

Del sintagma adjetivo núcleo al sintagma adjetivo

Bárbara Méndez

Universidad Nacional de Rosario
Facultad de Ciencias Agrarias
Grupo Infosur
barbaramendez555@hotmail.com

Abstract

In this paper I propose to conduct the automatic recognition of the adjective sintagma (AS). It is important to note that this study is part of my doctoral thesis, which focuses on the nucleus adjective sintagma (sadjn) within the framework of computational linguistics. Sadjn is the construction starting at the first category and finishing at its nucleus. The study of the AS involves studying the categories that may occur after the nucleus (adjective). I undertake two key tasks: the linguistic description of AS and its implementation in machine. For its implantation in machine I use two tools: Smorph and MPS, used for text processing. Smorph performs the morphological analysis and MPS works on local grammars. Both constitute a module, so that the output from Smorph is the input to MPS.

Keywords: Nucleus adjective sintagma, Adjective sintagma, Implantation in machine

Resumen

En este trabajo me propongo realizar el reconocimiento automático del sintagma adjetivo (SA). Cabe señalar que este estudio forma parte de mi tesis doctoral que tiene como objeto de investigación el sintagma adjetivo núcleo (sadjn) en el marco de la lingüística computacional. El sadjn es la construcción que empieza con la primera categoría y termina en su núcleo. El estudio del SA implica estudiar las categorías que pueden presentarse después del núcleo (adjetivo).

Realizo dos tareas fundamentales: la descripción lingüística del SA y la implantación en máquina de éste. Para la implantación en máquina me sirvo de dos herramientas informáticas: Smorph y MPS, que se utilizan para el procesamiento de textos. Smorph realiza el análisis morfológico y MPS trabaja sobre gramáticas locales. Ambas integran un módulo, de modo que la salida de Smorph es la entrada de MPS.

Palabras claves: Sintagma adjetivo núcleo, Sintagma adjetivo, Implantación en máquina.

1. INTRODUCCIÓN

En este trabajo me propongo realizar el reconocimiento automático del sintagma adjetivo (SA). Cabe señalar que este estudio forma parte de mi tesis doctoral que tiene como objeto de investigación el sintagma adjetivo núcleo (sadjn) en el marco de la lingüística computacional. El sintagma adjetivo núcleo es la construcción que empieza con la primera categoría y termina en su núcleo. El estudio del SA implica estudiar las categorías que pueden presentarse después del núcleo (adjetivo). A modo de ejemplo, el SA *tan contento con su trabajo*, está constituido por un sadjn (que se inicia con el cuantificador *tan* y termina en su núcleo *contento*) y por el sintagma preposicional *con su trabajo*.

Trabajo a partir de un corpus de 10.000 palabras de periódicos argentinos y de 5.000 palabras de textos narrativos que forma parte del banco de datos del grupo Infosur.

En trabajos anteriores (Méndez [1] [2]), he realizado la descripción acerca del sadjn, lo que me permitió su análisis automático. El reconocimiento automático del SA requiere de un estudio exhaustivo de éste. En primer lugar, realizaré una revisión del sintagma adjetivo núcleo, y propondré una clasificación del adjetivo basada fundamentalmente en la taxonomía presentada por Violeta Demonte (1982, y sobre todo, 1999) que divide dos clases de adjetivos atendiendo a su significado intrínseco: adjetivos asignadores de propiedades (calificativos y relacionales) y adjetivos adverbiales. Luego, el estudio se extenderá al SA, es decir que estudiaré las categorías que pueden ir después del adjetivo y presentaré los casos más frecuentes de SA.

Para la implantación en máquina me sirvo de dos herramientas informáticas: Smorph y MPS, que se utilizan para el procesamiento de textos. Smorph realiza el análisis morfológico y MPS trabaja sobre gramáticas locales. Ambas integran un módulo, de modo que la salida de Smorph es la entrada de MPS.

2. EL SINTAGMA ADJETIVO NÚCLEO (SADJN) Y EL SINTAGMA ADJETIVO (SA)

2.1 El sadjn

El sadjn constituye el segmento que se inicia con el primer elemento del sintagma adjetivo y finaliza en su núcleo. Lleva siempre como núcleo un adjetivo, puede aparecer sólo, por ejemplo *contento* o antecedido por modificadores preadjetivales, como *tan contento*. Las categorías que lo integran son: adjetivos, cuantificadores y adverbios.

El reconocimiento de las partes del sadjn posibilita estudiar morfológica y sintácticamente las categorías que intervienen, determinar el comienzo, el término y el núcleo del sadjn.

