PROPUESTA DE ETIQUETADO GRAMATICAL PARA LOS NUMERALES

A GRAMMAR LABELING PROPOSAL FOR NUMERALS

Aurora Martín de Santa Olalla Sánchez
Universidad Complutense

Recibido: 3 de marzo de 2019. Aceptado: 7 de septiembre de 2019.

Resumen: Esta propuesta se basa en la necesidad de contar con cuatro elementos mínimos para la construcción de una gramática de los numerales: una *tabla léxica,* donde estén los nombres de los números de una lengua que no puedan obtenerse mediante reglas, o sea los básicos, una *tabla de elementos vinculantes,* donde estén las partículas, preposiciones o conjunciones, que unan los elementos básicos para formar los derivados, los nombres compuestos de los números, una *tabla de rasgos tipológicos,* en la línea de las anteriores, si bien más amplia y explícita, a la que se dirija el puntero del programa de formalización para saber cuál es el tipo de la combinación, y unas *reglas combinatorias.*
Palabras clave: computación, corpus, etiqueta, léxico, numeral, regla, sintaxis, tipología.

Abstract: This proposal is based on the need to have four minimum elements for the construction of a grammar of the numerals: a lexical table, with the names of the numbers of a language that cannot be obtained through rules, that is, the basic ones, a table of binding elements, with the particles, prepositions or conjunctions that join the basic elements to form the derivatives, the names composed of the numbers, a table of typological features, in the line of the previous ones, although more extensive and explicit , to which the pointer of the formalization program is directed to know what the type of the combination is, and some combinatorial rules.
Key words: computation, corpus, label, lexicon, numeral, rule, syntax, tag, typology.

Hay notables intentos de formalización de la gramática de los numerales (Katwijk: 1965, 1968; Brainerd: 1966, 1968; Brandt Corstius: 1965, 1968; Hurford: 1975, 1980, 1987; Corbett: 1978; Bonet: 1981), donde pueden encontrarse discusiones y planteamientos previos. Esta propuesta se basa en la necesidad de contar con cuatro elementos mínimos para la construcción de esa gramática: una *tabla léxica,* donde estén los nombres de los números de una lengua que no puedan obtenerse mediante reglas, o sea los básicos, una *tabla de elementos vinculantes,* donde estén las partículas, preposiciones o conjunciones, que unan los elementos básicos para formar los derivados, los nombres compuestos de los números, una *tabla de rasgos tipológicos,* en la línea de las anteriores, si bien más amplia y explícita, a la que se dirija el puntero del programa de formalización para saber cuál es el tipo de la combinación, y unas *reglas combinatorias.* Estas reglas pueden ser de tipo booleano y limitarse a comprobar en las tablas las condiciones veritativas, para ir eligiendo los elementos que entran en una combinación; pero también pueden ser declarativas, pues en principio no parece haber nada que se oponga al desarrollo de un paquete de programas en PROLOG, por ejemplo, que establezcan las uniones y exclusiones posibles entre los elementos de las tablas. La presencia de fenómenos de fusión, por último, puede hacer aconsejable incluir, en algunas lenguas, un procedimiento de refinado del producto obtenido mediante la aplicación de las reglas combinatorias.

Es apremiante la necesidad de contar con un sistema estándar de etiquetas, membretes o marcas (*tags*), para la recuperación automática de la información. La propuesta que se presenta parte de Aurora Martín de Santa Olalla (1994) y permite un etiquetado coherente, manual, semiautomático o automático, de archivos digitales, diseñado especialmente para corpus de referencia. Para los conceptos de «archivos digitales», «corpus de referencia», «SGML», «DTD», etc. remitimos a Marcos Marín (1994).

La categoría de los numerales es seguramente la que presenta mayores fluctuaciones o vacilaciones de sus

miembros entre distintas clases y entre los distintos tipos que integran su misma clase, por lo que es necesario utilizar membretes dobles en los sistemas de etiquetado.

La semejanza formal y funcional entre los numerales y las clases de sustantivo y adjetivo (y pronombre, en algunos casos) queda suficientemente reflejada mediante el rasgo 'función'. Entre los numerales, sin embargo, existen muchas formas que, han pasado a adquirir un significado algo distinto del numeral original que justifica su codificación como miembro de la categoría que lo define en términos funcionales y morfosintácticos. El ejemplo más claro en este sentido lo constituye el ordinal primero/primera en algunas expresiones lexicalizadas donde no significa el primero en orden en oposición a un segundo o tercero:

Artículos de primera necesidad
Causa primera
Primera enseñanza

En estos casos se codificarán como adjetivos.

