Lenguas, lenguaje y lingüística

Contribuciones desde la Lingüística General

Adriana Gordejuela Senosiáin

Dámaso Izquierdo Alegría

Felipe Jiménez Berrio

Alberto de Lucas Vicente

Manuel Casado Velarde

(Eds.)

Título: Lenguas, lenguaje y lingüística.

Subtítulo: Contribuciones desde la Lingüística General.

Autor: A. Gordejuela Senosiáin, D. Izquierdo Alegría, F. Jiménez Berrio, A. de Lucas Vicente, M. Casado Velarde (eds.).

Editorial: Servicio de Publicaciones de la Universidad de Navarra.

Maquetación y corrección de textos: Apiedepágina.net

ISBN: 978-84-8081-478-2

Reservados todos los derechos de edición.

Queda prohibida la reproducción total o parcial de esta obra por cualquier método o procedimiento.

Lenguas, lenguaje y lingüística.

Contribuciones desde la Lingüística General

Adriana Gordejuela Senosiáin
Dámaso Izquierdo Alegría
Felipe Jiménez Berrio
Alberto de Lucas Vicente
Manuel Casado Velarde
(eds.)

ÍNDICE

PRAGMÁTICA NOMINAL EN LOS SUFIJOS DERIVADOS DE LA LENGUA ESPAÑOLA
Jaume Alavedra i Regàs **9**

LA DUPLICACIÓN PRONOMINAL EN LAS ORACIONES DE RELATIVO EN CATALÁN: UNA MARCA DE ESPECIFICIDAD
Cristina Albareda **21**

THE SYNTAX OF ELLIPSIS IN ARABIC FRAGMENT ANSWERS
Ali Algryani **35**

SELF-CONCEPT, EMOTIONAL INTELLIGENCE AND READING SKILL IN SECOND LANGUAGE ACQUISITION: IMPACT ON THE LEVEL OF ENGLISH?
Elena Alonso-Blanco, Manuel Soriano-Ferrer, Ángel López García-Molins **45**

THE INTERACTION OF EXTRAPOSITION FROM DP AND *RIGHT NODE RAISING* IN ENGLISH AND SPANISH
Marian Alves **53**

HACIA UNA SOCIOLINGÜÍSTICA DE LA ESCRITURA. LA ENCUESTA
Leopoldo Ildefonso Baliña García **65**

APORTES DEL CONCEPTO DE NORMA AL ESTUDIO DE LAS ACTITUDES LINGÜÍSTICAS
Rafael Alberto Barragán Gómez **79**

ANALITISMO FRENTE A LA NOMINALIDAD. ESTUDIO CONTRASTIVO POLACO-ESPAÑOL
Janusz Bień **91**

LA EVOLUCIÓN DE LOS VERBOS DE MOVIMIENTO EN CHUJ
Cristina Buenrostro **103**

LA PÉRDIDA DEL LÉXICO DIALECTAL: ALGUNOS DATOS DE CASTELLANOPARLANTES DE SAN SEBASTIÁN
Bruno Camus Bergareche, Sara Gómez Seibane **117**

TECHNICAL TRANSLATION, TERMINOLOGY AND THE ACCESSIBILITY OF SPECIALIZED KNOWLEDGE IN PORTUGUESE
Luis Cavaco-Cruz **129**

HACIA UN GLOSARIO DEL LÉXICO DE LA INFORMÁTICA Y LA INTERNET EN ESPAÑOL
Lirian Ciro, Neus Vila Rubio **141**

EL LINGÜISTA EN EL PARLAMENTO
Giovana de Sousa Rodrigues **153**

ALGUNOS MITOS SOBRE LA ADQUISICIÓN DE LA LENGUA MATERNA
Iván Enríquez Martínez **161**

MARCO LEGAL Y PLANIFICACIÓN LINGÜÍSTICA EN LAS COMUNIDADES BILINGÜES DE ESPAÑA
Gérard Fernández Smith, Luis Escoriza Morera **173**

CORSICAN EQUIVALENTS OF *IT*-CLEFT SENTENCES IN ENGLISH: AN OVERVIEW
Pierre-Don Giancarli **187**

DESCOMPOSICIÓN LÉXICO-CONCEPTUAL DE LOS VERBOS PARASINTÉTICOS CON PREFIJO *DES-*
Elisabeth Gibert Sotelo **203**

TWO KINDS OF MINIMAL ANSWERS TO *YES-NO* QUESTIONS IN CZECH AND SPANISH
Hana Gruet-Skrabalova **217**

SOBRE LA NATURALEZA HÍBRIDA DE LAS RELATIVAS LIBRES INDEFINIDAS
Edita Gutiérrez Rodríguez, Pilar Pérez Ocón **229**

THE LEXICALIZATION OF ENGLISH LOANWORDS INTO EGYPTIAN ARABIC
Walaa Hassan **243**

¿ES EL CONOCIMIENTO MORFOLÓGICO UN MECANISMO DETERMINANTE EN LA RECUPERACIÓN DEL LÉXICO DISPONIBLE?
Natividad Hernández Muñoz **259**

TRADUCCIÓN COMO MEDIACIÓN INTERCULTURAL: DELIMITACIÓN CONCEPTUAL Y DIMENSIONES DE UNA PRÁCTICA
Carlos Hernández Sacristán **269**

LOS CENTROS DE CARÁCTER CULTURAL EN LOS ESTUDIOS DE DISPONIBILIDAD LÉXICA: ANÁLISIS Y NUEVA PROPUESTA
María Herreros Marcilla **279**

ASPECTOS ANTROPOLÓGICOS SOBRE EL FENÓMENO DEL CAMBIO DE CÓDIGO EN COMUNIDADES BILINGÜES. EL CASO DE ESTUDIO DE LA COMUNIDAD DE HABLA ALICANTINA
José Iborra Torregrosa **291**

