

LINGÜÍSTICA

LINGÜÍSTICA Y MATEMÁTICAS

JOSÉ FERNANDO DOMENE VERDÚ

PUBLICACIONES
UNIVERSIDAD DE ALICANTE

LINGÜÍSTICA Y MATEMÁTICAS

JOSÉ FERNANDO DOMENE VERDÚ

LINGÜÍSTICA Y MATEMÁTICAS

AXIOMATIZACIÓN DE LA TEORÍA GRAMATICAL Y SU
APLICACIÓN A LA TIPOLOGÍA LINGÜÍSTICA

PUBLICACIONES DE LA UNIVERSIDAD DE ALICANTE

Publicaciones de la Universidad de Alicante
Campus de San Vicente s/n
03690 San Vicente del Raspeig
Publicaciones@ua.es
<http://publicaciones.ua.es>
Teléfono: 965903480
Fax: 965909445

© José Fernando Domene Verdú, 2010
© de la presente edición: Universidad de Alicante

ISBN eBook: 978-84-9717-115-1
Depósito legal: XX-xxxx-20xx

Diseño de portada: candela ink.
Corrección de pruebas: M.^a de Andrés Chico
Composición: Cristina Payá (www.ipstudio.es)
Impresión y encuadernación:

Cualquier forma de reproducción, distribución, comunicación pública o transformación de esta obra solo puede ser realizada con la autorización de sus titulares, salvo excepción prevista por la ley. Diríjase a CEDRO (Centro Español de Derechos Reprográficos, www.cedro.org) si necesita fotocopiar o escanear algún fragmento de esta obra.

ÍNDICE

PRÓLOGO.	9
1. INTRODUCCIÓN.	21
2. INTRODUCCIÓN HISTÓRICA Y FUNDAMENTOS DE LA LINGÜÍSTICA MATEMÁTICA	27
3. LA TEORÍA DE CONJUNTOS.	73
4. LA TEORÍA DE CONJUNTOS APLICADA A LA GRAMÁTICA DE LAS LENGUAS	79
5. PRINCIPIOS FUNDAMENTALES DE LA MATEMATIZACIÓN DE LA GRAMÁTICA.	85
5.1. Elemento gramatical	85
5.2. Cambio gramatical.	88
5.3. Grupo gramatical	91
5.4. Elemento gramatical complejo	100
5.5. Clasificaciones gramaticales	105
5.6. Sistema gramatical.	111
5.7. Estructura gramatical.	124
6. DEFINICIONES AXIOMÁTICAS DE LOS CONCEPTOS LINGÜÍSTICOS FUNDAMENTALES	133
6.1. Matemática, fonética y fonología.	133
6.2. Matemática y morfología	144
6.3. Matemática y sintaxis	149
6.4. Matemática y morfosintaxis.	165
6.4.1. Las funciones gramaticales.	165
6.4.2. Los casos.	169
6.4.3. Los cambios funcionales.	180

6.4.4. La clasificación de las funciones gramaticales	184
6.4.5. Los sistemas morfológicos	185
6.4.6. La clasificación de los sistemas morfológicos	191
6.4.7. Los subsistemas morfológicos	192
6.4.8. La estructura morfológica.	199
6.5. El método de análisis morfológico.	206
7. APLICACIÓN DEL MÉTODO MATEMÁTICO A LA GRAMÁTICA DE LAS LENGUAS.	229
7.1. La lingüística matemática	229
7.2. Fonética y fonología.	232
7.3. Morfología	242
7.4. Sintaxis.	245
7.5. Morfosintaxis.	247
7.6. La axiomatización de la morfología de las lenguas	252
7.6.1. Introducción	252
7.6.2. La clasificación de las lenguas	254
7.6.3. Lenguas analíticas o aislantes.	268
7.6.4. Lenguas flexivas externas	268
7.6.5. Lenguas flexivas internas	275
7.6.6. Lenguas aglutinantes con EMV simple	290
7.6.7. Lenguas aglutinantes con EMV compleja	291
7.7. La importancia de los símbolos en la axiomatización de la gramática de las lenguas	313
8. CONCLUSIONES	319
9. SÍMBOLOS UTILIZADOS	337
9.1. Símbolos generales	337
9.2. En la formulación matemática.	339
10. REFERENCIAS BIBLIOGRÁFICAS.	343

PRÓLOGO

En este imponente libro, José Fernando Domene Verdú nos ofrece la propuesta más detallada que conozco de axiomatización de los conceptos fundamentales del metalenguaje gramatical propio de la descripción lingüística contemporánea. Se trata de elaborar definiciones rigurosas de los conceptos y métodos de análisis básicos del metalenguaje especializado utilizado por los gramáticos y lingüistas para describir las lenguas naturales. Esas definiciones se basan en la teoría de conjuntos, empleada habitualmente en los enunciados teóricos de diversas disciplinas, las matemáticas entre ellas.

Es necesario poner de manifiesto en este prólogo que en el libro no se dice en ningún momento que las lenguas naturales y las lenguas formales como la teoría de conjuntos son de la misma naturaleza. De hecho, en algunos pasajes de esta obra, el autor insiste en las evidentes diferencias entre los lenguajes formales y las lenguas naturales y reconoce que no se pueden estudiar y tratar éstas como aquéllos. Cuando se habla de aplicar las matemáticas al análisis de las lenguas naturales, surgen diferentes puntos de vista posibles y numerosos malos entendidos. Por ello, conviene detenerse algo en el propósito de este libro.

En el caso concreto que nos ocupa, no estamos ante la idea de que las lenguas naturales, en su funcionamiento y propiedades, sean iguales o similares a objetos matemáticos o formales, en la línea de Chomsky y Miller¹ o Montague². Estamos, más bien, ante un proceso de matematización del metalenguaje gramatical utilizado para describir las lenguas. En todo estudio científico, los investigadores desarrollan una serie de conceptos y métodos para dar cuenta de las propiedades y el comportamiento del objeto

1. CHOMSKY N. y MILLER G. A., *El análisis formal de los lenguajes naturales*, Madrid, Alberto Corazón (1976).

2. MONTAGUE R., «Gramática Universal», en *Ensayos de Filosofía Formal*, MONTAGUE R., Madrid, Alianza, pp. 158-182.

de estudio. Esos conceptos y métodos no pertenecen a ese objeto de estudio, sino al lenguaje formal propuesto, desarrollado y utilizado para dar cuenta de él. Desde hace mucho tiempo, los gramáticos y lingüistas han utilizado las lenguas naturales con el fin de describir los conceptos y procedimientos utilizados para dar cuenta de las lenguas de manera científica. Esos conceptos y procedimientos son construcciones teóricas ideadas para describir y explicar las propiedades y los comportamientos del objeto de estudio; en este caso, las lenguas naturales. Sin embargo, estas lenguas naturales no son instrumentos adecuados para la actividad teórica; sus conceptos no son fijos y rígidos, sino cambiantes y adaptables y van adquiriendo un sentido u otro a través de su uso. Sin embargo, el gramático y el lingüista necesitan un metalenguaje preciso y riguroso, ya que se trata de realizar una descripción científica de un objeto natural. Por ello, es conveniente sustituir en lo posible la lengua natural por un lenguaje formal lo suficientemente explícito y riguroso que, además, y precisamente por ese carácter formal, no introduzca de forma subrepticia aspectos del *explanandum* en el *explanans*; para evitar la circularidad es fundamental que en los elementos utilizados para la definición de los conceptos no haya de forma más o menos implícita propiedades y procesos que han de ser descritos y explicados, tal como expone acertadamente el autor de esta obra en uno de sus párrafos. La utilización de un lenguaje formal sin contenido empírico alguno nos garantiza precisamente esto. La teoría de conjuntos, aquí recurrida para construir ese metalenguaje formal, es precisamente un lenguaje formal adecuado en las propuestas de un metalenguaje riguroso, preciso y deductivo. No debe olvidarse, por otro lado, que la teoría de conjuntos ha sido empleada también con bastante éxito para describir propiedades sustantivas de las lenguas. Por ejemplo, el uso de técnicas algebraicas para la descripción de la semántica de la cuantificación y la pluralidad en las lenguas naturales ha resultado ser extremadamente fructífera, tal como muestran los numerosos estudios con este enfoque de los que disponemos hoy en día³.

Este libro tiene dos partes principales bien diferenciadas que constituyen el núcleo de su aportación. En el capítulo sexto se establecen axiomáticamente algunos de los conceptos fundamentales del análisis gramatical entre los que se encuentran los conceptos de caso, función gramatical y sistema morfológico. En el capítulo siguiente, se hace una clasificación tipológica de

3. KEENAN E. L. y FALTZ L. M. *Boolean Semantics for Natural Language*, Dordrecht, Reidel, 1998; LANDMAN F., *Structures for Semantics*, Dordrecht, Kluwer, 1985; LINK G., *Algebraic Semantics in Language and Philosophy*, Stanford, CSLI, 1991; WINTER Y. *Flexibility Principles in Boolean Semantics*, Cambridge, The MIT Press, 2001, entre muchos otros.

los sistemas lingüísticos, que supone la más detallada tipología lingüística hecha hasta la fecha desde un punto de vista formal.

Estos dos capítulos, los principales y fundamentales de esta obra, van precedidos de una excelente revisión histórica de la lingüística formalmente orientada, denominada en el libro «lingüística matemática»⁴, de una introducción a la teoría de conjuntos y de un capítulo dedicado a los conceptos fundamentales previos de la axiomatización de la teoría gramatical.

El capítulo de la revisión histórica de la lingüística formal empieza con una descripción del desarrollo del método axiomático de la ciencia en general. El origen del enfoque formal de la lingüística en el siglo xx lo sitúa el autor acertadamente en el *Cours* de Saussure y en las aportaciones de Jakobson, Hjelmslev y Bloomfield y, más recientemente, de Harris, Chomsky y Montague. En este capítulo, se describe detalladamente el ambiente intelectual y tecnológico que hizo posible, entre los años 40 y 50 del siglo pasado, un desarrollo realmente espectacular de los métodos formales en su aplicación al análisis de las lenguas naturales. Las abundantes notas que puede encontrar el lector ofrecen, de forma detenida, toda la información más técnica necesaria para entender ese desarrollo sin que por ello se haga lenta y trabajosa la exposición histórica, que constituye el cuerpo principal del texto del capítulo. Uno de los aspectos interesantes de este libro es la atención realmente oportuna que se le da al cambio lingüístico y a su enfoque formal; ya desde este primer capítulo se toma en cuenta este aspecto crucial del funcionamiento de las lenguas naturales.

El tercer capítulo es una exposición de los conceptos e instrumentos básicos de la teoría de conjuntos, que serán utilizados en los siguientes capítulos del libro.

El capítulo cuarto ofrece definiciones formales de *elemento gramatical*, *clase gramatical*, *categoría gramatical*, *cambio gramatical*, *grupo y subgrupo gramaticales*, *sistema y subsistema gramaticales* y *estructura gramatical*. Ésta es una aportación muy notable que, sin duda, puede clarificar el uso metalingüístico de estos conceptos y, sobre todo, puede contribuir a eliminar la a veces insoportable ambigüedad con la que se utilizan estos términos, expresiones y conceptos cuando no hay detrás una definición precisa formalmente delimitada. Por supuesto, estas definiciones son totalmente formales y no tienen contenido empírico; pero se trata precisamente de esto, ya que

4. Este término es problemático porque se presta a diversas interpretaciones que no siempre son aceptadas por los autores que utilizan lenguajes formalizados para hacer una exposición rigurosa de los elementos gramaticales de las lenguas del mundo, tal como se hace en la presente obra.

lo que se propone es un conjunto de conceptos teóricos con los cuales se pueda llegar a un lenguaje común que facilite la descripción, comparación y evaluación de los análisis de los gramáticos; cosa muy difícil de hacer, si no imposible, cuando no se hacen explícitos los fundamentos teóricos de esos análisis gramaticales que han de evaluarse o compararse.

El quinto capítulo se ocupa de definir los elementos fundamentales de la axiomatización de la teoría gramatical y muestra una sección entera dedicada a los fundamentos metalingüísticos formales del cambio gramatical. Esto es algo realmente poco usual en los tratamientos de la lingüística formal o *lingüística matemática*, tal como la denomina el autor. El resto de secciones trata del grupo gramatical, que obedece a los aspectos formales de las relaciones de contigüidad, del elemento gramatical complejo, que introduce el aspecto asociativo de los planos del significante y del significado desde una perspectiva formal conjuntista, de las categorías gramaticales, concebidas como ciertos conjuntos y subconjuntos de elementos gramaticales, del sistema gramatical y de la estructura gramatical. Con todo ello, se hace una presentación conjuntista de los elementos formales que constituyen un sistema lingüístico en su constitución y en su funcionamiento dinámico. Las propiedades de la linealidad del significante, de las dos caras del signo lingüístico, de las relaciones sintagmáticas y paradigmáticas, de las clases gramaticales y de los cambios gramaticales se reformulan aquí en los términos de los conceptos y procedimientos característicos de la teoría de conjuntos. Todo ello supone alcanzar un nivel de explicitud formal en la determinación de las propiedades esenciales de las lenguas naturales que rara vez es presentado de esta forma tan extensa y detallada en las aportaciones relativas a la teoría lingüística con las que contábamos hasta ahora.

En el capítulo sexto se aplican los conceptos definidos en capítulos anteriores al análisis de los diversos niveles de la gramática de una lengua. Se comienza por la definición de los elementos y procesos fónicos. El autor nos da algunas definiciones formales de los cambios fonológicos más característicos, que constituyen una aportación muy interesante de este trabajo, dado que hasta ahora no se habían propuesto definiciones tan precisas y explícitas de algunos de los principales tipos de cambio fonológico. A continuación, se analizan y explican los conceptos y procesos morfológicos y se realizan nuevas formulaciones del cambio fonológico y fonético, además de la correspondiente al cambio morfológico propiamente dicho. En la sección de sintaxis de este capítulo se definen formalmente las nociones de elemento sintáctico, grupo sintáctico, sistema sintáctico (oración), función gramatical, estructura sintáctica, valencia verbal, satélite verbal y cambio sintáctico. En este capítulo se hace una interpretación psicologista de las reglas generativas

y transformacionales de la gramática chomskiana, que no coincide, en mi opinión, con el *status* teórico que tienen estos elementos en las propuestas generativistas. Las reglas transformacionales se ven como expresión de procesos mentales, pero parece más bien que los procesos transformacionales son meros recursos teóricos utilizados para explicar ciertas propiedades de las estructuras sintácticas. A pesar de que es una interpretación muy frecuente entre los lingüistas, la idea de que, por ejemplo, en una oración interrogativa como *a quién has visto* el hablante desplaza el pronombre interrogativo desde su posición canónica posverbal hacia la posición inicial, en el nudo *COMP*, de la oración, no se corresponde con el carácter teórico del recurso a una transformación de anteposición del objeto directo desde la posición posverbal a la inicial de la oración. En efecto, los generativistas no pretenden describir lo que los hablantes hacen de manera efectiva al hablar, es decir, la actuación lingüística, sino las propiedades formales de las construcciones gramaticales. Lo que pretenden es describir algunos aspectos de la competencia lingüística; para ello, recurren a conceptos teóricos como *desplazamiento*, *huella*, que no son más que denominaciones metafóricas a través de las que se quiere dar cuenta de algunas propiedades formales de las construcciones gramaticales⁵. Que esas metáforas se realicen de modo más o menos literal en la actuación, no es directamente relevante. Es decir, se puede hablar de la existencia de una regla de desplazamiento sin que esto implique necesariamente que el hablante, en su actuación lingüística, desplace un elemento de un lugar a otro. Por ello, la interpretación de que estos procesos teóricos corresponden a operaciones mentales, tal como se dice en esta obra, es una interpretación particular de las propuestas chomskianas que da un paso de forma independiente de esas propuestas y que, por tanto, no se pueden integrar en lo que constituye esa doctrina generativista. Por supuesto, la propuesta aquí hecha a propósito del carácter formal o matemático, tal como se dice en él, de las estructuras sintácticas sí que está en plena consonancia con las propuestas generativistas, en las que, como he dicho

5. «La relación entre los elementos desplazados y la posición original se denomina “movimiento”: se dice que un sintagma “se mueve” de una posición a otra dentro de la estructura. Esto no significa que se produzca un movimiento *real* cuando procesamos la oración; es decir, que el hablante piense en un elemento y después lo mueva mentalmente, incluso si esta implicación equivocada en una época dio origen a una serie de investigaciones psicolingüísticas. El movimiento es una relación abstracta entre dos formas de la oración, que se comportan como *si* se moviera algo. En otras palabras, el movimiento es una relación en la competencia –el conocimiento del lenguaje– y no un proceso de la actuación.» (COOK V. J., y NEWSON M. *Chomsky's Universal Grammar: An Introduction*, Oxford, Blackwell, 2007, p. 33 [cursivas de los autores]).

anteriormente, estos elementos son de carácter claramente teórico y sirven para caracterizar formalmente determinados aspectos y propiedades de la competencia lingüística. Por ello, utilizar la expresión de *proceso mental* a la hora de caracterizar formalmente la generación de las estructuras gramaticales y los procesos transformacionales no es, en mi opinión, adecuado, máxime cuando tenemos en cuenta que el término *generativo* se usa con la idea teórica de una caracterización formal algorítmica explícita, más que con la interpretación más habitual y frecuente de realización efectiva o ejecución mental o material de esos algoritmos formales⁶.

La parte de este capítulo dedicada a los casos gramaticales está muy bien expuesta y organizada; además, nos proporciona una visión de esta categoría muy completa y razonable, que será de mucha utilidad a los lingüistas y gramáticos en general. En esta sección se opta por hacer una definición de caso morfosintáctica en la que se tienen en cuenta las diversas expresiones posibles sintéticas y analíticas de esta categoría, frente a otros enfoques en los que sólo se habla de caso cuando una lengua presenta variaciones flexivas nominales que indican función sintáctica. Dado el carácter tipológico de la parte final de este libro parece que esta opción es la más razonable. Adicionalmente, en esta sección se opta por utilizar el término *caso* tanto como exponente de las relaciones sintácticas como semánticas, lo cual es un uso generalizado desde Fillmore en su *Gramática de los Casos* de los años 60⁷. Esta igualación podría estar justificada para quienes mantienen que hay lenguas en las que no hay funciones sintácticas, sino sólo semánticas tal como hacen Van Valin y LaPolla⁸. Por otro lado, cuando vamos más allá de las funciones sintácticas fundamentales (sujeto, objeto directo e indirecto) pasamos a otras funciones sintácticas que están claramente determinadas semánticamente, tales como *destinatario*, *instrumento*, *localización*, *tiempo*, etc. Por ello, parece que, en última instancia, la consideración de los criterios semánticos a la hora de estudiar los casos es más o menos inevitable. Se habla de *caso verbal* cuando las formas verba-

6. «*Generativo* significa que la descripción de una lengua dada por la gramática de un lingüista es rigurosa y explícita. [...] Una de las trampas más conocidas en las que cae la gente [...] consiste en usar el término *generativo* como sinónimo de *productivo* en vez de como sinónimo de *explícito y formal*.» (COOK V. J. y NEWSON M., *Chomsky's Universal Grammar. An Introduction*, op. cit., p. 35 [cursivas de los autores]).

7. FILLMORE, Ch. «The case for case» en *Universals in Linguistic Theory*, BACH E. y HARMS R. (eds.), Nueva York, Holt, 1968, pp. 1-88.

