

# Linguistische Arbeiten

494

Herausgegeben von Peter Blumenthal, Klaus von Heusinger,  
Ingo Plag, Beatrice Primus und Richard Wiese



# Variación sintáctica en español

Un reto para las teorías de la sintaxis

Editado por

Gabriele Knauer y Valeriano Bellosta von Colbe

Max Niemeyer Verlag  
Tübingen 2005



Bibliografische Information der Deutschen Bibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

ISBN 3-484-30494-4      ISSN 0344-6727

© Max Niemeyer Verlag GmbH, Tübingen 2005

<http://www.niemeyer.de>

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen. Printed in Germany.

Gedruckt auf alterungsbeständigem Papier.

Druck: Laupp & Göbel GmbH, Nehren

Einband: Nädle Verlags- und Industriebuchbinderei, Nehren

## Índice

*Gabriele Knauer/Valeriano Bellosta von Colbe*

La variación sintáctica como reto teórico:

Una introducción ..... 1

*Guido Mensching*

Variación sintáctica, lingüística de corpus y gramática generativa:

Teorías, métodos y problemas..... 13

*Josse De Kock*

Corpus, frecuencia y gramaticalidad:

Treinta construcciones contrincantes en tres corpus ..... 35

*Nicole Delbecque*

El análisis de corpus al servicio de la gramática cognoscitiva:

Hacia una interpretación de la alternancia lineal SV/VS ..... 51

*Alicia González de Sarralde*

Sobre la posición del sujeto en relación con las estructuras de la narración ..... 75

*Valeriano Bellosta von Colbe*

Variación sintáctica en »Role and Reference Grammar«:

La posición de los complementos en las oraciones ditransitivas..... 97

*Pedro Martín Butragueño*

La construcción prosódica de la estructura focal en español..... 117

*Amparo Morales*

La adquisición del lenguaje en niños puertorriqueños:

A propósito de la hipótesis del sujeto nulo..... 145

*Ulrich Detges*

La gramaticalización de los acusativos preposicionales en las lenguas iberorrománicas:

Una hipótesis pragmática ..... 155

*Eugeen Roegiest*

Variación pronominal en español:

El pronombre dativo entre sintaxis y semántica..... 175

*Rena Torres Cacoullós*

La perspectiva diacrónica en variación sincrónica:

El dativo de intensificación ..... 191

VI

*Irania Malaver*

Análisis comparativo de las expresiones adjetivales de edad  
en las hablas de Sevilla y de Caracas .....211

*Dexy Galué*

*Me acuerdo que...*: los verbos pronominales y el queísmo.....221

*Gabriele Knauer (Humboldt-Universität zu Berlin)*  
*Valeriano Bellosa von Colbe (Universität zu Köln)*

## La variación sintáctica como reto teórico: Una introducción

### 1. Introducción

Sin duda, una de las controversias más antiguas en la historia de la lingüística es la que tuvo y sigue teniendo lugar entre «anomalistas» y «analogistas». Mientras que los primeros ven en la lengua una realidad esencialmente irregular, a pesar de algunas aparentes regularidades, los segundos la conciben como esencialmente regular, aunque plagada de irregularidades. Tal controversia sigue viva hoy entre quienes conciben la lengua ante todo como el sistema abstracto que «emerge» del conjunto de preferencias que se da en la comunicación humana dentro de una comunidad («actuación» o «lenguaje E») y quienes la ven sobre todo como el conocimiento autónomo, en parte innato, de un determinado código que un ser humano ideal posee en su cognición («competencia» o «lenguaje I»; cf. Chomsky 1965; 1986). Los primeros, los «funcionalistas», son más propensos a admitir la existencia de variación sintáctica en una lengua y realizan sus observaciones a partir de grandes corpus de datos empíricos procedentes de la comunicación; mientras que los segundos, los «generativistas», prefieren pasarla por alto y se apoyan en la evidencia psicolingüística procedente de la introspección de uno o varios hablantes en su propia cognición (cf. Mensching en este volumen). Es interesante observar que incluso aquellas teorías que se dicen funcionalistas y parten, en principio, de la lengua como realidad que refleja igualmente la cognición y la comunicación humanas, son reacias a aceptar la existencia de variación sintáctica en una lengua y, a veces, a trabajar sistemáticamente con datos empíricos (cf. Martín Butragueño 1999). El resultado de esta tendencia es que el constante descubrimiento de fenómenos de variación en las lenguas a partir de datos empíricos apenas se refleja en las descripciones de esas lenguas y menos todavía en las teorías lingüísticas, que, a fin de cuentas, deberían explicar tales fenómenos.

Las gramáticas «descriptivas» han evitado emitir juicios de valor y se han limitado a su tarea de descripción, relacionando los fenómenos de variación sintáctica con factores extralingüísticos diatópicos y diastráticos (cf. Serrano 1999). Pero tanto las clásicas gramáticas «normativas» como las actuales gramáticas «explicativas» han entendido la variación sintáctica en una lengua bien como «opción» («libertad estilística» del hablante), bien como «irregularidad» («excepción» en el sistema), bien como «incorrección» («violación» del sistema por parte del hablante). El primero de los casos llega a confundirse a veces con «arbitrariedad» porque toca fenómenos generalizados que se consideran fuera de la gramática y de sus reglas y que están sometidos a las necesidades comunicativas del momento (p. ej. el «orden de palabras»; cf. Delbecque y González de Sarralde en este volumen). Los otros dos casos (cf. De Kock en este volumen) se encuentran dentro de la gramática y de sus reglas. Si la variación toca un fenómeno generalizado pero circunscrito a un determinado entorno sintáctico fácilmente descriptible, se habla de «irregularidad» (p. ej. el acusativo

preposicional; cf. Detges en este volumen). Si la variación toca un fenómeno particular, propio de un grupo determinado de hablantes, que no está circunscrito a un entorno sintáctico fácilmente descriptible, se habla de «incorrección» (p. ej. el dequeísmo; cf. Galué en este volumen). Mientras que las gramáticas «normativas» concebían tales «incorrecciones» como el producto de una deficiente alfabetización y escolarización, debida a factores socioeconómicos y políticos, y pretendían contribuir a remediarlas; las gramáticas «explicativas» los conciben como el resultado bien de *lapsus linguae* casuales y asistemáticos en los adultos, bien de una deficiente adquisición de la lengua en los niños, debida a la «pobreza del estímulo» ofrecido por los adultos. Muy distinto es, sin embargo, el caso de los fenómenos de variación externos a la sintaxis: su estudio conlleva el uso de etiquetas mucho menos estigmatizadas; se habla de «alófonos» en fonología, de «alomorfos» en morfología, de «sinónimos» en semántica o de «norma» en teoría lingüística (cf. Coseriu 1952).