LA RHINOGLOTTOPHILIA DESDE UNA PERSPECTIVA TIPOLÓGICA (CON UNA NOTA SOBRE LA LENGUA VASCA)
Iván Igartua 303

DIRECTIONALITY IN ADVANCED TONGUE ROOT HARMONY
Gary Linebaugh 315

EL BILINGÜISMO INDIVIDUAL: ENFOQUES SOBRE UN CONCEPTO
Lara Lorenzo Herrera 325

LA PREPOSICIÓN EN ESPAÑOL Y EN CHINO. DIFERENCIAS CONSERVADORAS ADITIVAS
Mª Azucena Penas Ibáñez, Jinbai Zhang 333

LENGUA Y CULTURA EN EL EPISTOLARIO DE PEDRO DE MUGICA A ANTONI M. ALCOVER
Maria Pilar Perea 345

INFLUENCIA DE LOS SONIDOS ADYACENTES Y LOS MÁRGENES DE DISPERSIÓN DE LAS VOCALES MEDIAS ANTERIORES DEL CATALÁN EN HABLA ESPONTÁNEA
Agnès Rius-Escudé, Francina Torras Compte 357

LA CATEGORIZACIÓN DEL CONOCIMIENTO ESPECIALIZADO: ANÁLISIS Y CLASIFICACIÓN DE LAS RELACIONES DE SIGNIFICADO ENTRE TÉRMINOS COHIPÓNIMOS
Mercedes Roldán Vendrell 369

LOS ESQUEMAS ENTONATIVOS DEL FRAGATINO: DESCRIPCIÓN Y COMPARACIÓN CON VARIEDADES ROMÁNICAS PRÓXIMAS
Lourdes Romera, Ana Ma. Fernández Planas, Wendy García-Elvira, Paolo Roseano, Josefina Carrera, Albert Ventayol, Eugenio Martínez Celdrán 389

LA INFLUENCIA DEL GÉNERO TEXTUAL EN LA PRODUCCIÓN DEL ELEMENTO PROSÓDICO
Asier Romero, Aintzane Etxebarria, Iñaki Gaminde, Urtza Garay 401

A VUELTAS CON EL YEÍSMO: PRODUCCIÓN FONÉTICA, PERCEPCIÓN CATEGORIAL Y CAMBIO
Assumpció Rost Bagudanch 417

ESTUDIO DE DISPONIBILIDAD LÉXICA EN APRENDICES ITALIANOS DE ESPAÑOL: ANÁLISIS CUANTITATIVO
Roberto Rubio Sánchez 429

SOME OBSERVATIONS ON MORPHOLOGICAL CASE IN OLD SAXON
Iker Salaberri 443

REPRESENTACIONES SOCIALES EN TORNO AL PROCESO DE EVALUACIÓN: ALUMNOS DE ELE EN LA CIUDAD DE MÉXICO
Yuritzy de la Paz Sánchez López **455**

ESTRATEGIAS LINGÜÍSTICAS PARA LA CODIFICACIÓN DE LA CAUSALIDAD EN TEXTOS ESCRITOS POR ESTUDIANTES UNIVERSITARIOS
Paola Sánchez Portilla, Celia Díaz Argüero **463**

LA INFLUENCIA DEL CONOCIMIENTO DE VARIAS LENGUAS EXTRANJERAS EN EL LÉXICO DISPONIBLE
Inmaculada Clotilde Santos Díaz **477**

EN BUSCA DE LA IDENTIDAD LINGÜÍSTICA. EL ESPAÑOL CARA A CARA CON EL CROATA
Anita Skelin Horvat, Maša Musulin **485**

CARACTERÍSTICAS ACÚSTICAS DE LA ASPIRACIÓN DE /-S/ IMPLOSIVA EN EL ESPAÑOL HABLADO EN MÁLAGA. HACIA LA RESILABIFICACIÓN PRESTIGIOSA DE UN SEGMENTO SUBYACENTE ENTRE LOS JÓVENES UNIVERSITARIOS
Matilde Vida-Castro **495**

MICROINFORMÁTICA. MODELOS DE VARIACIÓN INFORMATIVA MICROPARAMÉTRICA
Xavier Villalba, Sílvia Planas-Morales **507**

AINDA AS 'REFLEXÕES' SOBRE A *PRONUNCIAÇÃO* NO SÉC. XVIII PORTUGUÊS
Ana Paula Banza **517**

ASPECTOS DE LA MORFOLOGÍA PORTUGUESA EN EL SIGLO XVIII: LAS *REFLEXÕES* (1768/1842) DE FRANCISCO JOSÉ FREIRE
Maria Filomena Gonçalves **529**

DIDASCALIAS E IMPLICATURAS: UN ESTUDIO LINGÜÍSTICO-PRAGMÁTICO DEL LENGUAJE FEMENINO EN EL SIGLO XVII
Mª José Rodríguez Campillo **539**

EL CONTACTO LINGÜÍSTICO EN LA LITERATURA ALJAMIADA: ¿DÓNDE ESTÁ DIOS?
Juan Antonio Thomas **551**

PRESENTACIÓN DE LIBRO:
LA LINGÜÍSTICA EN ESPAÑA: 24 AUTOBIOGRAFÍAS INTELECTUALES
Xavier Laborda, Lourdes Romera, Ana Ma. Fernández Planas **563**

Gordejuela Senosiáin, A., Izquierdo Alegría, D., Jiménez Berrio, F., De Lucas Vicente, A. y M. Casado Velarde (eds.) (2015): ***Lenguas, lenguaje y lingüística. Contribuciones desde la Lingüística General***. Pamplona: Servicio de Publicaciones de la Universidad de Navarra. ISBN: 978-84-8081-478-2

ANALITISMO FRENTE A LA NOMINALIDAD. ESTUDIO CONTRASTIVO POLACO-ESPAÑOL

Janusz Bień
Universidad Católica de Lublin Juan Pablo II

Introducción

Según nuestra hipótesis principal, que se irá desarrollando en el cuerpo del texto[1], no existe ninguna correlación positiva entre el grado de analitismo de una lengua y su tendencia al estilo nominal. Este planteamiento se aleja de algunas opiniones que aparecen en la literatura lingüística según las cuales la tendencia analítica causa profundos cambios estructurales en las lenguas.