8. VALIN, VAN R. D. y LAPOLLA R., *Syntax. Structure, Meaning and Function*, Cambridge, Cambridge University Press, 1997.

les expresan los participantes en el suceso y su función en diversas lenguas del mundo. Esta terminología no es la utilizada habitualmente y pudiera ser confuso para gramáticos y no gramáticos. De todas formas, si se diferencia *caso nominal* y *caso verbal*, tal como se hace en el libro, no hay ningún peligro de confusión grave. Hay una sección detenida dedicada al concepto de cambio funcional, que es muy original y novedosa en la bibliografía sobre el cambio lingüístico y que puede servir de punto de partida de posteriores desarrollos de la axiomatización de la teoría del cambio sintáctico, que es uno de los ámbitos de la lingüística diacrónica menos formalizados hasta la fecha, aunque hoy en día contamos con aportaciones realmente sobresalientes como la de Roberts y Roussou⁹.

A continuación se proponen definiciones axiomáticas de *sistema morfológico* y de *paradigma morfológico* así como de los tipos y subtipos de estructura morfológica nocional y funcional, utilizando los conceptos fundamentales de la teoría de conjuntos, tal como se hace en el resto de capítulos y secciones del presente libro. Es una exposición muy detallada que abarca prácticamente la totalidad de parámetros de la estructura morfológica evidentes en las diversas lenguas del mundo, cosa que será de gran utilidad para el capítulo séptimo, en el que se trata la axiomatización de los conceptos fundamentales de la tipología lingüística. La subsección dedicada al análisis morfológico expone una serie de conceptos y principios que constituyen una auténtica metateoría de análisis morfológico, que puede utilizarse para continuar el trabajo metodológico ya clásico de autores como Z. Harris¹⁰. La exposición de los modelos formales de razonamiento inductivo es especialmente relevante desde este punto de vista, dado que no suele aparecer en las presentaciones de los métodos de análisis y descripción lingüísticos más habituales en nuestra bibliografía. A partir de esta exposición se da una serie de reglas y criterios inductivos de análisis morfológico que son, sin la menor duda, una guía muy interesante para la descripción de las lenguas del mundo y, sobre todo, de las menos estudiadas y conocidas (que, por desgracia constituyen la mayoría de las lenguas actualmente habladas en el planeta). Algu-

9. ROBERTS, I. y ROUSSOU A. *Syntactic Change. A minimalist approach to grammaticalization*. Cambridge, Cambridge University Press, 2003.

10. Véase sobre todo Harris Z. *Structural Linguistics*, Chicago, University of Chicago Press, 1960.

nas de estas reglas y criterios son aplicados al caso de la morfología verbal vasca para comprobar su relevancia y utilidad¹¹.

En esta sección del libro se ilustran además las definiciones axiomáticas de las unidades y procesos morfológicos de las lenguas mediante una detallada aplicación al problema de la datación cronológica relativa de algunos de los morfemas del verbo eusquérico, que no es más que un resumen de la tesis doctoral sobre el verbo vasco leída y defendida por el autor y de la cual el libro que estoy prologando es parte introductoria. Esta ilustración pone de manifiesto cómo la formulación explícita de los conceptos y procesos gramaticales hace posible el desarrollo de enfoques heurísticos y predictivos acerca de fenómenos lingüísticos no necesariamente asociados de modo directo a dichos conceptos y procesos, lo cual no es sorprendente, dado que la lengua es un sistema en el que todo está relacionado tal como reza el famoso dicho de A. Meillet¹².

El capítulo séptimo, el otro de los capítulos esenciales de la obra, está dedicado a la definición de los conceptos teóricos básicos de la tipología de las lenguas. En la primera sección se introduce brevemente a los conceptos de la teoría de conjuntos esenciales para entender el resto del capítulo; esto hace posible que un lector interesado en esta cuestión pueda seguir su contenido sin tener que haber estudiado previamente las páginas anteriores. A continuación, se formalizan, utilizando la teoría de conjuntos, los principios básicos de la correspondencia *fono-fonema-letra*. Estos principios se ilustran mediante la exposición de los sistemas fonológicos y ortográficos de algunas lenguas tales como el castellano, el euskera suletino, el euskera roncalés o el euskera estándar. Después se analiza la tipología morfológica y sintáctica. En esta última, mediante la teoría de conjuntos, se analizará la proyección de cadenas sintácticas de elementos léxicos en sus correspondientes estructuras sintácticas y semánticas.

El apartado sexto de este capítulo trata de la axiomatización de la morfosintaxis de las lenguas, lo que equivale a desarrollar una tipología morfo-

11. Este libro ha surgido del prólogo de una tesis doctoral del autor sobre el verbo vasco. Por ello, esta lengua ocupa un lugar prominente en él.

12. La idea de que la lengua es un sistema «où tout se tient» ha sido atribuida en más de una ocasión erróneamente a F. de Saussure. He aquí dos apariciones de esta expresión en dos obras de A. Meillet: «chaque fait linguistique fait partie d'un ensemble où tout se tient» (MEILLET A., *La méthode comparative en linguistique historique*, París, Honoré Champion, 1970, p. 12), «les différentes parties du système linguistique indo-européen forment un ensemble où tout se tient». (MEILLET A., *Introduction à l'étude comparative des langues indo-européennes*, Alabama, University of Alabama Press, 1964, p. IX)

lógica de las lenguas establecida sobre bases formales. Se realiza sobre los fundamentos formales de las definiciones de los elementos y procesos explicados y desarrollados anteriormente. Se empieza haciendo un breve pero denso repaso a la historia de las clasificaciones tipológicas de las lenguas basadas en la morfología, surgidas, como es sabido, en el siglo XIX, en Alemania. En ese momento se origina la famosa clasificación de las lenguas en aislantes, aglutinantes, flexivas o fusionantes, incorporantes y polisintéticas. La clasificación que se define, explica, ejemplifica y desarrolla en esta obra sigue esta tradición decimonónica. De esta manera, el autor distingue los siguientes tipos: aislante o analítico, con subtipos puro (miao), derivativo (tai), compositivo (vietnamita), aglutinante/incorporante (japonés); sintético, con los subtipos flexivo externo (español, italiano) e interno (árabe clásico), aglutinante simple (coreano), complejo (quechua) e incorporante (chucoto); y por último, el tipo polisintético (esquimal).

La clasificación propuesta es muy conservadora y tiene algunos de los inconvenientes de las tipologías clásicas decimonónicas basadas en la morfología. Una primera crítica consiste en la influencia que la escritura tiene en esta clasificación; influencia que, en mi opinión, ha sido subestimada y cuyo estudio crítico puede llevar a cuestionar los fundamentos de esta tipología clásica¹³. Un problema evidente es que se seleccionan de cada lengua unos rasgos y se marginan otros para de esta manera conseguir que todos los tipos y subtipos tengan representantes claros. Esto lo ha visto el propio autor, quien en la nota número 90 observa agudamente que, basándose en algunas formas vascas, podríamos decir que esta lengua es polisintética. Ello demuestra la labilidad de esta clasificación morfológica tradicional. No parece fácilmente justificable que el coreano esté en la clase de lengua sintética aglutinante simple y el japonés esté en la clase aislante aglutinante, cuando las dos lenguas tienen una estructura morfológica y sintáctica muy similar, que hace muy difícil justificar su clasificación en dos tipos diferentes. Por supuesto, aquí encontramos una propuesta global basada en algunos rasgos de las lenguas del mundo, que no puede entrar en todos los detalles de cada una de las lenguas que habrían de considerarse relevantes¹⁴. Hacerlo supondría embarcarse en un trabajo de varios volúmenes que habría de lle-

13. Véase MORENO CABRERA J. C, «Tipología Lingüística y escritura», en *La lengua y su naturaleza dinámica. Homenaje a Ivan Kanchev en su 70 aniversario*, Sofía, pp. 57-67 y «The Written Language Bias in Linguistic Typology», en *Cuadernos de Lingüística XV* IUOG, 2008.

14. El WALS puede ser aquí de gran ayuda: HASPELMATH M., y otros (eds.), *The World Atlas of Language Structures*, Oxford, Oxford University Press, 2005.

var muchos años. De todas maneras, en esta obra se pueden encontrar datos detallados sobre algunas de las lenguas de cada uno de los tipos, que pueden servir, sin duda, de punto de partida para desarrollar y perfeccionar la clasificación propuesta.

En estas comparaciones entre lenguas hay que andar con pies de plomo a la hora de hacer juicios sobre su complejidad o simplicidad. En ocasiones circulan por la bibliografía determinados puntos de vista y afirmaciones sobre las lenguas que pueden inducir al investigador a error. Un ejemplo paradigmático es el de la supuesta simplicidad y primitivismo de la lengua mura-piraha a la que se refiere el autor en la nota número 97 siguiendo alguna de las descripciones sesgadas e interesadas en presentar esta lengua como más primitiva o simple, que han sido muy criticadas por algunos lingüistas¹⁵. El caso contrario lo vemos en los juicios sobre la supuesta gran complejidad de algunos aspectos de determinadas lenguas. Sobre la base de sus sólidos conocimientos sobre la lengua vasca, el autor hace unas afirmaciones sobre el sistema verbal del vasco en la nota número 140 que podrían ser malinterpretadas por algún lector, incluso con formación lingüística. En esa nota se dice que el verbo vasco tiene más de 168000 formas distintas, sin contar las formas pasivas y las numerosas formas dialectales, algunas de las cuales expone el autor en páginas siguientes. Esto podría llevar a pensar a más de uno que esta lengua es de una complejidad enorme y que, por tanto, es una idioma mucho más difícil de aprender y de usar que las lenguas de su entorno; más aún, de ahí podría deducirse que buena parte de la situación sociolingüística actual y pasada de esta lengua está condicionada por esta gran dificultad. Sin embargo, si hacemos una evaluación sistemática de la morfología verbal vasca comparándola, por ejemplo, con la del castellano y hacemos caso omiso de algunos de los sesgos de carácter escriturista, quedándonos sólo con los datos del habla real, encontramos que la complejidad verbal del castellano y del vasco es muy similar y que no está justificada la idea de que el vasco es

15. Como muestra de esta polémica pueden consultarse los siguiente trabajos, que pueden encontrarse en Internet: EVERETT, D., «Cultural Constraints on grammar and cognition in Pirahã» en *Current Anthropology*, 46, 4, 2005, pp. 621-646; «Cultural constraints on grammar in Pirahã: A reply to Nevins, Pesetsky, and Rodrigues» 2007; NEVINS, A. y otros, «Piraha Exceptionality: A Reassessment», 2007. Por otro lado, es palmariamente falso que los mura-pirahã carezcan de mitos y leyendas, como se recoge en esa nota número 97 del libro. Algo similar se puede aducir respecto de lo que se dice en la nota número 150 a propósito de las lenguas makú.

mucho más complejo que el castellano en este aspecto¹⁶, como esgrime con contundencia más de uno.

Como aportación sobresaliente, entre otras a las que ya he hecho referencia, hay que mencionar la precisión y explicitud de las notaciones formales proporcionadas por el autor para realizar comparaciones tipológicas entre las lenguas. Esta regimentación notacional, en mi opinión, es una aportación de primer orden en un campo en el que coexisten numerosas notaciones provenientes de las diferentes tradiciones de la descripción y estudio de las distintas familias a las que pertenecen las lenguas incluidas en las comparaciones tipológicas; ello contribuye a crear malentendidos e interpretaciones fallidas que, muchas veces, vician de raíz los resultados y conclusiones de este tipo de actividad investigadora en el campo de la lingüística general. Esto lo resalta el autor precisamente en el apartado séptimo de este capítulo, donde da ejemplos concretos de lo que acabo de decir referidos al análisis e interpretación de las formas verbales del vasco.

Cierra la publicación con un capítulo de conclusiones en el que se resume y sintetiza todo el aparato axiomático explicado y aplicado a lo largo de toda la obra.

Contra lo que le pudiera parecer al lector al hojear el presente libro, el nivel de formalización de los razonamientos que aquí se llevan a cabo es bajo, dado que todas las demostraciones se realizan a través de explicaciones informales utilizando la lengua usual, sin seguir la rutina demostrativa de carácter algorítmico típica de las obras matemáticas, como, por ejemplo, las que tratan de la axiomática de teoría de conjuntos, cuyo uso de la lengua natural suele estar reducido a un mínimo¹⁷. Ello hace que el libro pueda ser leído fácilmente por cualquier lector que disponga sólo de unos conocimientos elementales de teoría de conjuntos; incluso aquellos que no la conozcan en absoluto pueden proceder a una lectura provechosa haciendo el pequeño esfuerzo que supone asimilar las definiciones de todos los conceptos relevantes proporcionados por el autor. Este lector puede estar seguro de que el autor ha definido todos y cada uno de los conceptos de teoría de conjuntos necesarios para seguir sus propuestas. Por ello, se puede afirmar que el li-

16. En efecto, basta tomar los pronombres proclíticos de las formas verbales del castellano (o de cualquiera de las otras lenguas romances) como constituyentes de la morfología verbal (y hay razones de peso para hacerlo) para convertir la complejidad del verbo castellano similar a la del verbo vasco. Véase la exposición al respecto de de Juan Carlos Moreno Cabrera, «Sobre la complejidad y dificultad de las lenguas. El caso del euskera» en *Revista de Lenguas y Literaturas catalana, gallega y vasca*, UNED, 2008, en prensa.

17. Véase, por ejemplo, MOSTERÍN, J., *Teoría axiomática de conjuntos*, Barcelona, Ariel, 1971.

bro está autocontenido en cuanto a los conocimientos necesarios para leerlo, consultarlo o utilizarlo.

En fin, estamos ante una obra original, útil y necesaria, que es un auténtico hito en la bibliografía del español de lingüística general y que espero que sirva de estímulo para el desarrollo de los estudios de formalización de la teoría gramatical en nuestro ámbito cultural.

Juan Carlos Moreno Cabrera
Catedrático de Lingüística general
Universidad Autónoma de Madrid

1. INTRODUCCIÓN

La aplicación de las matemáticas a las distintas ramas del saber, o lo que es lo mismo, a las distintas partes de la filosofía fue, sin duda, lo que dio origen a la ciencia moderna y lo que ha hecho posible su desarrollo vertiginoso en los últimos siglos. El precedente estuvo en la Grecia helenística con los matemáticos y astrónomos de la Escuela de Alejandría (Pitágoras, Euclides, Arquímedes, etc.), cuyos principios todavía siguen siendo válidos, pero se interrumpió en época romana y en la Edad Media por el protagonismo en la filosofía medieval de Platón y, a partir de Tomás de Aquino, también de Aristóteles, que rechazaba el uso de las matemáticas y la experimentación. En el Renacimiento, se recuperó y difundió la ciencia helenística en Europa, lo que propició la génesis de la teoría heliocéntrica y la aplicación de las Matemáticas –sobre todo de la geometría– y de la experimentación a la Física –principalmente a la astronomía– para demostrar dichas teorías con rigor, certeza y exactitud. El consiguiente enfrentamiento a los partidarios de la filosofía escolástica, que seguían defendiendo la teoría egocéntrica, fue inevitable. Copérnico en el siglo XVI; Kepler, Galileo y Descartes en el XVII; y finalmente Leibniz y Newton entre el XVII y el XVIII, dieron origen a la Física moderna gracias al método experimental y, sobre todo, a la utilización de las Matemáticas. En los siglos XVIII y XIX, su empleo se extendió también a otras ciencias naturales, como la Química, e incluso a algunas ciencias sociales como la Economía o la Sociología (Malthus, etc.).

La axiomatización de las matemáticas y de la lógica a partir del siglo XIX, realizadas por G. Boole y G. Frege principalmente, le confirió a las matemáticas el protagonismo absoluto como ciencia instrumental para ser utilizada en las demás ciencias y, como no, también de la Lingüística. En 1926, L. Bloomfield realizó un intento de matematización de la Lingüística, enunciando una serie de postulados, que no fueron tenidos en cuenta por los lingüistas. En los años 50, y como consecuencia de la axiomatización de los lenguajes formales –de los lenguajes de programación concretamente– se produjo el intento más serio de aplicación de las Matemáticas a la Lingüís-

tica, con Z. S. Harris y N. Chomsky principalmente, y apareció lo que luego se conoció como lingüística matemática, que, debidamente desformalizada por N. Chomsky, se aplicó a las lenguas naturales y dio lugar a la Gramática generativa y transformativa (GGT).

El fracaso de la lingüística matemática no fue obstáculo para el uso de la estadística en la Lingüística, al igual que había ocurrido o estaba ocurriendo en las ciencias sociales. El precedente estuvo en A. A. Markov, que en 1913 publicó un estudio estadístico de la novela *Eugenio Onéguin*. En lingüística histórica, por ejemplo, se utilizó el método léxico-estadístico, usado por J. H. Greenberg (1967) y que A. Tovar (1961) aplicó a la relación de la lengua vasca con diversas lenguas caucásicas (circasiano, avar, georgiano), aunque fue criticado por el mismo Tovar, porque las relaciones estadísticas entre las lenguas no permiten decidir si son debidas al parentesco genealógico, a la presencia de préstamos o, simplemente, al resultado del azar (Tovar, 1961: 258). En la actualidad, la estadística es de uso común en Lingüística, y se va generalizando prácticamente en todas las partes de la Gramática para explicar la variabilidad lingüística. En este sentido, se puede citar el «modelo cuantitativo» utilizado por W. Labov (1982), H. López Morales (1983), etc.

Sin embargo, la matematización de la Gramática según el modelo categorial sigue pendiente y todavía está sin resolver. No hay que olvidar que una ciencia cuyos métodos de demostración pertenecen a la lógica es una ciencia que está formalizada, y que la Matemática es la ciencia formalizada por excelencia. Por eso, se utiliza como ciencia instrumental en otras ciencias, permitiendo utilizar las demostraciones de tipo deductivo; le otorga fiabilidad a las teorías y leyes de esas ciencias, con demostración mediante razonamientos válidos o correctos basados en la lógica; y permite la simplificación de las ciencias matematizadas, por muy complejas que sean, al utilizar símbolos y fórmulas que agilizan los cálculos y procesos. Es evidente, por ello, que la aplicación de las Matemáticas a la Lingüística y la matematización de la Gramática es algo necesario, y que reportaría evidentes ventajas en la investigación lingüística.

Después de considerar algunos aspectos relativos al ejercicio de la investigación lingüística, creímos conveniente tomar una decisión desde el punto de vista metodológico. En efecto, hay que considerar que el número de datos a tener en cuenta supera la capacidad del investigador, y que la abundancia de ellos lleva a la formulación de hipótesis, sobre la base de unos datos pretendidamente representativos del fenómeno estudiado que nunca pueden ser contrastados de manera satisfactoria y fehaciente. Todo ello lleva, a su vez, a la coexistencia de hipótesis, explicaciones, argumentaciones, definiciones diversas sobre un mismo fenómeno. Asimismo, esta situación dificulta la

construcción de un edificio teórico general sobre la organización y funcionamiento de los sistemas lingüísticos en general y de cada lengua en particular. Por ello, después de considerar la inexistencia de un recurso o instrumento teórico-metodológico que proporcionara una base sólida al edificio, al mismo tiempo que posibilitara el tratamiento simplificado de los datos y su rápida consulta, y que finalmente ofreciera garantías de contrastabilidad de todas cuantas hipótesis pudieran formularse, decidimos que una manera de subsanar esta deficiencia radicaba en el empleo de una ciencia que desde hace siglos se ha venido configurando como ciencia auxiliar de algunas disciplinas: la ciencia matemática. No olvidemos que con el empleo o integración de las matemáticas, muchas ciencias han alcanzado su mayoría de edad.