## 2. Variación

Si nos acogemos al más sencillo de los modelos semióticos, el de Saussure (1916), comprobaremos que incluye de forma implícita un principio de economía, que sólo años más tarde se desarrollará explícitamente y se aplicará universalmente (cf. Jespersen 1924): a cada significante le corresponde un solo significado; a cada significado, un solo significante. Bien es sabido que la lengua quiebra esa biunivocidad pretendidamente ideal en la relación entre significante y significado en las dos direcciones posibles. La polisemia permite combinar un significante con varios significados al precio de complicar el descifrado del mensaje con la resolución de ambigüedades. La sinonimia en sentido fuerte permitiría combinar un significado con varios significantes al precio de complicar el cifrado del mensaje con la resolución de redundancias. Mientras que la polisemia se entiende como un fenómeno ventajoso para la economía de la lengua en todos sus niveles, la sinonimia recibe una valoración desigual, dependiente del nivel estudiado, cosa que se concreta en denominaciones distintas.

Una visión onomasiológica de la sinonimia nos permite descubrir paralelismos interesantes entre la primera y la segunda articulaciones del lenguaje (cf. Martinet 1960). Mientras que, en la primera articulación del lenguaje, el significado lingüístico se ve como «contenido» delimitado por su «valor» en el sistema; en la segunda, el significante lingüístico se puede reinterpretar como «forma», posición estructural o mero «valor» en el sistema («forma» del significado y del significante respectivamente para Hjelmslev 1943). Así, tanto el «contenido» como la «forma» se pueden concebir como significado con respecto a realidades distintas: el «contenido» es el significado para el significante; la «forma» es el significado para la sustancia fónica. Tal significado, cuando se asigna a varios significantes, abre un espacio de variación en la lengua que se concreta en el habla por medio de la selección de una de las posibles variantes. La lingüística no sólo se pregunta, pues, cuántas y cuáles son las variantes legítimas para una variable definida a través del significado, sino cómo y por qué se originan y, sobre todo, qué utilidad tienen dentro del sistema de la lengua.

La lingüística estructural ha aceptado y explicado la variación en la segunda articulación como empleo de alófonos en el habla, los cuales ocupan la misma posición en el sistema («no oposición»), lo que ha mantenido abierta la comunicación con otras disciplinas lingüísticas que estudian la variación del habla en diacronía y en sincronía. Sin embargo, aunque ha explicado la variación en la primera articulación de manera análoga, la ha valorado como un fenómeno desventajoso para la economía de la lengua, cuya presencia ha de restringirse al máximo. Sobre todo cuando los significantes son fonológicamente similares o están etimológicamente emparentados, la morfología estructuralista habla, de modo análogo a la fonología, de empleo de alomorfos en el habla, los cuales ocupan la misma posición en el sistema. Sin embargo, cuando no se dan estas condiciones, la semántica estructuralista sustituye el concepto de alomorfismo por el de sinonimia en sentido débil. Mientras que la sinonimia en sentido fuerte sería la combinación de un mismo haz de rasgos semánticos, es decir, de una misma posición dentro del sistema, con varios significantes, que, por lo tanto, no se oponen entre sí; la sinonimia en sentido débil sería la combinación de varios haces de semas idénticos y, por tanto, varias posiciones idénticas dentro del sistema, con sus correspondientes significantes. Es en este momento cuando aparece la dicotomía entre sinonimia total y parcial, es decir, la distinción entre haces idénticos en todos sus semas y haces idénticos en la mayoría de sus semas (bajo la influencia de los conocimientos léxico y enciclopédico y de la situación comunicativa); perspectiva que invita a una interpretación de la sinonimia en sentido débil como noción prototípica. De esta manera, el pretendido estudio de la sinonimia en sentido fuerte, es decir, de la variación, se convierte en la demostración de su imposibilidad y en el restablecimiento de la pretendida biunivocidad de significante y significado, que se apoya en el principio de economía mencionado al inicio. He aquí, pues, la «paradoja teórica de la variación»: el sentido fuerte del concepto de sinonimia invita al estudio de la variación en diacronía y sincronía; el sentido débil, lo bloquea. No podemos, pues, llamar «variación» a ambos sentidos; la variación referida a una misma «forma» prefiere el sentido fuerte de la sinonimia (en fonología y morfología) y se opone a la variación referida al «contenido», que prefiere el sentido débil de la sinonimia (en semántica y sintaxis). Por eso, sólo en este último caso, se estudian implícitamente la diferenciación y especialización semánticas de los significantes y la pérdida progresiva de sinonimia que resulta de estos procesos.

### 3. Variación sintáctica

La lingüística estructural europea se ocupó poco de una sintaxis que, hasta aquel punto, había vivido de una tradición semántica, centrada en el «monema». La lingüística estructural norteamericana se dedicó menos todavía a una variación sintáctica, que, inserta en el eje paradigmático, servía para la definición de clases sin distinción de significados. Esto propició la asunción inconsciente de la «paradoja teórica de la variación» y su perpetuación tras el «giro cognitivo y comunicativo» manifestado en los paradigmas teóricos generativista y funcional. Dos alternativas se presentaban. Por un lado, el paradigma generativista declara que el objeto propio de la lingüística es el estudio de la lengua I o competencia, tal como se

da en un hablante ideal, y que, por lo tanto, el habla o la actuación, tal como aparece en las manifestaciones lingüísticas de una comunidad real, al estar posiblemente sujeta a toda clase de perturbaciones, debe dejarse de lado. De este modo, la variación, si pertenece a la actuación, queda subsumida en las perturbaciones que le son propias; si pertenece a la competencia, produce tantas gramáticas (variación interlingüística) o tantas reglas transformacionales de aplicación optativa en una gramática (variación intralingüística) como variantes identificadas (cf. Chomsky 1965; Mensching, en este volumen).

Por otro lado, el paradigma funcionalista rechaza la posición anterior y declara que el habla o la actuación es el objeto propio de la lingüística en la medida en que ésta es el único vehículo para que un emisor comunique un mensaje a un receptor en una situación determinada. El habla se constituiría, por tanto, en la expresión de las »soluciones a las tareas presentes permanentemente en la actividad comunicativa« (Seiler 1973: 11; traducido en Moreno Cabrera 1987: 115). De este modo, la variación reflejaría la multiplicidad de soluciones a que puede llegar la cognición humana ante un problema comunicativo dado (sinonimia en sentido fuerte; variación interlingüística); problema que, en cada lengua, se plantearía de modos diversos según las situaciones comunicativas (sinonimia en sentido débil; variación intralingüística).

En ambos casos se observa la influencia de la »paradoja teórica de la variación«. En un primer paso, se plantea la variación sintáctica como si se tratara de sinonimia en su sentido fuerte: existen varias estructuras superficiales que corresponden a una única estructura profunda o varias construcciones que suponen la solución que la cognición ofrece a un único problema comunicativo. En un segundo paso, se investiga la variación sintáctica para demostrar la inviabilidad de este análisis y postular la existencia de sinonimia en sentido débil: las construcciones o estructuras superficiales se encuentran en distribución (libre o complementaria) dependiendo de factores internos o externos a la lengua que interactúan de modo bien probabilístico, bien regular. El uso de una variante concreta respondería, pues, a una o varias motivaciones formales o funcionales de configuración distinta a la correspondiente a la variante alternativa. En cualquier caso, quedaría restablecida la relación biunívoca entre significativo (construcción o estructura superficial) y significado (posición estructural o función comunicativa).

