

## **Sobre constituyentes inmediatos y construcciones sintácticas. El análisis distribucionalista de la cláusula**

La corriente lingüística denominada distribucionalismo intentó construir una teoría capaz de dar cuenta de los fenómenos sintácticos basándose exclusivamente en el concepto de 'distribución', esto condujo a una descripción de la organización de las expresiones lingüísticas en términos de constitución, prescindiendo totalmente de nociones funcionales y conectivas. El examen de algunos aspectos de la práctica distribucionalista nos revela, sin embargo, que los verdaderos primitivos no son la distribución y la constitución, sino las llamadas "construcciones sintácticas", que, aunque aparentemente se establecen a partir de marcas de tipo gramatical ("taxemas"), de hecho se definen por el valor referencial correspondiente.

En esta nota pretendemos mostrar que el análisis en constituyentes que propone el distribucionalismo, al menos en lo que se refiere a la división binaria de la unidad cláusula, está basado en última instancia en consideraciones de contenido representativo muy semejantes a las que se manejan en algunos manuales de gramática a propósito de la bipartición tradicional en "sujeto" y "predicado". De esta manera se cuestiona la idea de que el distribucionalismo ofrece un sustento formal al análisis de la cláusula en dos, y sólo dos, constituyentes inmediatos.

De una manera esquemática podemos decir que el distribucionalismo trabaja sobre un corpus amplio de enunciados y a partir de la comparación de los mismos los segmenta en unidades que a su vez no pueden ser segmentadas: los morfemas. El paso siguiente consiste en agrupar estos morfemas en clases distribucionales, clases formadas por todos aquellos morfemas que pueden aparecer en el mismo contexto, siendo el "contexto" la serie de morfemas que preceden y/o siguen a un morfema dado.

Con estas premisas, si nosotros queremos describir un determinado morfema de un enunciado nos veremos obligados a tener en cuenta todos y cada uno de los restantes morfemas que constituyen dicho enunciado. En palabras de Pittman, "one is almost compelled to regard every morpheme in an utterance as pertinent to the description of every other morpheme" (Pittman 1948, p. 275).

Resulta obvio que con este planteamiento las descripciones gramaticales serían excesivamente complicadas y, probablemente, inadecuadas para dar cuenta de la organización de las secuencias. Surge entonces el principio de los constituyentes inmediatos como un método para simplificar la descripción de las expresiones lingüísticas, ya que el análisis en constituyentes reduce drásticamente el contexto o entorno lineal que se ha de considerar en cada caso y, en consecuencia, facilita el establecimiento de clases distribucionales<sup>1</sup>.

El primer autor que formuló la teoría de los constituyentes inmediatos fue Bloomfield:

(...) it appears that every complex form is entirely made up (...) of morphemes. The number of these *ultimate constituents* may run very high. The form *Poor John ran away* contains five morphemes: *poor*, *John*, *ran*, *a-* (...), and *way*. However, the structure of complex forms is by no means as simple as this; we could not understand the forms of a language if we merely reduced all the complex forms to their ultimate constituents. Any English-speaking person who concerns himself with this matter, is sure to tell us that the *immediate constituents* of *Poor John ran away* are the two forms *poor John* and *ran away*, that each of these is, in turn, a complex form; that the immediate constituents of *ran away* are *ran*, a morpheme, and *away*, a complex form, whose constituents are the morphemes *a-* and *way*; and that the constituents of *poor John* are the morphemes *poor* and *John* (Bloomfield 1933, p. 161).

Sin embargo, esta formulación del principio de los constituyentes inmediatos no proporciona pautas que permitan analizar correctamente cualquier secuencia. Es más, la única referencia en ese sentido remite a la intuición del hablante, en contradicción con los presupuestos generales de la teoría distribucionalista, que se declara ajena a consideraciones de este tipo.