El español se caracteriza por un analitismo más elevado que el polaco y esta constatación se refiere tanto a la lengua general como a sus variantes funcionales. Por otro lado, los índices de nominalidad suponen una ligera inclinación del polaco por el estilo nominal en comparación con el español. Esta constatación va de alguna manera en contra de lo que dan a entender las características tipológicas de ambas lenguas.

Nuestras observaciones se basan en el censo de datos efectuado en dos corpus electrónicos: *Corpus del Español* (*CdE*) (http://www.corpusdelespanol.org/x.asp) y *Narodowy Korpus Języka Polskiego* (*NKJP*) (http://nkjp.pl) (muestras formalizadas de un millón de *tokens*) y en corpus manuales comparables, compuestos de cuatro tipos de textos funcionales (muestras reducidas a diez mil formas gráficas).

1. Definiciones del analitismo

El analitismo[2] lingüístico se entiende corrientemente como el hecho de repartir un significado en dos o más unidades mientras que la lengua dispone también de una solución sintética equivalente a nivel semántico que le permite empaquetar la información en una sola unidad lingüística. Este modo de ver el fenómeno es meramente semántico y, por un lado, se acerca al concepto de economía discursiva de la lengua, mientras que, por otro lado, se aleja de su enfoque formal del que hablaremos a continuación. En fin, el concepto de sintetismo/analitismo se reduce en esta versión corriente al número de elementos mediante los que una lengua pueda expresar el mismo significado. El fenómeno en cuestión se resalta con facilidad en el plano contrastivo, donde una lengua más flexiva, como el polaco, opone soluciones paradigmáticas (en el sentido clásico del término −un lexema con sus morfemas flexivos−) a expresiones morfosintácticas de una lengua como la es-

[1] El presente artículo se apoya casi enteramente en amplios fragmentos de otro trabajo nuestro anterior (Bień 2013: 183 y ss.).

[2] En el presente estudio, damos predilección a los vocablos *analitismo*/*sintetismo* en detrimento de los términos clásicos y más ampliamente usados *análisis/síntesis*; tratamos ambas parejas como sinonímicas.

pañola, menos flexiva y más analítica. Así pues, el polaco expresa mediante su sistema casual innumerables matices semánticos, lo que se corresponde, de manera más o menos sistemática, a expresiones preposicionales en español: de casa -*domu* (pol.) (genitivo), a María -*Marii* (dativo), con la mano -*ręką* (instrumental), etc. Por cautela, puntualicemos que se trata aquí de contextos concretos, puesto que tampoco es siempre así en español y no faltan ejemplos contrarios a su supuesta tendencia analítica en comparación con el polaco (u otras lenguas eslavas) (Bień 2010: 27-31, 2013: 198-199).

Saliendo del marco de estructuras nominales, el ejemplo que mejor ilustra la distinción formal entre los conceptos de sintetismo y analitismo es, sin duda, el modo de expresar los tiempos gramaticales en diferentes lenguas. Este tipo de ejemplos acentúa la supuesta diferencia entre las lenguas eslavas y las románicas en este plano. Por ejemplo, el polaco no posee tiempos perfectos compuestos, pero con sus formas sintéticas puede denotar varios valores gramaticales distintos[3]: *Wypił* (pol.) - *(Él) ha bebido* / *Wypiła* - *(Ella) ha bebido* / *Wypili* - *(Ellos) han bebido* / *Wypiły* - *(Ellas) han bebido*, etc. Los afijos flexivos de los que dispone el polaco (-*ił*, -*iła*, -*ili*, -*iły*) marcan el tiempo, el número, la persona e incluso el género gramaticales. El castellano, en cambio, en calidad de una lengua más analítica en estos contextos, expresa solo tres de esos valores con recursos morfológicos más numerosos (morfemas libres). Atendiendo a este concepto, se supone que las soluciones sintéticas se manifestarán con más frecuencia en la lengua con una flexión desarrollada.

1.1. Enfoque morfológico de Greenberg

Aunque han sido muchos los autores que indican este mecanismo, Greenberg (1960, 1968) es el primero en describir la manera exacta de cómo definir y evaluar el analitismo y el sintetismo morfológico de las lenguas que radica en el principio mayor: número de morfemas por palabra de promedio. Su método de análisis práctico de los datos lingüísticos es muy formal, pero los resultados numéricos son muy dispersos y varían de una publicación a otra. Existen varias dificultades a la hora de contabilizar este fenómeno en las lenguas, ya que se trata de un procedimiento que se escapa a cualquier formalización y no resulta muy factible aplicarle herramientas electrónicas, propias de la lingüística del corpus. Además, el concepto mismo de morfema no tiene, por motivos prácticos, un tratamiento igual en todas las lenguas. Es más, aparte de ser una unidad tipológicamente variable, su estatus varía también en función de la corriente lingüística que la define.

Por último, recuérdese que la aplicación del concepto morfológico del analitismo no ha sido unánime por quienes se han ocupado de este problema lingüístico[4].

Los resultados empíricos a los que llegó Greenberg han marcado los trabajos tipológicos y contrastivos durante casi medio siglo y siguen vigentes como punto de referencia también hoy en día. El lingüista americano sometió a análisis algunas lenguas del mundo con vistas a evaluar 11 índices y once ratios que, por falta de espacio, prescindimos de

[3] Greenberg (1968: 128) selecciona ejemplos ingleses (modo analítico) y latinos (modo sintético): "Thus the English construction *I shall go*, in which first person singular subject (future) and the concrete verbal meaning are each expressed by separate words, would be regarded as analytic, by contrast with the Latin translation equivalent ībō (*I shall go*) in which the root ī (*to go*) and *bō*, the future first person singular inflection, are contained in the same word".