La »paradoja teórica de la variación« ha conducido al paradigma generativista a desentenderse, en buena parte, del tratamiento teórico del tema hasta la introducción del acercamiento desarrollado por la Teoría de la Optimalidad (cf. p. ej. Prince/Smolensky 1993), una teoría que surgió a comienzos de los años noventa en el seno de la fonología y se extendió rápidamente a la sintaxis en diversas versiones del paradigma como la Gramática Léxico-Funcional (cf. p. ej. Bresnan 2001) o el Programa Minimalista (cf. p. ej. Müller 2000). Esta misma paradoja ha producido una profunda división en el paradigma funcional. Desde un punto de vista teórico, el tema también se ha resistido a una interpretación global hasta la aparición del acercamiento neurocognitivo de las motivaciones en competencia (ingl. *competing motivations*; cf. p. ej. DuBois 1985). Desde un punto de vista práctico, el tema pasó a ser dominio casi exclusivo de una pujante sociolingüística (cf. p. ej. 1966) que parte de la existencia de la variación como hecho indiscutible en el habla, sin preocuparse de su estatus dentro de una teoría de la lengua o de la sintaxis. Ve en ella la manifestación lingüística de múltiples factores extralingüísticos, entre los que destacan las condiciones e identidades sociales y culturales de determinados grupos de hablantes. La sociolingüística recurrió, por lo tanto, desde el primer momento, a la recogida, la organización, el análisis y la interpreta-

ción de gran cantidad de material empírico según la metodología propuesta por la naciente lingüística de corpus (cf. p. ej. Muller 1968). La simbiosis entre sociolingüística y lingüística de corpus desencadenó una avalancha de estudios variacionistas que, dependiendo del fenómeno estudiado, se basaban bien en los factores extralingüísticos defendidos por la primera, bien en los factores intralingüísticos descubiertos por la segunda a través de la aplicación informatizada de métodos estadísticos (cf. p. ej. Sankoff (ed.) 1978).

En la actualidad, la «paradoja teórica de la variación» se manifiesta en una creciente divergencia entre los estudios de carácter teórico procedentes de los dos paradigmas mencionados y los de carácter práctico surgidos de la simbiosis entre sociolingüística y lingüística de corpus. Los estudios teóricos siguen centrándose, al menos implícitamente, en la introspección de un hablante idealizado, sin que los avances teóricos y empíricos hayan encontrado un reflejo adecuado más que en excepciones (cf. p. ej. Bybee/Hopper (eds.) 2001). Los estudios prácticos continúan desbrozando parcelas de la lengua afectadas por la variación sin que tal actividad haya fructificado en modelos coherentes de variación compatibles con los paradigmas mencionados.

#### 4. Retos y respuestas

Este volumen se propone, por un lado, dar una muestra de los trabajos de hombres y mujeres de varias generaciones, procedentes de distintos países, trabajos que representan los distintos enfoques teóricos y metodológicos con los que se están investigando algunos de los problemas de variación sintáctica más candentes, y, por otro, contribuir a una reflexión plural sobre la relación dialéctica entre explicación teórica como hermenéutica de los datos y descripción empírica como banco de pruebas de la teoría.

En la primera parte del volumen, se presentan tres reflexiones metodológicas sobre el valor del uso de corpus en lingüística. El trabajo de *Guido Mensching* (Freie Universität Berlin) se sitúa en la perspectiva teórica del generativismo (Programa Minimalista) y constituye una apología de la lengua-I como objeto propio de la lingüística y de la introspección como método adecuado para su investigación. El corpus, como muestra representativa de la lengua-E, carece de peso metodológico en el generativismo. En el estudio de la situación actual de una lengua, su única función sería la de testimoniar la presencia de determinados conocimientos en las gramáticas de algunos de los hablantes de una comunidad. En el estudio de la evolución histórica de una lengua, esa función testimonial adquiriría obviamente mayor relieve debido a la falta de hablantes capaces de introspección. Finalmente, el estudio de la frecuencia de determinados fenómenos no aportaría datos ni sobre la importancia de un fenómeno dentro de la gramática ni sobre los cambios diacrónicos a los que se ve sometida la gramática, puesto que la frecuencia de uso de un fenómeno en el entorno de un niño no sería una indicación de un mejor o peor aprendizaje del mismo. La existencia de distintas variantes en la lengua-I sería un problema del lexicón; el motivo de la selección de una u otra en la lengua-E, un problema de la pragmática (intención comunicativa), influida por factores extralingüísticos. El generativismo se ocupa únicamente de justificar teóricamente el funcionamiento de cada una de las variantes dentro de una gramática idiolectal.

Muy otra es la posición de uno de los pioneros de la lingüística de corpus, *Josse De Kock* (Katholieke Universiteit Leuven). El autor escoge quince variables sintácticas y compara la frecuencia de uso de sus dos posibles variantes en el corpus de «Gramática Española: Enseñanza e Investigación» de literatura hispánica. De este modo, elabora una escala ordinal de la frecuencia de determinadas variantes y la compara con las que obtiene del corpus del «Proyecto de Estudio Coordinado de la Norma Lingüística Culta de las principales ciudades de Iberoamérica y de la Península Ibérica» (o, sencillamente corpus «Norma Culta»). El autor obtiene para una misma variante lingüística frecuencias y posiciones en la escala que difieren dependiendo de las variables extralingüísticas medio y lugar. El autor propone el uso de este método como criterio objetivo para discernir el grado de presencia de una variante en el sistema lingüístico, es decir, su «grado de gramaticalidad». En otras palabras, el juicio de gramaticalidad de un hablante dependerá siempre en mayor o menor medida del entorno lingüístico en el que haya aprendido la lengua de niño y la esté utilizando de adulto.

En una lúcida autocrítica, *Nicole Delbecque* (Katholieke Universiteit Leuven) recorre su propia trayectoria intelectual en el estudio de la posición del sujeto desde la lingüística cuantitativa probabilista hasta la lingüística cognoscitiva, dentro de la perspectiva teórica del funcionalismo. En la primera fase, se esboza una regla probabilista basada en la combinación de varias variables intralingüísticas (p. ej. longitud del constituyente, papel temático del argumento, función tópico del referente); la autora critica el método porque tiende a convertir el establecimiento de correlaciones en un fin en sí mismo. En la segunda fase, se elabora una explicación funcionalista apoyada en las nociones comunicativas que ofrece la lingüística del texto, entonces en boga (progresión temática, perspectiva funcional de la oración) que se completa con una teoría basada en la semántica del verbo (proceso, acción) y de su relación con el sujeto (agentividad). Ante la insuficiencia del planteamiento, se da un paso más y se intenta explotar el modelo de la lingüística cognoscitiva: mientras que en la anteposición se conceptualiza el sujeto como punto de partida de un «flujo de energía», en la posposición queda construido como su punto de llegada.

