

**Una gramática de estructura de frases
generalizada (GPSG) de la oración
española: el orden de los
constituyentes mayores**

Ursula Klenk

Seminar für Romanische Philologie
Universität Göttingen

Una gramática de estructura de frases generalizada (GPSG) de la oración española: el orden de los constituyentes mayores

Ursula Klenk
Universidad de Göttingen, Rep. Fed. de Alemania

1. Aspectos generales de la GPSG

El orden de los constituyentes mayores de la oración española, a saber sujeto, verbo finito, objetos y complementos circunstanciales, es relativamente libre. En la gramática generativa transformacional los órdenes posibles pueden ser descritos por medio de reglas transformacionales, de manera que uno de estos órdenes se considera la estructura profunda y los otros se derivan de éste mediante transformaciones. La ventaja de tal sistema es que con las estructuras profundas se da un punto de partida para la interpretación semántica. Pero hay graves desventajas para el procesamiento del lenguaje natural, porque parseres para gramáticas transformacionales son demasiado complejos, si no imposibles. Por eso, en la lingüística computacional, desde sus principios, se han buscado modelos sintácticos mejor manejables.

La GPSG (véase Gazdar/Klein/Pullum/Sag 1985, abreviado en lo siguiente GKPS) es monostratal en que partiendo del concepto de la gramática libre de contexto, pero modificándolo bajo varios aspectos, describe estructuras superficiales sin intercalar una estructura profunda. Coincidiendo con las reglas sintácticas hay reglas interpretativas, que añaden representaciones semánticas a los árboles generados por la sintaxis.

Un concepto central de la GPSG es el de la sintaxis de dominancia inmediata (ID) y de precedencia lineal (LP), abreviada aquí sintaxis ID/LP. Mientras que reglas libres de contexto representan juntos la dominancia inmediata de constituyentes y su orden lineal, una sintaxis ID/LP describe estas relaciones por dos conjuntos diferentes de reglas, llamadas respectivamente reglas de ID y reglas de LP. Una regla de ID tiene la forma

$$A \rightarrow A_1, \dots, A_n$$

(A, A_1, \dots, A_n símbolos), lo que quiere decir que A domina directamente las A_i es ($1 \leq i \leq n$) y que se admite cada orden posible de éstas. A parte de esto, las sintaxis ID tienen la misma forma que las libres de contexto. Las reglas de LP tienen la forma

$$A < B$$

(A, B símbolos), lo que indica que en todas reglas de ID, en cuyo lado derecho aparecen A y B juntas, A debe preceder a B para que resulte una estructura admitida. De este modo, las reglas de LP funcionan como filtros aplicados a los órdenes generados por las reglas de ID (una formalización de la sintaxis ID/LP se da en Klenk 1985).

Para lenguas que tienen orden relativamente libre de los constituyentes, una descripción en el formalismo ID/LP parece atractiva, sobre todo si se considera que hay parseres eficientes para tales sistemas.

Consideramos como constituyentes mayores el verbo finito y sus actantes (sujeto, objetos, complementos circunstanciales obligatorios), el complemento predicativo así como los complementos circunstanciales que no son actante o sub-constituyente de un actante del verbo (en lo siguiente: complementos libres; a éstos cuentan también frases subordinadas, gerundios y construcciones con infinitivo y participio, si tienen función circunstancial). Son, del punto de vista de la gramática generativa clásica, los constituyentes inmediatamente dominados por la S y la VP más altas de un árbol.

La sintaxis presentada en lo siguiente es limitada a frases no coordinadas y no negadas. Se entiende como sintaxis "pura" de la cual se derivan las estructuras superficiales, no considerando por de pronto aspectos semánticos.

2. Oraciones principales sin complementos libres

En estas oraciones el verbo finito se encuentra en la primera o en la segunda posición, p.ej.

- (1) Llama la atención la falta de recursos financieros que se anuncian para la realización del proyecto.
- (2) María lee un periódico.

Salvo con pronombres personales objetos y reflexivos otras posiciones parecen ser excluidas. Estos pronombres los consideramos partes de un complejo verbal especial que no tratamos en este artículo.

Partiendo de la segmentación clásica de la oración en frase nominal (NP) sujeto y frase verbal (VP) se necesitan las reglas siguientes de ID para los casos con sujeto explícito. *

- (3) S \rightarrow NP, VP
 - VP \rightarrow V
 - VP \rightarrow V, NP
 - VP \rightarrow V, PP
 - VP \rightarrow V, NP, PP
 - VP \rightarrow V, PP, PP
 - VP \rightarrow COP, PRED
- (PP: frase preposicional,
COP: verbo copulativo,
PRED: complemento predicativo)

El problema es que con este sistema hasta tres actantes pueden preceder al verbo y que no hay reglas de LP que dan precisamente los órdenes correctos. Además el sujeto no puede colocarse entre el verbo y el objeto. Necesitamos todavía otras reglas de ID para las oraciones con frases verbales discontinuas. Las da la metaregla

- (4) VP \rightarrow V, X \implies S \rightarrow NP, V, X

* En la GPSG los símbolos de categorías se representan según las convenciones de X-bar-theory. Como en este artículo no es preciso tratar la noción de cabeza (head), utilizamos las notaciones usuales de la gramática generativa.