[4] Algunos lingüistas no siguen de cerca la estela de su predecesor. Así, por ejemplo, Cowgill (1963: 120 y ss.) lleva a cabo un interesante estudio sobre los universales morfológicos en las lenguas del mundo y hace numerosas referencias tanto a la diversificación de los datos como al método según se procesen.

describir detalladamente aquí (Greenberg 1960). La tabla que insertamos a continuación contiene algunos datos numéricos selectos y resalta el rasgo mayor, el índice de síntesis y análisis:

	síntesis	aglutinación	composición	prefijación	sufijación
vietnamita	1,06	–	1,07	0,00	0,00
inglés	1,68	0,30	1.00	0,04	0,64
sánscrito	2,59	0,09	1,13	0,16	1,18
swahili	2,55	0,67	1,00	1,16	0,41
esquimal	3,72	0,03	1,00	0,00	2,72

Tabla I. Índices morfológicos de Greenberg (lenguas seleccionadas)[5].

Este recuento numérico permite al lingüista americano establecer una escala continua de dichos valores y etiquetar las lenguas según la ratio que mejor las caracterice. De hecho, se propusieron tres grupos de lenguas, cuyos límites radican en cierta medida en la convención:

– lenguas analíticas, cuya densidad morfemática es de 1 a 2,2
– lenguas sintéticas, en las que se establece de 2,2 a 3,0 morfemas por palabra
– lenguas polisintéticas, que sobrepasan la ratio de 3,0 morfemas por palabra de promedio y llegan a 3,72 (Greenberg 1968: 129).

La crítica más directa del planteamiento de Greenberg concierne sus muestras estadísticamente poco significativas. De hecho, todas las ratios evaluadas por el autor tienen como base 100 palabras gráficas, descompuestas, según varios criterios, en un número más o menos elevado de morfemas. Con la perspectiva actual se observa también que los mismos datos se citan repetidamente a través de los estudios al tiempo que faltan nuevas aportaciones numéricas más profundas. Una de las causas de esta limitación radica en la dificultad que conlleva el recuento de morfemas en un corpus compuesto por textos multilingües.

Los datos que aporta Gawełko (2001: 402-403) confirman *grosso modo* la tendencia común, pero sugieren un grado de sintetismo algo excesivo de las lenguas románicas. De las tres lenguas analizadas, −italiano, español y portugués−, el primero sería el más sintético, con 2,07 morfemas por palabra de promedio; entre el español y el portugués no se atestiguaría una diferencia muy marcada: 1,88 y 1,84 morfemas por palabra, respectivamente. No obstante, la muestra que utiliza el autor sigue siendo insuficiente para ser considerada como universalmente válida, pues se trata de la toma de 58 palabras para el portugués, 59 para el español y 69 para el italiano.

En Gawełko (2010: 51) se facilitan otros datos interesantes concernientes a las cuatro lenguas siguientes: latín, 1,77 morfemas por palabra; polaco, 1,67; italiano, 1,45; inglés, 1,15. El autor no da a conocer, sin embargo, el tamaño de la muestra que le permite calcular estos índices.

[5] La tabla recoge solo una parte de los datos facilitados por Greenberg (1960) y se inspira de un acortamiento metodológico que se hace en Moreno Cabrera (2003: 467).

Por su parte, Moreno Cabrera (2003) ofrece una interesante muestra comparativa de un texto en vasco y en español que constituye en sí un mini-corpus, sugiriendo datos numéricos sorprendentes. En resumen, estos reflejan los siguientes índices de análisis/síntesis[6]: en el texto vasco se recuentan 321 morfemas (tanto léxicos como gramaticales) repartidos en 163 palabras gráficas, lo que desemboca en un índice de análisis/síntesis de 1,97. En el texto español aparecen 353 morfemas que, divididos por el número de palabras (224), dan un índice de síntesis de 1,58 (Moreno Cabrera 2003: 471 y ss.).

1.2. Enfoque gramatical-formal

Algunos autores insisten, ante todo, en la necesidad de distinguir entre la definición formal y semántica del analitismo/sintetismo: "it is important to maintain the distinction between formal and semantic definitions of analyticity and syntheticity" (Danchev 1992: 26). Un enfoque puramente formal del fenómeno que consiste en contabilizar las marcas gramaticales libres, es decir, la frecuencia de las palabras funcionales en un texto, se observa actualmente en la investigación de Szmrecsanyi (2009, 2012) y otros estudios relacionados. Es precisamente en ellos donde se recogen consideraciones que preconizan la faceta formal del analitismo: "The formal definition […] of analyticity implies two or more language units whereas the definition of syntheticity implies one unit. This means analytic constructions are manifested in at least two free morphemes whereas synthetic constructions are manifested in one bound morpheme" (Ehret 2012: 1). Esta concepción del analitismo/sintetismo es absolutamente precisa y, aparentemente, es la única vía que permite llegar a unas medidas numéricas estables.

El método consiste en calcular la frecuencia de algunos elementos de la lengua, precisamente morfemas gramaticales libres, que son considerados como marcas definitorias del fenómeno:

> Specifically, we utilize a quantitative, language-internal measure of overt grammatical *analyticity*, defined as the text frequency of free grammatical markers, and a measure of overt grammatical *syntheticity*, defined as the text frequency of bound grammatical markers (Szmrecsanyi 2009: 319. La cursiva es del autor).

Aunque el trabajo discutido se lleva a cabo con el inglés y tiene una finalidad distinta de la nuestra, nos parece de interés enumerar todos los parámetros de analitismo que se proponen en él. En cierta manera, los tratamos como punto de partida para la elaboración de nuestra propia serie de parámetros, aplicables a los textos españoles y polacos. Son once los tipos de elementos funcionales (*synsemantic or function words*) que el autor toma en cuenta para establecer un índice de analitismo del inglés general y de sus variantes sociolingüísticas[7]: 1. Palabras introductorias de oraciones completivas (*complementizers* en la fuente; por ejemplo, *that*); 2. Conjunciones de coordinación; 3. Determinantes; 4. Marca del infinitivo inglés *to*; 5. Verbos modales (por ejemplo, *can*); 6. Partículas de negación (*not*); 7. Expresión existencial (*there*); 8.

