

Índice

- Capítulo 1: **Introducción**
págs. 2 a 15
- Capítulo 2: **El sintagma nominal núcleo desde 5P**
págs. 16 a 46
- Capítulo 3: **Las herramientas informáticas**
págs. 47 a 58
- Capítulo 4: **SMORPH y MPS en el snn español**
págs. 59 a 81
- Capítulo 5: **Resultados y conclusiones**
págs. 82 a 88
- Anexo
págs. 89 a 185
- Bibliografía
págs. 186 a 188

Análisis automático de textos: el sintagma nominal núcleo

Andrea Rodrigo

Universidad Nacional de Rosario
Facultad de Humanidades y Artes
Escuela de Posgrado
Tesis de Maestría en Teoría Lingüística y Adquisición del lenguaje 2006

Capítulo 1¹

Introducción

0. La elección del objeto de estudio

Dentro de la oración y en el texto en general, el nombre constituye un elemento insoslayable, en tanto eje sintáctico, junto al verbo y en tanto eje semántico, por la gran carga de significado que aporta. Desde la **lingüística informática**, entendida como la ciencia que trata la aplicación de los métodos computacionales al estudio del lenguaje natural, se torna una necesidad describir la estructura del sintagma nominal (nos limitaremos al sintagma nominal núcleo) con vistas a su formalización².

Un **sintagma núcleo** es un bloque casi inseparable que comienza en el inicio del sintagma y termina en el núcleo. Su extracción resuelve ambigüedades léxicas, que no se pueden eliminar en el nivel de la palabra.

2. La descripción lingüística del sintagma nominal núcleo (snn)

En este trabajo, se focaliza la descripción lingüística del snn³ observando cuáles son las categorías morfosintácticas que pueden oficiar de núcleo del snn, cuáles se encuentran al comienzo y cuáles ocupan una posición intermedia⁴. Se hacen hipótesis descriptivas sobre la configuración del snn y una vez formalizadas⁵ estas descripciones se ponen a prueba en el análisis automático de textos reales.

¹ * Esta tesis es deudora en primer término de los seminarios que dictó Gabriel G. Bès en 2004 y 2005 y en segundo lugar del trabajo realizado por el equipo del proyecto INFOSUR.

² Se hace necesario mencionar como antecedentes de la temática, las valiosas investigaciones del Dr. Gabriel G. Bès (1999)² en el estudio del sintagma verbal núcleo en francés (svn). No se puede dejar de mencionar tampoco la tesis doctoral de Salah Aït-Mokhtar (1998), dirigida por el Dr. Bès, en la que presentó el analizador morfológico SMORPH, de cuya aplicación al español este trabajo es una muestra. Por otra parte, la tesis doctoral de Carolina Hagège (2000) sobre el sintagma nominal núcleo en portugués (snn) es otro aporte crucial, así como el trabajo en conjunto elaborado por la Dra. Zulema Solana y el Dr. Gabriel Bès (2004), en lo que respecta al svn en español rioplatense². La investigación del snn integra el Proyecto Infosur: investigación y desarrollo dirigido por la Dra. Zulema Solana, con la asesoría científica del Dr. Gabriel G. Bès.

³ Me circunscribiré a los snn cuyo núcleo es núcleo del sujeto, tanto en oraciones principales como coordinadas y subordinadas, para resumir, snn sujeto.

⁴ Me valgo de un corpus compuesto por artículos de los diarios Página 12, La Nación y Clarín.

⁵ Se formalizan mediante las propiedades, que se explican más adelante.

La especificidad del snn, a diferencia de lo que sucede en otros sintagmas como el preposicional o el adjetivo, hace que sean varias las categorías candidatas a actuar como núcleo: el nombre, el pronombre, el adjetivo, el numeral y el determinante.

- *Ellos organizaron el picnic.* (pronombre)
- *Los pobres no mienten.* (adjetivo)
- *Dos se fueron a París.* (numeral)
- *Algunos no saben lo que es el amor.* (determinante)

Las categorías que se colocan al comienzo son el determinante, el numeral, el adjetivo y el nombre.

- *Fuentes periodísticas aseguraron que el hecho se produjo en pocas horas.* (nombre)
- *Distintos textos se refieren a un mismo tema.* (adjetivo)

3. El Paradigma 5P

Se aborda el snn desde los fundamentos del Paradigma 5P que reúne las contribuciones del GRIL (Groupe de Recherches dans les Industries de la Langue) bajo la dirección del Dr. Gabriel Bès⁶. La descripción lingüística obra en este planteo como fuente declarativa de los algoritmos informáticos. Se busca así la independencia entre los dos, de manera que un mismo algoritmo puede servir para distintas lenguas. La descripción del snn debe ser lo suficientemente exhaustiva y precisa como para poder formalizar.

De los cinco módulos del Paradigma 5P (Protocolos/P1, Propiedades/P2, Proyecciones/P3, Principios/P4 y Procesos/P5) esta investigación se detiene en los Protocolos y en las Propiedades. Los **Protocolos** son las observaciones, que deben ser intersubjetivas y registrables en un lenguaje explícito. Las observaciones se centran en las expresiones, no en la utilización efectiva de una lengua (hablar, escribir, memorizar un texto)⁷. Interesan las expresiones bien formadas, en el sentido de prácticamente establecidas y se entiende que los datos son independientes del dispositivo cognitivo.

Sobre esta base, se trabaja la descripción del snn y para su formalización se especifican las **Propiedades**. Las Propiedades pueden identificarse formalmente como axiomas y están dissociadas de la maquinaria algorítmica. Al analizar los datos lingüísticos se encuentran regularidades, estas regularidades se expresan en las Propiedades. Todo lo observado se debe poder deducir entonces de los axiomas. Los modelos son los objetos que satisfacen las Propiedades.

Tres clases de Propiedades entran en juego:

- **Propiedades de existencia**

⁶ cf. BÈS, G (1999), HAGÈGE, C (2000), BÈS et al (1999), COHEUR et al (2003)

⁷ cf. BÈS (2002) p. 62

- **Propiedades de linealidad**
- **Propiedades de flechado.**

Analizando un snn cualquiera como:

- *la buena acción*

vemos cómo operan estas Propiedades.

En principio, las **Propiedades de existencia** establecen:

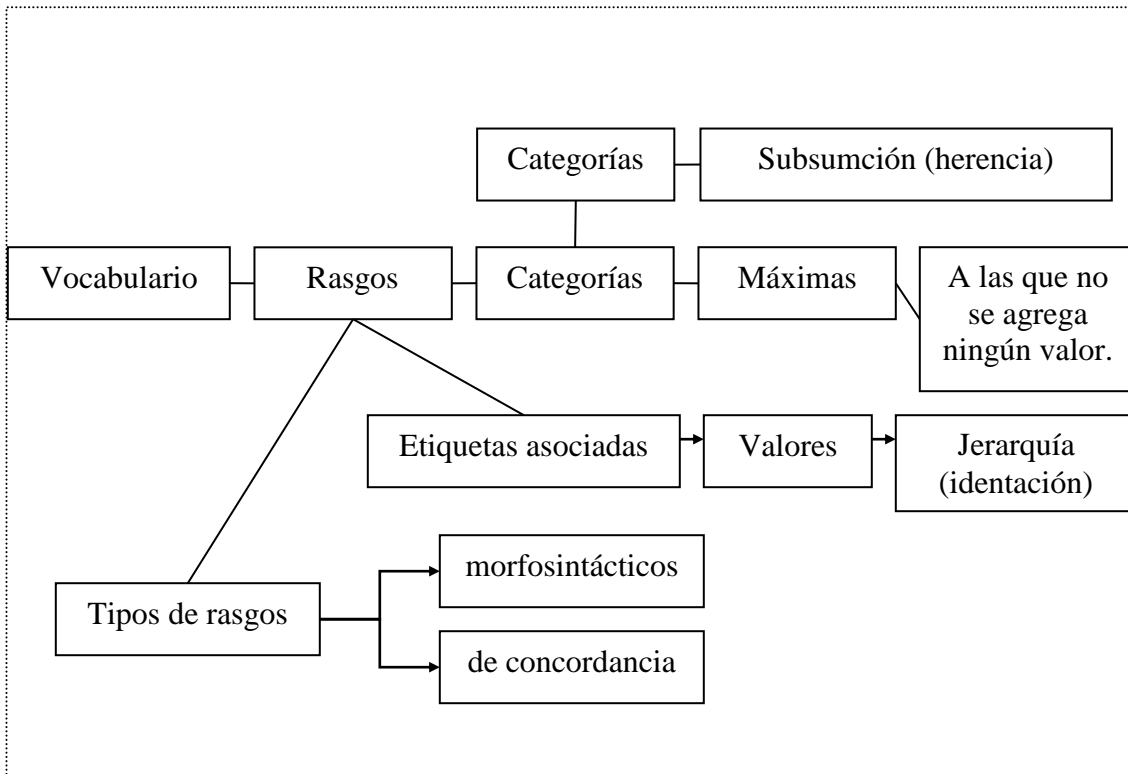
- cuál es el Vocabulario, en este caso, determinante, adjetivo, nombre,
- qué categorías morfosintácticas pueden ser núcleo, en este caso, el nombre,
- cuáles exigen la presencia de otras, en este caso, ni el artículo ni el adjetivo pueden aparecer solos, ambos necesitan la presencia del nombre para conformar el snn,
- cuáles excluyen la presencia de otras, así, en un snn como este, no pueden aparecer ni el pronombre ni otra clase de determinante como el demostrativo, por ejemplo, pero sí puede aparecer el adverbio (*la muy buena acción*).

Las **Propiedades de linealidad** tienen que ver con las relaciones de orden entre los elementos de los conjuntos definidos, marcando la precedencia de una categoría sobre otra. En lo que respecta a estas Propiedades, se observa que el adjetivo y el artículo preceden al nombre.

Y por último, las **Propiedades de flechado**, marcan las relaciones de flechaje. La noción de flechaje es muy cercana a la noción de dependencia. El rol de estas Propiedades es especificar el grafo con el cual se construye la representación semántica del conjunto de los pares en relación de flechado. En este caso, las categorías determinante y adjetivo flechan sobre el nombre, que a su vez, flecha sobre sí mismo.

Nociones fundamentales del Paradigma 5P

Solo en forma preliminar, se presentan aquí las nociones fundamentales de 5P, mostrando la relación entre ellas (serán desarrolladas más extensamente en el Capítulo 2):



rasgos: etiquetas asociadas a valores, que guardan una relación jerárquica

categorías: conjuntos de rasgos que entablan relaciones de subsumción

categorías máximas: categorías a las que no se pueden agregar ningún valor

relaciones de subsumción: hay relaciones de subsumción si, dado un conjunto A, el conjunto B es subconjunto de A.

relaciones de herencia: relaciones determinadas por las relaciones de subsumción.

4. Las herramientas informáticas

Se ha expresado cómo la descripción lingüística puede ser formalizada en virtud de las Propiedades de 5P esbozadas solo de modo introductorio. Sin embargo, queda por explicar cómo se llega a tratar automáticamente el snn ya descripto y formalizado desde 5P. Aquí me valgo de las herramientas SMORPH y MPS.

4.1 SMORPH

4.1.1 Caracterización

Es un analizador y generador morfosintáctico desarrollado en el GRIL por Salah Aït-Mokhtar, bajo la dirección del Dr. Gabriel Bès. Permite desarrollar analizadores para diferentes lenguas (francés, portugués, español, inglés y en menor medida, alemán). Concilia en una sola etapa la segmentación y la lematización de un texto, es decir, la pre-sintaxis. Realiza segmentación, análisis y generación morfológica de textos en un mismo paso. De un texto común se llega así a un texto lematizado donde se asocian la serie de unidades de base a la información lingüística. Esto se logra al asignar a cada elemento un nombre o lema determinado, como lo hace un diccionario.

Esta herramienta compila, minimiza y compacta la información lingüística para que esté disponible en un fichero binario. Los códigos fuente se dividen en cinco ficheros. Se agrupan para permitir una compilación **ejecutable**, o sea traducible al lenguaje de la computadora. En el fichero de **entradas**, se ingresan los ítems léxicos acompañados por un indicador de modelo. Este indicador de modelo es el enlace con el fichero de **modelos**, en el que se especifica la información morfológica, género y número y las terminaciones que se requieren en cada ítem.

Ej. Fichero de **entradas**:

niño @n4

Fichero de **modelos**:

@n4

-1

+o	masc/sg
+a	fem/sg
+os	masc/pl
+as	fem/pl.

En el fichero de modelos, se coloca el indicador de modelo, aquí “@n4”. A la derecha, con un signo “-” se indican los caracteres que se substraen, en este caso “-1” y a continuación con el signo “+” se declaran las terminaciones que se deben concatenar a “niñ” para conformar el masculino/singular, “+o”, el femenino/singular, “+a”, el masculino/plural, “+os”, el femenino/plural, “+as”. De esta manera, se conecta el fichero de modelos con el **fichero de terminaciones**, ya que éstas deben ser declaradas allí.

Ej. Fichero de **terminaciones**

+o
+a
+os
+as.

A su vez, en el fichero de **rasgos**, se organizan jerárquicamente las etiquetas, por ejemplo, nombre, adjetivo, etc. También se adicionan los rasgos de concordancia, género y número:

Ej: Fichero de **rasgos**

EMS etmorsin

 nom nombr

 det determ

 adj adje

TNOM tipnombr

 npr nprop

 nc ncom .

GEN género

 masc masc

 fem fem

NUM número

 sg sing

 pl plur

Lo anterior se interpreta de la siguiente manera: a la izquierda se lee, EMS, que significa etiqueta morfosintáctica, a la derecha simplemente un modo de recordar la abreviatura para el usuario, debajo de EMS se lee “nom” (nombre), “det” (determinante), “adj” (adjetivo) y así sucesivamente con las demás categorías que integran el snn. Luego se lee “TNOM” (tipo de nombre) e inmediatamente debajo, las clases de nombre con que se trabaja “npr” (nombre propio), “nc” (nombre común). Seguidamente aparecen los rasgos de concordancia: “GEN” (género) y debajo “masc” (masculino), “fem” (femenino), “NUM” (número), “sg” (singular), “pl” (plural)⁸.

Esto se completa con el **fichero de códigos ASCII**, en el que se especifican, entre otros, los caracteres separadores, las equivalencias entre Mayúsculas y Minúsculas y con el **fichero data**, que reúne la información de todos.

Para tratar el snn, se ha contado con la implantación en máquina realizada por el equipo de investigación del Proyecto **Infosur**, ingresé además 3000 nombres, 1000 adjetivos y además adverbios, pronombres, determinantes y numerales, a fin de que los textos sean analizados por SMORPH con una importante información lingüística disponible. Una vez que se declaran los datos en cada fichero, se coloca en el input un texto que va a ser analizado en la salida de SMORPH. Para ejemplificar esto, se presentará un fragmento del texto que aparece tratado en el Capítulo 5.

4.1.2 Ejemplo de Análisis morfosintáctico de un texto con la herramienta SMORPH

⁸ Aquí resumimos el fichero de rasgos, que se va a presentar de manera completa en el Capítulo 4.

4.1.2.1 Fragmento del texto⁹

Kirchner evoluciona bien, pero sigue internado en Santa Cruz Ya se le retiró el suero y está siendo medicado por vía oral. Su médico admitió que la recuperación llevará días o semanas. Pero no aclaró cuándo tendrá el alta.

4.1.2.2 Análisis de SMORPH

Aclaración: Las palabras que no han sido declaradas en el fichero “entradas” aparecen como “palabras desconocidas”, “mi” que refiere a “mot inconnue”.

'Kirchner'.
['Kirchner', 'TNOM', 'npr'].

'evoluciona'.
['evoluciona', 'mi'].

'bien'.
['bien', 'EMS', 'adv'].

','.
['cc', 'EMS', 'coma'].

'pero'.
['pero', 'mi'].

'sigue'.
['sigue', 'mi'].

'internado'.
['internado', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg', 'TADJ', 'adj1'].

'en'.
['en', 'EMS', 'prep'].

'Santa Cruz'.
['Santa_Cruz', 'TNOM', 'npr'].

'Ya'.

⁹ El texto se titula: LA SALUD DEL PRESIDENTE: LLEVA TRES DIAS HOSPITALIZADO y es de Clarín 11/04/04

['Ya', 'EMS', 'adv'].

'se'.

['se', mi].

'le'.

['le', mi].

'retiró'.

['retiró', mi].

'el'.

['el', 'TDET', 'art'].

'suero'.

['suero', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'sg'].

'y'.

['y', mi].

'está'.

['está', mi].

'siendo'.

['siendo', mi].

'medicado'.

['medicado', mi].

'por'.

['por', 'EMS', 'prep'].

'vía'.

['vía', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

'oral'.

['oral', 'TADJ', 'adj2'].

'.'.

['pf', 'EMS', 'ponc'].

'Su'.
['su', 'TDET', 'pos'].

'médico'.
['médico', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'sg'].

'admitió'.
['admitió', 'mi'].

'que'.
['que', 'mi'].

'la'.
['el', 'TDET', 'art'].

'recuperación'.
['recuperación', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

'llevará'.
['llevará', 'mi'].

'días'.
['día', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'pl'].

'o'.
['o', 'mi'].

'semanas'.
['semana', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'pl'].

'Pero'.
['Pero', 'mi'].

'no'.
['no', 'EMS', 'adv'].

'aclaró'.
['aclaró', 'mi'].

'cuándo'.
['cuándo', 'mi'].

'tendrá'.
['tendrá', 'mi].

'el'.
['el', 'TDET', 'art'].

'alta'.
['alta', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

'·'.
['pf', 'EMS', 'pnc']. ¹⁰

4.2 MPS

4.2.1 Caracterización

El **Módulo Post Smorph** (MPS) fue implantado en el GRIL por F. Abacci. Ajusta SMORPH con el analizador sintáctico. Es declarativo, se puede usar con cualquier salida de SMORPH y proporciona una entrada para cualquier analizador sintáctico. La entrada de MPS es la salida de SMORPH, en tanto, la salida de MPS es la segmentación de la entrada en oraciones, que son segmentadas en expresiones analizadas. Este módulo segmenta en oraciones y resuelve las ambigüedades por el contexto, agrupando en sintagmas núcleos las expresiones analizadas sucesivas. Actúa en la recomposición de los segmentos y realiza una desambigüización parcial en el trabajo de ajuste entre morfología y sintaxis.

Con MPS se efectúan ciertas tareas relevantes de la pre-sintaxis tomándose en cuenta para el análisis sintáctico algunas indicaciones de estructuración del documento. Se aplica MPS para la extracción de los snn. Para poder llegar a este resultado es preciso proporcionar **reglas de recomposición** (RCM) que indiquen las posibles combinaciones de categorías morfosintácticas que forman un snn y que surgen de la observación lingüística.

En lo que sigue, se presenta la aplicación de MPS sobre el fragmento visto más arriba. El análisis de SMORPH va a officiar como el input de este módulo. Cabe aclarar que solo van a aparecer concatenados los snn y cuando no concatena, MPS repite el análisis de SMORPH. Se resaltan los snn en posición sujeto: [Kirchner], [su médico], [la recuperación]. Lógicamente también van a aparecer

¹⁰ En el análisis de SMORPH falta aclarar las siguientes abreviaturas: TNOM, tipo de nombre, GEN, género, fem, femenino, NUM, número, sg, singular, pnc, punto, TDET, tipo de determinante, art, artículo, mi, palabra desconocida o no ingresada en el diccionario, adv, adverbio, masc, masculino, pos, posesivo, TADJ, tipo de adjetivo, adj2, adjetivo dos o relacional, prep, preposición, adj1, adjetivo uno (calificativo o participio), pnc, punto, coma, coma.

concatenados algunos snn que no están en posición sujeto. Al final de cada frase se lee: %fin de la phrase numero...

4.2.2 Análisis de MPS del texto presentado en el punto 4.1.2

'Kirchner'.
['Kirchner', 'EMS', 'snn'].

'evoluciona'.
['evoluciona', mi].

'bien'.
['bien', 'EMS', 'adv'].

','.
['cc', 'EMS', 'coma'].

'pero'.
['pero' , mi].

'sigue'.
['sigue' , mi].

'internado'.
['internado', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg', 'TADJ', 'adj1'].

'en'.
['en', 'EMS', 'prep'].

'Santa Cruz'.
['Santa_Cruz', 'EMS', 'snn'].

%fin de la phrase numero : 1

'Ya'.
['Ya' , 'EMS', 'adv'].

'se'.
['se' , mi].

'le'.
['le' , mi].

'retiró'.
['retiró' , mi].

'el suero'.
['el suero', 'EMS', 'snn'].

'y'.
['y', mi].

'está'.
['está', mi].

'siendo'.
['siendo', mi].

'medicado'.
['medicado', mi].

'por'.
['por', 'EMS', 'prep'].

'vía'.
['vía', 'EMS', 'snn'].

'oral'.
['oral', 'TADJ', 'adj2'].

'∴'.
['pf', 'EMS', 'ponc'].

%fin de la phrase numero : 2

'Su médico'.
['su médico', 'EMS', 'snn'].

'admitió'.
['admitió', mi].

'que'.
['que', mi].

'la recuperación'.
['el recuperación', 'EMS', 'snn'].

'llevará'.
['llevará', mi].

'días'.
['día', 'EMS', 'snn'].

'o'.
['o' , mi].

'semanas'.
['semana', 'EMS', 'snn'].

'Pero'.
['Pero' , mi].

'no'.
['no', 'EMS', 'adv'].

'aclaró'.
['aclaró' , mi].

'cuándo'.
['cuándo' , mi].

'tendrá'.
['tendrá' , mi].

'el alta'.
['el alta', 'EMS', 'snn'].

'!'.
['pf', 'EMS', 'ponc'].

%fin de la phrase numero : 3

Para emplear MPS, en el tratamiento automático del snn, se han formulado cincuenta y tres¹¹ reglas de recomposición (RCM). Para ejemplificar cómo se expresan estas reglas, se transcribe una de ellas.

4.2.3 Ejemplo de una de las reglas de recomposición de MPS para el snn:

Con la siguiente regla, puede indicarse un snn como: *'el segundo estante'*

S1 [L1, 'TDET', 'art']S2 [L2,'TNUME','num2'] S3 [L3, 'EMS','nom'] --> S1+S2+S3
[L1+L2+L3, 'EMS', 'snn'].

S1, S2, S3, son variables sobre las ocurrencias.
L1, L2, L3, son variables sobre los lemas.

¹¹ Es un espectro amplio de reglas que permitiría analizar cualquier texto.

EMS etiqueta morfosintáctica, allí aparece 'nom' la del nombre.

TDET: tipo de determinante, para 'art' artículo.

'TNUME': tipo de numeral y 'num2' correspondiente a numeral ordinal. Se trata de rasgos/valores

+ indica la operación de concatenación¹².

La regla dice que una ocurrencia analizada como 'art' seguida de otra analizada como 'num2' y de otra analizada como 'nom' se concatenan en snn

5. Consideraciones finales

La lingüística informática se ocupa de resolver diversos problemas como la búsqueda de información, la comparación de distintas fuentes de información, traducción automática, el resumen automático, etc.

Para que esto sea posible, se debe crear un ámbito interdisciplinario entre informáticos, lingüistas, lógicos y estadísticos. Al lograr que las herramientas informáticas y la modelización lingüística interactúen entre sí es posible no solo resolver los problemas actuales de información y clasificación de documentos, sino abordar los problemas futuros que se avizoran en la complejidad de la vida actual.

La pretensión de este trabajo es esbozar algunas líneas de análisis en torno al tratamiento automático del snn en español¹³. Se trata de dejar establecido aquí el modo de reconocimiento de los snn mediante análisis automático para que esto pueda usarse en recuperación y extracción de información, ámbito en el que las demandas se hacen por lo general mediante sintagmas nominales.

¹² En esta explicación sigo a BÈS- SOLANA (2004/b)

¹³ dando por sentado que en los textos reales con que se trabaja no hay desviaciones gramaticales de los fenómenos que analizo.

