

ELEMENTOS DE LINGÜÍSTICA



J. A. ARGENTE GIRALT • S. BALARI RAVERA

A. BASTARDAS BOADA • J. BUSQUETS RIGAT

R. Cerdà MASSÓ • M.ª T. ESPINAL FARRÉ

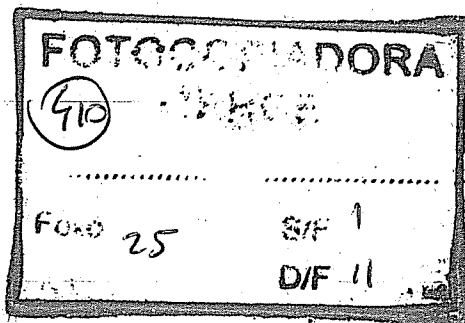
C. HERNÁNDEZ SACRISTÁN • Á. LÓPEZ GARCÍA

J. LLISTERRI BOIX • J. C. MORENO CABRERA

M. PRUÑONOSA TOMÁS • E. RIDRUEJO ALONSO

K. ROTAETXE AMUSATEGI • J. VIDAL VILLALBA

C. MARTÍN VIDE (ed.)



OCTAEDRO UNIVERSIDAD
TEXTOS

410--25

sobre procedimientos morfológicos, conceptos gramaticales y tipologías. Es una obra esencial de la que parten otros desarrollos posteriormente actual aunque sólo sea por los datos que aporta.

34), *Morfología generativa*, trad. J. Pazó. Alianza, Madrid, 1987.

fundamental por su exposición organizada y detallada de la perspectiva morfológica (morfología léxica).

38), «Inflection and derivation», *Linguistics*, 26, pp. 561-581.

claro y de síntesis sobre la distinción de estos mecanismos. Recomendado como ampliación.

A., S. (1990), *Fundamentos de morfología*. Síntesis, Madrid.

de introducción que sigue el método de la morfología léxica. Complemento de la de Scalise (1984).

(1986), *Aspectos sintáctico-semánticos en la sufijación*. Servicio de Ediciones, Universidad de Murcia, Murcia.

con buenos conocimientos en lingüística generativa, como mínimo, una buena perspectiva del estudio de la morfología en la teoría generativa en el campo de la sufijación con propuestas que adelantan en parte algunas actuales.

selección bibliográfica hay que añadir la referencia de la revista *Lexicologie*, donde se exponen trabajos sobre temas morfológicos sumamente convenientes para una ampliación en el análisis morfológico.

LA FRASE

por SERGIO BALARI RAVERA

Ojo: EL AUTOR REDEFINE CIERTOS SINTAGMAS O FRASES PREFOSICIONALES COMO FRASES NOMINALES.

En este capítulo presentamos algunas ideas y herramientas básicas para la construcción de un modelo de la sintaxis de las lenguas naturales. Después de situar la sintaxis en relación con otras disciplinas lingüísticas y en el marco de lo que podría ser un modelo general del lenguaje, nos ocupamos de tres aspectos esenciales que toda teoría sintáctica debe tratar: (a) las relaciones estructurales y la noción de constituyente, (b) las restricciones de subcategorización y la estructura argumental y (c) el problema del orden de palabras. Las secciones 2, 3 y 4 del capítulo se ocupan sucesivamente de estos aspectos, y al final de cada una se incluye un apartado especial en el que se presentan técnicas y herramientas formales para recoger los problemas empíricos descritos. Así se va construyendo progresivamente una pequeña gramática del castellano, capaz de recoger las restricciones de subcategorización de los núcleos, la concordancia sujeto-verbo y algunos aspectos del orden de palabras. Al final se incluyen unos ejercicios dedicados a continuar la labor de construcción de esta gramática y una breve bibliografía comentada.

1. INTRODUCCIÓN

Cuando los humanos usamos del lenguaje para comunicarnos, proferimos enunciados, que, entre otras muchas cosas, nos sirven para emitir juicios, hacer preguntas, dar órdenes o simplemente afirmar algo. El número de enunciados que podemos llegar a proferir a lo largo de nuestra vida es potencialmente infinito. Y, sin embargo, parece razonable suponer que la capacidad del lenguaje, localizada en alguna región de nuestro cerebro (con casi toda seguridad, en el hemisferio izquierdo), sólo dispone de recursos finitos para dar lugar a ese número potencialmente infinito de enunciados. Esta conclusión se sigue del hecho de que el lenguaje es sólo uno de los muchos procesos —como la memoria, el control motor o la percepción visual— que se llevan a cabo en nuestro cerebro, el cual, en tanto que objeto material regido por los mismos principios físicos que cualquier otro objeto material, tiene necesariamente una capacidad limitada de almacenamiento y procesamiento de datos.

En general, la solución que se ha propuesto a esta aparente paradoja es la de suponer que la capacidad del lenguaje consta, como mínimo, de un conjunto relativamente pequeño de reglas que determinan cuáles son las oraciones bien formadas de

una lengua, es decir, una *gramática*. Uno de los objetivos de la lingüística será, pues, dilucidar cómo se organiza y estructura la gramática de los humanos o, en otras palabras, elaborar un *modelo* de la gramática. Aquí, utilizamos «modelo» con el significado específico que se le concede en la filosofía de la ciencia, significado, por otra parte, no muy alejado de alguna de sus acepciones coloquiales. Así pues, suponemos la existencia de un determinado objeto real, la gramática, y, como investigadores, intentamos construir un modelo de ese objeto con la esperanza de acercarnos tanto como sea posible a la realidad.

Resulta de utilidad encuadrar la gramática en el seno de un modelo más general de *procesamiento* del lenguaje. De este modo, podremos presentar una idea más diáfana acerca de cuál es su papel y aclarar el significado de términos como el de *oración bien formada*, que dejamos sin definir un poco más arriba. Un esquema bastante simple de este modelo podría ser el de la figura 1:

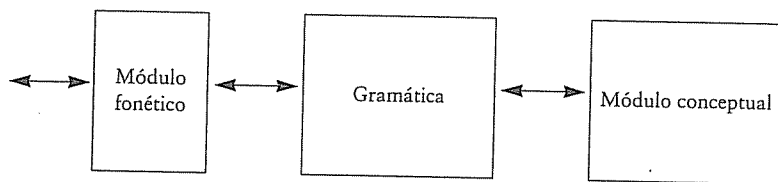


Figura 1

En el proceso de comprensión, el módulo fonético recibe información procedente del sistema auditivo, la depura, eliminando todo aquello que no es lingüísticamente relevante (como, por ejemplo, el ruido de fondo), y transforma la información restante en datos comprensibles para la gramática: una *representación fonética*. Ésta constituye la entrada de la gramática, cuyo cometido es producir una *representación semántica* que será interpretada en el módulo conceptual. El proceso de producción es el inverso: la *representación conceptual* producida en el módulo correspondiente es transformada por la gramática en una representación fonética, que el módulo fonético traduce en la información relevante para que los mecanismos articulatorios produzcan los sonidos apropiados. Como muestra el esquema, suponemos que cada módulo es capaz de operar en ambas direcciones. La direccionalidad del proceso depende únicamente de la naturaleza de la información de entrada. Ésta es la hipótesis más económica, ya que nos evita asumir, por ejemplo, la existencia de una gramática para la comprensión y otra gramática para la producción.

A la luz de este modelo, pues, podemos caracterizar la gramática como la entidad encargada de emparejar sonidos con significados. Esta caracterización precisa un poco más nuestra tarea de elaborar un modelo de la gramática: nuestro modelo será un conjunto de reglas que cumplan precisamente esa función de emparejar sonidos con significados. De acuerdo con esto y según hemos visto y veremos en este volumen (véanse los capítulos sobre fonología, morfología y semántica), algunas reglas de la gramática son fonológicas, otras reglas son morfológicas y finalmente otras son semánticas. Sin embargo, la fonología y la morfología no son suficientes para que ese emparejamiento entre sonidos y significados se produzca. Los enunciados que proferimos no suelen es-

tar formados por palabras aisladas, sino que generalmente incluyen grupos más o menos numerosos de palabras. Como ocurre con las unidades fonológicas y morfológicas, las palabras no se ordenan de forma arbitraria, sino que se organizan en unidades mayores siguiendo también unas reglas específicas.

Esas unidades mayores en que se organizan las palabras son las *oraciones* o *frases*, y el conjunto de reglas de la gramática encargadas de establecer cómo deben estructurarse las palabras dentro de una oración es lo que llamaremos *sintaxis*. No nos entretendremos aquí en dar una definición muy precisa de oración —para lo cual el lector puede consultar el capítulo 2 del manual de Hernanz & Brucart (1987)—; nos limitaremos simplemente a decir que es una entidad teórica, el resultado de la elaboración de un modelo de la gramática del hablante. Cuando construimos un modelo de la gramática, tratamos de establecer cuál es la forma y el grado de interacción de las diferentes reglas fonológicas, morfológicas, sintácticas, etc. Una manera de interpretar cuál es la función de las diversas reglas es concebirlas como una serie de instrucciones para construir una oración. Podríamos decir, pues, que una oración define una clase de la cual podemos encontrar numerosos ejemplos en la realidad, los enunciados; es decir, un enunciado como *Mañana lloverá* puede ser proferido por el hombre del tiempo un 14 de agosto de 1994, pero puede ser vuelto a proferir por el mismo individuo dos días más tarde, o puede haber sido proferido por cualquiera de los muchos hombres del tiempo que aparecen diariamente en los diversos canales de televisión. Cada una de estas preferencias de *Mañana lloverá* constituyen enunciados diferentes, pero todos nos remiten a una única oración.

Como hemos dicho, el número de oraciones que genera una gramática es enorme, quizá infinito, pero todas ellas son oraciones comprensibles para un hablante de la lengua; es decir, por muy largas y complejas que sean las oraciones, y por mucho que nos cueste interpretarlas, siempre las reconoceremos como oraciones de nuestra lengua, puesto que siguen fielmente las instrucciones de la gramática. Una gramática define, pues, sólo el conjunto de oraciones bien formadas o *gramaticales* de la lengua. En este sentido, es importante distinguir el concepto de gramaticalidad que acabamos de dar de otros conceptos que ocasionalmente se suelen asociar a las oraciones. *Gramaticalidad* es un concepto distinto de corrección en relación a una norma determinada; por ejemplo, una oración como *La di a María un libro* es normativamente incorrecta, pero gramatical en ciertos dialectos del español. Por otra parte, hallamos con frecuencia oraciones que por su longitud o complejidad nos resultan difíciles de interpretar, pero que no son agramaticales; en estos casos hablamos de *inaceptabilidad*, que es un concepto ligado al tiempo que necesitamos para procesar una oración determinada. La inaceptabilidad es, por tanto, un concepto útil para construir modelos de procesamiento, pero no necesariamente para construir modelos de la gramática.

El concepto de gramaticalidad será fundamental para nuestro estudio de la sintaxis. Efectivamente, si nuestro objetivo es el de construir un modelo de la gramática, la principal fuente de datos para aventurar hipótesis sobre los principios y reglas que ésta puede contener es precisamente nuestra capacidad, en tanto que hablantes, para emitir juicios sobre si una oración es gramatical o no. Con ello enunciamos un principio metodológico básico en lingüística: cuando emitimos un juicio de gramaticalidad estamos haciendo uso de nuestra capacidad del lenguaje, de nuestra gramática; por lo tanto, si nuestro objetivo es construir un modelo fiel de ella, entonces éste deberá recoger, con la mayor fidelidad posible, tales juicios. Ello significa, siguiendo con la idea de que las re-

glas de la gramática son comparables a instrucciones para construir una oración, que las reglas de nuestro modelo deberán ser capaces de construir aquellas oraciones que hayamos juzgado como gramaticales, pero no aquellas que hayamos juzgado como agramaticales. Nótese, por tanto, que daremos tanta importancia a los datos positivos (las oraciones gramaticales) como a los datos negativos (las oraciones agramaticales).

Con lo dicho hasta ahora, ya estamos en disposición de abordar el estudio de la sintaxis con garantías suficientes. No quisiéramos, sin embargo, poner el punto final a esta introducción sin antes mencionar algunos aspectos que se podrían echar de menos.

En primer lugar, hemos asumido que una de nuestras tareas como lingüistas será la de construir un modelo de la gramática, la cual hemos identificado con la capacidad del lenguaje. No obstante, esto no tiene por qué ser necesariamente así y, de hecho, sería injusto presentarlo como la única alternativa. El considerar la gramática como un objeto mental o psicológico no es una posición unánime en lingüística, aunque quizá sí sea mayoritaria. No es éste el lugar para presentar las diferentes posturas filosóficas que existen en la lingüística contemporánea —para lo cual me remito al libro de Katz (1985) y a los artículos allí recogidos—, pero téngase en cuenta que puede haber diferentes concepciones sobre qué tipo de objeto real es la gramática que estamos modelando.

En segundo lugar, y dada la concepción que hemos esbozado aquí, es evidente que cada gramática es un objeto individual y que cada individuo, en tanto que hablante, posee su propia gramática. De acuerdo con esta hipótesis, dos individuos que hablen una misma lengua harán uso de gramáticas idénticas o, en todo caso, muy parecidas, mientras que hablantes (monolingües) de lenguas diferentes poseerán gramáticas diferentes. Un aspecto importante de la investigación lingüística será, por tanto, el de determinar en qué se parecen y en qué se distinguen las gramáticas de hablantes diferentes, sean éstos hablantes de dialectos de una misma lengua o hablantes de lenguas distintas.

Este punto nos lleva al último asunto que queremos tratar aquí: el problema de la adquisición del lenguaje. Éste ha sido uno de los motores de la investigación lingüística contemporánea y el responsable, al menos en parte, de la concepción mentalista de la gramática que esbozamos más arriba. Cuando elaboramos un modelo de una gramática determinada —del español, por ejemplo—, no cabe duda de que estamos tratando con la gramática de un hablante maduro; un recién nacido no es un hablante del español ni de ninguna otra lengua, ya que es hablante potencial de cualquier lengua. Es lícito, pues, preguntarse cómo llegamos a ese estado final de la gramática desde nuestra condición de hablantes potenciales de cualquier lengua. No vamos a responder aquí a esta pregunta (no tenemos una respuesta definitiva), ni tampoco vamos a detenernos a exponer con detalle las hipótesis que se han formulado en este sentido. Pese a lo expuesto hasta aquí, concebimos este capítulo como una breve introducción a los métodos y herramientas útiles para construir modelos de la gramática (concretamente, de aquella parte que hemos denominado sintaxis), independientemente de cuál sea nuestra concepción de la misma. Creemos que en este nivel se puede ser (casi) neutral.

Con este objetivo en mente, hemos estructurado el presente artículo de la manera siguiente. Cada uno de los apartados que siguen trata un aspecto esencial para el estudio de la sintaxis de las lenguas naturales. El apartado 2 trata de la noción de constituyente e introduce las técnicas básicas para la representación de estructuras sintácticas; el apartado 3 trata de las funciones y relaciones gramaticales y de las diferentes estrategias a que recurren las lenguas naturales para expresarlas; finalmente, el apartado 4 trata el problema del orden de palabras. En cada apartado hemos intentado introducir con la

mayor claridad posible todas aquellas herramientas y conceptos teóricos esenciales para el estudio de la sintaxis, así como observaciones concernientes a las diferencias tipológicas que presentan las diversas lenguas del mundo. Sin embargo, no hemos querido perder de vista en ningún momento el objetivo final de todo lingüista, que es el de construir un modelo de la gramática, y para construir un modelo es preciso escribir reglas para analizar su funcionamiento y sus interacciones. Por este motivo, cada uno de los apartados incluye un subapartado final en el que se va construyendo paso a paso una pequeña gramática del español que incorpora en cada momento aquellos fenómenos que hemos estudiado en subapartados anteriores. Así esperamos cubrir tres de los aspectos fundamentales del estudio de la sintaxis: el descriptivo, el tipológico y el formal.