---

[6] En su enfoque morfológico, los dos fenómenos son totalmente simétricos puesto que se miden con el mismo parámetro de carácter continuo: densidad morfemática de la palabra.

[7] Al parecer, el grado de analitismo/sintetismo de una lengua puede también sufrir una variación dialectal, lo que se desprende principalmente del primer estudio del autor (Szmrecsanyi 2009: 352).

Pronombres personales; 9. Preposiciones; 10. Marcas de comparación y grado superlativo (*more* o *most*); 11. Verbos auxiliares (*do*, *have*, *be*) (Szmrecsanyi 2012: 2).

Gracias a su claridad y rigor formal, que permite un recuento de datos muy fiable, este modo de abordar el analitismo no parece despertar controversias. No obstante, creemos que algunos de sus puntos son discutibles. Así, los verbos modales no son, a nuestro juicio, puramente funcionales al modo de los verbos auxiliares por excelencia (los del tiempo o de la pasiva), sino que conllevan un sentido léxico, contextualmente muy limitado. Por este motivo, antes de proceder al censo de los datos, es recomendable aplicar un método de desambiguación que permita fijar los usos léxicos y funcionales, ya que solamente estos últimos deberían considerarse como parámetros analíticos.

No olvidemos tampoco que la propuesta de Szmrecsanyi se refiere a textos ingleses y no es posible aplicarla sin modificaciones a otras lenguas que no disponen necesariamente de todas las marcas analíticas que acabamos de enumerar. Al llevar a cabo un análisis parecido con las lenguas románicas o el polaco sería imprescindible hacer un repaso crítico de los parámetros y fijar unos criterios evaluativos propios.

El método de Szmrecsanyi, junto con el enfoque global que se presenta a continuación, manifiesta diferencias esenciales con el método morfo-cuantitativo de Greenberg. Aunque ambos son formales y se apoyan en mayor o menor medida en la disposición de los morfemas, lo que caracteriza al primero es sobre todo el deseo de desviar la atención del concepto teórico del morfema ("avoiding reference to the morpheme construct, which is theorically not unproblematic"). Como se observa en numerosas ocasiones en la investigación de Szmrecsanyi, el término sustituyente *analytic/synthetic marker* es mucho más general y puede designar en realidad a cualquier parte del discurso. En la práctica, sus características se pueden delimitar con evidencia: se trata siempre de elementos con alguna carga gramatical. Su pertenencia al grupo de elementos analíticos o sintéticos se establece con criterios funcionales que hemos definido más arriba: elementos libres y *asemánticos*, sin significado léxico (marcas de analitismo), y elementos cuyo significado gramatical se define por la vinculación que mantienen con su lexema base (marcas de sintetismo).

Otro aspecto importante desde el punto de vista metodológico que conviene precisar es que la ratio de analitismo/sintetismo no se evalúa con todas las categorías de la lengua (lo que sería ilógico puesto que las marcas de sintetismo y las de analitismo se excluyen)[8], sino con algunos tipos de formas seleccionadas que representan unidades más o menos gramaticalizadas.

1.3. Enfoque global del analitismo y fenómenos relacionados

La definición del analitismo/sintetismo lingüístico adquiere un carácter global en los estudios tipológicos de Gawełko (2001, 2010). Según el lingüista polaco, este fenómeno causa profundos cambios estructurales en las lenguas, por lo que su comprensión no debe

[8] Los parámetros de sintetismo en el método de Szmrecsanyi se reducen a tres tipos principales y luego se extienden a varios elementos concretos. Se trata de marcas gramaticales vinculadas (*bound grammatical markers*) que pueden ser desinencias verbales como *-s* (*he walks*); nominales como *-s*, marca del plural; y afijos en algunos adjetivos de grado comparativo o superlativo (*-er*/*-est* en: *smaller*/*smallest*). El último tipo de marcas de sintetismo lo constituye un grupo de alomorfos, muy frecuentes en la conjugación irregular de los verbos ingleses, como se observa, por ejemplo, en la pareja *sing-sang*. Los elementos que no entran en los recuentos son, entre otros, morfemas cero o palabras que se descalifican por ser a la vez marcas de analitismo y de sintetismo, como en verbos funcionales que contienen morfemas flexivos o son alomorfos (*has*).

limitarse a indicar diferencias morfosintácticas entre estructuras aisladas, ni solo a establecer el número de morfemas por palabra de promedio, sino que consiste en describir cualquier mecanismo sistemático de las lenguas, sea morfológico, sintáctico o estilístico, que contribuya a soluciones analíticas. En la frecuencia de algunas estructuras y partes del discurso concretas se materializan dichas tendencias de las lenguas, denominadas *parámetros de analitismo*:

> fuerte tendencia al estilo nominal, número considerable de palabras gráficas en la frase media, bajo índice de morfemas en una palabra media, bajo índice de morfemas en una palabra media, numerosas ocurrencias de pronombres personales sujeto, de artículos, de preposiciones, de inversiones (vs, ovs), de posposiciones del adjetivo, de frases pasivas, de verbos transitivos, de expresiones presentativas del tipo fr. *il y a* o esp. *hay*, y de enunciados conformes a las exigencias de la perspectiva funcional de la frase (Gawełko 2001: 397).

Se trata de uno de los estudios más holísticos sobre el tema donde el analitismo es considerado como causa mayor de varios fenómenos gramaticales y léxicos[9]: "Sin embargo, todas las tendencias señaladas por los lingüistas no son capaces de poner en duda la extraordinaria importancia de la tendencia analítica de las lenguas indoeuropeas, que conduce a un cambio de tipo de lengua" (Gawełko 2001: 395-396). Pese a todo, el enfoque global suscita algunas objeciones, o matizaciones, que en nada debilitan el planteamiento en su conjunto, bien fundamentado tanto en su lógica como en los recuentos numéricos concretos que presenta. Así pues, a falta de una definición concreta del analitismo, se podría objetar que diferentes mecanismos lingüísticos que se describen en las fuentes citadas pueden entenderse tanto como causas o efectos o como rasgos funcionales del fenómeno. La interpretación depende de la definición adoptada, morfológica o morfosintáctica. Ambas pueden coexistir siempre y cuando la teoría de la complementariedad de fenómenos lingüísticos, existentes en dos planos distintos, sea universalmente válida.

