

Una propuesta para el tratamiento automático de los adverbios

A proposal for the automatic processing of adverbs

Andrea Rodrigo
Grupo INFOSUR, U.N.R.
Rosario, Argentina
andreafrodrigo@yahoo.com.ar

Abstract

This work considers adverbs within the “Sintagma Adverbial Núcleo” (SAdvN). The range of combinations of adverbs is considered, with a focus on some SadvN, which allows to outline a first adverb classification based on the modifiers that the head adverb can take. According to the properties of the Paradigma 5P (Bés, G.G. 1999), their description is then formalized and loaded into computer by means of Smorph and Mps (computer programming tools) in order to corroborate the linguistic hypotheses which have been set out.

Key words: Spanish adverbs. Classification. Description. Head Adverb. Computational Linguistics.

Resumen

En este trabajo, se tratan los adverbios dentro del sintagma adverbial núcleo (sadvn). Se consideran las combinaciones posibles entre adverbios y se focaliza la atención sobre algunos sadvn, lo que permite establecer una primera clasificación de adverbios a partir de los modificadores que admite el núcleo. La descripción se formaliza según las propiedades del Paradigma 5P (Bès, G.G. 1999) y se implanta en máquina con Smorph y Mps (herramientas informáticas) a fin de corroborar las hipótesis lingüísticas planteadas.

Palabras claves: adverbios en Español, clasificación, descripción, núcleo, lingüística informática.

1. INTRODUCCION

Se realiza una propuesta para el tratamiento automático de los adverbios. Para ello, se toman en cuenta las nociones de sintagma núcleo y de sintagma adverbial núcleo (sadvn). Haciendo un breve recorrido teórico en el que se plantea cómo se han considerado los adverbios en la lingüística no computacional, se pasa a una propuesta de clasificación de los adverbios, a partir de la observación de textos reales. Luego de formalizar estos datos, se procesan automáticamente con las herramientas informáticas Smorph y Mps, presentándose una muestra hacia el final del trabajo.

2. LOS ADVERBIOS EN LA LINGÜÍSTICA NO COMPUTACIONAL

En la lingüística no computacional, los adverbios han sido tratados de manera extensa. Apoyándose en un criterio semántico- sintáctico tradicionalmente se los clasifica de acuerdo a su “función” como circunstancial de modo, lugar, tiempo, causa, medio o instrumento de la acción verbal [1]. Se entiende que los adverbios forman parte de construcciones propiamente adverbiales, modificando al verbo. Ej:

(1) *Llega tarde.*

(2) *Huyó despavoridamente,*

aunque también participan de construcciones adjetivas. Ej:

(3) *casi blanco*

(4) *muy severo*

Por su parte, y sin dejar de lado la función sintáctica, desde una postura semántico- pragmática, Ofelia Kovacci [2] distingue entre adverbios léxicos y pronominales, de acuerdo a si significan o no por sí mismos. La lista de los adverbios *léxicos* incluye a los **calificativos**, (*bien, mal*), los **en - mente** como *claramente*, los de **lugar** (*arriba, abajo*) por citar algunos de ellos. En tanto, en el caso de los *pronominales* se incluye a los **deícticos** espaciales (*aquí, allí, ahí*) y los temporales (*ahora, después*), entre otros.

Desde lo morfológico, se marca el hecho de que los adverbios no flexionan y en consecuencia no establecen relaciones de concordancia. Sin embargo, algunos adverbios sufren procesos derivativos admitiendo la terminación de diminutivo. Ej:

(5) *cerquita*

(6) *despacito*

Incluso algunos adverbios pueden constituirse con la partícula superlativa *-ísimo*. Ej:

(7) *tempranísimo*

En tanto, al tocar los adverbios en *-mente*, se toma en cuenta su formación. Esta se logra con la unión de un adjetivo concordando en género y número con la terminación sustantiva *-mente*. También como en (7), admiten el superlativo. Ej:

(8) *rapidísimamente*

Una peculiaridad fonológica: los adverbios en *-mente* constituyen uno de los pocos casos en los que no se respeta el principio de acentuación única que rige en español. Así en una palabra como *fácil - mente*, además del acento ortográfico del adjetivo *fácil* aparece un acento prosódico en la palabra *- mente*. Por último, desde lo semántico, se focaliza la posibilidad que tienen algunos adverbios en *- mente* de actuar ellos mismos como adverbios de grado precediendo a otro adverbio como *bien*. Ej:

(9) *totalmente bien*

(10) *completamente bien,*

Además se registran combinaciones de tres adverbios. Ej:

(11) *sensiblemente más lejos.*

Por todo lo expuesto, se observa que desde la lingüística no computacional se enfocan los adverbios desde diversos ángulos (semántico, sintáctico, morfológico y fonológico).