2. UNIDADES DE ANÁLISIS EN SINTAXIS

2.1. LOS CONSTITUYENTES INMEDIATOS

En el apartado anterior observamos que las oraciones son agrupaciones de palabras y que éstas se ordenan siguiendo unas reglas específicas. Esta observación, sin dejar de ser del todo cierta, es un tanto engañosa, ya que parece comportar la idea de que situar una palabra detrás de otra en un orden determinado es suficiente para construir una oración. Si así fuera, una oración como *Gabriel está en Roma* tendría una representación sintáctica como la de (1):

- (1) Gabriel + está + en + Roma,

Donde el símbolo «+» indica que dos palabras están yuxtapuestas (o *concatenadas*). En apariencia, ésta es la teoría sintáctica más simple, pues en principio sólo necesitamos una regla para concatenar palabras, además de una lista con todas las palabras de la lengua. Es sencillo, sin embargo, demostrar que esta primera intuición no es la correcta. En el apartado anterior también dijimos que la gramática debía permitirnos construir todas las oraciones bien formadas de la lengua, *pero* ninguna de las mal formadas. Es el segundo elemento de la coordinación el que es importante en este caso. Si tomamos la oración de (1) y alteramos el orden de sus elementos, veremos que producimos secuencias gramaticales pero también secuencias agramaticales:

- (2) a. Gabriel + está + en + Roma.
 b. Está + Gabriel + en + Roma.
 c. En + Roma + está + Gabriel.
 d. *Roma + Gabriel + está + en.
 e. *En + está + Roma + Gabriel.
 f. *Gabriel + en + está + Roma.

Ninguna de las secuencias (2.d-f) es gramatical en español (eso es precisamente lo que indica el asterisco que las precede y que utilizaremos siempre que queramos indicar agramaticalidad) y, no obstante, todas ellas están de acuerdo con la única regla de nuestra hipotética gramática. Por otro lado, las secuencias (2.a-c) sí son gramaticales en español. ¿Qué debemos hacer para que nuestra sintaxis permita (2.a-c) y prohíba (2.d-f)? Quizá con añadir nuevas reglas de concatenación más refinadas sea suficien-

te, aunque podemos adelantar que tal tarea resultaría infructuosa: pronto el número de reglas será tan enorme que tendremos serios problemas para manejar nuestra gramática. Quede para el lector la tarea de comprobarlo. No olvide que la gramática debe dar cuenta de *todas* las oraciones de la lengua y no solamente de *Gabriel está en Roma*.

Existen otros motivos, además de los aducidos, para rechazar nuestra primera teoría sintáctica. Si comparamos las secuencias gramaticales y agramaticales de (2) veremos que parece haber una manera de alterar el orden de los elementos sin crear secuencias agramaticales: tratando la secuencia *en Roma* como si fuera una unidad. Efectivamente, si procedemos de acuerdo con este criterio, siempre obtenemos una secuencia gramatical (aunque quizá no del todo aceptable):

- (3) a. Gabriel + está + [en + Roma].
- b. Está + Gabriel + [en + Roma].
- c. [En + Roma] + está + Gabriel.
- d. Gabriel + [en + Roma] + está.
- e. [En + Roma] + Gabriel + está.

Con los corchetes queremos indicar que *en Roma* es un elemento con estructura interna, pero capaz de funcionar como una unidad. Esto es una pista de que quizá sea preciso reconocer unidades intermedias entre la oración y la palabra, lo cual comporta necesariamente abandonar la idea de que la oración tiene una estructura lineal, ya que para obtener el resultado de (3) hemos tenido que construir previamente la unidad [en + Roma] mediante concatenación y, después, la hemos concatenado con las otras dos unidades. De hecho, hemos seguido un procedimiento parecido al que describimos en (4), donde O simboliza la oración:

- (4) I. en + Roma = A.
- II. Gabriel + está + A = O.

Podemos reconstruir estos pasos gráficamente, en forma de árbol invertido, de modo que las ramas más bajas del mismo representen el paso I en la construcción de la oración, hasta llegar al nodo superior, que representa la oración completa una vez aplicado el paso II (figura 2):

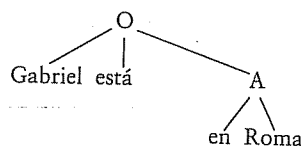


Figura 2

Si tomamos este árbol como una representación de la estructura de la oración, vemos que ésta no es únicamente lineal, sino también jerárquica, ya que nos vemos obligados a reconocer unidades de orden intermedio como la que aparece etiquetada como

A en la figura. Esta idea equivale a reconocer que dentro de una oración ciertos elementos mantienen un vínculo más estrecho que otros, y que se agrupan formando unidades complejas que a su vez pueden agruparse con otras unidades mayores formando una estructura jerárquica de varios niveles hasta formar una oración completa. A estas unidades sintácticas inferiores las denominaremos *constituyentes* de la oración. En adelante, además, adoptaremos los diagramas arbóreos como medio más práctico para representar las estructuras sintácticas, principalmente porque nos permiten visualizar con bastante claridad las relaciones lineales y jerárquicas que se establecen entre los diferentes constituyentes. Asimismo, cuando hablemos sobre los árboles, utilizaremos la terminología corriente en estos casos, y denominaremos *nodos* a aquellos puntos (generalmente marcados con una etiqueta) de los que nacen nuevas *ramas* y *hojas* a aquellos puntos de los que no nace ninguna nueva rama.

Hasta aquí, hemos demostrado la existencia de constituyentes. Sin embargo, dado que la determinación de la estructura de una oración es un aspecto importante en el estudio de la sintaxis, conviene establecer algunos métodos para determinar si una secuencia es o no un constituyente. Afortunadamente, los constituyentes poseen algunas propiedades que nos serán bastante útiles para identificarlos. De hecho, ya hemos visto una de esas propiedades en el ejemplo (3): cuando alteramos el orden de una oración simple, las piezas que movemos siempre son constituyentes. Así, el constituyente *en Roma* puede ocupar diversas posiciones dentro de la oración siempre que no rompamos su unidad ni alteremos el orden relativo de sus constituyentes. Más adelante, en el apartado 4 volveremos sobre este asunto.

Otra propiedad es que sólo un constituyente puede funcionar como respuesta a una pregunta. Así, cualquiera de las respuestas de (5) es aceptable para una pregunta como *¿Dónde está Gabriel?*, pero ninguna de las de (6) lo es:

- (5) a. En Roma.
- b. Está en Roma.
- c. Gabriel está en Roma.
- (6) a. *Está en.
- b. *Gabriel está en.
- c. *Gabriel está.

Obsérvese que esta prueba nos demuestra que la estructura de *Gabriel está en Roma* es más rica que la que le atribuimos en la figura 2, ya que *está en Roma* también es un constituyente (mientras que *Gabriel está* no lo es). Subsana mos, pues, el error con un nuevo árbol como el de la figura 3:

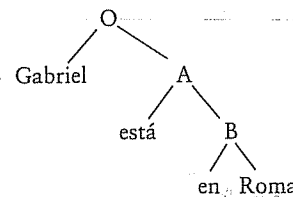


Figura 3

Finalmente, otro criterio que suele funcionar es el denominado criterio de la distribución. De acuerdo con este principio, todo constituyente es una unidad sintáctica compleja sustituible por otra unidad del mismo tipo pero más simple y, viceversa, un constituyente simple puede ser sustituido por otro de mayor complejidad. Sin profundizar de momento en el problema de qué queremos decir con «unidad del mismo tipo», noción que por el momento podemos mantener a un nivel intuitivo, vemos que, efectivamente, cada uno de los constituyentes que aparecen en la figura 3 es sustituible por otro elemento sin que la gramaticalidad de la oración se vea afectada. En (7) tenemos un pequeño ejemplo de ello, donde los elementos entre corchetes son constituyentes que han sustituido a otros:

- (7) a. Gabriel está en Roma.
 b. [Mi hermano] está en Roma. (*Gabriel por mi hermano*)
 c. [La [hermana de Laura]] está en Roma. (*Mi por La y hermano por hermana de Laura*)
 d. [La [hermana de Laura]] [ha estado] en Roma. (*está por ha estado*)
 e. [La [hermana de Laura]] [ha estado] en [Génova]. (*Roma por Génova*)
 f. [La [hermana de Laura]] [ha estado] [aquí]. (*en Génova por aquí*)
 g. [La [hermana de Laura]] [volverá]. (*ha estado aquí por volverá*)

La prueba de la distribución nos aporta nueva información sobre los constituyentes y nos permite, además, perfeccionar un poco más el sistema de representación que hemos empezado a utilizar. En primer lugar, observamos que el constituyente *la hermana de Laura* tiene dos constituyentes inmediatos, *la* y *hermana de Laura*, y no *la hermana* y *de Laura* como quizá pueda haber parecido en un primer momento. Hablamos de constituyentes inmediatos porque en *la hermana de Laura* también tenemos los constituyentes *la*, *hermana*, *de Laura*, *de* y *Laura*, pero no pertenecen al nivel jerárquico inmediatamente inferior como los anteriores. En segundo lugar, la prueba de la distribución nos demuestra que una sola palabra puede ser distribucionalmente equivalente a un constituyente más complejo, lo que indica la conveniencia de suponer que siempre existe un nivel jerárquico inmediatamente superior a los nodos del árbol ocupados por palabras. Un ejemplo de árbol que incorpora esta idea es el que tenemos en la figura 4:

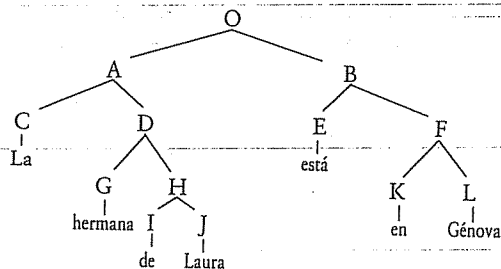


Figura 4

Hasta ahora hemos centrado nuestros esfuerzos en demostrar que la estructura sintáctica de una oración no es únicamente lineal, sino también jerárquica. Hemos aportado pruebas en favor de la existencia de los constituyentes y hemos presentado algunos criterios para identificarlos. También hemos introducido una convención para representar la estructura sintáctica de las oraciones: el diagrama arbóreo. Los árboles son útiles porque nos permiten visualizar las diferentes relaciones lineales y jerárquicas que se establecen entre los constituyentes de la oración; nos falta, sin embargo, la terminología necesaria para hablar de estas relaciones. ¿Qué relación existe, por ejemplo, entre el nodo A y el nodo H del árbol de la figura 4? ¿Y entre el nodo A y el B? Todo árbol puede describirse en términos de dos relaciones estructurales entre pares de nodos, una relación jerárquica y una relación lineal. La primera es la relación de *dominio* y la segunda es la relación de *precedencia*. Así, por ejemplo, el nodo A domina al nodo H (pero H no domina a A) y el nodo A precede al nodo B (pero B no precede a A); es decir, ambas relaciones poseen la propiedad *asimétrica* —de hecho, la relación de dominio es *antisimétrica*, ya que se considera conveniente asumir que la relación posee también la propiedad *reflexiva*; es decir que un nodo se domina a sí mismo—. Ambas relaciones poseen además la propiedad *transitiva*, es decir, si un nodo domina a otro, también domina a todos los nodos dominados por el segundo; lo mismo ocurre con la precedencia. Por ejemplo, A domina a D y D domina a H, entonces A domina también a H. Con esta información, podemos realizar una descripción completa del árbol de la figura 4 exclusivamente en términos de las relaciones de dominio y de precedencia. Obsérvese que, al establecer las propiedades de ambas relaciones, podemos hacer una descripción más económica; por ejemplo, no es necesario decir que A domina a H, ya que es suficiente con especificar que A domina a D y que D domina a H.

Hasta ahora nos hemos mantenido a un nivel bastante formal y, quizá, poco lingüístico. Es evidente, por ejemplo, que los símbolos que utilizamos para etiquetar los nodos del árbol son meras convenciones que no recogen el hecho de que un constituyente como *Gabriel* tiene la misma distribución que *la hermana de Laura*. Es decir, ambos constituyentes pertenecen a una misma clase o *categoría* sintáctica. ¿Cómo podemos incorporar esta idea? ¿De qué manera podemos identificar y caracterizar las diferentes categorías sintácticas? En el siguiente subapartado abordamos estas cuestiones.

2.2. CATEGORÍAS SINTÁCTICAS

Quizá la mejor manera de responder a las preguntas con que terminábamos el subapartado anterior consista en dar previamente respuesta a otra cuestión: ¿qué tipo de elementos son los que aparecen en las hojas del árbol? Estos elementos —también denominados *terminales* porque poseen la propiedad de no dominar a ningún otro elemento— hasta el momento han aparecido representados como palabras ortográficas. Está claro, no obstante, que esto no es más que una convención a la que recurrimos para simplificar nuestras representaciones sintácticas. Lo más natural, dado que estamos construyendo un modelo de la gramática, es que estos elementos sean palabras construidas de acuerdo con las reglas de los componentes fonológico y morfológico de la gramática, es decir, *palabras gramaticales*. Una palabra gramatical puede ser concebida como un conglomerado de información de varios tipos: fonológica, semántica, morfológica. Sabemos que toda palabra posee una representación fonológica y un sig-

nificado. Sabemos, también, que apelando a criterios morfológicos podemos asociar a cada palabra un cierto número de propiedades, como por ejemplo que pertenece a una determinada categoría y que tiene asociados unos determinados rasgos flexivos.