La Lingüística ha considerado la Matemática, en tanto que instrumento de formalización de datos lingüísticos con fines informáticos y estadísticos. Esta matemática aplicada a la Lingüística apenas ha tenido consecuencias de orden teórico y metodológico para ella. No ha sido así en el caso de la llamada *lingüística matemática* que desde los años 40 se fijó el objetivo de la elaboración de modelos matemáticos como aparato formal de una teoría lingüística. De esa forma, la Lingüística, gracias al lenguaje formal de base matemática, se dotó de un metalenguaje unívoco, con coherencia lógica y total explicitud; es decir, de un buen lenguaje científico.

De los distintos modelos matemáticos existentes, la Lingüística adoptó el modelo axiomático, que es de naturaleza algebraica y está en la base de la teoría de las gramáticas generativas de lenguajes formales, como el de la lógica y los lenguajes de la informática, con la intención de aplicarlo al estudio y descripción de la organización y funcionamiento de los sistemas lingüísticos. Lo cierto es que, con el tiempo, el instrumento de descripción formal y sus propiedades se trasladaron al objeto, el sistema descrito. Esto es, los investigadores vieron o intentaron ver en las lenguas las propiedades matemáticas de los lenguajes formales. La investigación ha demostrado que las lenguas naturales, tanto en lo que se refiere a su organización (pero de manera muy especial en su funcionamiento), poco tienen que ver con los lenguajes formales. Éstos son simples (sólo tienen sintaxis), regulares (no presentan excepciones) y unívocos (cada término corresponde a un solo significado y cada significado a un solo término); es decir, responden a la naturaleza del modelo algebraico que los creó mientras que las lenguas constituyen sistemas complejos, irregulares y con un cierto grado de ambigüedad propios de los grupos humanos que las configuraron y utilizan.

Siendo así que la organización y funcionamiento de un sistema lingüístico no obedece las mismas reglas que los lenguajes formales de naturaleza matemática, es obvio que la matemática no debe contemplarse como

un modelo a adoptar, sino «sólo» como un instrumento de fundamentación teórico-metodológica, esto es, como un instrumento que permite el manejo de un gran número de unidades lingüísticas para el análisis así como para la contrastación de hipótesis. En este sentido, cabe recordar que en el ámbito de las matemáticas, la «teoría de conjuntos» está en la base de todo el edificio matemático desde el siglo XIX, y su vocabulario constituye la parte fundamental del lenguaje de las matemáticas modernas. La axiomatización (esto es, el hacer explícitos los términos no definidos y los postulados sobre los que reposa una teoría) del lenguaje matemático, a principios del siglo XX, vino a subsanar las paradojas y antinomias que se derivan de la utilización de la noción intuitiva de conjunto como colección de elementos u objetos que satisfacen una propiedad dada.

En suma, decidí adoptar la teoría de conjuntos como punto de partida para la elaboración de los principios teóricos y metodológicos que sustentan el método de análisis, es decir, la determinación y definición de unidades y propiedades de la morfología de las lenguas. Sobre la base de la teoría de conjuntos he definido los conceptos lingüísticos básicos de análisis morfológico de las lenguas, así como de sus propiedades; esto es, he formado un conjunto de conceptos unívocos, con coherencia lógica y total explicitud. Estos conceptos se pueden expresar por medio de símbolos y fórmulas que simplifican y agilizan la argumentación y la exposición, a la vez que les otorgan fiabilidad, cuando se demuestran mediante razonamientos válidos o correctos basados en la lógica y las matemáticas.

Debido a las dificultades que presenta la aplicación de la lingüística matemática a algunos aspectos de las lenguas, presento a continuación una alternativa basada en la teoría de conjuntos, que es la parte más elemental de las Matemáticas. Para ello, es necesario empezar recordando los conceptos básicos de la teoría de conjuntos y definir a partir de ellos los conceptos básicos de la nueva alternativa de matematización que estarán presentes a lo largo de toda esta parte. A continuación se exponen los principios fundamentales de dicha matematización de la gramática y se definen los conceptos básicos de la misma aplicados a la gramática en general (elemento, cambio, grupo, clases, sistema y estructura gramaticales), para ir aplicándolos a lo largo del capítulo siguiente a cada una de las partes de la gramática (fonética y fonología, morfología, sintaxis y morfosintaxis) y al mismo método de análisis morfológico. Indudablemente, la teoría está hecha para ser aplicada a la realidad y, para ello, en el siguiente capítulo se resume y aplica a diversos ejemplos de la gramática de algunas lenguas como el castellano o el vasco, geográficamente muy cercanas pero muy alejadas tipológica y genéticamente. Dada la importancia de los símbolos y fórmulas en cualquier tipo

de matematización, y por lo tanto también en la de la Gramática, es necesario finalizar con una lista de los símbolos utilizados.

Quiero expresar mi agradecimiento al doctor Joaquín Gorrochategui Churruca y a la doctora María Helena Fernández Prat, por la confianza que depositaron en mí, por sus indicaciones y sugerencias relativas a los fundamentos epistemológicos de la lingüística general y matemática, que han constituido el fundamento para el correcto desarrollo de una idea primigenia, han servido para resolver mis dudas y me han abierto nuevos caminos en la investigación. A mi hermano, Joaquín, le agradezco su ayuda en todas las cuestiones relacionadas con la informática, materia en la que ha demostrado ser un auténtico especialista. Quiero expresar también mi agradecimiento al doctor Juan José Chao Fernández por haber alentado la puesta en práctica de una idea, así como por haber sido el vínculo con los mencionados profesores; al doctor Josep Vicent Guia Marín, por haber revisado y corregido el texto de esta publicación desde el punto de vista de las matemáticas; al doctor Francisco Gimeno Menéndez por todas las atenciones que siempre me ha dispensado, por su apoyo y, sobre todo, por su ayuda en la publicación de esta obra; y al doctor Juan Carlos Moreno Cabrera, por el interés que siempre ha demostrado por el presente trabajo y por haber escrito el prólogo de este libro con toda la atención y sabiduría de quien es un verdadero especialista en la materia. Finalmente, mi agradecimiento a la Universidad de Alicante por haber realizado la presente edición.

2. INTRODUCCIÓN HISTÓRICA Y FUNDAMENTOS DE LA LINGÜÍSTICA MATEMÁTICA

La ciencia se ha caracterizado siempre por fijar previamente las hipótesis que se postulan y verificar después las tesis que de ellas se derivan mediante dos caminos que validan las hipótesis formuladas, la experimentación y la demostración. Esto lo que se conoce como el método científico, que fue seguido ya en época helenística por la Escuela de Alejandría, con Euclides, Apolonio y, especialmente, por Arquímedes en el siglo III a.C. aunque no por Aristóteles, cuya filosofía fue la que se generalizó en la Edad Media. En los siglos XIII y XIV reapareció de la mano de J. Duns Scotus (1266-1308) y Guillermo de Ockham (1300-1349), fue enunciado teóricamente por Francis Bacon (1561-1626) y perfeccionado por John Stuart-Mill (1806-1873).

Sin embargo, en algunas materias la experimentación no fue suficiente para verificar las tesis que se derivaban de las hipótesis formuladas y fue necesaria la demostración, utilizada ya desde la época griega, en la que los conocimientos ya fueron adquiridos por procedimientos empíricos. La demostración, a su vez, necesitó un instrumento para efectuar los razonamientos que derivaban una conclusión de unas premisas. Ese instrumento de demostración fue la lógica clásica y el método deductivo, que fue fundamentado por la llamada Escuela de Atenas, con las contribuciones de Platón y Aristóteles sobre todo, y protagonizó la filosofía aristotélica imperante en la Edad Media. Por otra parte, ya los pitagóricos desarrollaron las Matemáticas y utilizaron la simbolización, asignándole números a las cosas. Los griegos aplicaron también las matemáticas a las ciencias naturales, principalmente a la Astronomía, ideando una compleja teoría de astronomía matemática basada en la construcción de elaborados modelos geométricos que pretendió dar una explicación cuantitativa precisa al esquema de los movimientos planetarios partiendo del supuesto del cosmos geocéntrico. El *Almagesto* de Ptolomeo representa la culminación de esta tradición de astronomía matemática y los modelos y métodos por él empleados configuraron la astronomía posterior. Y

también aplicaron las Matemáticas y la simbolización a las demás ciencias, de manera que otra obra fundamental, *Los Elementos* de Euclides, fue la base de estudio de las ciencias naturales durante toda la Edad Media y hasta la Edad Moderna. Los alquimistas medievales también utilizaban símbolos, al igual que en la cábala y la Astrología. El desinterés de los romanos por las Matemáticas y el protagonismo de la filosofía de Aristóteles, que tras su desarrollo y adaptación al cristianismo por Tomás de Aquino fue la oficial en la Iglesia y las Universidades, impidió el desarrollo científico durante toda la Edad Media en el mundo cristiano porque rehusaba el uso de las Matemáticas y de la experimentación. No ocurrió lo mismo en Oriente, donde los hindúes mantuvieron y desarrollaron la ciencia griega y los árabes la transmitieron a Occidente e inventaron la trigonometría. Ya en el siglo XVI se empezaron a utilizar como ciencia instrumental de las demás ciencias naturales. Y las Matemáticas, complementadas con el método experimental desarrollado por Bacon, les dieron a las demostraciones un rigor y una precisión de la que carecía la filosofía aristotélica. Esto, y el enfrentamiento a los partidarios de la filosofía de Aristóteles, que consideró las Matemáticas inaplicables al mundo físico real y que no se sostenía ante los nuevos métodos científicos, motivó el extraordinario desarrollo de la ciencia a partir del siglo XVI.

La matematización y el consiguiente desarrollo de la ciencia a partir del siglo XVI se puede explicar por cuatro causas históricas muy concretas: el auge del comercio marítimo, los avances técnicos, el Renacimiento y la invención de la imprenta. La conquista de Constantinopla por los turcos en 1453 interrumpió el comercio de la seda y las especias con el Extremo Oriente a través de la ruta por Asia Central y los países europeos se vieron en la necesidad de buscar nuevas rutas comerciales para recuperar este comercio con China y Japón. Estas nuevas rutas necesariamente tenían que ser marítimas, bien bordeando las costas de África, como hicieron los portugueses, o bien aventurándose a través del Atlántico, como hizo Colón al servicio de Castilla. El auge de la navegación para buscar nuevas rutas comerciales, que posibilitó los viajes a China y Japón y del descubrimiento de América en 1492, motivó el fomento de las técnicas de navegación, el desarrollo de la astronomía y el perfeccionamiento de las matemáticas, sobre todo de la aritmética, el álgebra y la trigonometría. A esto contribuyeron también los avances técnicos, sobre todo en el armamento tras la invención de la pólvora y la aparición de las armas de fuego, que requerían un estudio matemático del movimiento de los proyectiles. El Renacimiento despertó el interés por la antigüedad clásica y, por tanto, por los físicos y matemáticos griegos (Pitágoras, Euclides, Arquímedes, Ptolomeo, etc.). Y, por último, la invención de la imprenta por Gutenberg (imprimió la *Biblia Mazarina* en 1456) permitió la difusión de los

conocimientos científicos por toda Europa, de manera que en 1482 apareció el primer libro de matemáticas, que fue la versión latina de *Los Elementos* de Euclides, y a él siguieron en los años sucesivos las obras de los matemáticos de la Escuela de Alejandría, principalmente de Arquímedes.

A consecuencia de ello, las Matemáticas experimentaron un gran desarrollo en el siglo XVI, con Tartaglia, Cardano y Ferrari, en Italia; Vieta, en Francia; y Stiefel y Neper, en Alemania. Lo más significativo para la ciencia fue que en el siglo XVI se comenzaron a utilizar las Matemáticas como ciencia instrumental en las ciencias naturales, principalmente en la mecánica y la Astronomía, o sea, en la Física. Copérnico (1473-1543) utilizó las Matemáticas para realizar cálculos astronómicos y demostrar la teoría heliocéntrica, midiendo la distancia entre el Sol y los planetas conocidos entonces con una exactitud sorprendente. La formuló en 1513 partiendo de la teoría geocéntrica de Ptolomeo y la expuso en el libro titulado *De revolutionibus orbium coelestium*, considerando que los complejos recursos geométricos empleados por él tenían que ser entendidos meramente como medios de cálculo, y no como descripciones de los movimientos reales de los planetas. Sin embargo, no la pudo publicar hasta 1543 y su editor, el clérigo luterano Andreas Osiander (1498-1552), incluyó un prefacio en el que afirmaba que debía ser considerada como un método matemático para facilitar los cálculos astronómicos. Kepler (1571-1630) concibió la estructura de las órbitas planetarias como relaciones matemáticas, y buscó analogías entre ellas y la perfección de las estructuras geométricas, afirmando en su *Astronomia Nova* (1609) que las matemáticas constituyen el lenguaje de Dios y con él la naturaleza se manifiesta como obra del Creador. Galileo (1564-1642) fue quien revolucionó la Astronomía con sus observaciones telescópicas y el primero que utilizó en la Física el método infinitesimal basado en la idea de «paso al límite», cuyo precursor fue también Arquímedes, porque sus investigaciones así se lo exigieron, sobre todo para explicar y defender la teoría heliocéntrica de Copérnico. Galileo fue también el primero que defendió que la naturaleza está escrita en caracteres matemáticos, y actuó en consecuencia, creando una teoría de la naturaleza en la que los problemas físicos fueron concebidos en términos matemáticos. Mientras Aristóteles consideró las Matemáticas inaplicables al mundo físico real y su filosofía era la imperante, Galileo trataba de crear una física que estuviera basada en la matematización. Expuso su teoría en el libro *Sidereus Nuncius* (1610) y en *Il Saggiatore* (1623) afirmó: «La filosofía se halla escrita en el gran libro que está siempre abierto ante nuestros ojos (quiero decir, el universo); pero no podemos entenderlo si antes no aprendemos la lengua y los signos en que está escrito. Este libro está escrito en lenguaje matemático y los símbolos son triángulos, círculos u

otras figuras geométricas, sin cuya ayuda es imposible comprender una sola palabra de él y se anda perdido por un oscuro laberinto» (citado en Mínguez, 1986: 151).

A esta matematización de la física se debió el extraordinario avance que le proporcionó este investigador, así como sus seguidores y los investigadores posteriores (Kepler, Fermat, Pascal y Barrow, maestro de Newton), quienes hicieron importantes aportaciones en este sentido. Descartes (1596-1650) descubrió la geometría analítica a partir de la conjunción del álgebra con la geometría, defendió la aplicación de las Matemáticas a la Filosofía y afirmó que toda realidad estaba compuesta de dos sustancias: el espíritu (caracterizado por el pensamiento) y la materia (caracterizada por su extensión espacial). La rígida separación de materia y espíritu, y el énfasis sobre la materia en movimiento como base de la explicación científica, proporcionó un nuevo marco conceptual para la teorización científica, que tuvo una profunda influencia sobre la ciencia posterior. Fue fundamental para su época porque despejó el camino para el posterior estudio de la materia con independencia del espíritu, de la metafísica y de la idea de Dios, y por tanto de forma experimental y matemática. Y así, fueron Newton (1642-1727) y Leibniz (1646-1716) quienes desarrollaron el cálculo infinitesimal organizando los resultados hasta entonces obtenidos en un cuerpo de doctrina completo y eficaz al incorporar la noción de límite. Este desarrollo del cálculo infinitesimal, en las ramas del cálculo diferencial, descubierto por Leibniz en 1684 y por Newton en 1704, y del cálculo integral, y su aplicación a la física le permitió a Newton consumar la revolución que realizó en la ciencia de su tiempo. Newton fue el primero que separó la ciencia de la filosofía y la consideró independiente de la metafísica, renunciando a conocer los principios metafísicos de las cosas, aunque como creyente no llegó a prescindir del todo de la intervención divina. Defendió que los movimientos de la materia eran descriptibles por leyes matemáticas, pero necesitaban la intervención divina periódica porque, sin ella, el sistema planetario llegaría a desordenarse; la naturaleza se descompondría y necesitaría ser restaurada constantemente. Así, la ley de gravitación universal, que es su aportación más conocida, la demuestra matemáticamente pero la considera como un argumento a favor de la existencia de Dios, igual que Kepler decía que las matemáticas constituyen el lenguaje natural de Dios.

En Química, Dalton (1766-1844) ya utilizaba los símbolos y las Matemáticas en sus estudios sobre el número atómico y el peso atómico de los elementos. Las Matemáticas también se utilizaron en las ciencias sociales desde que el economista y sociólogo inglés Malthus publicó en 1798 su obra *Essay of the Principle of Population*, en la que postuló el principio de que

los medios de subsistencia crecen en progresión aritmética mientras que la población lo hace en progresión geométrica con la finalidad de justificar que la pobreza y las calamidades públicas eran inevitables y sólo la guerra, el hambre y las enfermedades las podían frenar.

Hasta el siglo XIX, las matemáticas fueron consideradas como un simple instrumento para las ciencias naturales y, por ello, fueron desarrolladas por astrónomos, físicos o geodestas, que sólo se preocuparon de las ramas de las matemáticas que consideraron aplicables a los estudios que ellos realizaban en el campo de las ciencias naturales. Por ello, a estos científicos, que sólo accidentalmente eran matemáticos, les preocupaba poco el rigor de sus teorías y les bastaba con su eficacia. En el siglo XIX se consideró la necesidad de dotar a las teorías científicas de un rigor que hasta entonces no tenían. En efecto, las conclusiones de cualquier razonamiento deductivo se obtenían como resultado de unas premisas que se suponían ciertas, pero hasta entonces nadie se preocupaba en demostrar la certeza de esas premisas iniciales, a pesar de que de esa certeza dependía la validez de la conclusión final. Todo empezó al fracasar los intentos de demostrar el postulado de Euclides y comprobarse que éste podía tener contradicciones. Como consecuencia de ello, se originó en matemáticas el movimiento axiomatizador, según el cual era necesario fijar previamente y de forma explícita los axiomas y las reglas para deducir cualquier teoría. El método axiomático fue utilizado ya por Euclides en Matemáticas y Aristóteles en Lógica y la formalización del lenguaje matemático y de la Lógica y las Matemáticas ya la realizaron Vieta y Leibniz en la Edad Moderna. No obstante, tras un primer intento de Riemann, fue David Hilbert (1862-1943) el creador en 1900 del método axiomático para las Matemáticas, según el cual, un sistema de axiomas de una rama de las Matemáticas debía estar formado por axiomas que fueran independientes (no deducibles unos de otros) y consistentes (carentes de contradicciones), además de ser completo (que se pudiera demostrar la verdad o la falsedad de todos los enunciados y sus consecuencias), aunque la completitud fue cuestionada por Kurt Gödel en 1931.

En el proceso de axiomatización de las Matemáticas fue imprescindible el desarrollo de la lógica simbólica, que fue iniciada por George Boole en 1854, creando al mismo tiempo un lenguaje formal en el que escribir la Lógica, o sea, formalizándola. La primera axiomatización absolutamente formalizada de la lógica elemental fue realizada por el lógico alemán Frege en 1879, en su obra *Begriffsschrift* (Conceptografía), y después se confeccionaron otros sistemas axiomáticos, como el de la aritmética elemental, elaborado por Giuseppe Peano en 1889, y el de la teoría de conjuntos. Entre los sistemas axiomáticos más antiguos de la lógica elemental, los principales son el

de Whitehead-Russell o sistema WR (1910), que J. Nicod logró reducir a un único axioma en 1917, el de Lukasiewicz o sistema C-N (1929), basado en la notación de los conectores mediante letras mayúsculas del alfabeto y el de Hilbert-Bernays o sistema HB (1934). Peano consiguió deducir la matemática de su tiempo con diez premisas de deducción y otras diez de naturaleza lógica general y su trabajo fue mejorado notablemente por Bertrand Russell y Whitehead en 1908, en cuya obra *Principia Mathematica* desarrollaron un método de trabajo que trataba de probar que toda la matemática se reduce a la lógica y que se conoció como «formalismo» o «logicismo matemático».