El resto del volumen presenta algunos avances en temas controvertidos dentro del campo de la variación sintáctica en español, desde una multitud de perspectivas teóricas y metodológicas. La segunda parte continúa la temática del estudio de Nicole Delbecque y se concentra sobre el «orden de palabras» y la estructura informativa. *Alicia González de Sarralde* (Instituto «Cervantes», Madrid) se concentra también en la variable posición del sujeto. Tras mostrar las dificultades de las explicaciones centradas en el nivel de la oración o del enunciado, toma partido por un modelo funcional cognitivista enraizado en la lingüística de texto: la teoría de la «*quaestio*». El método seguido también es peculiar: La autora selecciona veintinueve probandos madrileños, les muestra una misma película de vídeo de diez minutos de duración y les propone la tarea de relatar con sus propias palabras el argumento de la película. Con ello, la autora logra un corpus de relatos referidos al mismo estado de cosas, que puede analizar onomasiológica y semasiológicamente. El análisis onomasiológico revela las distintas estrategias de expresión de un estado de cosas cuando permanecen constantes el referente del sujeto y el grado de agentividad expresado por el verbo. El análisis semasiológico revela los distintos usos y funciones que se da a los ordenamientos «sujeto-verbo» y «verbo-sujeto». De este modo, la autora define cuatro funciones centrales de la posposición del sujeto: *flash back*, reanudación de una cadena de sucesos interrumpi-

da, especificación de referentes introducidos previamente por medio de sucesos conocidos y jerarquización de sucesos.

Valeriano Bellosta von Colbe (Universität zu Köln) explora las posibilidades de la »Gramática del Papel y de la Referencia« (ingl. *Role and Reference Grammar*) a la hora de explicar la posición de los complementos directos e indirectos en las oraciones ditransitivas, tomando como corpus el *Archivo de Textos Hispánicos de la Universidad de Santiago de Compostela* analizado por medio de la *Base de Datos Sintácticos del Español Actual*. El autor pasa revista a distintos factores semánticos (control, afectación), sintácticos (»peso sintáctico«) y pragmáticos (focalización) y esboza un tratamiento inspirado en la Teoría de la Optimalidad en el que el orden depende del resultado de una competición de factores que se hacen presentes en el momento mismo de la comunicación.

En un sugerente trabajo, Pedro Martín Butragueño (El Colegio de México), tomando como base teórica la Fonología Entonacional Generativa, explora la relación entre estructura prosódica, informativa y sintáctica en español a partir de grabaciones realizadas a cuatro probandos mexicanos y consigue llegar a una revisión de los postulados expuestos por Zubizarreta (1999) en la *Gramática descriptiva de la lengua española*. Por un lado, la variable prosódica foco (definida como una »prominencia tonal« especial) dispone de al menos tres variantes, a saber, »o marcar como H\* [sc. alto] un acento nuclear del que se esperaba que fuera L\* [sc. bajo], o hacer otro tanto con un acento no nuclear, o bien subdividir el grupo melódico introduciendo nuevos tonos de juntura«. Por otro lado, la variable informativa foco (definida como »la parte no presupuesta de la oración, es decir, la parte de información no compartida por hablante y oyente«) es un fenómeno que precisa de una clasificación previa. En principio, se distingue entre foco neutro y contrastivo. El foco neutro puede ser simple o múltiple y, en el primero de los casos, tiene distintos ámbitos sintáctico-informativos de aplicación (oración-enunciado, sintagma verbal-comentario, sintagma nominal). En teoría, »el foco neutro se identifica por medio del acento nuclear neutro, que es el tono que recae sobre la última sílaba tónica del grupo melódico, mientras que el foco contrastivo puede identificarse por medio de un tono enfático, susceptible de colocarse sobre cualquier sílaba acentuable.« Ni en uno ni en otro caso pueden hacerse atribuciones seguras. La variable informativa foco en sus distintas variantes no sólo se expresa por medio de la variante-variable prosódica foco sino también a través de la variante-variable sintáctica del orden de constituyentes. El autor examina la realización prosódica de las variantes posicionales del grupo formado por sujeto, verbo y complemento, sin olvidar los casos de dislocación a la izquierda y a la derecha, prestando especial atención a las posibles diferencias entonatorias entre tema y foco antepuestos.

En la tercera parte del volumen, se tocan los distintos problemas que ofrece la codificación variable de las »funciones sintácticas« de la oración (sujeto, complementos directo, indirecto y oblicuo, atributo). Con respecto al sujeto, Amparo Morales (Universidad de Puerto Rico, Río Piedras) nos presenta el parámetro del sujeto nulo, que según el paradigma generativista, es uno de los que manifiestan más claramente la variación interlingüística que da lugar a diversas gramáticas. Mientras que la mayoría de los dialectos del español se consideran de sujeto nulo, el español del Caribe muestra, en ocasiones, características típicas de una lengua de sujeto obligatorio (p. ej. uso de sujetos expletivos, anteposición del sujeto en oraciones interrogativas). El estudio de la adquisición del pretendido parámetro en niños puertorriqueños revela la importancia de los factores pragmáticos en la distinción de las personas. Por un lado, la adquisición de recursos sintácticos de referencia intratextual

(anáfora) es más lenta que la de los de referencia extralingüística (deíxis), por lo que el niño, en sus narraciones, debe ayudarse de la repetición de sujetos de tercera persona. Por otro lado, el estado cognitivo del niño es de un fuerte egocentrismo, el cual se manifiesta en la repetición de sujetos de primera y segunda persona como estrategia de autotematización contrastiva en el resto de las interacciones (cf. *infra*, Detges en este volumen).

*Ulrich Detges* (Ludwig-Maximilians-Universität München) considera los motivos históricos a los que se debe la gramaticalización del acusativo preposicional, es decir, de la variación en el marcado del complemento directo ( $\emptyset/a$ ). Para ello, el autor parte del uso focalizador de las preposiciones léxicas, típico del latín, y de la generalización del uso de *ad* a esquemas bivalentes y/o trivalentes (cf. *infra*, Roegiest en este volumen). Gracias a esta premisa, puede constatar que la preposición *a*, sucesora de las latinas, continúa tal uso en iberorromance medieval tanto en contextos remáticos (en posición final o inicial) como en contextos temáticos (contrastivos o no) y postular un cambio semántico-pragmático por el que la preposición pierde primero su »peso focalizador« y, después, su función temática. El autor ve el motivo para el uso focalizador en la función pragmática discursiva de autotematización contrastiva que los hablantes utilizan sobre todo para conquistar o defender su turno en la conversación. Las construcciones que realizan tal función se caracterizan por la anteposición de gran cantidad de material autoreferente (pronombres personales tónicos de primera persona con focalizador acompañados de verbos de opinión o sentimiento). Prolongando su análisis, el autor consigue mostrar el motivo común subyacente a tres fenómenos que aparecen así interrelacionados: la anteposición de complementos, la referencia anáforica a éstos por medio de un clítico pretendidamente redundante y el acusativo preposicional propiamente dicho.