Para la clasificación de los adjetivos me basé fundamentalmente en la propuesta de Violeta Demonte [3]. Distingue tres clases: adjetivos calificativos, adjetivos relacionales, adjetivos adverbiales. Cabe agregar que trabajo con otro tipo de adjetivos, los participiales. A continuación, presento la clasificación con la designación correspondiente para cada clase. Esta clasificación ha resultado pertinente para la posterior modelización e implantación del sadjn.

- adjetivo calificativo (adjc): *hermoso, inteligente, flaco, extraño.*
- adjetivo relacional (adjr): *nacional, cultural, institucional.*
- adjetivo adverbial (adva): *presunto, próximo, único.*
- adjetivo participial (adjp): *escrito, dormido, amado, cantado.*

En cuanto a los modificadores preadjetivales, distinguí en:

- cuantificadores (cuant): *mucho, bastante, demasiado, harto, algo, nada, poco, medio, un tanto, tan, cuan, muy, más, menos, un tanto, tan, cuan.*
- adverbios temporales (adv): *recién, siempre.*
- adverbios de modo 1 (adm1): adverbios terminados en- mente.
- adverbios de modo 2 (adm2): *mal, bien.*

La clasificación de las categorías me permitió indagar sobre las posibles combinaciones y plantear hipótesis sobre la configuración del sadjn. A continuación, presento algunas de las hipótesis realizadas:

- Hay dos tipos de adjetivos que no se combinan con modificadores preadjetivales, por lo tanto formarán sadjn de un constituyente, éstos son **los adjetivos relacionales (adjr) y los adjetivos adverbiales (adja).**
- Tanto los **adjetivos calificativos (adjc)** como los **participiales (adjp)** pueden estar antecidos por modificadores, sin embargo presentan diferentes estructuras.
- El sadjn puede estar constituido por uno, por dos o por tres constituyentes.

En la siguiente tabla presento los posibles casos de sadjn.

Tabla n°1: Tipos de sadjn

Sadjn	Estructura	Ejemplos
Sadjn1	adjc	<i>cálido</i>
Sadjn 2	adjr	<i>nacional</i>
Sadjn3	adja	<i>presunto</i>
Sadjn4	adjp	<i>escrito</i>
Sadjn5	cuant + adjc	<i>muy cálido</i>

Tabla nº1 (continuación): Tipos de sadjn

Sadjn	Estructura	Ejemplos
Sadjn6	cuant + adjp	<i>tan escrito</i>
Sadjn7	adv1 + adjp	<i>recién escrito</i>
Sadjn8	adv1 + adjc	<i>realmente hermosa</i>
Sadjn9	adv1 + adjp	<i>increíblemente escrito</i>
Sadjn10	Adv2 + adjp	<i>bien escrito</i>
Sadjn11	Cuant + adv1 + adjc	<i>tan increíblemente atractiva</i>
Sadjn12	Cuant + adv1 + adjp	<i>tan increíblemente escrito</i>
Sadjn13	Cuant + cuant + adjc	<i>mucho más atractiva</i>
Sadjn13	Adv1 + cuant + adjc	<i>realmente más atractiva</i>

2.2 El sintagma adjetivo (SA)

En la estructura del SA pueden reconocerse tres tipos de constituyentes: el núcleo, que siempre es un adjetivo; los modificadores del adjetivo y los complementos. Siguiendo a Bosque [3] los modificadores, excepto los modificadores comparativos, van antes del adjetivo, y por lo tanto forman parte del sadjn. En cuanto a los complementos, se adjuntan a la derecha del adjetivo. Parto de los siguientes SA:

(i-a) *preocupado por la situación actual*

(i-b) *inútil con las manos*

(ii) *inseguro para manejar*

(iii) *bueno por conveniencia*

(iv) *contento con el nacimiento del niño*

Como puede verse en los ejemplos anteriores, el SP puede presentar diferentes estructuras internas:

- Puede estar conformado por las preposiciones *a, con, para* y un sintagma nominal compuesto por determinante, nombre y adjetivo.

Sp1a: Prep. (a, con, para)+ Sn1 (det +nombre + adjetivo)

preocupado por la situación actual

- Puede estar conformado por las preposiciones *a, con, para* y un sintagma nominal compuesto por determinante y nombre.

Sp1b: Prep. (a, con, para)+ Sn (det +nombre)

inútil con las manos

- Puede estar constituido por las preposiciones *a, para* y un sintagma nominal conformado por un infinitivo sustantivado.

Sp2: Prep. (a, para)+ Sn (infinitivo)

inseguro para manejar

- Puede estar conformado por las preposiciones *por, y* y un sintagma nominal compuesto por un nombre.

