

GRAMÁTICA DE ESTRUCTURA SINTAGMÁTICA GENERALIZADA (GESG)*

Celso Vargas

ABSTRACT

In 1985 Uszkoreit and Peters published some mathematical results that deal with the weak generative power of linguistic frameworks which incorporate metarules, mainly, GESG. In this paper we describe, informally, these results. In the first section, we point out some of the theoretical motivations that articulated GESG. In the second, we describe, briefly, GESG. In the third, we discuss two investigations that matter the claim that natural languages are context-free. In the last, we describe the results of Uszkoreit and Peters (1985).

El modelo de Chomsky presentado en *ASPECTS OF THE THEORY OF SYNTAX* (1965) consta de varios componentes y subcomponentes. Entre sus componentes tenemos el componente base y el componente transformacional. La base matemática del componente base es bastante simple y precisa ya que incorpora al modelo una gramática libre de contexto. Sin embargo, las características del componente transformacional no eran igual de precisas. No obstante, cuando Peters y Ritchie (1969-1973) formalizaron la teoría de transformaciones siguiendo los lineamientos generales de *ASPECTS*, es decir, que incluía reglas de movimiento, reglas de sustitución y reglas de elisión, se mostró que una gramática que incluyera transformaciones de este tipo generaba una clase de lenguajes que es equivalente a la clase de lenguajes generados por los sistemas restrictivos ilimitados (SRI), o lo que es equivalente, a la clase de máquinas de Turing. Como señala Wasow, "ellos (Peters y Ritchie) probaron que las gramáticas de transformaciones estándar tenían el poder de las máquinas de Turing, esto es, pueden ser utilizados para formalizar cualquier procedimiento que, en principio, sea formalizable" (Wasow 1985: 198). Este resultado contrastaba fuertemente con lo esperado por Chomsky. En efecto, de acuerdo con Chomsky, una gramática para una lengua natural debe ser un procedi-

miento que nos permita modelar la habilidad de los hablantes de esa lengua para correlacionar significados y sonidos (véase Wasow 1985:198). Sin embargo, dadas las características encontradas a las gramáticas transformacionales no es posible hacer tal modelación. Como señala Gazdar:

En un reciente artículo, Levelt hace la siguiente observación: "si es el formalismo de *ASPECTS* el que constituye el conocimiento innato del niño sobre su lengua materna, entonces, dados los resultados de Peters y Ritchie (1973) y los de Gold (1967) sobre la definición técnica de aprendizaje, las lenguas naturales no son aprendibles dado que la clase es recursivamente enumerable" (Gazdar 1982: 132).

Esto significa que el modelo de transformaciones plantea problemas de decidibilidad. Es decir, dada una secuencia de símbolos no es posible saber si esa secuencia no pertenece a la gramática en cuestión. Estos resultados exigían que se buscaran los medios para restringir el poder generativo de la gramática. Al final de la década de los 70 y como resultado, en parte, de la prueba de Peters y Ritchie, encontramos un componente transformacional con una variedad de mecanismos que lo hacían prácticamente inmanejable. En efecto, se incorpora una teoría de filtros, reglas que introducen símbolos complejos, reglas de inserción léxica, reglas de redundancia léxica, convenciones de ligamiento, etc. Los intentos por superar estos inconvenientes culminan, en el caso de Chomsky, en el modelo *GOVERNMENT AND BINDING* (1981) en el que el componente transformacional se ve drásticamente reducido a una única transformación a saber: mover-alfa. En este nuevo modelo,

* Quiero agradecer al profesor Stanley Peters por el reporte "On Some formal Properties of Metarules" con base en el cual preparamos este artículo. También al profesor Guillermo Coronado por la lectura de un borrador anterior. Al profesor don Jack Wilson, profesor y maestro, por sus observaciones y comentarios.

como señala Wasow (1985) no se pone énfasis en la modelación de la habilidad de los hablantes, sino que, de acuerdo con Chomsky, las gramáticas deben ser modelos que nos permitan explicar cómo es que los hablantes de las lenguas naturales aprenden su lengua. En tanto que el modelo de 1965 ponía énfasis en la ADECUACION DESCRIPTIVA, el modelo de 1981 pone énfasis en la ADECUACION EXPLICATIVA.

Pero a su vez y como resultado de la discusión de la teoría generativo-transformacional surgen nuevos enfoques de la sintaxis. Principalmente dos: la GRAMATICA DE ESTRUCTURA SINTAGMATICA GENERALIZADA (GESG) desarrollada por Gazdar y Pullum a partir de 1981 y la GRAMATICA LEXICO-FUNCIONAL (GLF) desarrollada por Ronald Kaplan y Joan Bresnan alrededor de la misma fecha. (Para una exposición de estas teorías véase Sells 1985). Aquí nos ocuparemos exclusivamente del modelo GESG.

Ahora bien, existen diferencias importantes en el modelo transformacional y el modelo GESG. La primera de ellas es que no utiliza transformaciones, sino que todos los fenómenos son tratados sin necesidad de postular la existencia de un nivel profundo. En este sentido se dice que el modelo está basado en la estructura superficial (surface based). Una segunda diferencia es la siguiente: las reglas usuales mediante las cuales se construye una gramática libre de contexto son interpretadas de manera bastante diferente a como son interpretadas en el modelo transformacional. En efecto, en el modelo transformacional las reglas de una gramática son interpretadas como reglas de reescritura, en tanto que en el modelo GESG son interpretadas como condiciones de admisibilidad sobre secuencias de símbolos. Esta interpretación tiene muchas ventajas sobre la otra interpretación. En efecto, podemos construir gramáticas que no lo serían desde el punto de vista de las reglas de reescritura. Para utilizar el ejemplo de Sells (1985) el siguiente fragmento de una gramática dependiente de contexto no puede generar ninguna oración desde la interpretación reescritural:

1. a. S SN SV
- b. SN Juan/---SV
- c. SV camina / SN---

En efecto, desde el punto de vista reescritural este fragmento no puede generar ninguna oración ya que los contextos especificados se implican mutuamente. En cambio si las interpretamos desde

el punto de vista de que imponen condiciones de admisibilidad, estas reglas generan la oración "Juan camina" y especifican que únicamente esta oración puede ser generada.

Una tercera diferencia es que la GESG utiliza metarreglas en la descripción de los fenómenos para las cuales se utilizaban transformaciones. Una cuarta diferencia es que la GESG se expresa en el marco de ID/LP (immediate dominance/linear precedence). Esta condición impone ciertas restricciones sobre la clase de gramáticas, es decir, nos permite seleccionar un subconjunto de las gramáticas libres de contexto (véase Gazdar y Pullum 1981). Una quinta diferencia radica en el modo en que utiliza símbolos complejos. En efecto, en tanto que el modelo Transformacional introduce los símbolos complejos mediante reglas de especificación léxica que luego las reglas transformacionales o fonológicas retoman, la GESG la introduce directamente mediante reglas. Una sexta diferencia es el uso, central en la GESG, de la noción de cabeza (Head). Sobre estos aspectos volveremos después.