Así pues, utilizando diversos criterios morfológicos, asignamos las palabras gramaticales a las diferentes categorías: nombre, verbo, adjetivo, preposición, etc. De acuerdo con esto, podemos imaginar una representación esquemática de una palabra gramatical que contenga toda esta información asociada a la información fonológica correspondiente:

$$(8) \left[\begin{array}{l} \text{FONOLOGÍA = /ermana/} \\ \text{MORFOSINTAXIS = } \left[\begin{array}{l} \text{CATEGORÍA = N} \\ \text{FLEXIÓN = } \left[\begin{array}{l} \text{GÉNERO = fem} \\ \text{NÚMERO = sing} \end{array} \right] \end{array} \right] \end{array} \right]$$

En (8) ofrecemos una posible representación de la palabra gramatical «hermana» en forma de matriz de rasgos compleja. Como habrá podido comprobar el lector, hemos adoptado la convención de organizar la información de forma jerárquica y haciendo uso de un sistema de rasgos algo más elaborado del que se suele utilizar en las representaciones fonológicas. Así, tenemos dos grandes tipos de información lingüística, la fonológica y la morfosintáctica. El primer tipo, simplificando mucho, consideramos que es una representación fonológica (ilustrada aquí mediante una transcripción fonológica aproximada); el segundo es más complejo: la información que hemos denominado morfosintáctica puede clasificarse a su vez en información categorial (en este caso, la palabra es un Nombre) e información flexiva, la cual también es compleja y posee, en este caso, dos rasgos: el género y el número. Más adelante veremos cómo podemos sacar partido de esta manera de representar la información lingüística para establecer generalizaciones gramaticales; por el momento, nos basta con considerar que (8), o algo parecido a (8), es lo que ocupa la posición del elemento terminal «hermana» en la representación sintáctica de (7):

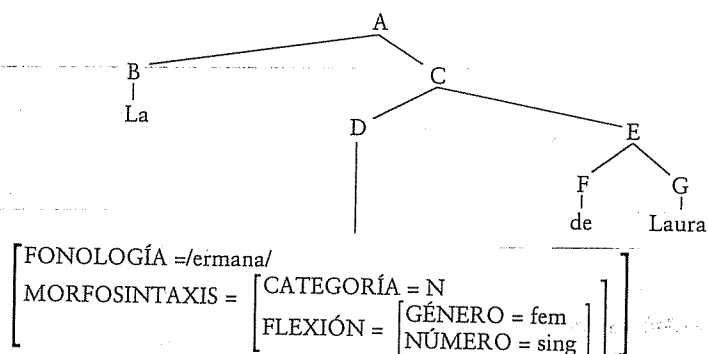


Figura 5

Como pudimos comprobar con la prueba de la distribución de (7), los diversos nodos en el árbol etiquetados con las letras A, C y D corresponden a elementos de una misma categoría, es decir, A es un elemento nominal como *Gabriel* o como *la hermana de Laura*; C es también nominal, ya que junto a la palabra *la* pueden aparecer elementos simples como *hermana* o elementos más complejos como *hermana de Laura*; finalmente, D, al ser el nodo que domina inmediatamente al elemento terminal, debe ser también un nominal. Parece, pues, como si parte de la información contenida en el terminal, en particular aquella información relacionada con la categoría, «subiera» árbol arriba siguiendo el camino D—C—A. Así pues, parece razonable que recojamos este hecho en nuestra representación asignando una etiqueta común a cada uno de los nodos; en este caso, suele adoptarse la convención de utilizar las etiquetas categoriales N para el nodo D, SN para el nodo C y también SN (por Sintagma Nominal) para el nodo A. Visto ahora desde esta perspectiva ligeramente distinta, quizá sea más intuitivo pensar que el camino N—SN—SN se proyecta desde el nodo terminal, como un árbol que germina y va extendiendo sus ramas a medida que crece. Siguiendo con nuestra metáfora, y adoptando la práctica habitual en sintaxis, podemos denominar *núcleo* al nodo desde el que se proyecta la estructura y *proyecciones* de este núcleo a cada uno de los nodos del camino N—SN—SN, donde SN es la *proyección máxima* de N. (N, en tanto que elemento preterminal, puede denominarse también núcleo aunque, según la explicación que aquí hemos dado, estrictamente no lo sea.)

Evidentemente, lo dicho para los nombres vale también para las demás categorías, adjetivo (A), verbo (V) y adverbio (Adv), de modo que también hablaremos de SA, SV y SAdv, con sus respectivos núcleos y proyecciones intermedias. Un aspecto importante de las estructuras sintácticas como la de la figura 5, al que dedicaremos más espacio en el siguiente apartado, es que define un patrón estructural que, aparentemente, todas las categorías siguen con bastante fidelidad. Así pues, si queremos ser precisos, el árbol de la figura 5 debería ser como lo representamos en la figura 6:

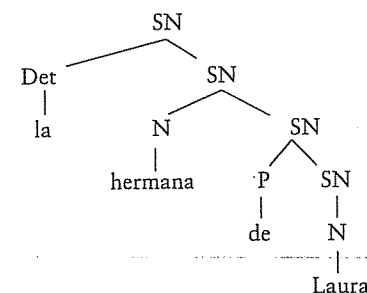


Figura 6

En este esquema introducimos dos nuevas etiquetas categoriales para los símbolos preterminales que dominan al artículo *la* y a la preposición *de*, respectivamente. Con la etiqueta Det (léase *determinante*) queremos designar todos aquellos elementos que acom-

pañan al nombre, como los artículos, los demostrativos, los numerales, los posesivos, etc., que se comportan de manera parecida en cuanto a su distribución dentro del SN; con la etiqueta P designaremos todos aquellos elementos dentro de la clase de las preposiciones.

Hay algunos aspectos de la figura 6, relacionados con estos elementos, que no podemos pasar por alto: entramos así en el terreno un tanto controvertido de la tipología de las categorías sintácticas, en el que nos detendremos un poco, aunque no podamos dedicarle la atención que merecería (para un excelente repaso de los problemas que suscita esta cuestión puede consultarse, no obstante, Bosque, 1989). Los determinantes, junto con las conjunciones subordinantes (o, por utilizar un nombre más común hoy en día, los *complementadores*, como *que* o *si*), las preposiciones, las conjunciones coordinantes (*y*, *ni*, *o*, *pero*, etc.), ciertos adverbios de grado (*más*, *menos*, *muy*, etc.) y algunos otros elementos forman la clase de lo que se ha venido a denominar *categorías menores* o *funcionales*, frente a la clase de las *categorías mayores* o *léxicas*, integrada por los nombres, los verbos, los adjetivos y los adverbios. Esta clasificación es controvertida (sin duda más de uno no estará de acuerdo con ella), aunque como en seguida veremos es posible justificarla.

Históricamente, todas las tradiciones gramaticales han supuesto la existencia de una distinción más o menos parecida a la que hemos establecido aquí entre categorías funcionales y léxicas (pues éstos son los términos que hemos decidido utilizar), basándose en criterios de muy diversa índole. Por ejemplo, la tradición grecolatina distinguía entre *categorías variables* (nombres, verbos, adjetivos) e *invariables* (preposiciones, adverbios, conjunciones), trazando así una frontera entre las categorías que poseían flexión y aquellas que no la poseían; otras tradiciones han querido distinguir entre *clases abiertas* y *clases cerradas* a partir de la observación de que en las lenguas ciertas clases de palabras son abiertas en el sentido de que admiten nuevas incorporaciones, mientras que otras clases de palabras son cerradas, tienen comparativamente pocos miembros y su número permanece fijo; en efecto, en las lenguas es normal que se creen (y que caigan en desuso) nombres o verbos, mientras que la aparición (o la desaparición) de una preposición es un fenómeno mucho más raro y que comporta un cambio radical en la estructura de la gramática. El problema de este tipo de clasificaciones es que tienen un alcance muy limitado, ya que las propiedades de ser variable o invariable o de ser una clase abierta o cerrada no son uniformes en las lenguas del mundo: por ejemplo, en inglés la clase de los adjetivos es abierta como en castellano, pero los adjetivos ingleses son invariables; en latín los artículos no existían, mientras que en castellano ésta es una clase cerrada de categorías variables; muchas lenguas africanas poseen una clase cerrada de adjetivos variables; las lenguas celtas (como el galés, el irlandés o el bretón) poseen clases cerradas de preposiciones variables, al contrario de lo que ocurre en la mayoría de lenguas, donde las preposiciones forman una clase cerrada de elementos invariables. Podríamos hallar un sinnúmero de ejemplos de este tipo.

Para los propósitos de este capítulo, pues, utilizaremos un criterio intuitivo a la hora de distinguir las categorías léxicas de las funcionales: las primeras son aquellas que poseen un contenido semántico, es decir, las palabras que pertenecen a una categoría léxica denotan entidades en el mundo (individuos, objetos, etc., los nombres, y acciones, estados, procesos, los verbos) o propiedades de estas entidades (adjetivos y adverbios); por otra parte, las categorías funcionales tienen un valor relacional y, por tanto, poseen un contenido eminentemente gramatical: son los elementos necesarios

para poner en relación los diferentes constituyentes de la oración; por utilizar una metáfora constructiva, si las categorías léxicas son los ladrillos de la sintaxis, las categorías funcionales son su cemento.

2.3. GRAMÁTICA I

Por ahora, en este apartado hemos desarrollado un primer esbozo de teoría sintáctica, identificando las unidades constitutivas de las oraciones (los constituyentes) y estableciendo una clasificación de las mismas dentro de distintas categorías sintácticas. Nos faltan, sin embargo, las herramientas formales necesarias para construir nuestro modelo de la gramática. Ya hemos visto qué forma queremos que tengan nuestras representaciones sintácticas: los diagramas arbóreos, pero nos falta hallar una manera de describir las reglas de la gramática responsables de construir esos árboles. Tenemos, no obstante, una pista de cómo pueden ser estas reglas. Recordemos que en (4) describimos un procedimiento para construir el árbol de la figura 2 y, de hecho, no precisamos más que refinar un poco más la manera de describir ese procedimiento para alcanzar nuestro objetivo haciendo uso de los símbolos que hemos ido introduciendo para representar las categorías sintácticas. Por ejemplo, en la figura 4 vemos que una oración tiene dos constituyentes inmediatos, un SN y un SV, lo cual, siguiendo la convención más común, puede expresarse mediante la regla de (9):

$$(9) \quad O \rightarrow SN \quad SV.$$

Como vemos, la regla expresa de forma lineal las dos relaciones estructurales fundamentales de todo diagrama arbóreo: el símbolo a la izquierda de la flecha representa la categoría que domina a las dos categorías a la derecha de la flecha; la relación de precedencia se expresa directamente mediante el orden lineal de los símbolos a la derecha de la flecha. Así pues, (9) es una regla para construir el árbol de (10):

$$(10) \quad \begin{array}{c} O \\ \diagdown \quad \diagup \\ SN \quad SV \end{array}$$

Reglas como la de (9) se conocen con el nombre de *reglas de estructura de frase* o *reglas de estructura sintagmática*. De forma similar, hemos visto que hay diversas posibilidades para construir un SN; las expresamos mediante reglas en (11):

- (11) a. $SN \rightarrow Det \quad SN.$
- b. $SN \rightarrow P \quad SN.$
- c. $SN \rightarrow N \quad SN.$
- d. $SN \rightarrow N.$

Y lo mismo podemos hacer con el SV:

- (12) a. $SV \rightarrow V \quad SN.$
- b. $SV \rightarrow V.$

Sólo nos faltan las reglas para introducir las palabras:

- (13) a. N → {hermana, casa, Roma, Laura, libro, pan, ...}.
 b. V → {estar, saltar, toser, comer, leer, ...}.
 c. P → {a, en, de, por, para, con, ...}.
 d. Det → {el, la, los, las, un, una, ...},

donde introducimos la convención de escribir entre corchetes las diferentes alternativas de que disponemos; es decir, la regla (13.a), por ejemplo, es una manera abreviada de expresar seis reglas, una para cada una de las palabras que aparecen entre corchetes. Asimismo, también para ahorrar espacio, escribimos sólo las formas en infinitivo de los verbos, pero ello debe interpretarse como una forma abreviada de representar todo el paradigma.

Hechas estas consideraciones, podemos comprobar que la minigramática que acabamos de construir es capaz de generar un buen número de oraciones gramaticales del castellano; no todas, porque aún es incompleta, pero sí oraciones como: *La hermana de Laura está en Roma, La casa de la hermana de Laura está en Roma, Laura lee un libro, etc.* ¿Estamos seguros, sin embargo, de que no genere oraciones agramaticales?

3. FUNCIONES, RELACIONES GRAMATICALES Y SUBCATEGORIZACIÓN

Nuestra gramática tiene, entre muchos otros defectos, uno bastante grave. De entre las oraciones agramaticales que genera destacamos las de (14):

- (14) a. *La casa la hermana Laura está en Roma.
 b. *Laura tosió un libro.
 c. *Laura estaba.
 d. *Laura come de pan.
 e. *Laura está casa.

Cada uno de los ejemplos de (14) ilustra un caso ligeramente distinto de una serie de fenómenos relacionados entre ellos y que podríamos caracterizar intuitivamente de la manera siguiente: en cada una de las oraciones de (14) no se respeta alguna de las restricciones que un núcleo impone sobre el número y la naturaleza de los elementos que pueden acompañarlo.

Abundemos un poco en esta cuestión estudiando alguno de los casos con más detalle. Por ejemplo, en (14.a) el problema está en la secuencia *la casa la hermana Laura*, que, pese a poderse construir utilizando las reglas de la gramática, no es un SN bien formado en castellano. ¿Por qué? Porque hemos omitido por dos veces la preposición *de*, que es obligatoria en esta situación en que queremos expresar ciertas relaciones entre los núcleos y los elementos que los acompañan. Por ejemplo, el nombre *hermana* (como muchos otros nombres relacionales como *padre, madre, amigo, etc.*) exige la presencia de un SN marcado con *de*. Lo mismo ocurre entre el nombre *casa* y el SN mediante el cual expresamos el poseedor de la casa: *la casa de la hermana de Laura*.

En (14.b) tenemos un problema ligeramente distinto pero, como veremos, también relacionado con el anterior. Aquí —y siguiendo con una caracterización intuitiva del

problema— lo que ocurre es que tenemos demasiados elementos: el verbo *toser* puede ir acompañado por un SN que exprese *quién* ha tosidó, así como podría ir acompañado por constituyentes que expresaran *cuándo, dónde o por qué* lo hizo. Sin embargo, nunca puede ir acompañado por un SN que exprese *qué* ha tosidó alguien; eso no tiene sentido. Por otra parte, en (14.c) tenemos el problema contrario: el verbo *estar* exige la expresión obligatoria del lugar en donde alguien está (o bien una descripción del estado que se le atribuye al sujeto).

Finalmente (14.d) y (14.e) también muestran las dos caras de un mismo problema. En (14.d), el verbo *comer* va acompañado por el SN que expresa *qué* ha comido Laura (nótese incidentalmente que ese SN es opcional), pero va marcado con la preposición *de*, lo cual es imposible en castellano. Por otra parte, en (14.e) el SN que indica en dónde ha estado Laura no lleva la preposición *en*, que es obligatoria para expresar esta relación.

¿Cómo podemos recoger todas estas observaciones para incorporarlas a nuestra gramática? A fin de responder esta pregunta debemos ampliar un poco más nuestro conocimiento sobre la sintaxis de las lenguas naturales con conceptos como los de *valencia y subcategorización*.

3.1. VALENCIA Y RESTRICCIONES DE SUBCATEGORIZACIÓN

Una de las primeras observaciones que podemos hacer en relación a los ejemplos de (14) es que los núcleos deben combinarse con un número determinado de elementos para dar lugar a un sintagma bien formado. Así, *toser* necesita un SN (pero no más, ni menos), *estar* necesita dos SSNN (pero no más, ni menos), etc. Es decir, estableciendo una analogía con la química, podemos decir que los núcleos, como los átomos, tienen una propiedad léxica a la que también denominamos *valencia* y que indicará el número de elementos con los que debe combinarse ese núcleo para construir un sintagma bien formado. De acuerdo con esto, podemos decir que *toser* es un verbo de valencia uno, mientras que *estar* es un verbo de valencia dos. Desde este punto de vista; pues, podemos interpretar un verbo como *besar*, por ejemplo, como la expresión de una *relación* entre un par de elementos, los participantes en la acción de besar: el besador y el besado. Estos elementos son los *argumentos* de la relación. Así, un verbo de valencia uno es un verbo que requiere un argumento, un verbo con valencia dos es un verbo que requiere dos argumentos y así hasta llegar a cuatro, que es el número máximo de argumentos que se supone que puede tener un verbo (o cualquier otro núcleo). En (15) tenemos algunos ejemplos de verbos de valencias distintas:

- (15) a. llover (valencia cero).
 a'. Llueve.
 b. toser (1: TOSEDOR) (valencia uno).
 b'. Gabriel tose.
 c. besar (1: BESADOR, 2: BESADO) (valencia dos).
 c'. Laura besó a Gabriel.
 d. dar (1: DADOR, 2: DADO, 3: RECEPTOR) (valencia tres).
 d'. Laura dio un disco a Gabriel.
 e. vender (1: VENDEDOR, 2: VENDIDO, 3: COMPRADOR, 4: PRECIO) (valencia cuatro).
 e'. Gabriel vendió un coche a Javier por 80.000 pesetas.