Capítulo 2

El sintagma nominal núcleo desde 5P

1. Los Sintagmas Núcleos

Un sintagma núcleo es un bloque casi inseparable, razón por la cual su extracción no presenta demasiados problemas. He aquí algunos sintagmas núcleos:

(i) *Los he mirado.*

Se trata de un sintagma verbal núcleo (svn)¹⁴. Entre el clítico *los* y el auxiliar *he*, entre *he* y *mirado* (verbo en participio, que obra como núcleo del sintagma) no puede intercalarse nada. Esta imposibilidad es taxativa en el primer caso, ya que el clítico se coloca antes del auxiliar, como en (i), o antes del verbo como en (ii):

(ii) *Los vi.*

en cambio, respecto de la intercalación entre el auxiliar y el participio, puede darse la presencia de adverbios o expresiones adverbiales de carácter incidental (*Juan había, ayer, terminado su trabajo, cuando... o Juan había, hace tiempo, conocido a esa mujer*). Además no se puede modificar el orden de los elementos; en cambio, si se tiene un sintagma verbal (SV) como:

(iii) *Entregué los guardapolvos a Juan.*

el orden de los elementos que lo componen puede alterarse:

(iv) *Entregué a Juan los guardapolvos (y a Pedro los portafolios).*

En cambio, al abordarse un sintagma nominal núcleo¹⁵:

(v) *las blancas playas*

En este caso, *las* lo inicia y *playas* marca su final. Se puede intercalar algún elemento entre *las* y *blancas*

(vi) *las muy blancas playas*

¹⁴ Para el sintagma verbal núcleo en francés cf. BÈS Gabriel (1999) y para el svn en español BÈS, G., SOLANA, Z. (2004/a)

¹⁵ Para el snn en portugués cf. HAGÈGE, C. (2000). En tanto, el snn en español se trata en: SOLANA, Z., RODRIGO, A. (2005)

Si se trata de un sintagma nominal (SN) puede extenderse más allá del núcleo.

(vii) *las blancas playas del sur*

Seguidamente, se introduce el sintagma adjetivo núcleo (*san*).

(viii) *tan blando*

Lo inicia el adverbio *tan* y lo finaliza el adjetivo *blando*. El orden es fijo, de modo que entre el adverbio y el núcleo (el adjetivo) no puede intercalarse ningún elemento.

Si se trata de un sintagma adjetivo (SA) pueden continuarlo otros elementos:

(ix) *tan blando por dentro*.

Se presenta en lo que sigue, un sintagma preposicional núcleo (*snn*).

(x) *en la comisaría*

Lo inicia la preposición *en* (estos sintagmas núcleo- los preposicionales- son los únicos que tienen el núcleo obligatoriamente al comienzo) y lo finaliza el próximo sintagma núcleo (en este caso el *snn*).

Los sintagmas núcleos se aproximan a los *chunks* cf. Abney (1991) Abney aduce que un *parsing* del texto en *chunks* ofrece menor dificultad y esgrime por un lado, razones prosódicas y psicolingüísticas que apoyan la existencia de los *chunks* y, por el otro, muestra las ventajas que se crean en lo que concierne a la resolución de algunas ambigüedades léxicas, que no pueden eliminarse en el nivel de la palabra, pero sí en el nivel de los *chunks*. Los *chunks* se agrupan de acuerdo a esquemas prosódicos, el acento más fuerte de la oración cae en un *chunk* y entre los *chunks* se van haciendo pausas. A su vez cabe agregar que las relaciones de orden entre *chunks* son relativamente libres y están mediadas por el léxico, mientras que las relaciones dentro del *chunk* son muy estrictas.

2. El Paradigma 5P

5P es la metodología que se utiliza para el tratamiento de lenguas naturales. Esta propuesta cristaliza las contribuciones del GRIL¹⁶ e intenta resolver las dificultades de interacción entre lingüistas¹⁷, informáticos y lógicos interesados en las lenguas naturales¹⁸. Representa un desafío: el de trabajar no en paralelo, sino

¹⁶ Groupe de Recherches dans les Industries de la Langue

¹⁷ Y tomando a la lingüística como ciencia de lo real empírico. ref: Seminario de Lingüística Informática Dr. Gabriel Bès 2005

¹⁸ En este Capítulo, se toma como bibliografía fundamental el texto BÈS (1999)

interactuando entre lo lingüístico, lo informático y lo lógico. Resulta de la complementación de dos puntos de vista: los aspectos puramente declarativos concernientes a las características de las lenguas naturales y los aspectos algorítmicos del tratamiento automático. Adapta el estudio de las lenguas naturales al esquema de las ciencias empíricas (se incluye en esta clasificación física, biología, química y se excluye matemáticas y lógica en tanto ciencias formales), entendiendo que la ciencia debe ser realizada por relaciones lógicas organizadas¹⁹. 5P intenta describir y formalizar el lenguaje natural de modo que sirva de fuente de información a la maquinaria algorítmica. No es una simple teoría y no está completamente cerrado, ya que tanto los problemas como las soluciones se presentan en términos no dogmáticos. Se plantea aquí la interacción entre la observación, la formulación de hipótesis²⁰, las deducciones a partir de estas mismas hipótesis y la verificación de las deducciones según los datos. El Paradigma 5P está integrado por diferentes módulos y el orden no es totalmente arbitrario, según se ve a continuación:

- **P1:** P de Protocolos; un Protocolo es la representación de un dato obtenido por un Observador (explícitamente modelado), es qué se observa, qué se ha decidido observar. Los Protocolos se relacionan con la sistematicidad de las observaciones,
- **P2:** P de Propiedades; las Propiedades son las descripciones formales de carácter axiomático, pero también deducciones obtenidas y testeadas con respecto a los Protocolos, tienen que ver con el conocimiento lingüístico declarativo. Un conjunto finito de Propiedades especifica en intensidad²¹ un conjunto de secuencias de expresiones de una lengua determinada. Un modelo es una secuencia que satisface un conjunto de Propiedades.
- **P3:** P de Proyecciones; las Proyecciones son generalizaciones sobre las Propiedades o sobre un subconjunto de Propiedades de una lengua natural, expresan regularidades descriptivas en un sistema lingüístico determinado
- **P4:** P de Principios; se identifican con las leyes o restricciones generales del lenguaje sobre las Proyecciones, que son válidas para todas las lenguas o para un conjunto de ellas. Las Proyecciones y los Principios se ubican en un nivel descriptivo más alto con respecto a las Propiedades.
- **P5:** P de Procesos; un Proceso es un procedimiento que está implantado en máquina o que puede ser implantado y con el que se tratan las secuencias de las lenguas naturales. Los Procesos, en tanto herramientas computacionales capaces de integrar el cálculo con los diferentes niveles descriptivos, constituyen el enlace efectivo entre las descripciones y la tecnología aplicada al lenguaje natural. Especifican la maquinaria algorítmica que permite analizar o producir un enunciado en una lengua particular.

¹⁹ El Dr. Gabriel Bès resalta en sus escritos la importancia de la concepción de ciencia de Bochensky (1958). Este alude a la idea de *connaissance intéressante*.

²⁰ Que deben formularse del modo más general posible

²¹ Entendiéndola como el conjunto de las propiedades, las características o atributos esenciales de un objeto, que le dan una identidad específica y lo distinguen de cualquier otro y sabiendo que cuanto mayor es la intensidad de un concepto, más limitada es su extensión.

Ninguno de estos módulos está subordinado al tratamiento automático, ya que 5P separa lo descriptivo y lo declarativo, respecto de los algoritmos y de los procesos. Se evita en consecuencia que las gramáticas jueguen un doble rol, en tanto descripción y en tanto fuente declarativa de los algoritmos informáticos²².

2.1 Aspectos formales de las categorías

Se supone un Vocabulario de rasgos, es decir, un conjunto no vacío y finito de rasgos. Un rasgo es una etiqueta asociada a uno o más valores. Por ejemplo, la etiqueta NUMERO (o en abreviación, NUM), posee los valores 'sg' y 'pl', la etiqueta GÉNERO (o en abreviación, GEN), los valores 'm' y 'f'. Una categoría es un conjunto finito no vacío de valores de rasgos (o de manera abreviada, de valores). Un solo valor de cada etiqueta es admitido en cada categoría. Por convención, no se utiliza un mismo valor para dos etiquetas diferentes, de manera que se pueden describir las categorías por la sola enumeración de sus valores. En general, un rasgo con sus valores se escribe:

< ETIQUETA; {valor1, ... , valorn}>

Supongamos el Vocabulario de rasgos V que sigue, en donde EMS es la abreviación de *Etiqueta Morfo-Sintáctica*, $TDET$ es la abreviación de *Tipo de DETerminante*, $TNOM$ es la abreviación de *Tipo de NOMbre* y 'det', 'n', 'adj', 'art', 'indf', 'nc', 'npr' anotan los valores correspondientes, respectivamente, a determinante, nombre, adjetivo, artículo, indefinido, nombre común y nombre propio (según se aclara en la nota 24).

$V = \{ \langle EMS; \{det, n, adj\} \rangle, \langle TDET; \{art, indf\} \rangle, \langle TNOM; \{nc, npr\} \rangle \}$

Para organizar los valores en una jerarquía y expresar las relaciones de herencia se utiliza la indentación. Así, por ejemplo, mediante

```
det
  art
  indf
n
  nc
  npr
adj
```

se expresa que si en una categoría definida mediante este esquema se tiene el valor 'art', se tiene también el valor 'det'. Se definen entonces las categorías [det, art], [det, indf], [n, nc], [n, npr], [adj].

²² Ref: Seminario de Lingüística Informática Bès (2005)

Dado el Vocabulario de rasgos V , las categorías así definidas son *categorías máximas*. En general, una categoría máxima es un conjunto de valores tal que, dado un Vocabulario de rasgos, ningún otro valor puede agregarse. Es decir, si a V se agregan los rasgos de GEN y NUM y se obtiene V' , las categorías [det, art], [det, indf], [n, nc], [n, npr], [adj] no son máximas, ya que habría que definir, por ejemplo:

[det, art, m, pl]

categoría máxima que corresponde al significante gráfico *los*.

Se obtendrían así, entre otras, las categorías máximas [det, art, m, sg], [det, art, m, pl], [det, art, f, sg], [det, art, fem, pl]. Las categorías son, formalmente, conjuntos de valores, y, por lo tanto, el orden de notación de los valores no es pertinente. De modo que [det, art, f, sg] = [sg, f, art, det].

Respecto a la *relación de subsumción* entre categorías, ya se dijo en el Capítulo 1 que un conjunto A de valores subsume un conjunto B , si A es un subconjunto de B . Así, por ejemplo, [det] subsume [det, indf] o [det, f]; [f] subsume [det, art, f, sg], [art, f, pl]...

Las relaciones de subsumción son elementos claves para expresar las Propiedades (sobre las que me detendré en el punto 2.2), pero también son indispensables para la organización del Vocabulario de rasgos, de las relaciones de herencia y de las categorías máximas que éstas determinan. Las Propiedades se expresan sobre conjuntos de valores. Van a ser válidas sobre todas las categorías subsumidas por el conjunto de valores mencionados en una Propiedad. Así, cuando se afirma que la categoría [det] precede la cat [n], significa que toda categoría con el valor [det] precede toda categoría con el valor [n], de tal manera que, según el ejemplo de V' precedente, los artículos, los indefinidos, en singular o plural, en masculino o femenino, deben preceder a los nombres, sean propios o comunes.

2.1.1 Especificación lingüística de las categorías

Se distinguen dos tipos de rasgos: los morfosintácticos y los de concordancia. Los rasgos morfosintácticos son los que siguen²³.

<EMS; {det, n, adj, num, pron, adv}>
 <TDET; {art, indf, dem, pos, tod, amb, send, cad, ciert, tant, mism}>
 <TN; {nc, npr}>
 <TADJ; {adj1, adj2}>
 <TNUM; {num1, num2}>
 <TPRON; {prpers, prindf, prdem, prpos}>

²³ Aclaración de las abreviaturas: det (determinante), art (artículo), indf (indefinido), dem (demostrativo), pos (posesivo), tod (todo), amb (ambos), send (sendos), cad (cada), ciert (cierto), tant (tanto), mism (mismo), n (nombre), nc (nombre común), npr (nombre propio), adj (adjetivo), num (numeral), pron (pronombre), adv (adverbio), prpers (pronombre personal), prindf (pronombre indefinido), prdem (pronombre demostrativo), prpos (pronombre posesivo).

Los rasgos morfosintácticos se organizan en la jerarquía que sigue, en donde solo se notan los valores de los rasgos.

```

det
  art
  indf
    indf1
      [indf1a, indf1b]
    indf2
      [indf2a, indf2b, indf2c]
  dem
  pos
  tod
  amb
  send
  cad
  ciert
  tant
  mism
n
  nc
  npr
    [npr1, npr2, npr3, npr4]
adj
  adj1
  adj2
num
  num1
  num2
pron
  prpers
  prindf
  prdem
  prpos
adv

```

Los rasgos de concordancia son los que siguen:

```

<GEN; {m, f, neu, _}>
<NUM; {sg, pl, _}>

```

El símbolo ‘_’ anota una variable que, intuitivamente, expresa “cualquier valor”. Por ejemplo, mediante `[GEN=_]` utilizado en la categoría del numeral *tres*, se expresa que éste puede usarse tanto con un masculino como con un femenino (*tres libros, tres flores*).

La combinatoria de valores de concordancia necesaria para completar la descripción de las categorías del snn es la que sigue. Por convención ad-hoc, el primer valor del par es de GEN y el segundo de NUM; a la derecha de cada par se lo ejemplifica.

[m, sg]	<i>el, libro</i>
[f, sg]	<i>la, flor</i>
[_, sg]	<i>bastante, su</i>
[m, pl]	<i>los, libros</i>
[f, pl]	<i>las, flores</i>
[_, pl]	<i>tres, bastantes, sus</i>
[f, _]	<i>crisis</i>
[m, _]	<i>lunes</i>
[neu, sg]	<i>lo</i>

Una categoría máxima es la unión de una categoría formada por todos los rasgos morfosintácticos que resultan de aplicar la jerarquía precedentemente presentada, con uno de los pares posibles de concordancia. Así, la categoría [det, tant] formará con los pares de concordancia, las categorías máximas:

[det, tant, m, sg]	<i>tanto</i>
[det, tant, f, sg]	<i>tanta</i>
[det, tant, m, pl]	<i>tantos</i>
[det, tant, f, pl]	<i>tantas</i>

Del mismo modo, la categoría [det, amb] formará sus categorías máximas.

[det, amb, f, pl]	<i>ambas</i>
[det, amb, m, pl]	<i>ambos</i>

Es claro que no todas las categorías de rasgos morfosintácticos se combinan con todos los pares posibles de rasgos de concordancia.

2.1.2 Observaciones lingüísticas e inventario de categorías máximas

En las secciones 2.1.2.1 a 2.1.2.6., se presentan las principales observaciones sobre las categorías morfosintácticas que, respectivamente, incorporan los valores 'det', 'n', 'adj', 'num', 'pron', 'adv'. Bajo la rúbrica *Inventario*, en forma de matriz, se especifican los valores de concordancia que deben asociarse a cada categoría morfosintáctica para obtener las categorías máximas, y para las clases morfológicas en inventarios cerrados, se especifican los significantes gráficos respectivos.

2.1.2.1 Determinantes²⁴

²⁴ No se toca 'det' interrogativo, ej: '¿Qué persona trajo el libro?'

Los 'det' tienen en común las siguientes características:

- (i) Constituyen clases cerradas.
- (ii) Pueden referirse a otro elemento, que es el núcleo del snn; y solo en ese caso se refieren o modifican al núcleo del snn al cual pertenecen.
- (iii) Pueden ser núcleo del snn, cumpliendo una *función pronominal*.

A continuación, se discriminan los tipos de 'det':

2.1.2.1.1 'art':

- (i) Nunca pueden ser núcleo del snn.

Inventario

m, sg	f, sg	m, pl	f, pl	_, sg	_, pl	_, _	neu,sg
<i>el</i>	<i>la</i>	<i>los</i>	<i>las</i>				<i>lo</i>

2.1.2.1.2 'indf':

- (i) Pueden ser núcleo del snn
- (ii) Pueden aparecer combinados entre ellos con restricciones.

2.1.2.1.2.1 'indf1':

- (i) Están al comienzo del snn.
- (ii) No se combinan dos 'indf1' entre sí.
- (iii) Se dividen en 'indf1a' y 'indf1b'.

2.1.2.1.2.1.1 'indf1a'²⁵:

- (i) Modifican a 'n', 'adj' y 'num' (en el caso de 'num1' cuando se refiere al nombre del número). Ej. *un hombre, ningún hombre, cualquier hombre*.
- (ii) Excepto *un*, se combinan con 'indf2b'. Ej. *cualquier otro, algún otro, ningún otro*.
- (iii) No pueden ser núcleos del snn.

Inventario²⁶

²⁵ Las posibilidades combinatorias son las que me resolvieron a separar en subclases de indefinidos.

²⁶ No incluyo los indefinidos relativos que pueden introducir subordinadas como *quienquiera, cuantos, etc*

m, sg	f,sg	m,pl	f,pl	_,sg	_,pl	_, _	neu,sg
<i>un, cualquier, algún, ningún</i>							

2.1.2.1.2.1.2 'indf1b':

- (i) Modifican a 'n'. Ej. *Mucha sal es dañina para la salud.*
- (ii) No pueden ser núcleo del snn.

Inventario

m, sg	f, sg	m,pl	f,pl	_, sg	_,pl	_, _	neu,sg
	<i>mucha poca, demasiada</i>						

2.1.2.1.2.2 'indf2':

- (i) Pueden ser núcleo del snn.
- (ii) Por sus diferentes posibilidades combinatorias se los divide en 'indf2a', 'indf2b' e 'indf2c'.

2.1.2.1.2.2.1 'indf2a':

- (i) Pueden modificar a 'n' y, salvo *mucho, demasiado y bastante* a 'adj'. Ej. *Muchos chicos fueron a la fiesta./ una loca, alguna loca, ninguna loca/ unas universitarias, algunas universitarias, muchas universitarias, demasiadas universitarias, varias universitarias.*
- (ii) *algunos, muchos, demasiados, bastantes y varios* pueden modificar a 'num1' cuando se refieren al nombre del número. Ej. *Muchos dos no te salvan esta jugada.*
- (iii) *unos y unas* pueden modificar a 'num1'. Ej. *Unos dos niños se quedaron atrás.*
- (iv) Pueden modificar a 'num2', salvo *mucho, demasiado y bastante*. Ej. *Demasiados primeros juntos hacen desastre en el patio.*
- (v) Admiten combinaciones con 'indf2b': *alguna otra, algunos otros, algunas otras, ninguna otra, muchos otros, muchas otras, demasiados otros, demasiados otros, varios otros, varias otras.*
- (vi) Existe ambigüedad de asignación categorial en *una* con el numeral.

Inventario

m, sg	f, sg	m, pl	f, pl	_, sg	1	2	3
<i>mucho</i> <i>demasiado</i>	<i>una</i> <i>alguna</i> <i>ninguna</i>	<i>unos</i> <i>algunos</i> <i>muchos</i> <i>demasiados</i> <i>varios</i>	<i>unas</i> <i>algunas</i> <i>muchas</i> <i>demasiadas</i> <i>varias</i>	<i>bastante</i>			

En el cuadro anterior, por razones de espacio:

- 1) _,pl
- 2) _, _
- 3) neu,sg

2.1.2.1.2.2.2 'indf2b':

- (i) Se anteponen a 'n', a 'adj' y a 'num2'. Ej. *otro hombre, otra pequeña, el otro tercer año.*
- (ii) Con respecto a 'num1': se anteponen *otros* y *otras*, *otro* solo al nombre del número y *otra* no se antepone. Ej. *Otros dos niños van al jardín. Otro uno me salió en el mazo.*
- (iii) Se posponen a 'dem'. Ej. *Estos otros no van en barco.*
- (iv) *otros* y *otras* se posponen a *tantos* y *tantas*. Ej. *Tantos otros sufrieron la devaluación.*
- (v) Se posponen a 'pos'. Ej. *Mis otros libros van en el estante.*

Inventario

m, sg	f, sg	m, pl	f, pl	_, sg	_,pl	_, _	neu,sg
<i>otro</i>	<i>otra</i>	<i>otros</i>	<i>otras</i>				

2.1.2.1.2.2.3 'indf2c':

- (i) Salvo *poco*, se anteponen a 'n' y a 'adj'. Ej. *Pocos hombres cumplieron. Pocas estudiantes trajeron el uniforme.*
- (ii) Se posponen a 'indf2a' en plural, excepto a *varios*: Ej: *unos pocos, algunos pocos, *varios pocos.*
- (iii) Cuando 'indf2c' se combina con 'indf2b' se anteponen y posponen excepto **poco otro*. Ej. *pocos otros, otros pocos, pocas otras, otras pocas.*
- (iv) *poco(s)* puede estar antecedido de los 'adv' *muy, bastante, demasiado, excesivamente*. Ej *muy pocos, bastante poco, demasiado poco.*

Inventario

m, sg	f, sg	m, pl	f, pl	_, sg	_,pl	_, _	neu,sg
<i>poco</i>		<i>pocos</i>	<i>pocas</i>				

2.1.2.1.3 'dem':

- (i) No pueden ser núcleo del snn²⁷.
- (ii) Pueden estar al comienzo del snn.
- (iii) Se refieren o modifican al núcleo del snn al que pertenecen.
- (iv) Modifican a 'n', 'adj', 'num2' y, salvo los femeninos singulares, a 'num1'.Ej. *Estos hombres no saben lo que quieren. Este sabio murió sin honor. Este cuarto es para la abuela* (refiriéndose a una torta). *Estos dos no trabajan.*
- (v) No se puede combinar un 'dem' con otro 'dem'.
- (vi) Se anteponen a 'indf2b'. Ej. *Esos otros no me gustan.*

Inventario

m, sg	f, sg	m, pl	f, pl	_, sg	_,pl	_, _	neu,sg
<i>este</i>	<i>esta</i>	<i>estos</i>	<i>estas</i>				
<i>ese</i>	<i>esa</i>	<i>esos</i>	<i>esas</i>				
<i>aquel</i>	<i>aquella</i>	<i>aquellos</i>	<i>aquellas</i>				

2.1.2.1.4 'pos'

- (i) Pueden estar al comienzo del snn²⁸. Ej. *Nuestra casa está en venta.*
- (ii) No pueden ser núcleo.
- (iii) Las formas *nuestro, nuestra, nuestros, nuestras* son ambiguas en cuanto a su asignación categorial, son 'pos' cuando inician el snn y 'prpos' cuando son núcleo, en cuyo caso no están precedidas por otra categoría²⁹.
- (iv) Se anteponen a 'indf2b'. Ej. *Tu otro ojo está morado.*

Inventario

m, sg	f, sg	m, pl	f, pl	_, sg	_, pl	_, _	neu,sg
<i>nuestro</i>	<i>nuestra</i>	<i>nuestros</i>	<i>nuestras</i>	<i>mi</i>	<i>mis</i>		
<i>vuestro</i>	<i>vuestra</i>	<i>vuestros</i>	<i>vuestras</i>	<i>tu</i>	<i>tus</i>		
				<i>su</i>	<i>sus</i>		

²⁷ Esta clasificación está hecha de acuerdo a la norma anterior de tildación de la Academia.

