	23	16	8	2	2						5	64,08	1,59
SL (2)	49	32	27	16	13	4	1					6	34,50	2,28
SL (3)	117	12	8	3	1	1						5	82,39	0,98
SL (4) 187	121	52	8	3	2	1						4	64,70	1,32
SL(5)	58	49	32	20	13	8	4	3				7	31,01	2,27
SL(6)	91	67	12	5	4	2	1					6	48,66	1,70
SL(7)	97	40	21	16	7	2	2	1	1			8	51,87	1,87
SL(8)	132	21	15	14	3	1	1					6	70,58	1,41
SL(9)	95	47	19	9	7	6	3	1				7	50,80	1,87
SL(10)	164	13	5	2	2	1						5	88,77	0,75
SL(11) 152	134	10	3	2	2	1						5	88,15	0,74
SL(12)	89	53	7	1	1	1						5	58,55	1,32
SL(13)	139	9	3	1								3	91,44	0,57
SL(14)	109	32	6	5								3	71,71	1,27
SL(15) 145	128	17										1	88,27	0,75
SL(16)	110	22	5	5								3	75,86	1,14
SL(17)	128	12	3	2								3	89,51	0,69
SL(18)	120	16	4	2	1	1	1					6	82,75	0,90
SL (19)	71	51	5	3	1	1	1					6	50,34	1,48
SL (20)	113	17	5	4	4	1	1					6	77,93	1,14
SL (21)	100	25	7	6	2	2	1	1	1			8	68,96	1,45
SL (22)	67	39	28	7	3	1						5	46,20	1,62
L (1) 142	61	38	22	20	1							4	42,95	1,96
L (2)	71	22	15	13	12	9						5	50	2,12
L (3)	96	42	2	1	1							4	67,60	1,13
L (4)	133	8	1									2	93,66	0,55
L (5)	88	34	10	4	4	1	1					6	62,67	1,52
L (6)	106	25	8	2	1							4	74,64	1,15
L (7)	64	34	21	10	8	2	2	1				7	42,25	2,00
L (8)	85	39	10	4	4							4	59,15	1,57
L (9) 187	89	32	21	18	13	8	3	2	1			8	47,59	2,08
L (10)	130	19	15	10	3	3	2	2	2	1		9	69,51	1,44
L (11)	152	19	9	5	2							4	81,28	1,01
L (12)	49	46	39	26	15	9	2	1				7	26,20	2,37
L (13) 152	52	38	32	23	2	2	1	1	1			8	35,86	2,1
L (14)	95	24	12	11	4	4	2					6	65,51	1,67

L (15)	76	62	9	4	1							4	52,41	1,59
L (16)	104	16	9	8	6	5	1	1	1	1		9	69,33	1,52
L (17)	88	43	16	3	1	1						5	60,68	1,49
L (18)	97	32	12	8	2	1						5	66,89	1,68
L (19)	105	35	8	1	1	1	1					6	72,41	1,25
L (20)	82	34	17	8	5	4	2					6	53,28	1,84
L (21)	50	32	25	19	7	6	5	4	2	1	1	10	32,89	2,20
L (22)	141	8	1	1	1							4	92,76	0,5
L (23) 145	70	52	17	2	2	1	1					6	44,27	1,62
L (24)	82	35	26	1	1							4	56,55	1,54
L (25)	125	10	5	1	1	1	1	1				7	86,20	0,85
L (26)	78	58	5	2	1	1						5	53,79	1,36
L (27)	70	44	20	6	2	1	1	1				7	48,27	1,86
L (28)	84	33	18	3	2	2	2	1				7	57,93	1,60
L (29)	85	32	15	4	3	3	2	1				7	58,62	1,64
L (30)	103	25	13	2	1	1						5	71,03	1,28
L (31)	123	16	3	2	1							4	84,82	0,81

B. Tabla de valores para el cálculo de χ^2

Frecuencias a partir de los datos de la tabla 30 de contingencia (página 248)		
Valor (frecuencia observada)	Cálculo a partir de los datos de la tabla	Resultado (frecuencia teórica)
2623	$\frac{3073.9823}{13915}$	2169
450	$\frac{3073.4092}{13915}$	903
2028	$\frac{2630.9823}{13915}$	1856
602	$\frac{2630.4092}{13915}$	773
2347	$\frac{3503.9823}{13915}$	2472
1156	$\frac{3503.4092}{13915}$	1030
2825	$\frac{4709.9823}{13915}$	3324
1884	$\frac{4709.4092}{13915}$	1384