Aquí X es una variable que representa secuencias de constituyentes (posiblemente vacías). (4) aplicado a (3) genera las reglas de ID

- (5a) S → NP, V
 S → NP, V, NP
 S → NP, V, PP
 S → NP, V, NP, PP
 S → NP, V, PP, PP

Además se añade

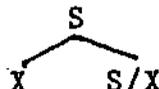
- (5b) S → NP, COP, PRED

Con las reglas de LP

- (6) V < X , COP < X , donde X ∈ {NP, PP, PRED}

que se aplican a (3) y (5), resulta que el verbo está solamente en primera o segunda posición. Pero como (3) y (5) juntos dan todas las cadenas sobre {V, NP, PP, COP, PRED} que (5) genera por sí solo, el sistema es sintácticamente ambiguo, y eso también si tenemos en cuenta las reglas de (6). Tal sistema no se desea, porque la ambigüedad es puramente sintáctica sin tener función alguna. Además, los casos donde un objeto precede al verbo, no se generan. Por todo eso proponemos suprimir la categoría VP y tomar como punto de partida solamente (5). Las funciones sintácticas (sujeto, objeto etc.) pueden ser representadas mediante sub-índices en los símbolos categoriales.

Pero ahora debemos establecer todavía reglas que generan el verbo en la posición segunda. Proponemos una solución utilizando las llamadas categorías con SLASH, como se propuso también para oraciones alemanas (Uszkoreit 1982, Russell 1985). Partimos de estructuras que tienen la forma



donde X es un constituyente diferente de V y S/X es el "resto de la oración". La notación A/B donde A y B son símbolos categoriales representa un constituyente de categoría A de la cual un constituyente de categoría B ha sido eliminado, dejando un hueco (gap) de categoría B. A (5) se añaden de este modo las reglas

- (7) S → NP, S/NP
 S → PP, S/PP
 S → PRED, S/PRED

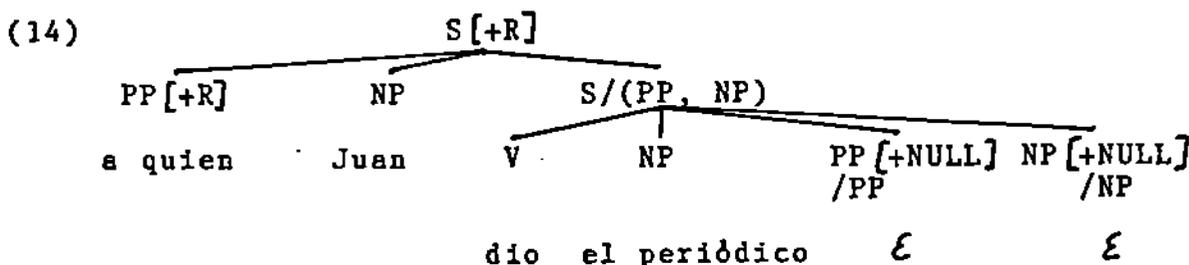
y a (6) la regla

- (8) X < S/X

A las S/X se aplican las reglas de ID que extienden S, salvo las de (7) que inducirían categorías con SLASHes iterados lo que GKPS excluye y de lo cual resultarían estructuras no correctas del español. Para obtener estructuras bien formadas, S/X debe tener un sub-constituyente con el rasgo +NULL (introducido por reglas), el cual genera la palabra vacía (ξ). +NULL implica /X que se traslada reiteradamente de un nudo al nudo padre hasta encontrar

- (13) $S[+R] \rightarrow X[+R], S/X$
 $S[+R] \rightarrow AV[+R], S$
 $S[+R] \rightarrow X[+R], Y, S/(X,Y), (AVC)$ } donde $X, Y \in \{NP, PP\}$
 y un símbolo sin +R
 tiene -R
- $X[+R] < Z[-R]$, donde $X \in \{NP, PP, AV\}$, $Z \in \{NP, PP, AVC, S, S/W\}$ y W es una secuencia de hasta 2 símbolos categoriales.

Según (13), la frase relativa de (12) recibe la estructura



5. Semántica

Hemos establecido nuestro sistema excluyendo el tratamiento de aspectos semánticos. Pero la admisión de sub-índices para las funciones gramaticales hace posible de introducir reglas interpretativas que estructuran las oraciones en predicados y argumentos, lo que puede ocurrir p. ej. en un nivel funcional semejante al de la gramática lexico-funcional. Pero eso no da una representación semántica completa. No se resuelve p.ej. el problema de representar sintácticamente los dominios de cuantificadores y de palabras de negación de tal manera que una representación semántica de estas estructuras se adjunte fácilmente. Eso es, sin embargo, un problema general de todas las sintaxis monostratales conocidas, no solamente de la GPSG. Pero eso ya es un otro tema.

Literatura:

- Gazdar, G. / Klein, E. / Pullum, G. / Sag, I. (1985): Generalized Phrase Structure Grammar. Basil Blackwell, Oxford.
- Klenk, U. (1985): ID/LP-Syntaxen für Wortstellung im Spanischen. In: Klenk (ed.), Kontextfreie Syntaxen und verwandte Systeme, Niemeyer, Tübingen.
- Russell, G. (1985): A GPS-Grammar for German Word Order. In Klenk (ed.), op. cit.
- Uszkoreit, H. (1982): German Word Order in GPSG. In: Flickinger, D. / Macken, M. / Wiegand, N. (eds.), Proceedings of the First West Coast Conference on Formal Linguistics. Stanford.