²⁸ El único caso en que los 'pos' no están al comienzo del snn. es cuando están anteceditos por 'tod'

²⁹ cf. el punto 2.1.2.5.1, donde se tratan los 'prpos'

2.1.2.1.5 'tod':

- (i) Pueden ser núcleo del snn. Ej. *Todo pasa, Todos aceptaron.*
- (ii) Siempre están al comienzo del snn.
- (iii) Cuando van con 'art', 'dem' y 'pos' se anteponen inmediatamente. Ej. *Todo este amor es para vos. Todas mis muñecas están en el ropero.*
- (iv) Pueden ir con 'indf2b' precedido de 'det', salvo 'otro', en el caso de *otra*, no necesitan ningún elemento anterior. Ej: *Toda otra cuestión está fuera de discusión. Todos los otros se tomaron el yogur. Todas las otras van a confesarse temprano. *Todo el otro se tomó el yogur.*
- (v) Cuando están seguidos inmediatamente por 'n', éste va en singular. Ej. *Todo hombre que se precie mantiene a su mujer.*
- (vi) Seguidos de 'art' pueden modificar a 'n', 'adj', 'num' o 'indf2b'. Ej. *Todos los perros van al cielo. Todas las pequeñas compran una Barbie. Todos los quintos se van a Bariloche.*
- (vii) Pueden anteceder a 'prdem'. Ej. *Todos éstos no caben en la repisa.*
- (viii) Anteceden a los 'prpers' plurales. Ej. *todos ellos*

Inventario

m, sg	f, sg	m, pl	f, pl	_, sg	_, pl	_, _	neu,sg
<i>todo</i>	<i>toda</i>	<i>todos</i>	<i>todas</i>				

2.1.2.1.6 'amb':

- (i) Pueden ser núcleo del snn. Ej. *Ambos llegaron.*
- (ii) Siempre están al comienzo del snn.
- (iii) Van siempre en plural.
- (iv) Pueden modificar a 'n' o 'adj': Ej. *Ambos niños llegaron en bicicleta. Ambas partes fueron buenas. Ambos flacos se tiñeron de rubio.*

Inventario

m, sg	f, sg	m, pl	f, pl	_, sg	_, pl	_, _	neu,sg
		<i>ambos</i>	<i>ambas</i>				

2.1.2.1.7 'send':

- (i) No pueden ser núcleo.
- (ii) Siempre están al comienzo del snn.
- (iii) Van siempre en plural.
- (iv) Modifican a 'n'. Ej. *sendas bicicletas*

Inventario

m,sg	f,sg	m, pl	f, pl	_, sg	_, pl	_, _	neu,sg
		<i>sendos</i>	<i>sendas</i>				

2.1.2.1.8 'cad':

- (i) No puede ser núcleo.
- (ii) Va en singular con 'n', 'adj' y con 'num2'. Ej. *cada hombre, cada pequeño. Cada tercero es un desastre nuevo.*
- (iii) 'cad' puede ser masculino o femenino, singular o plural, por eso se expresa con dos guiones bajos.
- (iv) No puede aparecer solo.

Inventario

m,sg	f,sg	m, pl	f, pl	_, sg	_, pl	_, _	neu,sg
						<i>cada</i>	

2.1.2.1.9 'ciert':

- (i) No pueden ser núcleo.
- (ii) Van con 'n' o 'adj'. Ej. *Cierta mujer llegó a ese pueblo. Cierta parlamentaria dio que hablar en las revistas.*

Inventario

m, sg	f, sg	m, pl	f, pl	_, sg	_, pl	_, _	neu,sg
<i>cierto</i>	<i>cierta</i>	<i>ciertos</i>	<i>ciertas</i>				

2.1.2.1.10 'tant':

- (i) Pueden ser núcleo. Ej. *Tantos se autoconvocaron en los cacero-lazos.*
- (ii) Pueden ir con 'n', 'adj'. Ej. *Tantos buenos hombres quedan solteros. Tanto político le hace mal al país.*
- (iii) En plural, pueden ir con 'indf2b'. Ej. *Tantos otros sufrieron hambre.*

Inventario

m, sg	f, sg	m, pl	f, pl	_, sg	_, pl	_, _	neu,sg
<i>tanto</i>	<i>tanta</i>	<i>tantos</i>	<i>tantas</i>				

2.1.2.1.11 'mism':

- (i) Pueden ser núcleo. Ej. *Las mismas no participaron del encuentro.*
- (ii) Deben ser precedidos por alguna categoría como 'dem' o por 'art' o por 'pos'. Ej. *Estos mismos se guardan en el armario. Los mismos hombres subieron al barco. Sus mismos amigos fueron al velorio.*

Inventario

m, sg	f, sg	m, pl	f, pl	_, sg	_, pl	_, _	neu,sg
<i>mismo</i>	<i>misma</i>	<i>mismos</i>	<i>mismas</i>				

2.1.2.2 Nombres

Trato separadamente a los 'n' comunes y a los 'n' propios. Ambos tienen dos características en común:

- (i) Constituyen un inventario abierto.
- (ii) Cuando aparece 'n' siempre es núcleo de snn.

2.1.2.2.1 Nombres comunes

La enumeración de las características que siguen se refiere a la posición sujeto del snn, que es el objeto de estudio de esta investigación:

- (i) Pueden ir antecidos de 'det', 'adj1', 'num' y 'adv'. Ej. *La casi totalidad de los costos subió por la inflación. Esas buenas chicas se casan primero. Mis tres hermanos se recibieron en julio.*
- (ii) 'n' en plural puede ir solo, sin que lo anteceda ningún elemento Ej. *Fuentes del gobierno de la ciudad explicaron que...* (Clarín)
- (iii) 'nc' en singular no puede ir solo.

2.1.2.2.2 Nombres propios

Se dividen en 'npr1', 'npr2', 'npr3' y 'npr4'

2.1.2.2.2.1 'npr1'

- (i) son los de persona simple, van solos o con 'art' acompañado por 'adj1', o 'num' Ej. *Carlos, el famoso Carlos*³⁰ o precedidos por 'dem' Ej. *Aquel Pérez sí que entendía de tango.*³¹

³⁰ Salvo registros propios de la oralidad, marcados sociolingüísticamente.

³¹ No trato casos como *El Pérez de quien te hablé.*

2.1.2.2.2 'npr2'

(i) son los de persona complejos, están formados por uno o más nombres de pila y un apellido y las restricciones respecto de la combinatoria son las mismas que las planteadas para 'npr1'.

2.1.2.2.3 'npr3'

(i) son los de institución, llevan siempre 'art', 'indf', 'dem', 'cad', 'ciert', 'amb', 'num', ya sea solos, o acompañados de 'adj1'. Ej. *la Corte Suprema de Justicia, la famosa Corte Suprema de Justicia.*

2.1.2.2.4 'npr4'

(i) son los geográficos, tienen comportamiento distinto según los lemas: *Asia/ el Asia, Argentina/ la Argentina, Chile/ *el Chile*, pero todos pueden comportarse como 'npr1'. Ej. *el Chile de Pinochet.*

2.1.2.3 Adjetivos

- (i) Pertenecen a una clase abierta.
- (ii) Distingo dos tipos de 'adj': los 'adj1' y los 'adj2'

2.1.2.3.1. 'adj1'

- (i) 'adj1' corresponde a los llamados tradicionalmente calificativos. Incluye a los participios.
- (ii) Pueden ir entre 'det' y 'n', entre 'num' y 'n'. Ej. *el hermoso día, dos hermosos días.*
- (iii) En plural, pueden iniciar el snn. Ej. *Confiables fuentes del gobierno anunciaron que...*
- (iv) Pueden ser núcleo si están anteceditos de 'det' o de 'num'. Ej. *dos tontos, los débiles.*
- (v) Pueden estar modificados por un 'adv'. Ej. *su muy querido amigo*

2.1.2.3.2 'adj2'

- (i) 'adj2' corresponde a los llamados tradicionalmente relacionales.
- (ii) Cuando aparecen, siempre son núcleo del snn.
- (iii) Van con 'det' o con 'num'. Ej. *dos constitucionales, los universitarios.*
- (iv) No pueden ir entre 'det' y 'n'. Ej. **la constitucional reforma*
- (v) No pueden estar modificados por un 'adv'.

2.1.2.4 Numerales

Se consideran solamente los tradicionalmente llamados cardinales ('num1') y ordinales ('num2').

- (i) Pertenecen a una clase abierta.
- (ii) Pueden ser núcleo del snn, en el caso de 'num1' sin necesidad del acompañamiento de otra categoría y en el caso de 'num2' acompañados de 'det'. Ej. *Tres llegaron tarde. El segundo llegó tarde.*
- (iii) Pueden estar al comienzo del snn. Ej. *dos niños* u ocupar el lugar siguiente al determinante. Ej. *los dos niños.*

Por las observaciones realizadas, no solo conviene diferenciar dos subclases 'num1' (cardinales) y 'num2' (ordinales) sino también tratarlos como a una clase separada de los 'det', entre otras razones, porque son los únicos que no constituyen una clase cerrada.

2.1.2.4.1 'num1'

- (i) Corresponde a los llamados tradicionalmente cardinales. Ej. *dos, tres, cuatro*
- (ii) Van solos o antecidos de 'det' y pueden presentarse delante de 'n'. Ej. *dos días, los dos días, los dos Pérez*³².
- (iii) Pueden ser núcleo. Ej. *Dos pasaron. Los tres están en casa.*

2.1.2.4.2 'num2'

- (i) Corresponde a los ordinales. Ej. *primero, segundo, tercero*
- (ii) No pueden aparecer solos ni iniciar el snn, esto es lo que los diferencia de 'num1'. Ej. **segundos vinieron*
- (iii) Pueden estar antes de 'n' seguidos de 'det'. Ej. *la sexta fila.*
- (iv) Pueden ser núcleo. Ej. *El segundo llegó tarde.*

2.1.2.5 Pronombres

- (i) Son siempre núcleo.
- (ii) Pueden comenzar el snn.³³

Distingo cuatro tipos de 'pron': los personales ('prpers'), los indefinidos ('prindf'), los demostrativos ('prdem') y los posesivos ('prpos').

2.1.2.5.1 'prpers'

- (i) Los plurales pueden ir antecidos por 'tod'. Ej. *todos nosotros*
- (ii) Los 'prpers', dentro del snn, no exigen rasgos de concordancia.

³² Aquí, no se tratan las variaciones en la expresión de cardinales por ejemplo: *una docena* en lugar de *doce* y la particularidad de poder estar seguidos por partitivo, como una *docena de lápices*.

³³ Solo el 'prpos' no inicia el snn, porque está antecido por 'art' u otros 'det'.

Inventario:

m, sg	f, sg	m, pl	f, pl	_, sg	_, pl	_, _	neu,sg
<i>él</i>	<i>ella</i>	<i>nosotros</i> <i>vosotros</i>	<i>nosotras</i> <i>vosotras</i>	<i>yo</i> <i>tú</i> <i>vos</i> <i>usted</i>	<i>ustedes</i>		

2.1.2.5.2 'prindf'

- (i) Los 'prindf', dentro del snn, no exigen rasgos de concordancia.
(ii) No pueden estar precedidos ni por 'tod' ni por ninguna otra categoría. Ej
Nadie vino.

Inventario

m, sg	f, sg	m, pl	f, pl	_, sg	_, pl	_, _	neu,sg
<i>algo</i> <i>uno</i> <i>alguno</i> <i>alguien</i> <i>ninguno</i> <i>cualquiera</i>		<i>nadie</i> <i>nada</i>					

2.1.2.5.3 'prdem'³⁴

- (i) Pueden ir anteceditos por 'tod'.
(ii) Los 'prdem', dentro del snn, no exigen rasgos de concordancia.

Inventario

m, sg	f, sg	m, pl	f, pl	_, sg	_, pl	_, _	neu,sg
<i>éste</i> <i>ése</i> <i>aqué</i>	<i>ésta</i> <i>ésa</i> <i>aquélla</i>	<i>éstos</i> <i>ésos</i> <i>aquéllos</i>	<i>éstas</i> <i>ésas</i> <i>aquéllas</i>				<i>esto</i> <i>eso</i> <i>aquello</i>

2.1.2.5.4 'prpos'

³⁴ Según lo que acotamos en la nota 23, esta clasificación está hecha de acuerdo a la norma anterior de tildación de la Academia.

- (i) Van precedidos de un 'det'. Ej. *el mío*, se trata de las formas no apocopadas.
 (ii) Las formas *nuestro, nuestra, nuestros, nuestras* son ambiguas en cuanto a su asignación categorial, son 'pos' cuando inician el snn y 'prpos' cuando son núcleo.

Inventario

m, sg	f, sg	m, pl	f, pl	_, sg	_, pl	_, _	neu,sg
<i>mío</i>	<i>mía</i>	<i>míos</i>	<i>mías</i>				<i>mío</i>
<i>tuyo</i>	<i>tuya</i>	<i>tuyos</i>	<i>tuyas</i>				<i>tuyo</i>
<i>suyo</i>	<i>suya</i>	<i>suyos</i>	<i>suyas</i>				<i>suyo</i>
<i>nuestro</i>	<i>nuestra</i>	<i>nuestros</i>	<i>nuestras</i>				<i>nuestro</i>
<i>vuestro</i>	<i>vuestra</i>	<i>vuestros</i>	<i>vuestras</i>				<i>vuestro</i>

2.1.2.6 Adverbios

'adv'

- (i) Pueden ir antes de un 'adj1'. Ej. *La muy hermosa no sabe lo que es ser fea*.
 (ii) Pueden ir antes de 'n': Son muy pocos los 'n' que admiten 'adv'. Ej. *la casi totalidad de sus costos*.
 (iii) Los 'adv' *muy, bastante y demasiado* pueden ir antes de *pocos*³⁵. Ej. *Muy pocos asistieron a la fiesta*.

2.2 Las Propiedades

Hemos visto consideraciones generales sobre el Paradigma 5P, los aspectos formales de las categorías, la especificación lingüística de las categorías y también observaciones lingüísticas que nos muestran cómo se puede abordar el snn desde esta perspectiva. Sin embargo, me he detenido muy brevemente en las Propiedades, uno de los puntos esenciales que es preciso tratar. Aquí voy a considerarlas en particular, en consecuencia van a volver a reiterarse conceptos claves como **categoría, valor, subsumción**.

Lo primero que debe señalarse es que las Propiedades (P2) pueden ser desde el punto de vista formal identificadas con axiomas, no son una explicación ni una teoría y están completamente disociadas de la maquinaria algorítmica. Se tienen que especificar todos los detalles de las construcciones en función de las categorías y de las entradas que se han definido en el léxico. Cuando ese material está bien descripto pueden determinarse regularidades.

Las Propiedades están asociadas a los Procesos que deberán poder reconocer en los textos las expresiones que satisfacen estas Propiedades. Son presentadas

³⁵ No se tratan los 'adv' que inician sintagma nominal como *Solo la niña lloró*.

en un formalismo calculable³⁶ e interpretable que puede ser testeado. Ante todo, la finalidad del cálculo y del formalismo es siempre un saber. El formalismo evita las especificaciones tácitas, en un sistema axiomático, formalmente desarrollado, las consecuencias se deducen de los axiomas elegidos de una manera definida³⁷. Como ya se dijo, las Propiedades son identificadas con axiomas y no constituyen en sí mismas reglas de producción. Al tratarse de un sistema "formalmente calculable", se comprende entonces un cálculo expresado por operaciones donde se evidencian las capacidades matemáticas y formales. A esto debo añadir sin embargo la noción de "efectivamente calculable", ya que no solo se habla de un cálculo formal, sino de un cálculo pasible de ser efectuado en condiciones realistas y efectivas por una máquina existente³⁸. El sistema entero es visto así como axiomático y modular, entendiendo que los modelos, en tanto listas de Categorías Máximas asociadas a significantes, constituyen objetos que satisfacen diferentes clases de Propiedades. En términos más sencillos, se dice que toda Propiedad debe tener al menos un modelo que la exprese y que a su vez todo modelo no debe contradecir ninguna de las Propiedades y debe tener al menos una expresión lingüística.

Las Propiedades³⁹ se expresan con independencia las unas de las otras y en cualquier orden y se refieren a las relaciones entre categorías⁴⁰. Las categorías son conjunto de rasgos⁴¹ y el orden de los valores no es significativo. Cada rasgo en una categoría es una etiqueta asociada a un solo valor⁴². Existen rasgos booleanos, con valores + o - :

Ej.

verbo [+ , -];
nombre [+ , -],

y también existen rasgos con más de dos valores o con dos valores que no son + o - :

Ej

género [masc, fem, _];
persona [1, 2, 3].

³⁶ Sobre este punto, BÈS (1999) retoma a BOCHENSKY (1958), quien plantea la posibilidad de hacer una abstracción del significado de los signos y considerar exclusivamente la forma gramatical. El formalismo constituye así un método que utiliza reglas y donde sólo se toma en cuenta la forma gráfica de los signos. De este modo, la sintaxis de las expresiones se presenta con los elementos que permiten que se comprenda la interpretación semántica formal.

³⁷ cf. BOCHENSKY (1958), p. 90-91. Ref. nota anterior

³⁸ cf. BÈS (1999), p. 281

³⁹ Las Propiedades que se formulan aquí para el snn en español, fueron implantadas en el útil informático diseñado por Xerox, xfst, por Celina Beltrán, ref. Seminario de Lingüística Informática Bès (2005).

⁴⁰ La noción de categoría se toma de GAZDAR. ref: Seminario de Lingüística Informática Bès (2005)

⁴¹ 5P no trabaja con signos monádicos, con signos, sino con conjuntos de etiquetas-valores. ref: Seminario de Lingüística Informática Bès 2005

⁴² cf. BÈS (1999), p. 290

Rasgos, categorías y léxico, son entonces objetos básicos que no guardan relaciones de linealidad y que no ofrecen ningún tipo de combinatoria entre las categorías o entre las entradas léxicas. Finalmente, es preciso señalar que la subsumción⁴³ es la relación básica que conecta modelos y Propiedades.

Las condiciones de adecuación externa del sistema son las siguientes:

- 1) Toda serie de palabras en un texto escrito se supone bien formada si se conforma con las condiciones de observación definidas previamente y constituye un protocolo de observación.
- 2) Toda Categoría Máxima de todo modelo debe poder estar asociada con al menos una palabra del Vocabulario⁴⁴.

Hay tres clases de Propiedades:

- 1) Existencia
- 2) Linealidad
- 3) Flechado

2.2.1 Propiedades de Existencia

Las Propiedades de existencia expresan las categorías que uno reconoce en un modelo y las relaciones de implicación de existencia entre estas categorías. No se marca ninguna relación de orden.

Determinan:

- (i) cuáles son las categorías que se pueden utilizar en los modelos,
- (ii) cuáles pueden ser núcleo,
- (iii) cuáles aparecen una sola vez,
- (iv) cuáles exigen o excluyen la presencia de otras,
- (v) cuáles concuerdan entre sí.

Hay seis clases de Propiedades de existencia (los términos ubicados entre paréntesis son los predicados con los que van a ser nombradas):

2.2.1.1 Vocabulario (*amod*)

Desde el momento en que se tiene un lenguaje, este lenguaje es un vocabulario que constituye un conjunto finito de objetos. *Amod* representa las categorías que uno puede utilizar en todos los modelos. Cada símbolo del modelo es subsumido por el Vocabulario. En el snn en español, el vocabulario se formula así:

⁴³ Hay subsumción si el conjunto de rasgos de una categoría A es incluido en el conjunto de rasgos de una categoría B. cf. BÈS (1999), op. cit., p. 292.

⁴⁴ cf. BÈS (1999), op. cit. p. 289.

P1: amod (snn-esp, [det, n, adj, num, pron, adv]).⁴⁵

Formulación intuitiva: 'det', 'n', 'adj', 'num', 'pron' y 'adv' son las categorías que aparecen en el snn.

2.2.1.2 Núcleo, (*oblig* y *obligdi*)

Se utiliza para expresar la categoría que será considerada como el núcleo del modelo. El primero (*oblig*) es empleado cuando hay una categoría posible como núcleo, como por ejemplo en el caso del sintagma adjetivo núcleo (san) y el segundo (*obligdi*), cuando hay más de una categoría posible, cada una disyuntivamente posible con respecto a las otras. En el caso del snn en español, más de una categoría puede ser núcleo, si bien siempre es una sola por cada modelo, por lo tanto se expresa con *obligdi*. (Para notar el núcleo se utiliza el símbolo ° antepuesto a la categoría). La Propiedad que expresa qué categorías pueden ser núcleo del snn en español es la siguiente:

P2: obligdi (snn-esp, [indf1, indf2, tod, amb, tant, mism, n, adj, num, pron,])⁴⁶.

Formulación intuitiva: 'indf1', 'indf2', 'tod', 'amb', 'tant', 'mism', 'n', 'adj', 'num', 'pron' pueden ser núcleos del snn.

Ej:

- *la niña* ('nc')
- *la soñada* ('adj1')
- *dos* ('num1')

En Bès (1999), se compara lo que sucede en dos snn, en portugués y en francés⁴⁷. En español, el 'n', el 'num1' y el 'prdem', pueden ser núcleo del snn:

Ej:

- *Estos hombres vinieron.*
- *Dos vinieron.*

⁴⁵ cf. SOLANA-RODRIGO (2005) p. 75

⁴⁶ cf. SOLANA-RODRIGO (2005) p. 75

⁴⁷ "Por ex. pour décrire le syntagme nominal noyau du français - snnfr- et du portugais- snnp-on a besoin, respectivement, des *obligdi* suivants

obligdi (snnfr, (n, card, ...)).

obligdi (snnpt, [n, card, dem, ...]).

Dans les deux langues, on peut avoir un syntagme nominal noyau avec un nom ou un cardinal comme noyau. Mais une différence entre les deux langues est que le français n'admet pas un démonstratif comme noyau, alors que le démonstratif peut être noyau en portugais. Exemples:

Ces ° garçons son venus

*Ces son venus

Estes ° rapazes vieram

°Estes vieram" cf. BÈS (1999), p. 296

- *Éstos vinieron.*

2.2.1.3 Unicidad, (uniq)

No hay dos símbolos asociados al modelo subsumidos por uno y el mismo símbolo. Unicidad se refiere a la exigencia de utilización única de una categoría en un modelo. La Propiedad se formula así:

P3: unic (snn-esp, [art, indf2a, indf2b, indf2c, dem, pos, tod, amb, send, cad, ciert, tant, mism, nc, adj2, num2, pron])⁴⁸.

Formulación intuitiva: Se expresa que solo hay un 'art', un 'indf2a', un 'indf2b', un 'indf2c', un 'dem', un 'pos', un 'tod', un 'amb', un 'nc', un 'adj2', un 'num2', un 'pron' en un snn.

Ej.

- **El los niños*
- **Estos esos niños*

2.2.1.4 Propiedades de Exigencia, (exig)

Se refieren a las exigencias de inclusión de un sub-conjunto de categorías en un modelo, por la presencia en el mismo modelo de otro sub-conjunto de categorías. Expresan la exigencia de la presencia de un subconjunto de categorías en un modelo, motivada por la presencia de otro subconjunto de categorías.

P+1: exig (snn-esp, [[art], [n], [adj], [indf2b], [mism], [num], [prpos]])⁴⁹.

Formulación intuitiva: La presencia de 'art' requiere a un 'n', o un 'adj', o un 'indf2b', o 'mism' o un 'num' o un 'prpos'.

Ej:

- *el espejo* ('art' + 'n')
- *la hermosa* ('art' + 'adj1')
- *las otras* ('art' + 'indf2b')
- *el mismo* ('art' + 'mism')
- *las dos* ('art' + 'num')
- *las mías* ('art' + 'prpos')

P+2: exig (snn-esp, [[indf1], [n]]).

Formulación intuitiva: La presencia de 'indf1' requiere a un 'n'.

Ej:

- *Un/cualquier/algún/ningún niño cruza solo la calle.*

⁴⁸ cf. SOLANA-RODRIGO (2005) p. 75

⁴⁹ En lo que sigue, me baso en SOLANA-RODRIGO (2005) p. 75

- *Mucha/poca/demasiada sal es nociva para la salud.*

P+3: exig (snn-esp, [[indf1a], [[indf2b], [adj], [num]]]).

Formulación intuitiva: La presencia de 'indf1a' requiere a 'indf2b', 'adj' o 'num'.

Ej:

- *Cualquier otro es mejor que vos.*
- *Un sabio no hace milagros.*
- *Ningún dos te va a salvar la jugada.*

P+4: exig (snn-esp, [[pos], [[n], [adj]]]).

Formulación intuitiva: La presencia de 'pos' requiere a 'n', 'adj'.

Ej:

- *mi casa*
- *nuestra regional*

P+5: exig (snn-esp, [[prpos], [art]]).

Formulación intuitiva: La presencia de 'prpos' requiere a 'art'.

Ej:

- *El mío no estaba.* (con 'art')

P+6: exig (snn-esp, [[ciert], [[n], [adj]]]).