El método axiomático se aplicó también en otras ciencias, como en la aritmética por Giuseppe Peano en 1889, en la mecánica general por Hermes en 1938, en la genética por Henry Woodger y en la teoría económica por Von Neumann y Morgenstern en 1948, con lo que la matematización y la axiomatización se extendieron también a las ciencias sociales. En efecto, las ecuaciones diferenciales se aplican en cálculos de crecimiento de una población biológica, por ejemplo, y una serie de técnicas matemáticas (teoría de juegos, etc.) que se conocen con el nombre de investigación operativa se han utilizado para la toma de decisiones. La Sociología también incorporó la teoría de sistemas y la teoría de juegos para desarrollar modelos, ya a partir de los años 60.

Los axiomas son, por tanto, las proposiciones que fijan las reglas del juego para la demostración de los teoremas. Esto no fue suficiente para su conocimiento profundo porque, para ello, es necesario conocer también los elementos con los que se trabaja, las causas de elección de cada camino de demostración y lo que sugirió la conjetura de la que nació el teorema. Esto se consiguió mediante la determinación de las estructuras básicas de las matemáticas, que son tres: las algebraicas (basadas en las leyes de composición), las de orden (basadas en las relaciones binarias de orden) y las topológicas (basadas en las nociones de límite y continuidad). Fue el paso siguiente a la axiomatización de las Matemáticas y recibió el nombre de método axiomático estructural. Apareció en los años 20 e influyó en todas las demás ciencias y áreas del conocimiento, añadiendo el adjetivo «estructural». A principios del siglo xx nació la gramática estructural: el *Curso de Lingüística general* de Saussure se publicó en 1916 y, por influencia de él, se fundó en 1926 el Círculo Lingüístico de Praga (con Trubetzky y Jakobson), que originó el estructuralismo europeo, y en 1933 Leonard Bloomfield publicó su libro *Language*, que fue el inicio del estructuralismo norteamericano. El mismo Bloomfield intentó aplicar por primera vez a la Lingüística el método axiomático de las matemáticas en «A set of Postulates for the Science of Language» (1926: 153-164). También Louis Hjelmslev, en su obra *Prolegomena to a Theory of Language* (1963), hizo una propuesta de formalización

matemática basada en el álgebra, que formaba parte de la teoría lingüística que denominó Glosemática¹⁸.

En lo que se refiere a la utilización directa en Lingüística de las Matemáticas, V. I. Buniakovski ya había indicado en 1847 la posibilidad de utilizar las Matemáticas en Lingüística, lo mismo que hizo Saussure en 1894 al decir que «les quantités du langage et leurs rapports sont régulièrement exprimables, dans leur nature fondamentale, par de formules mathématiques». Después precisó más esta idea diciendo acertadamente que «L'expression simple sera algébrique ou elle ne sera pas» (citado por Marcus-Nicolau-Steti, 1978: 52-53). Algo más de medio siglo después, la rama de las Matemáticas que intervino de forma determinante en el origen de la lingüística matemática fue precisamente el álgebra. Sin embargo, la rama de las Matemáticas que primero se utilizó en Lingüística fue la estadística, sin duda porque es la parte más fácil de utilizar en cualquier disciplina, siempre que existan elementos cuantitativos que puedan contarse o, lo que es lo mismo, que sean enumerables. Ciertamente, la primera vez que se aplicaron las Matemáticas a la Lingüística fue en un estudio estadístico sobre la frecuencia de la aparición de las letras en la novela *Eugenio Onegin*, publicado en 1913 por A. A. Markov con el título *Essai d'un recherche statistique sur le texte du roman Eugène Onéguin*. Un trabajo análogo lo realizó C. Shannon sobre la lengua inglesa, publicándolo en 1951 con el título *Prediction and Entropy in Printed*

18. «El uso de los símbolos glosemáticos de las distintas funciones pueden ilustrarlo con los siguientes ejemplos, en los que a y b representan cualesquiera terminales, v una función variable y c un terminal constante: FUNCIÓN: $a \phi b$; RELACIÓN: $a R b$; CORRELACIÓN: $a : b$; DETERMINACIÓN: $v \Rightarrow c$ ó $c \Leftarrow v$; SELECCIÓN: $v \rightarrow c$ ó $c \leftarrow v$; ESPECIFICACIÓN: $v \vdash c$ ó $c \vDash v$; INTERDEPENDENCIA: $c \leftrightarrow c$; SOLIDARIDAD: $c \sim c$; COMPLEMENTARIEDAD: $c \nmid c$; CONSTELACIÓN: $v | v$; COMBINACIÓN: $v - v$; AUTONOMÍA: $c \dagger v$. El número de terminales, por supuesto, no se reduce a dos» (Hjelmslev, 1963: 64). Los tipos de funciones las resume en la misma página, en la siguiente tabla:

función		relación (conexión)	correlación (equivalencia)
cohesión reciprocidad	}	determinación	selección
		interdependencia	solidaridad
		constelación	combinación
		especificación	complementariedad
		autonomía	

Dentro de la glosemática, Hjelmslev definió un lenguaje diciendo que «A LANGUAGE (symbol: $Ly^o g^o S$) is the Paradigmatic of a Denotative Semiotic whose Paradigms are Manifested by all Purports», <<http://www.glossematica.net>>, def. 38, por ejemplo. Esto es una formalización, pero no una axiomatización, porque carece del aparato matemático que le sirve de fundamento.

English; en la actualidad la utilización de la estadística está muy extendida en los estudios lingüísticos.

Se trata, pues, de estudios estadísticos de textos literarios y consisten en el uso de la estadística en la investigación lingüística como ciencia instrumental, igual que ocurre prácticamente en el resto de ciencias y en todas las disciplinas que estudian elementos o fenómenos enumerables. Los aspectos lógico-matemáticos de los lenguajes formales, en cambio, han llegado a afectar a la esencia misma de la Lingüística, que concibe la gramática de las lenguas naturales y el lenguaje como algo susceptible de ser matematizado, axiomatizado o, si se quiere con más imprecisión terminológica, formalizado. Por ello, supone un paso mucho mayor y más trascendental que cualquier otro avance producido en la lingüística teórica porque, mientras que todos ellos se produjeron como consecuencia de los avances en la investigación lingüística y siempre dentro de la Lingüística misma, la aparición de las matemáticas en el ámbito de las lenguas naturales trascendió el ámbito de la Lingüística y supuso el uso de una metodología totalmente ajena a ella. No obstante, con el tiempo se produjo el ensamblaje teórico de ambas que cristalizó en una teoría lingüística concreta, la GGT.

Por lo tanto, en un momento en que el ámbito de la lingüística general estaba dominado por la lingüística estructuralista, aparecieron los primeros ordenadores y, con ellos, los lenguajes de programación capaces de hacerlos funcionar. Al mismo tiempo, apareció una corriente, la lingüística matemática, que fusionó ambas tendencias gracias a la influencia de los lenguajes informáticos en la lingüística estructural. La lingüística matemática, por lo tanto, fue el resultado de la influencia de los ordenadores y los lenguajes de programación en la lingüística estructural imperante en aquel momento histórico. A su vez, los lenguajes informáticos son lenguajes formales y están basados en el lenguaje formal de la lógica, creado a finales del siglo XIX y principios del XX como consecuencia del proceso de axiomatización de la lógica clásica y de las matemáticas en general y está en el origen de la lógica simbólica o matemática. Por su parte, las gramáticas generativas formales se utilizaron como modelo para describir las lenguas naturales siguiendo una corriente de pensamiento procedente del campo de la Lógica y, por ello, la lógica simbólica axiomatizada está en la base de la axiomatización de los lenguajes formales empleados por la lingüística matemática como modelo de descripción de la organización y funcionamiento de las lenguas naturales¹⁹.

19. Hay lingüistas que piensan y piensaron que las lógicas no clásicas (lógicas formales, lógica simbólica, lógica matemática) «ofrecen la posibilidad de formalizar aspectos o usos

Sin embargo, la aplicación de la Lógica a la Lingüística no es algo nuevo, sino que continúa una tradición que se remonta a la lógica clásica y a la Edad Moderna (Leibniz, Locke, la *Grammaire Générale et Raisonnée* o *Gramática de Port-Royal*, bajo la influencia de Descartes, etc.) y está basada esencialmente en la concepción del lenguaje formal de la lógica como universal y subyacente de todas las lenguas. Esta idea sigue presente y cobra más fuerza entre los lógicos de finales del XIX y principios del XX, como G. Frege, B. Russell, L. Wittgenstein y R. Carnap, que consideran que sólo la lógica simbólica puede aclarar la naturaleza de determinados problemas planteados por el lenguaje. El matemático David Hilbert expresó en 1918 la necesidad indiscutible de que todas las ciencias evolucionaran hacia la axiomatización (Hilbert, 1918: 415). Y esta aplicación de la lógica simbólica, y especialmente del método axiomático, a la Lingüística influyó también en los lingüistas estructuralistas en fechas muy tempranas. Sólo ocho años después de esta afirmación de Hilbert, en 1926, fue nada menos que Bloomfield quien realizó el primer intento de formular un sistema axiomático para la Lingüística, siete años antes de su obra fundamental *Language* (1933), y lo publicó en un artículo titulado «A set of postulates for the science of language» (Bloomfield, 1926: 153-156). En él presentó 77 axiomas y definiciones, pero éstas no cumplen las condiciones que se exigen dentro de un sistema axiomático y, además, el sistema carece de las reglas deductivas para la derivación de frases, por lo que no se puede considerar como un sistema axiomático en sentido estrictamente matemático. Tampoco lo es el sistema formulado por K. Bühler en 1933 ni los postulados de B. Bloch para el análisis fonológico en 1948. Sin embargo, todos ellos indican la preocupación de los lingüistas por aplicar la lógica a la Lingüística ya dentro del estructuralismo. Es en los años 50 cuando se construyen para las lenguas naturales los primeros sistemas axiomáticos que cumplen estrictamente los presupuestos matemáticos, y ello

del lenguaje que antes parecían escapar a la lógica, mostrando la capacidad que la lógica formal tiene de ir estando a la altura que la complejidad del lenguaje exige. El aprovechamiento que de esas lógicas están haciendo los lingüistas –es decir, los científicos *naturales* del lenguaje– es la mejor confirmación de su utilidad. Lejos de pensar que el lenguaje ordinario, por su complejidad, exige estadios “informales”, asistemáticos, en los que la lógica formal desempeñe a lo sumo un papel de ideal de precisión inalcanzable, los lingüistas actuales creen que, precisamente por su complejidad, el lenguaje ordinario demanda una *lógica formal*, una reconstrucción *formal* cada vez más refinada, como la que permiten las lógicas no-clásicas» (Deaño, 1980: 321). Así, «también conviene tener en cuenta las importantes aplicaciones de la lógica simbólica a la lingüística (N. Chomsky y R. Montague) y a la informática, particularmente el desarrollo de la inteligencia artificial (A. Newell, H. Simon)» (Garrido, 1991: 26).

se debió naturalmente a las investigaciones sobre la traducción automática por ordenador (Rohrer, 1978: 24).

En esta aplicación de la lógica simbólica a la traducción automática, y por extensión a la lingüística, ejercieron una influencia decisiva los principales teóricos de la lógica simbólica, quienes defendieron abiertamente que «los diversos cálculos lógicos pueden entenderse como modelo para las lenguas naturales» (Rohrer, 1978: 20). Así, R. Carnap defendió en 1937 que de la mejor manera que se puede estudiar la sintaxis de una lengua para la traducción automática por ordenador es comparándola con un lenguaje artificial²⁰. En esta misma obra había definido previamente la sintaxis lógica de un lenguaje formal como la exposición sistemática de las reglas formales que gobiernan las expresiones lingüísticas de un lenguaje²¹. Como se ve, Carnap ya utilizó en 1937 el concepto de reglas que gobiernan la sintaxis lógica de un lenguaje («formal rules which govern it»), que toma sin duda de las reglas de reescritura de los sistemas axiomáticos de la lógica y es el que caracteriza una gramática generativa. Siguiendo esta misma línea, «E. Post pedía en 1943 que las reglas de formación de un cálculo lógico se equiparasen formalmente a las reglas de deducción» (Rohrer, 1978: 25), lo que permitió la aplicación a las lenguas naturales de la teoría de las funciones recursivas, la consideración de éstas como sistemas combinatorios y el consiguiente desarrollo de la lingüística matemática, también llamada combinatoria o algebraica.

Después se prosiguió en la equiparación de las lenguas naturales a los lenguajes formales y en 1950 «P. Rosenbloom cree igualmente que no existe ninguna diferencia de principio entre los lenguajes simbólicos de los lógicos y los lenguajes naturales» (Rohrer, 1978: 21) y afirma que la única diferencia entre ambos es la mayor complejidad, la imprecisión y la indefinición de las reglas sintácticas de las lenguas naturales²². Y, además, A. Church en 1951 llegó «a suponer que los resultados que se obtienen en la construcción de cálculos se pueden aplicar sin modificación a las lenguas naturales» (Rohrer,

20. «The syntactical property of a particular word-language, such as English, or of particular classes of word-languages, or of a particular sub-language of a word-language, is best represented and investigated by comparison with a constructed language which serves as a system of reference» (Carnap, 1937: 8).

21. «By the *logical syntax* of a language, we mean the formal theory of the linguistic forms of that language –the systematic statement of the formal rules which govern it» (Carnap, 1937: 1).

22. «Words are also symbols, and that ordinary language merely differs from these other symbolic languages in that its rules of syntax are very complicated and never stated precisely and explicitly (Rosenbloom, 1950: 96).

1978: 21)²³. Y Bar-Hillel propuso en 1954 seguir desarrollando en Lingüística el método de Carnap (Bar-Hillel, 1954: 230-237) y el mismo Carnap, que ya había descrito la sintaxis de la lógica, desarrolló al mismo tiempo también la semántica de la misma, y ello lo expresa en 1962 diciendo que «we need, in addition to a purely formal analysis of language, an analysis of the signifying function of language – in other words, a theory of meaning and interpretation» (Carnap, 1962: IX).

Rohrer equipara la sintaxis a la semántica porque «con el mismo derecho con que incluimos en la sintaxis las reglas de formación que combinan rasgos semánticos, podemos incluirlas también en la semántica. En estas reglas no existe ni una referencia al hablante ni a lo designado por las expresiones. Son puramente formales y, según Carnap, pertenecen por ello a la sintaxis. Por otra parte, estas reglas unen rasgos semánticos y combinaciones de rasgos, por lo cual podría considerárselas totalmente como reglas semánticas y llamar semántica generativa a esta parte de la formación» (Rohrer, 1978: 17). Sintaxis y semántica se entendieron, pues, equiparables²⁴.

Siguiendo a Carnap, la lógica simbólica y la teoría semántica de la Lógica se aplicó directamente a la Lingüística en los años 70 de la mano de Richard Montague y de Patrick Suppes (Moreno Cabrera, 1985), que en los últimos años han intentado «hacer encajar una teoría sintáctica con una teoría semántica de las lenguas naturales» (Acero-Bustos-Quesada, 1989: 145),

23. «Although all the foregoing account has been concerned with the case of a formalized language, I would go on to say that in my opinion there is no difference in principle between this case and that one of the natural languages» (Church, 1951: 100-112). Y después lo expresa más o menos en los mismos términos que Rosenbloom: «The difference of a formalized language from a natural language lies not in any matter of principle, but in the degree of completeness that has been attained in the laying down of explicit syntactical and semantical rules and the extent to which vaguenesses and uncertainties have been removed from them» (Church, 1951: 100-112).

24. «Nuestro modelo de gramática tiene así la siguiente constitución. Contiene una semántica generativa que corresponde, sobre todo, a la parte de formación de un cálculo de predicados. Las unidades léxicas no se clasifican en sustantivos, adjetivos o verbos. En su lugar tenemos una separación de predicados y variables. Los predicados se subclasifican, a su vez, según el número de sus posibles argumentos. A esto se añaden cuantificadores, que posiblemente se pueden considerar, sin embargo, como predicados, sin que formen necesariamente una clase separada. Por otra parte, a los componentes semánticos de la base se añaden los axiomas usuales del cálculo de enunciados, axiomas para los cuantificadores, axiomas de la lógica modal y, posiblemente, otros axiomas cuya función no se ha establecido hasta ahora con claridad. Con ayuda de las reglas de deducción se pueden determinar las relaciones de equivalencia y consecuencia entre fórmulas generadas por la estructura profunda. Un sistema de reglas transforma las fórmulas de predicados lógicos en estructuras superficiales bien formadas.» (Rohrer, 1978: 17)

aplicando directamente a la Lingüística la relación entre sintaxis y semántica existente en la lógica de enunciados o lógica de primer orden²⁵. La fundamentación matemática de Montague se basa en un homomorfismo entre un álgebra sintáctica y un álgebra semántica fundada en los valores de verdad (V y F)²⁶, pero la de Suppes es más sencilla y consiste en «asociar con cada

25. Así, «Richard Montague was the first to seriously propose and defend the thesis that the relation between syntax and semantics in a natural language such as English could be viewed as not essentially different from the relation between syntax and semantics in a formal language of first order logic» (Partee-Meulen-Wall, 1990: 317).

26. Desde que Boole demostró en 1847 la relación entre la Lógica y el álgebra, las funciones de verdad y los valores de verdad de las proposiciones lógicas se pueden expresar también mediante esta parte de las Matemáticas. Así, el conjunto $A = \{\text{Verdadero, Falso}\} = \{V, F\} = \{0, 1\}$ formado por los dos valores de verdad es un álgebra de Boole binaria, que es un álgebra de Boole (A, \vee, \wedge) cuyo conjunto $A = \{0, 1\}$ está formado por dos únicos elementos, el 0 y el 1. Un álgebra de Boole (A, \vee, \wedge) , a su vez, es un cuerpo cuyas dos leyes de composición interna son la disyunción no exclusiva (\vee) y la conjunción (\wedge) o, lo que es lo mismo, un conjunto A dotado de dos l.c.i. \vee y \wedge se dice que es un álgebra de Boole si se verifica que las dos leyes son asociativas y conmutativas, cada una de ellas es distributiva respecto a la otra, tiene elemento neutro y todo elemento $a \in A$ tiene un elemento complementario $a' \in A$ y sólo uno. Un álgebra A es un conjunto A en el que se define una o más leyes de composición f_1, f_2, \dots, f_n y se representa por $A = \langle A, f_1, f_2, \dots, f_n \rangle$.

Un álgebra $A = \langle A, f_1, f_2, \dots, f_n \rangle$ se puede relacionar con otra álgebra $B = \langle B, g_1, g_2, \dots, g_n \rangle$ mediante una correspondencia $F: A \rightarrow B$ definida entre ellas: «Relations between algebras may be described by functions mapping one algebra in another; $F: A \rightarrow B$ » (Partee-Meulen-Wall, 1990: 253). Si esa correspondencia entre dos álgebras es una aplicación (o sea, que todo elemento de A tiene imagen y sólo una imagen), se denomina un morfismo (u homomorfismo), que se define como toda aplicación $f: A \rightarrow B$ entre dos álgebras (A, \star) y (B, \square) tal que $f(a \star b) = f(a) \square f(b), \forall a \in A, \forall b \in B$. Si la aplicación es inyectiva (algún elemento de B no tiene antiimagen en A pero los que tienen antiimagen sólo tienen una) se llama monomorfismo, si es suprayectiva (todo elemento de B tiene antiimagen en A y alguno puede tener más de una) es un epimorfismo (u homomorfismo) y si es biyectiva (todo elemento de B tiene antiimagen y sólo una antiimagen en A), es un isomorfismo.