Para estudiar los criterios que el distribucionalismo maneja en el análisis en constituyentes nos centraremos en la exposición que hace Wells en su artículo "Immediate Constituents", que, como el propio autor señala, tiene como objetivo "to replace by a unified, systematic theory the heterogeneous and incomplete methods hitherto offered for determining IMMEDIATE CONSTITUENTS" (Wells 1947, p. 81).

Según Wells, el criterio fundamental para determinar los constituyentes de una secuencia es la noción de "expansión", basada en el hecho de que "sometimes two sequences occur in the same environments though they have different internal structures" (ibíd., p. 83). Así, dadas dos secuencias A y B pertenecientes a la misma clase distribucional, A será expansión de B -que será "modelo" de A- si está formada por lo menos por el mismo número de morfemas y no pertenece a la

---

<sup>1</sup> "a good analysis in terms of ICs usually reduces the total possible environmental factors of a given morpheme or sequence of morphemes to one: in other words, it states that the only pertinent environment of a given IC is its concomitant (the other IC)" (Pittman 1948, p. 275).

misma "clase secuencial"<sup>2</sup>, o, lo que es lo mismo, no presenta la misma estructura interna que B.

A la hora de efectuar el análisis se nos pueden plantear dos situaciones:

(i) que la secuencia conste sólo de dos morfemas, con lo cual la segmentación no presentará dificultades;

(ii) que la secuencia conste de más de dos morfemas. En tal caso hay que recurrir al concepto de expansión: para efectuar el análisis de una secuencia A de este tipo hemos de encontrar en el corpus una secuencia B de la que A sea expansión y cuyo análisis no ofrezca dudas; el caso más sencillo es que B conste sólo de dos morfemas. Naturalmente, cada uno de los constituyentes de B ha de ser modelo de uno de los dos segmentos en que resultará dividida A.

Wells da un ejemplo del proceso de análisis (cfr. *ibíd.*, pp. 83 y ss.). Para la secuencia *The king of England opened Parliament* toma como modelo *John worked*, siendo *the king of England* expansión de *John* y *opened Parliament* expansión de *worked*: la primera segmentación da como resultado *The king of England / opened Parliament*. Para establecer los constituyentes inmediatos de *the king of England* hay que buscar una secuencia con dos morfemas (como señala Wells, "*John is no help here because it is a single morpheme*" ;*ibíd.*, p. 83); Wells toma como modelo *the king*, lo que le permite dividir *the / king of England*, puesto que *king of England* es expansión de *king*. Por su parte *opened Parliament* se divide sin problemas en *opened / Parliament*. Podría objetarse en este último caso, en vista de que la secuencia está realmente formada por tres morfemas (*open*, *-ed* y *Parliament*), que sería posible un análisis *open / -ed Parliament*, pero esta propuesta queda excluida por "the principle that word divisions should be respected" (*ibíd.*, p. 84)<sup>3</sup>.

Como, desafortunadamente, no hay en el artículo de Wells análisis en constituyentes inmediatos para cláusulas más complejas que la anterior, recurriremos a los ejemplos que maneja Hockett (1958, p. 178; vers. esp., pp. 172 y 179). De acuerdo con las representaciones de Hockett, para una cláusula como *Juan siempre bebe agua* el análisis sería:

<sup>2</sup> La definición de clase secuencial ("sequence-class") que da Wells es la siguiente:

"Given a sequence S, a sequence-class to which S belongs is defined as the class of all sequences whose first morpheme belongs to the same morpheme-class as the first morpheme of S, whose second morpheme belongs to the same class as the second morpheme of S, and so on" (Wells 1947, pág. 82).

<sup>3</sup> Más adelante establece matizaciones a este principio:

"Instead of proclaiming, therefore, that every word in every language must be a constituent of any sequence in which it occurs as a part, the most we may say is that every word should be so regarded unless it engenders a conflict or complication in the description of the language" (Wells 1947, pp. 101-102).

Juan	siempre	bebe	agua

Veamos ahora si es posible, manejando el principio de la expansión, justificar otra estructura constitutiva. Admitamos provisionalmente la segmentación *Juan / siempre bebe agua* y quedémonos con el segmento *siempre bebe agua*.