En su corta historia el modelo GESG ha pasado por dos etapas. La primera de ellas se extiende desde su formulación a principios del 80 hasta 1984. La segunda etapa se extiende desde 1984 hasta el presente. El tránsito de una etapa a otra está condicionada, en gran parte, por los problemas que enfrentó desde el momento de su postulación y que culminan con los resultados de Uszkoreit y Peters (publicados en 1985, pero conocidos, sin duda, antes de su publicación). Como resultado de esto, la GESG ha buscado nuevos medios para restringir el modelo o reformularlo enteramente. Una de estas posibilidades es el modelo propuesto por Pollard (1984) (su tesis doctoral) y trabajos ulteriores y que se denomina HEAD-DRIVEN GRAMMAR. Otra de las posibilidades exploradas aparece en el libro de Gazdar, Pullum, Klein y Sag (1985), GENERALIZED PHRASE STRUCTURE GRAMMAR. En este trabajo nos centraremos exclusivamente en esta primera etapa.

Uno de los aspectos más llamativos de estos modelos es su posibilidad o factibilidad de ser implementados computacionalmente. En efecto, Shieber (1986) desarrolla un formalismo denominado UNIFICATION-BASED FORMALISM en el que muestra cómo podemos expresar diferentes modelos lingüísticos dentro de ese formalismo (el modelo de GOVERNMENT AND BINDING no puede ser expresado en un formalismo. De acuerdo con el Dr. Maxwell, en comunicación personal, es

necesario hacer reformas algo sustanciales al modelo de Chomsky si queremos expresarlo en dicho formalismo. Sin embargo, hasta el momento Chomsky y colaboradores han hecho caso omiso de estas sugerencias).

1— Motivación teórica del modelo GESG.

Las gramáticas son procedimientos para especificar lenguajes (lenguas en el caso de las gramáticas para las lenguas naturales). Podemos construir lenguajes utilizando los siguientes recursos: Dado un conjunto A, que denominamos conjunto base, podemos construir un nuevo conjunto A^* y que se denomina monoide libre, del siguiente modo:

1. a. $\{ \}$ (el conjunto vacío) pertenece a A^* .
- b. Si x pertenece a A, entonces, también pertenece a A^* .
- c. Si x e y pertenecen a A^* , entonces, xy pertenece a A^* .

Mediante la aplicación de estas reglas podemos construir un conjunto que tiene infinito número de secuencias, aun cuando el conjunto base sea finito. En el caso de las lenguas naturales el vocabulario es finito. Ahora bien, podemos considerar un lenguaje como cualquier subconjunto de A^* . Hay varias formas de especificar o seleccionar estos lenguajes o subconjuntos; uno de ellos es mediante gramáticas.

Una gramática hace uso de un vocabulario auxiliar, denominado vocabulario no terminal (V_n), un conjunto de reglas que me indican cómo construir secuencias o expresiones; un símbolo especial y que se denomina axioma o punto de parte, y que indica que toda derivación en esa gramática debe comenzar con ese símbolo. Es usual utilizar S (entencia) u O(ración). Además, desde luego, hace uso del vocabulario que forman las secuencias. A este vocabulario se denomina vocabulario terminal (V_t). Así pues, una gramática es un cuádruple $\langle V_n, V_t, S, R \rangle$ con las características anteriormente especificadas. Si estuviéramos hablando de una gramática del español, entonces, VN incluiría al menos las categorías siguientes: SN, SV, V, O, SP y otras más.

La forma usual de representar una regla es la siguiente:

AXB ----AYB

y que podemos leer: en el contexto A_B reemplácese X por Y. El contexto puede ser vacío.

Dependiendo del tipo de regla que tengamos son las gramáticas y lenguajes generados por ellas. Hay lenguajes y gramáticas más complejas que otras. La clasificación de los lenguajes de acuerdo con su complejidad es conocida como JERAR—QUIA DE CHOMSKY. De acuerdo con esta clasificación:

Lenguajes regulares \langle Lenguajes libres de contexto \langle lenguajes indizados \langle Lenguajes dependientes de contexto \langle Lenguajes recursivos \langle lenguajes recursivamente enumerables (véase Vargas (1986) para una caracterización de esta jerarquía). En la presentación anterior de la jerarquía el símbolo ' \langle ' indica que el lenguaje de la izquierda es menos complejo que el lenguaje de la derecha. Lo que muestra la prueba de Peters y Ritchie es que una gramática de transformaciones del tipo propuesto en el 65 tiene la capacidad de generar lenguajes que son recursivamente enumerables.

Tanto el modelo de 1965 como el modelo de la GESG, como indicamos anteriormente, utilizan reglas independientes de contexto. Dentro de la jerarquía anterior, las gramáticas libres de contexto ocupan una posición bastante privilegiada. Por un lado, además de ser bastante complejas, poseen una base matemática bastante simple y precisa. Esto posibilita que sean formalizables y decidibles con relativa facilidad (véase Brainerd 1972 sobre estos aspectos).

Independientemente de sus propiedades matemáticas, las gramáticas libres de contexto poseen, desde el punto de vista teórico, otros atractivos. Gazdar (1981; 1982) señala las siguientes:

1— Si las lenguas naturales son libres de contexto, esto nos permitiría imponer una restricción bastante fuerte respecto a la clase de las gramáticas que es necesario considerar como candidatos posibles para las lenguas naturales, esto es, es necesario considerar solamente esta clase de gramáticas cuya caracterización, como indicamos, es bastante precisa.

2— Nos permitiría explicar de manera bastante plausible el hecho de la relativa facilidad con que los hablantes de una lengua "interiorizan" las estructuras sintácticas de esa lengua. En efecto, "las oraciones de una lengua libre de contexto son analizables (parsable) en un tiempo que es proporcional al cubo de la longitud de la oración sino menos" (Gazdar 1981:155).

3— En el modelo generativo-transformacional existen suficientes problemas con los intentos de

incorporar una teoría semántica al modelo, sea ésta de condiciones de verdad o no. Estas dificultades derivan de varios factores. Uno de ellos radica en la distinción entre estructura profunda y estructura superficial. Es difícil "acomodar" una teoría semántica que sea lo suficientemente adecuada para dar cuenta del significado de las oraciones de una lengua natural en ambos niveles. En efecto, hay evidencia de que algunas transformaciones pueden introducir cambios en el significado de las oraciones, con lo cual el significado de una oración en el nivel profundo es diferente del significado en el nivel superficial. Hay varios mecanismos por medio de los cuales las transformaciones pueden introducir cambios de significado. Uno de estos aspectos es la pasivización. Un ejemplo clásico de esto es el siguiente:

2. a. En esta aula todo estudiante habla dos idiomas.
- b. Dos idiomas son hablados por todo estudiante en esta aula.

La oración 2.b. es la correspondiente pasiva de 2.a. Sin embargo, en tanto que 2.a. no implica que los idiomas que cada estudiante habla sean los mismos, 2.b. sí lo implica.

Ahora bien, en el modelo transformacional la necesidad de una semántica para la estructura superficial se hace evidente con la concordancia entre tiempo y aspecto. En efecto, las siguientes oraciones (tomados de Thomason 1974) tendrían la misma estructura profunda, pero diferirían respecto al valor de verdad:

3. a. Un glaciar cubrió la tierra.
- b. Un glaciar cubre la tierra.

En tanto que 3.a. es verdadera (según la información geológica disponible), 3.b. es claramente falsa. Se puede objetar, y esto es correcto, que la diferencia en significado entre 3.a. y 3.b. puede establecerse si consideramos que el tiempo y el aspecto de ambas oraciones es diferente. Sin embargo, dado que son las transformaciones las que introducen la concordancia entre tiempo y aspecto la diferencia de significado tiene que establecerse en el nivel de la estructura superficial. Una de las consecuencias del modelo transformacional del 65 es que nos lleva a "duplicar" de algún modo el componente semántico.