C. Cuadro de respuestas comparativas entre los participantes que han repetido en dos momentos distintos (para calcular el coeficiente kappa de Cohen)

En la tabla se recoge la comparación entre los resultados de la primera ocasión en que participaron, letra normal; y la segunda ocasión, letra en negrita. En la primera fila se recogen los datos del modelo, y en la segunda las clases (que se van repitiendo como recordatorio). A su vez, en la primera columna se recoge el número asignado de manera aleatoria a los 11 participantes que, amablemente, han repetido el test.

MODE LO del test	de, para	en, en	hacia, a, en	para en	de, de	a, a, a	a, en, de, con	para	De, a, de	de, por, por
Tipo de preposic ión →	SL SL	L SL	L F L	SL L	SF SL	F L SL	F L SL SL	L	SF F F	SF F F
1	de, para	en, de	hacia, a, en	para, en	de, de	a, a, a	a, en, de, con	para	en, a, de	sobre, en, por
	de, para	en, de	hacia, a, en	para, en	de, de	a, en, tras	a, en, de, entre	para	de, a, de	entre, en, por
2	de, para	con, por	hacia, a por	para, en	de, de	a, a, a	a, en, de, con	para	de, a, de	sobre, por, ante, por
	de, para	con, por	hacia, a, por	para, para	de, de	a, para, a	a, en, de, con	para	de, a, de	en, por, por
	SL SL	L SL	L F L	SL L	SF SL	F L SL	F L SL SL	L	SF F F	SF F F
3	de, para	con, por	hacia, a, en	para, en	de, con	a, para, a	a, en, de, con	para	de, a, de	en, por, por
	de, para	con, por	hacia, a, en	para, para	de, de	a, para, a	a, en, de, con	para	de, a, de	sobre, en, por
	SL SL	L SL	L F L	SL L	SF SL	F L SL	F L SL SL	L	SF F F	SF F F
4	sobre, para	con, por	hacia, a, en	para, en	de, sobre	a, durante, por	a, durante, de, con	durante	de, a, de	sobre, en, por

	de, para	con, por	hacia, a, por	para, para	de, de	a, a, a	a, por, de, con	para	de, a, de	sobre, por, por
	SL SL	L SL	L F L	SL L	SF SL	F L SL	F L SL SL	L	SF F F	SF F F
5	de, para	mediante, por	hacia, a, por	de, para	de, de	a, a, a	a, por, de, con	para	de, hasta, de	de, por, por
	de, para	en, de	hacia, a, por	para, para	de, de	a, a, a	por, en, de, con	para	de, hasta de	de, por, por
6	de, para	con, por	hacia, a, por	para, para	de, de	a, en, por	a, en, de, con	para	de, de, para	para, en, por
	de, para	con, de	hacia, a, en	para, en	de, de	a, a, a	a, en, de, con	para	de, de, hast a	para, en, por
7	de, para	con, de	por,a, por	para, de	de, de	a, para, por	a, durante, de, entre	para	de, a, para	de, por, por
	de, para	en, por,	hacia, a, en	para, para	de, de	a, a, a	a, durante, de, con	para	de, a, de	sobre, por, por
8	de, para	en, por	hacia, a, en	para, en	de, de	a, a, a	a, en, de, con	hasta	de, a, de	de, por, por
	de, para	en, por	hacia, a, en	para, de	de, de	a, a, a	a, en, de con	para	de, a, de	de, por, de
9	de, para	en, por	hacia, a, por	para, en	de, de	a, a, a	a, en, de, con	hasta	de, a, de	de, por, por
	de, a	con, por	hacia, a, por	para, para	de, de	a, hacia, por	a, por, de, con	para	de, a, de	de, ante, por
10	de, a	en, por	a, a, por	para, para	de de	a, a, a	a, en, de, con	para	de, a, de	de, por, por,
	de, para	con, en	a, a, por	para, para	de, para	a, para, a	a, en, de, con	para	de, a, de	en, por, por
11	de, para	en, de	hacia, para, por	para, de	de, a	a, a, por	en, para, por, hacia	para	en, por, de	de, en, de
	de, para	en, de	hacia, a, por	para, en	de, de	a, en, por	a, en, por, en	por	de, por, de	de, en, por