Aquí hemos utilizado una notación pseudológica para expresar el hecho de que el verbo denota una relación con un determinado número de argumentos, los cuales aparecen numerados entre paréntesis inmediatamente después del verbo. También damos una caracterización semántica de estos argumentos sobre cuál es su papel en esa relación; es decir, toda acción de besar comporta la presencia de un besador y un besado. Esta caracterización no es muy interesante, ya que no establece ninguna generalización sobre el tipo de participantes que pueden intervenir en una acción determinada, pero es suficiente para nuestros propósitos; véase, sin embargo, el capítulo 7 (§ 2.1.1 Relaciones temáticas o de predicación), donde se introduce la noción de papel temático como un intento de establecer generalizaciones en este nivel.

Es evidente que la noción de valencia, al establecer una primera tipología de los núcleos, nos abre la vía para resolver el problema de nuestra gramática, ya que en (14.b) y en (14.c) no estamos respetando la valencia de los verbos *toser* y *estar*. Sin embargo, como también se demuestra en (14), la valencia es un concepto demasiado neutro, y por tanto insuficiente, para dar cuenta de todos los problemas que esos ejemplos nos plantean. Por ejemplo, los verbos *comer* y *estar* tienen dos argumentos, pero imponen restricciones distintas sobre su naturaleza tanto semántica como sintáctica. En particular, el segundo argumento de estos dos verbos tiene que ser distinto en cada caso: el de *estar* debe denotar un lugar y debe ser un SN marcado con la preposición *en* (cf. *Laura está en Roma*); el de *comer* debe denotar una comida y debe ser un SN sin ninguna marca especial (cf. *Laura come pan*). Es decir, todo núcleo, además de las restricciones semánticas que impone sobre la naturaleza de sus argumentos (el tipo de entidades en el mundo que esos argumentos deben denotar), también impone restricciones sintácticas sobre el tipo de sintagma que debe utilizarse para expresar ese argumento. Llamamos a éstas *restricciones de subcategorización*. Así, *comer* es un verbo con dos argumentos que *subcategoriza* un SN y, opcionalmente, otro SN (cf. *Gabriel está comiendo*); *estar* es un verbo con dos argumentos que *subcategoriza* un SN y un SN marcado con *en*; el nombre *hermana* es un nombre con un argumento que *subcategoriza* un SN marcado con *de* (en 3.3. veremos cómo podemos expresar estas ideas formalmente en nuestra gramática).

Es importante tener en cuenta que toda esta información léxica, que podemos denominar *estructura argumental* del núcleo, forma parte de la información morfosintáctica de toda palabra gramatical almacenada en el léxico. Nos queda por examinar ahora de qué manera se refleja esta información en las estructuras sintácticas.

3.2. RELACIONES GRAMATICALES, DIÁTESIS Y VOZ

Tomemos algunos ejemplos de verbos castellanos cuya valencia sea superior a cero:

- (16) a. Gabriel tose mucho.
 b. El tren llegó tarde.
 c. Laura besó a Gabriel en la boca.
 d. Gabriel come pan con moderación.
 e. Laura regaló un disco a Gabriel por su cumpleaños.
 f. Gabriel vendió el coche a Javier por 80.000 pesetas antes de comprarse uno nuevo.

Como vemos, en cada una de estas oraciones aparecen constituyentes que, tanto si son obligatorios como si son opcionales, obedecen las restricciones que el verbo impone sobre la naturaleza de sus argumentos; aparecen también otros elementos que nunca son obligatorios y que no se corresponden con argumentos del verbo, pero que aportan nueva información sobre la acción denotada por el verbo: la sitúan en el espacio o en el tiempo (*tarde, en la boca, por su cumpleaños, antes de comprarse uno nuevo*) o especifican el modo en que ésta se lleva a cabo (*mucho, con moderación*). Es decir, cada constituyente desempeña una *función* diferente dentro de las representaciones sintácticas. En el apartado anterior ya identificamos una de las funciones que puede llevar a cabo un constituyente dentro de un sintagma, la función de *núcleo*, que siempre es desempeñada por una palabra. Ahora acabamos de identificar las otras dos funciones básicas que puede desarrollar un constituyente: la función de *complemento* y la función de *modificador* o *adjunto*. Todos aquellos sintagmas que acompañan al núcleo en una estructura sintáctica y que se corresponden con los argumentos de ese núcleo los denominaremos *complementos* (de modo que, como veremos más adelante, esta noción de complemento es un poco más amplia, ya que también incluye al sujeto); por otra parte, aquellos sintagmas que, sin corresponderse con ningún argumento, aporten información adicional sobre la entidad denotada por el núcleo los denominaremos, indistintamente, *modificadores* o *adjuntos*. Estas definiciones son deliberadamente vagas porque es muy difícil trazar una frontera nítida entre complementos y adjuntos. Por ejemplo, los adjuntos siempre son opcionales, pero no todos los complementos son siempre obligatorios (en (17.d), *a Javier* debe interpretarse como el comprador):

- (17) a. Gabriel vendió el coche.
 b. Gabriel vendió el coche a Javier.
 c. Gabriel vendió el coche por 80.000 pesetas.
 d. *Gabriel vendió a Javier por 80.000 pesetas.

Tampoco podemos decir, por ejemplo, que todos los constituyentes que expresan un contenido espacio-temporal o modal sean adjuntos, ya que algunos verbos tienen argumentos de este tipo (cuando un elemento entre paréntesis va precedido de un asterisco queremos indicar que su presencia es obligatoria):

- (18) a. Laura está *(en Roma).
 b. Gabriel puso el libro *(encima de la mesa).
 c. Este libro se lee *(muy rápido).

Quizá la prueba más fiable para distinguir los adjuntos de los complementos es la prueba de la sustitución por la proforma *hacerlo*. Dicha proforma tiene la propiedad de sustituir al verbo y todos sus complementos, pero no necesariamente a los adjuntos. Compárese (19.a) con (19.b):

- (19) a. *Gabriel se tomó un café con leche en casa y Laura lo hizo un té en el bar.
 b. Gabriel se tomó un café con leche en casa y Laura lo hizo en el bar.

Como demuestra (19.b), la proforma *hacerlo* tiene, en este caso, la interpretación de «tomarse un café con leche», lo cual explica la imposibilidad en (19.a) de expresar el

complemento *un té*. Sin embargo, el adjunto *en el bar* sí puede expresarse, pese a la presencia de la proforma. En general, pues, cuando en una estructura coordinada sustituyamos parte del SV por *hacerlo*, los constituyentes cuya expresión sea imposible (como ocurre con *un té* en (19.a)) serán complementos, mientras que aquellos que sí puedan aparecer junto a la proforma serán adjuntos. Esta prueba sólo funciona, sin embargo, con algunos verbos de acción, pero no con verbos de estado o estativos, por ejemplo:

(20) *Gabriel estuvo en Roma en agosto y Laura lo hizo en octubre.

Como dijimos, no existe un conjunto definitivo de pruebas que nos permita distinguir claramente entre complementos y adjuntos, aunque las pruebas que hemos presentado son suficientes para establecer esta distinción en la mayoría de los casos. Centremos, pues, nuestra atención en los complementos y precisemos un poco más su papel en las estructuras sintácticas.

En primer lugar nos interesa justificar la distinción entre el concepto de argumento, en tanto que noción eminentemente léxica, y el concepto de complemento, en tanto que noción eminentemente sintáctica. Hemos visto que toda palabra gramatical en el léxico tiene una valencia, es decir, un número determinado de argumentos. Esos argumentos, por lo general, se expresan explícitamente en la sintaxis mediante complementos, como quedó ejemplificado en los ejemplos de (15) y de (16). Sin embargo, esto no siempre es así, ya que ciertos verbos permiten que sus argumentos permanezcan implícitos en la sintaxis, sin que aparezca ningún complemento:

- (21) a. Gabriel siempre está comiendo.
 b. Gabriel vendió el coche.
 c. Gabriel vendió el coche a Javier.
 d. Gabriel vendió el coche por 80.000 pesetas.

En (21.a) el segundo argumento de *comer* (lo comido) no se expresa mediante ningún complemento, aunque no por ello dejamos de interpretar que hay algo que Gabriel come. Lo mismo ocurre en (21.b-d), donde tenemos diferentes grados de explicitud en cuanto a los argumentos de *vender*: en (21.b) *Gabriel vendió el coche*, no se nos dice ni a quién ni por cuánto, pero necesariamente tuvo que haber un comprador y un precio; en (21.c) conocemos el nombre del comprador, aunque no el valor de la transacción, que necesariamente tuvo que existir si se trató de una venta (*Gabriel vendió el coche*, no lo regaló); finalmente, en (21.d) se expresa el precio pero no el comprador, aunque necesariamente tuvo que haber uno, pues toda venta implica una compra. (Nótese que por este mismo motivo podríamos argumentar que en muchos casos los adjuntos locativos y temporales son, de hecho, argumentos que con frecuencia permanecen implícitos, ya que toda acción se produce necesariamente en un espacio y en un tiempo determinados.)

Otra observación relevante como justificación de la necesidad de distinguir entre argumentos y complementos es el hecho de que los argumentos de un verbo no siempre se expresan mediante un complemento del mismo tipo:

- (22) a. Gabriel vendió el coche.
 b. El coche fue vendido por Gabriel.
 c. Se venden coches de ocasión.

- d. Gabriel rompió el jarrón.
 e. El jarrón se rompió.
 f. Gabriel bailó mucho durante la fiesta.
 g. Durante la fiesta se bailó mucho.

Por ejemplo, en (22.a) y (22.b) tenemos dos oraciones que expresan exactamente el mismo contenido, pero su sintaxis es diferente, ya que los argumentos del verbo aparecen expresados mediante complementos distintos, aunque siguen interpretándose como vendedor y vendido. En el resto de casos tenemos situaciones parecidas, aunque con la diferencia de que algunos argumentos, en vez de expresarse mediante complementos diferentes, permanecen implícitos, como ocurre, por ejemplo, en (22.g), donde no aparece ningún complemento.

A la vista de estos datos, a los que retornaremos más adelante, debemos concluir que el hecho de que un verbo tenga una valencia determinada no garantiza que todos y cada uno de sus argumentos se expresen de forma explícita en la sintaxis mediante un constituyente pleno. Al llegar a esta conclusión, hemos dado con otro de los problemas fundamentales del estudio de la sintaxis de las lenguas naturales: el problema de la proyección de la estructura argumental de un núcleo en una estructura sintáctica; es decir, dado que las propiedades léxicas de un núcleo (es decir, su valencia y la tipificación semántica de sus argumentos) permanecen constantes, ¿cómo podemos dar cuenta del hecho de que estas propiedades léxicas puedan dar lugar a estructuras sintácticas diferentes? Para poder responder a esta pregunta, debemos profundizar un poco más en la noción de complemento y en otros conceptos relacionados con ella que todavía no hemos introducido.

Un aspecto importante de lo que acabamos de observar, no obstante, es que, si bien un verbo con una estructura argumental determinada puede dar lugar a estructuras sintácticas diferentes, esta relación no es aleatoria, sino que es predecible gracias a una propiedad común a casi todas las estructuras argumentales, que tienden a proyectarse en una estructura sintáctica que podríamos considerar como más básica, de la cual podemos derivar todas las demás. Esta propiedad se ilustra mejor con verbos de valencia dos (adaptamos los ejemplos que siguen del libro de Ll. Gràcia, *La teoria temàtica*. Publicacions de la UAB, Bellaterra, 1989):

- (23) a. Juan [SV le rompió el brazo].
 b. Juan [SV se rompió el brazo].

El verbo *romper* tiene dos argumentos; el segundo siempre se interpreta como lo ROTO, que en (23) se expresa mediante el SN complemento *el brazo*. El primer argumento también se proyecta como un SN, pero su interpretación está determinada por todo el SV en conjunto (que, siguiendo otra convención muy común, representamos encerrado entre corchetes etiquetados; un sistema de representación que es más económico que el de los diagramas arbóreos). Nótese que, según el benefactivo (o quizá malefactivo, en este caso) que aparezca dentro del SV, el SN se interpreta como el ROMPEDOR (23.a), es decir, el que ejecuta la acción de romper el brazo de alguien, o como el MALEFICIARIO (23.b), es decir, aquel que se ha roto el brazo. El que la interpretación del complemento que expresa el primer argumento de un núcleo dependa del conjunto formado por el verbo más todos los demás complementos es un hecho bas-

tante sistemático que ha llevado a proponer una distinción fundamental entre ese primer argumento, al que se denomina *argumento externo*, y el resto de argumentos, los *argumentos internos*. Nótese que esta distinción entre argumento externo y argumento interno se ve avalada, además de por casos como el de (23), por el hecho de que, estadísticamente, en las *expresiones idiomáticas* es mucho más frecuente que el argumento interno se exprese mediante un complemento fijo, mientras que el complemento que expresa el argumento externo permanece abierto:

- (24) a. SN estiró *la pata*.
 b. SN hace *novillos*.
 c. SN le busca *las cosquillas* a alguien.
 d. SN le saca *las castañas* del fuego a alguien.

Asimismo, existe una serie de construcciones, conocidas por el nombre de *construcciones de verbo-soporte*, que se forman con un *verbo débil* (casi vacío de contenido semántico) más un argumento interno fijo, pero nunca un argumento externo fijo:

- (25) a. SN dio *un beso* a alguien (cf. SN besó a alguien).
 b. SN le tiene *miedo* a alguien (cf. SN teme a alguien).
 c. SN le puso *la zancadilla* a alguien (cf. SN zancadilleó a alguien).

Ello implica, por tanto, que la tipificación semántica de los complementos se produce en la estructura sintáctica, y no en el léxico, y en virtud de la *posición* que el SN ocupa en dicha estructura sintáctica. Así, para expresar este hecho, en un verbo de valencia dos, por ejemplo, el argumento externo se expresa mediante un SN hermano del SV, mientras que el argumento interno se expresa mediante un SN hermano del núcleo V. Evidentemente, la relación estructural que, en cada caso, mantiene el complemento con su núcleo es distinta (figura 7): son complementos que expresan *relaciones gramaticales* distintas, el *sujeto* el primero, y el *complemento* (u *objeto*) *directo* el segundo:

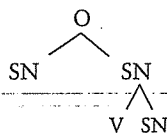


Figura 7

Las relaciones gramaticales y la manera que tiene la lengua de expresarlas superficialmente es la pista más directa de que disponemos para identificar el papel semántico que recibe cada complemento. Por este motivo las lenguas utilizan una serie de estrategias para identificar la relación gramatical que un complemento mantiene con su núcleo y para señalar cuándo la correlación entre papeles semánticos y relaciones gramaticales no sigue la estrategia «más básica». A fin de evitar giros lingüísticos innecesarios es conveniente poner un nombre a cada una de estas dos estrategias, que, siguiendo el manual

de Moreno Cabrera (1991), denominaremos *diátesis* y *voz*. La primera hace referencia a la expresión de los papeles semánticos a través de una relación gramatical u otra. Así por ejemplo, la diátesis básica del verbo *cerrar* consiste en expresar el CERRADOR como sujeto y el CERRADO como complemento directo, los cuales, en castellano, se identifican por el hecho de que el sujeto es un SN que concuerda con el verbo en número y persona y que tiende a preceder al verbo, mientras que el complemento directo es un SN que, si es animado, va marcado por la preposición *a*. Por otra parte, la voz hace referencia a la forma morfológica superficial del verbo, que, en el caso que nos ocupa, es la voz activa. Así, *Gabriel cerró la puerta* es una oración que expresa la diátesis básica del verbo *cerrar* a través de la voz activa. Cualquier alteración de la diátesis básica se verá acompañada por una alteración de la voz; en (26) tenemos ejemplos de las cuatro diátesis posibles del verbo *cerrar* expresadas con sus correspondientes voces:

- (26) a. Gabriel cerró la puerta (voz activa).
 b. La puerta fue cerrada por Gabriel (voz pasiva).
 c. La puerta se cerró (voz media).
 d. Laura hizo a Gabriel cerrar la puerta (voz causativa).