Ahora bien, si pudiéramos trabajar sin transformaciones, es decir, si fuera posible la construcción

de una gramática libre de contexto que sea a la vez lo suficientemente adecuada para tratar todos los "fenómenos" para los cuales se utilizaban las transformaciones y, suficientemente restrictiva, entonces es posible incorporar una teoría semántica. El programa lingüístico de Gazdar es construir una teoría sintáctica que tenga estas propiedades. En el esbozo de 1981; 1982 Gazdar establece para cada regla sintáctica su correspondiente regla semántica.

2— Descripción breve del modelo.

La GESG comienza definiendo un conjunto de categorías sintácticas básicas y, a partir de este conjunto define otro conjunto de categorías sintácticas, llamadas derivadas. Luego se definen conjuntos de reglas básicas y derivadas y, finalmente, se establece un conjunto de metarreglas cuya función, entre otras, es permitimos tratar fenómenos sintácticos para los cuales se utilizaban las transformaciones. Como señalamos anteriormente, las reglas sintácticas en el modelo GESG son interpretadas como reglas que imponen condiciones de admisibilidad sobre las secuencias de símbolos mas que procedimientos de reescritura. Esta característica es transmitida a las metarreglas. Como indican Uszkoreit y Peters:

Las gramáticas contienen reglas para generar oraciones. Las metarreglas son enunciados acerca de estas reglas. Son mecanismos metagramaticales que pueden ser usados para generar reglas de la gramática o para codificar ciertas relaciones entre ellas tales como redundancia en su forma (Uszkoreit y Peters 1985:1).

Además de estos elementos, la GESG utiliza otros procedimientos, entre ellos: la utilización de la teoría X-barras en el establecimiento de relaciones sintácticas, la utilización de símbolos complejos y el formato ID/LP (immediate dominance/linear precedence (véase Gazdar y Pullum 1981). Así pues, la GESG puede ser definida como un séptuple $\langle V_t, V_n, D(V_n), S, R, D(R), \Rightarrow \rangle$, donde V_t es el vocabulario terminal, V_n el vocabulario no terminal, $D(V_n)$ es el vocabulario no terminal derivado, S es el axioma del sistema, R es el conjunto de reglas básicas, $D(R)$ es el conjunto de reglas derivadas y \Rightarrow el conjunto de metarreglas.

Quisiéramos ilustrar lo que hemos dicho hasta el momento. Dado un conjunto de categorías sintácticas básicas (V_n), construimos otro conjunto de categorías sintácticas derivadas del siguiente modo:

Si X e Y son categorías básicas, entonces, X/Y es una categoría sintáctica derivada. Como indicamos más arriba, $D(Vn)$ denota el conjunto de estas nuevas categorías. Si $Vn = \{ A, B \}$, entonces $D(Vn) = \{ A/B, A/A, B/A, B/B \}$. Como puede verse si Vn posee más de un elemento entonces $D(Vn)$ es siempre mayor que Vn, específicamente $D(Vn) > Vn$. Si Vn es finito, como en el caso de las lenguas naturales, entonces $D(Vn)$ lo es también.

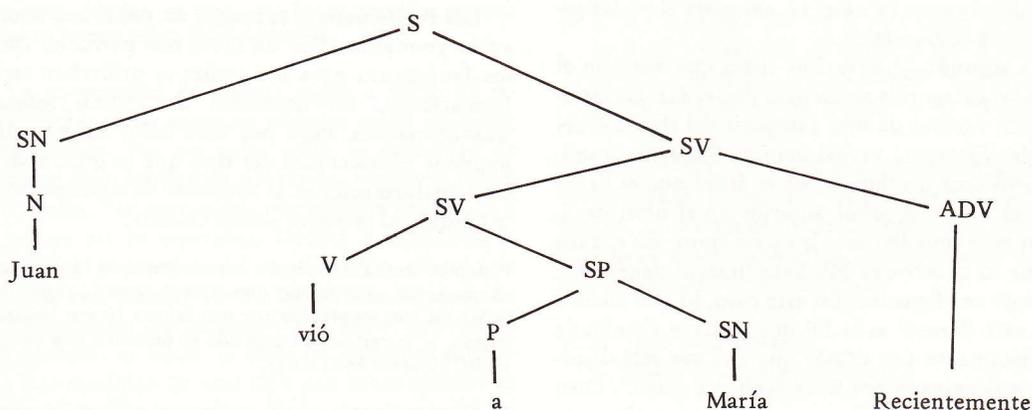
Ahora bien, el rango o dominio de cada una de las categorías sintácticas derivadas puede ser otra(s) categoría(s) derivada(s) o una categoría básica y, al revés, cada categoría sintáctica básica puede dominar inmediatamente sobre una categoría sintáctica derivada. Sin embargo, no todas las

categorías sintácticas derivadas se utilizan en una gramática determinada. El conjunto de categorías sintácticas derivadas que se utilizan viene especificada por las metarreglas. Las categorías sintácticas derivadas juegan un papel fundamental en el modelo de la GESG, ya que nos permiten dar cuenta de los fenómenos para los cuales se utilizaban las transformaciones. Consideremos las siguientes oraciones:

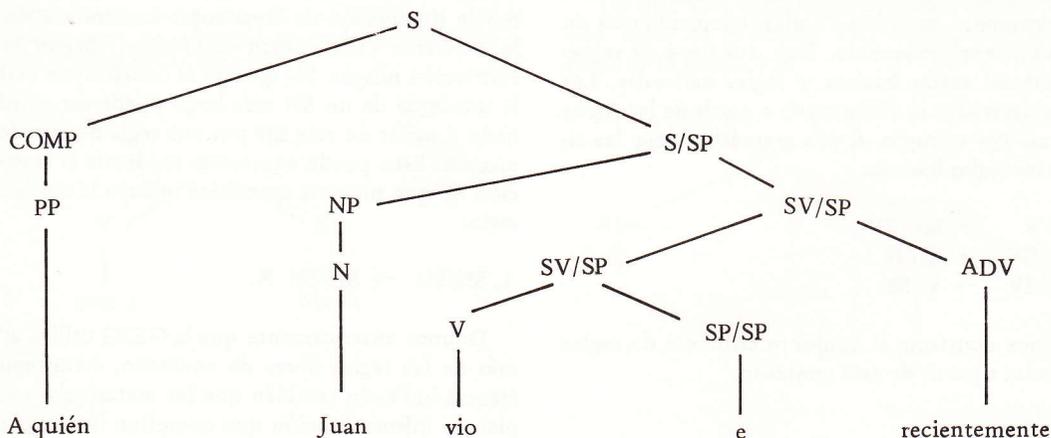
1. a. Juan vió a María recientemente
- b. A quien vió Juan recientemente?.

Según lo que hemos dicho hasta ahora, la GESG asociaría las siguientes estructuras a estas oraciones:

1.a.



1.b.



Hay varias observaciones que hacer respecto a este proceso derivacional. La primera observación tiene que ver con el uso del símbolo 'e' (vacío). En una gramática que interprete sus reglas sintagmáticas como reglas de reescritura no es posible incorporar reglas que dominen sobre vacío, ya que este tipo de gramáticas tienen el poder de un sistema restrictivo ilimitado (véase Chomsky 1977 (citado en Gazdar 1982: 138). Sin embargo, si interpretamos las reglas como se hace en la GESG como condiciones de admisibilidad el problema desaparece: esto: "no es verdadero [...] bajo el modo de interpretación de admisibilidad. El único efecto sobre la capacidad generativa de una gramática de estructura sintagmática que permita producciones-e (sea libre de contexto o dependiente de contexto) [...] es que el lenguaje que consiste de la secuencia vacía puede ser generado" (Gazdar 1982:138). Bajo esta interpretación entonces, la incorporación de producciones-e (vacías) no aumenta el poder generativo de la gramática.