Formulación intuitiva: La presencia de 'ciert' requiere a 'n' o 'adj'.

Ej:

- *Ciertos hombres van a desaparecer.* (con 'n')
- *Cierta universitaria fue a la asamblea.* (con 'adj')

P+7: exig (snn-esp, [[cad], [[adj,sg], [n,sg], [num2,sg]]]).

Formulación intuitiva: La presencia de 'cad' requiere a 'adj' o 'n' o 'num2'. Se indica además el número que debe ser siempre singular.

Ej:

- *Cada tonto va a esa fiesta.*
- *Cada doctor atiende en su respectivo consultorio.*
- *Cada segundo debe ir con su celador.* (refiriéndose a un curso)

P+8: exig (snn-esp, [[send], [n,p]]).

Formulación intuitiva: La presencia de 'send' requiere a 'n' pero en plural.

Ej:

- *Sendas bicicletas.*

P+9: exig (snn-esp, [[nc, sg], [det, sg]]).

Formulación intuitiva: La presencia de 'nc' singular requiere a 'det' también en singular.

Ej:

- * *Niño cruza la calle*
- *El niño cruza la calle.* (con 'art')

Las seis Propiedades de exigencia que siguen incorporan la variable NEANT que significa la ausencia de toda otra categoría.

P+10: exig (snn-esp, [[^otod], NEANT]).

P+11: exig (snn-esp, [[^oamb], NEANT]).

P+12: exig (snn-esp, [[^otant], NEANT]).

Formulación intuitiva: Implica que cuando 'tod' o 'amb' o 'tant' son núcleos de snn., no pueden estar acompañados por otra categoría.

Ej:

- *Todo me sale mal.*
- *Ambos van en la cajonera.*
- *Tantos vinieron a la fiesta.*

P+13: exig (snn-esp, [[prindf], NEANT]).

P+14: exig (snn-esp, [[prpers], NEANT]).

P+15: exig (snn-esp, [[prdem], NEANT]).

Formulación intuitiva: Implica que los 'pron', excepto el 'prpos'. no pueden estar acompañados por otra categoría.

Ej:

- *Ella sabe que digo la verdad.* ('prpers')
- *Algo está pasando acá.* ('prindf')
- *Esto no va más.* ('prdem')

2.2.1.5 Concordancia, (*exigac*)

Las mayúsculas notan la concordancia⁵⁰. Las exigencias de concordancia no siempre introducen necesidad de coexistencia. Simplemente se afirma que en el caso de que existieran tales categorías en el modelo, entonces se deberán acomodar a las exigencias de concordancia que se estipulan en la regla. El formalismo permite notar más de dos categorías, lo cual no es contradictorio sobre las exigencias de concordancia⁵¹. En el snn en español, se estipula la siguiente regla de concordancia.

⁵⁰ En BÈS (1999) se señala que hay tres tipos de concordancia. p. 298

⁵¹ Pour décrire le syntagme nominal noyau du français- snnfr-et du portugais-snnpt-on a besoin d'exigac très simples, comme p. ex.:

exigac (snnfr, [[artGN],[nGN]]).

Pconc: exigac (snn-esp, [[detGN], [nGN], [adjGN], [numGN], [pronGN]])⁵².

Ej:

- los (masc, pl) *perfumes* (masc, pl).
- los (masc, pl) *buenos* (masc, pl) *perfumes* (masc, pl)

Del mismo modo, se puede plantear la exigencia de no concordancia⁵³.

Las Propiedades de exigencia se complementan con las Propiedades de exclusión que son su contraparte.

2.2.1.6 Propiedades de Exclusión, (*exclu*)

Expresan las exigencias de exclusión de un subconjunto de categorías en un modelo motivada por la presencia en el mismo modelo de otro subconjunto de categorías. *Exclu* combina implicación y negación. En el snn en español, las Propiedades de exclusión son las siguientes:

P~1: exclu (snn-esp, [[⁰det], [n], [⁰adj], [⁰num], [pron]])⁵⁴

Formulación intuitiva: 'det' como núcleo excluye a 'n' y a 'pron' y también a 'adj' y a 'num' cuando son núcleos del snn, a su vez éstos se excluyen entre sí.

Ej:

- *Otra vino.*
- **Otra ella vino.*
- **Ella niña vino.*

P~2: exclu (snn-esp, [[amb], [art], [dem], [pos], [cad], [prpers], [prindf], [prdem]]).

Formulación intuitiva: 'amb' excluye a 'art', 'dem', 'pos', 'cad', 'prpers', 'prindf', 'prdem' y a su vez éstos se excluyen entre sí.

Ej:

- **Ambos ellos niños*
- **Ambos estos niños*
- **Cada estos niños.*

En Bès(1999), se compara el snn en francés y en portugués. En portugués, no es posible un 'pos' sin un 'art' o un 'dem', pero en francés ocurre lo contrario⁵⁵. En español, al igual que en francés, 'pos', 'art' y 'dem' se excluyen entre sí en el snn.

Cet exigac indique simplement qu'un article s'accorde en genre et nombre avec le nom.

cf. BÈS(1999) p. 299

⁵² cf SOLANA-RODRIGO p. 76

⁵³ cf. BÈS (1999) p. 299

⁵⁴ cf. SOLANA-RODRIGO (2005) p. 76. Lo mismo vale para todas las reglas de exclusión que presento.

⁵⁵ P ex, dans le description du syntagme nominal noyau du portugais (snnpt) on a le exig suivant

P~3: exclu(snn-esp, [[amb],[indf]]).

Formulación intuitiva: 'amb' excluye a 'indf'.

Ej:

- **ambos unos*
- **unos ambos*
- **ambos algún*

P~4: exclu(snn-esp, [[amb], [tod], [send], [tant]]).

Explicación intuitiva: 'amb' excluye a 'tod', a 'send' y a 'tant' y a su vez éstos se excluyen entre sí.

Ej:

- **ambos todos*
- **sendos ambos*
- **todos sendos*

P~5: exclu(snn-esp, [[amb], [prpos]]).

Formulación intuitiva: 'amb' excluye la presencia de 'prpos'.

Ej:

- **ambos míos*
- **ambas nuestras*

P~6: exclu(snn-esp, [[dem], [indf1], [indf2a], [tant]]).

Formulación intuitiva: 'dem' excluye a 'indf1', a 'indf2a' y a 'tant' y a su vez éstos se excluyen entre sí.

Ej:

- **este cualquier*
- **una poca*
- **bastantes estos*

P~7: exclu(snn-esp, [[dem], [cad], [prpers], [prindf], [prdem]]).

Formulación intuitiva: 'dem' excluye a 'cad', a 'prpers', a 'prindf' y a 'prdem' y a su vez éstos se excluyen entre sí.

Ej:

- **cada este*
- **él algo*

exig (snnpt, [[pos], [artd], [dem]]).

car en portugai (du Portugal) on ne peut pas avoir dans le syntagme nominal noyau un possessif san un artile défini ou un démonstratif, alors qu'une Proprieté analogue n'est pas adéquate pour les deux possessifs du francais. Aninsi en portugais on observe:

Este meu livro

- *éstos este*

P~8: exclu(snn-esp, [[pos], [art], [dem], [cad]]).

Formulación intuitiva: 'pos' excluye a 'art', a 'dem' y a 'cad' y a su vez éstos se excluyen entre sí.

Ej:

- **el mi*
- **estos cada*
- **la cada*

P~9: exclu(snn-esp, [[pos], [ciert], [pron]]).

Formulación intuitiva: 'pos' excluye a 'ciert' y a 'pron' y a su vez éstos se excluyen entre sí.

Ej:

- **cierto mi*
- **alguien cierto*
- **mi nosotros*

P~10:exclu(snn-esp, [[num1], [tod], [send], [ciert], [tant], [prpers], [prindf], [prdem]].).

Formulación intuitiva: 'num1' excluye a 'tod', 'send', 'ciert', 'tant', 'prpers', 'prindf' y 'prdem' y a su vez éstos se excluyen entre sí.

Ej:

- **dos todos*
- **ciertos sendos*
- **yo tantos*

P~11: exclu(snn-esp, [[art], [ciert], [tant], [[prpers], [prindf], [prdem]]).

Formulación intuitiva: 'art' excluye a 'ciert', a 'tant', a 'prpers', a 'prindf' y a 'prdem' y a su vez éstos se excluyen entre sí.

Ej:

- **cierta la*
- **la cierta*
- **algo esto*

P~12: exclu(snn-esp, [[cad], [tod], [send], [mism], [adv]]).

Formulación intuitiva: 'cad' excluye a 'tod', a 'send', a 'mism' y a 'adv' y a su vez éstos se excluyen entre sí.

Ej:

- **cada todo*
- **sendos mismos*

- **tan cada*

P~13: exclu(snn-esp, [[cad], [pron]]).

Formulación intuitiva: 'cad' excluye a 'pron'.

Ej:

- **yo cada*
- **cada éstos*

2.2.2 Propiedades de linealidad

Si las Propiedades de existencia definen los conjuntos de categorías susceptibles de formar los modelos, estas Propiedades introducen relaciones de orden entre los elementos de los conjuntos definidos. Formalmente, las Propiedades de linealidad transforman los conjuntos en listas. El orden en la notación de las categorías en el conjunto no es pertinente. Las Propiedades de linealidad del snn en español son las siguientes:

P<1 <snn-esp, { }, precede (det, {n, adj, num, pron, adv})>⁵⁶.

Formulación intuitiva: 'det' siempre precede a 'n', a 'adj', a 'num', a 'pron' y a 'adv'.

Ej:

- *un abogado* (precede a 'n')
- *muchos tontos* (precede a 'adj')
- *otros dos* (precede a 'num')
- *los míos* (precede a 'pron')
- *la casi totalidad* (precede a 'adv' y a 'n').

P<2 <snn-esp, { }, precede (art, {indf})>.

Formulación intuitiva: 'art' siempre precede a 'indf'.

Ej:

- *los pocos*
- *la mucha*
- *la otra*

P<3 <snn-esp, { }, precede (indf1, {indf2})>.

Formulación intuitiva: 'indf1' siempre precede a 'indf2'.

Ej:

- *Un poco es suficiente.*
- *Ninguna otra besa como vos.*
- *Algunos pocos fueron al recital.*

P<4 <snn-esp, { }, precede (adj, {n})>.

⁵⁶ cf. SOLANA-RODRIGO (2005) p. 76

Formulación intuitiva: ‘adj’ siempre precede a ‘n’.

Ej:

- *Un buen muchacho cuida a su novia.*

P<5 <snn-esp, { }, precede (tod, {n, det, adj})>.

Formulación intuitiva: ‘tod’ siempre precede a ‘n’, a ‘det’ y a ‘adj’.

Ej:

- *Todo hombre tiene una debilidad.*

- *Toda mi vida fue un milagro.*

- *Todos los tontos van al cielo.*

P<6 <snn-esp, { }, precede (num, {n})>.

Formulación intuitiva: ‘num’ siempre precede a ‘n’.

Ej:

- *los dos hombres*

- *primer grado*

P<7 <snn-esp, { }, precede (num1, {adj})>.

Formulación intuitiva: ‘num1’ siempre precede a ‘adj’.

Ej:

- *Dos hermosas mujeres llegaron al pueblo.*

P<8 <snn-esp, { }, precede (adv, {n, adj, num1})>.

Formulación intuitiva: ‘adv’ siempre precede a ‘adj’ y a ‘num1’.

Ej:

- *el tan preciado regalo*

2.2.3 Propiedades de flechado

Las Propiedades de existencia y las Propiedades de linealidad especifican conjuntos de listas. Las Propiedades de flechado dan cuenta de las relaciones entre los elementos dispuestos en estas listas. La noción de flechado es muy cercana a la noción de dependencia⁵⁷. El rol de estas Propiedades es especificar el grafo con el cual se construye la representación semántica del conjunto de los pares en relación de flechado. Si muchas categorías flechan sobre otras categorías, las categorías que flechan y la categoría sobre la que ellas flechan son consideradas como los argumentos de una función que hace posible calcular la

⁵⁷ en tanto relación entre dos elementos de una frase. cf. BÈS (1999), p. 302.

representación semántica asociada a esas categorías⁵⁸. El núcleo flecha en sí mismo por una convención. En el snn, por ejemplo, se organiza la información que va a flechar sobre el verbo. Las condiciones generales indican que todo grafo debe ser conectado, con excepción del reflexivo y que no hay cruces de flechas en los sintagmas núcleos.

Las Propiedades de flechaje dejan ordenados los elementos en listas, se expresan en relaciones de orden. Estas Propiedades se presentan escritas de manera ordenada, pero el orden de escritura no es significativo. A la derecha del antislash (\) uno nota las categorías que no pueden subsumir alguna categoría en el modelo a la izquierda del antislash⁵⁹. Con las propiedades de flechado, se pueden calcular las relaciones semánticas. Si varias categorías flechan sobre una (generalmente el núcleo) las que flechan y la que recibe el flechado obran como los argumentos de una función a partir de la cual puede calcularse la representación semántica que debe estar asociada a estas categorías, *flecha x, y [y]* expresa *x flecha sobre y*; $\{x\}$ expresa la exigencia de ausencia de *x*.

En el snn en español, cito las siguientes Propiedad de flechaje:

F1 <snn-esp,{_}, flecha (det, [n])>⁶⁰.

Formulación intuitiva: ‘det’ flecha sobre ‘n’.

F2 <snn-esp,{_}, flecha (det, [prpos])>.

Formulación intuitiva: ‘det’ flecha sobre ‘prpos’.

F3 <snn-esp,{_}, flecha(det, [mism])>.

Formulación intuitiva: ‘det’ flecha sobre ‘mism’.

F4 <snn-esp,[\{n}], flecha (det, [adj])>.

Explicación intuitiva: en ausencia de ‘n’, ‘det’ flecha sobre ‘adj’ que es el núcleo.

F5 <snn-esp,[\{[n], [adj]}], flecha (det, [num])>.

Formulación intuitiva: en ausencia de ‘n’ y de ‘adj’, ‘det’ flecha sobre ‘num’.

F6 <snn-esp,[_], flecha (num, [n])>.

Formulación intuitiva: ‘num’ flecha sobre ‘n’.

F7 <snn-esp,[_], flecha (num, [adj])>.

Formulación intuitiva: ‘num’ flecha sobre ‘adj’.

F8 <snn-esp,[\{[n],[adj]}], flecha (num1, [num2])>.

Formulación intuitiva: en ausencia de ‘n’ y de ‘adj’, ‘num1’ flecha sobre ‘num2’.

⁵⁸ sobre argumentos se cita a la Semántica formal de Kamp, ref: Seminario de Lingüística Informática Bès 2005

⁵⁹ El antislash es la negación, sacamos del conjunto tal categoría.

⁶⁰ cf. SOLANA-RODRIGO(2005), p. 77

F9 <snn-esp,[_], flecha (cad, [n])>.

Formulación intuitiva: ‘cad’ flecha sobre ‘n’.

F10 <snn-esp,[{[n], [adj]}], flecha (cad, [num])>.

Formulación intuitiva: en ausencia de ‘n’ y de ‘adj’, ‘cad’ flecha sobre ‘num’.

F11 <snn-esp,[_], flecha (ciert, [n])>.

Formulación intuitiva: ‘ciert’ flecha sobre ‘n’.

F12 <snn-esp,[_], flecha (amb, [n])>.

Formulación intuitiva: ‘amb’ flecha sobre ‘n’.

F13 <snn-esp,[_], flecha (adj, [n])>.

Formulación intuitiva: ‘adj’ flecha sobre ‘n’.

F14 <snn-esp,[_], flecha (art, [prpos])>.

Formulación intuitiva: ‘art’ flecha sobre ‘prpos’.

F15 <snn-esp,[_], flecha (indf, [n])>.

Formulación intuitiva: ‘indf’ flecha sobre ‘n’.

F16 <snn-esp,[_], flecha (indf, [adj])>.

Formulación intuitiva: ‘indf’ flecha sobre ‘adj’.

F17 <snn-esp,[{[n],[adj]}], flecha (indf, [num2])>.

Formulación intuitiva: en ausencia de ‘n’ y de ‘adj’, ‘indf’ flecha ‘num2’.

F18 <snn-esp,[{[n],[adj]}], flecha (indf1, [indf2])>.

Formulación intuitiva: en ausencia de ‘n’ y de ‘adj’, ‘indf1’ flecha sobre ‘indf2’.

Capítulo 3

Las herramientas informáticas

0. Introducción

En este Capítulo, se realiza una presentación acerca de los útiles con que trabajo en esta tesis. Se trata de SMORPH y de MPS, dos herramientas informáticas que se usan para el procesamiento de textos. SMORPH realiza el análisis y generación morfológicos y MPS (módulo Post-Smorph) trabaja sobre gramáticas locales. Ambas integran una arquitectura local, de modo que la salida de SMORPH es la entrada de MPS. El planteo intenta acercar a la lingüística como ciencia con el tratamiento automático, lo lingüístico se describe, se formaliza y sirve de fuente para lo informático: a esto se lo llama *exigencia de declaratividad*⁶².

*“Se trata entonces de herramientas que deben ser declarativas, en las cuales la explicitación de la información utilizada deber ser explícita y perfectamente delimitada con respecto a los procesos algorítmicos usados para el tratamiento automático y que reciban en entrada corpus efectivos”.*⁶³

Al conformarse como útiles declarativos entonces, la información lingüística se obtiene a partir de las descripciones lingüísticas.

En lo que sigue, se describen ambas herramientas para dar algunos criterios que nos permitan comprender de qué manera se concibieron y cómo podemos llegar a trabajar con ellas para arribar al tratamiento automático del snn en español.

1. SMORPH

SMORPH (en francés S[egmentation] et MORPH[ologie]) fue desarrollado en el GRIL por Salah Aït-Mokhtar, bajo la dirección del Dr. Gabriel Bès⁶⁴. Es un módulo destinado al tratamiento de morfología y algunos aspectos de la sintaxis. SMORPH posee la capacidad de obtener análisis mediante síntesis. Es una herramienta flexible que permite desarrollar analizadores para diferentes lenguas (francés, portugués, español, inglés y en menor medida, alemán). Está dirigido por reglas que recorren una base de datos, dicha base está compuesta de información lexical. Funciona independientemente de los datos lingüísticos y se concibió en una plataforma primaria de desarrollo de base UNIX.

⁶² BÈS (2002).

⁶³ RODRIGO MATEOS-BÈS (2004) p.2

⁶⁴ SMORPH ha sido concebido y realizado en C++ en un entorno Unix/Linux; cf.: por Salah Aït-Mokhtar (1998). Ha sido llevado a un entorno Windows y sus mensajes traducidos al español por Daniel Guillot (Cedia Consultores, Mendoza).

La novedad de este analizador y generador morfosintáctico es que concilia en una sola etapa la segmentación y la lematización de un texto⁶⁵, es decir, la pre-sintaxis. A diferencia de lo que ocurre con otros analizadores, este programa permite segmentación (tokenización), análisis y generación morfológica de textos en un mismo paso⁶⁶. Estos procesos son los que conducen hacia un tratamiento lingüístico más elaborado. La segmentación tiene por finalidad delimitar las unidades en los textos, es decir, separar al texto en frases, párrafos⁶⁷. En tanto, el análisis morfológico consiste en atribuir etiquetas a las unidades. El pasaje de un texto común a un texto lematizado supone asociar la serie de unidades de base a la información lingüística⁶⁸. Esta asociación se logra al asignar a cada elemento un nombre o lema determinado, tal como lo hace, por ejemplo, un diccionario. Los verbos se declaran así en infinitivo, los nombres, en masculino-singular.

De modo que se parte de un texto, que se presenta como una secuencia de caracteres ASCII⁶⁹ (es decir, numerizado), se determinan todas las sub-secuencias lingüísticas significativas que componen el texto, y por último éstas van a asociarse con las descripciones morfológicas correspondientes. Es decir, que en la salida de SMORPH se obtiene la segmentación en oraciones y el análisis de las expresiones de cada oración. Así:

<secuencia de Códigos ASCII ['lema', LRV]>

Ejemplo: 'abogados'.

['abogado', 'EMS'⁷⁰, 'nom', 'GEN', 'masc', 'NUM', 'pl'].

Uno de los objetivos más importantes en la concepción de esta herramienta es la obtención de independencia entre los componentes lingüísticos y la maquinaria algorítmica. El código fuente es independiente del soporte máquina.

La **segmentación** de un texto no es una operación tan simple como parece a primera vista. No siempre la falta de espacios indica que se trata de una sola unidad, es lo que ocurre con los prefijos, sufijos, etc. Los prefijos y los sufijos se distinguen en entradas caso por caso.

Ej: *anti*⁷¹

⁶⁵ La segmentación y el análisis morfológico serán realizadas en un mismo proceso que ahora accede a todas las informaciones del diccionario. AĪT-MOKHTAR, Salah (1998), p. 20.

⁶⁶ Para que la segmentación y el análisis morfológico se hagan en un solo paso, a cada caracter se le asigna una entrada y las variantes no definidas son eliminadas del análisis.

⁶⁷ AĪT-MOKHTAR, Salah, (1998), p. 8

⁶⁸ “L’association d’informations linguistiques aux unités est appelée *analyse morphologique* ou *lemmatisation*“. AĪT-MOKHTAR, Salah, (1998), p. 7

⁶⁹ ASCII, son las siglas de American Standard Code por Information Interchange (Estándar American de Codificación para el Intercambio de Información). ASCII es un estándar que representa caracteres y símbolos en forma electrónica. Apunta a una plataforma de comunicación global.

⁷⁰ La secuencia de códigos ASCII es *abogados*. El lema o EMS (etiqueta morfosintáctica) es nom (nombre) y la LRV (lista de rasgos valores) nos dicen: GEN, género: masc (masculino) y NUM, número: pl (plural).

⁷¹ El símbolo ^ indica que se concatena con algo que sigue después, se coloca a la derecha para los prefijos y a la izquierda para los sufijos.

Examinando por ejemplo la palabra *antinflamatorio*, vemos que en una sola unidad lingüística concurren dos unidades: prefijo + adjetivo. Pero a su vez, el prefijo *anti* forma palabras como *anticuerpo*, *antibalas* donde se reúne a una nombre. Si tomamos un sufijo como *eña*, lo declararíamos así:

Ej: *^eña*

Y tenemos tanto *malagueña*, *formoseña*, *panameña*, todos nombres adjetivos, puesto que pueden asumir ambos usos. Es decir, la riqueza morfológica de la lengua española nos da una variedad de combinaciones posibles que es preciso registrar en entradas para que puedan ser reconocidas por el analizador.

Por otra parte, la presencia de espacios no siempre delimita unidades⁷² ¿Qué pasa con los nombres propios, por ejemplo? En ocasiones constan de dos palabras ¿Cuál es la solución propuesta, entonces? El guión bajo o apóstrofe intermedio es el rasgo que une las ocurrencias y hace que SMORPH las lea como una sola unidad.

Ej: *Buenos_Aires* /npr⁷³,

indicándose en un costado que se trata de un solo nombre propio (compuesto en este caso). Otro ejemplo: en *una docena de huevos*, observamos que en “una docena de” se indican cuántos de la clase huevos se consideran, aquí el blanco entre las palabras no señala el límite entre unidades diferentes, puesto que hay una sola unidad.

Estudiemos además el punto. Regularmente está marcando el fin de la oración⁷⁴, sin embargo esto no siempre es lo que ocurre. Miremos ahora las siglas o las abreviaturas. En esas ocurrencias, un punto seguido de una mayúscula no es indicativo de que se haya concluido una frase. Estamos ante un caso de ambigüedad en la segmentación⁷⁵. Por lo tanto, son muchas las cuestiones que debemos tomar en cuenta a la hora de proceder a la segmentación textual.

El **análisis morfológico** consiste en atribuir información lingüística a las unidades textuales. Basándose en un diccionario, el analizador atribuye información morfológica a las unidades. El analizador reconoce las formas flexionadas a partir de los lemas. Ejemplo, con el lema *casa* reconoce *casas*. La infraestructura informática efectúa un análisis que descompone el texto en secuencias gráficas y atribuye información morfosintáctica a las respectivas etiquetas. Para recoger y tratar los datos, se extrae la información relevante de los distintos ficheros, la cual se coloca después en la memoria.