Para este segmento un distribucionalista propondría un análisis *siempre / bebe agua*, basándose en la existencia en el corpus de secuencias como *siempre bebe* (*Juan siempre bebe*), de manera que *bebe agua* sería expansión de *bebe* (cfr. Wells *ibíd.*, p. 83: "*opened Parliament* is an expansion of *worked*"). Pero es igualmente posible una segmentación del tipo *siempre bebe / agua* si nos basamos en la existencia de secuencias como *bebe agua* (*Juan bebe agua*), es decir, considerando *siempre bebe* expansión de *bebe*.

Para resolver disyuntivas como esta, Wells, consciente de que el principio de la expansión no se cumple de manera absoluta en prácticamente ningún caso<sup>4</sup>, sostiene que entre varios posible análisis hemos de elegir aquel en que los constituyentes resultantes puedan aparecer en un mayor número de contextos diferentes y, además, puedan ser reemplazados por segmentos de clases secuenciales lo más variadas posible (cfr. *ibíd.*, pp. 87-88). Pero si intentamos aplicar este criterio a nuestro análisis no encontramos diferencias apreciables entre ambas alternativas.

Ante tal situación no parece existir otra solución que la segmentación en tres constituyentes: *siempre / bebe / agua*, y ello sin salirnos del marco de Wells. La condición que impone para reconocer más de dos constituyentes inmediatos en un segmento es la siguiente:

GIVEN A CONSTITUTE CONSISTING OF THREE CONTINUOUS SEQUENCES A, B AND C, THEN, IF NO REASON CAN BE FOUND FOR ANALYZING IT AS AB/C RATHER THAN A/BC, OR AS A/BC RATHER THAN AB/C, IT IS TO BE ANALYZED INTO THREE CORRELATIVE ICS, A/B/C. Similarly, four ICs may be recognized when no analysis into three ICs is recommended, and so on (Wells 1947, p. 103).

Volvamos ahora sobre la primera segmentación de la cláusula. Según los ejemplos que manejan los distribucionalistas, una cláusula como *Juan lee manuales de botánica* se segmentará *Juan / lee manuales de botánica* tomando como modelo *Juan lee*. Aunque pueda parecer descabellado, existe la posibilidad de anali-

<sup>4</sup> "it is rare, if it happens at all, that one sequence is an absolute equivalent of a sequence belonging to an entirely different sequence-class. On this account our interest centers in non-absolute expansions, and the problem is to find expansions that approximate as closely as possible to being absolute" (Wells 1947, p. 87).

zar binariamente esa secuencia en *Juan lee / manuales de botánica*, considerando *Juan lee* expansión de *lee*, a partir de secuencias como *lee manuales de botánica*. Se podría objetar que *Juan lee* no es posible en todos los contextos en que *lee* puede aparecer (*María lee novelas policiacas / \*María Juan lee novelas policiacas*), pero la propuesta distribucionalista presenta también problemas de este tipo: *lee manuales de botánica* no se da en los mismo contextos que *lee* (*María lee novelas policiacas / \*María lee manuales de botánica libros policiacos*)<sup>5</sup>.

De los ejemplos anteriores se deduce que los mismos criterios manejados para justificar el análisis distribucionalista pueden servir también para apoyar otras propuestas de estructura jerárquica de los constituyentes. Esto nos hace pensar que la concepción que el distribucionalismo tiene de las relaciones constitutivas no puede ser explicada simplemente a partir de los criterios que se emplean en la segmentación de las secuencias, sino que tal concepción está determinada por otro tipo de factores.

En mi opinión, la clave se encuentra en afirmaciones como las siguientes:

When a word contains three or more morphemes, it is usually necessary to analyze it into two and only two IMMEDIATE CONSTITUENTS, either or both which may be susceptible of further analysis. Only in this way is it possible to reveal the morphological structure of the word; a mere listing of the morphemes, with no indication of their ordering, would tell us nothing about the constructions involved (Bloch-Trager 1942, p. 67, subr. mío).