El concepto holístico del analitismo nos parece defendible en el caso de que algunos mecanismos se traten como parámetros independientes, aunque cada uno de ellos marque una tendencia en el mismo sentido. Una visión global del analitismo reúne tanto la frecuencia de elementos sintácticos selectos como algunas soluciones derivadas del orden de los constituyentes. Sin embargo, algunas de ellas no parecen estar directamente relacionadas con el analitismo, desde nuestra perspectiva del problema.

2. Conceptos de la nominalidad y métodos de evaluación

Algunas unidades de la lengua pueden llevar carga nominal o verbal sin que sus características definitorias se correspondan con las denominaciones clásicas de las partes del discurso. Efectivamente, en algunas ocasiones, los elementos tradicionalmente conceptuados como nominales pueden conllevar rasgos verbales, y viceversa, los elementos nominales se comportan como verbales. De hecho, para medir el grado de nominalidad de

[9] El hecho de vincular el analitismo con el tipo de léxico (Gawełko 2010: 42) radica en la presencia de verbos genéricos en las expresiones analíticas con verbo como soporte o en otras parecidas. No obstante, nos parece difícil establecer una relación directa y regular entre el tipo de léxico y el analitismo en la totalidad del corpus; en nuestro sentido del término, son las palabras funcionales y asemánticas, privadas de cualquier carga léxica, las que constituyen los fundamentos definitorios del fenómeno.

un texto es necesario aplicar un parámetro de evaluación cuantitativa riguroso, que se zafe de usos marginales y categorialmente dudosos de las partes del discurso. Nos referimos, entre otros, al comportamiento transcategorial de algunas formas no personales del verbo, de los *nomina actionis* o de los adverbios. Frente a ello, es necesario proceder a unos acortamientos metodológicos y a una preselección de los elementos que sean objeto de recuentos y de análisis comparativos. Dos métodos de recuento numérico de los ejemplos del corpus parecen ser aplicables para establecer las tendencias nominales de las lenguas:

1. Calcular todos los elementos nominales y verbales posibles del corpus y definir luego cuál es la proporción numérica entre ellos.
2. Calcular la relación numérica solamente entre los elementos centrales de las dos tendencias opuestas: sustantivos y verbos.

Para establecer el índice de nominalidad en el presente trabajo hemos optado por el segundo enfoque, por motivos prácticos que nos parecen obvios: los corpus electrónicos aportan una gran cantidad de datos sobre la relación verbal-nominal que sería un error metodológico no tener en cuenta. Luego, la nominalidad calculada en los corpus electrónicos equilibrados, tratados como muestras representativas de la lengua general, es confrontada con los índices obtenidos a partir de ambos corpus manuales de discursos especializados.

Conviene preguntarse si la evaluación de la ratio de nominalidad ha de basarse solamente en los verbos flexivos o en todas las formas verbales del corpus. Las formas no personales funcionan en su mayoría como verbos predicativos y en principio no se oponen categorialmente a las variantes flexivas (Bień 2013: 68-102). Es cierto también que una parte de las formas infinitivas y participiales no pueden recibir otra interpretación que sustantiva por su alto grado de lexicalización, por lo que estas no se contarán como verbales ni en el corpus manual, ni en el corpus electrónico. Ambos métodos, y también algún método intermedio, se aplican a través de la literatura lingüística, así que los índices aportados por diversos autores no siempre han de reflejar fielmente las mismas tendencias estilísticas de la lengua general y de los lenguajes especializados. De ello viene la necesidad de establecer el grado de nominalidad de los textos con la mayor precisión posible. Para alcanzarla, hemos considerado que el método científicamente más útil y transparente que permite hacer más comparaciones entre ambas lenguas es el de calcular dos índices de nominalidad: el primero (NOM1), que se basa en los sustantivos clásicos y los verbos en uso personal, y el segundo (NOM2), que toma en cuenta todas las formas sustantivas y verbales.

2.1. Nominalidad. Datos de los corpus electrónicos y manuales

Los números que se deducen de las anotaciones que ofrece el *Corpus del Español* (CdE) permiten establecer el índice extendido de nominalidad (NOM2) a 1,3751, que es sensiblemente más bajo que el índice equivalente obtenido a partir del corpus electrónico polaco (NKJP: *Narodowy Korpus Języka Polskiego*). La evaluación del índice básico (NOM1) exige el recurso a una descripción más exacta de las anotaciones que permita delimitar los usos flexivos y no flexivos de los verbos. Al tener a nuestra disposición los datos que ofrece el corpus, el método más simple que se puede aplicar consiste en descartar ocurrencias

correspondientes a las formas no personales del verbo y los lemas que no se corresponden a los sustantivos plenos (por ejemplo, abreviaciones). Así, la proporción numérica entre los sustantivos plenos y las formas personales del verbo en la muestra formalizada de un millón de *tokens*, elevaría el índice de nominalidad (NOM1) a 2,0147. Asimismo, para la evaluación del primer índice de nominalidad (NOM1) en el corpus electrónico polaco NKJP se han contado las ocurrencias de formas sustantivas plenas y las ocurrencias de formas flexivas del verbo, y el segundo índice de nominalidad (NOM2) se ha calculado sobre todas las formas sustantivas y verbales.