Finalmente, entre los ejemplos de doble asignación de membrete de más difícil desambiguación figura la que se produce para la forma un/una entre los membretes correspondientes a la categoría artículo indefinido y los correspondientes a la categoría numeral cardinal y consecuentemente la que se produce entre el pronombre indefinido y el numeral cardinal.

Ya en el *Esbozo*, la RAE señalaba que *un[{o,a}]* es numeral cuando se contrapone a otro de mayor valor numérico explícito o implícito. En la forma plural *unos/unas* es exclusivamente indefinido (artículo o pronombre) excepto cuando denota el guarismo con que se representa la unidad, cuando significa identidad o semejanza y cuando se refiere a nombres que solo se emplean o que en determinadas acepciones solo se emplean en plural.

Se recogen a continuación las etiquetas propuestas en Aurora Martín de Santa Olalla (1994) dentro del esquema

global de codificación morfosintáctica del español para etiquetado automático o semiautomático de corpus de referencia.

Para una mejor comprensión se da tan sólo una mínima tabla de equivalencias; pero es preciso estudiar las codificaciones de rasgos para comprender todo el proceso (Martín de Santa Olalla: 1994). Las codificaciones plenas figuran a la izquierda del guión. A la derecha de éste se sitúan las abreviadas. **NUC1** es abreviatura de **NUCMSAJ**, en el primer ejemplo de numeral cardinal, *infra*. No se olvide que el mismo código, p. ej. **F**, puede significar rasgos diferentes según su posición en el conjunto.

AJ adjetivo
AV adverbio
C cardinal (en NUC)
F fraccionario (en NUF), tb. femenino
M multiplicativo (en NUM), tb. masculino
N nombre
NU numeral
P plural
R ordinal (en NUR), tb. pronombre
S singular

NUMERALES CARDINALES

NUCMSAJ-NUC1 Un caballo (en oposición a dos/tres/cuatro/caballos)

NUCMSN-NUC2 El seis de infantería ligera (ejemplo de Bello)

Ha salido premiado el 15.500
Desde su planificación en 1750

NUCMSR-NUC3 Dame uno

NUCFSAJ-NUC4 Una tarta (en oposición a dos/tres/cuatro tartas)

NUCFSN-NUC5 La nueve de este libro (la lección nueve)

NUCFSR-NUC6 Dame una

NUC6PR-NUC7 Dame tres (Si apareciera sin contexto que nos permitiera identificar el género)

NUCMPAJ-NUC8 Doscientos coches
Unos tirantes (= un par de tirantes, en oposición a dos/tres /cuatro ... pares de tirantes)

NUCMPN-NUC9 Los ochos y los nueves

NUCMPR-NUC10 Dame doscientos
¿Quieres libros? Si dame tres

NUCFPAJ-NUC11 Doscientas casas
Unas gafas (= un par de gafas, en oposición a dos/tres/cuatro ... pares de gafas)

NUCFPN-NUC12 Las nueves de estos libros (las lecciones nueve)

NUCFPR-NUC13 Dame doscientas
¿Quieres calculadoras? Si dame tres

NUMERALES ORDINALES

NURMSAJ-NUR1 El primer niño de la clase

NURMSN-NUR2 El primero de la clase

NURFSAJ-NUR3 La primera niña de la clase

NURFSN-NUR4 La primera de la clase
NURMPAJ-NUR5 Los primeros niños de la clase

NURMPN-NUR6 Los primeros

NURFPAJ-NUR7 Las primeras niñas de la clase

NURFPN-NUR8 Las primeras

NU00AV-NUR9 Primero llegó María

NUMERALES FRACCIONARIOS

NUFMSAJ-NUF1 Medio melocotón

NUFMSN-NUF2 El diez por ciento

NUFFSAJ-NUF3 Media naranja

NUFFSN-NUF4 La mitad de la tarta

NUFMPAJ-NUF5 Dos medios bollos

NUFMPN-NUF6 Los diez por ciento restantes

NUFFPAJ-NUF7 Dos medias naranjas

NUFFPN-NUF8 Las mitades de la tarta

NUMERALES MULTIPLICATIVOS

NUMMSAJ-NUM1 parto quíntuple

NUM6SAJ-NUM2 componente doble

NUMMSN-NUM3 El doble de ración

NUMFSAJ-NUM4 imagen doble

NUMFSN-NUM5 una dupla

NUMMPAJ-NUM6 partos quíntuples

NUMMPN-NUM7 Los dobles

NUMFPAJ-NUM8 Las dobles parejas

NUMFPN-NUM9 Las duplas

EL NUMERAL (%FICHERO FSD)