**D. Tablas de contingencia de algunas comparaciones entre 11
participantes que han respondido en dos ocasiones**

Preposiciones funcionales				Cálculo del kappa de Cohen
Sí = coincidencia				
No = desacuerdo				
El número indica el participante				
		1		0,58
		Sí	No	
2	Sí	5	1	
	No	1	5	
		1		1
		Sí	No	
3	Sí	6	0	
	No	0	6	
		2		0,58
		Sí	No	
4	Sí	5	1	
	No	1	5	
		2		1
		Sí	No	
5	Sí	6	0	
	No	0	6	
		3		

		Sí	No	1
6	Sí	6	0	
	No	0	6	
		3		0,58
		Sí	No	
7	Sí	5	1	
	No	1	5	
		4		0,58
		Sí	No	
8	Sí	5	1	
	No	1	5	
		4		0,67
		Sí	No	
9	Sí	5	0	
	No	1	5	
		5		0,67
		Sí	No	
10	Sí	5	0	
	No	1	5	
		5		0,66
		Sí	No	
11	Sí	4	0	
	No	2	6	
		6		0,58
		Sí	No	

1	Sí	5	1	
	No	1	5	
		6		0,58
		Sí	No	
2	Sí	5	1	
	No	1	5	
		7		0,67
		Sí	No	
3	Sí	5	0	
	No	1	5	
		7		0,58
		Sí	No	
4	Sí	5	1	
	No	1	5	
		8		0,58
		Sí	No	
5	Sí	5	1	
	No	1	5	
		8		0,67
		Sí	No	
6	Sí	5	0	
	No	1	5	
		9		0,67
		Sí	No	
7	Sí	5	0	

	No	1	5	
		9		0,58
		Sí	No	
8	Sí	5	1	
	No	1	5	
		10		0,58
		Sí	No	
9	Sí	5	1	
	No	1	5	
		10		0,66
		Sí	No	
11	Sí	4	0	
	No	2	6	
		11		0,66
		Sí	No	
1	Sí	4	0	
	No	2	6	
PROMEDIO TOTAL DE FUNCIONALES = 0,64				
PREPOSICIONES SEMIFUNCIONALES				CÁLCULO KAPPA DE COHEN
Total de preposiciones				
Sí = coincidència				
No = desacuerdo				
El número indica el participante				

		1		0
		Sí	No	
2	Sí	1	1	
	No	1	1	
		1		0
		Sí	No	
3	Sí	1	1	
	No	1	1	
		2		1
		Sí	No	
4	Sí	2	0	
	No	0	2	
		2		1
		Sí	No	
5	Sí	2	0	
	No	0	2	
		3		1
		Sí	No	
6	Sí	2	0	
	No	0	2	
		3		1
		Sí	No	
7	Sí	2	0	
	No	0	2	
		4		

		Sí	No	1
8	Sí	2	0	
	No	0	2	
		4		1
		Sí	No	
9	Sí	2	0	
	No	0	2	
		5		1
		Sí	No	
10	Sí	2	0	
	No	0	2	
		5		1
		Sí	No	
11	Sí	2	0	
	No	0	2	
		6		0
		Sí	No	
1	Sí	1	1	
	No	1	1	
		6		1
		Sí	No	
2	Sí	2	0	
	No	0	2	
		7		1
		Sí	No	

3	Sí	2	0	
	No	0	2	
		7		1
		Sí	No	
4	Sí	2	0	
	No	0	2	
		8		1
		Sí	No	
5	Sí	2	0	
	No	0	2	
		8		1
		Sí	No	
6	Sí	2	0	
	No	0	2	
		9		1
		Sí	No	
7	Sí	2	0	
	No	0	2	
		9		1
		Sí	No	
8	Sí	2	0	
	No	0	2	
		10		1
		Sí	No	
9	Sí	2	0	

	No	0	2	
		10		1
		Sí	No	
11	Sí	2	0	
	No	0	2	
		11		0
		Sí	No	
1	Sí	1	1	
	No	1	1	
PROMEDIO TOTAL DE SEMISEMIFUNCIONALES = 0,80				