Los dos estudios siguientes están consagrados a las formas pronominales asociadas tradicionalmente a la »función« de complemento indirecto (*le/lles*). En efecto, la polisemia de estas formas (*le/lles* = complemento directo, marco contextual) genera, a primera vista, sinonimia (variación) en las »funciones« (p. ej. complemento directo = {*lo/la/los/las*}, {*le/lles*}). *Eugeen Roegiest* (Universiteit Gent) dedica su estudio al ya clásico tema del leísmo, la variación entre los pronombres personales acusativo y dativo con referentes humanos en la »función« de complemento directo. El autor realiza un análisis cuantitativo y cualitativo basándose en un corpus literario de narradores españoles contemporáneos. Desde un punto de vista cuantitativo, los dialectos estudiados se pueden ordenar formando un continuo en cuyos extremos hallamos el leísmo marcado y el leísmo atenuado en las formas del masculino. Desde un punto de vista cualitativo, el autor examina algunos entornos sintáctico-semánticos que podrían favorecer el leísmo atenuado y el leísmo en las formas del femenino. El primer entorno examinado, los predicados y las construcciones que alternan entre un esquema sintáctico bivalente (con complemento directo) y uno trivalente, podría favorecer el leísmo en la realización de los esquemas bivalentes bajo el influjo del otro esquema posible (predicados de comunicación y manipulación, construcciones factitivas y de percepción directa con infinitivo). Esto le conduce a subrayar la característica común a todos los predicados y construcciones de este entorno: el alto grado de agentividad y subjetividad del referente del elemento pronominalizado que sería responsable de su interpretación como beneficiario, papel temático situado entre el de agente y el de paciente en las jerarquías implicativas temáticas. Esta explicación también es válida para el segundo entorno estudiado, los predicados que alternan entre dos esquemas sintácticos bivalentes uno con complemento directo y otro con complemento indirecto (verbos de sentimiento y expresión

de sucesos espontáneos). Según el autor, «marcamos el experimentante como acusativo cuando en una estructura biactancial está dominado por un sujeto interpretado como intencional y controlador, lo marcamos como dativo cuando domina por sus rasgos inherentes y su función temática reactiva al estímulo inactivo objeto del suceso emotivo y no interpretado como intencional».

El trabajo de *Rena Torres Cacoullos* (University of New Mexico, Albuquerque) se puede considerar un estudio de la variación en la expresión de «funciones» no argumentales. La autora muestra cómo *lelles* ha ido adquiriendo paulatinamente nuevas funciones en el español de México y de Nuevo México. En el marco de la Gramática Cognoscitiva y desde la perspectiva de la Teoría de la Gramaticalización, la autora postula que las construcciones en las que aparecen las formas estudiadas han sufrido un doble proceso de pérdida de transitividad (destransitivización) y de democión a marco contextual (con pérdida de referencialidad) que contrasta con los procesos opuestos de incremento de transitividad (en «construcciones de explotación total») y de elevación a participante (en construcciones de dativo ético y de dativo de afectación). La autora analiza e interpreta dos fenómenos en apoyo de su hipótesis. Para ella, el primer proceso se ve reflejado por la desaparición del léismo entre los siglos XVI y XIX. El segundo proceso se puede diagnosticar a partir de la proliferación de la copresencia de *le* y objeto indirecto nominal unida a la paradójica pérdida de la concordancia entre ambos y al aumento de la coreferencia con sintagmas preposicionales sin marca de objeto indirecto que supondrían la pérdida de los valores referencial y argumental de *le* respectivamente.

Con el trabajo de *Irania Malaver* (Universidad Central de Venezuela, Caracas) se concentra la atención en las oraciones copulativas y en la «función sintáctica» de atributo. Como es bien sabido, en este caso se da variación entre las cópulas *ser* y *estar* que puede afectar al significado del atributo. En el caso estudiado, los atributos adjetivales referentes al concepto de edad, tal alternancia parece obedecer a una oposición gramaticalizada en español entre clasificación y caracterización (cf. ingl. *stage level* vs. *individual level predicates*). La autora, inspirándose en estudios recientes, subraya el papel del discurso en la selección de la cópula en el habla de Caracas: mientras que las oraciones copulativas colocadas en primer plano discursivo hacen uso de tal oposición, las colocadas en el trasfondo la neutralizan. La comparación con el habla de Sevilla correspondiente al proyecto «Norma Culta» revela que el habla de Sevilla utiliza únicamente *ser* en los contextos de trasfondo, mientras que el de Caracas emplea indistintamente *ser* y *estar*. La autora concluye que, «[d]ada la presencia de la distinción entre primer plano y trasfondo en los dos dialectos, parece claro que el dialecto sevillano la mantiene opaca al usar en ambos casos *ser*, mientras que el dialecto caraqueño la hace emerger al oponer en el primer plano E[xpresiones de] E[dad] tipológicas [*scil.* clasificación] y cronológicas [*scil.* caracterización] y neutralizar esta misma oposición en el trasfondo».

El volumen concluye con un estudio variacionista clásico concerniente a la hipotaxis. *Dexy Galué* (Universidad Central de Venezuela, Caracas) estudia la variación entre presencia y ausencia de la preposición *de* ante proposiciones subordinadas sustantivas cuando el verbo de la proposición principal la selecciona y rige de acuerdo con la norma (queísmo). Para ello se vale de una muestra procedente del corpus *Estudio sociolingüístico del habla de Caracas 1987* que incluía hablantes de ambos sexos distribuidos en tres niveles socioeconómicos y cuatro grupos etarios. La autora somete las variables de tipo de referente del sujeto de la proposición principal, de interferencia de elementos lingüísticos entre verbo de

la proposición principal y frontera de la subordinada, de presencia de segmentos fónicos similares a la preposición y de contexto estructural a un análisis probabilístico con el programa *GoldVarb2.0*. Puesto que únicamente las dos últimas variables citadas junto con el nivel socioeconómico resultan ser estadísticamente significativas, la autora concentra su atención en la explicación del contexto estructural más influyente: los verbos pronominales. Al mismo tiempo que subraya la presión analógica ejercida por verbos de significado similar (*acordarse (de/Ø) que* vs. *recordarse (de/Ø) que* vs. *recordar que*), propone la interesante hipótesis de que la pérdida de la preposición conduciría al reanálisis de la función sintáctica de la conjunción subordinante *que*: »la preposición precede la cláusula y delimita la frontera oracional, [...] la omisión de *de* hace que el subordinante *que* se separe de la cláusula y se adjunte al verbo pronominal ocupando el lugar de la preposición; con ello se redefine la organización sintagmática de los constituyentes«. No cabe duda de que las implicaciones de este análisis pueden ser de gran envergadura y precisan de un estudio más detallado.