La primera tarea morfológica es determinar cuáles categorías morfológicas se van a reconocer. Para salvaguardar la independencia de los componentes, ya sea lingüísticos como informáticos, se trabaja dividiendo las categorías en dos clases: reales (creadas por la lingüística) e internas (usadas internamente

⁷² AÏT-MOKHTAR, Salah, (1998), p. 15

⁷³ El uso de guión bajo significa que puedo tener un número indefinido de rasgos entre las ocurrencias que están unidas.

⁷⁴ Al respecto, cabe citar el trabajo de BELTRÁN- BÈS (2005) Aquí se trata la problemática de la obtención de límites de oraciones, que es una etapa básica y clave del análisis automático de textos.

⁷⁵ AÏT-MOKHTAR, Salah, (1998), p. 18

por el programa). Estas categorías se agrupan en ficheros separados. El componente lingüístico define qué tipo de información debe estar registrada en el diccionario. A partir de allí, se construye una serie de reglas que permitan el reconocimiento de palabras flexionadas y otro tipo de palabras morfológicamente complejas como los diminutivos. Por su parte, en el nivel informático, el programa se ocupa de extraer los morfemas definidos por las reglas reconstruyendo las diversas formas de encontrar el lema en el diccionario.

En la construcción de un analizador, se tiende a implementar la información de modo que sea lo menos ambigua posible.

¿De qué forma se implementa SMORPH? La información lingüística es compilada y después minimizada y compactada para que esté siempre disponible en un fichero binario. Los códigos fuente son divididos en varios ficheros para obtener mejor maniobrabilidad y comprensión. Los ficheros son agrupados después en una performance que permite la compilación⁷⁶ ejecutable. Las reglas y el diccionario son desarrollados de acuerdo a la información lexical, lo que permite una cobertura lingüística que contempla tanto la flexión verbal como los accidentes morfológicos del nombre y del adjetivo, por ejemplo. La información morfológica dada por el analizador se presenta en forma de un conjunto de rasgos a los cuales están asociados valores. El lingüista define cuáles son los rasgos/valores que quiere utilizar en función de lo que está trabajando en ese momento. No es necesario declarar rasgos y valores en un fichero aparte.

Ej

Categoría	Etiqueta Morfosintáctica
n	nombre
a	adjetivo... etc

¿Qué es lo que se declara en SMORPH? Hay diferentes informaciones lingüísticas que se declaran en SMORPH:

- las definiciones ASCII
- los rasgos/valores (traits)
- los modelos donde se conjugan las formas flexivas y los rasgos/valores
- las entradas (el diccionario⁷⁷) que relacionan el lema⁷⁸ con los modelos y el lema con los rasgos/valores.
- las terminaciones, que son las formas flexivas

⁷⁶ Entendemos por “compilar” la traducción a un lenguaje accesible a la computadora.

⁷⁷ Un solo diccionario contiene información lingüística y tipográfica. La información se registra en forma declarativa y modular. Se trata de información sobre el alfabeto, sobre las formas de la lengua: prefijos, sufijos, la indicación del lema correspondiente y los rasgos. AÏT-MOKHTAR, Salah, (1998), p. 25

⁷⁸ cf. Nota a la página 20 de RODRIGO, J y BÈS, Gabriel (2001) “No debe confundirse la ocurrencia amar que es una forma verbal en infinitivo, con el lema *amar* que es una etiqueta arbitraria elegida para caracterizar un conjunto de raíces en relación con un conjunto de anteposiciones o terminaciones. La tradición hace que se elija la forma verbal del infinitivo para representar los lemas verbales, la forma singular de los nombres para los lemas nominales, las formas adjetivales singulares en masculino para representar los lemas adjetivales, etc., pero un lema NO es ni un infinitivo, ni un nombre o adjetivo en singular”.

1.1 Las definiciones ASCII

Consisten en la declaración de subconjuntos de caracteres ASCII particulares. Estas definiciones hacen posible la resolución de los problemas de segmentación.

1.1.1. Los caracteres espacios delimitan unidades cuando se encuentran dentro de una entrada lexical registrada en el diccionario.

Ej:
 ESPACIOS:
 \32 #blanco
 ESPACIOS/⁷⁹

1.1.2 Los caracteres separadores delimitan las unidades, pero no aparecen en la salida del analizador a diferencia de los que ocurre con los caracteres ASCII.

Ej
 SEPARADORES:
 \10\32 #espacios
 SEPARADORES/⁸⁰

1.1.3 Los potens son las equivalencias posibles entre caracteres. Un carácter en minúscula es equivalente a uno en mayúscula o a un carácter en mayúscula con codificación diferente.

Ej:
 POTENS:
 A a .
 B b .
 C c .
 D d .
 E e .
 F f .
 G g .
 POTENS/

1.1.4 Los alters son las alternativas entre los caracteres, llegan a reconocer un segmento con error ortográfico y se utilizan en la colocación de acentos ortográficos.

Ej.
 ALTERS:

⁷⁹ El declarar el blanco entre palabras hace que SMORPH sólo tome en cuenta este caracter para asignarle la función de reconocimiento de las ocurrencias. Vemos que se coloca dos veces la palabra "espacios" es simplemente para marcar donde comienzan y dónde terminan los caracteres declarados como tales.

⁸⁰ Podemos desambigüizar ocurrencias simplemente atendiendo a si la forma es seguida o no por un carácter separador. En la tesis doctoral de AÏT-MOKHTAR, se da el ejemplo de la palabra *sur*. En español, podemos pensar en la palabra *sobre* (como preposición) o integrando una palabra compuesta como *sobretudo*. Ver AÏT-MOKHTAR, Salah, (1998), p. 17

a	á .
e	é .
i	í .
o	ó .
u	ú .

ALTERS/

1.2 Los rasgos/valores

Es donde se vuelcan todas las etiquetas con que se va a trabajar: nombre, verbo, adjetivo, etc., pero también género, número, es decir, la información que se va a necesitar. Así por ejemplo, tenemos la etiqueta morfosintáctica (EMS) v (verbo), pero a su vez en tipo de modo verbal (TMVE) pp (participio), lo que nos muestra la amplia variedad de información que se declara.

Ej

EMS	etmorsin
v	verbo

TMVE	tipmv
pp	part . ⁸¹

1.3 Los modelos

Son los lemas de las unidades lingüísticas a partir de los que opera un mecanismo de reconocimiento o generación por flexión o derivación. Dichos modelos se registran en el diccionario. Aquí se describen diferentes modelos de flexión. Cada modelo posee un identificador, los operadores que sustraen caracteres (de modo de obtener la raíz) y operadores que concatenan los caracteres asociados a los rasgos/valores.

Tomemos el nombre *niño*, observamos que el operador de sustracción nos marca -1 lo que significa que hay sustracción de un carácter al lema. Se concatenan los caracteres -o, -a, -os y -as para formar el masculino singular, el femenino singular y el masculino y el femenino plural respectivamente.

Ej:

@n4	-1
+o	nom/masc/sg
+a	nom/fem/sg
+os	nom/masc/pl
+as	nom/fem/pl .

Aquí @n4 es el identificador de modelo, -1 es el operador de sustracción y + o, +a, +os y + as (operadores que concatenan) y nom (nombre), masc (masculino), fem (femenino), sg. (singular) y pl. (plural) son los valores.

Veamos ahora otro nombre como *general*:

⁸¹ Al declarar rasgos/valores se utiliza un punto final para indicar que se ha terminado la lista. Conviene aclarar que lo que se escribe a la derecha etmorsin (etiqueta morfosintáctica), por ejemplo, no es leído por SMORPH, sólo es una indicación para el usuario.

Ej:

@n5	-0	
+@		nom/masc/sg
+es		nom/masc/pl .

Aquí @n5 es el identificador de modelo, -0 es el operador de sustracción, que aquí es *ceró*, porque no se sustrae ningún carácter al lema y +@ (operador que concatena, un conjunto vacío, puesto que no es preciso adicionar ningún carácter) y +es (operador que concatena). En tanto, nom (nombre), masc (masculino), sg. (singular) y pl. (plural) son los valores.

De esta manera, también se tratan las formas verbales, los adjetivos e incluso la derivación morfológica. También se contemplan las contracciones como *al* y *del* (preposición + artículo).

1.4 Las entradas lexicales⁸²

Se pueden definir de distintas maneras, ya sea que se incluyan lema (o forma canónica), identificador de modelo, la raíz, la descripción lingüística, la terminación, etc. Veamos una forma de declarar en entradas, aquí nombres comunes, con el indicador de modelo.

Ej:

<lema>	<modelo>
baba ⁸³	@n2 .
babero	@n1 .
babor	@n26 .
babosa	@n2 .
bacalao	@n1 .
batche	@n1 .
bachiller	@n7 .
bachillerato	@n1 .
bacilo	@n1 . ⁸⁴

Ejemplifiquemos también a los nombres propios⁸⁵:

Ej

<lema>	<valores>
Alejandro	/npr .
Antonio	/npr .
Angel	/npr .
Azar	/npr .
Buccolini	/npr .

⁸² Es conveniente especificar en las entradas del diccionario si se trata o no de formas ambiguas. AĪT-MOKHTAR, Salah, (1998), p. 19

⁸³ Las formas son notadas en minúscula, a excepción de las siglas o nombres propios. AĪT-MOKHTAR, Salah, (1998) p. 20

⁸⁴ Al declarar las entradas, se utiliza un punto al finalizar cada una de ellas.

⁸⁵ En virtud de la normalización de las mayúsculas, se entiende que si el segmento contiene una mayúscula, se trata de un nombre propio lexical. AĪT-MOKHTAR, Salah, (1998), p.

Buenos_Aires ⁸⁶/npr .
 Buonomo /npr .

como otra de las formas posibles de ingresar las entradas, incorporando los valores. Una tercera forma es:

Ej

<lema>	<ocurrencia>	<valores>
el	el	/art.
el	la	/art.

1.5 Las terminaciones⁸⁷

Se declaran las utilizadas en los modelos de flexión en el fichero separador.

Ej:

- @ . ⁸⁸
- o .
- a .
- os .
- as . ⁸⁹

También es posible marcar lo que se llaman terminaciones distinguidas, es cuando tenemos una particularidad lingüística, la cual se anota con una indicación especial⁹⁰.

Ej.

- ción nom .
- ciones nom .

SMORPH utiliza los cinco tipos de información declarativa y de allí genera y compila todas las expresiones analizadas posibles con sus listas de rasgos /valores (más de una lista cuando hay ambigüedades) y finalmente da por resultado todas las expresiones compiladas. A través del análisis, compara las secuencias de entrada con cada una de las expresiones compiladas. Hay tres propiedades que nos hablan de la capacidad de SMORPH:

- 1) Si la expresión x no está registrada va a analizar un elemento supletorio como y, si es que se declaró debidamente que x puede ser reemplazado por y. Ejemplo, mayúscula y minúscula (a A) o en la colocación del acento ortográfico (a á)

⁸⁶ Ya dije antes que las formas compuestas que contienen separadores se indican con guiones bajos entre sí.

⁸⁷ En tanto secuencias de caracteres utilizados en los modelos.

⁸⁸ El símbolo @ va a aparecer tanto en modelos como en terminaciones. En modelos, marca que no se concatena ningún carácter, en terminaciones, una terminación vacía

⁸⁹ También en este caso, como en entradas, a cada terminación declarada sigue un punto.

⁹⁰ SMORPH deduce las expresiones que resultan al aplicar la morfología derivacional mediante las terminaciones distinguidas.

- 2) En una palabra como *abiótico*, por ejemplo, no se analiza preposición más adjetivo (a + biótico) sino como una totalidad, desde el momento en que el analizador prioriza la secuencia de códigos ASCII más larga.

Antes de proseguir con la propiedad número tres, hagamos una pequeña digresión explicativa. El principio de la más larga forma⁹¹ (de que se habla en la propiedad dos) propone que en el momento de realizar la segmentación, se reenvíe la serie compuesta de sub-segmentos a las más largas formas reconocidas⁹². A su vez, una segmentación válida produce solamente el segmento de la forma reconocida como la más larga si y sólo si la forma no sea ambigua en segmentación y contenga sub-ocurrencias.

- 3) Con el símbolo ^ se marca la posibilidad de concatenar un sufijo a otra expresión como en el caso de los enclíticos, en comerlo, por ejemplo.

Por último es importante decir que cuando SMORPH encuentra una palabra desconocida⁹³, es decir, no definida en el diccionario, el proceso no se detiene. Anotemos aquí cómo se registran en la salida de SMORPH las palabras desconocidas, para dar cuenta de cómo el analizador prosigue su tarea:

Ej:

'En'.
['En', mi⁹⁴].

'un'.
['un', mi].

'clima'.
['clima', mi].

'tenso,'.
['tenso,', mi].

Habiendo culminado estas consideraciones teóricas sobre SMORPH y antes de pasar revista al módulo MPS y a la relación entre ellos, cierro este punto señalando los límites de SMORPH:

- 1) No resuelve ambigüedades
- 2) No concatena los análisis de salida.⁹⁵

2. MPS

⁹¹ AÏT-MOKHTAR, Salah, (1998) p. 37.

⁹² La prospección de las ocurrencias debe hacerse de izquierda a derecha, a diferencia de lo que sucede con otros analizadores.

⁹³ Para entender cómo procede SMORPH frente a las palabras desconocidas, ver AÏT-MOKHTAR, Salah, (1998) p. 39

⁹⁴ mi, representa palabra desconocida.

⁹⁵ BÈS-SOLANA (2004/b) p. 6.

El Módulo Post Smorph, fue especificado por C. Hagège, G. Bès y F. Abbaci con la participación de J. Rodrigo e implantado en el GRIL por F. Abacci⁹⁶. MPS se encarga de ajustar SMORPH con el analizador sintáctico. Es perfectamente declarativo, se puede usar con cualquier salida de SMORPH y produce una entrada para cualquier analizador sintáctico. Hace tratamiento de gramática local, con él es posible modificar la segmentación de un texto hecha por el analizador morfológico, el contenido y el formato de la información asociada con el segmento. Como ya dijimos, 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 es segmentada en expresiones analizadas. Observamos que en la salida de SMORPH se mantienen las ambigüedades, entonces es función de este módulo segmentar en oraciones y resolver las ambigüedades por el contexto, agrupando en sintagmas núcleos las expresiones analizadas sucesivas. Por lo tanto, este módulo actúa en la recomposición de los tokens (agrupa o desagrupa algunos tokens previamente delimitados) y realiza una desambigüización parcial en el trabajo de ajuste entre morfología y sintaxis. Así en:

- 1) *La mira*
- 2) *La señora,*

SMORPH leerá tanto para 1) como para 2) las mismas posibilidades para ese *La*,

- a) artículo femenino singular
- b) clítico.

En MPS, en el trabajo de agrupación se formará el sintagma verbal núcleo para 1), es decir, la opción b) y el sintagma nominal núcleo para 2), o sea, la opción a).

Vemos entonces que, con este módulo, es posible efectuar ciertas tareas relevantes de la pre-sintaxis tomándose en cuenta para el análisis sintáctico algunas indicaciones de estructuración del documento. Todas las transformaciones que se realizan para adaptar la salida del analizador morfológico (SMORPH) con la entrada del analizador sintáctico se indican en el fichero de manera explícita.

¿Cómo se implementa MPS? Este útil requiere el fichero de salida del analizador morfo-sintáctico (SMORPH) en formato PROLOG⁹⁷ y además tres ficheros. Uno donde se indica cuáles son las unidades lingüísticas que **separan** las frases, otro que especifica las reglas que van a modificar la segmentación inicial hecha por SMORPH, efectuando la **recomposición** y un último fichero que contiene las reglas que permiten el ajuste entre morfología y sintaxis, es decir, la **correspondencia**.

2.1 Fichero de separación de frases

⁹⁶ MPS fue implantado en C++ en un contexto Windows por Faiza Abacci

⁹⁷ Se trata del lenguaje de programación que requiere la versión de MPS que estamos usando.

Tradicionalmente las series que se van a formar son los sintagmas, entonces el conjunto de caracteres ASCII que se estimen como separadores (símbolos de puntuación como punto, coma) y que a su vez aparecen en la salida de SMORPH son los que se declaran en MPS, como sigue a continuación:

```
%punto final%
'.'
['pf', 'EMS', 'ponc'].
```

2.2 Fichero de recomposición

Modifica la segmentación de SMORPH. Hay dos posibilidades básicas:

2.2.1 el reagrupamiento de un nuevo segmento, por ejemplo en una frase nominal como *el jefe*.

```
%1art+n da snn%
```

```
S1 [L1, 'TDET', 'art'] S2 [L2, 'EMS', 'nom'] --> S1+S2 [L1+L2, 'EMS', 'snn'].
```

```
%el jefe%
```

2.2.2 la separación de un segmento en dos segmentos como en el caso de *del* en *de+el*

La noción de **compatibilidad** nos dice que la descripción presente en la salida de SMORPH debe ser compatible con la descripción lingüística presente en la regla.

Etiquetas de SMORPH -----> Recomposición de los
segmentos

Parte izquierda compatibilidad Parte derecha

La recomposición permite realizar las tareas de pre-sintaxis, puesto que se reagrupa la serie de significantes en un segmento único. El significante resultante se obtiene por concatenación. Las reglas de recomposición actúan para eliminar las ambigüedades. En efecto, las reglas se aplican siempre y cuando haya **compatibilidad de descripción** entre al menos una de las descripciones lingüísticas asociadas a cada segmento de la salida de SMORPH con una de las descripciones de la regla. En la práctica, las reglas de recomposición incluyen *variables* que favorecen la generalización de las reglas aplicándose a un ensamble de significantes (si las variables presentes en la regla pueden unificarse con las constantes presentes en la salida de SMORPH). Las variables se expresan en mayúsculas y no se encuadran por

apóstrofes, a diferencia de las *constantes*⁹⁸. Gracias a la utilización de las variables en las reglas, pueden recomponerse, por ejemplo, los prefijos⁹⁹.

Por último, diremos que el orden de las reglas es pertinente, puesto que afecta el análisis resultante¹⁰⁰.

2.3 Fichero de correspondencia

Las tareas de correspondencia se aplican después de la recomposición. Aquí se trabaja con la información asociada a los segmentos (que pueden o no estar recompuestos). Esta información es modificada generando una información con un nuevo formato. La correspondencia, por medio de un ensamble de reglas, describe también una parte izquierda y una parte derecha. Las condiciones de aplicación de una regla de recomposición son similares a las condiciones de aplicación de una regla de correspondencia. Pero las reglas de correspondencia (a diferencia de las reglas de recomposición) no contienen en su parte izquierda ninguna variable. Puede aparecer el rasgo 'LEM', cuyo valor es una suma de caracteres. La aplicabilidad de las reglas de correspondencia depende de diversas circunstancias:

- si el rasgo 'LEM' está omitido en la parte izquierda de la regla, la información contenida en esta parte es resultado de una recomposición.
- si el rasgo 'LEM' aparece en la parte izquierda de la regla, el valor del rasgo se identifica con el lema presente en la descripción de la salida de SMORPH o con la descripción del resultado de la recomposición.

Al aplicarse una de las reglas de correspondencia, la descripción lingüística de la parte izquierda se reemplaza por la descripción de la parte derecha.

⁹⁸ Incluyo en *constantes*, el nombre de los rasgos, los valores de los rasgos, los significantes, los lemas

⁹⁹ En el Capítulo 4, se explica con más detalle lo relativo a la aplicación de MPS.

¹⁰⁰ BÈS-SOLANA (2004/b), p.10

Capítulo 4

SMORPH y MPS en el snn español

0. Introducción

SMORPH parte de una serie de datos declarados para luego delimitar en el texto el segmento asociado a la información morfosintáctica. Este Capítulo se refiere en primera instancia al fichero de modelos, a cómo se fue creando y a cuáles fueron las razones que pesaron a la hora de optar por uno u otro camino en el momento de trabajar el nombre común. Se prosigue con el fichero de entradas en relación con los modelos, y con el fichero de rasgos haciendo una breve alusión al fichero de terminaciones. Al final se toma MPS, incluyendo el modo en que se trabajaron las reglas para tratar el snn en español con algunos ejemplos. Se termina con las reglas elaboradas para MPS.

1. Aplicación de SMORPH en el snn en español

1.1 Los modelos en SMORPH para el nombre común

1.1.1 Introducción

El modo de trabajo fue el siguiente¹⁰⁰: se ingresaron cerca de 3000 nombres en entradas¹⁰¹ para los que se implementaron cuarenta y seis modelos. Estos modelos se crearon tomando en cuenta características morfológicas y observando que en algunas terminaciones la presencia de la tilde no es idéntica en plural y en singular (modelos 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 y 36). Aún cuando el diferenciar modelos de acuerdo a la tilde implicó un aumento en la cantidad de modelos, esto es imprescindible si se pretende que se detecten eventualmente errores ortográficos, pues uno de los fines que se persiguen es colaborar en la preparación de útiles informáticos que sirvan en la enseñanza de español como lengua extranjera. En el caso de encontrarse dos formas posibles, como *arterioscleorosis* y *arterioesclerosis*, se ingresaron las dos. También se tomaron palabras compuestas como *taparrabos*, *autopista*, *buscavidas*, y palabras con un prefijo fuera de la palabra como *ex_gobernador*, las que se escribieron sin espacios y unidas por un guión bajo, para que SMORPH reconozca la forma entera. La convención adoptada es ésta: cuando existe un mismo modelo que se sigue para los masculinos, los femeninos y los ambiguos los modelos correspondientes a cada uno se ubican en ese orden, por ejemplo los modelos 1 (para los masculinos), 2 (para los femeninos) y 3 (para los ambiguos), en los tres

¹⁰⁰ Debieron ingresarse también adjetivos, determinantes, numerales, adverbios, de modo de completar el input para poder trabajar el snn.

¹⁰¹ Pienso que se trata de una muestra significativa, pero sin duda es probable que al ingresar más nombres puedan precisarse más modelos.

modelos se usa @, es decir desinencia cero, para el singular y se agrega -s para el plural. A continuación, se presentan ejemplos acompañados de una breve descripción de cada modelo:

1.1.2 Modelo número uno

@n1	-0	
+@		nom/masc/sg
+s		nom/masc/pl .

Con este modelo, se cubren casos como *barrilete*, *queso*, *guiso* donde para formar el plural se necesita solo es preciso concatenar -s, mientras que para formar el singular no es necesario sustraer ni concatenar ningún carácter al lema, por eso se utiliza el signo @.

1.1.3 Modelo número dos

@n2	-0	
+@		nom/fem/sg
+s		nom/fem/pl .

Con este modelo, se cubren casos como *barrera*, *escalera*, *ensalada* donde para formar el plural solo se necesita concatenar -s, mientras que para formar el singular no es necesario concatenar ni sustraer ningún carácter al lema, por eso se utiliza el signo @. Este modelo, a diferencia del modelo uno, cubre los nombres femeninos.

1.1.4 Modelo número tres

@n3	-0	
+@		nom/_/sg
+s		nom/_/pl .

Con este modelo, se cubren casos como *ábside*, *accionista*, *acróbata* donde para formar el plural solo se necesita concatenar -s, mientras que para formar el singular no es necesario sustraer ni concatenar ningún carácter al lema, por eso se utiliza el signo @. Este modelo, a diferencia del modelo uno y del modelo dos cubre los nombres de naturaleza ambigua, es decir, los que pueden ser tanto femeninos como masculinos, por eso se utiliza _, para indicar con una variable que puede ser tanto uno u otro valor.

1.1.5 Modelo número cuatro

@n4	-1	
+o		nom/masc/sg
+a		nom/fem/sg
+os		nom/masc/pl

+as nom/fem/pl .

Este modelo abarca casos como *becerro*, *bisnieto*, *abogado* donde se distinguen el masculino singular sustrayendo un carácter al lema y sumando -o, el masculino plural con el agregado de -os, y en el caso del femenino singular y el femenino plural se concatena -a y -as respectivamente.

1.1.6 Modelo número cinco

@n5 -0
 +@ nom/masc/sg
 +es nom/masc/pl .