El componente semántico de la lógica también se puede definir matemáticamente como un álgebra $B = (B, \square)$, en la que el conjunto $B = \{0, 1\}$ que la define es el de los valores de verdad de las proposiciones, verdadero (1) y falso (0). La relación entre el componente sintáctico, que es un álgebra $A = (A, \star)$, y el componente semántico de la Lógica, que es otro álgebra $B = (B, \square)$, se puede definir también matemáticamente por medio de la relación entre las dos álgebras. Y de la misma manera que el conjunto de los valores de verdad de una proposición se relaciona con el conjunto de los valores de verdad de su proposición equivalente, definiendo una correspondencia denominada función de verdad, el componente sintáctico y el componente semántico también definen entre sí una correspondencia y, en este caso concreto, una aplicación entre las dos álgebras (A, \star) y (B, \square) , que es lo que se conoce como un homomorfismo, que expresa algebraicamente el concepto lógico de *interpretación semántica*, utilizado ya por Tarski en su teoría de la verdad.

Este homomorfismo $f: A \rightarrow B$ entre dos álgebras $A = (A, \star)$ y $B = (B, \square)$ que define el concepto de la interpretación semántica o relación entre la sintaxis y la semántica de

regla de producción de una gramática generativa un conjunto de reglas o funciones semánticas» (Acero-Bustos-Quesada, 1989: 145)²⁷.

los lenguajes formales, representadas respectivamente por esas mismas dos álgebras $\mathbf{A} = (A, \star)$ y $\mathbf{B} = (B, \square)$, se conoce como el principio de composicionalidad (*The Principle of Compositionality*), en virtud del cual el significado de un enunciado es una aplicación definida entre el significado de sus partes y el conjunto de las reglas sintácticas por las cuales éstas se combinan: «The meaning of a complex expression is a function of the meaning of its parts and of the syntactic rules by which they are combined» (Partee-Meulen-Wall, 1990: 318). Dicho principio no es más que la aplicación de la definición de homomorfismo a la relación entre la semántica y la sintaxis de un lenguaje formal: $f(a \star b) = f(a) \square f(b)$, $\forall a \in A, \forall b \in B$. Fue el americano Richard Montague el que desarrolló el principio de composicionalidad considerando a la sintaxis y a la semántica de la Lógica como álgebras y el que aplicó el concepto matemático de homomorfismo a la relación entre el álgebra de la sintaxis y el álgebra de la semántica.

Para Montague, el concepto de interpretación semántica, que ya fue utilizado por Tarski en su teoría de la verdad y pertenece evidentemente a la Lógica, es un homomorfismo definido entre dos álgebras, una sintáctica \mathbf{A} y otra semántica \mathbf{B} .

El *álgebra sintáctica* es $\mathbf{A} = \langle A, F_0, \dots, F_4 \rangle$, tal que el conjunto A está formado por todas las proposiciones lógicas moleculares bien formadas a partir de un conjunto $X_0 = \{p, q, r, p', q', \dots\}$ de las proposiciones atómicas, y las operaciones sintácticas F_0, \dots, F_4 son las relaciones entre esas proposiciones atómicas definidas por los cinco conectores lógicos ($\neg, \wedge, \vee, \rightarrow, \vdots$) de tal manera que, siendo α y β dos proposiciones lógicas cualesquiera, dichas operaciones vendrían definidas así: $F_0(\alpha) = \neg\alpha$, $F_1(\alpha, \beta) = \alpha \wedge \beta$, $F_2(\alpha, \beta) = \alpha \vee \beta$, $F_3(\alpha, \beta) = \alpha \rightarrow \beta$, $F_4(\alpha, \beta) = \alpha : \beta$. Evidentemente, el conjunto X_0 no es otro que el V (alfabeto o vocabulario) del sistema axiomático definido anteriormente y el A , no es otro que el F (lenguaje formal) de ese mismo sistema axiomático (Serrano, 1977: 188-189).

El *álgebra semántica* es $\mathbf{B} = \langle B, G_0, \dots, G_4 \rangle$, tal que el conjunto está formado por los dos valores de verdad, $B = \{0, 1\}$, que pueden tener las proposiciones lógicas y las operaciones semánticas G_0, \dots, G_4 son las que corresponden a las operaciones sintácticas F_0, \dots, F_4 . La interpretación semántica asigna un valor de verdad a cada proposición atómica y valores de verdad a las proposiciones moleculares, de tal forma que establece una relación entre las proposiciones del conjunto A y los dos valores de verdad del conjunto $B = \{0, 1\}$ que se puede definir como una aplicación entre los dos conjuntos y entre todas las operaciones de cada uno de ellos, o sea, entre las dos álgebras \mathbf{A} y \mathbf{B} . Y una aplicación entre dos álgebras es un homomorfismo, por lo que el concepto de interpretación semántica de las proposiciones lógicas se puede definir como un homomorfismo entre el álgebra sintáctica \mathbf{A} y el álgebra semántica \mathbf{B} .

La semántica de la Lógica y el principio de composicionalidad vienen a completar la definición de gramática formal, que se refería sólo al componente sintáctico, y le aporta el componente semántico que, como se ha podido comprobar, está basado en las funciones de verdad de las proposiciones lógicas. Por lo tanto, la semántica de la Lógica y el principio de composicionalidad completan la axiomatización de los lenguajes formales, que es la esencia de la lingüística matemática.

27. Así, por ejemplo, en la oración típica de la lógica aristotélica «todo hombre es mortal», la extensión semántica de «hombre» se refiere al conjunto de todos los hombres y se representa por la letra H , la palabra «mortal» se refiere al conjunto de todos los seres mortales y se representa por la letra M , y la relación entre el significado de ambas es la función semántica de inclusión representada por la fórmula lógica $H \subset M$. Esta teoría se basa en

La Lógica, por tanto, se aplicó claramente a la Lingüística, sobre todo a partir de la década de los años 50²⁸. La gramática de los lenguajes formales se tomó como modelo para la de las lenguas naturales, entendiendo como modelo «el proceso por el cual, de las informaciones extraídas a partir de la investigación sobre A y de la correspondencia entre A y B, se obtienen ciertas informaciones sobre B», siendo A un objeto o fenómeno «sometido a una

reducir todo el lenguaje a conjuntos y, por ello, recibe el nombre de semántica conjuntista. Sin embargo, «el problema de la semántica conjuntista es que es demasiado general y abstracta. [...] Pero esa semántica no nos da ulterior información de carácter más concreto sobre las entidades conjuntistas asociadas. Todo lo que hacemos es representarlas de algún modo más sencillo; por ejemplo, mediante letras, como hemos hecho anteriormente. Para muchos fines este tipo de análisis semántico tan abstracto no basta» (Acero-Bustos-Quesada, 1989: 148).

28. La lógica simbólica axiomatizada está en la base de los lenguajes formales de la informática empleados por la lingüística matemática como modelo de descripción de la organización y funcionamiento de las lenguas naturales. Por esta razón, como es obvio, a la hora de abordar el estudio y la comprensión de los principios teórico-metodológicos que fundamentan la lingüística matemática, es necesario conocer siquiera algunos rudimentos de disciplinas como la Lógica, el álgebra, los lenguajes de programación... Los manuales más difundidos no alcanzan a exponer con detalle todos y cada uno de estos conocimientos. Maurice Gross, en *Modelos matemáticos en lingüística* (1972), expone los fundamentos informáticos de la lingüística matemática, pero omite el fundamento lógico-matemático de esta disciplina. En *Introducción a la lingüística matemática* (1972), Gladkij y Melchuk exponen los logros de la lingüística matemática en la U.R.S.S., así como sus carencias y defectos al aplicarse a las lenguas naturales, pero dan por sabidos los fundamentos matemáticos, tanto lógicos como algebraicos de dicha disciplina, y no hacen referencia alguna a conceptos fundamentales de la informática. Sebastián Serrano, en *Elementos de lingüística matemática* (1975) y en *Lógica, lingüística y matemáticas* (1977), expone ampliamente y con precisión los fundamentos algebraicos de la lingüística matemática, aunque obvia la base lógica de dicha disciplina. Los lingüistas rumanos Marcus, Nicolau y Steti, en *Introducción a la lingüística matemática* (1978), describen el papel de la lingüística matemática en el contexto de la lingüística general, pero nada dicen de su fundamentación matemática y nada de su origen informático. Las publicaciones de Zellig Harris –*Methods in Structural Linguistics* (1951), *Structural Linguistics* (1969), *Structures mathématiques du langage* (1971)– constituyen una excepción que se explica por el hecho de que estos trabajos se sitúan en el inicio del desarrollo de la lingüística matemática. Ciertamente, Harris presenta y desarrolla un modelo de lingüística matemática propio, basado casi exclusivamente en las transformaciones. Finalmente, el libro de Partee, Meulen y Wall, titulado *Mathematical Methods in Linguistics* (1990), es posiblemente una de las publicaciones más completas porque expone con detalle tanto los fundamentos algebraicos como los lógicos, e incluso los informáticos, de la lingüística matemática, además de recoger las aportaciones de Richard Montague sobre la matematización de la semántica. Sólo el libro de Maurice Gross expone de forma explícita y clara la relación entre la lingüística matemática y la lingüística generativa.

prueba como sustitutivo de otro objeto o fenómeno B, con el cual A mantiene una cierta correspondencia» (Marcus-Nicolau-Steti, 1978: 75-77)²⁹.

A pesar de que se han demostrado las afinidades entre el lenguaje formal y el lenguaje natural, las gramáticas formales son sólo un modelo para las lenguas naturales y sólo cumplen algunas propiedades de éstas³⁰. Y, aunque

29. La problemática de esta aplicación de las gramáticas generativas formales a las lenguas naturales la describe muy bien Lyons (1993: 108): «Una gramática generativa consiste en una especificación matemática precisa de la estructura gramatical de las oraciones que ella misma genera. Esta definición no limita la aplicabilidad de la gramática generativa a las lenguas naturales. De hecho, tampoco implica que la gramática generativa sea en absoluto pertinente para describir lenguas naturales. Los conjuntos de sintagmas caracterizados como lenguas por las gramáticas generativas son lo que los lógicos denominan *lenguas formales*. Todo sintagma posible está o no bien formado; no existen sintagmas en un estado intermedio o indeciso a este respecto. Además, todo sintagma bien formado presenta una estructura totalmente determinada, definida por la descripción estructural que le asigna la gramática. No está claro que las lenguas naturales sean formales en este sentido del término. Muchos lingüistas sostendrían que no lo son. Pero esto no significa que las lenguas formales no puedan utilizarse como modelo de las lenguas naturales. Basta con que la propiedad de la gramaticalidad, aun cuando no quede totalmente determinada, lo sea empíricamente dentro de unos límites razonables, y también que las demás propiedades estructurales del modelo puedan identificarse en la lengua natural a la que sirve de modelo la lengua formal en cuestión».

30. El concepto de «modelo» está definido en la lógica de predicados y, para determinarlo con rigor de forma matemática, es necesario fijar previamente los conceptos de sistema, tipo de similaridad, homologación, interpretación, satisfacción y teoría. Un sistema se podría definir como un cuádruple $\langle A; R_1 \dots R_n; f_1 \dots f_m; a_1 \dots a_l \rangle$ formado por los siguientes cuatro conjuntos:

- Un conjunto A no vacío, llamado universo del sistema, que agrupa a la totalidad de los elementos que forman parte de él.
- Un conjunto de propiedades o relaciones monarias $\{R_1 \dots R_n\}$ de los elementos de A, que definen subconjuntos A_1, \dots, A_n dentro de él y que interesa distinguir, o bien un conjunto de relaciones n-arias $A \times A \times \dots \times A$, definidas en ese mismo universo A y que forman n-tuplas ordenadas o cadenas (a_1, a_2, \dots, a_n) con los elementos de él, que también interesa distinguir.
- Un conjunto $\{f_1, \dots, f_m\}$ de funciones o aplicaciones $f: A \times A$ definidas en dicho universo A que interesa distinguir.
- Un subconjunto de elementos $\{a_1, \dots, a_l\}$ de dicho universo que interesa distinguir.

Los conjuntos de elementos distinguidos y de propiedades o relaciones distinguidas puede ser el conjunto vacío (Mosterín, 1989: 122).

El tipo de similaridad de un sistema puede definirse como una terna ordenada $\langle r_1 \dots r_n; s_1 \dots s_m; l \rangle$ compuesta por la aridad R_i de las relaciones, por la aridad s_i de las funciones y por el número l de elementos distinguidos del sistema (Mosterín, 1989: 123). Considerando que un lenguaje formal cualquiera es un sistema y, por tanto, se comporta como tal en relación a otros sistemas, se puede afirmar que «un lenguaje formal y un sistema se llaman homólogos cuando ambos tienen el mismo tipo de similaridad» (Mosterín, 1989: 128). Siendo L un lenguaje formal homólogo de un sistema A, una interpretación J de L en A es una aplicación $J: L \rightarrow A$ tal que $y = J(x)$ para cualquier elemento $x \in L$ e $y \in A$ (Mosterín,

Lyons justifica la aplicabilidad del modelo, deja claro también que «tampoco implica que la gramática generativa sea en absoluto pertinente para describir lenguas naturales» porque «todo sintagma bien formado presenta una estructura totalmente determinada, definida por la descripción estructural que le asigna la gramática» y, evidentemente, «no está claro que las lenguas naturales sean formales en este sentido del término. Muchos lingüistas sostendrían que no lo son» (Lyons, 1993: 108). Uno de estos lingüistas que criticaron desde el principio la equiparación de las lenguas naturales con los lenguajes formales fue Noam Chomsky, a pesar de que al principio se ocupó de la ló-

1989: 129). Por lo tanto, «Con esto estamos en disposición de definir recursivamente la satisfacción de cualquier fórmula por la interpretación \mathbf{J} sobre \mathbf{A} » (Mosterín, 1989: 130). A partir del concepto de satisfacción definido anteriormente, se puede definir una teoría como un conjunto Γ de fórmulas, proposiciones lógicas, enunciados o sentencias de un lenguaje formal tal que todas y cada una de las proposiciones ϕ de dicho lenguaje formal se deduce de ese conjunto Γ (Mosterín, 1989: 131). Y un sistema es un modelo de una teoría determinada si satisface todas y cada una de las fórmulas o proposiciones de ésta, o sea, si se define una aplicación entre el lenguaje formal al que pertenecen las proposiciones de dicha teoría y un sistema homólogo a éste: «Las sentencias que constituyen una teoría \mathbf{T} pueden ser interpretadas sobre un sistema homólogo cualquiera \mathbf{A} . Con ello se convierten en proposiciones que pueden ser verdaderas o falsas. Si todas son verdaderas, decimos que \mathbf{A} es un modelo de \mathbf{T} . Resumiendo: \mathbf{A} es un modelo de \mathbf{T} si y sólo si para cada $\phi \in \mathbf{T}$: \mathbf{A} sat ϕ » (Mosterín, 1989: 135).

Como se puede comprobar, la diferencia existente entre las lenguas naturales y los lenguajes formales, que ha sido expresada matemáticamente y que abarca también a los elementos de las lenguas naturales que no existen en los lenguajes formales (como la morfología o la fonología), hace que no tengan el mismo tipo de similaridad $\langle r_1 \dots r_n; s_1 \dots s_m; l \rangle$, que a consecuencia de ello no son homólogos y que no exista una aplicación $\mathbf{J}: \mathbf{L} \rightarrow \mathbf{A}$, tal que $y = \mathbf{J}(x)$ para cualquier elemento $x \in \mathbf{L}$ e $y \in \mathbf{A}$, entre las relaciones, funciones y elementos distinguidos de unas y de otros. Tan sólo existe una correspondencia entre los sistemas \mathbf{L} y \mathbf{A} , pero esta correspondencia $\mathbf{J}: \mathbf{L} \rightarrow \mathbf{A}$ no es una aplicación porque existen elementos de \mathbf{L} cuya imagen no es un conjunto unitario o es el conjunto vacío, por lo que no cumplen la definición de aplicación. Si esta correspondencia no es una aplicación, el sistema \mathbf{A} que forman los lenguajes formales no satisface algunos enunciados ϕ de teoría de las gramáticas transformativas. Y, si el sistema que forman los lenguajes formales no satisface una parte de la teoría de las gramáticas transformativas de las lenguas naturales, entonces, los lenguajes formales y las gramáticas generativas que los describen no son un modelo válido para las lenguas naturales, porque no se ajustan a la definición matemática de modelo. Puesto que la consideración de los lenguajes formales como modelo de las lenguas naturales es, en definitiva, el principio que fundamenta la aplicación de la lingüística matemática a las lenguas naturales, el hecho de que los lenguajes formales no sean un modelo válido para éstas implica que la lingüística matemática planteada en estos términos no es aplicable totalmente a las lenguas naturales. Se puede aplicar a las partes regulares de sus gramáticas, porque éstas son comparables a los lenguajes formales, pero no tanto a las irregularidades que presentan, porque éstas son consecuencia de la historia de cada lengua en particular y, por ello, no son generadas mecánicamente mediante reglas por los hablantes, sino aprendidas por ellos.

gica simbólica con fines lingüísticos y así lo reflejó en 1953 en su primera publicación, «Systems of Syntactic Analysis», que apareció en una revista de lógica llamada *Journal of Symbolic Logic* (Chomsky, 1953: 242-256). Y lo reiteró dos años más tarde en *Logical Structure of Linguistic Theory* (Chomsky, 1975) al postular una estructura-S somera o superficial, afirmando «una estructura-S es análoga a una expresión en los sistemas convencionales de la lógica, y es convertida en una representación de la “forma lógica” por medio de reglas muy simples» (Chomsky, 1955a: 231-232)³¹.

Sin embargo, muy pronto cambió de actitud y pasó a criticar la idea de Carnap de estudiar la sintaxis de las lenguas naturales comparándola con la de los lenguajes formales, tomándose la de éstos como modelo de la de las lenguas naturales, y en 1955 rechazó la propuesta de Bar-Hillel de seguir desarrollando en la Lingüística el método de Carnap, que había sido expuesta un año antes. Según Chomsky, la Lógica sólo sería útil para formular una teoría lingüística general, pero no para decidir qué tipos de sistemas forman el objeto de la lingüística ni cómo han de describirse³². Por ello, sólo utiliza la lógica simbólica para axiomatizar las principales teorías sintácticas del estructuralismo norteamericano. Y, basándose en los trabajos de los teóricos de la lógica simbólica E. Post, P. Rosembloom y W. Quine, desarrolló su modelo matemático y la definición matemática de gramática generativa (Rohrer, 1978: 27-34), expuesto por primera vez en 1955 en *The Logical Structure of Linguistic Theory* y después, en 1963, en «Introduction to the formal analysis of natural languages» en colaboración con G. Miller.