La aportación original más interesante de los estudios aquí reunidos estriba en el redescubrimiento y en la nueva valoración del papel del discurso y de la frecuencia de uso en el origen, extensión, gramaticalización y desambiguación de construcciones alternativas o variantes más o menos marcadas.

La mayor parte de las contribuciones aquí reunidas constituyen versiones corregidas y actualizadas de algunas de las conferencias que fueron presentadas y discutidas en el marco de la sección 8 del XIII Congreso de la Asociación Alemana de Hispanistas (Universität Leipzig, 08.–11.03.2001). Agradecemos al profesor Alfonso de Toro y al Centro de Investigaciones sobre Iberoamérica de la Universidad de Leipzig su apoyo en todo lo referente a la organización, a la Asociación Alemana de Hispanistas, a la Sociedad Alemana para el Fomento de la Investigación (*Deutsche Forschungsgemeinschaft*) y a la Embajada de España en Alemania el apoyo económico prestado a algunos de los conferenciantes. Tenemos una deuda de gratitud con los profesores Alberto Gil (Saarbrücken), Daniel Jacob (Colonia), Franz-Josef Klein (Siegen), Ingrid Neumann-Holzschuh (Ratisbona) y Beatrice Primus (Colonia) por haber revisado los manuscritos presentados. Una deuda pareja hemos contraído con el profesor Peter Blumenthal (Colonia), que, tras revisar el volumen entero, ha apoyado su inclusión en la colección *Linguistische Arbeiten*, de cuya sección de lingüística románica es responsable.

## Bibliografía

- Bresnan, Joan (2001): *Lexical-Functional Syntax*. – Malden, Oxford: Blackwell (=Blackwell Textbooks in Linguistics).
- Bybee, Joan / Paul Hopper (2001): *Frequency and the Emergence of Linguistic Structure*. – Amsterdam, Philadelphia: Benjamins (=Typological Studies in Language 45).
- Chomsky, Noam A. (1965): *Aspects of the Theory of Syntax*. – Cambridge, MS: The MIT Press.
- (1986): *Knowledge of Language: Its Nature, Origin and Use*. – New York, Westport, London: Praeger (=Convergence).

- Coseriu, Eugenio (1952): Sistema, norma y habla. – En: *Revista de la Facultad de Humanidades y Ciencias* (Montevideo) 9, 113–181.
- DuBois, John (1985): Competing Motivations. – En: J. Haiman (ed.): *Iconicity in Syntax: Proceedings of a Symposium on Iconicity in Syntax, Stanford, June 24–26, 1983*, 343–365. Amsterdam, Philadelphia: Benjamins (=Typological Studies in Language 6).
- Hjelmslev, Louis (1943): *Omkring sprogteoriens grundlaeggelse*. – København: Munksgaard.
- Jespersen, Otto (1924): *The Philosophy of Grammar*. – London: Allen & Unwin.
- Labov, William (1966): *The Social Stratification of English in New York City*. – Washington: Center for Applied Studies.
- Martín Butragueño, Pedro (1999): ¿Es funcional la variación sintáctica? – En: Serrano (ed.) 1999: 221–235.
- Martinet, André (1960): *Éléments de linguistique générale*. – Paris: Colin (=Collection Armand Colin 349).
- Moreno Cabrera, Juan Carlos (1987): *Fundamentos de sintaxis general*. – Madrid: Síntesis (=Textos de apoyo: Lingüística 4).
- Muller, Charles (1968): *Initiation à la statistique linguistique*. – Paris: Larousse (=Langue et langage).
- Müller, Gereon (2000): *Elemente der optimalitätstheoretischen Syntax*. – Tübingen: Stauffenburg (=Stauffenburg Linguistik 20).
- Prince, Alan/Paul Smolensky (1993): *Optimality Theory: Constraint Interaction in Generative Grammar*. – New Brunswick, NJ: Rutgers University (=Rutgers Center for Cognitive Science: Technical Report 2).
- Sankoff, David (ed.) (1978): *Linguistic Variation: Models and Methods*. – New York: Academic Press.
- Saussure, Ferdinand de (1916;1998): *Cours de linguistique générale*. Charles Bailly/Albert Sechehaye/Albert Riedlinger/Tullio de Mauro (eds.) – Paris: Payot (=Grande bibliothèque Payot).
- Seiler, Hansjakob (1973): Das Universalienkonzept. – En: H. Seiler (ed.): *Linguistic Workshop*. Vol. 1: Vorarbeiten zu einem Universalienprojekt, 6–19. München: Fink (=Structura 4).
- Serrano, María José (1999): Nuevas perspectivas en variación sintáctica. – En: Serrano (ed.) 1999: 11–49.
- (ed.) (1999): *Estudios de variación sintáctica*. – Madrid, Frankfurt a. M.: Iberoamericana, Vervuert.
- Zubizarreta, María Luisa (1999): Las funciones informativas: tema y foco. – En: I. Bosque, V. Demonte (eds.): *Gramática descriptiva de la lengua española*, Vol. 3: Entre la oración y el discurso. Morfología, 4215–4244. Madrid: Espasa–Calpe (=Colección »Nebrija y Bello«).



Guido Mensching (*Freie Universität Berlin*)

## Variación sintáctica, lingüística de corpus y gramática generativa: Teorías, métodos y problemas

### 1. Introducción

En los preparativos del presente libro<sup>1</sup>, se planteaba la siguiente cuestión: A pesar de que exista un número cada vez mayor de corpus y de estudios basados en ellos, los resultados de tales estudios no han atraído la atención esperable en aquellos investigadores dedicados al desarrollo de teorías sintácticas. Aquí me voy a centrar en este problema desde el punto de vista del programa minimalista, el modelo más reciente de la gramática generativa (Chomsky 1992 ss.). Para la discusión anterior sobre la variación sintáctica y su relación con la gramática generativa, véase, entre otros, el artículo de Godard (1992), aún basado en el modelo chomskyano anterior (»régimen y ligamiento«). Algunos aspectos a que Godard alude se tratan más detenidamente aquí y otros aparecen bajo una luz diferente en el programa minimalista. Además, Godard no tiene en cuenta explícitamente la cuestión de los corpus.

Este artículo<sup>2</sup> se divide en los apartados siguientes: En el apartado 2 se van a introducir algunas nociones básicas del programa minimalista. El tema del apartado 3 son los problemas de corpus en cuanto al concepto de lengua interna (*lengua-I* como opuesto al de la lengua externa, *lengua-E*), mientras que el apartado 4 trata de los efectos del problema de la actuación frente a la competencia. En el apartado 5 se examinarán, desde el punto de vista minimalista, algunos aspectos sobre el problema de la »regla variable«, que juega un papel importante en los estudios de corpus de índole sociolingüística. Por, último, bajo 6, se resumen los resultados y se intentará contestar a la pregunta siguiente: ¿Qué es lo que los corpus o la lingüística de corpus pueden aportar a la gramática generativa?