PREPOSICIONES SEMILÉXICAS				CÁLCULO KAPPA DE COHEN
Total de preposiciones				
Sí = coincidencia				
No = desacuerdo				
El número indica el participante				
		1		0,55
		Sí	No	
2	Sí	7	2	
	No	2	7	
		1		0,44
		Sí	No	
3	Sí	7	3	

	No	2	6	
		2		0,33
		Sí	No	
4	Sí	7	4	
	No	2	5	
		2		0,11
		Sí	No	
5	Sí	7	6	
	No	2	3	
		3		0,44
		Sí	No	
6	Sí	6	2	
	No	3	7	
		3		0,44
		Sí	No	
7	Sí	6	2	
	No	3	7	
		4		0,33
		Sí	No	
8	Sí	4	1	
	No	5	8	
		4		0,22
		Sí	No	
9	Sí	4	6	
	No	5	3	

		5		0,22
		Sí	No	
10	Sí	6	4	
	No	3	5	
		5		0,22
		Sí	No	
11	Sí	6	4	
	No	3	5	
		6		0,55
		Sí	No	
1	Sí	7	2	
	No	2	7	
		6		0,55
		Sí	No	
2	Sí	7	2	
	No	2	7	
		7		0,44
		Sí	No	
3	Sí	6	2	
	No	3	7	
		7		0,33
		Sí	No	
4	Sí	4	7	
	No	5	2	
		8		

		Sí	No	0,55
5	Sí	8	3	
	No	1	6	
		8		0,66
		Sí	No	
6	Sí	7	1	
	No	2	8	
		9		0,44
		Sí	No	
7	Sí	6	2	
	No	3	7	
		9		0,55
		Sí	No	
8	Sí	6	1	
	No	3	8	
		10		0,22
		Sí	No	
9	Sí	5	3	
	No	4	6	
		10		0,11
		Sí	No	
11	Sí	5	4	
	No	4	5	
		11		0,33
		Sí	No	

1	Sí	5	2	
	No	4	7	
PROMEDIO TOTAL DE SEMILÉXICAS 0,38				

PREPOSICIONES LÉXICAS				CÁLCULO KAPPA DE COHEN
Total de preposiciones Sí = coincidencia No = desacuerdo El número indica el participante				
		1		0,75
		Sí	No	
2	Sí	8	2	
	No	0	6	
		1		0,75
		Sí	No	
3	Sí	8	2	
	No	0	6	
		2		0,25
		Sí	No	
4	Sí	6	4	
	No	2	4	
		2		0,25
		Sí	No	
5	Sí	6	4	

	No	2	4	
		3		0,55
		Sí	No	
6	Sí	6	2	
	No	2	6	
		3		0,25
		Sí	No	
7	Sí	6	4	
	No	2	4	
		4		0,25
		Sí	No	
8	Sí	4	6	
	No	4	2	
		4		0
		Sí	No	
9	Sí	4	4	
	No	4	4	
		5		0,25
		Sí	No	
10	Sí	4	6	
	No	4	2	
		5		0,12
		Sí	No	
11	Sí	4	5	
	No	4	3	

		6		0,85
		Sí	No	
1	Sí	6	0	
	No	2	8	
		6		0,55
		Sí	No	
2	Sí	6	2	
	No	2	6	
		7		0,25
		Sí	No	
3	Sí	4	2	
	No	4	6	
		7		0
		Sí	No	
4	Sí	4	4	
	No	4	4	
		8		0,25
		Sí	No	
5	Sí	6	4	
	No	2	4	
		8		0,55
		Sí	No	
6	Sí	6	2	
	No	2	6	
		9		

		Sí	No	0
7	Sí	4	4	
	No	4	4	
		9		0,25
		Sí	No	
8	Sí	6	4	
	No	2	4	
		10		0,12
		Sí	No	
9	Sí	4	3	
	No	4	5	
		10		0,25
		Sí	No	
11	Sí	5	3	
	No	3	5	
		11		0,62
		Sí	No	
1	Sí	8	3	
	No	0	5	
PROMEDIO TOTAL DE LÉXICAS 0,54				

E. Tablas de datos sobre el coeficiente kappa de Cohen para la comparación de 11 participantes en dos momentos distintos

Las cifras indican la concordancia y la discordancia entre la primera ocasión (primera fila) y la segunda (bis) que han hecho los participantes seleccionados.