Una segunda observación tiene que ver con el uso de las categorías sintácticas derivadas. La interpretación normal de una categoría del tipo SV/SN es la de 'Sintagma verbal menos Sintagma nominal'. Esto crea un 'hueco' en el árbol que es llenado en el caso del árbol anterior en el nivel de S. Los 'huecos' pueden ser de varios tipos. En el caso anterior lo constituye PP. Este 'hueco' debe subir hasta que sea llenado. En este caso, lo que viene a llenar este 'hueco' es la PP que aparece dominada inmediatamente por COMP que a su vez está directamente dominada por S, es decir, "a quién". Otro de los aspectos que debemos señalar es que las producciones-e aparecen dominadas por la categoría sintáctica X/X, en el ejemplo anterior por SP/SP.

Antes de decir algo sobre las metarreglas veamos las reglas que utiliza la GESG. Como dijimos anteriormente se utilizan reglas independientes de contexto exclusivamente. Hay dos tipos de reglas sintácticas: reglas básicas y reglas derivadas. Las reglas derivadas se construyen a partir de las reglas básicas. Por ejemplo, si una gramática tiene las siguientes reglas básicas:

2. a. S → SN SV
- b. SN → det N
- c. SV → V SN

podemos construir el conjunto siguiente de reglas derivadas a partir de este conjunto.

3. a. S/SN → SN SV/SN
- b. S/SN → SN/SN SV
- c. SN/SN → det N/SN
- d. SV/SN → V SN/SN
- e. S/SV → SN SV/SV
- f. SN/SV → det N/SV
- g. SV/SV → V SN/SV

En general, si $S \rightarrow Y Z W$ es una regla básica, entonces, $X/K \rightarrow Z/K W$ es una regla derivada, donde Z y K son miembros de Vn. Debemos indicar que algunas de las posibilidades de combinación indicadas en 3 no aparecerían en ninguna gramática. Ejemplos de este tipo son las reglas 3.c y 3.g. En efecto, no es posible que en una gramática categorías de la forma X/X aparezcan al lado izquierdo de las reglas, por la siguiente razón: X/X es equivalente a X menos X, es decir, e (vacío) y vacío no puede dominar sobre ninguna categoría.

Las reglas derivadas juegan un papel importante en la gramática. Por un lado, nos permiten tratar los fenómenos para los cuales se utilizaban transformaciones, sin necesidad de utilizar ninguna transformación. Pero por otro lado, nos permiten imponer restricciones del tipo que se imponían en las transformaciones al formular su descripción estructural. Al respecto señala Gazdar:

Podemos formalizar ciertas restricciones, si lo deseamos, simplemente estipulando que ciertos tipos de reglas derivadas no son empleadas por una lengua (o por cualquier lengua, si la restricción sugerida se pretende que sea universal) (Gazdar 1981:161).

Gazdar ilustra cómo podemos utilizar las reglas derivadas para imponer restricciones sobre el tipo de secuencias admitidas en una lengua con varios tipos de restricciones encontradas en la década de los 70. Una de estas restricciones es conocida como la Restricción de Ross sobre la ramificación a la izquierda (Left Branch Condition). Según esta restricción ningún SN que sea el constituyen más a la izquierda de un SN más largo puede ser reordenado a partir de este SN por una regla de transformación. Esto puede expresarse mediante la restricción de que ninguna gramática incluya la siguiente regla:

4. SN/SN → SN/SN X.

Dijimos anteriormente que la GESG utiliza además de las reglas libres de contexto, metarreglas. Hemos indicado también que las metarreglas cumplen la misma función que cumplían las transfor-

maciones en el modelo transformacional. Siguiendo a Uszkoreit y Peters podemos definir una meta-regla como un acople (template) de dos reglas A y B (ambas libres de contexto) y que denotamos $A \Rightarrow B$, con la interpretación siguiente:

Si una gramática contiene una regla de la forma A, también contiene la regla correspondiente de la forma B. Como esta interpretación sugiere las reglas de una gramática son cerradas bajo la aplicación de las metarreglas (Uszkoreit y Peters 1985:1).

Las metarreglas son, al igual que el componente transformacional, una extensión del conjunto de reglas de estructura sintagmática base. Pero como señala Sells (1985) más allá de esta similitud no se puede ir. En efecto, existen diferencias fundamentales entre una transformación y una metarregla.

Veamos ahora la relación entre las metarreglas y las transformaciones. Las transformaciones mapean (map) árboles en nuevos árboles, en tanto que las metarreglas mapean reglas en nuevas reglas (Sells 1985:92).

La proyección de nuevas reglas a partir de reglas básicas tiene enormes ventajas sobre las transformaciones. Por un lado, evita la necesidad de postular una estructura profunda, necesaria de algún modo. Por otro lado, nos permite restringir, o al menos así lo esperaban Gazdar y asociados, la capacidad generativa del modelo. En efecto, de acuerdo con Gazdar (1981;1982) la incorporación de metarreglas no incrementa el poder generativo del modelo, es decir, la clase de lenguajes generados por modelos de este tipo son subconjuntos de la clase de los lenguajes libres de contexto. De acuerdo con Gazdar:

[...] si uno agrega metarreglas a una gramática de estructura sintagmática, entonces, uno únicamente alarga, de un modo gobernado por reglas, el conjunto de las reglas de estructura sintagmática que uno está utilizando, pero la gramática como un todo sigue siendo de estructura sintagmática (Gazdar 1982:160).

Con propósitos ilustrativos y sin pretender que al análisis sea enteramente correcto, veamos el tipo de reglas que están involucradas en la derivación de las oraciones 1.a y 1.b de este apartado. En la oración 1.a están involucradas las siguientes reglas:

5. a. S \rightarrow SN SV
- b. SV \rightarrow SV ADV
- c. SV \rightarrow V SP
- d. SP \rightarrow Prep SN
- e. SN \rightarrow (det.) N

En la construcción de la oración 1.b a partir de 1.a están involucradas (no hablamos de derivar en el sentido de Chomsky 1965, sino más de ampliar el conjunto de oraciones aceptables una vez que ciertas condiciones han sido satisfechas) las siguientes metarreglas:

6. a. S \rightarrow SN SV \Rightarrow S \rightarrow COMP S/SP
- b. S \rightarrow SN SV \Rightarrow S/SP \rightarrow SN SV/SP
- c. SV \rightarrow SV ADV \Rightarrow SV/SP \rightarrow SV/SP ADV
- d. SV \rightarrow V SP \Rightarrow SV/SP \rightarrow V SP/SP

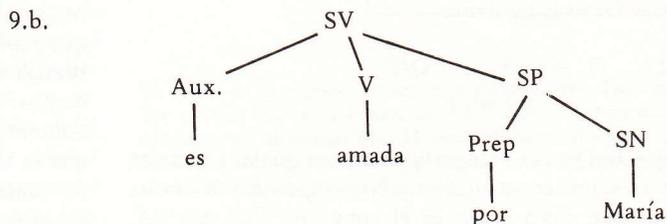
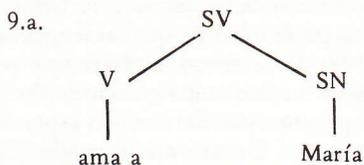
Ahora bien, no en todas las metarreglas están involucradas categorías sintácticas derivadas. En efecto, el tratamiento de las oraciones pasivas no involucra estas categorías. Para ver esto consideremos las siguientes oraciones.