Nótese que, en la voz activa, el verbo no lleva ninguna marca morfológica especial, el argumento externo se corresponde con el sujeto el cual concuerda con el verbo, mientras que el argumento interno se expresa como complemento directo. En la voz pasiva, el verbo adopta la forma de una perífrasis con SER + PARTICIPIO, el argumento externo se expresa como un complemento preposicional (opcional) marcado con *por*, mientras que el argumento interno se expresa como sujeto. En la voz media, el verbo siempre va precedido por el morfema *se* y se caracteriza por la no expresión de forma explícita del argumento externo. Finalmente, en la voz causativa, el verbo adopta la forma HACER + INFINITIVO y se caracteriza por la aparición de un complemento adicional, expresado como sujeto (y con un papel semántico que podríamos caracterizar como CAUSANTE o INDUCTOR de la acción), mientras que el argumento interno permanece inalterado y el argumento externo se expresa mediante una relación gramatical que no habíamos mencionado hasta ahora (y de la que nos ocuparemos un poco más adelante): el *complemento indirecto*. Nótese que las voces media y causativa coinciden con una alteración de la estructura argumental, la supresión del argumento externo en el primer caso y el añadido de un nuevo argumento en el segundo.

Ahora podemos analizar con un poco más de detenimiento los ejemplos de (22), que repetimos aquí como (27):

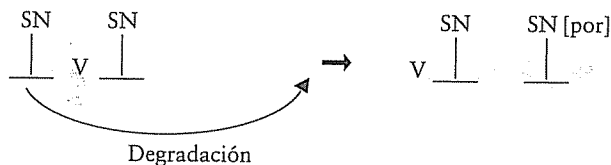
- (27) a. Gabriel vendió el coche.
 b. El coche fue vendido por Gabriel.
 c. Se venden coches de ocasión.
 d. Gabriel rompió el jarrón.
 e. El jarrón se rompió.
 f. Gabriel bailó mucho durante la fiesta.
 g. Durante la fiesta se bailó mucho.

En (27.a-c) tenemos tres ejemplos con el verbo *vender* en los que diferentes alteraciones de la diátesis se expresan mediante voces diferentes y tienen consecuencias lige-

ramente diferentes en lo que a la expresión de los complementos respecta. (27.a) se correspondería con la diátesis más básica en voz activa, donde el VENDEDOR aparece como sujeto y el VENDIDO como complemento directo. La voz pasiva del verbo en (27.b) indica una alteración de la diátesis en que el VENDIDO se expresa como sujeto y el VENDEDOR aparece como complemento preposicional marcado con *por*. Finalmente, en (27.c) el VENDEDOR permanece implícito, mientras que el VENDIDO se expresa mediante un SN que concuerda con el verbo en voz media y es, por tanto, el sujeto (estas construcciones son muy comunes en las lenguas románicas, particularmente en castellano, y suelen denominarse *pasivas reflejas*). El par (27.d-e) muestra también un fenómeno muy común en castellano, con una estructura básica en voz activa donde el sujeto expresa un argumento externo que podríamos caracterizar genéricamente como CAUSANTE, mientras que el complemento directo expresa un argumento interno que podríamos caracterizar como AFECTADO. La construcción en voz media (27.e), por otra parte, deja implícito el CAUSANTE, mientras que el AFECTADO pasa a ocupar su lugar en la posición de sujeto (a menudo estas construcciones se denominan *anticausativas*, precisamente porque el causante permanece implícito y desconocido). Finalmente, en el par (27.f-g) tenemos el caso de un verbo de valencia uno (los constituyentes *mucho* y *durante la fiesta* son adjuntos) que en la voz activa expresa su único argumento como sujeto. Sin embargo, si este argumento permanece implícito, el verbo aparece en la voz media sin ningún complemento que lo acompañe; nótese que la oración tiene una interpretación genérica, sin hacer referencia a una persona o personas en concreto: estas construcciones son las que tradicionalmente se han denominado *impersonales reflejas*.

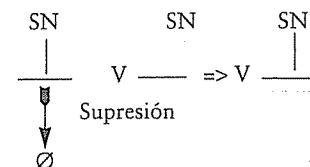
Como vemos, apelando a los conceptos de diátesis y de voz podemos dar cuenta de una amplia gama de construcciones de las lenguas naturales. De hecho, si observamos con más detalle los diferentes casos de alteración de la diátesis vemos que todos ellos comportan alguna alteración de la estructura argumental de un verbo que afecta esencialmente al argumento externo: sea su *supresión* o su *degradación* a argumento interno, la cual puede producirse de forma espontánea o por la introducción de un nuevo argumento externo. Tomemos como ejemplo las construcciones pasiva y pasiva refleja (27.b-c). La primera podríamos analizarla como el resultado de un proceso de alteración de la valencia de *vender* en que el argumento externo es degradado a argumento interno de forma espontánea, de modo que en la sintaxis este argumento «degradado» se proyecta como un complemento preposicional marcado con *por*, mientras que el segundo argumento interno pasa a expresarse como sujeto; esquemáticamente, podríamos representar este proceso como en (28), donde sólo representamos los argumentos que se ven afectados por el proceso:

(28) Degradación del Argumento Externo:



Por otra parte, la pasiva refleja, en la que el argumento externo nunca puede expresarse como complemento (al menos normativamente, ya que algunos dialectos del castellano aceptan un complemento con *por* en estas construcciones), podría analizarse como consecuencia de la supresión de este argumento, lo cual permite la «promoción» del segundo argumento interno a sujeto. De nuevo, podemos intentar esquematizar el proceso:

(29) Supresión del Argumento Externo:



De momento, nos hemos ocupado fundamentalmente de intentar caracterizar los procesos de alteración de la valencia y, en menor medida, las consecuencias sintácticas que éstos traen consigo. Por ejemplo, hemos definido la voz como la estrategia morfológica que utilizan las lenguas para indicar que un verbo ha visto alterada su valencia, es decir, un cambio en la expresión de la diátesis. No hemos profundizado, sin embargo, en la estrategia que utilizan las lenguas para indicar qué tipo de alteración se ha producido; nótese que toda alteración de la diátesis comporta una alteración de las relaciones gramaticales y, por tanto, es importante detenerse un instante en las estrategias que utilizan las lenguas para identificar las distintas relaciones gramaticales: la *concordancia* y el *caso*.

Ya observamos más arriba que en castellano el sujeto (cuando lo hay) siempre concuerda con el verbo, mientras que ningún otro complemento lo hace; la concordancia es, pues, una manera de identificar cuál es el sujeto de una oración. En general, podemos definir la concordancia como la covariación de determinados rasgos gramaticales (persona, número, género, etc.) que se produce entre un núcleo y sus complementos. En castellano, sujeto y verbo concuerdan en número y persona, lo cual se expresa mediante la flexión del verbo en función de la persona y el número del sujeto; en árabe, sujeto y verbo concuerdan en número, persona y género:

- (30) a. huwa ja-?-a
 él llegar-3a.MASC.SG
 «él llegó».
- b. hiya ja-?-ati
 ella llegar-3a.FEM.SG
 «ella llegó».

En suahelí (o swahili, lengua bantú hablada en el África Oriental) tanto el sujeto como el complemento directo concuerdan con el verbo en género (o *clase nominal*; en suahelí hay seis) y número (los números romanos indican la clase nominal; PERF es un morfema de perfecto):

- (31) m-toto a-me-ki-leta ch-akula
 III^a.SG-niño III^a.SG-PERF-I^a.SG-comprar I^a.SG-comida
 «el niño ha comprado la comida».

Así, al marcar morfológicamente nombres y verbos se identifica, sin ambigüedades, cuál es cada complemento; por ejemplo, en (31) el prefijo *a-* identifica exclusivamente al sujeto que va marcado con el prefijo *m-*, ya que ambos pertenecen a la misma clase y al mismo número; asimismo, el prefijo *ki-* y el prefijo *ch-* permiten identificar exclusivamente el complemento directo.

Además de la concordancia, y no necesariamente como alternativa a ella, las lenguas recurren a la estrategia de marcar los complementos con un morfema que nos permite identificarlos directamente como complementos de uno u otro tipo. Esta marca, el caso, es muy común en las lenguas del mundo e incluso el castellano conserva un sistema de casos en los pronombres personales:

- (32) a. Yo lo vi.
 b. Tú la viste.
 c. Él me vio.
 d. Ella te vio.

Los pronombres *yo* y *me*, *tú* y *te*, *él* y *lo*, *ella* y *la*, respectivamente, son de la misma persona y número (y, los de tercera persona, también género), pero están en casos diferentes: *yo*, *tú*, *él* y *ella* sólo pueden ser sujetos en (32), están en caso *nominativo*; los pronombres *me*, *te*, *lo*, *la* sólo pueden ser complementos directos, están en caso *acusativo*. En castellano, los pronombres personales pueden aparecer en un tercer caso, el *dativo*, y permiten identificar al complemento indirecto:

- (33) a. Gabriel le dio un libro (a Laura).
 b. Gabriel les dio un libro (a sus amigos).

Los pronombres *le* y *les* son complementos indirectos en dativo y expresan el tercer argumento del verbo *dar*, es decir, el RECEPTOR. Es difícil hacer una caracterización sistemática del complemento indirecto, aunque éste siempre se expresa mediante un SN humano marcado por la preposición *a* con un papel semántico que podríamos caracterizar genéricamente como beneficiario/mafeiciario. Es muy fácil de distinguir cuando también hay complemento directo (como ocurre en (33)), pero los complementos directos también van marcados con *a* cuando son humanos, de modo que es muy difícil distinguirlos de complementos indirectos en verbos de valencia dos, ya que hay algunos verbos cuyo segundo argumento se expresa como un complemento indirecto y no como un complemento directo. En estos casos, la prueba más fiable es la de la pronominalización (aunque no siempre, ya que *le*, por ejemplo, puede funcionar como complemento directo masculino humano, y otros dialectos del castellano permiten aún más variación):

- (34) a. Gabriel vio a Laura.
 b. Gabriel la/*le vio.
 c. Eso preocupa mucho a Gabriel.
 d. Eso *lo/le preocupa mucho.

En otras lenguas, el sistema de casos se extiende a todos los SSNN, además de los pronombres. En alemán, por ejemplo, nominativo, acusativo y dativo se expresan en los determinantes (y ocasionalmente en los nombres también):

- (35) der Lehrer hat den Brief dem Student-em gegeben
 el.NOM profesor ha el.AC carta el.DAT estudiante-DAT dado
 «el profesor ha dado la carta al estudiante».

Como se muestra en (35), el caso en alemán se expresa en las diferentes formas del artículo y sólo algunos nombres lo expresan también morfológicamente; el sujeto concuerda además en número y persona con el verbo. En japonés, el caso se expresa mediante una marca morfológica en el nombre (PRET indica el morfema de pretérito; nótese que en japonés ningún complemento concuerda con el verbo):

- (36) a. Ken-ga Naomi-wo misome-ta
 Ken-NOM Naomi-AC enamorarse-PRET
 «Ken se enamoró de Naomi».
- b. Ken-ga Naomi-ni kyuukonsi-ta
 Ken-NOM Naomi-DAT proponer.en.matrimonio-PRET
 «Ken propuso en matrimonio a Naomi».

Estos sistemas de casos son muy simples, con tres o cuatro casos, pero hay lenguas con sistemas mucho más complejos que permiten expresar aquellas funciones sintácticas que en castellano normalmente se expresan mediante preposiciones; por ejemplo, el euskera tiene dieciséis casos que permiten expresar varios tipos de complementos y adjuntos locativos, direccionales, benefactivos, comitativos (que indican con quién se realiza una determinada acción), etc.

La concordancia y el caso, como hemos visto, son dos estrategias que no se excluyen mutuamente, aunque tampoco es necesario que aparezcan juntas. Sin embargo, hay lenguas que poseen sistemas de casos o de concordancia muy empobrecidos e incluso inexistentes, por lo que deben recurrir a una tercera estrategia para identificar las relaciones gramaticales: el orden de palabras. Efectivamente, cuando no se dispone ni de caso ni de concordancia para identificar las relaciones gramaticales, las lenguas tienden a poseer un orden de palabras fijo, de modo que la posición del SN es lo que lo identifica como sujeto o complemento directo. El inglés es, en gran medida, una lengua de orden fijo:

- (37) a. John loves Mary
 «John ama a Mary».
- b. Mary loves John
 «Mary ama a John».

Los SSNN no llevan ninguna marca, aparte de su posición relativa al verbo, que permita identificarlos como sujeto o complemento directo, de modo que si alteramos el orden de ambos complementos, la interpretación de la frase cambia. Otra lengua de orden fijo es el dakota (una lengua india de Norteamérica):

- (38) a. Wičhása ki mathó wꜜ kté
 hombre el oso un mató
 «el hombre mató un oso».
- b. Mathó wꜜ wičhása ki kté
 oso un hombre el mató
 «un oso mató al hombre».

De nuevo, cualquier alteración del orden relativo de los complementos comporta un cambio en la interpretación de la frase, lo cual no ocurre con lenguas como el japonés, el alemán o el castellano, en las que el orden de palabras es relativamente libre, gracias a que el caso y la concordancia permiten identificar las relaciones gramaticales sin que haya apenas ambigüedades. Pero las variaciones de orden no son totalmente libres ni vacías de contenido, como veremos en el último apartado de este capítulo. Antes, sin embargo, intentaremos incorporar algunas de las ideas que hemos desarrollado en nuestra gramática.

3.3. GRAMÁTICA II

El primer punto que trataremos es el de la valencia y las restricciones de subcategorización. Dejaremos de lado, sin embargo, el tratamiento de las restricciones semánticas, ya que no es éste un asunto que hayamos tratado con todo el detalle que merece (pero véase el capítulo 7, §2.1.1).