Los adverbios han sido estudiados también en lo que se refiere a las **relaciones con otras clases de palabras**, como por ejemplo con la preposición. Desde el punto de vista semántico, preposiciones y adverbios expresan relaciones espaciales y temporales [3]. Existen además cuestiones sintácticas como el entorno en que aparecen, que es similar en algunos casos, aunque la distribución es diferente. En lo que respecta al nombre, también hay similitudes, ambos son término de preposición y en algunos usos admiten posesivo (Ej: *detrás suyo*), aunque los adverbios no pueden construirse con artículos o con demostrativos. Finalmente, la relación con el adjetivo es estrecha, no sólo por ser parte constitutiva de los adverbios en *-mente* en su forma femenina, sino por el hecho de que tanto adverbios como adjetivos pueden aceptar modificadores.

En cuanto a las **combinaciones**, no constituyen el criterio que se utiliza para clasificarlos. Se destacan las siguientes:

- (12) *muy cerca*
- (13) *tan lejos*
- (14) *muy tiernamente*
- (15) *muy ocasionalmente*
- (16) *más frecuentemente*
- (17) *más precisamente*
- (18) *casi anualmente*
- (19) *más tarde*
- (20) *recién ahora*
- (21) *bien circularmente*
- (22) *bien lejos*
- (23) *menos ortográficamente*
- (24) *muy difícilmente*
- (25) *muy poco amablemente*
- (26) *bastante bien*
- (27) *bastante tarde*
- (28) *algo tarde*
- (29) *demasiado frecuentemente*
- (30) *bastante excepcionalmente*
- (31) *poco probablemente*
- (32) *demasiado pronto*

Resumiendo: dentro de la lingüística no computacional, se busca describir y clasificar desde un criterio semántico-pragmático apuntando al comportamiento sintáctico. En cuanto a las combinaciones entre adverbios, se las toma más a título enumerativo que clasificatorio. Esto es lo que se va a intentar en este trabajo, el que se va a situar en la lingüística computacional.

3. LOS SINTAGMAS ADVERBIALES NÚCLEOS (sadvn)

Se procura considerar los adverbios desde la noción de sintagma núcleo, que ha sido trabajada por Bès en el estudio del sintagma verbal núcleo en francés [4] y en la tesis doctoral de Hagège sobre el sintagma nominal en portugués [5]. A su vez el sintagma nominal núcleo en español fue tomado por Solana-Rodrigo [6] y en mi tesis de maestría [7]. Un **sintagma núcleo** constituye un verdadero bloque que como tal puede extraerse mediante reglas muy precisas y cuya característica más saliente radica en el orden relativamente fijo de sus elementos internos (comienzo, núcleo), entendiéndose que el límite final se encuentra marcado por el núcleo.

Dentro de los sintagmas núcleo, se encuentran los **sintagmas adverbiales núcleo** (sadvn), en los que se centra este trabajo. Específicamente se van a tratar los sadvn constituidos por:

- un adverbio. Ej:

(33) *ciertamente*

- dos adverbios. Ej:

(34) *tan sólo*

(35) *muy decididamente*

(36) *nunca más*

(37) *completamente bien*

(38) *inmediatamente después*

(39) *casi inmediatamente*

Se excluyen las combinaciones de tipo:

(40) *poco amablemente*

(41) *bastante más lejos*

y con el intensivo mismo [8][9] :

(42) *Iré a tu casa ahora mismo,*

así como la negación, que merece un tratamiento más detallado.

4. CLASIFICACIÓN DE LOS ADVERBIOS

Al trabajar con textos reales, se observan las combinaciones entre adverbios. De acuerdo a cómo se presentan estas combinaciones y a los adverbios que modifican o son modificados por otros adverbios dentro del sadvn, se ofrece una propuesta de clasificación de los adverbios que se expone en la tabla del punto 4.2. Ahora bien es preciso primero adelantar las categorías que se observan en el cuadro.