Aunque en las fuentes electrónicas consultadas se facilitan los datos finales recogidos por el motor de búsquedas, no faltan dudas acerca de cómo aprovecharlos plenamente en nuestro estudio. En efecto, las anotaciones que se encuentran en el corpus polaco NKJP y las del CdE no son absolutamente paralelas, de lo que surge la necesidad de homologar las etiquetas y los datos que se almacenan bajo ellas. Asimismo, durante el censo, el motor de búsquedas puede rechazar formas que no cumplen con sus requisitos o, por otro lado, aceptar formas que no deberían compartir el lema o la categoría flexiva que se les asigna. El hecho se debe tanto a las imprecisiones gráficas ocurridas durante el procesamiento electrónico de los textos como a algunos problemas técnicos imposibles de superar. A modo de ejemplo, los usos funcionales y léxicos del verbo se confunden y nunca se distinguen en las anotaciones de ambos corpus utilizados. Así pues, los datos que ofrecen las fuentes electrónicas necesitan un repaso crítico, una selección previa y en varios casos una rectificación manual.

La tabla siguiente recoge los datos numéricos concernientes a los índices de nominalidad básico (NOM1) y ampliado (NOM2) en ambas lenguas:

Corpus	Formas flexivas de verbos	Verbos totalidad	Sustantivos ordinarios	Sustantivos totalidad	NOM1	NOM2
NKJP	98.038	133.028	282.561	287.408	2,882	2,161
CdE	87.633	128.386	176.550	176.550	1,375	2,015

Tabla II. Nominalidad NOM1 y NOM2en polaco y en español; corpus NKJP/CdE, muestra equilibrada de un millón de *tokens*.

El análisis llevado a cabo en un corpus manual no presenta mayores dificultades técnicas en comparación con los corpus automáticos. Desde el punto de vista formal, surgen las mismas dudas que durante el trabajo con los motores de búsquedas en los corpus electrónicos; no obstante, la perspectiva humana las supera de manera mucho más natural. Puesto que varios aspectos formales del censo se han comentado en las líneas anteriores, nos limitamos a insertar en las tablas los resultados finales del análisis.

Tipo de lenguaje	Código del corpus	Formas flexivas de verbos	Formas no flexivas de verbos	Sustantivos	NOM1	NOM2
narrativa	C.1	1.228	460	2.225	1,812	1,334
	K.1	1.389	375	2.558	1,842	1,450
textos jurídicos	C.2	564	413	2.686	4,762	2,749
	K.2	572	522	3.978	6,955	3,636

textos científicos	C.3	627	474	2.851	4,547	2,589
	K.3	595	524	3.142	5,280	2,808
prensa	C.4	806	324	2.562	3,179	2,267
	K.4	1.004	289	3.455	3,441	2,672

Tabla III. Nominalidad (NOM1 y NOM2) en polaco y español; corpus comparable de textos funcionales.

3. Analitismo. Datos de los corpus electrónicos y de las muestras manuales

Para calcular los índices analíticos en ambas lenguas, nos hemos atenido a siete tipos de parámetros que manifiestan un menor o mayor paralelismo de uso en ambas lenguas: pronombres, preposiciones, conjunciones, determinantes, marcas de grado analítico, verbos auxiliares y marcas de negación.

Para elaborar nuestras hipótesis cuantitativas concernientes a los índices del analitismo en las dos lenguas, hemos procesado, en primer lugar, los datos provenientes de las fuentes electrónicas mencionadas arriba: *Corpus del Español* (CdE) y *Narodowy Korpus Języka Polskiego* (NKJP).

Parámetro	Frecuencia	Índices de analitismo por mil
total pronombres	39.820	39,820
total preposiciones	119.195	119,195
total conjunciones	74.812	74,812
total determinantes	16.569	16,569
total grado analítico	4.359	4,359
total verbo auxiliar	9.473	9,473
total negación	9.343	9,343
total marcas analíticas	**273.571**	**273,571**

Tabla IV. Analitismo en español; corpus CdE; muestra equilibrada de un millón de *tokens*.

Parámetro	Frecuencia cruda	Índices de analitismo por mil
total pronombres	42.187	34,707
total preposiciones	116.751	96,051
total conjunciones	62.684	51,567
total grado analítico	1.257	1,034
total determinantes	18.256	15,019
total verbo auxiliar	3.002	2,470
total negación	2.299	1,891
total marcas analíticas	**246.436**	**202,742**

Tabla V. Analitismo en polaco, corpus NKJP, muestra equilibrada de un millón de *tokens*.

La extensión de cada muestra de textos manuales funcionales es de 10 mil palabras gráficas. En cada caso se trata de corpus comparables, con la excepción del corpus de narrativa que es de tipo paralelo (la lengua base es principalmente el español). Los cuatro corpus

analizados se caracterizan por la mayor diversificación de géneros posible. El corpus administrativo-jurídico consta de fragmentos de constituciones estatales, leyes parlamentarias, códigos penales, reglamentos internos o estatutos institucionales. El corpus científico abarca artículos académicos de materias variadas: lingüística, medicina, psicología o tecnología. Los textos periodísticos provienen en su totalidad de títulos nacionales o locales de prensa de información (versiones electrónicas). Finalmente, el corpus de narrativa se compone de fragmentos de cuatro novelas en versión bilingüe (texto base y traducción). Los cuatro corpus funcionales abarcan numerosos textos cortos y sus referencias bibliográficas exigen demasiado espacio para poder ser citadas en el presente artículo[10].

Tipo de lenguaje	Lengua del corpus	Índice de analitismo /mil
narrativa	esp	334,1
	pol	284,5
textos jurídico-administrativos	esp	279,9
	pol	200,7
textos científico-académicos	esp	293,6
	pol	194,8
prensa	esp	283,2
	pol	197,4

Tabla VI. Índices de analitismo en polaco y en español; corpus comparable y paralelo de textos funcionales.