<range feature=numeral>
<descr> este rasgo especifica la pertenencia a la categoría gra matical numeral y la diferenciación en tipos o clases.
</descr>
<f.structure=T>
</range>
<range feature=tipo>
<descr> este rasgo especifica los distintos tipos o clases de numerales en español. De acuerdo con ellos se establecerán los distintos rasgos aplicables a cada uno de los tipos o clases.
</descr>
<f.structure=C>

<f. structure=O><f.structure=F>
<f. structure=M>
</range>
<range feature=cardinal>
<descr>describe los rasgos y valores del numeral cardinal en español
</descr>
<f.structure name=G><f.structure name=N>
<f.structure=U>
</range>
<range feature=género>
<descr>da cuenta de los distintos valores del rasgo género para el numeral cardinal.Codificaremos como “invariante o cualquiera” (=6) los numerales cardinales con fun ción de pronombre en aquellos con textos en los que no es posible iden tificar el género de estos.
</descr>
<atom name=M><atom name=F>
<atom name=6>
</range>
<range feature=número>
<descr>da cuenta de los distintos valores del rasgo número para el numeral cardinal.
</descr>
<atom name=S><atom name=P>
</range>
<range feature=función>
<descr>da cuenta de los diferentes comportamientos funcionales en términos de categorías gramaticales que pueden adoptar los numerales cardinales.
</descr>
<atom name=N><atom name=AJ>
<atom name=P>
</range>
<range feature=ordinal>
<descr>describe los rasgos y valores de los numerales ordinales.
</descr>
<f.structure
name=G><f.structure=N><f.structure=U>
</f.structure>
<range feature=género>
<descr>da cuenta de los distintos valores para el rasgo género de los numerales ordinales.

</descr>
<atom name=M><atom name=F>
</range>
<range feature=número>
<descr>da cuenta de los distintos valores para el rasgo número de los numerales ordinales.
</descr>
<atom name=S><atom name=P>
</range>
<range feature=función>
<descr>da cuenta de los diferentes comportamientos funcionales en términos de categorías gramaticales que pueden adoptar los numerales ordinales.
</descr>
<atom name=N><atom name=AJ>
<atom name=AV>
</range>
<range feature=fraccionario>
<descr>da cuenta de los distintos rasgos y valores del numeral fraccionario.
</descr>
<f.structure=G><f.structure=N><f.structure=U>
</range>
<range feature=género>
<descr>da cuenta del valor género para el numeral fraccionario.
</descr>
<atom name=M><atom name=F>
<range feature=número>
<descr>da cuenta de los distintos valores para el rasgo número de los numerales fraccionarios.
</descr>
<atom name=S><atom name=P>
</range>
<range feature=función>
<descr> da cuenta de los diferentes comportamientos funcionales en términos de categorías gramaticales que pueden adoptar los numerales fraccionarios.
</descr>
<atom name=N><atom name=AJ>
</range>
<range feature=multiplicativo>
<descr>da cuenta de los distintos rasgos y valores que se dan en los numerales multipli cativos.

</descr>
<f.structure name=G>
<f.structure name=N><f.structure=U>
</range>
<range feature=género>
<descr>da cuenta de los distintos valores para el rasgo género del cardinal multiplicativo. Codificaremos como "invariante o cualquiera" (=6) los numerales multiplicativos ("doble", "triple", ...) invariantes en la expresión del género.
</descr>
<atom name=M><atom name=F>
<atom name=6>
</range>
<range feature=número>
<descr>da cuenta de los distintos valores del rasgo número de los cardinales multiplicativos.
</descr>
<atom name=S><atom name=P>
</range>
<range feature=función>
<descr>da cuenta de los diferentes comportamientos funcionales en términos de categorías gramaticales que pueden adoptar los numerales multiplicativos.
</descr>
<atom name=N><atom name=AJ>
</range>