Whenever two (or, rarely, more) forms are spoken together, as constituents of a complex form, the grammatical features by which they are combined, make up a construction (Bloomfield 1933, p. 169).

Parece, pues, que lo que subyace al fenómeno de la constitución es el concepto de "construcción"; las relaciones constitutivas dejan de ser una noción auto-suficiente ya que su descripción en las secuencias concretas precisa de la definición previa de las construcciones que las determinan.

En el distribucionalismo las construcciones se identifican y clasifican en diversos tipos a partir de cuatro especies de rasgos<sup>6</sup>:

The free forms (words and phrases) of a language appear in larger free forms (phrases), arranged by taxemes of modulation, phonetic modification, selection, and order. Any meaningful, recurrent set of such taxemes is a *syntactic construction* (ibíd., p. 184).

<sup>5</sup> Es lo mismo que sucede con el ejemplo que maneja Wells, pues no siempre que es posible *worked* puede aparecer *opened Parliament*:

*He worked his promotion / \*He opened Parliament his promotion*

*This land has not been worked for many years / \*This land has not been opened Parliament for many years*

*She worked the problem out / \*She opened Parliament the problem out*

*He worked on the car for 3 days / \*He opened Parliament on the car for 3 days.*

<sup>6</sup> Estos rasgos reciben diferentes denominaciones: "taxemes" en Bloomfield (1933, pp. 163 y ss. y 184), "features of arrangement" en Bloch-Trager (1942, p. 72), "kinds of syntactical linkage" en Hockett (1958, p. 214).

En las lenguas de nuestro entorno son los taxemas de selección, junto con los de orden, los que resultan más rentables para el establecimiento de las distintas construcciones sintácticas. Los rasgos de selección se refieren a la clase distribucional de los constituyentes. Así, *drink milk* y *fresh milk* presentan construcciones diferentes -la primera es una construcción "action-goal" y la segunda "attribute-head"- debido a que *drink* y *fresh* pertenecen a distintas clases distribucionales: *drink* a la de los verbos transitivos y *fresh* a la de los adjetivos (cfr. Bloomfield 1933, p. 165).

Entre los taxemas de selección, Bloomfield distingue tres subtipos: la concordancia ("concord or congruence"), el régimen ("government") y la referencia cruzada ("cross-reference") (cfr. *ibíd.*, pp. 191-193). A partir de estos rasgos el distribucionalismo establece las construcciones sintácticas propias de cada lengua. Concretamente en inglés, los autores de esta corriente se refieren a construcciones como:

"actor-action"	<i>John / ran away</i>
"action-goal"	<i>drink / milk</i>
"attribute-head"	<i>fresh / milk</i>
"relation-axis"	<i>beside / John</i>

Lo que se consigue mediante el análisis de una secuencia en sus constituyentes sucesivos es aislar los segmentos que tienen alguna "función"<sup>7</sup> en alguna de las construcciones definidas por la gramática. El principio de la expansión cumple aquí un papel fundamental. Si se asignan a la construcción actor-acción expresiones como *John worked*, donde *John* "llena" la posición de (o tiene la función de) 'actor' y *worked* la de 'acción', y resulta que *the king of England* es expansión de *John*, por tanto puede aparecer en las mismas posiciones que *John*, y *opened Parliament* es expansión de *worked* porque pertenece a su misma clase distribucional, habrá que concluir que *The king of England opened Parliament* es también una construcción actor-acción cuyos constituyentes inmediatos *the king of England* y *opened Parliament* funcionan respectivamente como actor y como acción.

De todo lo anterior se extrae que el concepto de construcción determina el de constitución, lo cual es perfectamente lógico desde una perspectiva distribucional, teniendo en cuenta que las funciones de los elementos se identifican con sus posiciones, con su distribución, y que el análisis en constituyentes inmediatos proporciona un marco adecuado para describir la distribución de las unidades, ya que reduce el contexto pertinente para la descripción de una unidad en una secuencia concreta al segmento de la misma secuencia que resulta ser co-constituyente de la unidad en cuestión.