El problema principal que tenía nuestra gramática era que, dada una regla de estructura de frase, cualquier palabra gramatical podía ocupar la posición del núcleo siempre que cumpliera la sola restricción de pertenecer a la misma categoría. Así, el verbo *toser* podía aparecer como núcleo de un SV que tuviera un complemento directo, y el verbo *estar* podía aparecer como núcleo de un SV que no tuviera ningún complemento. Éste es, precisamente, el tipo de situaciones que queremos evitar. Para conseguirlo sólo necesitamos ampliar ligeramente el mismo principio que utilizábamos en §2.3 para asegurar que un núcleo de una determinada categoría se insertara dentro de un sintagma de la misma categoría; es decir, si en una regla aparecía el símbolo V, con ello queríamos decir que sólo un verbo podía aparecer como núcleo de ese sintagma. Ahora lo que queremos es poder decir además que el verbo que debe aparecer como núcleo debe ser de un tipo (o *subcategoría*) determinado. ¿Cómo podemos expresar que un verbo pertenece a una subcategoría determinada? Mediante el uso de rasgos. Recordemos que cuando desarrollamos una representación para las palabras gramaticales, consideramos necesario otorgarles una estructura interna compleja, de modo que la información gramatical se organizara en diferentes tipos. Sin embargo, a lo largo del capítulo, los símbolos que hemos utilizado para etiquetar los nodos de los diagramas arbóreos eran atómicos, no tenían estructura interna. Supongamos ahora que la tienen, es decir, que los nodos que aparecen en las representaciones sintácticas son matrices de rasgos, con una estructura idéntica a la de las palabras gramaticales. Supongamos además que toda palabra gramatical posee un rasgo especial que nos indica a qué subcategoría pertenece (llamémoslo SUBCAT) y que su valor puede ser cualquier número natural positivo (es decir, 1, 2, 3, etc.). De acuerdo con esto, podríamos decir que el verbo *toser*, por ejemplo, pertenece a la subcategoría 1 y que sólo puede insertarse en aquellas reglas en que pueden aparecer núcleos que pertenecen a esta subcategoría; es decir, nuestras reglas ahora serían como en (39):

- (39) SV → V[1],

donde, para ahorrar espacio, escribimos V[1] para abreviar la matriz de rasgos de (40):

$$(40) \left[\begin{array}{l} \text{FON} = ? \\ \text{MORFSIN} = \left[\begin{array}{l} \text{CAT} = \text{V} \\ \text{FLEX} = ? \\ \text{SUBCAT} = 1 \end{array} \right] \end{array} \right]$$

En (40) abreviamos un poco el nombre de los rasgos para que la matriz no crezca demasiado y se haga completamente ilegible, e introducimos un nuevo símbolo (?) para representar aquella información que desconocemos o que no nos interesa especificar. (Por ejemplo, queremos que en (39) se pueda insertar cualquier verbo cuya subcategoría sea 1 y, por eso, no especificamos cuál ha de ser su fonología; ? es pues un símbolo que puede tomar cualquier valor en el momento en que éste se conozca, es una *variable*, como la *x* de una ecuación. Más adelante perfeccionaremos nuestra manera de escribir las variables.)

Con esta técnica podemos escribir reglas que recojan todas las restricciones de subcategorización que los núcleos imponen sobre sus complementos si, además de especificar la subcategoría a la que pertenece el núcleo, expresamos también mediante rasgos las propiedades sintácticas de los complementos, como, por ejemplo, la preposición con que deben ir marcados. Empecemos, pues, por introducir el marcaje preposicional en la gramática. Supongamos que todo SN, en virtud del hecho de que puede estar marcado por una preposición, posee un rasgo en el que este marcaje queda reflejado. Llamemos a este rasgo MARCP (abreviatura de MARCA PREPOSICIONAL) y supongamos que sus valores pueden ser los nombres de las preposiciones (es decir, A, DE, PARA, POR, EN, CON, etc.) además de un valor especial NIL, con el que indicaremos que el SN no tiene ninguna marca preposicional. (NIL es una palabra común en los lenguajes de programación para expresar precisamente lo que aquí queremos expresar, es decir, la ausencia de marca o el valor cero de algún rasgo. Podríamos igualmente haber tomado otras palabras como NO, NADA o NINGUNA; nótese, sin embargo, que NIL *no es una variable*, ya que las variables indican cualquier valor y aquí lo que queremos expresar es un valor concreto, el no-valor.) Siguiendo la convención que introdujimos con SUBCAT, con MARCP también abreviemos nuestras representaciones de las categorías escribiendo SN[A] o SN[NIL] para indicar cuál es el valor de MARCP en cada caso; cuando necesitemos expresar más de un rasgo los separaremos por comas: P[40, POR] representa una preposición con valor 40 para SUBCAT y valor POR para MARCP. Para escribir la regla que introduce las preposiciones y las marcas preposicionales utilizaremos además un pequeño truco que nos permitirá asegurar que cada preposición marca debidamente al SN al que acompaña:

- (41) SN[Prep] → P[40, Prep] SN[NIL],
 donde Prep = {A, DE, POR, EN, PARA, CON, NIL, etc.}.

Aquí Prep es una variable como ? más arriba, aunque la interpretación que aquí le estamos dando es un poco más refinada, ya que al escribir Prep en ambos lados de la regla lo que estamos diciendo es que si el valor de Prep de la preposición es A, por ejemplo, también lo será el valor de Prep de SN. Además, no queremos que Prep tome va-

lores extraños, sino sólo los valores posibles que puede tomar MARCP: ésta es precisamente la función de la cláusula que acompaña a la regla; esta cláusula es una *restricción*, ya que restringe el tipo de valores que puede tomar Prep a una única clase, y consideraremos que nunca puede ser violada. Finalmente, nótese que el SN que aparece como hermano de la preposición está explícitamente marcado como NIL; así impedimos que en esa posición aparezcan SSNN marcados con otro valor que no sea NIL e impedimos que la gramática genere secuencias como **a de Juan*, **por en la mesa*, etc.

Ahora podemos escribir reglas para el SV que recojan las restricciones de subcategorización de cada verbo; nótese que consideramos que los complementos indirectos son hermanos del complemento directo:

- (42) a. SV → V[1].
 b. SV → V[2] SN[NIL].
 c. SV → V[3] SN[A].
 d. SV → V[4] SN[EN].
 e. SV → V[5] SN[NIL] SN[A].

Sin embargo, esto no es del todo satisfactorio, ya que no recoge algunos de los aspectos que hemos tratado en el texto. Por ejemplo, no recoge la diferencia entre un SN[A] en acusativo o en dativo, ni que en verbos como *comer* el complemento directo es opcional, y no obliga a suponer que un verbo que puede llevar tanto complementos directos animados con *a* como complementos directos no animados sin preposición es en realidad dos verbos, cada uno de una subcategoría diferente (V[2] y V[3]). Intentemos resolver pues cada uno de estos problemas, considerando en primer lugar el problema del verbo *comer*. Estos verbos son interesantes porque muestran un comportamiento distinto según si el complemento directo, cuando está en singular, lleva determinante o no lo lleva:

- (43) a. Gabriel comió pan.
 b. *Gabriel comió el pan.
 c. Gabriel se comió el pan.
 d. *Gabriel se comió pan.

Parece que la presencia/ausencia del determinante es signo de una diferencia más profunda, una diferencia de caso:

- (44) a. *Gabriel lo comió (pan).
 b. Gabriel se lo comió (el pan).

Es decir, un SN como *pan* en posición de complemento directo no puede pronominalizarse en acusativo. Esto recuerda mucho una *diferencia* que también existe en catalán, aunque en esta lengua sí que se puede pronominalizar el complemento directo carente de artículo, pero con un pronombre distinto:

- (45) a. En Gabriel va menjar pa.
 b. En Gabriel en va menjar.
 c. En Gabriel es va menjar el pa.
 d. En Gabriel se'l va menjar.

El pronombre *en* en catalán está en caso partitivo, no acusativo, lo cual nos hace sospechar que quizá en castellano *pan* en (43.a) también está en partitivo y no en acusativo, aunque no exista un pronombre en ese caso; nótese que los verbos con complementos directos partitivos no pueden pasivizarse:

- (46) a. *Pan fue comido por Gabriel.
 b. *Pa va ser menjat per en Gabriel.

En italiano la diferencia también existe, aunque en esta lengua se puede marcar el SN con una preposición para indicar que está en partitivo:

- (47) a. Gabriele ha mangiato pane/del pane.
 b. Gabriele ne ha mangiato.
 c. Gabriele si è mangiato il pane.
 d. Gabriele se lo è mangiato.

Podemos concluir, pues, a la vista de los datos comparativos, que en castellano también hay un caso partitivo, aunque éste nunca se distinga del acusativo más que por la presencia/ausencia del artículo. Nótese, además, que son los verbos que permiten una lectura partitiva del complemento directo singular los que pueden omitirlo, dejando implícito su segundo argumento:

- (48) a. Gabriel comió.
 b. Gabriel bebió.
 c. *Gabriel se comió.
 d. *Gabriel se bebió.

Hechas estas consideraciones, podemos decir que los verbos de dos argumentos pueden expresar dicho argumento como un complemento directo partitivo o como un complemento directo en acusativo. En el primer caso, el complemento es opcional, en el segundo es obligatorio y debe ir marcado con *a* si es animado. ¿Qué ocurre con los verbos de tres argumentos? Su complemento directo puede ser acusativo o partitivo:

- (49) a. Gabriel dio pan a Laura.
 b. Gabriel dio el pan a Laura.

Introduzcamos, pues, estas diferencias en las reglas de (43):

- (50) a. SV → V[1] (p.ej.: toser, estornudar, dormir).
 b. SV → V[2] (SN[NIL, PART]) (p.ej.: comer, beber).
 c. SV → V[3] SN[AvNIL, AC] (p.ej.: ver, comerse, beberse).
 d. SV → V[4] SN[EN] (p.ej.: estar, poner).
 e. SV → V[5] SN[NIL, ACvPART] SN[A, DAT] (p.ej.: dar, regalar, entregar).

En (50) hemos introducido un nuevo símbolo, «v», con el que indicamos una disyunción de valores. Así, por ejemplo, el valor de MARCP del SN de (50.c) puede ser A o NIL. El paréntesis indica opcionalidad. Además, siguiendo la convención, abrevia-

mos los rasgos de caso expresando directamente el valor de caso en vez de CASO = PART, por ejemplo.

Nos quedan por resolver las reglas que introducen los complementos del N y las reglas que introducen los determinantes. Las primeras son sencillas, pues son idénticas a las de los verbos:

- (51) a. SN → N[20] (p.ej.: Laura, Roma).
 b. SN → N[21] SN[DE] (p.ej.: casa, libro, pan).
 c. SN → N[22] SN[DE] (p.ej.: hermana, padre, amigo).

En (51.c) decimos que el complemento es obligatorio pese a la posibilidad de tener SSNN como *sus padres* o *mis hermanas*. Sin embargo, es evidente que los pronombres posesivos expresan precisamente ese único argumento que tienen estos nombres relacionales de forma paralela a como *lo* o *la* expresan el argumento de un verbo; podemos, pues, presuponer la existencia de un paralelismo entre ambos tipos de pronominalización, aunque no resolveremos aquí el problema del posicionamiento de estos pronombres, ya que es un problema complejo que va más allá de los objetivos de este capítulo. Esta observación sí nos permite, no obstante, hacer una breve reflexión sobre el caso de los complementos nominales. Es evidente que, al pronominalizar como posesivos y no como acusativos, dativos o partitivos, su caso debe ser otro. Por analogía con otras lenguas (como el alemán) que poseen un caso explícito para marcar los complementos nominales, lo denominaremos genitivo, e incluiremos esta información en nuestras reglas:

- (52) a. SN → N[21] SN[DE, GEN] (p.ej.: casa, libro, pan).
 b. SN → N[22] SN[DE, GEN] (p.ej.: hermana, padre, amigo).

Pasemos ahora a las reglas para introducir determinantes. Quizá lo mejor es intentar averiguar cuál es el caso del SN que es hermano del determinante. Esto sólo podemos averiguarlo de forma indirecta, por ejemplo, a través del catalán. En esta lengua, un determinante (con la excepción del artículo determinado) puede aparecer solo, mientras que el SN al que acompaña aparece pronominalizado:

- (53) a. Me n'he menjat un.
 b. Me n'he menjat tres.
 c. Me n'he menjat algun.

En castellano, esto también es posible, aunque no aparece ningún pronombre:

- (54) a. Me he comido uno.
 b. Me he comido tres.
 c. Me he comido alguno.

Parece, pues, que el caso del SN que acompaña al determinante tiene que ser partitivo, pero el SN resultante no lo es ni lo puede ser, ya que un SN partitivo o está en plural (*Javier lee libros*) o no lleva determinante (*Gabriel bebe agua*); por el contrario, si lleva determinante sólo puede ser acusativo:

- (55) a. Javier lee libros vs. Javier (se) lee los libros.
 b. *Javier los lee vs. Javier (se) los lee,

donde el asterisco de (55.b) debe interpretarse en el sentido de que no puede ser la pronominalización de (55.a). Para recoger este hecho podemos escribir una regla como la siguiente:

- (56) SN[¬PART] → Det SN[PART],

donde ¬PART es un valor negado (el símbolo «¬», en lógica, significa «no»). Con ello expresamos algo parecido a «cualquier caso con la excepción del partitivo».

Finalmente, para completar nuestra gramática, nos falta ocuparnos del sujeto y, en particular, de la relación de concordancia que mantiene con el SV.

El sujeto siempre va marcado en nominativo, lo que se puede expresar de la misma manera que hemos expresado la asignación de los demás casos. En cuanto a la concordancia, debemos echar mano de una estrategia que ya utilizamos anteriormente para el marcaje preposicional: las variables. Sabemos que en castellano sujeto y verbo concuerdan en persona y número, lo que, en términos formales, puede interpretarse como identidad de los rasgos de persona y número de sujeto y verbo:

- (57) O → SN[NOM, PERS=Pe, NUM=Num] SV[PERS=Pe, NUM=Num],
 donde Pe = {1, 2, 3} y Num = {SING, PLU}.

Como siempre, al escribir la regla abreviamos para ahorrar espacio, aunque es preciso no olvidar que los rasgos de persona y número forman parte de un rasgo de nivel superior que denominamos FLEXIÓN.

4. EL ORDEN DE PALABRAS

Gran parte de nuestra discusión sobre las alteraciones de la valencia y las consiguientes variaciones de las relaciones gramaticales, la asignación de caso y la concordancia llevaba implícita la idea de que existe una fuerte tendencia en las lenguas a expresar la diátesis básica a partir de la que se derivan las demás diátesis posibles con un determinado orden de los constituyentes. Por ejemplo, en castellano, los complementos de un verbo tienden a disponerse de forma que el sujeto preceda al verbo y éste a su vez a los demás complementos, cuyo orden relativo, si es que hay más de un complemento, suele ser: complemento directo, seguido de complemento indirecto, seguido de complementos preposicionales. Es decir, cada diátesis tiende a expresarse con un orden de palabras preferido o *no marcado* (por utilizar un término común en la jerga lingüística). Sin embargo, también vimos que lo que en castellano no es más que una tendencia en otras lenguas (como el inglés o el dakota) es una imposición de la gramática, es decir, es el propio orden de los constituyentes lo que nos permite identificar un complemento como sujeto o complemento directo y cualquier alteración del mismo provoca una oración imposible en esa lengua o un cambio en la interpretación de la misma.

Al hablar de órdenes no marcados, no obstante, necesariamente nos vemos obligados a considerar aquellos órdenes que son marcados y a preguntarnos cuál es su fun-

ción en la lengua, si es que la tienen. Por ejemplo, en castellano, sabemos que podemos alterar el orden no marcado con bastante libertad y, en ocasiones, dichas variaciones de orden tienen consecuencias en cuanto a la información que comunica una oración. Consideremos (58):

- (58) a. Gabriel bebe agua.
b. AGUA bebe Gabriel.