Sp3: prep + Sn (nombre)

Bueno por conveniencia

-Puede estar conformado por dos sintagmas preposicionales del tipo 1 como:

Sp4: Sp1(prep + Sn1)+ Sp1 (prep + Sn1)

contento con el nacimiento del niño

Tabla nº2: Tipos de Sp

Tipos de Sp	Estructura	Ejemplos
Sp1	Sp1 a Prep. (a, con, para) + Sn1 (det +nombre + adjetivo)	<i>preocupado por la situación actual</i>
	Sp1b Prep. (a, con, para)+ Sn (det +nombre)	<i>inútil con las manos</i>
Sp2	Prep. (a, para)+ Sn (infinitivo)	<i>inseguro para manejar</i>
Sp3	prep + Sn (nombre)	<i>Bueno por conveniencia</i>
Sp4	Sp1(prep + Sn1)+ Sp1 (prep + Sn1)	<i>contento con el nacimiento del niño</i>

3. IMPLANTACIÓN EN MÁQUINA

3.1. Smorph

SMORPH es un segmentador y lematizador desarrollado por Ait-Mokhtar [5], permite analizar morfológicamente, y da como salida la asignación categorial y morfológica correspondiente a cada palabra, de acuerdo con los rasgos que se declaran.

Es una herramienta declarativa: la información utilizada está separada de la maquinaria algorítmica, esto hace que se la pueda adaptar al uso que quiera darse, con el mismo software se puede tratar cualquier lengua si se le cambia la información lingüística. En Smorph deben declararse cinco tipos de información:

Entradas:

Las entradas representan el diccionario lingüístico. Las palabras pueden aparecer directamente con la indicación de los rasgos morfológicos o con la indicación del modelo morfológico al que pertenecen. En el caso de los adjetivos, se declara el lema, que por convención se expresa con la forma masculina y la indicación del modelo morfológico al que pertenece. Los cuantificadores y adverbios, preposiciones que no tienen variaciones morfológicas, se señala su carácter de adverbios y de indefinidos.

Modelos:

En los modelos se señala la estructura morfológica. En este caso, sólo los adjetivos necesitan de esta información.

Terminaciones

Son una serie de caracteres que expresan un rasgo o conjunto de rasgos, por ejemplo *o, a, os, as* para *atractivo, atractiva, atractivos, atractivas*.

Rasgos

Se debe especificar la etiqueta que se desea utilizar, por ejemplo la etiqueta morfosintáctica; EMS que contiene los valores, entre otros, de adj, adv y det.

Son una forma de expresar digitalmente el texto, se deben declarar los códigos que se desean utilizar, por ejemplo el espacio, el signo de interrogación, etc.

3.2. MPS

MPS significa módulo post-smorph. Por medio de reglas que deben ser declaradas, MPS va a trabajar con dos o más estructuras de datos de la salida de Smorph en formato Prolog2 y va a dar en salida otro formato según el analizador que se vaya a utilizar. MPS ejecuta dos funciones Principales que son la *recomposición* y la *correspondencia*. La recomposición puede ser de dos tipos diferentes: *reagrupamiento* y *división*.

En la función de reagrupamiento, MPS mediante determinadas reglas puede expresar los valores de entrada (sobre dos o más estructuras de datos de salida de Smorph) y los valores de salida sobre la estructura agrupada.

Las reglas que se declaren con la función de división van a provocar el efecto inverso, a fin de obtener en la salida una secuencia de entidades que sea análoga a las que Smorph asigna a las reglas que se declaren con la función de correspondencia van a operar sobre una sola estructura de datos a la salida de Smorph y van a poder modificarla en otra estructura de datos. Estas reglas permiten formular en Smorph descripciones básicas, generales y adaptarlas después a la exigencia de cada analizador o de cada aplicación, o enriquecerlas con nuevos pares de <etiqueta=valor>.

3.3. Resultados de Smorph y MPS

Los siguientes ejemplos de sadjn y SA pertenecen al corpus antes mencionado.

casi instantáneo

muy agudo

tan asombradamente inteligente

demasiado subjetiva

siempre dormida

sobradamente justificados

ajena a la interna sindical

carente de sorpresas

tan permisivos con los docentes

presentes en el palco oficial

presente en la sala principal

contento con los resultados

reunidos en el paraíso

siempre dormida

A continuación presento la salida de Smorph de dos ejemplos:

muy agudo

'muy'.

['muy', 'EMS', 'Cuan'].

'agudo'.

['agudo', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg'].

ajena a la interna sindical

'ajena'.

['ajeno', 'EMS', 'adjc', 'GEN', 'fem', 'NUM', 'sg'].

'a'.

['a', 'EMS', 'prep'].

"la'.

["'una' ,].

'interna'.

['interno', 'EMS', 'adjc', 'GEN', 'fem', 'NUM', 'sg'].

['internar', 'EMS', 'v', 'MODOV', 'ind', 'PERS', '3a', 'NUM', 'sg', 'TPO', 'pres'].