4.1. Categorías con que se trabaja

adv

adv1 Ej: *alrededor, actualmente, enfrente*

adv2

adv2a Ej: *abajo, aproximadamente, lejos*

adv2b Ej: *absolutamente, casi, menos, muy*

adv2c Ej: *admirablemente, inmediatamente, más*

4.2. Tabla con la clasificación de los adverbios

En lo que sigue, se muestra una propuesta de clasificación de los adverbios

Tabla 1: Clasificación de los adverbios.

| Clase de adverbio | son núcleo | pueden llevar modificadores | pueden ser modificadores | ejemplos | |
|-------------------|------------|-----------------------------|--------------------------|---|--|
| adv1 | + | - | - | - <i>acaso</i> - <i>actualmente</i> - <i>alrededor</i> - <i>anoche</i> - <i>apenas</i> - <i>asimismo</i> - <i>ciertamente</i> - <i>detrás</i> - <i>enfrente</i> - <i>teóricamente</i> - <i>únicamente</i> | - <i>Actualmente no estoy estudiando</i> - <i>Vivían enfrente.</i> - <i>Apenas la toqué.</i> |
| adv2a | + | + | - | - <i>abajo</i> - <i>abiertamente</i> - <i>acá</i> - <i>adelante</i> - <i>adentro</i> - <i>afuera</i> - <i>intensamente</i> - <i>justamente</i> - <i>lejos</i> - <i>mal</i> - <i>sólo</i> - <i>superficialmente</i> - <i>temprano</i> - <i>tranquilamente</i> | - <i>Tan sólo buscaba una solución.</i> - <i>Traelo más acá.</i> - <i>Me siento más adelante.</i> - <i>Se comporta muy superficialmente.</i> |
| adv2b | +/- | - | + | - <i>casi</i> - <i>considerablemente</i> - <i>excesivamente</i> - <i>menos</i> - <i>muy</i> - <i>perfectamente</i> - <i>prácticamente</i> - <i>recién</i> - <i>tan</i> - <i>ya</i> | - <i>Te visita casi siempre.</i> - <i>Llegó considerablemente lejos.</i> - <i>Me siento perfectamente bien.</i> - <i>Ya entonces tocaba el piano.</i> - <i>Ella lo supo recién ahora.</i> |

| | | | | | |
|--------------|-----|---|---|---|--|
| adv2c | +/- | + | + | - <i>absolutamente</i> - <i>completamente</i> - <i>inmediatamente</i> - <i>más</i> | - <i>Estoy casi absolutamente convencido de que era ella.</i> - <i>Me siento absolutamente bien.</i> - <i>Fue tan inmediatamente como pudo.</i> - <i>Lo hizo inmediatamente después.</i> - <i>No voy nunca más.</i> - <i>Lo oye más claramente.</i> |
|--------------|-----|---|---|---|--|

4.2.1. La noción de núcleo

Antes de pasar a la descripción de cada una de las clases nombradas en el punto anterior, es preciso aclarar la noción de núcleo con que se trabaja aquí. Es necesario señalar que la noción de núcleo está ligada a la idea de dependencia de algunos elementos respecto de otros. El núcleo de un sintagma determina no sólo qué categorías pueden aparecer dentro de ese contexto, sino también la distribución que van a adoptar en tanto modificadores, es decir, a la izquierda. Incorporando algunas de las nociones trabajadas por Abney en lo que respecta a los patrones prosódicos ligados a la entonación, también llamados chunks [10], se entiende que el núcleo se va a marcar con un acento más fuerte dentro del sintagma.

4.2.2. Adv1

Los adv1 siempre van a ser núcleos dentro del sadvn, esta es la característica fundamental de este grupo de adverbios. La segunda característica se refiere a que siempre van a estar solos dentro del sadvn, es decir, que no van a admitir ningún tipo de modificador. Son ejemplos de este grupo:

(43) *únicamente*

(44) *acaso*

(45) *acerca*

4.2.3. Adv2

Los adv2 constituyen una clase de adverbios que, a diferencia de los adv1, pueden o no estar solos en el sadvn. Este grupo se va a subdividir de acuerdo a si modifican al núcleo o ellos mismos son núcleo del sintagma adverbial.

4.2.3.1. Adv2a

Los adv2a son el grupo más numeroso de los adverbios. Presentan como característica saliente el hecho de que pueden admitir un modificador. Este modificador es por lo común un adv2b como

muy o *tan*, aunque también son posibles otras combinaciones con algunos de los llamados adverbios “pronominales”. Ej:

- (46) *muy profundamente*
- (47) *poco profundamente*
- (48) *muy poco profundamente*
- (49) *tan sólo*
- (50) *muy tarde*

En todos los casos, los adv2a, a semejanza de los de la clase anterior son siempre núcleos del sadvn.