<sup>7</sup> El término "función" tiene aquí el valor que le atribuye Bloomfield: "The positions in which a form can appear are its functions or, collectively, its function" (Bloomfield 1933, p. 185).

Hay, sin embargo, razones para dudar de que las construcciones se definan realmente a partir de los taxemas y que las funciones puedan identificarse sólo con criterios distribucionales.

La construcción actor-acción, que determina la estructura bipartita de la cláusula, se justifica para el inglés por taxemas como el orden, la clase distribucional de las unidades, la concordancia, etc., pero, contradictoriamente, se define por su significado, en términos similares a los que podemos encontrar en la gramática tradicional:

The meaning of the construction is roughly this, that whatever is named by the substantive expression is an actor that performs the action named by the finite verb expression (Bloomfield 1933, p. 183)

Sólo con una definición de este tipo puede afirmarse la existencia de la construcción actor-acción en otras lenguas, como hace Bloomfield para el francés, alemán, español, italiano, ruso (cfr. *ibíd.*, pp. 172-173), e incluso chino (*ibíd.*, p. 199). Si, efectivamente, las construcciones se definiesen por series recurrentes de taxemas, sería bastante improbable que dos lenguas distintas presentasen las mismas construcciones: en inglés, por ejemplo, el orden es pertinente en la construcción actor-acción, mientras que en español y en latín apenas tiene relevancia en el análisis bipartito de la cláusula.

Asimismo, de este modo, a pesar de que teóricamente las "funciones" del distribucionalismo se identifican con las "posiciones", al definir las construcciones por su contenido representativo, las funciones que se manejan tienen también un carácter semántico-referencial (por ejemplo, el "actor" es "that performs the action named by the finite verb expression"; Bloomfield 1933, p. 183).

Con esta perspectiva, las diferencias en la concepción de la lengua y en los métodos de análisis que existen entre el distribucionalismo, una corriente que nació como reacción contra el mentalismo imperante en gramática a principios de siglo, y esta misma gramática fuertemente condicionada por presupuestos lógico-semánticos deben ser matizadas. En lo que se refiere al aspecto concreto que hemos examinado en estas páginas, el arraigo del planteamiento tradicional de la estructura de la cláusula en el análisis binario distribucionalista es manifiesto. Y no resulta difícil encontrar pruebas de ello:

The common name for such bipartite favorite sentence forms [v. gr. la construcción actor-acción] is predications. In a predication the more object-like component is called subject, the other part the predicate (Bloomfield 1933, p. 173).

[en chino] The favorite sentence-construction is one of *subject* and *predicate*, much like the English actor-action construction (*ibíd.*, p. 199).

## REFERENCIAS BIBLIOGRAFICAS

- Bloch, E. y G. L. Trager (1942), *Outline of Linguistic Analysis*, Linguistic Society of America, Waverly Press, Baltimore, 1942.
- Bloomfield, L. (1933), *Language*, Holt, Rinehart and Winston, Nueva York, 1933.
- Hockett, C. F. (1958), *A Course in Modern Linguistics*, McMillan, Nueva York, 1958. Vers. esp. de E. Gregores y J. A. Suárez, *Curso de lingüística moderna*, EUDEBA, Buenos Aires, 1979<sup>4</sup>.
- Pittman, R. C. (1948), "Nuclear structures in Linguistics", *Language*, 24, 1948-49, pp. 287-292. Cito por la reed. en M. Joos (ed.), *Readings in Linguistics I. The Development of Descriptive Linguistics in America 1925-56*, The Univ. of Chicago Press, Chicago, 1957, 1966<sup>4</sup>, pp. 275-278.
- Wells, R. S. (1947), "Immediate constituents", *Language*, 23, 1947, pp. 81-117.

VICTORIA VAZQUEZ

Universidade de Santiago de Compostela