Con este modelo, se cubren nombres masculinos como *abridor*, *berenjena*, *temor* donde para formar el plural solo se necesita concatenar -es, mientras que para formar el singular no es necesario sustraer ni concatenar ningún carácter al lema, por eso se utiliza el signo @.

1.1.7 Modelo número seis

@n6 -0
 +@ nom/fem/sg
 +es nom/fem/pl .

Con este modelo, se cubren casos como *actitud*, *barbaridad*, *sociedad* donde para formar el plural se necesita solo concatenar -es, mientras que para formar el singular no es necesario sustraer ni concatenar ningún carácter al lema, por eso se utiliza el signo @. Este modelo, a diferencia del modelo cinco, cubre los nombres femeninos.

1.1.8 Modelo número siete

@n7 -0
 +@ nom/_/sg
 +es nom/_/pl .

Este modelo cubre casos como *bachiller*, *fiscal* donde para formar el plural solo se necesita concatenar -es, mientras que para formar el singular no es necesario sustraer ni concatenar ningún carácter al lema, por eso se utiliza el signo @. Este modelo, a diferencia del modelo cinco y del modelo seis cubre los nombres de naturaleza ambigua, es decir, los que pueden ser tanto femeninos como masculinos, por eso se utiliza _, para indicar con una variable que puede ser tanto uno u otro valor.

1.1.9 Modelo número ocho

@n8 -0

+@	nom/masc/sg
+es	nom/masc/pl
+a	nom/fem/sg
+as	nom/fem/pl .

Este modelo combina el modelo uno, el modelo cinco y el modelo cuatro, ya que para formar el masculino plural solo se necesita concatenar -es, mientras que para formar el masculino singular no es necesario sustraer ni concatenar ningún carácter al lema, por eso se utiliza el signo @, mientras que para distinguir el femenino singular se concatena -a y en el caso del femenino plural se concatena -as. En este modelo, se cubren casos como *bedel, gobernador, opositor*.

1.1.10 Modelo número nueve

@n9	-0	
+@		nom/masc/_ .

Este modelo abarca nombres masculinos como *lunes, viernes, jueves* donde tanto para formar el plural como para formar el singular no es necesario restar ni concatenar ningún carácter al lema, por eso se utiliza el signo @ y el _ para anotar que esto sucede tanto en plural como en singular.

1.1.11 Modelo número diez

@n10	-0	
+@		nom/fem/_ .

Este modelo abarca nombres femeninos como *antítesis, bronquitis, artritis* donde tanto para formar el plural como para formar el singular no es necesario sustraer ni concatenar ningún carácter al lema, por eso se utiliza el signo @ y _ para anotar que esto sucede tanto en plural como en singular.

1.1.12 Modelo número once

@n11	-0	
+@		nom/_/sg
+@		nom/_/pl .

Este modelo cubre casos como *aguafiestas, buscavidas, mandamás* donde tanto para formar el plural como para formar el singular no es necesario restar ni concatenar ningún carácter al lema, por eso se utiliza el signo @. Este modelo, a diferencia del modelo nueve y del modelo diez abarca los nombres de naturaleza ambigua, es decir, los que pueden ser tanto femeninos como masculinos, por eso se utiliza _, para indicar con una variable que puede ser tanto uno u otro valor.

1.1.13 Modelo número doce

@n12 -2
 +ón nom/masc/sg
 +ones nom/masc/pl .

Con este modelo, se cubren nombres masculinos como *bodegón*, *acordeón*, *buzón* donde para formar el plural se deben sustraer dos caracteres y sumar *-ones*, mientras que para formar el singular se concatena *-ón*. Este modelo se incluye para solucionar la presencia de la tilde en el singular y no en el plural.

1.1.14 Modelo número trece

@n13 -2
 +ón nom/fem/sg
 +ones nom/fem/pl .

Este modelo cubre nombres femeninos como *agresión*, *violación*, *presión* donde para formar el plural se deben sustraer dos caracteres y sumar *-ones*, mientras que para formar el singular se concatena *-ón*. Este modelo se incluye para solucionar la presencia de la tilde en el singular y no en el plural.

1.1.15 Modelo número catorce

@n14 -2
 +ón nom/masc/sg
 +ones nom/masc/pl
 +ona nom/fem/sg
 +onas nom/fem/pl .

Este modelo abarca nombres como *anfitrión*, *bufón*, *patrón* donde para formar el masculino plural se deben sustraer dos caracteres y concatenar *-ones*, mientras que para formar el singular se concatena *-ón*. En tanto, en el femenino se agrega *-ona* y *-onas* para el singular y plural respectivamente. Este modelo se incluye para solucionar la presencia de la tilde solamente en el masculino singular.

1.1.16 Modelo número quince

@n15 -2
 +ón nom/masc/sg
 +ones nom/masc/pl
 +onesa nom/fem/sg
 +onesas nom/fem/pl .

Con este modelo, se cubren casos como *barón*, donde para formar el masculino plural se deben sustraer dos caracteres y concatenar *-ones*, mientras que para formar el femenino singular se concatena *-onesa*. En tanto, en el femenino plural se concatena *-onesas*. Este modelo se incluye para solucionar la presencia de la tilde solamente en el masculino singular

1.1.17 Modelo número dieciséis

@n16 -2
 +án nom/masc/sg
 +anes nom/masc/pl .

Con este modelo, se cubren nombres masculinos como *ademán*, *alacrán*, *volcán*, donde para formar el plural se deben sustraer dos caracteres y concatenar *-anes*, mientras que para formar el singular se concatena *-án*. Este modelo se incluye para solucionar la presencia de la tilde en el singular y no en el plural.

1.1.18 Modelo número diecisiete

@n17 -2
 +én nom/masc/sg
 +enes nom/masc/pl .

Este modelo abarca nombres masculinos como *badén*, *andén*, *satén*, donde para formar el plural se deben sustraer dos caracteres y concatenar *-enes*, mientras que para formar el singular se concatena *-én*. Este modelo se incluye para solucionar la presencia de la tilde en el singular y no en el plural.

1.1.19 Modelo número dieciocho

@n18 -2
 +ín nom/masc/sg
 +ines nom/masc/pl .

Este modelo cubre nombres masculinos como *bergantín*, *patín*, *cojín*, donde para formar el plural se deben sustraer dos caracteres y concatenar *-ines*, mientras que para formar el singular se concatena *-ín*. Este modelo se incluye para solucionar la presencia de la tilde en el singular y no en el plural.

1.1.20 Modelo número diecinueve

@n19 -2
 +ín nom/masc/sg
 +ines nom/masc/pl
 +ina nom/fem/sg
 +inas nom/fem/pl .

Este modelo es una variante del anterior donde se cubren nombres como *benjamín*, donde para formar el masculino plural se deben sustraer dos caracteres y concatenar *-ines*, mientras que para formar el singular se concatena *-ín*. En tanto, en el femenino se concatena *-ina* e *-inas* para el singular y plural

respectivamente. Este modelo se incluye para solucionar la presencia de la tilde solamente en el masculino singular y es similar al modelo catorce.

1.1.21 Modelo número veinte

@n20	-2	
+ún		nom/masc/sg
+unes		nom/masc/pl .

Con este modelo, se cubren nombres masculinos como *atún*, *betún*, donde para formar el plural se deben sustraer dos caracteres y sumar *-unes*, mientras que para formar el singular se concatena *-ún*. Este modelo se incluye para solucionar la presencia de la tilde en el singular y no en el plural.

1.1.22 Modelo número veintiuno

@n21	-2	
+ís		nom/masc/sg
+ises		nom/masc/pl .

Este modelo cubre nombres masculinos como *anís*, donde para formar el plural se deben sustraer dos caracteres y concatenar *-ises*, mientras que para formar el singular se agrega *-ís*. Este modelo se incluye para solucionar la presencia de la tilde en el singular y no en el plural.

1.1.23 Modelo número veintidós

@n22	-2	
+ús		nom/masc/sg
+uses		nom/masc/pl .

Con este modelo, se cubren nombres masculinos como *autobús*, *patatús* donde para formar el plural se deben sustraer dos caracteres y sumar *-uses*, mientras que para formar el singular se concatena *-ús*. Este modelo se incluye para solucionar la presencia de la tilde en el singular y no en el plural.

1.1.24 Modelo número veintitrés

@n23	-4	
+omen		nom/masc/sg
+ómenes		nom/masc/pl .

En este modelo, se cubren nombres masculinos como *abdomen*, donde para formar el plural se deben sustraer cuatro caracteres y concatenar *-ómenes*, mientras que para formar el singular se concatena *-omen*. Este modelo, a diferencia de los modelos doce, trece, catorce, quince, dieciséis, diecisiete,

dieciocho, diecinueve, veinte, veintiuno, veintidós se incluye para solucionar la presencia de la tilde en el plural y no en el singular.

1.1.25 Modelo número veinticuatro

@n24 -4
 +umen nom/masc/sg
 +úmenes nom/masc/pl .

Este modelo abarca nombres masculinos como *albumen*, *cardumen*, *resumen* donde para formar el plural se deben sustraer cuatro caracteres y sumar *-úmenes*, mientras que para formar el singular se concatena *-umen*. Este modelo, de manera similar al modelo veintitrés, se incluye para solucionar la presencia de la tilde en el plural y no en el singular.

1.1.26 Modelo número veinticinco

@n25 -6
 +égimen nom/masc/sg
 +egímenes nom/masc/pl .

En este modelo, se abarca nombres masculinos como *régimen* donde para formar el plural se deben sustraer seis caracteres al lema y concatenar *-egímenes*, mientras que para formar el singular se concatena *-égimen*. Este modelo, a diferencia de los modelos veintitrés y veinticuatro, se incluye para solucionar la presencia de la tilde en distintos lugares en el singular y en el plural.

1.1.27 Modelo número veintiséis

@n26 -0
 +@ nom/masc/sg .

Este modelo cubre nombres masculinos singulares como *boldo*, *abor* donde para formar el singular no es necesario sustraer ni concatenar ningún carácter, por eso se utiliza el signo @. Estos nombres no se utilizan en plural.

1.1.28 Modelo número veintisiete

@n27 -0
 +@ nom/fem/sg .

Con este modelo, se cubren nombres femeninos singulares como *beneficencia*, *salud* donde para formar el singular no es necesario sustraer ni concatenar ningún carácter, por eso se utiliza el signo @. Estos nombres no se utilizan en plural.

1.1.29 Modelo número veintiocho

@n28 -1
 +z nom/masc/sg
 +ces nom/masc/pl .

Con este modelo, se cubren nombres masculinos como *barniz*, *tapiz*, *arroz* donde para formar el plural se necesita sustraer un carácter al lema y concatenar -ces, mientras que para formar el singular es necesario concatenar -z.

1.1.30 Modelo número veintinueve

@n29 -1
 +z nom/fem/sg
 +ces nom/fem/pl .

Este modelo abarca nombres femeninos como *actriz*, *bisectriz*, *vez* donde para formar el plural se necesita sustraer un carácter al lema y concatenar -ces, mientras que para formar el singular es necesario concatenar -z.

1.1.31 Modelo número treinta

@n30 -1
 +z nom/_/sg
 +ces nom/_/pl .

Este modelo cubre nombres como *aprendiz* donde para formar el masculino singular o femenino singular es necesario concatenar -z. Ahora bien, en el caso del plural se usa _, para indicar que puede asumir cualquiera de los dos valores de género, para formarlos se concatena -ces.

1.1.32 Modelo treinta y uno

@n31 -0
 +@ nom/masc/pl .

Este modelo abarca nombres masculinos plurales como *anales*, *ambages*, *taparrabos* donde para formar el plural no es necesario sustraer ni concatenar ningún carácter, por eso se utiliza el signo @. Estos nombres no se utilizan en singular. Es la contrapartida del modelo veintiséis.

1.1.33 Modelo treinta y dos

@n32 -0
 +@ nom/fem/pl .

Este modelo abarca nombres femeninos plurales como *bombachas*, *postrimerías* donde para formar el plural no es necesario sustraer ni sumar ningún

carácter, por eso se utiliza el signo @. Estos nombres no se utilizan en singular. Es la contrapartida del modelo veintisiete.

1.1.34 Modelo treinta y tres

@n33	-0	
+@		nom/masc/sg
+es		nom/masc/pl
+esa		nom/fem/sg
+esas		nom/fem/pl .

Con este modelo, se cubren nombres como *abad* donde para formar el masculino plural solo se necesita concatenar *-es*, mientras que para formar el masculino singular no es necesario sustraer ni concatenar ningún carácter al lema, por eso se utiliza el signo @. En tanto, para formar los femeninos, se concatena *-esa* y *-esas*, singular y plural respectivamente. Este modelo es una variante de los modelos cinco, seis y quince.

1.1.35 Modelo treinta y cuatro

@n34	-0	
+@		nom/masc/sg
+s		nom/masc/pl
+sa		nom/fem/sg
+sas		nom/fem/pl .

Este modelo abarca nombres como *alcalde*, *duque*, *conde* donde para formar el masculino plural solo se necesita concatenar *-s*, mientras que para formar el masculino singular no es necesario sustraer ni concatenar ningún carácter al lema, por eso se utiliza el signo @. En tanto, para formar los femeninos, se concatena *-sa* y *-sas*, singular y plural respectivamente. Este modelo es una variante del modelo treinta y cuatro.

1.1.36 Modelo treinta y cinco

@n35	-2	
+és		nom/masc/sg
+eses		nom/masc/pl
+esa		nom/fem/sg
+esas		nom/fem/pl .

Este modelo cubre nombres como *marqués* donde para formar el masculino plural se necesita sustraer dos caracteres al lema y concatenar *-eses*, mientras que para formar el masculino singular se concatena *-és*. En tanto, para formar los femeninos, se concatena *-esa* y *-esas*, singular y plural respectivamente. Este modelo es una variante de los modelos treinta y cuatro y treinta y cinco. Se incluye

para solucionar la presencia de la tilde en el masculino singular. Este modelo se aplicaría para nombres adjetivos como *libanés*, *maltés*, *japonés*.

1.1.37 Modelo treinta y seis

@n36	-4	
+amen		nom/masc/sg
+ámenes		nom/masc/pl .

Este modelo cubre nombres como *examen* donde para formar el masculino plural se necesita sustraer cuatro caracteres al lema y concatenar *-ámenes*, mientras que para formar el masculino singular se concatena *-amen*.

1.1.38 Modelo treinta y siete

@n37	-5	
+argen		nom/masc/sg
+árgenes		nom/masc/pl .

Este modelo cubre nombres como *margen* donde para formar el masculino plural se necesita sustraer cinco caracteres al lema y concatenar *-árgenes*, mientras que para formar el masculino singular se concatena *-argen*.

1.1.39 Modelo treinta y ocho

@n38	-4	
+igen		nom/masc/sg
+ígenes		nom/masc/pl .

Este modelo cubre nombres como *origen* donde para formar el masculino plural se necesita sustraer cinco caracteres al lema y concatenar *-ígenes*, mientras que para formar el masculino singular se concatena *-igen*.

1.1.40 Modelo treinta y nueve

@n39	-2	
+ás		nom/masc/sg
+áses		nom/masc/pl .

Este modelo cubre nombres como *compás* donde para formar el masculino plural se necesita sustraer dos caracteres al lema y concatenar *-ases*, mientras que para formar el masculino singular se concatena *-ás*.

1.1.41 Modelo cuarenta

@n40	-1	
+e		nom/masc/sg

+es	nom/masc/pl
+a	nom/fem/sg
+as	nom/fem/pl .

Este modelo cubre nombres como *nene* donde para formar el masculino plural se necesita sustraer un carácter al lema y concatenar *-es*, mientras que para formar el masculino singular se concatena *-e*. En tanto está la posibilidad de concatenar *-a* y *-as* para formar el femenino singular y el femenino plural, respectivamente.

1.1.42 Modelo cuarenta y uno

@n41	-4	
+imen		nom/masc/sg
+ímenes		nom/masc/pl .

Este modelo cubre nombres como *crimen* donde para formar el masculino plural se necesita sustraer cuatro caracteres al lema y concatenar *-ímenes*, mientras que para formar el masculino singular se concatena *-imen*.

1.1.43 Modelo cuarenta y dos

@n42	-4	
+ogan		nom/masc/sg
+órganes		nom/masc/pl .

Este modelo cubre nombres como *eslogan* donde para formar el masculino plural se necesita sustraer cuatro caracteres al lema y concatenar *-órganes*, mientras que para formar el masculino singular se concatena *-ogan*.

1.1.44 Modelo cuarenta y tres

@n43	-5	
+oquin		nom/masc/sg
+óquines		nom/masc/pl .

Este modelo cubre nombres como *esmoquin* donde para formar el masculino plural se necesita sustraer cinco caracteres al lema y concatenar *-óquines*, mientras que para formar el masculino singular se concatena *-oquin*.

1.1.45 Modelo cuarenta y cuatro

@n44	-2	
+és		nom/masc/sg
+eses		nom/masc/pl .

Este modelo cubre nombres como *interés* donde para formar el masculino plural se necesita sustraer cinco caracteres al lema y concatenar *-eses*, mientras que para formar el masculino singular se concatena *-és*.

1.1.46 Modelo cuarenta y cinco

@n45	-4	
+agen		nom/fem/sg
+ágenes		nom/fem/pl .

Este modelo cubre nombres como *imagen* donde para formar el femenino plural se necesita sustraer cuatro caracteres al lema y concatenar *-ágenes*, mientras que para formar el masculino singular se concatena *-agen*.

1.1.47 Modelo cuarenta y seis

@n46	-1	
+y		nom/masc/sg
+yes		nom/masc/pl
+ina		nom/fem/sg
+inas		nom/fem/pl .

Este modelo cubre nombres como *rey* donde para formar el masculino plural se necesita sustraer un carácter al lema y concatenar *-yes*, mientras que para formar el masculino singular se concatena *-y*. En tanto, para formar el femenino singular y plural respectivamente se concatena, *ina* e *inas*.

1.1.48 Comentarios acerca de la observación de la relación de entradas con los respectivos modelos

Al observar la relación de los modelos con las entradas, se nota que:

- los modelos con mayor cantidad de nombres comunes son los modelos uno y dos, es decir, los que implican el agregado de una *-s* para pluralizar, en su versión masculina y femenina. Estos abarcan en conjunto poco más de la mitad del total,

- los modelos que agregan *-es* para pluralizar (cinco, seis y siete) le siguen en orden de importancia, correspondiendo la mayoría al modelo masculino cinco,

- los nombres que no tienen plural (modelos veintiséis y veintisiete) aparecen también con un gran número,

- entre los modelos que presuponen modificación de la tilde, se destacan el doce y el trece (los que agregan *-ón*, *-ones*). Entre estos últimos, merece destacar que los femeninos superan a los masculinos, en ese único caso.

1.2 Los rasgos en el snn en español

Los rasgos que voy a usar en esta propuesta son¹⁰²:

EMS	etmorsin
adv	adver
nom	nombr
det	determ
adj	adje
prep	prepos
pron	pronom
num	numer
coma	coma
ponc	ponctu .

GEN	género
masc	masc
fem	fem
_ ¹⁰³	masc o fem
neut	neutro .

NUM	numero
sg	sing
pl	plur
_ ¹⁰⁴	sing o plur .

TNOM	tipnombr
npr	nprop
nc	ncom .

TDET	tipdet
art	artic
dem	demos
pos	poses
indf	indef
amb	ambos
tod	todo
send	sendos
cad	cada
ciert	ciertos
tant	tantos
mism	mismos
con	contrac . #contracción : al, del...

TINDF	tipindf
-------	---------

¹⁰² Para las abreviaturas de los rasgos, ver nota 24 del Cap. 2

¹⁰³ El _ anota una variable, puede ser tanto masculino como femenino.

¹⁰⁴ También aquí el _ anota una variable, en este caso puede ser tanto singular como plural.

indf1 indef1
indf2 indef2 .

TINDF1 tipindf1
indf1a indef1a
indf1b indef1b .

TINDF2 tipindf2
indf2a indef2a
indf2b indef2b
indf2c indef2c .

TNUM tipnumer
num1 numer1
num2 numer2 .

TADJ tipadj
adj1 adjet1
adj2 adjet2 .

TPRON tippron
prpers pronpers
prindf pronindf
prdem prondem
prpos pronpos .

1.3 Las terminaciones en el snn en español

Distingo entre terminaciones y terminaciones distinguidas, las que necesitan una información adicional.

1.3.1 Terminaciones

@ .
o .
a .
os .
as .
s .
es .
z .
ces .
za .
és .
eses .
esa .
esas .

1.3.2 Terminaciones distinguidas

ón nom .
 ones nom .
 ona nom .
 onas nom .
 onesa nom .
 onesas nom .
 omen nom .
 ómenes nom .
 égimen nom .
 egímenes nom .
 án nom .
 anes nom .
 én nom .
 enes nom .
 ín nom .
 ines nom .
 ina nom .
 inas nom .
 ún nom .
 unes nom .
 ís nom .
 ises nom .
 ús nom .
 uses nom .
 umen nom .
 úmenes nom .

Se concluye aquí con la exposición acerca de cómo se aplicó SMORPH para tratar el snn en forma automática, dando cuenta de los ficheros de entradas, modelos, rasgos y terminaciones. En lo que sigue, vamos a ver qué es lo que ocurre con MPS.

2. Aplicación de MPS para el snn en español

Como ya se mencionó en el Capítulo 3, la entrada para MPS la proporciona el software SMORPH (segmentación y morfología), que trata el análisis presintáctico en una sola etapa. La herramienta informática MPS¹⁰⁵ actúa en una organización modular con SMORPH. Recibe en entrada una salida SMORPH (en formato Prolog) y va a dar en salida otro formato según el analizador que se vaya a utilizar; MPS puede modificar las estructuras de datos que se reciben en la entrada. Ya

¹⁰⁵ MPS ha sido extendido en Pasmó (cf. HAGÉGE, Carolina, NAMEDE, Nuno & PAULO, Joana, Pasmó-análise morfológica, Technical report, INESC-ID, Lisboa), que posee otras funcionalidades.

vimos que las funciones principales que ejecuta MPS son: la Recomposición (Reagrupamiento o División) y la Correspondencia.

En virtud de la característica de **declaratividad** de este módulo, ya indicada en el Capítulo 3, es posible expresar mediante reglas los valores de entrada (sobre dos o más estructuras de datos de la salida SMORPH) y los valores de salida sobre la estructura reagrupada. Las reglas que se declaren con la función de división van a provocar el efecto inverso. Como ya se vio, con ellas se tratan las contracciones (por ejemplo, una ocurrencia 'del' en español¹⁰⁶).

Para la detección de snn, recorro a las reglas de recomposición (fundamentalmente de reagrupamiento) mediante las cuales se especifica qué categorías pueden reagruparse constituyendo un snn. Aquí se presentan algunos pocos ejemplos de estas reglas y, a continuación, las reglas en su totalidad. Estas reglas, pensadas para la extracción, son las que permiten quitar la ambigüedad de determinadas ocurrencias lingüísticas.

A título ilustrativo, se muestran algunas reglas de formación de snn (la entrada de la regla se escribe antes de la flecha y la salida después)¹⁰⁷. En las páginas siguientes, aparecen todas las reglas implementadas para el snn.

a) Con esta regla pueden señalarse un snn como por ejemplo: *'esa querida amiga'*

S1 [L1, 'TDET', 'dem'] S2 [L2, 'EMS', 'adj1'] S3 [L3, 'EMS', 'nom'] --> S1+S2+S3
[L1+L2+L3, 'EMS', 'snn'].

Esto se explica de la siguiente manera:

S1, S2, S3, son variables sobre las ocurrencias.

L1, L2, L3, son variables sobre los lemas en el sentido en que se usa en lingüística computacional.

EMS etiqueta morfosintáctica, allí aparecen 'adj' la del adjetivo y 'nom' la del nombre.

TDET: tipo de determinante, para 'dem' demostrativo. Se trata de rasgos/valores + indica la operación de concatenación¹⁰⁸.