RASGOS DE LA CATEGORÍA NUMERAL

<! ENTITY T-C "<feature
name=TIPO><atomic>CARDINAL</atomic>
</feature>">

<! ENTITY T-R "<feature
name=TIPO><atomic>ORDINAL</atomic>
</feature>">

<! ENTITY T-F "<feature
name=TIPO><atomic>FRACCIONARIO</atomic>
</feature>">

<! ENTITY T-M "<feature
name=TIPO><atomic>MULTIPLICATIVO</atomic>
</feature>">

<! ENTITY U-AJ“<feature
name=FUNCIÓN><atomic>ADJETIVO</atomic>
</feature>”>

<! ENTITY U-N “<feature
name=FUNCIÓN><atomic>NOMBRE</atomic>
</feature>”>

<! ENTITY U-R “<feature
name=FUNCIÓN><atomic>PRONOMBRE</atomic>
</feature>”>

<! ENTITY U-AV“<feature
name=FUNCIÓN><atomic>ADVERBIO</atomic>
</feature>”>

EL NUMERAL (%DEFINICIÓN DE ENTIDADES)

<! ENTITY NUCMSAJ “<f.struct bar=1>&C-NU; T-C; G-M; N-S; U-AJ;</f.struct>”>

<! ENTITY NUCMSN “<f.struct bar=1>&C-NU; T-C; G-M; N-S; U-N;</f.struct>”>

<! ENTITY NUCMSR “<f.struct bar=1>&C-NU; T-C; G-M; N-S; U-R;</f.struct>”>

<! ENTITY NUCFSAJ “<f.struct bar=1>&C-NU; T-C; G-F; N-S; U-AJ;</f.struct>”>

<! ENTITY NUCFSN “<f.struct bar=1>&C-NU; T-C; G-F; N-S; U-N;</f.struct>”>

<! ENTITY NUCFSR “<f.struct bar=1>&C-NU; T-C; G-F; N-S; U-R;</f.struct>”>

<! ENTITY NUCMPAJ “<f.struct bar=1>&C-NU; T-C; G-M; N-P; U-AJ;</f.struct>”>

<! ENTITY NUCMPN “<f.struct bar=1>&C-NU; T-C; G-M; N-P; U-N;</f.struct>”>

<! ENTITY NUCMPR “<f.struct bar=1>&C-NU; T-C; G-M; N-P; U-R;</f.struct>”>

<! ENTITY NUCFPAJ “<f.struct bar=1>&C-NU; T-C; G-F; N-P; U-AJ;</f.struct>”>

<! ENTITY NUCFPN “<f.struct bar=1>&C-NU; T-C; G-F; N-P; U-N;</f.struct>”>

<! ENTITY NUCFPR "<f.struct bar=1>&C-NU; T-C; G-F; N-P; U-R;</f.struct>">

NUMERALES ORDINALES

<! ENTITY NURMSAJ "<f.struct bar=1>&C-NU; T-R; G-M; N-S; U-AJ;</f.struct>">

<! ENTITY NURMSN "<f.struct bar=1>&C-NU; T-R; G-M; N-S; U-N;</f.struct>">

<! ENTITY NURFSAJ "<f.struct bar=1>&C-NU; T-R; G-F; N-S; U-AJ;</f.struct>">

<! ENTITY NURFSN "<f.struct bar=1>&C-NU; T-R; G-F; N-S; U-N;</f.struct>">

<! ENTITY NURMPAJ "<f.struct bar=1>&C-NU; T-R; G-M; N-P; U-AJ;</f.struct>">

<! ENTITY NURMPN "<f.struct bar=1>&C-NU; T-R; G-M; N-P; U-N;</f.struct>">

<! ENTITY NU00AV "<f.struct bar=1>&C-NU; T-R; G-0; N-0; U-AV;</f.struct>">

NUMERALES FRACCIONARIOS

<! ENTITY NUFMSAJ "<f.struct bar=1>&C-NU; T-F; G-M; N-S; U-AJ; </f.struct>">

<! ENTITY NUFMSN "<f.struct bar=1>&C-NU; T-F; G-M; N-S; U-N; </f.struct>">

<! ENTITY NUFFSAJ "<f.struct bar=1>&C-NU; T-F; G-F; N-S; U-AJ; </f.struct>">

<! ENTITY NUFFSN "<f.struct bar=1>&C-NU; T-F; G-F; N-S; U-N; </f.struct>">

<! ENTITY NUFMPAJ "<f.struct bar=1>&C-NU; T-F; G-M; N-P; U-AJ; </f.struct>">

<! ENTITY NUFMPN "<f.struct bar=1>&C-NU; T-F; G-M; N-P; U-N; </f.struct>">

<! ENTITY NUFFPAJ "<f.struct bar=1>&C-NU; T-F; G-F; N-P; U-AJ; </f.struct>">

<! ENTITY NUFFPN "<f.struct bar=1>&C-NU; T-F; G-F; N-P; U-N; </f.struct>">

NUMERALES MULTIPLICATIVOS

<! ENTITY NUMMSAJ "<f.struct bar=1>&C-NU; T-M; G-M; N-S; U-AJ;</f.struct>">

<! ENTITY NUMMSN "<f.struct bar=1>&C-NU; T-M; G-M; N-S; U-N;</f.struct>">