7. a. Juan ama a María.
- b. María es amada por Juan.

la siguiente metarregla da cuenta de esta relación:

8. SV \rightarrow V SN \Rightarrow SV \rightarrow aux V SP

Lo que esta metarregla establece es que si hay una rama de la forma 9.a., entonces existe una rama de la forma 9.b.:



3. Lenguas naturales y gramaticales libres de contexto.

Chomsky en SYNTACTIC STRUCTURES mostró de manera concluyente que las lenguas naturales no pueden ser analizadas de manera satisfactoria mediante una gramática regular, esto es, que las lenguas naturales no son regulares. Chomsky (1964) y Postal (1964) argumentan que las gramáticas libres de contexto son inadecuadas por sí mismas para tratar las lenguas naturales. En general argumentan que las gramáticas de estructura sintagmática son inadecuadas para analizarlas, es necesario extenderlas en algún sentido. Sin embargo, no hubo una preocupación fuerte por mostrar, si es que esto se puede hacer, desde un punto de vista definitivo la inadecuación de las gramáticas de estructura sintagmática. La cuestión de si las lenguas naturales son libres de contexto o no, sigue siendo una cuestión abierta. Sin embargo, han aparecido varios trabajos tendientes a evaluar dicha afirmación. Uno de estos trabajos es el de Bresnan, Kaplan, Peters y Zeanen (1982). Otro trabajo en este sentido es el de Higgenbotham (1984).

Podemos mostrar que las lenguas naturales no son libres de contexto si podemos mostrar una o las dos de las siguientes posibilidades: a- existen algunos 'fenómenos' en una lengua natural particular que no pueden ser tratados de manera satisfactoria mediante una gramática libre de contexto; b- mostrando que las lenguas naturales tienen propiedades diferentes a las propiedades de las gramáticas libres de contexto. Bresnan et. al. (1982) se ocupa del primer aspecto; Higgenbotham (1984) del segundo. Si estos resultados son correctos, entonces, tenemos que evaluar el modelo GESG.

Una gramática para las lenguas naturales debe cumplir ciertas condiciones: por un lado, debe ser descriptivamente adecuada, es decir, debe dar cuenta adecuada de un corpus dado. Por otro lado, debe ser explicativamente adecuada, esto es, debe ser capaz de asignar las descripciones estructurales adecuadas a las oraciones del corpus. A la capacidad que tiene una gramática de generar un corpus determinado se denomina capacidad generativa débil, es decir, el conjunto de secuencias que aparecen como líneas últimas en la derivación debe coincidir con el corpus. Y la capacidad de asignar descripciones estructurales adecuadas al corpus, se denomina capacidad generativa fuerte. En este sentido, dos gramáticas pueden tener la misma capacidad generativa débil pero diferir respecto a la capacidad generativa fuerte. De igual manera una gra-

mática puede ser descriptivamente adecuada pero explicativamente inadecuada.

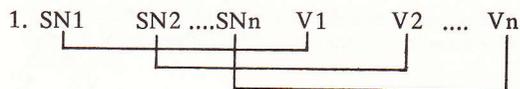
Pues bien, Bresnan et. al. (1982) argumentan que una gramática libre de contexto es descriptivamente adecuada pero explicativamente inadecuada y, en este sentido, las lenguas naturales no son fuertemente libres de contexto. Podemos argumentar que las lenguas naturales no son fuertemente libres de contexto, si encontramos al menos una lengua natural para la cual cualquier gramática libre de contexto, por razones intrínsecas a las gramáticas libres de contexto, no pueda asignar las descripciones estructurales adecuadas. Una lengua de este tipo, de acuerdo con Bresnan et. al. es el holandés:

Mostraremos en este artículo que el holandés es justamente una lengua tal, y así, que las lenguas naturales en general no son fuertemente libres de contexto. Esto no implica, sin embargo, que una descripción adecuada de una lengua natural requiere el poder pleno de una gramática transformacional: también mostraremos que las construcciones problemáticas del holandés son fuertemente generadas por una gramática léxico-funcional (Bresnan et. al. 1982:614).

Nos interesa aquí principalmente el primer aspecto, es decir, que el holandés no es fuertemente generado por una gramática libre de contexto. Por ello dejaremos de lado el enfoque léxico-funcional (sobre esto véase Sells 1985).

En el análisis propuesto por Bresnan et. al. no se utilizan metarreglas. Sin embargo, esto no tiene incidencia fuerte sobre nuestra discusión, ya que la GESG supone que la introducción de metarreglas no incrementa el poder generativo de la clase de los lenguajes generados por las gramáticas libres de contexto.

El holandés admite series de dependencias cruzadas con una estructura similar a la siguiente:



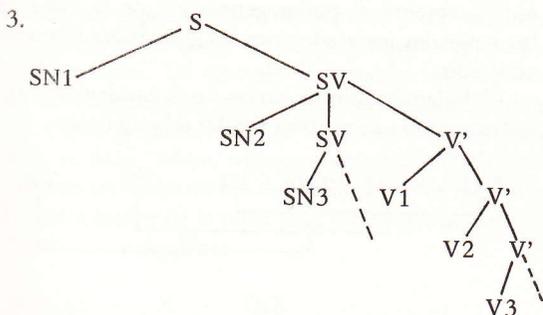
El verbo en la primera posición es formalmente distinguible ya que marca la concordancia en tiempo, persona y número con el primer SN. El verbo de la última posición es distinguido de los otros por restricciones de subcategorización (Bresnan et. al. 1982:615).

Estructuras de este tipo pueden ser generadas por una gramática libre de contexto que incluya símbolos complejos, como la siguiente:

2. a. $S \rightarrow \begin{matrix} SN \\ \left[\begin{matrix} \alpha \text{ pl} \\ \beta \text{ pers} \end{matrix} \right] \end{matrix} \begin{matrix} S' \\ \left[\begin{matrix} \alpha \text{ pl} \\ \beta \text{ pers} \\ + n \end{matrix} \right] \end{matrix} V$
- b. $S' \rightarrow \begin{matrix} SN \\ \left[\begin{matrix} \alpha \text{ pl} \\ \beta \text{ pers} \\ + n \end{matrix} \right] \end{matrix} \begin{matrix} S' \\ \left[\begin{matrix} \alpha \text{ pl} \\ \beta \text{ pers} \\ + n \end{matrix} \right] \end{matrix} V$
- c. $S' \rightarrow \begin{matrix} SN \\ \left[\begin{matrix} \alpha \text{ pl} \\ \beta \text{ pers} \\ + n \end{matrix} \right] \end{matrix} \begin{matrix} S'' \\ \left[\begin{matrix} \alpha \text{ pl} \\ \beta \text{ pers} \\ + n \end{matrix} \right] \end{matrix} V$
- d. $S'' \rightarrow \begin{matrix} SNn \\ \left[\begin{matrix} \alpha \text{ pl} \\ \beta \text{ pers} \\ + n \end{matrix} \right] \end{matrix} \begin{matrix} V \\ \left[\begin{matrix} \alpha \text{ pl} \\ \beta \text{ pers} \end{matrix} \right] \end{matrix}$

donde α significa + 0 - plural, β primera, segunda o tercera persona, y n es el número de verbos que pueden ser subcategorizados en holandés.