4.2.3.2. Adv2b

Los adv2b, a diferencia de los adv1, pueden o no ser núcleo del sadvn, dependiendo de si aparecen solos o en combinación con otros adverbios. Modifican a los adv2a. Ej:

- (51) *sumamente bien*
- (52) *prácticamente mal*

4.2.3.3. Adv2c

Los adv2c, como los 2b, pueden o no ser núcleo del sadvn. Sin embargo, a diferencia de los adv2b tanto modifican como pueden ser modificados por otro adverbio indistintamente. Se subdividen de acuerdo a qué tipo de adverbios se combinan con ellos. Los adverbios que aparecen a la izquierda de los adv2c, modificándolos, son los adv2b. Ej:

- (53) *muy inmediatamente*
- (54) *tan inmediatamente*

En cambio, los adverbios que aparecen a la derecha de los adv2c, ocupando el lugar del núcleo, son los adv2a, es decir que los adv2c los modifican. Ej:

- (55) *inmediatamente después*
- (56) *inmediatamente adelante*

Hasta aquí la descripción de los adverbios y la clasificación propuesta, en lo que sigue se explican brevemente las herramientas informáticas que se utilizan en este trabajo.

5. Las herramientas informáticas

5.1. Smorph y Mps

Se utilizan para procesar textos. **Mps** completa el análisis morfológico realizado por **Smorph**, en este caso conformando los sadvn. Estas herramientas integran una arquitectura local: la entrada de **Mps** es la salida de **Smorph**.

Smorph fue desarrollado en el GRIL¹ por Salah Aït-Mokhtar, bajo la dirección del Dr. Gabriel Bès [11]. SMORPH ha sido concebido y realizado en C++ en un entorno Unix/Linux, ha sido llevado a un entorno Windows y sus mensajes traducidos al español por Daniel Guillot (Cedia Consultores, Mendoza). Realiza la segmentación de las unidades en los textos y luego las analiza

¹ Groupe de Recherches dans les Industries de la Langue

morfológicamente, asignándoles etiquetas. Las etiquetas poseen la información lingüística que se va a atribuir a las unidades textuales. Es preciso que la información esté compactada en un fichero de tipo binario para que esté disponible. Los códigos fuente se dividen en cinco ficheros (definiciones ASCII, rasgos/valores, entradas lexicales, modelos, terminaciones) que se van a agrupar en una compilación² ejecutable, es decir se traduce a un lenguaje accesible a la computadora. En el caso del adverbio, no posee accidentes morfológicos, por lo tanto, no se construyen modelos de adverbio, tampoco se completa el fichero de terminaciones. Pero sí se va a trabajar con el **fichero de rasgos/valores**, allí se indican las etiquetas, de acuerdo con las categorías que aparecen en 4.1: adv1, adv2. Estas son las etiquetas morfosintácticas (EMS) y en tipos de adv2 (TADV2) se marcan los adv2a, 2b y 2c. En cuanto al **fichero de entradas**, el lema (la entrada tal como la muestra un diccionario) va a asociarse con los rasgos/valores ya establecidos. Ej:

| | |
|--------------------|-----------|
| (57) <lema> | <valores> |
| <i>actualmente</i> | /adv1 . |
| <i>muy</i> | /adv2b . |

Mps fue especificado por C. Hagège, G. Bès y F. Abacci con la participación de J. Rodrigo e implantado en el GRIL por F. Abacci. MPS fue implantado en C++ en un contexto Windows por Faiza Abacci. Ajusta Smorph con el analizador sintáctico. Como ya se dijo, la entrada de **Mps** es la salida de **Smorph**, en tanto, la salida de **Mps** es la segmentación de la entrada en oraciones, cada una de las cuales se segmenta en expresiones analizadas. **Mps** requiere el fichero de salida del analizador morfo-sintáctico (**Smorph**) en formato Prolog y además la presencia de tres ficheros: uno de **separación** de frases, otro que **recompone** la segmentación hecha por **Smorph** y el último que realiza la **correspondencia** entre morfología y sintaxis. De estos tres ficheros y por razones de espacio, sólo se ejemplifica el fichero de recomposición, indicando aquí una de las reglas que se implementan para el reagrupamiento de los sadvn:

(58) Con la siguiente regla (la entrada de la regla se escribe antes de la flecha y la salida después)[12], puede indicarse un sadvn como '*muy precisamente*':

%4 adv2b+adv2a da sadvn%

S1 [L1, 'TADV2', 'adv2b'] S2 [L2, 'TADV2','adv2a'] --> S1+S2 [L1+L2, 'EMS', 'sadvn'].