De esta forma, el uso de las gramáticas generativas formales como modelo de las lenguas naturales y la consiguiente influencia en la lingüística teórica de la lingüística matemática debidamente desformalizada queda patente en la intención de Chomsky al decir que «la gramática de una lengua concreta debe ser suplantada por una gramática universal que explique el aspecto creativo del uso lingüístico y exprese las profundas regularidades que, por ser universales, no aparecen en la gramática popularmente dicha» (Chomsky, 1965: 8). Con el término «gramática universal» se refiere a la idea desarrollada ya por Descar-

31. «La tesis doctoral de Noam Chomsky (*The Logical Structure of the Linguistic Theory*, 1955) supuso la primera aplicación de ideas procedentes de la teoría de los sistemas formales de la lógica al estudio de problemas de las lenguas naturales.» (Moreno Sandoval, 2001: 227)

32. «The correct way to use the insights and techniques of logic is in formulating a general theory of linguistic structure. But this fact does not tell us what sort of systems form the subject matter for linguistics, or how the linguist may find it profitable to describe them» (Chomsky, 1955b: 45).

tes, por la *Grammaire Général et Raisonnée* o *Gramática de Port-Royal* (1660) y por Leibniz, y basada esencialmente en el lenguaje formal de la lógica como universal y subyacente de todas las lenguas, e igualmente se refiere sin duda a la idea de característica universal de Leibniz en la expresión «lengua *caracterizada* por la gramática». Esta idea tuvo continuación en los siglos XIX y XX, sobre todo por parte de los lógicos, como G. Frege, B. Russel, L. Wittgenstein y R. Carnap, quienes mantenían que «sólo la Lógica puede aclararnos la naturaleza de determinados problemas y sólo en ella son formulables las correspondientes respuestas de un modo preciso» (Acero-Bustos-Quesada, 1989: 24). Siguiendo esta idea, y tomando como modelo las gramáticas formales descritas por la lingüística matemática, fue como Chomsky intentó explicar «las profundas regularidades» de la gramática de las lenguas naturales, que formarían parte de lo que él llamó estructura profunda. Sin embargo, este modelo de las gramáticas formales, y concretamente de las gramáticas transformativas, no explicaba tan bien todas las irregularidades de las lenguas naturales, si bien es cierto que tampoco lo pretendía.

Por ello, las reglas gramaticales de la gramáticas generativas aplicadas a las lenguas naturales (gramática transformativa) han sufrido continuas modificaciones de las reglas y hasta varias reformulaciones de la teoría, la más famosa de ellas realizada por el propio Chomsky en 1965 en su obra *Aspects of the Theory of Syntax*³³.

33. Así, «tanto las reglas de la estructura de constituyentes como también las reglas T se modifican tan pronto como se consideren nuevos hechos. Chomsky propuso en 1965 un nuevo análisis que, a su vez, ha sido modificado por Lakoff, etc. Efectivamente, en la historia de la GT no conocemos una regla que no se haya modificado más de una vez» (Rohrer, 1978: 66-67). Por ejemplo, el elemento auxiliar (abreviado Aux), que engloba a todos los accidentes gramaticales del verbo de una oración, ha sido analizado de muy diversas maneras y, mientras algunos autores lo incluyen dentro del sintagma verbal mediante las reglas $O \rightarrow SN + SV$ y $SV \rightarrow Aux + VM$, otros lo independizan totalmente de él mediante la regla $O \rightarrow SN + Aux + SV$ (Fraser, 1965) mediante lo que él denomina «convención de transportabilidad» (*transportability convention*). M. Bierwisch postuló como primaria para el alemán la regla $SV \rightarrow SN + V$, en lugar de la normal $SV \rightarrow V + SN$, porque permite una formulación más sencilla de las reglas transformativas al cambiar la posición del verbo (Bierwisch, 1963). Sin embargo, esta estructura primaria teórica fue rebatida por J. Ross, en 1967. Por eso, «si bien ciertos fenómenos sintácticos, como la posición del verbo en las oraciones subordinadas y en las oraciones con formas verbales de infinitivo, se pueden formular más fácilmente, surgen, por otra parte, nuevas dificultades y complicaciones. En el estado actual de la investigación, el orden sujeto, verbo, objeto parece valer como primario también para el alemán» (Rohrer, 1978: 68). En efecto, «the insufficiently restricted state the theory allowed an even more disturbing conclusion to be drawn: Given the sort of data that linguists ordinarily consider relevant, the Universal Base Hypothesis (the conjecture that the grammars of all natural languages have the same base rules) could not be proved false» (Partee-Meulen-Wall, 1990: 557), debido entre otras cosas a que «from

La auténtica utilización de las Matemáticas en las ciencias del lenguaje no llegó, por lo tanto, hasta la década de los años 40 y de los 50, con la axiomatización de los lenguajes formales utilizados en la informática y como consecuencia de la aparición de los ordenadores y de los lenguajes de programación. Esta axiomatización se utilizó como modelo para las lenguas naturales con el fin de posibilitar la traducción automática por ordenador y, tras el fracaso de ésta en los años 60, comenzó a utilizarse el nombre de lingüística matemática para esta nueva disciplina de las ciencias del lenguaje³⁴. Aparece como consecuencia de una innovación tecnológica, pero que al desarrollarse y perfeccionarse necesitó crear unos medios adecuados y cómodos para hacer funcionar una máquinas cuya complejidad excedía la de cualquier otra máquina conocida hasta entonces. Por ello, los medios mecánicos o electrónicos usuales que hacían funcionar las demás máquinas, tales como botones o interruptores, resultaban insuficientes para los ordenadores y se hizo necesario idear otros medios más sofisticados para comunicarse con ellos, como crear unos lenguajes artificiales con los que darles las órdenes oportunas.

Los lenguajes artificiales capaces de hacer funcionar los ordenadores reciben el nombre de *lenguajes de programación* y están especialmente diseñados para impartir órdenes al computador, por lo que deben tener unas características especiales y una sintaxis sencilla pero muy rigurosa para que resulte comprensible a la máquina. Ese rigor y esa sencillez en la sintaxis no se encuentran en las lenguas naturales, pero sí en el *lenguaje formal* de la lógica matemática y de las Matemáticas mismas, especialmente del álgebra, razón por la cual dicho lenguaje formal se tomó como modelo para los lenguajes de programación. La fundamentación matemática o axiomatización de tales lenguajes de programación dio origen a la lingüística matemática y se tomó, a su vez, como *modelo* para las lenguas naturales³⁵.

this it is relatively straightforwaed to show that there are an infinite number of such trivial “universal bases”» (Partee-Meulen-Wall, 1990: 559).

34. En efecto, «como disciplina constituida, la lingüística matemática aparece muy tarde [...], como consecuencia del desarrollo vertiginoso de los ordenadores electrónicos» (Marcus-Nicolau-Steti, 1978: 41).

35. Para que las lenguas naturales pudieran ser utilizadas por el ordenador tenían que ser «comprendidas» por él o, lo que es lo mismo, adaptadas al lenguaje propio de los ordenadores, que era el lenguaje de programación. Para ello, era necesario formalizar las lenguas naturales de una forma adecuada a los lenguajes de programación, que son los únicos que entiende el ordenador. Dado que todos los lenguajes de programación están basados en el lenguaje formal de la Lógica, la única manera de conseguirlo era adaptando las lenguas naturales al lenguaje formal de la Lógica, porque éste era el único «comprensible» para el ordenador. Para ello, fue necesario estudiar las características de ambas clases de len-

La gramática de un lenguaje formal fue definida como un sistema axiomático semithueano y las reglas de reescritura que «generaban» los enunciados de los lenguajes formales se equipararon a las reglas de inferencia de los sistemas axiomáticos, por lo que tales gramáticas se denominaron *gramáticas generativas*³⁶.

guajes o, lo que es lo mismo, conocer sus respectivas gramáticas, para definir después la gramática de las lenguas naturales mediante las características de los lenguajes formales utilizados en la lógica. En otras palabras, la solución consistió en tomar como *modelo* la gramática de los lenguajes formales para definir la gramática de las lenguas naturales. Y esto fue posible gracias a la *teoría de modelos*, definida por A. Tarski dentro de su teoría de la verdad, que fue expuesta en el artículo «El concepto de verdad en los lenguajes formalizados», presentado por él en 1931 a la Sociedad Científica de Varsovia (Garrido, 1991: 220).

36. «Una *gramática formal* es un tipo particular de *sistema axiomático*, que es una estructura matemática definida por el siguiente cuadruple: $\langle V, F, A, R \rangle$ donde V es un conjunto finito de símbolos llamado *alfabeto* o vocabulario; F es un lenguaje formal sobre V , llamado conjunto de *fórmulas* bien formadas; A es un subconjunto de F , llamado conjunto de los *axiomas*; R es un conjunto de *reglas* de derivación» (Serrano, 1977: 188-189): $SA = \langle V, F, A, R \rangle$. Pues bien, el conjunto V , que es el conjunto de símbolos del lenguaje formal de la lógica, y el F , que es el conjunto de fórmulas lógicas bien formadas que constituyen un lenguaje formal sobre V , son una parte del sistema axiomático, la del *lenguaje formal*, que es distinta de la otra parte, que es la del sistema deductivo. La parte del lenguaje formal también se puede definir matemáticamente por separado, porque el conjunto F es también un álgebra en la que se definen unas operaciones que determinan cuáles de las fórmulas lógicas de F están bien formadas. Una *gramática generativa* se definió, entonces, como un *sistema axiomático semithueano*, $G = \langle V_T, V_N, S, R \rangle$, que está formado por «un vocabulario fundamental o terminal, un vocabulario auxiliar, un símbolo inicial y una serie de reglas de reescritura» (Gladkij y Melchuk, 1972: 56), de tal manera que el «símbolo inicial» es el *axioma inicial* de los sistemas axiomáticos y, como en éstos, pertenece al vocabulario auxiliar. Por lo tanto, los conjuntos que lo forman son V_T, V_N, S y R , siendo R el conjunto de reglas de escritura que actúa sobre la unión de los conjuntos V_T y V_N , y siendo S un elemento de V_N . Las *reglas gramaticales* serían *funciones recursivas* de tipo semithueano y, por tanto, *aplicaciones* $f: \Sigma \rightarrow \Sigma$, tal que $\Sigma = V_T \cup V_N$ y cuyos pares ordenados presentan la forma (ψ, ω) , representada habitualmente por $\psi \rightarrow \omega$ y que se lee « ψ se reescribe ω »: $G = \langle V_T, V_N, S, R \rangle$, tal que $R = (x_1, x_2, \dots, x_{n-1}, x_n) \Rightarrow x_1, x_2, \dots, x_{n-1} \rightarrow x_n$. Teniendo en cuenta que un *álgebra A* es un conjunto A en el que se define una o más *leyes de composición* f_1, f_2, \dots, f_n y se representa por $A = \langle A, f_1, f_2, \dots, f_n \rangle$, una gramática formal es, por tanto, un álgebra y, más concretamente, un *monoide* definido en la unión de V_T y V_N , o sea en $V_T \cup V_N$, teniendo en cuenta que $S \in V_N$ y que las leyes de composición interna u operaciones asociadas forman el conjunto R de las reglas de reescritura. El conjunto R de las reglas de reescritura representa la concatenación de los elementos de $V_T \cup V_N$ y se puede representar por el símbolo $\widehat{\quad}$, con lo cual $\widehat{\quad}$ sería la ley de composición interna (por ejemplo, $\widehat{a \quad b \quad c \quad \dots \quad n}$), que es asociativa y tiene elemento neutro, pero no simétrico, por lo que el conjunto $\Sigma = V_T \cup V_N$ en el que se define no es un grupo, sino sólo un monoide. La *concatenación*, que es la ley de composición interna $\widehat{\quad}$ definida en el monoide, no sería más que una *relación binaria de orden* definida en el conjunto $\Sigma = V_T \cup V_N$, que forma una serie de *cadena*s cuyos elementos pueden ser también, a su vez, *cadena*s (Partee, Meulen,

Dado que la lingüística matemática apareció «como consecuencia del desarrollo vertiginoso de los ordenadores electrónicos» (Marcus-Nicolau-Steti, 1978: 41), su evolución estuvo condicionada por la de éstos y, por tanto, se desarrolló en función de los avances técnicos que registraron los ordenadores y el campo de la ingeniería informática³⁷. De esta forma, los tipos de lenguajes definidos estuvieron determinados por el tipo de ordenador para el que estaban destinados, de tal manera que la complejidad de estos lenguajes aumentaba en proporción directa a la de las máquinas que los debían utilizar. Principalmente, estos tipos de lenguajes son tres: los lenguajes de estados finitos o *finite state languages*, los lenguajes libres de contexto o *context-free languages* y los lenguajes dependientes del contexto o *context-sensitive languages*. Éstos toman los nombres del tipo de gramáticas que los generan, gramáticas de estados finitos o *finite state grammars* (GEF), gramáticas independientes del contexto o *context-free grammars* (GIC) y gramáticas dependientes del contexto o *context-sensitive grammars* (GDC)³⁸.

Wall, 1990: 434). Por lo tanto, una gramática formal es un monoide en el que se define una ley de composición interna denominada concatenación, que es asociativa y tiene elemento neutro, pero carece de simétrico, por lo que el conjunto $\Sigma = V_T \cup V_N$ en el que se define no llega a ser un grupo.

37. El uso de las gramáticas generativas de los lenguajes formales, y concretamente de los lenguajes de programación, como modelo para estudiar las lenguas naturales realizado por Chomsky ya fue señalado por Matthews: «Sus primeros críticos relacionaron muchas veces su trabajo con los proyectos de traducción mecánica. Una gramática generativa puede ser adecuada para una computadora pero, según decían, para ninguna otra cosa. En sus respuestas, Chomsky negó que tuviera interés ninguno por ese campo. Pero, ¿importa eso realmente? El primer experimento de traducción (el llamado “experimento de Georgetown”) es de 1954, la importante obra inédita de Chomsky, de 1955: ¿hemos de ver una simple coincidencia? Con seguridad, las dos cosas pertenecen al mismo movimiento intelectual. En otros temas, los pioneros habían empezado justo entonces a pensar en términos de computación: más tarde, esas formas de pensar alcanzarían su apogeo en las universidades de lengua inglesa de los años 60. Estamos ahora en las décadas siguientes. Podemos ver que gran parte de aquello no era más que simple embriaguez. ¿No lo apreciaron correctamente los críticos de Chomsky en ese momento?» (Matthews, 1983: 141).
38. Toda *gramática independiente de contexto* (GIC) se puede definir también como una *gramática dependiente de contexto* (GDC) en la que ambos elementos del contexto de un determinado símbolo que cambia en virtud de una regla son elementos vacíos, o sea, como una gramática cuyas reglas de reescritura son $\alpha A \beta \rightarrow \alpha w \beta$, tal que $A \in V_A$; $\alpha, \beta, w \in (V_T \cup V_N)$ pero $\alpha \wedge \beta = \emptyset$. Por lo tanto, toda gramática IC se puede considerar como un caso particular de gramática DC y satisface todas las condiciones de una gramática DC, por lo que se puede afirmar que los lenguajes IC son, a su vez, lenguajes DC, pero no al contrario. Al mismo tiempo, las gramáticas regulares son, a su vez, un caso particular de gramáticas IC, ya que sólo se diferencian de ellas en que tienen una regla final $P_1 \rightarrow e$, con lo que los lenguajes regulares son, a su vez, lenguajes IC. Por esta razón, los tres tipos de gramática se pueden definir muy bien en orden inverso:

Esta axiomatización de las gramáticas de los lenguajes formales de la lógica y de los lenguajes de programación, se aplicó a las lenguas naturales con el objeto de axiomatizarlas para que pudieran ser comprendidas y reproducidas por el ordenador y posibilitar así la traducción automática³⁹. Los 3 tipos de gramáticas generativas mencionadas (GEF, GIC y GDC) explican perfectamente los lenguajes formales pero, al utilizarlas como modelo para

- **Gramáticas dependientes del contexto o GDC:** Aquellas cuyas reglas de reescritura son del tipo $\alpha A\beta \rightarrow \alpha w\beta$, tal que $A \in V_A$; $\alpha, \beta, w \in (V_T \cup V_A)$; $\alpha \vee \beta \neq \emptyset$ y $\beta \neq \epsilon$. Puede ocurrir que $\alpha = \emptyset$, en cuyo caso la regla sería $A\beta \rightarrow w\beta$, o que $\beta = \emptyset$, en cuyo caso sería $\alpha A \rightarrow \alpha w$.
- **Gramáticas independientes del contexto o GIC:** Aquéllas cuyas reglas de reescritura son del tipo $\alpha A\beta \rightarrow \alpha w\beta$, tal que $A \in V_A$; $\alpha, \beta, w \in (V_T \cup V_A)$; pero $\alpha = \beta = \emptyset$, o sea, del tipo $A \rightarrow w$. Por tanto, son aquellas gramáticas dependientes del contexto en cuyas reglas de reescritura $\alpha A\beta \rightarrow \alpha w\beta$, $\alpha = \beta = \emptyset$, o sea, del tipo $A \rightarrow w$.
- **Gramáticas regulares (GR) o de estados finitos:** Aquéllas cuyas reglas de reescritura son del tipo $\alpha A\beta \rightarrow \alpha w\beta$, tal que $A \in V_A$; $\alpha, \beta, w \in (V_T \cup V_A)$; $\alpha = \beta = \emptyset$ y existe una regla final $P_1 \rightarrow \epsilon$, o sea, que son del tipo $A \rightarrow w$ tal que $\exists(P_1 \rightarrow \epsilon)$, siendo ϵ el elemento neutro. Por tanto, son aquellas gramáticas dependientes del contexto en cuyas reglas de reescritura $\alpha A\beta \rightarrow \alpha w\beta$, $\alpha = \beta = \emptyset$ y $\exists(P_1 \rightarrow \epsilon)$ o, lo que es lo mismo, son aquellas gramáticas independientes del contexto en cuyas reglas de reescritura $A \rightarrow w$, $\exists(P_1 \rightarrow \epsilon)$.

Los tres tipos de gramática se resumirían de forma más sintetizada utilizando sólo símbolos matemáticos, de la siguiente manera:

GDC = $\langle V_T, V_A, P_0, R \rangle$:

$R = \alpha A\beta \rightarrow \alpha w\beta$: $A \in V_A$; $\alpha, \beta, w \in (V_T \cup V_A)$; $\alpha \vee \beta \neq \emptyset$; $\beta \neq \epsilon$.

$\alpha = \emptyset \Rightarrow (R = A\beta \rightarrow w\beta)$, $\beta = \emptyset \Rightarrow (R = \alpha A \rightarrow \alpha w)$.

GIC = $\langle V_T, V_A, P_0, R \rangle$:

$R = \alpha A\beta \rightarrow \alpha w\beta$: $A \in V_A$; $\alpha, \beta, w \in (V_T \cup V_A)$; $\alpha = \beta = \emptyset$.

$R = A \rightarrow w$.

GR = $\langle V_T, V_A, P_0, R \rangle$:

$R = \alpha A\beta \rightarrow \alpha w\beta$: $A \in V_A$; $\alpha, \beta, w \in (V_T \cup V_A)$; $\alpha = \beta = \emptyset$; $\exists(P_1 \rightarrow \epsilon)$.

$R = A \rightarrow w$; $\exists(P_1 \rightarrow \epsilon)$.

Los tres tipos de gramática tienen una parte común, que consiste en la propia fórmula definitoria y en el tipo de reglas de reescritura:

$$\mathbf{G} = \langle V_T, V_A, P_0, R \rangle: R = \alpha A\beta \rightarrow \alpha w\beta: A \in V_A; \alpha, \beta, w \in (V_T \cup V_A).$$

Y de esta definición se deduce claramente que toda gramática formal es un sistema axiomático semithueano, y ello se puede demostrar con la simple comparación de las respectivas definiciones matemáticas:

Sistema semithueano = $\langle A, B, S, P \rangle$: $P = \alpha x\beta \rightarrow \alpha y\beta$.