Es evidente que (58.a) y (58.b) comunican exactamente un mismo contenido, es decir, que Gabriel bebe agua; sin embargo (58.b) proferida con una entonación especial, poniendo el énfasis sobre el SN *agua* (lo cual hemos expresado utilizando versalitas), transmite más información además del mero hecho de que Gabriel bebe agua. En efecto, al alterar el orden de los constituyentes y la entonación, (58.b) introduce un elemento de contraste que no está presente en (58.a) y nos comunica que es agua lo que Gabriel está bebiendo y no cualquier otra cosa. Así pues, (58.a) y (58.b), pese a comunicar exactamente el mismo contenido, no transmiten la misma información y se utilizarían en situaciones diferentes.

No podemos detenernos aquí a estudiar las diferentes funciones comunicativas que poseen ciertas alteraciones del orden de palabras, ya que ello escapa en gran medida al dominio de la sintaxis y pertenece más bien al terreno del estudio del lenguaje en relación a diferentes situaciones de uso o pragmática. Las lenguas recurren, no obstante, a diferentes estrategias sintácticas para expresar esas funciones (incluso las lenguas con órdenes más rígidos), de modo que, sin realizar un análisis pormenorizado de dichas funciones, sí que podemos estudiar los mecanismos sintácticos que participan en su expresión. Sin embargo, antes debemos estudiar con mayor detalle el concepto de orden básico y las funciones comunicativas que éste lleva asociadas.

4.1. ORDEN BÁSICO Y CATEGORÍAS PRAGMÁTICAS

Existe, pues, más de una manera «de decir la misma cosa» y cada manera de decir la comporta transmitir una información adicional sobre lo que queremos decir. Existe también, sin embargo, una manera no marcada «de decir una cosa», lo cual nos permite, por otra parte, afirmar que hay más de una manera de decir la. Esta manera no marcada es lo que denominamos *orden básico de palabras* (o de constituyentes), donde «básico» es un concepto un tanto difuso relacionado con factores tales como la frecuencia de uso (a mayor frecuencia, menos marcado), su asociación con una serie de parámetros que nos permiten decir que es la forma más neutra (por ejemplo ser una oración declarativa frente a ser una oración interrogativa o imperativa), etc.; véase Croft (1990, cap. 4).

Generalmente, el orden básico caracteriza el orden relativo del sujeto (S), el verbo (V) y el complemento directo (O) —aunque a veces se interpreta el símbolo O como representante de todos los demás complementos—, lo cual da lugar a seis órdenes básicos posibles: SOV, SVO, VSO, VOS, OVS, OSV. Hasta el momento, sólo se ha podido confirmar la existencia de los cinco primeros tipos de orden (no se conoce, pues, ninguna lengua que sea OSV), con una preferencia mayoritaria por los tres primeros y, en particular, los dos primeros. Ejemplos:

- (59) a. Hasan öküz-ü aldi (turco: SOV)
Hasan buey-AC compró
«Hasan compró el buey».
- b. Gabriel vio la película (castellano: SVO)
- c. Lladdodd y ddraig y dyn (galés: VSO)
mató el dragón el hombre
«el dragón mató al hombre».
- d. Nahita ny mpianatra ny vehivahy (malgache: VOS)
vio el estudiante la mujer
«la mujer vio al estudiante».
- e. Toto yahosiye kamara (hixkaryana: OVS)
Hombre cogiólo jaguar
«el jaguar cogió al hombre».

Todas estas oraciones son declarativas, sus verbos están en la voz activa, suelen proferirse con una entonación neutra, sin enfatizar ningún constituyente en particular y, según demuestran diversos estudios sobre cada una de las lenguas, ejemplifican el orden más frecuente en textos o diálogos. Evidentemente, en algunos de estos casos este orden será el único posible para una oración declarativa (como ocurre en galés), mientras que en otros será simplemente un orden preferido (como ocurre en castellano); no debemos confundir el hecho de que todas las lenguas tienen un orden básico con el hecho de que para algunas lenguas ese orden básico es obligatorio en muchos contextos.

Hay otro sentido en el cual podemos decir que dichas oraciones presentan un orden básico que se refiere a cómo se distribuye la información que éstas comunican. Toda oración, en el contexto de un discurso, puede analizarse de acuerdo con las dos categorías pragmáticas de *información dada* o *tema* e *información nueva* o *rema*. El tema es aquello de lo que trata la oración y es, por tanto, información conocida tanto por el hablante como por el oyente; el rema, en cambio, es lo que se enuncia acerca de ese tema y coincide, por tanto, con la información nueva que el hablante aporta sobre el tema (también es frecuente utilizar el término *foco* para designar la información nueva). Cada lengua tiene diferentes estrategias para distinguir tema y rema, y la más importante de estas estrategias (además de la entonación) es el orden de palabras, de modo que el orden básico coincide también con una estructuración básica de tema y rema. Por ejemplo, en castellano el orden SVO es también básico en el sentido de que el tema suele coincidir con el sujeto y el rema con todo el SV (es decir, el verbo y los demás complementos), y este orden coincide con la generalización de que, en esta lengua, la información dada siempre precede a la información nueva. En cuanto a la estructura del rema, que casi siempre es complejo, podemos decir que éste suele contener unos elementos de enlace con el tema (generalmente el verbo) y un elemento más prominente, sobre el cual recae un ascenso de la entonación, que suele colocarse al final.

Estas observaciones nos permiten dar cuenta de un buen número de los órdenes

posibles en una lengua con un orden relativamente libre como el castellano. Así, por ejemplo, en esta lengua el sujeto no siempre precede al verbo, sino que puede ir al final:

- (60) a. Ha llegado Gabriel.
b. Gabriel ha llegado.

En (60.a) el sujeto va al final y, de acuerdo con lo dicho, forma parte del rema. En efecto, esta oración podría ser la respuesta a una pregunta como *¿Qué ha pasado?*, de modo que todos los constituyentes de (60.a) son información nueva (no hay, pues, tema) y, para evitar que el sujeto se analice como tema, éste se sitúa al final; alternativamente, (60.a) podría ser la respuesta a una pregunta como *¿Quién ha llegado?*, donde el verbo es información dada (el tema; es decir, se sabe que ha llegado alguien, pero no quién), mientras que el sujeto es el rema y se sitúa al final respetando el orden tema-remas. Desde este punto de vista, (60.a) es ambigua, aunque todo hablante del castellano sabe distinguir una interpretación de otra en función de la entonación, que, en muchas ocasiones, suele ser la única pista, junto con el contexto, que nos permite distinguir una lectura de otra. Por otra parte, (60.b) sólo podría ser la respuesta a una pregunta como *¿Qué ha pasado con Gabriel?* (o *¿Qué ha hecho Gabriel?*), donde el sujeto es tema y precede al verbo, que es rema; alternativamente, en este caso el sujeto podría también omitirse, de modo que la respuesta podría ser simplemente *Ha llegado*. De modo similar, podemos dar cuenta del orden de los constituyentes dentro del rema si consideramos que el elemento más prominente es siempre el elemento más oblicuo, en el sentido de que cuanto más débil es el lazo con el verbo más oblicuo es un elemento. Así, los adjuntos son más oblicuos que los complementos y, dentro de los complementos, los complementos preposicionales son más oblicuos que el complemento indirecto, que, a su vez, es más oblicuo que el complemento directo:

- (61) a. Gabriel regaló un disco a Laura por su cumpleaños.
b. ?Gabriel regaló a Laura un disco por su cumpleaños.
c. ?Gabriel regaló por su cumpleaños un libro a Laura.
d. ??Gabriel regaló por su cumpleaños a Laura un libro,

donde los interrogantes indican un menor grado de aceptabilidad de los órdenes que escapan a la jerarquía de oblicuidad dentro del rema. Evidentemente, la prominencia puede alterarse si, por ejemplo, uno de los elementos es mucho más complejo (más largo) que los demás, en cuyo caso la jerarquía de oblicuidad no se respeta:

- (62) a. Gabriel regaló a Laura por su cumpleaños un disco pirata de Nirvana grabado en 1990.
b. ??Gabriel regaló a Laura un disco pirata de Nirvana grabado en 1990 por su cumpleaños.
c. ??Gabriel regaló un disco pirata de Nirvana grabado en 1990 a Laura por su cumpleaños.

Así, apelando a categorías pragmáticas como tema y rema podemos dar cuenta de las variaciones de orden que se observan en las distintas lenguas, principalmente en

aquellas en que el caso o la concordancia hacen que no sea preciso fijar un orden de los constituyentes para identificar las diferentes relaciones gramaticales. La estructura sintáctica es, pues, el mecanismo básico a que recurren las lenguas para expresar: (I) las relaciones semánticas que se establecen entre un predicado y sus argumentos; (II) las relaciones funcionales o gramaticales que se establecen entre un núcleo y sus complementos y adjuntos; y (III) la *estructura informacional* de una oración (la información nueva y la información dada).

Profundicemos un poco más, pues, en la relación entre estructura sintáctica y estructura informacional (utilizamos el neologismo «informacional» en lugar del término «funcional» que suele utilizarse en estos casos, porque nos parece más apropiado ya que, en cierto modo, la función de cualquier enunciado es siempre la de transmitir información —cierta o falsa—; al hablar de estructura informacional recogemos con más precisión el hecho de que lo que nos ocupa es cómo se organiza la información en las oraciones y los enunciados, es decir, en los vehículos de transmisión de la información).

4.2. ESTRUCTURA INFORMACIONAL Y ESTRUCTURA SINTÁCTICA

La estructura sintáctica puede codificar, pues, la estructura informacional de una oración, lo cual se lleva a cabo mediante alteraciones del orden básico de palabras, el cual representa, a su vez, una organización básica de las categorías pragmáticas tema y rema. En relación a esta estructura hemos estudiado el comportamiento del sujeto cuando éste es tema (anteposición u omisión) y cuando es rema o forma parte de la información remática (posposición). Sin embargo, en los casos estudiados o el tema era el sujeto o no había tema o lo era el verbo, pero ¿qué ocurre cuando el tema es cualquier otro complemento o adjunto? Evidentemente la única alternativa posible es que se produzca una alteración de la estructura sintáctica, ya que la estructura informacional no es directamente observable, sólo lo es, de hecho, a través de la organización de los constituyentes en la oración; no olvidemos que es la estructura sintáctica la que codifica la estructura informacional y no al revés. Por tanto, en castellano se tenderá a respetar el orden tema-remas en la medida de lo posible y, cuando el tema sea el complemento directo, o cualquier otro complemento o adjunto, éste se antepondrá al verbo. El castellano posee una estructura sintáctica especial para expresar este tipo de tematizaciones por anteposición: la dislocación con clítico, que consiste en desplazar hacia la izquierda un complemento, reiterándolo con un pronombre personal en el caso correspondiente (un clítico):

- (63) a. [TEMA El disco, Gabriel] [REMA lo regaló a Laura para su cumpleaños].
b. [TEMA A Laura, Gabriel] [REMA le regaló un disco para su cumpleaños].
c. [TEMA El disco, a Laura, Gabriel] [REMA se lo regaló para su cumpleaños].
d. [TEMA El disco, a Laura] [REMA se lo regaló Gabriel para su cumpleaños].

En (63) hemos indicado entre corchetes los dos componentes de la estructura informacional. Como se puede observar, más de un constituyente puede ser tema y la tematización de los complementos se marca con un clítico en el rema (las negritas y los subrayados pretenden expresar esta relación). Por otra parte, cuando el sujeto no es

tema, éste se pospone al verbo, como ocurre en (63.d). Nótese que los ejemplos de (63) se proferirían sin apenas diferencias con oraciones declarativas en cuanto a la entonación, con la única salvedad de que los temas se suelen separar por breves pausas (que hemos indicado con comas) para distinguirlos los unos de los otros. En principio, la tematización por anteposición es posible con cualquier elemento, incluidos los adjuntos, aunque, en estos casos, el castellano no dispone de los pronombres adecuados para reiterarlos y sólo la entonación nos sirve de ayuda para distinguir estas anteposiciones sin clítico de otras anteposiciones que en seguida examinaremos. Efectivamente, en castellano la anteposición de un constituyente no es patrimonio exclusivo de la tematización, sino que también se produce cuando el rema se corresponde con un único constituyente. En estos casos, se altera la tradicional estructura tema-remata en favor de una estructura remata-tema marcada con una entonación especial:

- (64) a. GABRIEL regaló un disco a Laura por su cumpleaños.
 b. UN DISCO regaló Gabriel a Laura por su cumpleaños.
 c. A LAURA regaló Gabriel un disco por su cumpleaños.
 d. POR SU CUMPLEAÑOS regaló Gabriel un disco a Laura.

Hemos utilizado versalitas para marcar los constituyentes rematizados y como indicación de que la entonación es marcada. Nótese que en estos casos el clítico no puede aparecer y que la anteposición aquí se utiliza para expresar un contraste, es decir, que fue Gabriel, y no otra persona, quien le regaló un disco a Laura por su cumpleaños, que fue un disco, y no otra cosa lo que Gabriel le regaló a Laura por su cumpleaños, etc. Un aspecto fundamental de la tematización por anteposición es que, a diferencia de la tematización, sólo se puede anteponer un constituyente (lo cual explica, por otra parte, que la posposición del sujeto sea obligatoria en estos casos):

- (65) *UN DISCO, A LAURA, GABRIEL regaló por su cumpleaños.

Otro aspecto interesante de las anteposiciones en castellano es que tematización y rematización pueden coexistir, dando lugar a organizaciones más complejas de la estructura básica tema-remata, donde el tema se divide en dos partes que envuelven al rema, que queda ocupando una posición central:

- (66) a. [TEMA Un disco, a Laura,] [REMA GABRIEL] [TEMA se lo regaló para su cumpleaños].
 b. [TEMA Gabriel, a Laura,] [REMA UN DISCO] [TEMA le regaló para su cumpleaños].

A la luz de estas observaciones de la relación entre estructura sintáctica y estructura informacional, podemos considerar desde una nueva perspectiva la función de los procesos de alteración de la valencia, así como otros procesos en los que el orden de palabras básico se ve alterado que todavía no hemos considerado. En efecto, recordemos que todo proceso de alteración de valencia puede entenderse, a grandes rasgos, como un proceso que comporta la degradación o la supresión de un argumento externo y la consiguiente promoción de un argumento interno; por ejemplo, la pasiva comportaría la degradación del argumento externo, el cual se proyecta entonces como un complemento preposicional, al tiempo que se promociona el argumento interno, que

se proyecta como sujeto. La pasiva es, pues, un tipo de tematización, aunque siguiendo una estrategia sintáctica distinta de la mera anteposición hacia la izquierda. ¿Existe alguna otra diferencia fundamental entre una tematización por pasivización y una tematización por anteposición? Sí, ya que la pasivización es un proceso que podríamos calificar de *local*, en el sentido de que sólo puede darse dentro del dominio definido por un predicado y sus argumentos; por otra parte, la anteposición no lo es, de modo que entre el constituyente antepuesto y el clítico que lo reitera pueden mediar tantos dominios locales como se desee:

- (67) a. Gabriel vendió el coche.
 b. El coche fue vendido por Gabriel.
 c. El coche, Gabriel lo vendió.
 d. El coche, cuentan que Laura dice que Gabriel lo vendió.
 e. *El coche es contado que Laura dice que vendió por Gabriel.

En (67.e) hemos intentado construir un ejemplo de pasiva no-local, lo que, evidentemente, es imposible; (67.d), por otra parte, muestra un caso de anteposición no local, que es perfecta. Casos como los de (67.d) no son exclusivos de la tematización, sino que también la rematización puede ser no local:

- (68) UN COCHE cuentan que Laura dice que vendió Gabriel.