<! ENTITY NUM6SAJ "<f.struct bar=1>&C-NU; T-M; G-6; N-S; U-AJ;</f.struct>">

<! ENTITY NUMFSAJ "<f.struct bar=1>&C-NU; T-M; G-F; N-S; U-AJ;</f.struct>">

<! ENTITY NUMFSN "<f.struct bar=1>&C-NU; T-M; G-F; N-S; U-N;</f.struct>">

<! ENTITY NUMMPAJ "<f.struct bar=1>&C-NU; T-M; G-M; N-P; U-AJ;</f.struct>">

<! ENTITY NUMMPN "<f.struct bar=1>&C-NU; T-M; G-M; N-P; U-N;</f.struct>">

<! ENTITY NUMFPAJ "<f.struct bar=1>&C-NU; T-M; G-F; N-P; U-AJ;</f.struct>">

<! ENTITY NUMFPN "<f.struct bar=1>&C-NU; T-M; G-F; N-P; U-N;</f.struct>">

BIBLIOGRAFÍA

BONET, Sebastià 1981 «Els numerals catalans,» *Anuario de Filología,* VII, Universidad de Barcelona, 509-528.

BRANDT Corstius, H. 1965 «Automatic Translation of Numbers into Dutch,» *Foundations of Language,* 1, 59-62.

BRANDT Cortius, H. (ed.) 1968 *Grammars for number names.* Dordrecht: Reidel.

BRAINERD, B. 1966a «Grammars for Number Names,» *Foundations of Language,* 2, 109- 133. (Versión revisada, Brainerd: 1968).

BRAINERD, B. 1966b «Two Grammars for Chinese Number Names,» *Canadian Journal of Linguistics,* 12, 33-51.

BRAINERD, B. 1967 «A Transformational-Generative Grammar for Rumanian Numerical Expressions», *Cahiers de Linguistique Théorique et Apliquée,* 4, 35-45. (reimp. Brandt Corstius: 1968, 41-52).

BRAINERD, B. 1968 «On the Syntax of Certain Classes of Numerical Expressions,» en Brandt Corstius (ed) 9-40. (Versión revisada de Brainerd: 1966a).

BRAINERD, B. 1973 «Mathematical and behavioral foundations of number,» *Journal of General Psychology,* 88, 221-281.

BRAINERD, B. (ed.) 1982 *Children's Logical and Mathematical Cognition,* N. York: Springer.

BRAINERD, B. y F.C.C. Peng 1968 «A Syntactic Comparison of Chinese and Japanese Numerical Expressions,» en Brandt Corstius (ed.) 53-81.

CORBETT, G.G. 1978a «Universals in the Syntax of Cardinal Numerals,» *Lingua,* XLVI, 61-74.

CORBETT, G.G. 1978b «Numerous Squishes and Squishy Numerals in Slavonic,» *International Review of Slavic Linguistics,* 3, 43-73.

HURFORD, James R. 1975 *The Linguistic Theory of Numerals,* Cambrigde (Ing.) N. York: Cambridge University Press, *Cambridge Studies in Linguistics,* 16.

HURFORD, James R. 1979 «Numerals and the Homogeneity of Description and Explanation,» *Lingua,* 48, 35-42.

HURFORD, James R. 1980 «A Note on Corbett's Numeral Universal,» *Lingua,* L.3, 247-248.

HURFORD, James R. 1987 *Language and Number. The emergence of a cognitive system,* Oxford: B. Blackwell.

KATWIJK, A. van 1965 «A Grammar of Dutch Number Names,» *Foundations of Language,* 1, 51-58.

KATWIJK,, A. van 1968 «A Functional Grammar of Dutch Number Names,» en Brandt Corstius (ed), 1-8.

MARCOS MARÍN, Francisco 1904 *Informática y Humanidades.* Madrid: Gredos.

MARTÍN DE SANTA OLALLA SÁNCHEZ, Aurora 1994 *Una propuesta de codificación morfosintáctica para corpus de referencia en lengua Española.* Tesis doctoral de la Universidad Autónoma de Madrid, dirigida por Francisco Marcos Marín.