Gramática formal = $\langle V_T, V_A, P_0, R \rangle$: $R = \alpha A\beta \rightarrow \alpha w\beta$: $A \in V_A$; $\alpha, \beta, w \in (V_T \cup V_A)$.

$\langle A, B, S, P \rangle = \langle V_T, V_A, P_0, R \rangle$

$A = V_T$; $B = V_A$; $S = P_0$; $P = R$

39. En efecto, «hasta los años 50 no se construyen sistemas axiomáticos, que cumplan los presupuestos matemáticos, para dominios parciales de las lenguas naturales. Tales modelos fueron desarrollados, en parte, en el marco de las investigaciones sobre la traducción automática» (Rohrer, 1978: 24).

las lenguas naturales, se comprobó que eran incapaces de describirlas⁴⁰. Por ello, fue necesario realizar algunas adaptaciones de las gramáticas generativas añadiendo conceptos que no existían en los lenguajes formales. Dado que la lingüística estructural imperante en aquel momento resultaba insuficiente para adaptarla a la informática porque no explicaba cómo se producía el lenguaje y cómo lo debía producir el ordenador, se tuvo que completar y perfeccionar añadiéndole las reglas gramaticales que explicaban la producción de oraciones. Para ello, también se recurrió a la lógica simbólica y, concretamente, a los sistemas axiomáticos y a las reglas de reescritura de las fórmulas lógicas, que eran concebidas como reglas de inferencia. Chomsky estableció que las reglas gramaticales se expresaran mediante las reglas de reescritura⁴¹, que pueden ser reglas sintagmáticas o de ramificación y reglas

40. La lingüística matemática, que nació simultáneamente en la U.R.S.S. y en Estados Unidos a mediados de los años 50 como consecuencia de la aparición de los ordenadores, se ha aplicado a la matematización de los lenguajes formales, pero no ha sido tan eficaz a la hora de aplicarse a los lenguajes naturales. En efecto, la lingüística matemática ha conseguido resultados con las lenguas formales o artificiales, pero no con las lenguas naturales, y ello es reconocido por los mismos especialistas. Así, el mismo Harris (1971) definió matemáticamente el concepto de gramática o de lengua como un «sistema abstracto» («le systéme abstrait»), y así tituló el capítulo séptimo en el que incluye y explica extensamente esa definición. Pero Gladkij y Mel'čuk en su *Introducción a la lingüística matemática* (1972: 94) lo reconocen abiertamente: «... no podemos ofrecer al lector una exposición bien trabada de resultados precisos en lo tocante a la descripción de los lenguajes naturales empleando las gramáticas generativas de tal o cual clase. Tendremos que limitarnos a consideraciones de carácter general y a un reducido número de hechos conocidos».

41. Las reglas de reescritura propuestas por Chomsky para las lenguas naturales son en realidad pares ordenados del tipo (x, y) pertenecientes al grafo de la aplicación $f: X \rightarrow Y$ definida entre un conjunto finito X y un conjunto finito Y y que por convenio se representan por medio de la flecha que simboliza dicha aplicación ($X \rightarrow Y$). Este tipo de regla es una aplicación porque es una correspondencia definida entre los conjuntos finitos X e Y en la que todo elemento de X tiene imagen y sólo una imagen, de manera que $y = f(x), \forall x \in X$. «Esta regla se lee así: el símbolo X ha de volverse a escribir como el símbolo Y . La flecha entre X e Y no es otra que una instrucción, una orden que nos obliga a no detenernos en X : el paso a Y es inevitable. Hablando de un modo más técnico diremos: X se reescribe Y .» (Tusón, 1995: 161)

Uno de los dos componentes de esta aplicación puede ser a su vez compuesto y estar formado por dos elementos, uno del conjunto Y y otro de un tercer conjunto Z , también finito, en cuyo caso este elemento imagen de la aplicación sería a su vez un par ordenado formado por un elemento cualquiera de cada uno de esos dos conjuntos o, lo que es lo mismo, por un par ordenado perteneciente al producto cartesiano $Y \times Z$ de los conjuntos Y y Z . La aplicación sería entonces $X \rightarrow Y \times Z$, en la que uno de los dos conjuntos que la definen es el producto cartesiano de dos conjuntos. Y toda aplicación en la que uno de los dos conjuntos que la definen es el producto cartesiano de dos conjuntos constituye una ley de composición, inversa en este caso (f^1) porque el producto cartesiano es el elemento imagen de la aplicación f . En Lingüística, esta ley de composición $X \rightarrow Y \times Z$

de inserción léxica, y añadió el concepto de transformación, que ya había sido formulado por el estructuralista Harris para explicar las irregularidades y ambigüedades de las lenguas naturales concibiendo las reglas transformativas también como reglas de reescritura⁴². Sin embargo, las reglas sintagmáti-

se representa de la forma $X \rightarrow Y + Z$. De esta forma, siendo $y \in Y$ y $z \in Z$, tenemos que $(y, z) = f^1(x)$, $\forall x \in X$.

Esta última regla $X \rightarrow Y \times Z$ puede tener dos variantes. La primera de ellas consiste en que uno de los dos componentes de un par ordenado del producto cartesiano $Y \times Z$ sea el elemento vacío \emptyset , o sea, que $(\emptyset, z) = f(x)$, $\emptyset \in Y$, $\forall z \in Z$. Cuando un elemento $y \in Y$ tiene el valor $y = \emptyset$, la ley de composición $X \rightarrow Y \times Z$, tal que $y = \emptyset$, es en realidad la aplicación $X \rightarrow Z$ porque puede ocurrir que $(\emptyset, z) = z = f(x)$, o sea la regla anterior. La ley de composición correspondiente sería $X \rightarrow Y \times Z$: $\emptyset \in Y$, y en Lingüística se representa por $X \rightarrow (Y) Z$, en el que (Y) sería opcional. La otra variante consiste en que el elemento imagen de la aplicación puede pertenecer al conjunto Y o al Z de forma excluyente, en cuyo caso el elemento imagen de la aplicación sería un elemento w perteneciente, o bien a Y o bien a Z , tal que $w = f(x)$ y $(w = y) \Delta (w = z)$. En este caso, se trata también de una aplicación $X \rightarrow (Y \Delta Z)$ o, lo que es lo mismo, $(X \rightarrow Y) \Delta (X \rightarrow Z)$, siendo Δ el símbolo de la disyunción exclusiva. Entonces, $w = f(x)$, tal que $w \in Y \Delta Z$ o, lo que es lo mismo, $w \in (A - B) \cup (B - A)$. En Lingüística, se representa por $X \rightarrow \{Y, Z\}$.

La primera regla enunciada, $X \rightarrow Y$, también podría considerarse una variante de la segunda, $X \rightarrow Y \times Z$, siempre que uno de los dos conjuntos de elementos imagen fuera el conjunto vacío. Así, si $Z = \{\emptyset\}$, tenemos que $X \rightarrow Y \times Z$ equivale a $X \rightarrow Y$, por lo que dicha regla se puede enunciar también de la forma $X \rightarrow Y \times Z$: $Z = \{\emptyset\}$. Así, todas las reglas de reescritura pueden considerarse como leyes de composición y representarse con la fórmula $X \rightarrow Y \times Z$, siendo las demás reglas distintas variantes de ésta que sólo difieren en los valores de cada uno de los conjuntos que la forman. Estas reglas de reescritura se completan sustituyendo cada una de las categorías finales por palabras concretas del vocabulario terminal de una lengua. Esta sustitución de categorías gramaticales por palabras del léxico de una lengua se denomina *inserción léxica*.

42. Harris (1971: 69) define una transformación como una aplicación biunívoca $\phi_{21}: A_1 \rightarrow A_2$ entre dos subconjuntos A_1 y A_2 del conjunto A de oraciones gramaticales, de una lengua. El subconjunto A_2 es el de las imágenes $a_2 = \phi(a_1)$ del A_1 , o sea, el de las oraciones que han sufrido alguna transformación, y el A_1 es el de las antiimágenes $a_1 = \phi^{-1}(a_2)$ de A_2 , de tal manera que la aplicación inversa es $\phi^{-1}: A_2 \rightarrow A_1$. Cada elemento $a_2 = \phi(a_1)$ del conjunto A_2 se diferenciará de su antiimagen $a_1 = \phi^{-1}(a_2)$ de A_1 dependiendo del tipo de transformación ϕ_{21} que haya sufrido, y estas transformaciones pueden ser básicamente de tres tipos, permutaciones, adiciones y omisiones. «Une relation transformationnelle entre deux ensembles de propositions A_1 et A_2 peut être décrite comme une application biunivoque ϕ de A_1 sur A_2 , telle que chaque proposition de A_2 est identique à son image inverse dans A_1 , à l'exception de quelques additions, permutations ou omissions qui accompagnent ϕ . On peut vérifier linguïstiquement que, étant donnée une transformation $A_1: A_2$, les permutations, additions et omissions qui différencient une séquence de morphèmes (ou proposition) de A_2 de son image inverse dans A_1 (qui leur est associée par chaque transformation) sera appelée la trace ϕ_{21} de la transformation, et peut être considérée comme une opération qui, à partir des séquences de morphèmes de A_1 , permet d'obtenir celles de A_2 . Nous écrivons: $A_2 = \phi_{21} A_1$ ou: $\phi_{21}: A_1 \rightarrow A_2$ ou: $\phi_{21}^{-1}: A_2 \rightarrow A_1$ » (Harris, 1971: 69). Por tanto, en la simbología generativa, estas transformaciones se podrían representar como las reglas

cas generaban todas las oraciones posibles, tanto las correctamente formadas o gramaticales como las formadas incorrectamente o agramaticales, por lo que fue necesario limitar su capacidad generativa introduciendo el concepto de subcategorización –que no era más que el análisis componencial formulado por Katz y Fodor en 1963– y estableciendo las consiguientes reglas de subcategorización⁴³.

de reescritura de las gramáticas formales, utilizando la formalización de las operaciones o leyes de composición, de la forma $X \rightarrow Y$ y $X \rightarrow Y \times Z$, siendo $X \rightarrow Y$ un caso particular de $X \rightarrow Y \times Z$ en el que $Z = \{\emptyset\}$.

Sin embargo, estas reglas transformativas no se pueden aplicar a todas las oraciones de una lengua, sino sólo a algunas de ellas, y para conocer el subconjunto de las oraciones o cadenas de una lengua para las que sirve cada regla de transformación, hay que recurrir al concepto de *analizabilidad*, que fue formulado por Chomsky en 1961 en su artículo «On the notion “rule of grammar”». Téngase en cuenta que los marcadores sintagmáticos, que son los que definen la estructura sintáctica de una oración, se representan mediante diagramas arbóreos y que éstos están formados por elementos no terminales o categorías gramaticales, que forman los nódulos, y por elementos terminales, que son los que sustituyen a los no terminales y forman las oraciones. Pues bien, sean t_1, \dots, t_n los elementos terminales en los que se puede dividir una oración o cadena terminal t , sea Q el marcador sintagmático de dicha cadena y T una regla transformativa que le afecta. Se cumple siempre que cada elemento t_i está dominado en Q por un nódulo o elemento no terminal A_i (una categoría gramatical). Entonces, la cadena t es *analizable* como un par ordenado $(t_1, \dots, t_n; A_1, \dots, A_n)$ en relación a Q . El primer componente t_1, \dots, t_n de este par ordenado, formado por todos los elementos terminales de la cadena t se denomina un *análisis propio* (*proper analysis*) de t en relación a Q y a la regla transformativa T . El segundo componente A_1, \dots, A_n , formado por todos los elementos no terminales o nódulos, se denomina *índice estructural* o *descripción estructural* y es el que especifica a qué cadenas se les puede aplicar la regla transformativa T . Los cambios que se producen en un marcador sintagmático Q al aplicar una regla transformativa T se llaman *cambios estructurales* y cada uno de los pasos aislados en los que se puede definir un cambio estructural recibe el nombre de *transformación elemental* (*elementary transformation*).

43. Si el conjunto inicial de constituyentes inmediatos de la aplicación que define toda regla gramatical no es el producto cartesiano, sino un subconjunto de él, fue necesario establecer un criterio que permitiera limitar la aplicación de las reglas gramaticales y decidir si una regla gramatical es correcta o no en cada caso, o sea, si los constituyentes inmediatos a los que afecta pueden generar o no una oración bien formada en una lengua. Y, como el hecho de que las oraciones estén bien formadas o no lo estén depende, como se ha visto en los ejemplos, del significado de sus constituyentes, ese criterio tenía que ser semántico, lo que obligó a valorar y desarrollar la semántica, tan marginada por el estructuralismo americano, como una parte fundamental de la Gramática. En efecto, fueron J. Fodor y J. Katz los que introdujeron en 1963 en la gramática generativa el *análisis componencial* en su versión conceptualista, que consiste en que el significado de un lexema se puede descomponer en una serie de *elementos conceptuales atómicos*. Por ejemplo, el concepto «yegua» se puede descomponer en los conceptos atómicos «animado», «adulto» y «hembra» (Acero-Bustos-Quesada, 1989: 65-66). De esta manera, una oración estará bien formada si los elementos conceptuales atómicos de sus constituyentes son compatibles entre sí y no

Surgió, así, una concepción dual de la gramática de las lenguas naturales, basada en dos estructuras sintácticas distintas que Chomsky, siguiendo una corriente muy generalizada en las distintas ciencias y basándose en la lógica de Frege, denominó *estructura profunda* y *estructura superficial*. La estructura profunda está formada por el componente sintáctico de la Gramática, que consta de las reglas sintagmáticas, de inserción léxica y de subcategorización, y por el componente semántico; mientras que la estructura superficial está formada por las reglas transformativas, que constituyen el componente transformacional, y por el componente fonológico. Se originó, así, un cuarto tipo de gramática generativa, llamado *gramática transformativa*, que se añadió a los tres ya existentes en la lingüística matemática y se utilizó desde el primer momento en la lingüística teórica aplicándose a las lenguas naturales con el nombre de gramática generativa y transformativa o generativismo. Sin

lo estará en el caso contrario. Teniendo en cuenta que las reglas gramaticales describían en realidad las relaciones entre categorías gramaticales, lo que hace el análisis componencial es establecer las categorías incluidas en otra categoría, o sea, subcategorizar. Por ello, a la aplicación del análisis componencial a los constituyentes se le llamó *subcategorización* y, a las reglas que establecen las asociaciones de constituyentes que generan oraciones gramaticales, *reglas de subcategorización*. De esta manera, se pudieron compensar las inadecuaciones de las reglas gramaticales, complementándolas con la subcategorización, que fue incorporada por Chomsky a la gramática generativa en 1965 en *Aspects of the Theory of Syntax*, tomándola del análisis componencial de Fodor y Katz (1963).

La subcategorización es el único fenómeno de la Gramática que convierte a las reglas gramaticales en reglas dependientes del contexto o *context-sensitive rules* ya que, a causa de ella, la aplicación de una regla gramatical cualquiera está condicionada por los rasgos semánticos de los constituyentes vecinos de aquellos a los que afecta, de tal manera que «the use of context-sensitive rules is restricted to subcategorization and lexical insertion» (Gross, 1972: 146-147). Las reglas de subcategorización se dividen, a su vez, en «dos tipos de reglas: las reglas de selección (*selectional rules*) y las reglas de subcategorización (*strict subcategorization rules*). Estos dos tipos de reglas son las únicas reglas de la base dependientes del contexto; todas las demás reglas de la base son reglas independientes del contexto» (Rohrer, 1978: 75). Las reglas de selección de Rohrer son las reglas de inserción léxica de otros autores como Gross (*lexical insertion*) y a las reglas sintagmáticas, Rohrer las denomina «reglas de ramificación».

El concepto de subcategorización también pertenece a la lingüística matemática y es, por tanto, complementario de las reglas gramaticales, que a su vez proceden de las reglas de reescritura de la definición matemática de gramática generativa. Así, los elementos sintácticos que intervienen en las reglas gramaticales no serían los simples constituyentes aislados, sino secuencias de ellos, sintagmas o, en términos matemáticos, n-tuplas ordenadas. Aun así, la aplicación inversa $f^{-1}: X \rightarrow Y \times Z$ que define matemáticamente las reglas gramaticales no estaría definida por el producto cartesiano $Y \times Z$, sino por un subconjunto G de él ($G \subset Y \times Z$), porque no todas las n-tuplas ordenadas de constituyentes tendrían imagen no vacía, sino sólo algunas de ellas. Por ello, no pueden representarse como un operación $X \rightarrow Y \times Z$, porque no es una ley de composición, sino sólo como una simple aplicación $X \rightarrow Y + Z$.

embargo, la gramática transformativa seguía presentando inadecuaciones a la hora de describir las lenguas naturales, que fueron solventadas en parte por Peters y Ritchie en 1973 con la hipótesis de su carácter recursivamente enumerable. Con ello, se produjo la revolución teórica más importante de la lingüística actual.

La contrastación de una teoría científica, que según el cuarto procedimiento de Popper⁴⁴ (1967: 32) consiste en «contrastarla por medio de la aplicación empírica de las conclusiones que pueden deducirse de ella», hizo enunciar sucesivas hipótesis *ad hoc* con la finalidad de poder adaptar las gramáticas generativas a las lenguas naturales. Una de esas hipótesis fue la misma teoría de las gramáticas transformativas. Puesto que la gramática transformativa tampoco fue capaz de describirlas completamente, se tuvieron que hacer algunos ajustes mediante otras hipótesis como la de Peters y Ritchie (1973) sobre el carácter recursivamente enumerable de las mismas. Estas hipótesis *ad hoc* han conseguido la descripción de las lenguas naturales y sus gramáticas al ir añadiendo más tipos de reglas gramaticales a las originalmente formuladas para los lenguajes formales, que son las reglas de reescritura aplicadas de forma recursiva. Ellas se convirtieron en las reglas sintagmáticas, a las que se añadieron las reglas de subcategorización, de inserción léxica y las reglas transformativas, cada una de ellas correspondiente

44. Karl Raimund Popper (1902-1994), especialmente en su obra *La lógica de la investigación científica* (*Logik der Forschung*, 1934, traducida al inglés en 1959), propone la búsqueda de un llamado *criterio de demarcación* entre la ciencia y la filosofía que permita decidir si una afirmación ha de ser estudiada y discutida dentro de la ciencia o, por el contrario, se sitúa en el campo más especulativo de la filosofía, alejándose así del positivismo dominante en el Círculo de Viena, que sólo distinguía entre las proposiciones contrastables, que son el objeto de la ciencia, y las que no lo son, que quedan en el campo de la filosofía o de la Teología. Según él, el conocimiento científico no avanza confirmando nuevas leyes, sino descartando leyes que contradicen la experiencia, que es lo que Popper llama *falsación*. Así, la labor del científico consiste en criticar leyes y principios de la naturaleza para reducir así el número de teorías compatibles con las observaciones experimentales de las que se dispone. El criterio de demarcación se define entonces como la capacidad de una proposición de ser refutada o falsada y sólo se admitirán como proposiciones científicas aquellas para las que sea conceptualmente posible un experimento o una observación que las contradiga. Por ejemplo, si se formula una ley que diga que todos los cisnes son blancos, esa ley es científica porque en cualquier momento puede aparecer un cisne negro que la contradiga. Por el contrario, la afirmación «Dios existe» no es científica porque no se puede contradecir mediante experimento ni observación alguna. La filosofía de Popper ha sido superada por otros autores importantes en el campo de la filosofía de la ciencia, entre ellos Thomas Samuel Kuhn (1922-1996) con su obra *La estructura de las revoluciones científicas* (1962). Hay que mencionar también a los españoles J. A. Díez y Carlos U. Moulines (1997), especialmente su obra *Fundamentos de filosofía de la ciencia*, y a J. A. Valor Yébenes, sobre todo la *Metodología de la investigación científica* (2000).

a un tipo de reglas de reescritura diferente (independientes del contexto o dependientes del contexto). A esas reglas, hay que añadir las reglas fonológicas, que fueron enunciadas a imagen y semejanza de las reglas sintácticas y, más concretamente, de las de subcategorización. La gramática transformativa consiste así en una serie de añadidos que han ido completando sucesivamente la teoría original y que han ido solucionando sobre la marcha los problemas que ésta iba presentando al ser aplicada⁴⁵.