Sin embargo, a pesar de que esta gramática genera las dependencias cruzadas del holandés, falla en asignar las descripciones estructurales adecuadas. Lo que se debe mostrar es que no existe una gramática libre de contexto que pueda asignar las descripciones estructurales adecuadas a este tipo de dependencias. Para ello los autores discuten cual es la estructura correcta de estas dependencias y establecen la siguiente:



Sin embargo, esta estructura no puede ser fuertemente generada por ninguna gramática libre de contexto. La prueba de ello es que depende de un algoritmo que para las gramáticas libres de contexto estableciera Thatcher (Thatcher 1967 y 1973, ambos citados en Bresnan et. al). De acuerdo con este

resultado para cada gramática libre de contexto existe una máquina de estados finitos con una pila (pushdown) de memoria tal que el conjunto generado por una gramática libre de contexto es aceptado por la máquina y,

para cada conjunto de árboles aceptable, existe una constante n tal que cualquier árbol en el conjunto que tenga una longitud mayor que n puede ser particionado en tres partes t_1 , t_2 y t_3 donde la longitud del subárbol t_2 - t_3 es menor que n y cualquier árbol formado por iteración de la parte t_2 , también pertenece al conjunto (Bresnan et. al. 1982:621).

En la estructura arbórea 3, existe una correspondencia entre los SNs y los Vs, y cuando se viola esta correspondencia se obtienen oraciones agramaticales. Pues bien, partiendo de que las dependencias del holandés tienen este carácter, la iteración de la parte t_2 nos lleva a violar tal correspondencia. Por tanto, tales dependencias no pueden ser generadas fuertemente por ninguna gramática libre de contexto (véase Bresnan et. al para más detalles).

Por otro lado, Higgenbotham (1984) argumenta que el inglés no es libre de contexto. Para probar esto hace uso de algunas de las propiedades matemáticas de los lenguajes libres de contexto, especialmente, de aquella según la cual los lenguajes libres de contexto son cerrados respecto a la intersección con los lenguajes regulares. Es decir, dado un lenguaje regular y un lenguaje libre de contexto, su intersección es siempre un lenguaje libre de contexto. Higgenbotham construye un lenguaje regular y al intersecarlo con el inglés obtiene un lenguaje que no es libre de contexto. Esto prueba, según Higgenbotham que el inglés no es libre de contexto. No nos vamos a extender en los detalles de esta prueba. Lo que nos interesa señalar es que esta prueba es algo artificiosa ya que depende del supuesto de que el inglés admite un número ilimitado de incrustaciones, cosa que en el inglés ordinario y en cualquier lengua es difícil de encontrar. Creemos que la prueba de que las lenguas naturales no son libres de contexto debe proceder en alguno de los dos siguientes sentidos: por un lado, analizando y evaluando los modelos lingüísticos que respaldan esa afirmación o, por el otro, buscando evidencia lingüística (empírica) de que ciertos fenómenos lingüísticos no pueden ser tratados mediante una gramática libre de contexto. En el primer sentido, ubicamos las investigaciones de Uszkoreit y Peters, y en el segundo sentido, las investigaciones del tipo de Bresnan et al.

Ahora bien, suponiendo que estos resultados son correctos, podemos esperar de la GESG alguna de las dos cuestiones siguientes:

1— Que este modelo es inadecuado en el sentido de que es incapaz de generar adecuadamente algunos fenómenos lingüísticos como la serie de dependencias cruzadas del holandés y, por tanto, del holandés;

2— Que la incorporación de metarreglas incrementa el poder generativo de la gramática, es decir, los lenguajes generados por las GESGs no son exclusivamente libres de contexto.

Los resultados de Uszkoreit y Peters (1985) muestran que el poder generativo débil de los modelos que incorporan el tipo de metarreglas caracterizadas hasta el momento recorren toda la jerarquía de Chomsky, es decir, que estos modelos plantean los mismos problemas de decidibilidad que plantea el modelo de Chomsky (1965).

4. Los resultados de Uszkoreit y Peters.

Desde la formulación de la GESG se notó que la introducción de variables esenciales incrementaba el poder generativo de la gramática, aunque no se conocía el carácter de este incremento. Esto entre otras cosas debido a que no se había formalizado un modelo gramatical con metarreglas y las propiedades de las metarreglas no se conocían totalmente. Se buscó, sin embargo, restringir el poder generativo de la GESG permitiendo la aparición de una única variable esencial en las metarreglas.

Entendemos que Aravind Joshi ha hecho la siguiente conjetura: usando una sola variable esencial no se incrementa el poder generativo más allá de los lenguajes libres de contexto —aun si un número infinito de reglas son derivadas. Esta conjetura fue aparentemente aceptada por la GESG como un hecho establecido. Una única variable esencial llegó a ser considerada 'segura'. (Uszkoreit y Peters 1985: 4).

Los resultados principales de Uszkoreit y Peters tienen que ver con las gramáticas con una sola variable esencial. Muestran que esta conjetura es falsa y que una gramática con una sola variable esencial posee un poder generativo débil (solo se ocupan del poder generativo débil) va más allá de los lenguajes libres de contexto.

Ahora bien, el carácter de los resultados de estos autores es bastante general, y no solo afectan a la GESG sino también a otros modelos lingüísticos que utilicen metarreglas, entre ellos el ANNOTATED PHRASE STRUCTURE GRAMMAR de-

sarrollado por Konolige. Para garantizarse la generalización de los resultados denominada GRAMÁTICA DE ESTRUCTURA SINTAGMÁTICA CON METARREGLAS (Metarule Phrase Structure Grammar) y que abreviaremos como ESM. Gramáticas en las que aparece una única variable esencial serán referidas como ESM/1.

Se define una gramática ESM como un séxtuple $\langle V_n, V_t, S, R, V_v, \Rightarrow \rangle$ donde V_n es el vocabulario no terminal, V_t es el vocabulario terminal, S es el axioma del sistema, R es un conjunto de reglas libres de contexto, V_v es el conjunto de variables esenciales cuyo rango es $(V_n \cup V_t)^*$, y \Rightarrow es una relación diádica que toma como argumentos reglas libres de contexto, es decir, subconjunto de R . En el caso de las gramáticas ESM/1, el conjunto V_v consta de una única variable esencial.

Como indicamos anteriormente, los resultados de Uszkoreit y Peters tienen que ver exclusivamente con el poder generativo débil y sus implicaciones para los modelos lingüísticos con metarreglas. Presentaremos primero los resultados generales más importantes y luego daremos ejemplos ilustrativos de estos resultados. De acuerdo con Uszkoreit y Peters.

(i) las gramáticas ESM/1 generan todos los lenguajes recursivamente enumerables.

Dado que las gramáticas ESM generan solo lenguajes recursivamente enumerables, se sigue que las gramáticas ESM/1 generan exactamente la clase de los lenguajes recursivamente enumerables.

(ii) Gramáticas ESM sin símbolos fantasma generan lenguajes que, si bien son infinitos, son 'aritméticamente densos'.

(iii) Gramáticas ESM/1 sin símbolos fantasma generan algunos lenguajes no recursivos. Dado que muchos lenguajes dependientes de contexto, y en verdad, muchos lenguajes indizados no son 'aritméticamente densos', las gramáticas ESM sin símbolos fantasma generan una clase de lenguajes que no pueden ser comparados con la clase de lenguajes dependientes de contexto o con la clase de lenguajes recursivos. De hecho,

(iv) gramáticas ESM/1 sin símbolos fantasma generan todos los lenguajes libres de contexto (Uszkoreit y Peters 1985: 5-6).