4. Conclusiones

Un análisis de varios tipos de corpus manuales y electrónicos, con vistas a establecer las tendencias *estilísticas* de las lenguas lleva a conclusiones, importantes en nuestra opinión, desde el punto de vista de la lingüística contrastiva y tipológica. Las tres hipótesis clave que se desprenden de nuestro estudio son las siguientes:

1. Los índices de nominalidad, que podemos denominar estándar (NOM1, evaluado con formas nominales y verbales plenas) y ampliado (NOM2, calculado con la totalidad de las formas nominales y verbales) suponen una ligera inclinación de la lengua polaca por la tendencia nominal en comparación con el español. No obstante, las cifras obtenidas en los grandes corpus electrónicos y en las muestras manuales de textos funcionales no permiten postular que entre ambas lenguas existen notables diferencias al respecto; en el corpus paralelo de textos de narrativa y en la lengua general los índices son los más bajos y los menos dispersos. Normalmente, las diferencias concernientes a la nominalidad en polaco y en español se mantienen o crecen junto con la especialización del lenguaje, alcanzando los máximos valores y denotando las mayores diferencias en los textos administrativo-jurídicos.

Estas tendencias de ambas lenguas hacen suponer que las disparidades estilísticas existentes entre ellas nunca descansan exclusivamente en causas funcionales o diastráticas. La nominalidad extremadamente alta que caracteriza al lenguaje jurídico y, en menor medida, al lenguaje científico, incluso sugiere que cualquier otra razón pierde validez aquí.

[10] Las citamos en su integridad en Bień (2013: 322-329).

Sin embargo, dadas las notables diferencias estructurales entre las lenguas, las posibles causas de estas tendencias pueden ser variadas.

2. La lengua española es más analítica que la lengua polaca, disparidad que se mantiene tanto en textos generales como en los discursos especializados. Llaman la atención los índices de analitismo en el lenguaje literario, que son bastante más elevados que en los demás corpus funcionales. Contribuye a ello, sobre todo, la alta frecuencia de pronombres personales, marcas naturalmente explotadas en los enunciados de valor deíctico, la gran cantidad de preposiciones y el frecuente uso de los verbos auxiliares temporales, que son también un recurso natural en las partes narrativas de los textos periodísticos. Observadas las tendencias muy firmes de ambas lenguas, se puede suponer que estos índices serán generalmente más elevados en las lenguas románicas que en las lenguas eslavas; no obstante, para comprobarlo es imprescindible llevar a cabo un estudio más global. Nuestro enfoque del analitismo, que sirve como punto de comparación con la tendencia nominal, se justifica por la repartición naturalmente discontinua de las marcas analíticas dentro del sistema lingüístico: algunas de ellas se relacionan exclusivamente con verbos, otras exclusivamente con sustantivos y algunas pueden relacionarse con ambas clases de palabras a la vez.

3. No existe relación directa que pueda expresarse en valores numéricos constantes entre la tendencia analítica de la lengua y su grado de nominalidad. Esta hipótesis, que consideramos la más importante del presente artículo, se aparta de algunas opiniones que aparecen en la literatura lingüística y se basa en los resultados que resumen las dos conclusiones precedentes. El español es una lengua más analítica que el polaco y esta tendencia se mantiene en cualquier tipo de corpus, general o especializado. No obstante, la repartición de los índices numéricos no es paralela con la variación funcional de los lenguajes, es decir, no resulta posible establecer una tendencia constante entre el índice analítico creciente y el grado de especialización de los textos.

Somos conscientes de que algunas de estas hipótesis, siempre mejorables, pueden suscitar polémicas. En un estudio de investigación, nos parece que es una ventaja: las discusiones posteriores servirán para desarrollar algunas cuestiones clave que se han ido esbozando en el presente trabajo. A fin de cuentas, con el presente estudio pretendemos solamente aportar pistas para un estudio más global, que abarque más lenguas y tome en cuenta más criterios de análisis.

Bibliografía

Bień, J. (2010): "Aspectos económicos y redundantes en la oración española", en J. Sueiro Justel et al. (eds.): *Actas del III Congreso Internacional de Lingüística Hispánica (CILHIS): Novos investigadores en linguas e lingüística*. Vigo: Axac, 23-38.

— (2013): *El estilo nominal en español y en polaco*. Lublin: Polihymnia.

CdE = Davies, Mark (2002-): *Corpus del Español: 100 million words, 1200s-1900s*. http://www.corpusdelespanol.org.

Cowgill, W. (1963): "A search for universals in Indo-european diachronic morphology", en J.H. Greenberg (ed.): *Universals of Language*. Cambridge, Mass.: The MIT Press, 114-141.

Danchev, A. (1992): "The evidence for analytic and synthetic developments in English", en M. Rissanen et al. (eds.): *History of Englishes: New methods and interpretations in historical linguistics*. Nueva York: Mouton de Gruyter, 25-41.

Ehret, K.L. (2012): *Analyticity and syntheticity in East African English and British English: a register comparison*. Friburgo: Albert-Ludwigs-Universität. http://www.freidok.uni-freiburg.de/volltexte/5804/pdf/KLEhret_BA_Thesis.pdf.

Gawełko, M. (2001): "Sobre la tendencia analítica de algunas lenguas romances: español, italiano y portugués", *Revista Española de Lingüística*, 31/2, 393-412.

—— (2010): "Sur la tendance analytique du polonais", *Lingua Posnaniensis*, 52, 39-54.

Greenberg, J.H. (1960): "A quantitative approach to the morphological typology of language", *International Journal of American Linguistics*, 26, 178-194.

—— (1968): *Anthropological Linguistics*. Nueva York: Random House.

Moreno Cabrera, J.C. (2003): "Síntesis y análisis en las lenguas: crítica de la tipología morfológica clásica y de algunas de sus aplicaciones sincrónicas y diacrónicas", *Estudios de Lingüística de la Universidad de Alicante*, 17, 465-504.

nkjp = *Narodowy Korpus Języka Polskiego* (*The National Corpus of Polish*). http://nkjp.pl/.

Szmrecsanyi, B. (2009): "Typological Parameters of Intralingual Variability: Grammatical Analyticity Versus Syntheticity in Varieties of English", *Language Variation and Change*, 21, 319-353.

—— (2012): "Analyticity and syntheticity in the history of English", en T. Nevalainen, E.C. Traugott (eds.): *The Oxford Handbook of the History of English*. Nueva York: Oxford University Press, 654-665.