Sin embargo, las transformaciones fueron cuestionadas por algunos lingüistas, y como consecuencia de ello surgieron otras alternativas. Como una escisión de la gramática generativa de Chomsky, surgieron las llamadas gramáticas de unificación y rasgos, que fueron desarrolladas por antiguos discípulos o colegas como Postal (la gramática relacional), Bresnan y Kaplan (LFG), Gazdar y Pullum (GPSG) y Pollard y Sag (HSPG). Todas ellas rechazan el concepto de *regla transformacional* y lo sustituyen por una operación universal, llamada *unificación*, que opera sobre *estructuras de rasgos* (Moreno Sandoval, 2001). Otras teorías gramaticales que se incluyen dentro de la lingüística matemática son las gramáticas categóricas⁴⁶, que han desarrollado

45. La lingüística matemática ha seguido manteniendo todo el aparato matemático de los lenguajes formales, correspondiente a la axiomatización de los lenguajes de la Lógica y de la informática, y lo ha seguido aplicando a las lenguas naturales siguiendo la corriente de pensamiento procedente de la Lógica. En efecto, la influencia de la axiomatización de los lenguajes de programación y de la incipiente traducción automática por ordenador en la posterior GGT es evidente (Chomsky, 1957a). Esta línea fue aceptada y desarrollada por Chomsky en un primer momento –a principios de la década de los 50– y abandonada sólo en parte por él a partir de 1955, cuando criticó las ideas de Carnap de tomar los lenguajes formales como modelo de la de las lenguas naturales y rechazó la propuesta de Bar-Hillel (1954) de seguir desarrollando el método de Carnap en la Lingüística (Chomsky, 1955b). Sin embargo, siguió utilizando la lógica simbólica para axiomatizar las teorías sintácticas del estructuralismo norteamericano y se basó en los trabajos de los teóricos de la lógica simbólica E. Post, P. Rosembloom y W. Quine para desarrollar su modelo matemático y la definición matemática de gramática generativa (Rohrer, 1978: 27-34), que expuso por primera vez en 1955 (Chomsky, 1955a).

46. Las gramáticas categóricas o categoriales fueron desarrolladas por el polaco Ajdukiewicz y después por Bar-Hillel, Lambek y otros. «En un sistema categórico existen justamente dos categorías gramaticales *fundamentales*, oración y nombre, que indicaremos a base de Σ y n , respectivamente. A todos los datos léxicos que no son nombres se les da, en el léxico, una clasificación categórica *derivada* con arreglo a la capacidad que presentan para combinarse entre sí o con una de las categorías fundamentales en la estructura de los constituyentes de las oraciones» (Lyons, 1968: 238). Las gramáticas categóricas se basan fundamentalmente en el concepto de «supresión», y se ayudan de las operaciones aritméticas, que toman como modelo: «En el tratamiento “quasi-aritmético” que hace Bar-Hillel, la clasificación categórica de elementos como *ran* puede expresarse como una “fracción” cuyo denominador indique con qué otra categoría puede combinarse *ran*, etc., y cuyo

un formalismo propio⁴⁷, y las gramáticas de dependencias⁴⁸, que presentan ciertas semejanzas con las categóricas⁴⁹ (Lyons, 1968: 237-242).

Los primeros problemas surgieron al aplicar las gramáticas transformacionales al procesamiento del lenguaje natural, ya que están concebidas para explicar la generación del lenguaje, mientras que la tarea más frecuente es la de analizar el lenguaje⁵⁰. Así, quedaban cuestionadas tanto las transformaciones como la estructura profunda ya que, el componente semántico fue equiparado a la estructura profunda, con lo que sólo quedó un nivel sintáctico: la estructura superficial⁵¹. Ya en los años 70, «destaca el trabajo de Peters y Ritchie (1973), que demostró que la teoría estándar formalmente equivalía

numerador denote la categoría de la construcción resultante. Así, la clasificación léxica de *ran*, etc., como Σ/n indica que tales elementos se combinan con nombres para formar oraciones. Una vez dada esta clasificación para *ran* y admitido que John es un nombre, sabemos que *John ran* es una oración gramaticalmente bien formada. Este hecho puede determinarse automáticamente por medio de una regla de “supresión” semejante a las que se emplean en aritmética, es decir, del mismo modo que $y \times x/y = x$, también $n \cdot \Sigma/n = \Sigma$ (esto es, ‘suprimimos’ el numerador y el denominador cuando son idénticos y, en este caso, nos quedamos con Σ , que indica que la expresión es una oración: el punto se utiliza aquí para expresar concatenación lineal)» (Lyons, 1968: 238-239).

47. «Otra forma de dar cuenta de las relaciones sintagmáticas entre los elementos de cada categoría compleja es definir estas categorías complejas por su función respecto de las categorías que las constituyen y respecto de otras categorías superiores. En este sentido, se ha desarrollado de una forma espectacular un formalismo categorial para dar cuenta de las categorías gramaticales.» (Moreno Cabrera, 1991: I, 106)
48. «Las relaciones de dependencia, cuando se representan directamente, se investigan independientemente de las estructuras de constituyentes que las realizan. Sin embargo, hoy día todavía no se ha desarrollado una teoría de la dependencia que goce de cierto acuerdo en la comunidad lingüística. Por ello, los estudios enunciados en esta corriente no son tan conocidos y citados como los que utilizan constituyentes.» (Moreno Cabrera, 1991: I, 90)
49. «Conviene llamar la atención sobre la “gramática de dependencia”, que presenta ciertas semejanzas con la categórica.» (Lyons, 1968: 238)
50. «Con todo ello, fue desde el campo aplicado del procesamiento del lenguaje natural de donde partieron las críticas más fundamentadas al modelo transformacional. Los primeros sistemas automáticos de los años 60, buscando modelos en la lingüística teórica, intentaron aplicar las ideas de Chomsky y sus colaboradores. El principal problema práctico que encontraron fue que la gramática transformacional está pensada como un proceso que parte de la estructura profunda para llegar a una estructura superficial. Dado que la tarea más frecuente de un procesador automático es la de analizar oraciones, el lingüista computacional tenía que invertir el proceso generativo de las transformaciones. Esto supuso enormes complicaciones.» (Moreno Sandoval, 2001: 15)
51. «Para Chomsky, la estructura profunda pertenecía a la sintaxis y es el punto de entrada para la interpretación semántica. Para los semantistas generativistas, en cambio, la estructura profunda era en sí misma la representación semántica, asumiendo que sólo hay un nivel sintáctico: la estructura superficial.» (Moreno Sandoval, 2001: 13)

a una máquina de Turing y, por tanto, la complejidad de procesamiento de las gramáticas transformacionales hace prohibitivo su uso en aplicaciones prácticas. Esto motivó la búsqueda de alternativas computables dentro de las gramáticas generativas. Por otra parte, desde una perspectiva psicolingüística se hacía evidente que un modelo tan costoso de computar era incompatible con el hecho de que los hablantes no necesitan un tiempo exponencial para procesar una oración» (Moreno Sandoval, 2001: 15).

La primera alternativa a las gramáticas transformacionales fue la gramática relacional (Perlmutter y Postal, 1977), que recuperó las funciones o relaciones gramaticales como elementos primitivos⁵². A partir de la gramática relacional y de las críticas de Peters y Ritchie (1973) a la gramática transformacional y al excesivo poder de las transformaciones, sugieron las gramáticas de unificación (Gazdar y Mellish, 1989). Siguiendo a Peters y Ritchie (1973), consideran que las gramáticas de las lenguas naturales son en realidad gramáticas independientes de contexto (GIC), y por eso parten de ellas, y no de las gramáticas dependientes de contexto (GDC), como hace la gramática transformacional, de manera que en las gramáticas relacionales la estructura sintáctica de las lenguas está representada mediante GIC⁵³. A partir de ellas, se utilizan los rasgos⁵⁴ y la unificación⁵⁵, que constituyen

-
52. «Se trata básicamente de una teoría sintáctica sobre la estructura de cláusulas que toma las relaciones gramaticales (sujeto, objeto directo, etc.) como elementos primitivos. Estas relaciones gramaticales, base de una teoría universal de la sintaxis, no pueden ser definidas universalmente en términos de estructuras de constituyentes y de orden lineal, como se asume en la gramática transformacional.» (Moreno Sandoval, 2001: 14)
53. «Desde el punto de vista de la actuación lingüística las gramáticas independientes de contexto son preferibles a las gramáticas transformacionales porque explican el procesamiento de emisiones lingüísticas de una manera más simple y eficiente. Desde el punto de vista de la competencia, también se podría argüir que cuanto más sencilla formalmente sea una lengua, más fácilmente podrá ser aprendida. Las teorías basadas en la unificación defienden que las lenguas naturales se puedan describir con gramáticas independientes de contexto, aumentadas en algunos casos con extensiones formales, como por ejemplo la utilización de rasgos» (Moreno Sandoval, 2001: 27).
54. «Un rasgo es un par atributo-valor donde: – un atributo es un símbolo (átomo o elemento constante sin estructura interna), que da nombre al rasgo, – un valor es o bien un símbolo o bien una estructura de rasgos (símbolo complejo con una estructura jerárquica interna)» (Moreno Sandoval, 2001: 41). Por su parte, «una estructura de rasgos es una función parcial de los atributos a sus valores» (Moreno Sandoval, 2001: 55), y «una función es una relación binaria unívoca que hace corresponder (*asigna*) a cada elemento del primer conjunto un solo elemento del segundo conjunto; además todos los elementos del primer conjunto están relacionados con algún elemento del segundo» (Moreno Sandoval, 2001: 247).
55. «La unificación es un proceso que combina la información de dos o más estructuras de rasgos para obtener una nueva estructura de rasgos que contenga toda la información de

el fundamento de ellas. Las gramáticas de unificación y rasgos rechazan el modelo de transformación, así como la estructura profunda, estableciendo un «nivel sintáctico único, sin representación profunda» (Moreno Sandoval, 2001: 14). Frente a las transformaciones, da mucha importancia al lexicismo o lexicalismo, ya que «algunos fenómenos (por ejemplo, las nominalizaciones) se explican mucho mejor con reglas léxicas que con transformaciones sintácticas» (Moreno Sandoval, 2001: 14).

La gramática transformacional nació tomando como base la lengua inglesa, y pronto surgieron problemas cuando se intentó aplicar a otras lenguas, especialmente a las que tipológicamente eran muy diferentes de ella⁵⁶. La gramática léxico-funcional (*Lexical Functional Grammar*, LFG) fue desarrollada por Bresnan y Kaplan (1982) debido a «la dificultad para adaptar el modelo transformacionista estándar a las lenguas con una estructura muy diferente a la del inglés (problema señalado por la gramática relacional)» (Moreno Sandoval, 2001: 79). Se caracteriza fundamentalmente por su carácter lexicista⁵⁷ y por que recupera las funciones sintácticas, igual que ya hizo la gramática relacional⁵⁸. La gramática de estructura sintagmática generalizada (*Generalized Phrase Structure Grammar*, GPSG), desarrollada por Gazdar, Pullum, Klein, Sag (1985) y otros lingüistas, también rechaza las transformaciones y se distingue de las demás porque incorpora el nivel semántico junto al sintáctico, prescindiendo de todos los demás niveles (fo-

ellas. La noción de unificación está estrechamente relacionada con la operación de unión de conjuntos, salvo que en el caso de la unificación el conjunto resultante debe ser una función» (Moreno Sandoval, 2001: 64).

56. Como norma general de cualquier teoría científica en general, y lingüística en particular, «la teoría debe tener en cuenta una gran cantidad de lenguas y no forzar a unas a entrar en el patrón de otras, en especial el inglés y las lenguas indoeuropeas» (Moreno Sandoval, 2001: 14).
57. «Cuando un determinado fenómeno lingüístico tiene varias interpretaciones, predomina la tendencia a preferir las soluciones en el lexicón y a reducir el papel de la gramática. El ejemplo clásico es el de interpretar las construcciones pasivas no como reglas de movimiento desde una posición estructural activa a otra pasiva (como hace la Gramática Transformacional), sino como un proceso léxico que relaciona las formas pasivas con las formas activas de los verbos» (Moreno Sandoval, 2001: 79).
58. «Las funciones sintácticas son modelos primitivos de la sintaxis. Se parte de la idea (prelucida por la gramática relacional) de que la sintaxis no trata únicamente de estructura en el sentido de reglas y árboles sintagmáticos, sino también de funciones (relaciones entre elementos). LFG toma los conceptos tradicionales como sujeto (SUJ), objeto (OBJ), etc., y los redefine, al tiempo que crea algunos nuevos, diferentes a los de la terminología tradicional.» (Moreno Sandoval, 2001: 80)

nológico, morfológico, pragmático)⁵⁹, y se caracteriza sobre todo por el uso de las *metarreglas*⁶⁰. Una evolución de la GPSG es la gramática sintagmática nuclear (*Head-Driven Phrase Structure Grammar*, HPSG), en la que Pollard y Sag (1987, 1994) eliminaron el carácter puramente «independiente de contexto» y las metarreglas⁶¹, que habían sido muy criticadas. Se caracteriza principalmente por haber añadido el nivel semántico a los niveles sintáctico y semántico de la GPSG⁶².

Como consecuencia de las críticas a la gramática transformacional y de la influencia que empezaron a tener las gramáticas de unificación y rasgos, Chomsky creó en los años 90 el programa minimalista, en el que introduce las principales características de las gramáticas de unificación y rasgos⁶³. Introduce el componente computacional, el concepto funcionalista de economía, el uso de rasgos y la combinación de la información de los rasgos, que es equivalente a la operación de unificación⁶⁴. Ha eliminado la estructura

-
59. La GPSG se debió a «la necesidad de eliminar las transformaciones de la gramática generativa, dejando gran parte del peso descriptivo en las reglas sintagmáticas (de ahí el nombre de esta teoría). Esto da como resultado un único nivel de estructura sintáctica». También se debió a «la necesidad de desarrollar un componente semántico plenamente formalizado, a la manera de la semántica de modelos. Dicho componente asume algunos fenómenos que la gramática transformacional atribuía a la sintaxis. Así, pues, Gazdar, Pullum, Klein, Sag y otros pocos se propusieron desarrollar una nueva teoría cuya sintaxis fuera no transformacional y cuya semántica estuviera basada en la de Montague» (Moreno Sandoval, 2001: 113). La GPSG «sólo considera los niveles sintáctico y semántico (es una teoría gramatical sin fonología)», y en ella «se rechaza la necesidad de transformaciones como mecanismo de descripción lingüística y se propone una versión ligeramente aumentada de las gramáticas independientes de contexto» (Moreno Sandoval, 2001: 126).
60. «Las metarreglas reflejan generalizaciones entre reglas gramaticales.» (Moreno Sandoval, 2001: 117)
61. «HPSG es una evolución de GPSG a partir de una implementación realizada por Pollard en Hewlett-Packard a mediados de los 80. Su innovación consistió en eliminar el carácter puramente “independiente de contexto” y las metarreglas.» (Moreno Sandoval, 2001: 131)
62. En la SPG, «el signo se representa a través de una matriz de rasgos con al menos dos atributos, PHON y SYNSEM. El valor de PHON incluye la interpretación fonética y fonológica. El atributo SYNSEM lleva toda la información referida a la sintaxis y semántica del signo (en la versión de 1987 este atributo estaba separado en dos, SYNTAX y SEMANTICS)» (Moreno Sandoval, 2001: 133).
63. «Al fin y al cabo, Chomsky es un maestro en la adaptación de las ideas de sus críticos.» (Moreno Sandoval, 2001: 204)
64. «Una de las innovaciones más destacadas es la aparición de un componente computacional que hace de interfaz con los otros dos componentes, el sistema fonológico y el sistema conceptual. Las estructuras se crean en el nivel computacional, pero para que sean válidas tienen que poder ser interpretadas por los interfaces que acceden a los sistemas externos,

profunda y la estructura superficial, muy criticadas desde los años 60, que ha sustituido por derivaciones, pero ha seguido siendo transformacional⁶⁵, a pesar de las críticas de que habían sido objeto las transformaciones en las dos décadas anteriores.

Las gramáticas de unificación no han conseguido solucionar definitivamente los problemas que presenta la aplicación de las gramáticas generativas a la descripción y funcionamiento de las lenguas naturales. Igual que la gramática transformacional, son incompletas y tienen importantes limitaciones, que afectan a la lingüística histórica, la dialectología y la sociolingüística⁶⁶, es decir, a las múltiples variaciones que presentan las lenguas a través del tiempo (diacronía), del espacio (dialectos) y de la sociedad (sociolectos), ya que su objeto de estudio es fundamentalmente la sincronía y, dentro de ella, la sintaxis. Sin embargo, ésta ha sido una limitación de las gramáticas generativas en general⁶⁷. Estos aspectos citados anteriormente no encajan muy bien en el concepto de gramática generativa, y esta inadecuación de las gramáticas generativas a la descripción y explicación de las lenguas naturales se puede explicar por el método deductivo que utilizan y que se encuentra en su misma esencia⁶⁸. De esta manera, se intentan encajar las características y los elementos particulares de las lenguas naturales en un modelo teórico

el fonológico y el semántico. Dentro del componente computacional (simplificado, el nivel sintáctico) se sustituyen las tradicionales estructura profunda y superficial por derivaciones. Las condiciones que se aplicaban en cada uno de los niveles mencionados han sido sustituidas por operaciones que funcionan a lo largo de toda la derivación. [...] En ocasiones, varias derivaciones compiten y aquí recurren a la teoría de la oportunidad: se elige la óptima, es decir, la más económica.» (Moreno Sandoval, 2001: 204)

65. «Podemos concluir que a pesar de los intentos de Chomsky por adaptar los principales presupuestos de las gramáticas de unificación y rasgos a su nuevo modelo, las innovaciones están limitadas por las posturas tradicionales de la escuela chomskyana: el programa minimalista sigue siendo transformacional. Y ésta sigue siendo la diferencia con respecto a las gramáticas de unificación» (Moreno Sandoval, 2001: 205).

66. «Las gramáticas de unificación son teorías parciales sobre el lenguaje y las lenguas naturales. Como tales, son incompletas ya que no pueden dar cuenta de todos los fenómenos relacionados con su objeto de estudio (nada se dice, por ejemplo, sobre la variación –diacrónica, dialectal, sociolectal– en el uso lingüístico).» (Moreno Sandoval, 2001: 1905)

67. «Otro de los puntos olvidados de las gramáticas generativas ha sido que las lenguas cambian en el tiempo y en el espacio y que no forman un todo homogéneo incluso en un mismo hablante. A estos aspectos siempre se les ha considerado periféricos, relacionados con el habla y la actuación, y como consecuencia se ha dado una falta de interés por los aspectos diacrónicos, dialectales y sociolingüísticos en las gramáticas generativas.» (Moreno Sandoval, 2001: 200)

68. «Estos modelos se caracterizan por intentar reflejar la competencia del hablante de manera deductiva.» (Moreno Sandoval, 2001: 197)