Quisiéramos ilustrar algunos de los resultados recién citados. Para probar el primer resultado se hace lo siguiente: dada una gramática recursiva, o lo que es lo mismo, sistema irrestricto, se muestra que existe un mecanismo para convertir dicha gramática en una gramática ESM/1, y que el lenguaje generado por la gramática recursiva es subconjunto del lenguaje generado por la gramática ESM/1. Como indicamos anteriormente, toda derivación de una gramática debe comenzar con el símbolo S

y tal que, el lenguaje generado por una gramática es el conjunto de secuencias w , tales que $S \rightarrow w$. El procedimiento de conversión de una gramática recursiva en una gramática ESM/1 sigue el siguiente proceso: sea $G = \langle Vt, Vn, S, R \rangle$ una gramática recursiva. Para construir la ESM/1 correspondiente agréguese a Vn los símbolos S' y A' y tómese la siguiente regla básica:

$$1. S' \rightarrow A' S A'$$

Y para toda regla de G , $X \rightarrow Y$ constrúyase la correspondiente metarregla:

$$2. S' \rightarrow A' X B \Rightarrow S' \rightarrow S' Y B.$$

Para toda a que pertenece a $Vt \cup Vn \cup \{A'\}$ establézcase la siguiente metarregla:

$$3. S' \rightarrow A' a B \Rightarrow S' \rightarrow A' Ba.$$

Finalmente, la siguiente metarregla nos permite eliminar los símbolos fantasma:

$$4. S' \rightarrow A' B A' \Rightarrow S' \rightarrow B.$$

Nótese que la única variable esencial utilizada en este proceso es B . No existe otra. Por lo tanto, la gramática en cuestión es ESM/1. Dado que las metarreglas son cerradas respecto a la aplicación de las reglas básicas, en este caso, la regla básica 1, y dado que ESM/1 no agrega nada a G , excepto los símbolos S' y A' que son, respectivamente, el nuevo axioma y el símbolo fantasma, y la variable significativa B , se sigue que el lenguaje generado por G es también generado por ESM/1. De ahí que el lenguaje generado por G sea subconjunto del lenguaje generado por ESM/1. Esto pone de manifiesto el poder generativo de las gramáticas que utilizan metarreglas con una única variable esencial.

La siguiente gramática ESM/1 nos permite generar un lenguaje libre de contexto. Sea $Vt = (a, b)$, $Vn = (X)$ y la siguiente regla básica (B) y metarreglas (M):

$$5. B. S \rightarrow AabA$$

$$M.1. S \rightarrow AaXA \Rightarrow S \rightarrow AXAaa$$

$$M.2. S \rightarrow AbX \Rightarrow S \rightarrow AXbb$$

$$M.3. S \rightarrow AAX \Rightarrow S \rightarrow AXA$$

$$M.4. S \rightarrow AXA \Rightarrow S \rightarrow X$$

Como puede verse el único símbolo fantasma utilizado es A y la única variable esencial o significativa de la gramática es X .

El lenguaje generado por esta gramática es $a^{2^n} b^{2^n}$ que es un subconjunto de $a^n b^n$ y que es libre de contexto. Veamos como funcionan las reglas de esta gramática. Las secuencias $ab, aabb, aaaabbbb$, pertenecen al lenguaje generado por esta gramática dado que:

$$6. S \rightarrow AabA \rightarrow ab \text{ (aplicando primero la regla B y luego M4).}$$

$$7. S \rightarrow AabA \rightarrow AbAaa \rightarrow AAaabb \rightarrow AaabbA \rightarrow aabb \text{ (reglas B, M1, M2, M3, M4).}$$

$$8. S \rightarrow AabA \rightarrow AbAaa \rightarrow AAaabb \rightarrow AaabbA \rightarrow AabbAaa \rightarrow AbbAaaaa \rightarrow AbAaaaabb \rightarrow AAaaaabbbb \rightarrow AaaaabbbbA \rightarrow \text{(reglas B, M1, M2, M3, M1, M1, M2, M2, M3, M4).}$$

Nótese que en todos los casos se debe comenzar con las reglas básicas. Las metarreglas se aplican únicamente si existen los inputs apropiados y estos vienen dados, inicialmente, por las reglas básicas. Es en este sentido en el que se dice que las metarreglas son cerradas con respecto a la aplicación de las metarreglas.

Si cambiamos la regla básica y agregamos la siguiente metarregla:

$$9. B'. S \rightarrow AabcA$$

$$M.5. S \rightarrow AcX \Rightarrow S \rightarrow AXcc$$

entonces, el lenguaje generado por esta gramática es dependiente

de contexto. En efecto, genera $a^{2^n} b^{2^n} c^{2^n}$.

Por ejemplo, la secuencia $aaaabbbbcccc$ se obtiene mediante la siguiente derivación:

$$10. S \rightarrow AabcA \rightarrow AbcAaa \rightarrow AcAaabb \rightarrow AAaabbcc \rightarrow AaabbccA \rightarrow AabbccAaa \rightarrow AbbccAaaaa \rightarrow AbccAaaaabb \rightarrow AccAaaaabbbb \rightarrow AcAaaaabbbbcc \rightarrow AAaaaabbbbcccc \text{ (reglas B', M1, M2, M5, M3, M1, M1, M2, M2, M5, M5, M3, M4).}$$

Podríamos construir otras gramáticas ESM/1 para ilustrar el poder generativo de estas gramáticas. Sin embargo, los ejemplos que hemos construi-

do muestran que el poder generativo de estas gramáticas va más allá de las gramáticas libres de contexto. Esto significa que las gramáticas ESM/1 tienen la capacidad de generar toda la jerarquía de Chomsky. La prueba de este resultado sigue en general el siguiente proceso: dad una gramática con tales y tales características (recursiva, dependiente de contexto, independiente de contexto, regular) y que genera el lenguaje L, construimos una gramática ESM/1 con un símbolo fantasma que genera un lenguaje tal que L es subconjunto del generado por ESM/1.

Por otro lado, los autores prueban que si restringimos las gramáticas ESM en el sentido de no permitir la aparición de símbolos fantasma (pero sí variables esenciales), la clase de lenguajes generados por estas gramáticas es un subconjunto propio de la clase de lenguajes recursivos y no la clase entera. Aun con esta restricción estas gramáticas generan lenguajes que no son libres de contexto. Tal es el caso de los lenguajes aritméticamente, los indexados (indexed), etc. Este subconjunto también puede ser generado por la clase de las gramáticas ESM/1 con un símbolo fantasma.

Ahora bien, de la formalización anterior se derivan consecuencias importantes para los modelos lingüísticos que utilizan metarreglas, en especial para el modelo GESG. Estas consecuencias son de dos tipos: por un lado, respecto a la afirmación de que las lenguas naturales son libres de contexto, y por el otro, respecto a la fertilidad de este tipo de modelos.

Respecto al primer aspecto, los resultados anteriores nos permiten cuestionar fuertemente la base teórica y la evidencia empírica sobre la que descansa la afirmación de que las lenguas naturales son libres de contexto. En efecto, se debe reconsiderar el éxito de la GESG en el tratamiento de los procesos sintácticos en el sentido en que han sido utilizados para apoyar tal afirmación.