Hasta ahora, hemos centrado nuestra atención en oraciones declarativas y en las posibilidades que nos da la sintaxis de codificar la estructura informacional a través de alteraciones del orden de palabras de las mismas. No querríamos acabar, no obstante, sin considerar, aunque sea brevemente, las relaciones entre estructura sintáctica e informacional en otros tipos de oraciones no declarativas, concretamente, en las interrogativas. Consideremos, en primer lugar, cuál es la función de una oración interrogativa. Evidentemente, preguntar, es decir, solicitar información sobre algo que se desconoce. Nuestro desconocimiento sobre una determinada situación puede ser muy variado: podemos ignorar si una determinada afirmación es verdadera o si una determinada situación ha llegado a producirse, en cuyo caso esperamos que se nos responda con un sí o con un no y para preguntar utilizamos lo que se denomina una *interrogativa total*:

- (69) ¿Ha llegado Gabriel?

Nótese que en las interrogativas totales, además de una entonación especial, el sujeto suele ir pospuesto en una configuración idéntica al caso en que el sujeto formaba parte del rema. En sentido estricto, aquí no podemos hablar de información nueva e información dada, ya que, en cierto modo, toda la oración expresa información desconocida; parece, sin embargo, que la información desconocida se comporta como la información nueva. Veamos lo que ocurre con otro tipo de oraciones interrogativas. Cuando nuestro desconocimiento abarca únicamente a uno de los elementos de una situación, entonces recurrimos a las *interrogativas parciales* para solicitar información sobre este punto. Así, podemos preguntar sobre el sujeto, el complemento directo, etc.:

- (70) a. ¿Quién ha regalado un disco a Laura por su cumpleaños?
 b. ¿Qué ha regalado Gabriel a Laura por su cumpleaños?
 c. ¿A quién ha regalado Gabriel un disco por su cumpleaños?
 d. ¿Cuándo ha regalado Gabriel un disco a Laura?

Si comparamos las interrogativas de (70) con las rematicaciones de (64) vemos que su estructura es idéntica, con la única diferencia de que ahora el elemento desplazado es un pronombre interrogativo. Ello parece indicar que, efectivamente, la información desconocida se comporta igual que la información nueva (o quizá que aquella es un tipo especial de información nueva), lo que explicaría que las oraciones interrogativas, dada su naturaleza, siempre poseen una estructura sintáctica en la que se codifica un proceso de rematicación.

Aquí acabamos, y sólo nos falta intentar introducir en nuestra gramática los mecanismos necesarios para tratar las alteraciones del orden de palabras. En este caso, sin embargo, deberemos quedarnos en un nivel intuitivo en muchos casos, ya que, con los desplazamientos no locales, hemos entrado en un territorio de gran complejidad formal, lo cual nos obliga a ser un poco menos precisos.

4.3. GRAMÁTICA III

Hasta el momento, las reglas de estructura de frase que hemos utilizado expresaban directamente las relaciones de dominio y de precedencia, de modo que el orden relativo de los constituyentes queda directamente fijado por la regla. Sin embargo, hemos visto que el orden relativo de los complementos, por ejemplo, puede ser variable en función de diversos factores. Existe una técnica bastante simple para dar cuenta de las relaciones de orden locales, que no comporta escribir una regla para cada orden posible. Supongamos, por tanto, que la información concerniente a las relaciones de dominio se expresa mediante un tipo de regla, mientras que la información concerniente a la precedencia se expresa mediante otro tipo de regla. Por ejemplo, la regla que introduce el sujeto y el SV podemos ahora escribirla como en (71.a), más la regla de precedencia de (71.b):

- (71) a. $O \rightarrow SN[NOM], SV.$
 b. $TEMA < REMA.$

En (71.a) separamos por una coma los constituyentes a la derecha de la regla para indicar que no están ordenados. En (71.b) tenemos un nuevo tipo de regla, que expresa una restricción de orden: el tema debe preceder al rema (el símbolo «<» significa «precede a»). De modo parecido podemos dar cuenta del orden relativo de los constituyentes dentro del SV:

- (72) a. $SN \rightarrow V[S], SN[NIL, AC \vee PART], SN[A, DAT].$
 b. $NÚCLEO < COMPLEMENTOS.$
 c. $COMPLEMENTO[-oblicuo] < COMPLEMENTO[+oblicuo].$

Con esta técnica podemos ir dando cuenta de un buen número de restricciones de orden, aunque tiene un defecto casi insalvable que la hace muy poco conveniente para

nuestros propósitos: las restricciones de orden que impone son absolutas y, precisamente, hemos podido comprobar que una restricción como la de (72.c), por ejemplo, no describe más que una tendencia que no tiene por qué respetarse cuando intervienen otros factores como la complejidad de un complemento. Otro problema es que la regla (71.b) permite la posposición del sujeto, pero lo sitúa al final del SV y no entre el verbo y los complementos como ocurre cuando no es un tema.

Algunos de estos problemas los podemos resolver, en parte, si alteramos nuestra interpretación de las reglas de precedencia en tanto que restricciones absolutas y consideramos, por el contrario, que son indicativas de una tendencia. Podemos, por ejemplo, ordenarlas de acuerdo con un orden de preferencia, de modo que pueda haber reglas contradictorias cuya aplicación se efectúe de acuerdo con el orden establecido. Una posibilidad sería, pues, la que esbozamos en (73):

- (73) $\left\{ \begin{array}{l} \text{Tema} < \text{Rema} \\ \text{Núcleo} < \text{Complementos} \\ \text{Complemento} [-oblicuo] < \text{Complemento} [+oblicuo] \\ \text{Complemento} [-complejo] < \text{Complemento} [+complejo] \end{array} \right\}$

Así pues, lo que tenemos en (73) es, de hecho, una única regla de precedencia compleja, en la que se indican las diferentes tendencias en la lengua ordenadas en función de qué tendencias son las preferidas. Por ejemplo, las dos últimas reglas son potencialmente contradictorias, pero la primera representa la tendencia preferida, la menos marcada, mientras que la segunda recogería aquellos casos, también posibles, que escapan a la opción preferida.

Sin embargo, sea cual sea la opción que adoptemos, no parece posible hoy por hoy que un análisis del orden de palabras basado únicamente en reglas de preferencia sea capaz de dar cuenta de todos los casos posibles. Para acabar este capítulo quisiéramos introducir, aunque sea de forma intuitiva, un nuevo tipo de regla, más poderoso que cualquiera de los que hemos utilizado hasta ahora, y que podría ser una vía de solución para muchas de las cuestiones relacionadas con el orden de palabras que hemos estudiado aquí: las *transformaciones*.

Una transformación es una regla capaz de derivar una representación sintáctica a partir de otra representación sintáctica, generalmente «desplazando» constituyentes a una posición distinta de la que ocupaban, aunque algunos autores han considerado o consideran transformaciones con la capacidad de suprimir o introducir estructura durante la derivación de la nueva representación. En principio, una transformación tiene el poder de efectuar cualquier operación sobre la representación de entrada, lo cual la convierte en un arma de doble filo si nuestro objetivo es el de generar única y exclusivamente las oraciones bien formadas de una lengua; si se hace uso de transformaciones, es imprescindible, por tanto, idear mecanismos capaces de restringir su poder. Por lo general, la estrategia más común para restringir el poder de las transformaciones consiste, precisamente, en suponer que sólo son capaces de realizar desplazamientos de constituyentes, dejando, además, una *huella* de su aplicación. Para aclarar mejor este punto, considérese la figura 8, donde, gracias a las transformaciones, damos cuenta de la colocación del pronombre interrogativo y de la «inversión» sujeto-verbo:

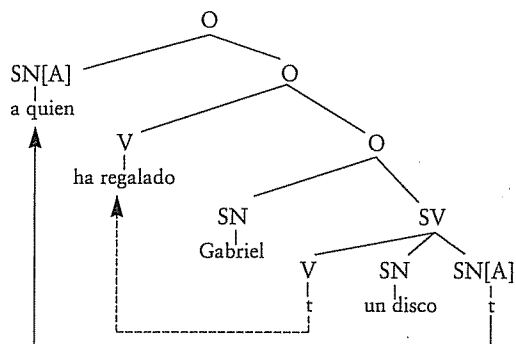


Figura 8

Las flechas indican la dirección del desplazamiento fruto de la aplicación de una transformación, y lo que hemos simbolizado como *t* son las huellas que ese desplazamiento ha dejado en la posición que ocupaba originalmente el constituyente desplazado (utilizamos *t* porque es el símbolo más comúnmente utilizado por los lingüistas; es una abreviatura del inglés «trace» [huella]). Las huellas constituyen una clase especial de «palabras» cuyo contenido fonológico es nulo (es decir, son mudas) y cuyo contenido gramatical está en función del contenido gramatical del constituyente que está relacionado con ellas; pertenecen a una clase de elementos denominados genéricamente *categorías vacías*. Nótese que las huellas cumplen una función muy importante, ya que nos permiten conservar información como el caso o el papel semántico de un constituyente desplazado: por ejemplo, el constituyente *a quién*, pese a haberse desplazado, sigue respetando las restricciones de subcategorización que le impondría el verbo (que también se ha desplazado) si se encontrara en la posición ocupada ahora por la huella. Las huellas son, pues, una indicación de que parte de la información gramatical (pero no la fonológica) de un constituyente está «en dos sitios a la vez» después de haber aplicado una transformación. Para facilitar la representación de estructuras sobre las que se han aplicado transformaciones y para indicar este hecho, en vez de flechas se suelen utilizar subíndices, como se muestra en la figura 9:

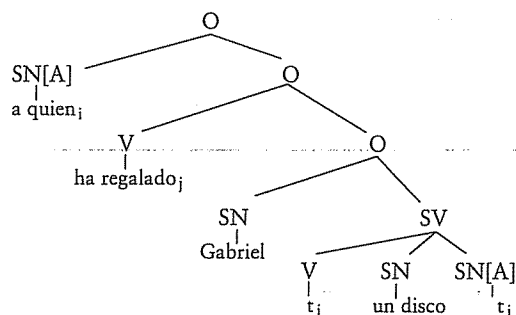
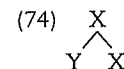


Figura 9

Gracias a este mecanismo de transformaciones, huellas y subíndices, podemos desplazar un constituyente a una posición tan alejada de su lugar de origen como queramos, sin nunca perder la información relevante, ya que los subíndices garantizan que se mantenga la conexión entre el elemento desplazado y su huella. Pasemos a considerar ahora un detalle que todavía no hemos comentado: la colocación de los elementos desplazados en la nueva estructura. Como se puede apreciar tanto en la figura 8 como en la figura 9, situamos estos elementos en una posición externa al nodo O original y creamos un nuevo nodo O que domina a éste y al sintagma desplazado. Es decir, hemos creado una subestructura como la que representamos genéricamente en (74):



Estas estructuras, en las que un nodo X domina a otro nodo X idéntico y a un nodo Y, se denominan *estructuras de adjunción* y, desde el punto de vista formal, podemos crearlas de dos maneras distintas. Por un lado, podemos suponer que es la propia aplicación de la transformación la que las crea, es decir, que la transformación, además de desplazar un elemento y dejar una huella, tiene la capacidad de crear el espacio estructural que el elemento desplazado ocupará al final del desplazamiento. Alternativamente, podemos suponer que dicho espacio estructural existía previamente, definiendo una serie de «puntos de aterrizaje» que se llenarán sólo si se desplaza algún elemento. En principio, ambas alternativas son equivalentes desde el punto de vista formal, aunque la segunda tiene alguna potencial ventaja como, por ejemplo, el hecho de quitar a las transformaciones el poder de construir estructura. Asimismo, nos da la posibilidad de prever el número de «puntos de aterrizaje», que podemos limitar a dos, tres, cuatro, etc., mientras que en la primera alternativa es potencialmente infinito. En cualquier caso, el inclinarnos por una u otra opción es, en última instancia, algo que debemos determinar empíricamente, observando los datos.

Para acabar, podemos hacer una última observación sobre los dos mecanismos que hemos examinado en esta sección, las reglas de precedencia y las transformaciones. Como habrá podido observarse, en el fondo, ambos mecanismos no son tan diferentes como en un principio podría parecer. En efecto, si interpretamos las reglas de precedencia como mecanismos con el poder de ordenar constituyentes dentro de un dominio local (el definido por un nodo madre y sus nodos hijos), vemos que éstas se comportan igual que si fueran un tipo especial de transformaciones, cuyo dominio de aplicación estuviera restringido precisamente a ese espacio estructural; las transformaciones, por el contrario, serían reglas de reordenación de constituyentes capaces de superar ese dominio local en determinadas circunstancias. Es posible, pues, que la combinación de ambos tipos de reglas con un conjunto relativamente pequeño de reglas de estructura de frase nos permita dar cuenta de los diferentes órdenes de palabras de cada lengua.

5. CONCLUSIÓN

En este capítulo hemos desarrollado algunos aspectos básicos de los que debe ocuparse toda teoría de la sintaxis de las lenguas naturales. Partiendo de las nociones más

elementales de estructura jerárquica y de constituyente, hemos ido introduciendo algunos de los problemas empíricos más relevantes y, con ellos, hemos ido presentando algunos mecanismos formales para dar cuenta de los mismos en la construcción de un modelo de la gramática. Al hacer esto, hemos optado por un enfoque bastante ecléctico, tomando prestadas de aquí y de allá ideas y herramientas formales que suelen considerarse más propias de un modelo teórico o de otro. Sin embargo, esto no es una introducción a un modelo teórico concreto, sino un intento de presentar una metodología general de análisis de los datos empíricos y de la construcción paulatina de un modelo formal capaz de dar cuenta de los mismos. Es éste el motivo por el cual hemos intentado separar el tratamiento de los problemas descriptivos de todo cuanto concierne a la formalización de esos problemas utilizando un lenguaje y unas herramientas teóricas determinados. El lector no tiene que olvidar, sin embargo, que dicha separación no es más que una ilusión y, en todo caso, una simplificación necesaria que se justifica solamente por las exigencias pedagógicas y expositivas que impone un texto como el presente. No existe una sintaxis neutral (como no existe una ciencia neutral) ni ésta acaba con los problemas y las herramientas que aquí hemos presentado; los límites los pone la imaginación y el ingenio del lingüista.

EJERCICIOS

1. A lo largo de este capítulo hemos ido construyendo una pequeña gramática del castellano; sin embargo, ésta no es, ni mucho menos, completa. El objetivo de este ejercicio es el de completarla partiendo de nuevos datos. La lista de datos es la siguiente:

1. Juan rompió la ventana con un martillo.
2. Juan cogió la bola.
3. Juan entendió el teorema.
4. Hicimos una foto a María.
5. Mario fue encarcelado por el juez a fin de evitar la alarma social.
6. Los asistentes obsequiaron a María con joyas.
7. María recibió un anillo de diamantes de Juan.
8. Juan dio a María un anillo de diamantes.
9. La madre nutre al hijo con su leche.
10. Los asistentes sorprendieron a María (con un regalo).
11. María sorprendió al ladrón (en la cocina).
12. Juan envió un paquete a María.
13. Las declaraciones del ministro sorprendieron a Pilar.
14. El juez combate por la justicia.
15. Carlos siempre ha combatido la tiranía.
16. Juan combate con Pedro voluntariamente.
17. Juan combate a Pedro voluntariamente.
18. Rogelio rodó colina abajo para divertirse.
19. Rogelio rodó colina abajo tras ser alcanzado por un disparo.
20. Fernando pulió la bandeja.