['internar', 'EMS', 'v', 'MODOV', 'imper', 'PERS', '2a', 'NUM', 'sg', 'TPO', 'pres'].

'sindical'.

['sindical', 'EMS', 'adjr', 'GEN', '_', 'NUM', 'sg'].

A partir de los datos obtenidos de las salidas de Smorph, utilicé la herramienta MPS con el fin de extraer los SA. Para esto debo realizar las reglas correspondientes para cada caso de SA. Por ejemplo, tomando el SA *ajena a la interna sindical*, la regla consiste en que un adjetivo calificativo (adjc) y un sintagma preposicional del tipo 1b (conformando por una preposición, un determinante, un nombre y un adjetivo) forman un SA1.

Reglas utilizadas para los ejemplos del corpus.

cuant + adjc → sadjn5 (de acuerdo a la tabla n°1)

Ej: *casi instantáneo*

cuant + advm1+ adjc → sadjn11

Ej.: *tan asombradamente inteligente*

adv2 + adjp → sadjn10

Ej.: *sobradamente justificado*

adv2 + adjp → sadjn 7

Ej: *siempre dormida*

sadjn 1(adjc) + Sp1a (prep + n+ adj) → SA1

Ej: *ajena a la interna sindical*

sadjn 1(adjc) + Sp1b(preop + snn1)→SA2

Ej.: *contento con la situación*

sadjn 4 (adjc) + Sp3 (preop + infinitivo)→S3

Ej.: *carente de sorpresas*

sadjn 5 (cuant + adjc) +Sp1(preop + n)→S5

tan permisivos con los docentes'

sadjn 4 (adjp) + Sp1 (preop + n)→S6

Ej: *reunidos en el paraíso*

Salida de Mps:

['casi instantáneo', 'EMS', 'sadjn5'].
['muy agudo', 'EMS', 'sadjn5'].
['sobradamente justificar', 'EMS', 'sadjn10'].
['tan asombradamente inteligente', 'EMS', 'sadjn11'].
['siempre dormida', 'EMS', 'sadjn7'].
['demasiado subjetiva ', 'EMS', 'sadjn5'].
['ajena a la interna sindical ', 'EMS', 'SA1'].
['carente de sorpresas ', 'EMS', 'SA3'].
['presentes en el palco oficial ', 'EMS', 'SA1'].
['presentes en el sala principal ', 'EMS', 'SA1'].
['contento con la situación ', 'EMS', 'SA2'].
['tan permisivos con los docentes', 'EMS', 'SA5'].
['reunidos en el paraíso', 'EMS', 'SA6']4

Aclaraciones:

EMS: etiqueta morfosintáctica; nom: nombre; adjc: adjetivo calificativo; adjr adjetivo relacional; v: verbo; gen: género, masc: masculino, fem: femenino; num: número; sg: singular; pl: plural; MODOV: modo verbal; ind: indicativo; PERS: persona; TPO: tiempo.

4. CONCLUSIÓN

En este artículo realicé la descripción y la implantación en máquina del sadjn y del SA. Estudié las diferentes estructuras que presentan ambas estructuras. En lo que respecta al sadjn, retomé algunas cuestiones ya estudiadas y propuse una nueva clasificación para los adjetivos. En cuanto al SA, estudie los sintagmas preposicionales que van después del núcleo.

En cuanto a la implantación en máquina, trabajé con un corpus elaborado por el equipo infosur. A partir de los datos obtenidos de la salida de Smorph utilicé la herramienta MPS. En función de las exigencias de esta herramienta declaré reglas de reagrupamiento para los sadjn y los SA.

Referencias:

- [1] Méndez B, (2008): “Descripción e implantación del sadjn”, en Revista INFOSUR N°2, edición digital: <http://www.infosurrevista.com.ar>.
- [2] Méndez, B (2009): “Formalización e implantación en máquina del sintagma adjetivo núcleo, en Revista INFOSUR N°3, edición digital: <http://www.infosurrevista.com.ar>.
- [3] Demonte, Violeta: „El adjetivo: Clases y usos. La posición del adjetivo en el sintagma nominal“. En: Bosque, Ignacio / Demonte, Violeta (eds.): Gramática descriptiva de la lengua española. Madrid: Real Academia Española / Espasa Calpe, 1999, vol. 1, pp. 130-200.
- [4] Bosque I. (1999) “El sintagma adjetival. Modificadores y complementos del adjetivo. Adjetivo y participio. “En: Bosque, Ignacio / Demonte, Violeta (eds.): Gramática descriptiva de la lengua española. Madrid: Real Academia Española / Espasa Calpe, Vol. 1, 217-310.
- [5] Ait-Mokhtar S. (1998): L’analyse presyntaxique en une seule étape. Tesis de Doctorado, GRIL, Univ. Blaise Pascal.