Respecto al segundo de los aspectos, dado el gran poder generativo de GESG, un modelo gramatical para las lenguas naturales debe evitar el uso de metarreglas en el sentido que han sido utilizadas hasta 1984, o de restringir el uso de estos procedimientos. Uszkoreit y Peters señalan algunos intentos de restringir el uso de metarreglas dentro de la GESG. El primero de ellos sugerido en 1982 por Gazdar consiste en redefinir las metarreglas permitiendo la aparición de una única variable esencial, pero que el rango de esta variable admite solo un número finito de instanciaciones, esto es, que pueda ser especificada mediante un número finito de

reglas. O lo que es lo mismo eliminar la aparición de variables esenciales en favor de las variables insenciales (en el sentido señalado anteriormente). Sin embargo, una solución como esta tiene consecuencias teóricas importantes que no deben dejarse de lado. Supongamos para utilizar el ejemplo de Uszkoreit y Peters, que una lengua VSO posee la siguiente metarregla:

$$11. \quad SV \rightarrow V U \Rightarrow SV \rightarrow V SV \quad U$$

donde el rango de $U = \{ \{ \}, SN, S', SN, SV' \}$, donde S' y SV' son expresiones de la teoría X-barra (X').

La diferencia de significado entre una metarregla con una variable abreviatoria (inesencial) y la correspondiente metarregla con variables esenciales puede verse como un contraste entre intensidad-extensión. La metarregla (11) en nuestro ejemplo supone expresar el hecho (entre otros) que todos los SVs pueden aparecer en oraciones SVO como constituyentes discontinuos. Si U es una variable abreviatoria, la frase "todos los SVS" refiere solo a la lista específica de los SVs conocidos en la gramática hasta el momento de escribirse la gramática. Si U es una variable esencial, la metarregla establece un enunciado más fuerte, esto es, que actualmente todos los SVs en el lenguaje independientemente de nuestro conocimiento o de la cobertura actual de la gramática, participan en el fenómeno sintáctico (Uszkoreit y Peters 1985: 17).

Así pues, un intento como este conlleva la consecuencia, desagradable para una gramática de pérdida de generalizaciones.

El segundo intento por restringir el poder generativo de las GESGs consiste en restringir el papel que desempeñan las metarreglas en la gramática pero dejándolas tal y como han sido formuladas hasta el momento. Esta idea fue sugerida por Thompson desde 1982 y consiste en "cerrar las reglas de estructura sintagmática de la gramática respecto a las metarreglas", es decir, deben ser 'finitamente cerradas'. Se dice que una gramática es finitamente cerrada con respecto a las metarreglas si cumple al menos una de las siguientes condiciones:

12. Para cualquier metarregla $M = A \Rightarrow B$
 1. B es una regla básica;
 2. B es una regla derivada sin la intervención de M en el proceso derivacional anterior.

Aunque el concepto de 'finitamente cerrado' hace a las metarreglas 'seguras' no es una solución fundada lingüísticamente. Muchas otras estipulaciones pueden mantener el conjunto de las reglas finito. Por ejemplo, podemos imponer un límite pequeño sobre la longitud de las secuen-

cias que una variable esencial puede tomar como rango. Esta estipulación sería ad hoc como lo es la solución finitamente cerrada, pero no requeriría una redefinición del funcionamiento de la metarregla; sin embargo, simplificaría el cálculo de las reglas de estructura sintagmática a partir del conjunto de reglas básicas (Uszkoreit y Peters 1985:19).

Lo que se requiere son restricciones lingüísticamente motivadas que nos permitan redefinir las metarreglas. Sin embargo, en este nivel se encuentran otros inconvenientes, entre ellos, el que aparte de las motivaciones que permitirían la introducción de metarreglas no se tiene claridad respecto al tipo de 'fenómenos' lingüísticos que deben ser tratados con las metarreglas, así como los que se encontrarían al intentar redefinir las metarreglas. Gazdar, Klein, Sag y Pullum (1985) ofrecen una reformulación del modelo. Dejan las metarreglas tal y como han sido introducidas hasta el momento pero restringen su aplicación mediante un conjunto de restricciones generales, entre ellas, LEXICAL HEAD CONSTRAINT que limita la aplicación de las metarreglas a la subcategorización léxica. Esto es posible porque en esta nueva reformulación de la GESG se establecen en el léxico muchas de las especificaciones y restricciones. Otro de los intentos por salir de estos problemas es la reformulación que hace Pollard (1984 y ss), en el que elimina del todo la aparición de metarreglas.

Uno de los aspectos interesantes es que estas reformulaciones, cuya evaluación desde el punto de vista de la formalización, está por hacerse, dejan abierta la cuestión de si las lenguas naturales son libres de contexto. Además del trabajo de 1985, Gazdar y asociados se han ocupado del tema en varios reportes técnicos de investigaciones realizadas en CENTER FOR THE STUDY OF LANGUAGE AND INFORMATION de la Universidad de Stanford. Entre ellos "Computationality relevant properties of Natural Languages" (Report-24) y "Applicability of Indexed Grammar to Natural Languages" (Report-34).

BIBLIOGRAFIA

- Brainerd, Bernard (1972) INTRODUCTION TO THE MATHEMATICS OF LANGUAGE STUDY. American Elsevier Publishing Co. Inc. New York.
- Bresnan, Joan et. al (1982) "Cross-serial Dependencies in Dutch" en: LINGUISTIC INQUIRY (13) 613-635.
- Chomsky, Noam (1965) ASPECTS OF THE THEORY OF SYNTAX. The MIT Press Cambridge.
- _____ (1981) LECTURES OF GOVERNMENT AND BINDING. Foris, Dordrecht.
- Gazdar, Gerard (1981) "Unbounded Dependencies and Coordinate Structure" en LINGUISTIC INQUIRY (12) 155-184.
- _____. (1982) "Phrase Structure Grammar" en, Jacobson y Pullum (1982) THE NATURE OF SYNTACTIC REPRESENTATION. D. Reidel Publishing Co.; 133-186.
- Gazdar, Gerard y Pullum, Jeffrey (1981) "Subcategorization, Constituent Order, and the notion of "head" en, Moortgat et. al. (1981) THE SCOPE OF LEXICAL RULES. Foris, Dordrecht.
- Gazdar, Gerard et. al. (1985) GENERALIZED PHRASE STRUCTURE GRAMMAR. Cambridge Mass., Harvard University Press.
- Higgenbotham, James (1984) "English is Not a Context-Free Language" en: LINGUISTIC INQUIRY (15)125-134.
- Peters, Stanley y Ritchie, Robert (1967; 1973) "On the Generative Power of Transformational Grammars" en INFORMATION SCIENCES (6) 49-83.
- Pollard, Carl (1984) GENERALIZED PHRASE STRUCTURE GRAMMAR, HEAD GRAMMARS AND NATURAL LANGUAGES. Tesis doctoral, Universidad de Stanford.
- Sells, Peters (1985) LECTURES ON CONTEMPORARY SYNTACTIC THEORIES. Center for the Study of Language and Information. Universidad de Stanford.
- Shieber, Stuart (1986) AN INTRODUCTION TO UNIFICATION-BASED APPROACHES TO GRAMMAR. Center for the Study of Language and Information, Universidad de Stanford.
- Uszkoreit, Hans y Peters, Stanley (1985) ON SOME FORMAL PROPERTIES OF METARULES. Report-85-43, Center for the Study of Language and Information, Universidad de Stanford.
- Vargas, Celso (1986) "La aproximación chomskyana a la gramática y algunas de sus implicaciones filosóficas" en, REVISTA DE FILOSOFIA DE LA UNIVERSIDAD DE COSTA RICA (60) 183-193.
- Wasow, Thomas (1985) "Poscript" en Sells (1985) 193-205.