---

<sup>1</sup> Me refiero sobre todo a la sección correspondiente en el XIII Congreso de la Asociación Alemana de Hispanistas en Leipzig; véase la introducción de este libro.

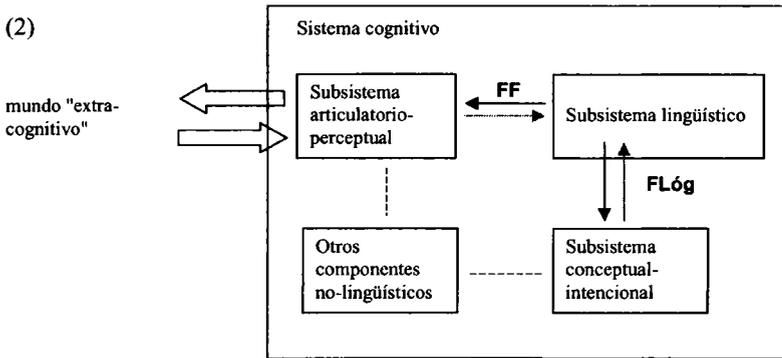
<sup>2</sup> Les agradezco a Natascha Pomino y a Eva Remberger la lectura previa de este artículo y sus consejos, y a María Jesús Beltrán la corrección.



## 2. El programa minimalista

Para plantear el problema de la variación lingüística y de la lingüística de corpus en la gramática generativa, vamos a partir de algunos conceptos básicos del programa minimalista, que se pueden ilustrar partiendo del esquema siguiente:

(2)



El sistema lingüístico interactúa con los dos componentes *articulatorio-perceptual* y *conceptual-intencional*, que no son lingüísticos. Las flechas son los »interfaces«, es decir, aquellos componentes que entregan la información lingüística a los módulos no lingüísticos, y viceversa, la reciben de estos módulos, dependiendo de si hablamos del análisis o de la producción de enunciados. Se llaman Forma Lógica (FLóg)<sup>3</sup> y Forma Fónica (FF) respectivamente.<sup>4</sup> En una manera muy simplificada, la producción de una frase se efectúa en los pasos ejemplificados en el esquema (1) a la izquierda.

Según la hipótesis nativista, el sistema lingüístico contiene muy poca información cuando el niño nace.<sup>5</sup> Esta información, llamada *gramática universal* (GU), consiste en propiedades básicas que caracterizan las lenguas naturales, como lo son por ejemplo:

- mecanismos y reglas básicas sobre la construcción de un léxico mental
- mecanismos que dirigen el flujo de la información entre los diferentes componentes tanto lingüísticos y no lingüísticos

<sup>3</sup> Sigo la convención introducida en Chomsky (1999c) de usar la abreviatura *Flóg*, ya que FL se usa para la facultad del lenguaje.

<sup>4</sup> Los procesos que llevan a la Forma Fónica son, según Chomsky (1995), muy diferentes de aquellos que llevan a la estructura sintáctica »visible« (derivación) e »invisible« (Flóg). Este hecho es relevante aquí. Según Serrano (1994) sería »absurdo creer que tras la acertada aportación laboviana al estudio de lenguaje en correlación con la sociedad, el método sólo tuviese utilidad en el terreno de la fonología. Aún más absurdo sería creer que a otros niveles distintos de ese los cambios no estuviesen implicados con fenómenos sociales, puesto que esto contradiría la unidad del lenguaje«. Según Chomsky, como acabamos de ver, no hay tal unidad.

<sup>5</sup> Sobre algunos resultados recientes de la genética en cuanto al hallazgo de un posible »gen lingüístico« (el gen FOXP2) véanse Pinker (2001) y Lai e.a. (2001).

– un mecanismo para construir estructuras sintácticas binarias

Según la hipótesis minimalista, gran parte de los aspectos restantes que tocan el funcionamiento de la sintaxis, incluyendo, por ej., el desplazamiento (*movement*) de constituyentes, siguen de propiedades que son, quizás, más generales aún, como las condiciones de economía. Cuando el niño aprende el idioma materno, la gramática universal se va enriqueciendo con informaciones específicas, de forma que el niño reconstruye mentalmente un lexicón y una gramática basados en el idioma en cuestión.<sup>6</sup> Esta lengua mental es la lengua interna (lengua-I) de esta persona. La lengua-I de cada hablante es diferente y además cambia constantemente.<sup>7</sup> Hablando de la variación lingüística esto explica naturalmente la existencia de los idiolectos. Al contrario, la lengua de un grupo de hablantes o de una comunidad lingüística son lenguas externas (lenguas-E). El campo de estudio de la gramática generativa son exclusivamente las lenguas-I. El concepto de lengua-I frente a la lengua-E fue introducido por Chomsky (1986), aunque pertenecía ya a la gramática generativa desde sus principios de forma implícita. Naturalmente, sigue existiendo también la distinción entre competencia y actuación. Estas dicotomías (lengua-I – lengua-E / competencia – actuación) se tratan en los apartados siguientes.

### 3. Los corpus y el problema de la lengua-E

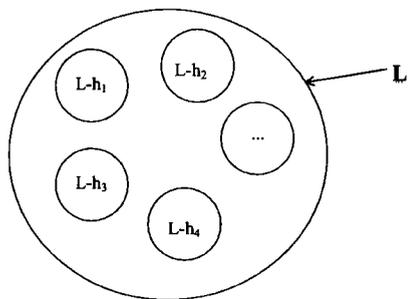
Si tenemos varios hablantes de un idioma, llamémoslos  $h_1, h_2, h_3, h_4, h_5, \dots$ , podríamos decir algo así: sus lenguas tienen algo (por lo general mucho) en común. El conjunto de las reglas y palabras de ellos forman la variedad X, por ejemplo, la lengua de la familia F ( $L_F$ ), la lengua de la zona geográfica G ( $L_G$ ), de un grupo social S ( $L_S$ ) o, incluso, si se trata de un país P,  $L_P$ . Se podría llegar a la descripción de estas lenguas o variedades formando la adición de las características lingüísticas de los hablantes (3a) o bien la intersección entre ellas (3b) (ver página siguiente):

El problema es que ninguna de estas lenguas existe en la realidad: no están almacenadas en ninguna parte; son construcciones hechas por algunos lingüistas, gramáticos o lexicógrafos; y de ninguna forma tienen justificación cognitiva: no se encuentran dentro de ningún hablante, son exteriores y por lo tanto Chomsky las llama lenguas-E. Las lenguas-E no forman el objeto de estudio del lingüista generativo, que se concentra exclusivamente en las lenguas-I.