%*muy precisamente*%

Esto se explica de la siguiente manera:

S1, S2 son variables sobre las ocurrencias.

L1, L2 son variables sobre los lemas en el sentido en que se usa en lingüística computacional.

EMS etiqueta morfosintáctica, allí aparece el sadvn.

TADV: tipo de adverbio para 'adv2a' y 'adv2b'.

+ indica la operación de concatenación

² Se entiende por "compilar" la traducción a un lenguaje accesible a la computadora.

5.2. Análisis de Mps

En un texto real³ se aplican las herramientas **Smorph** y **Mps**, conformándose los siguientes sadvn:

'tan profundamente'.

['tan profundamente', 'EMS', 'sadvn'].

'inmediatamente después'.

['inmediatamente después', 'EMS', 'sadvn'].

'más allá'.

['más allá', 'EMS', 'sadvn'].

'más tarde'.

['más tarde', 'EMS', 'sadvn'].

% fin de la phrase numero : 1

6. CONCLUSIONES

Habiendo aplicado las herramientas informáticas en un texto real, se llega a un 100% de margen de cobertura y de precisión en el reconocimiento de los sadvn, después de haber establecido reglas para su constitución, del tipo como la que aparece en 5.1. En el punto 5. 2, se ve cómo, a través de estas reglas, se concatenan las clases de adverbio, los adv2b y los adv2a, por un lado y los adv2c y los adv2a, por otro, para constituir los sadvn.

Referencias

- [1] Real Academia Española. Esbozo de una nueva gramática de la Lengua Española. Espasa Calpe, XXII edición, Madrid, 2005.
- [2] Kovacci, O. El adverbio. En Bosque y Demonte, Gramática Descriptiva de la Lengua Española, Espasa Calpe, Madrid, 1999.
- [3] Pavón Lucero, M.V. Clases de partículas: Preposición, Conjunción y Adverbio. En Bosque y Demonte, Gramática Descriptiva de la Lengua Española, Espasa Calpe, Madrid, 1999.
- [4] Bès G.G. La phrase verbale noyau en français. En Recherches sur le français parlé, Volume 15, 1999.

³Clarín, 29/01/08

- [5] Hagège C. Analyse syntaxique automatique du portugais. Tesis de Doctorado GRIL, Univ. Blaise Pascal, 2000.
- [6] Solana Z., Rodrigo A. El sintagma nominal núcleo. En Desarrollo, implementación y utilización de modelos para el procesamiento automático de textos. Compilado por Víctor Castel, Trabajos de las Segundas Jornadas de Lingüística Informática: Modelización e Ingeniería, Facultad de Filosofía y Letras, Universidad Nacional de Cuyo, 2005.
- [7] Rodrigo, A. Análisis automático de textos, el sintagma nominal núcleo. Tesis de Maestría, Escuela de Posgrado, Facultad de Humanidades y Artes, UNR, (trabajo inédito), 2006.
- [8] Eguren, L. Pronombres y adverbios demostrativos. Las relaciones deícticas. En Bosque y Demonte, Gramática Descriptiva de la Lengua Española, Espasa Calpe, Madrid, 1999.
- [9] Peregrín Otero. Pronombres reflexivos y recíprocos. En Bosque y Demonte, Gramática Descriptiva de la Lengua Española, Espasa Calpe, Madrid, 1999.
- [10] Abney, S. Parsing by Chunks. En Berwick et al. Principle-Base Parsing. Kluwer Academic Publishers. Dordrecht, 1991.
- [11] Aït-Mokhtar, S. L'analyse presyntaxique en une seule étape. Tesis de Doctorado, GRIL, Univ. Blaise Pascal Clermont-Ferrand, 1998.
- [12] Bès G.G., Solana Z. Análisis morfológico y gramáticas locales: introducción y una aplicación concreta. En Jornadas Argentinas de Lingüística Informática Modelización e Ingeniería, JALIMI, Rosario, 2004.