La regla dice que una ocurrencia analizada como 'dem' seguida de otra analizada como 'adj1' y de otra analizada como 'nom' se concatenan en snn.¹⁰⁹

b) Con la siguiente regla, pueden indicarse un snn como: *'el segundo estante'*

S1 [L1, 'TDET', 'art'] S2 [L2, 'TNUME', 'num2'] S3 [L3, 'EMS', 'nom'] --> S1+S2+S3
[L1+L2+L3, 'EMS', 'snn'].

¹⁰⁶ Ver Capítulo 3.

¹⁰⁷ Ya habíamos hecho una previa presentación de estas reglas en el Capítulo 3.

¹⁰⁸ En esta explicación sigo a BÈS- SOLANA (2004/b)

¹⁰⁹cf. Las indicaciones sobre el trabajo de MPS "MPS compara una oa en la entrada a analizar con el primer elemento de cada una de sus reglas. Si encuentra una regla cuyo primer elemento subsume el oa, verifica el oa siguiente con el elemento siguiente en la entrada de la regla, y así siguiendo. MPS aplica la primer regla en cuya entrada todos los elementos subsumen un oa en la entrada a analizar" cf. BÈS-SOLANA (2004/b)

Aquí queda por explicar

'TNUME': tipo de numeral y 'num2' correspondiente a numeral ordinal.

La regla dice que una ocurrencia analizada como 'art' seguida de otra analizada como 'num2' y de otra analizada como 'nom' se concatenan en snn.

Ya se ha mostrado cómo se logra la extracción de snn en español y de qué manera trabaja MPS. Sigo con las cincuenta y tres reglas que implemento para trabajar con este módulo.

Las reglas para MPS

A continuación, se exponen las reglas de recomposición (RCM) elaboradas para la detección del snn sujeto:

%1 art+adj1+npr2 da snn%

S1 [L1, 'TDET', 'art'] S2 [L2, 'EMS', 'adj1'] S3 [L3, 'TNOM', 'npr'] --> S1+S2+S3 [L1+L2+L3, 'EMS', 'snn'].

%la lejana Asia, el caudaloso Paraná%

%2art+npr2 da snn%

S1 [L1, 'TDET', 'art'] S2 [L2, 'TNOM', 'npr'] --> S1+S2 [L1+L2, 'EMS', 'snn'].

%el Asia%

%3 npr+npr da snn%

S1 [L1, 'TNOM', 'npr'] S2 [L2, 'TNOM', 'npr'] --> S1+S2 [L1+L2, 'EMS', 'snn'].

%Carlos Juárez%

%4npr da snn%

S1 [L1, 'TNOM', 'npr'] --> S1 [L1, 'EMS', 'snn'].

%Carlos%

%5art+adv+adj1+n da snn%

S1 [L1, 'TDET', 'art'] S2 [L2, 'EMS', 'adv'] S3 [L3, 'EMS', 'adj1'] S4 [L4, 'EMS', 'nom'] -> S1+S2+S3+S4 [L1+L2+L3+L4, 'EMS', 'snn'].

%el más notable elogio, la aún abierta crisis %

%6indf+adv+adj1+n da snn%

S1 [L1, 'TDET', 'indf'] S2 [L2, 'EMS', 'adv'] S3 [L3, 'EMS', 'adj1'] S4 [L4, 'EMS', 'nom'] --> S1+S2+S3+S4 [L1+L2+L3+L4, 'EMS', 'snn'].

% un mal entendido asistencialismo%

%7indf+adj+adj da snn%

S1 [L1, 'TDET', 'indf'] S2 [L2, 'EMS', 'adj'] S3 [L3, 'EMS', 'adj'] --> S1+S2+S3 [L1+L2+L3, 'EMS', 'snn'].

% un potencial sangrado%

%8pos+adv+adj1+n da snn%

S1 [L1, 'TDET', 'pos'] S2 [L2, 'EMS', 'adv'] S3 [L3, 'EMS', 'adj1'] S4 [L4, 'EMS', 'nom']
--> S1+S2+S3+S4 [L1+L2+L3+L4, 'EMS', 'snn'].

%su demasiado visible interés%

%9art+adv+adj1 da snn%

S1 [L1, 'TDET', 'art'] S2 [L2, 'EMS', 'adv'] S3 [L3, 'EMS', 'adj1'] --> S1+S2+S3
[L1+L2+L3, 'EMS', 'snn'].

%la muy tonta, lo verdaderamente grave%

%10art+adv+adj1+n da snn%

S1 [L1, 'TDET', 'art'] S2 [L2, 'EMS', 'adv'] S3 [L3, 'EMS', 'adj1'] S4 [L4, 'EMS', 'nom']
--> S1+S2+S3+S4 [L1+L2+L3+L4, 'EMS', 'snn'].

%el recién aprobado presupuesto%

%11art+num1+num2+n da snn%

S1 [L1, 'TDET', 'art'] S2 [L2, 'TNUME', 'num1'] S3 [L3, 'TNUME', 'num2'] S4 [L4, 'EMS', 'nom'] --> S1+S2+S3+S4 [L1+L2+L3+L4, 'EMS', 'snn'].

%las dos primeras casas%

%12dem+num1+num2+n da snn%

S1 [L1, 'TDET', 'dem'] S2 [L2, 'TNUME', 'num1'] S3 [L3, 'TNUME', 'num2'] S4 [L4, 'EMS', 'nom'] --> S1+S2+S3+S4 [L1+L2+L3+L4, 'EMS', 'snn'].

%esas dos primeras casas%

%13pos+num1+num2+n da snn%

S1 [L1, 'TDET', 'pos'] S2 [L2, 'TNUME', 'num1'] S3 [L3, 'TNUME', 'num2'] S4 [L4, 'EMS', 'nom'] --> S1+S2+S3+S4 [L1+L2+L3+L4, 'EMS', 'snn'].

%sus dos primeras casas%

%14art+num1+num2da snn%

S1 [L1, 'TDET', 'art'] S2 [L2, 'TNUME', 'num1'] S3 [L3, 'TNUME', 'num2'] --> S1+S2+S3 [L1+L2+L3, 'EMS', 'snn'].

%los dos primeros%

%15pos+adv+adj1+n da snn%

S1 [L1, 'TDET', 'pos'] S2 [L2, 'EMS', 'adv'] S3 [L3, 'EMS', 'adj'] S4 [L4, 'EMS', 'nom'] --> S1+S2+S3+S4 [L1+L2+L3+L4, 'EMS', 'snn'].

%su muy recordado padre%

%16art+prpos da snn%

S1 [L1, 'TDET', 'art'] S2 [L2, 'TPRON', 'prpos'] --> S1+S2 [L1+L2, 'EMS', 'snn'].

%la mía%

%17art+adj1+adj1 da snn%

S1 [L1, 'TDET', 'art'] S2 [L2, 'EMS','adj1'] S3 [L3, 'EMS','adj1'] --> S1+S2+S3
 [L1+L2+L3, 'EMS', 'snn'].
 %la pequeña tonta%

%18art+adj1+adj2 da snn%
 S1 [L1, 'TDET', 'art'] S2 [L2, 'EMS','adj1'] S3 [L3, 'EMS','adj2'] --> S1+S2+S3
 [L1+L2+L3, 'EMS', 'snn'].
 %la agitada departamental%

%19art+adj2 da snn%
 S1 [L1, 'TDET', 'art'] S2 [L2, 'EMS','adj2']--> S1+S2 [L1+L2, 'EMS', 'snn'].
 %la departamental%

%20indf2 + indf2 da snn%
 S1 [L1, 'TDET', 'indf2'] S2 [L2, 'TDET', 'indf2'] --> S1+S2 [L1+L2, 'EMS', 'snn'].
 %algunos pocos, muchas otras%

%21dem+adj1+n da snn%
 S1 [L1, 'TDET', 'dem'] S2 [L2, 'EMS','adj1'] S3 [L3, 'EMS','nom'] --> S1+S2+S3
 [L1+L2+L3, 'EMS', 'snn'].
 %aquella antigua casa%

%22pos+num1+n da snn%
 S1 [L1, 'TDET', 'pos'] S2 [L2, 'TNUME','num1'] S3 [L3, 'EMS','nom'] --> S1+S2+S3
 [L1+L2+L3, 'EMS', 'snn'].
 %tus dos casas%

%23pos+num2 +n da snn%
 S1 [L1, 'TDET', 'pos'] S2 [L2, 'TNUME','num2'] S3 [L3, 'EMS','nom'] --> S1+S2+S3
 [L1+L2+L3, 'EMS', 'snn'].
 %su primera casa%

%24dem+num1+n da snn%
 S1 [L1, 'TDET', 'dem'] S2 [L2, 'TNUME','num1'] S3 [L3, 'EMS','nom'] --> S1+S2+S3
 [L1+L2+L3, 'EMS', 'snn'].
 %aquellas dos casas%

%25dem+num2+n da snn%
 S1 [L1, 'TDET', 'dem'] S2 [L2, 'TNUME','num2'] S3 [L3, 'EMS','nom'] --> S1+S2+S3
 [L1+L2+L3, 'EMS', 'snn'].
 %aquella primera casa%

%26art+num1+adj1+n da snn%
 S1 [L1, 'TDET', 'art'] S2 [L2, 'TNUME', 'num1'] S3 [L3, 'EMS','adj1'] S4 [L4,
 'EMS','nom'] --> S1+S2+S3+S4 [L1+L2+L3+L4, 'EMS', 'snn'].
 %los dos pequeños niños%

%27art+adj1+n da snn%

S1 [L1, 'TDET', 'art'] S2 [L2, 'EMS', 'adj1'] S3 [L3, 'EMS', 'nom'] --> S1+S2+S3
[L1+L2+L3, 'EMS', 'snn'].

%la antigua casa%

%28indf2+adj1+n da snn%

S1 [L1, 'TDET', 'indf2'] S2 [L2, 'EMS', 'adj1'] S3 [L3, 'EMS', 'nom'] --> S1+S2+S3
[L1+L2+L3, 'EMS', 'snn'].

%una hermosa casa%

%29art+num1+n da snn%

S1 [L1, 'TDET', 'art'] S2 [L2, 'TNUME', 'num1'] S3 [L3, 'EMS', 'nom'] -->
S1+S2+S3 [L1+L2+L3, 'EMS', 'snn'].

%las tres casas%

%30art+num2+n da snn%

S1 [L1, 'TDET', 'art'] S2 [L2, 'TNUME', 'num2'] S3 [L3, 'EMS', 'nom'] --> S1+S2+S3
[L1+L2+L3, 'EMS', 'snn'].

%la primera casa%

%31art+n da snn%

S1 [L1, 'TDET', 'art'] S2 [L2, 'EMS', 'nom'] --> S1+S2 [L1+L2, 'EMS', 'snn'].

%la casa%

%32pos+n da snn%

S1 [L1, 'TDET', 'pos'] S2 [L2, 'EMS', 'nom'] --> S1+S2 [L1+L2, 'EMS', 'snn'].

%su casa%

%33dem+n da snn%

S1 [L1, 'TDET', 'dem'] S2 [L2, 'EMS', 'nom'] --> S1+S2 [L1+L2, 'EMS', 'snn'].

%esta casa%

%34indf+n da snn%

S1 [L1, 'TDET', 'indf'] S2 [L2, 'EMS', 'nom'] --> S1+S2 [L1+L2, 'EMS', 'snn'].

%una, alguna casa%

%35num1+n da snn%

S1 [L1, 'TNUME', 'num1'] S2 [L2, 'EMS', 'nom'] --> S1+S2 [L1+L2, 'EMS', 'snn'].

%dos casas%

%36toda+n da snn%

S1 [L1, 'TDET', 'tod'] S2 [L2, 'EMS', 'nom'] --> S1+S2 [L1+L2, 'EMS', 'snn'].

%toda casa%

%37cierta+n da snn%

S1 [L1, 'TDET', 'ciert'] S2 [L2, 'EMS', 'nom'] --> S1+S2 [L1+L2, 'EMS', 'snn'].

%cierta casa%

%38cada+n da snn%

S1 [L1, 'TDET', 'cad'] S2 [L2, 'EMS', 'nom'] --> S1+S2 [L1+L2, 'EMS', 'snn'].

%cada casa%

%39ambos+n da snn%

S1 [L1, 'TDET', 'amb'] S2 [L2, 'EMS', 'nom'] --> S1+S2 [L1+L2, 'EMS', 'snn'].

%ambas casas%

%40dem+num1 da snn%

S1 [L1, 'TDET', 'dem'] S2 [L2, 'TNUME', 'num1'] --> S1+S2 [L1+L2, 'EMS', 'snn'].

%estas dos%

%41indf da snn%

S1 [L1, 'TDET', 'indf'] --> S1 [L1, 'EMS', 'snn'].

%algunos%

%42dem da snn%

S1 [L1, 'TDET', 'dem'] --> S1 [L1, 'EMS', 'snn'].

%estos%

%43num1 da snn%

S1 [L1, 'TNUME', 'num1'] --> S1 [L1, 'EMS', 'snn'].

%dos%

%44todos da snn%

S1 [L1, 'TDET', 'tod'] --> S1 [L1, 'EMS', 'snn'].

%todos%

%45ambos da snn%

S1 [L1, 'TDET', 'amb'] --> S1 [L1, 'EMS', 'snn'].

%ambos%

%46art+num1 da snn%

S1 [L1, 'TDET', 'art'] S2 [L2, 'TNUME', 'num1'] --> S1+S2 [L1+L2, 'EMS', 'snn'].

%las dos%

%47rt+num2 da snn%

S1 [L1, 'TDET', 'art'] S2 [L2, 'TNUME', 'num2'] --> S1+S2 [L1+L2, 'EMS', 'snn'].

%la primera%

%48 art+adj da snn%

S1 [L1, 'TDET', 'art'] S2 [L2, 'EMS', 'adj'] --> S1+S2 [L1+L2, 'EMS', 'snn'].

%la tonta%

%49adj1+n da snn%

S1 [L1, 'EMS', 'adj1'] S2 [L2, 'EMS', 'nom'] --> S1+S2 [L1+L2, 'EMS', 'snn'].

%hermosas casas%

%50send+n da snn%

S1 [L1,'TDET','send'] S2 [L2, 'EMS', 'nom'] --> S1+S2 [L1+L2, 'EMS', 'snn'].
 %sendas casas%

%51art+mism+n da snn%

S1 [L1,'TDET','art'] S2 [L2, 'TDET', 'mism'] S3 [L3, 'EMS', 'nom'] --> S1+S2
 [L1+L2, 'EMS', 'snn'].
 %las mismas casas%

%52n da snn%

S1 [L1,'EMS','nom'] --> S1 [L1, 'EMS', 'snn'].
 %casas%

%53antos da snn%

S1 [L1,'TDET','tant'] --> S1 [L1, 'EMS', 'snn'].
 %tantos%

Capítulo 5

Resultados y conclusiones

0. Introducción

En este Capítulo, previo análisis automático de un texto periodístico^{110 111} por las herramientas informáticas ya presentadas teóricamente en el Capítulo 3 y en cuanto a su aplicación en el Capítulo 4, se exponen en una primera instancia los resultados obtenidos, luego se expresan en un cuadro los márgenes de cobertura y de precisión de MPS al analizar los snn sujeto. En una segunda instancia, se presentan las conclusiones.

1. Observaciones

1.1 Texto trabajado

LA SALUD DEL PRESIDENTE: LLEVA TRES DIAS HOSPITALIZADO

[Kirchner] evoluciona bien, pero sigue internado en Santa Cruz

Ya se le retiró el suero y está siendo medicado por vía oral. [Su médico] admitió que [la recuperación] llevará días o semanas. Pero no aclaró cuándo tendrá el alta.

[Lanús] jugó a favor de la recuperación del Presidente. Lo dijo ayer [el médico] personal de Néstor Kirchner al destacar su “evolución satisfactoria” de la fuerte descompostura estomacal por la que debió ser internado el jueves a la noche. De todos modos, afirmó que todavía no puede anticipar cuándo le darán el alta.

Más allá de la broma, [el médico] [Luis Buonomo] señaló que [el triunfo] del viernes de Racing, equipo del que [Kirchner] es hincha, sirvió para levantarle el ánimo y destacó que no ha evidenciado ningún tipo de complicaciones ni sobresaltos.

Pese a esto, [la fecha] en la que [Kirchner] podrá dejar el Hospital Regional de Río Gallegos sigue siendo una incógnita. Es un tema delicado, porque [la salud]

¹¹⁰ Clarín 11/04/04

¹¹¹ El análisis completo se expone en el Anexo final.

presidencial es una cuestión de Estado pero también a causa de la personalidad de Kirchner, que se resiste a aceptar mansamente las indicaciones médicas. Es por eso que [el médico] reiteró ayer que "si no hay complicaciones, [los tiempos] de recuperación para un paciente común son entre 7 y 10 días, porque [ese] es el período en el que se puede producir un nuevo y potencial sangrado".

Pero reconoció que por la personalidad y responsabilidad de Kirchner existe [una negociación]. [El Presidente] pretende dejar el hospital lo antes posible y [el médico] lo único que se animó a decir es que [el alta] no será hoy. "Las Pascuas las va a pasar acá", dijo.

Para referirse a la evolución del Presidente, [el médico] dijo que ya se le retiró el suero y que se lo está medicando por vía oral. De todos modos, agregó que la recuperación] completa llevará días o semanas.

[La idea] es que esté desestresado y que si vuelve a trabajar en los próximos días lo haga con la menor cantidad de presión posible. [Mi recomendación] serían 30 días de reposo psicofísico, pero sé que no va a ser posible, afirmó [Buonomo] resignado.

[El Presidente] sufrió una gastroduodenitis erosiva aguda con hemorragia, una lesión sangrante en el duodeno, producida por un antiinflamatorio recetado para calmarle un dolor de un tratamiento dental. Fue internado de urgencia el jueves a la noche en El Calafate, donde pensaba descansar durante el fin de semana largo, y el viernes fue trasladado a Río Gallegos, donde recibió una transfusión.

[Los dos partes] médicos de ayer -a las 10 y a las 19- destacaron su evolución favorable. Allí se aseguró que pasó una noche sin complicaciones.

Lo acompañaron [algunos amigos] porque [su esposa] [la senadora] Cristina Fernández de Kirchner, se fue a dormir a su casa, ubicada a sólo cuatro cuadras del hospital. Ayer lo fue a visitar después del mediodía pero entró por una puerta lateral para eludir a los periodistas que hacían guardia en la acceso principal.

[El primero] en visitarlo ayer fue el secretario de Legal y Técnica, Carlos Zanini. También estuvieron a su lado [su hijo] [Máximo] - [su hija] [Florencia] está en Buenos Aires - [su secretario] privado, [Daniel Muñoz], y [el gobernador] de Santa Cruz, [Sergio Acevedo].

No pudieron verlo [muchos] más porque tiene las visitas restringidas. Más allá de la guardia periodística, en la puerta del hospital no hubo manifestaciones y en Río Gallegos se respiró el clima de un sábado más.

[El médico] personal de Kirchner señaló que se le recetó un plan de nutrición para este tipo de patologías y reconoció que [la nutricionista] del hospital "no es Petrona C. de Gandulfo". [El menú] permitido incluye nada más que carnes desgrasadas, verduras hervidas y gelatinas.

Frente a una pregunta puntual, y para que no se empiecen a hablar tonterías, [Buonomo] aclaró que [Kirchner] "nunca tuvo una afección estomacal, con el único antecedente de un colon irritable".

Siempre ejerció y está ejerciendo el mando en contacto permanente con el Gabinete y tomando decisiones, dijo [el médico]. Sin embargo, [los funcionarios] que lo visitaron ayer afirmaron que lo mantienen alejado del teléfono.

Hoy a las 10 habrá un nuevo parte médico. A las 11, [el obispo] de Santa Cruz, [Alejandro Buccolini], dará una misa para los internados del hospital, por la Pascua. [Nadie] se animó a anticipar si participará [el Presidente].

1.2 Descripción del texto

Se trata de un texto periodístico, seleccionado al azar entre los que constituyen el corpus de INFOSUR y en el que se da por sentado que no hay desviaciones gramaticales de ningún tipo. Consta de setecientos treinta y un palabras y de treinta y seis oraciones. Sintácticamente, posee variedad de oraciones tanto simples como compuestas. Cabe mencionar que en determinado momento del texto aparece una incidental, marcada con guiones:

“estuvieron a su lado su hijo Máximo – su hija Florencia está en Buenos Aires – su secretario privado”...

Y además, se registran numerosas aposiciones:

- “su esposa, la senadora”,
- “su hijo Máximo”,
- “su hija Florencia”,
- “su secretario privado Daniel Martínez”,
- “el gobernador de Santa Cruz, Sergio Acevedo”,
- “el obispo de Santa Cruz, Alejandro Buccolini”

Entre corchetes se observan los snn núcleos a los que se ha abocado esta tesis, aquellos cuyo núcleo es núcleo del sujeto tanto en oraciones principales como subordinadas o coordinadas y que, según se ve en el Anexo final, son detectados por MPS previo paso por SMORPH. En general, se ubican en posición antepuesta al verbo, excepto en:

- “afirmó Buonomo resignado”
- “existe una negociación”
- “estuvieron a su lado su hijo Máximo...etc”
- “no pudieron verlo muchos”
- “lo acompañaron algunos amigos”
- “dijo el médico”
- “si participará el Presidente”

Los snn que reúnen estas características, aclarando que además se consideran las aposiciones y el núcleo sujeto de la incidental, en total cuarenta y nueve, son los siguientes:

[Kirchner]
[Su médico]
[la recuperación]
[Lanús]
[el médico]
[el médico]
[Luis Buonomo]
[el triunfo]
[Kirchner]
[la fecha]
[Kirchner]
[la salud]
[el médico]
[los tiempos]
[ese]
[una negociación]
[El Presidente]
[el médico]
[el alta]
[el médico]
[la recuperación]
[La idea]
[Mi recomendación]
[Buonomo]
[El Presidente]
[Los dos partes]
[algunos amigos]
[su esposa]
[la senadora]
[El primero]
[su hijo]
[Máximo]
[su hija]
[Florencia]
[su secretario]
[Daniel Muñoz]
[el gobernador]
[Sergio Acevedo]
[muchos]
[El médico]
[la nutricionista]
[El menú]

[Buonomo]
 [Kirchner]
 [el médico]
 [los funcionarios]
 [el obispo]
 [Alejandro Bucciolini]
 [Nadie]
 [El presidente]

1.3 Observaciones de la aplicación de SMORPH

A partir de los datos informatizados, se analizó por SMORPH el texto presentado en el punto 1.1. En este análisis, se observan las siguientes cuestiones:

- la totalidad de las entradas ingresadas, son analizadas en base a la información declarada,
- los nombres propios y modos preposicionales y adverbiales declarados con guiones bajos intercalados, son reconocidos como una unidad,

Ej, 'Santa Cruz'.
 [Santa_Cruz 'TNOM','npr'],

'a causa de'.
 [a_causa_de 'EMS','npr'],

- SMORPH predice algunas ocurrencias no ingresadas de acuerdo a sus características morfológicas y en base a los datos que posee:

Ejemplo:
 'recomendación'.
 ['recomendación', 'EMS','nom'],

- arroja ambigüedades como:

'médico'.
 ['médico', 'TADJ','adj2'].
 ['médico', 'EMS','nom', 'GEN','masc', 'NUM','sg'],

Es el caso de aquellos nombres que pueden actuar como adjetivos.

Ej. - *El médico admitió que los resultados no son buenos.* (nombre)
 - *El tratamiento médico no dio resultado.* (adjetivo)

1.4 Observaciones acerca del análisis de MPS

Respecto del análisis de MPS, se destaca:

- algunos puntos como en *Petrona C. de Gandulfo* los toma como puntos finales y no como tratándose de una sigla,
- cuando no concatena, repite el análisis de SMORPH.

2. Resultados

Se representa en el cuadro que sigue el margen de cobertura y precisión en el reconocimiento de los snn en la posición sujeto¹¹², MPS detecta el 100% de los snn que se marcaron entre corchetes,

snn (sujeto) analizar ¹¹³	a	Cobertura (reconocimiento)	Precisión (no incorrectos)
- 49	-	49 (100%)	- 49 (100%)

3. Conclusiones y perspectivas

Este trabajo representa una contribución para el análisis automático del snn en español y puede ser útil en el campo de la enseñanza del español como lengua extranjera, a la hora de mejorar los útiles didácticos. Al respecto, cabe nombrar el trabajo del equipo **INFOSUR** concerniente a la morfología verbal del español, cf. SOLANA et al (2006), donde se informatizaron cerca de 4000 entradas verbales.