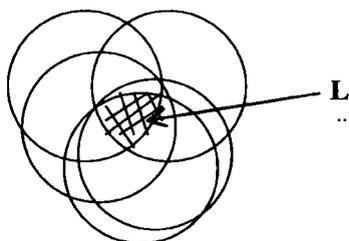
<sup>6</sup> La facultad del lenguaje (FL) es el producto de tres componentes (véase Chomsky 2001): (i) características específicas de la »lengua« con que el niño nace (= lengua en su estado inicial,  $E_0$ ), (ii) características determinadas por la necesidad de comunicación con los interfaces, (iii) características determinadas por el funcionamiento de sistemas orgánicos en general.

<sup>7</sup> »FL tiene un estado inicial  $E_0$ , expresión de los genes. [...] FL experimenta cambios de estado mediante disparadores e influencias de formación por parte del entorno. Si la FL de Jones está en el estado L, decimos que Jones posee (habla, sabe, ...) la lengua(-I) L. [...] Entendemos la gramática universal (GU) como la teoría del estado inicial y las gramáticas particulares como teorías de los estados obtenidos.« (Chomsky 1998; trad. en Chomsky 1999c).

(3) a.



b.



Es interesante observar que el concepto de lengua-E tiene semejanzas con la *langue* de Saussure, que es una entidad colectiva y no individual; véase Serrano (1994). Godard (1992) deja abierta la relación entre la lengua de una comunidad lingüística y la de un individuo. En una breve nota (1992: nota 6) se limita a citar a Bickerton (1973), para quien la competencia de un locutor es un subconjunto del sistema de reglas de la comunidad, y la visión contraria defendida por Trudgill (1983), quien muestra que un locutor individual no tiene conocimiento ni activo ni pasivo del conjunto de estas reglas. Este punto de vista concuerda aún con el modelo generativista más reciente y se acerca bastante a los conceptos de lengua-I y lengua-E respectivamente.

La lengua-I es, por lo tanto, diferente en cada hablante. El porqué de este hecho es muy fácil: una persona tiene una lengua-I particular porque la aprendió así, en gran parte en los primeros años de su vida, es decir porque el *input* lingüístico proporcionado por los otros individuos tiene una forma determinada. Cada niño se enfrenta con personas diferentes, y por tanto también con lenguas ligeramente diferentes. Todo ello explica, naturalmente, que exista la variación lingüística, y también explica el que ciertos grupos, por ejemplo sociales, tengan rasgos comunes. Pero la investigación de estos rasgos, y de cómo se determinan por razones sociales, políticas etc. es, en todo caso, la tarea de la sociolingüística, pero no es la tarea de la gramática generativa, pues en realidad son cuestiones que tocan la lengua externa. Los investigadores en el campo de la variación sintáctica critican de vez en cuando la gramática generativa por no prestar demasiada atención a estos aspectos. Esta crítica no es lícita, porque está exigiendo a una rama de la investigación algo que declaradamente no pertenece a su campo. De una forma algo exagerada, se podría decir que sería como si se reprochara a un oculista el hecho de no poder explicar por qué la gente interpreta el color verde como esperanza en algunas culturas y en otras como otra cosa.

Un corpus, tal y como se entiende por regla general en la lingüística de corpus y en los estudios sobre la variación sintáctica, contiene enunciados de varios hablantes, hecho debido a razones estadísticas y al objetivo de tales estudios, que consiste precisamente en detectar y examinar la variación lingüística en varios hablantes y no en uno solo. Está claro que las regularidades sintácticas que se pueden extraer a partir de un tal corpus reflejan una lengua-E. También las evaluaciones estadísticas, siempre y cuando se basen sobre más de un sólo hablante llevan a enunciados sobre la lengua-E. Vamos a ver un ejemplo ficticio:

(4) Un corpus determinado muestra la construcción 1 en un 15% de los casos, y la construcción 2 en un 85%.

Este enunciado, en sentido estricto, no es válido para la gramática generativa. Hay que tener en cuenta en primer lugar la pregunta siguiente: ¿Todas las personas que aparecen en el corpus usan ambas construcciones o no? Si es el caso, el primer resultado para la gramática generativa es el hecho de que ambas construcciones existen en las lenguas-I de estos hablantes, conclusión que a primera vista parece trivial porque es evidente. La gramática generativa tiene que explicar, entonces, los mecanismos puramente sintácticos que permiten a estos hablantes generar ambas construcciones. El porcentaje bajo de la construcción 1 se atribuirá a factores ajenos al sistema lingüístico, radicados sobre todo en el sistema conceptual-intencional, y por lo tanto se deben investigar en estudios de pragmática, dialectología, sociolingüística u otros. Volvemos a este problema más adelante.

Pongamos entonces, que el corpus muestra que todos los hablantes entrevistados usan ambas construcciones. El problema siguiente que se le plantea al investigador generativista es que la investigación no se puede parar allí, entre otras cosas porque, aunque las dos construcciones sean aparentemente las mismas en todos los hablantes, las razones sintácticas y sus estructuras no son necesariamente las mismas. Sigamos con nuestro ejemplo ficticio, siempre de forma muy simplificada, y pongamos que las dos construcciones en cuestión son las siguientes:

- (5) a. construcción 1: oración con verbo finito y sujeto expresado  
b. construcción 2: oración con verbo finito y sujeto no expresado

En la gramática generativa, el sujeto no-expresado es un pronombre fonológicamente vacío pero, sin embargo, existente en la estructura sintáctica (el elemento *pro*). Los principios universales del lenguaje predicen que si una lengua dispone de *pro* para expresar el sujeto, entonces esta lengua permitirá también la inversión de sujeto y predicado.<sup>8</sup>

- (5) c. construcción 3: inversión (tipo *Llegó Juan*)

Al detectar que la construcción 2 está presente en la lengua de todos los hablantes examinados, habría que comprobar, entonces, si la construcción 3 está presente también, cosa que sería de esperar. Pero en este momento puede haber varias soluciones de las cuales dependerá la estrategia por seguir:

- Caso 1: todos los hablantes usan también la construcción 3. Esto confirmaría que el sujeto inexpressado es el elemento *pro*.
- Caso 2: hay una parte de los hablantes que no usan en absoluto la construcción 3.<sup>9</sup> Esto podría significar varias cosas:
  - hay otras propiedades lingüísticas en las lenguas-I de los hablantes de este grupo que impiden la inversión de verbo y sujeto, independientes de la existencia o no-existencia de *pro*.
  - el sujeto inexpressado de este grupo de personas no corresponde a *pro*.

<sup>8</sup> Esta es la consecuencia del así llamado parámetro del sujeto nulo o »*pro-drop*«; véase, por ej. Haegeman (1994:454–457), Müller y Riemer (1998:158–164); para un punto de vista crítico cf. Bondre-Beil (1994). En el minimalismo, este principio tiene probablemente una explicación mucho más fácil, quizás sin tener que recurrir a un parámetro.

<sup>9</sup> Téngase siempre en cuenta que estamos hablando de una lengua ficticia y no de una variedad del español. Como observó uno de los colegas presentes en mi ponencia en Leipzig, esta situación no sería posible en ninguna variedad del español, hecho muy probablemente cierto.