Evaluación comparada de los participantes con ellos mismos							
Tabla de contingencia para obtener los coeficientes kappa de Cohen							
Clases funcionales y semifuncionales							
Número del participante	Funcionales		Coefficientes kappa	Número	Semifuncionales		Coefficientes kappa
	Sí	No			Sí	No	
1 / 1 bis	Sí	No	0,83	1 / 1 bis	Sí	No	0,50
Sí	6	1		Sí	1	1	
No	0	5		No	0	2	
2 / 2 bis	Sí	No	0,83	2 bis	Sí	No	1
Sí	5	0		Sí	2	0	
No	1	6		No	0	2	
3 / 3 bis	Sí	No	0,83	3 bis	Sí	No	1
Sí	6	1		Sí	2	0	
No	0	5		No	0	2	
4 / 4 bis	Sí	No	0,66	4 / 4 bis	Sí	No	1
Sí	5	1		Sí	2	0	
No	1	5		No	0	2	
5 / 5 bis	Sí	No	0,82	5 / 5 bis	Sí	No	1
Sí	5	2		Sí	2	0	
No	0	4		No	0	2	
6 / 6 bis	Sí	No	0,47	6 / 6 bis	Sí	No	1
Sí	5	3		Sí	2	0	
No	1	3		No	0	2	
7 / 7 bis	Sí	No	0,83	7 / 7 bis	Sí	No	1
Sí	5	0		Sí	2	0	
No	1	6		No	0	2	
8 / 8 bis	Sí	No	1	8 / 8 bis	Sí	No	1
Sí	6	0		Sí	2	0	
No	0	6		No	0	2	
9 / 9 bis	Sí	No	0,66	9 / 9 bis	Sí	No	1
Sí	5	1		Sí	2	0	
No	1	5		No	0	2	
10 / 10 bis	Sí	No	1	10 / 10 bis	Sí	No	1
Sí	6	0		Sí	2	0	
No	0	6		No	0	2	
11 / 11 bis	Sí	No	0,66	11 / 11 bis	Sí	No	1
Sí	5	1		Sí	2	0	
No	1	5		No	0	2	
Promedio			0,79	Promedio			0,95

Evaluación comparada de los participantes con ellos mismos
Tabla de contingencia para obtener los coeficientes kappa de Cohen
Clases semiléxicas y léxicas

Número del participante	Semiléxicas		Coeficientes kappa	Número	Léxicas		Coeficientes kappa
	Sí	No			Sí	No	
1 / 1 bis	Sí	No	0,44	1 / 1 bis	Sí	No	1
Sí	7	3		Sí	8	0	
No	2	6		No	0	8	
2 / 2 bis	Sí	No	0,55	2 / 2 bis	Sí	No	0,37
Sí	7	2		Sí	6	3	
No	2	7		No	2	5	
3 / 3 bis	Sí	No	0,44	3 / 3 bis	Sí	No	0,75
Sí	6	7		Sí	6	0	
No	3	2		No	2	8	
4 / 4 bis	Sí	No	0,22	4 / 4 bis	Sí	No	0,12
Sí	4	2		Sí	4	3	
No	5	7		No	4	5	
5 / 5 bis	Sí	No	0,66	5 / 5 bis	Sí	No	0,25
Sí	6	0		Sí	4	2	
No	3	9		No	4	6	
6 / 6 bis	Sí	No	0,55	6 / 6 bis	Sí	No	0,62
Sí	7	2		Sí	6	1	
No	2	7		No	2	7	
7 / 7 bis	Sí	No	0,44	7 / 7 bis	Sí	No	0,37
Sí	7	3		Sí	6	3	
No	2	6		No	2	5	
8 / 8 bis	Sí	No	0,88	8 / 8 bis	Sí	No	0,37
Sí	8	0		Sí	6	3	
No	1	9		No	2	5	
9 / 9 bis	Sí	No	0,55	9 / 9 bis	Sí	No	0,12
Sí	6	1		Sí	4	3	
No	3	8		No	4	5	
10 / 10 bis	Sí	No	0,11	10 / 10 bis	Sí	No	0,50
Sí	5	4		Sí	6	2	
No	4	5		No	2	6	
11 / 11 bis	Sí	No	0,22	11 / 11 bis	Sí	No	0,25
Sí	5	3		Sí	5	3	
No	4	6		No	3	5	
Promedio			0,46	Promedio			0,42