Una propuesta similar se prevé para el nombre, a partir de algunas de las cuestiones presentadas aquí. En efecto, la formalización de la descripción lingüística que propone el Paradigma 5P y la real libertad a la hora de presentar los datos, instaura un verdadero ida y vuelta, que hace revisar estos datos una y otra vez, de modo de dar cuenta de todas las combinaciones posibles. Esto provoca un afinamiento de esa descripción al tiempo que revaloriza y redefine el rol del lingüista en el tratamiento automático de textos. Esta tesis se enmarca en esta postura, poniendo de realce la especificidad del snn frente a otros sintagmas núcleos como el verbal (svn)¹¹⁴ o el adjetivo (san), simultáneamente se busca obtener capacidad predictiva respecto de los snn que pueden encontrarse en un texto en posición de sujeto.

¹¹² Los snn reconocidos por MPS son más que los que figuran en el cuadro que sigue, pero sólo contabilizamos los snn en posición sujeto.

¹¹³ Cabe aclarar aquí que reconoce tres snn incorrectos, pero no están en posición sujeto, en los tres casos, se trata de el ítem *médico*, el cual es ambiguo, puede ser tanto nombre como adjetivo, según se ve en el punto 1.3 de este Capítulo.

¹¹⁴ En SOLANA et al (2006), se toma al verbo teniendo en cuenta un ulterior tratamiento automático. Para el svn en particular, ver BÈS G.G., SOLANA Z. (2004/c)

Lejos de ser algo acabado sólo se pretendió dar una muestra de cómo las herramientas informáticas aquí utilizadas aplicadas al español posibilitan otra mirada sobre nuestra lengua. Las herramientas han mostrado su ductilidad como para adaptarse a una diversidad de datos. Esto se evidencia en la creación de sólo 36 modelos para 3000 nombres comunes. Sobre este particular, es pertinente mencionar también cómo, a partir de la creación de los modelos se muestran características propias del snn en español, como por ejemplo, el hecho de que los modelos con mayor cantidad de lemas nominales correspondan a los que agregan –s para constituir el plural.

Las peculiaridades del snn hacen que sea sumamente rica cualquier aproximación y que a su vez sea imposible no entrecruzarse con el adjetivo, el demostrativo, el pronombre, todos posibles candidatos a ser también núcleo del snn, por nombrar sólo algunos de los más usuales. Este entrecruzamiento hace que se hayan tenido que recortar muchas cuestiones, queda pendiente así un estudio más detallado de los determinantes y de los numerales, del mismo modo la investigación más a fondo de la conformación del snn en las construcciones interrogativas y en las subordinadas. Además este trabajo se ha centrado en la posición sintáctica sujeto, cabría detenerse en cómo se comporta el snn en otras posiciones. No menos importante es verificar qué sucede con el SN en su totalidad, cómo se encadena con, por ejemplo, el sintagma preposicional núcleo.

Como vemos son tantas e infinitas las posibilidades que se avizora un panorama muy provechoso para transitar. Quedan pues caminos de análisis que muy bien pueden tomarse como acciones futuras y que apuntan a seguir perfeccionando conocimientos en un área de investigación de la cual este aporte es sólo una mínima muestra.

Anexo

1. Análisis de SMORPH

Aclaración: Las palabras no declaradas en el archivo “entradas”: verbos, auxiliares, clíticos, conjunciones, aparecen con la leyenda “mi” que significa “mot inconnue”, palabra desconocida en francés.

'LA'.
['el', 'TDET', 'art'].

'SALUD'.
['salud', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

'DEL'.
['DEL', mi].

'PRESIDENTE':.
['presidente:', 'EMS', 'nom', 'GEN', '_', 'NUM', 'sg',].

'LLEVA'.
['LLEVA', mi].

'TRES'.
['tres', 'TNUM', 'num1'].

'DIAS'.
['día', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'pl'].

'HOSPITALIZADO'.
['HOSPITALIZADO', mi].

'
'
[' ', mi].

'Kirchner'.
['Kirchner', 'TNOM', 'npr'].

'evoluciona'.
['evoluciona', mi].

'bien'.
['bien', 'EMS', 'adv'].

'
'
['cc', 'EMS', 'coma'].

'pero'.
['pero', mi].

'sigue'.
['sigue', mi].

'internado'.
['internado', 'EMS','adj', 'GEN','masc', 'NUM','sg', 'TADJ','adj1'].

'en'.
['en', 'EMS','prep'].

'Santa Cruz'.
['Santa_Cruz', 'TNOM','npr'].

'Ya'.
['Ya', 'EMS','adv'].

'se'.
['se', mi].

'le'.
['le', mi].

'retiró'.
['retiró', mi].

'el'.
['el', 'TDET','art'].

'suero'.
['suero', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'y'.
['y', mi].

'está'.
['está', mi].

'siendo'.
['siendo', mi].

'medicado'.
['medicado', mi].

'por'.
['por', 'EMS', 'prep'].

'vía'.
['vía', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

'oral'.
['oral', 'TADJ', 'adj2'].

'.'.
['pf', 'EMS', 'ponc'].

'Su'.
['su', 'TDET', 'pos'].
'médico'.
['médico', 'TADJ', 'adj2'].
['médico', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'sg'].

'admitió'.
['admitió', 'mi'].

'que'.
['que', 'mi'].

'la'.
['el', 'TDET', 'art'].

'recuperación'.
['recuperación', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

'llevará'.
['llevará', 'mi'].

'días'.
['día', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'pl'].

'o'.
['o', 'mi'].

'semanas'.
['semana', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'pl'].

'Pero'.
['Pero', 'mi'].

'no'.
['no', 'EMS', 'adv'].

'aclaró'.
['aclaró', 'mi'].

'cuándo'.
['cuándo', 'mi'].

'tendrá'.
['tendrá', 'mi'].

'el'.
['el', 'TDET', 'art'].

'alta'.
['alta', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

'!'.
['pf', 'EMS', 'ponc'].

'Lanús'.
['Lanús', 'TNOM', 'npr'].

'jugó'.
['jugó', 'mi'].

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

'favor'.
['favor', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'sg'].

'de'.
['de', 'EMS', 'prep'].

'la'.
['el', 'TDET', 'art'].

'recuperación'.

['recuperación', 'EMS','nom', 'GEN','fem', 'NUM','sg'].

'del'.

['del', mi].

'Presidente.'

['presidente.', 'EMS','nom', 'GEN','_', 'NUM','sg'].

'Lo'.

['el', 'TDET','art'].

'dijo'.

['dijo', mi].

'ayer'.

['ayer', mi].

'el'.

['el', 'TDET','art'].

'médico'.

['médico', 'TADJ','adj2'].

['médico', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'personal'.

['personal', 'TADJ','adj2'].

'de'.

['de', 'EMS','prep'].

'Néstor'.

['Néstor', 'TNOM','npr'].

'Kirchner'.

['Kirchner', 'TNOM','npr'].

'al'.

['al', mi].

'destacar'.

['destacar', mi].

'su'.
 ['su', 'TDET','pos'].

"evolución".
 ["evolución', 'EMS','nom'].

'satisfactoria".
 ['satisfactoria"', mi].

'de'.
 ['de', 'EMS','prep'].

'la'.
 ['el', 'TDET','art'].

'fuerte'.
 ['fuerte', 'EMS','adj', 'GEN','masc', 'NUM','sg', 'TADJ','adj1'].

'descompostura'.
 ['descompostura', 'EMS','nom', 'GEN','fem', 'NUM','sg'].

'estomacal'.
 ['estomacal', 'TADJ','adj2'].

'por'.
 ['por', 'EMS','prep'].

'la'.
 ['el', 'TDET','art'].

'que'.
 ['que', mi].

'debió'.
 ['debió', mi].

'ser'.
 ['ser', mi].

'internado'.
 ['internado', 'EMS','adj', 'GEN','masc', 'NUM','sg', 'TADJ','adj1'].

'el'.
['el', 'TDET', 'art'].

'jueves'.
['jueves', 'EMS', 'nom', 'GEN', 'masc', 'GEN', '_'].

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

'la'.
['el', 'TDET', 'art'].

'noche'.
['noche', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

'!'.
['pf', 'EMS', 'ponc'].

'De todos modos'.
['de_todos_modos', 'EMS', 'adv'].

'!'.
['cc', 'EMS', 'coma'].

'afirmó'.
['afirmó', mi].

'que'.
['que', mi].

'todavía'.
['todavía', mi].

'no'.
['no', 'EMS', 'adv'].

'puede'.
['puede', mi].

'anticipar'.

['anticipar', mi].

'cuándo'.
 ['cuándo', mi].

'le'.
 ['le', mi].

'darán'.
 ['darán', mi].

'el'.
 ['el', 'TDET', 'art'].

'alta'.
 ['alta', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

':'.
 ['pf', 'EMS', 'ponc'].

'Más allá de'.
 ['más allá de', 'EMS', 'prep'].

'la'.
 ['el', 'TDET', 'art'].

'broma'.
 ['broma', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

','.
 ['cc', 'EMS', 'coma'].

'el'.
 ['el', 'TDET', 'art'].

'médico'.
 ['médico', 'TADJ', 'adj2'].
 ['médico', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'sg'].

'Luis'.

['Luis', 'TNOM','npr'].

'Buonomo'.

['Buonomo', 'TNOM','npr'].

'señaló'.

['señaló', mi].

'que'.

['que', mi].

'el'.

['el', 'TDET','art'].

'trunfo'.

['trunfo', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'del'.

['del', mi].

'viernes'.

['viernes', 'EMS','nom', 'GEN','masc', 'GEN','_'].

'de'.

['de', 'EMS','prep'].

'Racing'.

['Racing', 'TNOM','npr'].

','.

['cc', 'EMS','coma'].

'equipo'.

['equipo', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'del'.

['del', mi].

'que'.

['que', mi].

'Kirchner'.
 ['Kirchner', 'TNOM','npr'].

'es'.
 ['es', mi].

'hincha'.
 ['hincha', 'EMS','nom', 'GEN','_', 'NUM','sg'].

','
 ;'
 ['cc', 'EMS','coma'].

'sirvió'.
 ['sirvió', mi].

'para'.
 ['para', 'EMS','prep'].

'levantarle'.
 ['levantarle', mi].

'el'.
 ['el', 'TDET','art'].

'ánimo'.
 ['ánimo', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'y'.
 ['y', mi].

'destacó'.
 ['destacó', mi].

'que'.
 ['que', mi].

'no'.
 ['no', 'EMS','adv'].

'ha'.
 ['ha', mi].

'evidenciado'.
 ['evidenciado', mi].

'ningún'.
['ningún', 'TINDF1','indf1a'].

'tipo'.
['tipo', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'de'.
['de', 'EMS','prep'].

'complicaciones'.
['complicación', 'EMS','nom', 'GEN','fem', 'NUM','pl'].

'ni'.
['ni', mi].

'sobresaltos'.
['sobresalto', 'EMS','nom', 'GEN','masc', 'NUM','pl'].

':'.
['pf', 'EMS','ponc'].

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

'esto'.
['esto', 'TPRON','prdem'].

':'.
['cc', 'EMS','coma'].

'la'.
['el', 'TDET','art'].

'fecha'.
['fecha', 'EMS','nom', 'GEN','fem', 'NUM','sg'].

'en'.
['en', 'EMS','prep'].

'la'.
['el', 'TDET','art'].

'que'.
['que', mi].

'Kirchner'.
['Kirchner', 'TNOM','npr'].

'podrá'.
['podrá', mi].

'dejar'.
['dejar', mi].

'el'.
['el', 'TDET','art'].

'Hospital'.
['hospital', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'Regional'.
['Regional', 'TNOM','npr'].

'de'.
['de', 'EMS','prep'].

'Río Gallegos'.
['Río_Gallegos', 'TNOM','npr'].

'sigue'.
['sigue', mi].

'siendo'.
['siendo', mi].

'una'.
['una', 'TINDF2','indf2a'].

'incógnita'.
['incógnita', 'EMS','nom', 'GEN','fem', 'NUM','sg'].

'.'
 ['pf', 'EMS', 'ponc'].

'Es'.
 ['Es', mi].

'un'.
 ['un', 'TINDF1', 'indf1a'].

'tema'.
 ['tema', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'sg'].

'delicado'.
 ['delicado', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg', 'TADJ', 'adj1'].

'.'
 ['cc', 'EMS', 'coma'].

'porque'.
 ['porque', mi].

'la'.
 ['el', 'TDET', 'art'].

'salud'.
 ['salud', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

'presidencial'.
 ['presidencial', 'TADJ', 'adj2'].

'es'.
 ['es', mi].

'una'.
 ['una', 'TINDF2', 'indf2a'].

'cuestión'.
 ['cuestión', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

'de'.
['de', 'EMS', 'prep'].

'Estado'.
['Estado', 'TNOM', 'npr'].

'pero'.
['pero', 'mi'].

'también'.
['también', 'EMS', 'nom'].

'a causa de'.
['a_causa_de', 'EMS', 'prep'].

'la'.
['el', 'TDET', 'art'].

'personalidad'.
['personalidad', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

'de'.
['de', 'EMS', 'prep'].

'Kirchner'.
['Kirchner', 'TNOM', 'npr'].

','.
['cc', 'EMS', 'coma'].

'que'.
['que', 'mi'].

'se'.
['se', 'mi'].

'resiste'.
['resiste', 'mi'].

'a'.

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

'aceptar'.
['aceptar', mi].

'mansamente'.
['mansamente', mi].

'las'.
['el', 'TDET', 'art'].

'indicaciones'.
['indicación', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'pl'].

'médicas'.
['médicas', 'TADJ', 'adj2'].
['médico', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'pl'].

':'.
['pf', 'EMS', 'ponc'].

'Es'.
['Es', mi].

'por'.
['por', 'EMS', 'prep'].

'eso'.
['eso', 'TPRON', 'prdem'].

'que'.
['que', mi].

'el'.
['el', 'TDET', 'art'].

'médico'.
['médico', 'TADJ', 'adj2'].
['médico', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'sg'].

'reiteró'.
['reiteró', mi].

'ayer'.
['ayer', mi].

'que'.
['que', mi].

'si'.
['si', mi].

'no'.
['no', 'EMS', 'adv'].

'hay'.
['hay', mi].

'complicaciones'.
['complicación', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'pl'].

','
['cc', 'EMS', 'coma'].

'los'.
['el', 'TDET', 'art'].

'tiempos'.
['tiempo', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'pl'].

'de'.
['de', 'EMS', 'prep'].

'recuperación'.
['recuperación', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

'para'.
['para', 'EMS', 'prep'].

'un'.
['un', 'TINDF1', 'indf1a'].

'paciente'.
['paciente', 'EMS', 'nom', 'GEN', '_', 'NUM', 'sg'].

'común'.
 ['común', 'EMS','adj', 'NUM','sg', 'TADJ','adj1'].

'son'.
 ['son', mi].

'entre'.
 ['entre', 'EMS','prep'].

'7'.
 ['7', mi].

'y'.
 ['y', mi].

'10'.
 ['10', mi].

'días'.
 ['día', 'EMS','nom', 'GEN','masc', 'NUM','pl'].

','.
 ['cc', 'EMS','coma'].

'porque'.
 ['porque', mi].

'ese'.
 ['ese', 'TDET','dem'].

'es'.
 ['es', mi].

'el'.
 ['el', 'TDET','art'].

'período'.
 ['período', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'en'.
 ['en', 'EMS','prep'].

'el'.
['el', 'TDET', 'art'].

'que'.
['que', mi].

'se'.
['se', mi].

'puede'.
['puede', mi].

'producir'.
['producir', mi].

'un'.
['un', 'TINDF1', 'indf1a'].

'nuevo'.
['nuevo', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg', 'TADJ', 'adj1'].

'y'.
['y', mi].

'potencial'.
['potencial', 'EMS', 'adj', 'NUM', 'sg', 'TADJ', 'adj1'].

'sangrado".'.
['sangrado".'.', mi].

'Pero'.
['Pero', mi].

'reconoció'.
['reconoció', mi].

'que'.
['que', mi].

'por'.
['por', 'EMS', 'prep'].

'la'.
['el', 'TDET', 'art'].

'personalidad'.

['personalidad', 'EMS','nom', 'GEN','fem', 'NUM','sg'].

'y'.

['y', mi].

'responsabilidad'.

['responsabilidad', 'EMS','nom', 'GEN','fem', 'NUM','sg'].

'de'.

['de', 'EMS','prep'].

'Kirchner'.

['Kirchner', 'TNOM','npr'].

'existe'.

['existe', mi].

'una'.

['una', 'TINDF2','indf2a'].

'negociación'.

['negociación', 'EMS','nom', 'GEN','fem', 'NUM','sg'].

'El'.

['el', 'TDET','art'].

'Presidente'.

['presidente', 'EMS','nom', 'GEN','_', 'NUM','sg'].

'pretende'.

['pretende', mi].

'dejar'.

['dejar', mi].

'el'.

['el', 'TDET','art'].

'hospital'.

['hospital', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'sg'].

'lo'.

['el', 'TDET', 'art'].

'antes'.

['antes', 'EMS', 'adv'].

'posible'.

['posible', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg', 'TADJ', 'adj1'].

'y'.

['y', 'mi'].

'el'.

['el', 'TDET', 'art'].

'médico'.

['médico', 'TADJ', 'adj2'].

['médico', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'sg'].

'lo'.

['el', 'TDET', 'art'].

'único'.

['único', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg', 'TADJ', 'adj1'].

'que'.

['que', 'mi'].

'se'.

['se', 'mi'].

'animó'.

['animó', 'mi'].

'a'.

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

'decir'.

['decir', 'mi'].

'es'.
['es', mi].

'que'.
['que', mi].

'el'.
['el', 'TDET', 'art'].

'alta'.
['alta', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

'no'.
['no', 'EMS', 'adv'].

'será'.
['será', mi].

'hoy.'.
['hoy.', mi].

""Las'.
[""Las', mi].

'Pascuas'.
['Pascuas', 'TNOM', 'npr'].

'las'.
['el', 'TDET', 'art'].

'va'.
['va', mi].

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

'pasar'.
['pasar', mi].

'acá",'.
['acá"', mi].

'dijo.'.

'dijo'.
['dijo.', mi].

'Para'.
['para', 'EMS','prep'].

'referirse'.
['referirse', mi].

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

'la'.
['el', 'TDET','art'].

'evolución'.
['evolución', 'EMS','nom', 'GEN','fem', 'NUM','sg'].

'del'.
['del', mi].

'Presidente'.
['presidente', 'EMS','nom', 'GEN','_', 'NUM','sg'].

','
,':
['cc', 'EMS','coma'].

'el'.
['el', 'TDET','art'].

'médico'.
['médico', 'TADJ','adj2'].
['médico', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'dijo'.
['dijo', mi].

'que'.
['que', mi].

'ya'.
['ya', 'EMS','adv'].

'se'.

['se', mi].

'le'.

['le', mi].

'retiró'.

['retiró', mi].

'el'.

['el', 'TDET','art'].

'suero'.

['suero', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'y'.

['y', mi].

'que'.

['que', mi].

'se'.

['se', mi].

'lo'.

['el', 'TDET','art'].

'está'.

['está', mi].

'medicando'.

['medicando', mi].

'por'.

['por', 'EMS','prep'].

'vía'.

['vía', 'EMS','nom', 'GEN','fem', 'NUM','sg'].

'oral'.

['oral', 'TADJ','adj2'].

':'

['pf', 'EMS','ponc'].

'De todos modos'.
 ['de_todos_modos', 'EMS', 'adv'].

','
 ','
 ['cc', 'EMS', 'coma'].

'agregó'.
 ['agregó', mi].

'que'.
 ['que', mi].

'la'.
 ['el', 'TDET', 'art'].

'recuperación'.
 ['recuperación', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

'completa'.
 ['completa', mi].

'llevará'.
 ['llevará', mi].

'días'.
 ['día', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'pl'].

'o'.
 ['o', mi].

'semanas'.
 ['semana', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'pl'].

'La'.
 ['el', 'TDET', 'art'].

'idea'.
 ['idea', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

'es'.
 ['es', mi].

'que'.
 ['que', mi].

'esté'.

['esté', mi].

'desestresado'.

['desestresado', mi].

'y'.

['y', mi].

'que'.

['que', mi].

'si'.

['si', mi].

'vuelve'.

['vuelve', mi].

'a'.

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

'trabajar'.

['trabajar', mi].

'en'.

['en', 'EMS', 'prep'].

'los'.

['el', 'TDET', 'art'].

'próximos'.

['próximo', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'pl', 'TADJ', 'adj1'].

'días'.

['día', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'pl'].

'lo'.

['el', 'TDET', 'art'].

'haga'.

['haga', mi].

'con'.

['con', 'EMS', 'prep'].

'la'.
['el', 'TDET', 'art'].

'menor'.
['menor', 'EMS', 'adj', 'NUM', 'sg', 'TADJ', 'adj1'].

'cantidad'.
['cantidad', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

'de'.
['de', 'EMS', 'prep'].

'presión'.
['presión', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

'posible'.
['posible', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg', 'TADJ', 'adj1'].

'::'.
['pf', 'EMS', 'ponc'].

'Mi'.
['mi', 'TDET', 'pos'].

'recomendación'.
['recomendación', 'EMS', 'nom'].

'serían'.
['serían', 'mi'].

'30'.
['30', 'mi'].

'días'.
['día', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'pl'].

'de'.
['de', 'EMS', 'prep'].

'reposo'.
 ['reposo', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'sg'].

'psicofísico'.
 ['psicofísico', 'TADJ', 'adj2'].

','
 ['cc', 'EMS', 'coma'].

'pero'.
 ['pero', mi].

'sé'.
 ['sé', mi].

'que'.
 ['que', mi].

'no'.
 ['no', 'EMS', 'adv'].

'va'.
 ['va', mi].

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

'ser'.
 ['ser', mi].

'posible'.
 ['posible', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg', 'TADJ', 'adj1'].

'afirmó'.
 ['afirmó', mi].

'Buonomo'.
 ['Buonomo', 'TNOM', 'npr'].

'resignado.'.
 ['resignado.', mi].

'El'.
 ['el', 'TDET', 'art'].

'Presidente'.
 ['presidente', 'EMS', 'nom', 'GEN', '_', 'NUM', 'sg'].

'sufrió'.
 ['sufrió', 'mi].

'una'.
 ['una', 'TINDF2', 'indf2a'].

'gastroduodenitis'.
 ['gastroduodenitis', 'EMS', 'nom', 'GEN', 'fem', 'GEN', '_'].

'erosiva'.
 ['erosivo', 'EMS', 'adj', 'GEN', 'fem', 'NUM', 'sg', 'TADJ', 'adj1'].

'aguda'.
 ['agudo', 'EMS', 'adj', 'GEN', 'fem', 'NUM', 'sg', 'TADJ', 'adj1'].

'con'.
 ['con', 'EMS', 'prep'].

'hemorragia'.
 ['hemorragia', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

'una'.
 ['una', 'TINDF2', 'indf2a'].

'lesión'.
 ['lesión', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

'sangrante'.
 ['sangrante', 'mi].

'en'.
 ['en', 'EMS', 'prep'].

'el'.
 ['el', 'TDET', 'art'].

'duodeno'.
['duodeno', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'producida'.
['producida', mi].

'por'.
['por', 'EMS','prep'].

'un'.
['un', 'TINDF1','indf1a'].

'antiinflamatorio'.
['antiinflamatorio', 'TADJ','adj2'].

'recetado'.
['recetado', 'EMS','adj', 'GEN','masc', 'NUM','sg', 'TADJ','adj1'].

'para'.
['para', 'EMS','prep'].

'calmarle'.
['calmarle', mi].

'un'.
['un', 'TINDF1','indf1a'].

'dolor'.
['dolor', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'de'.
['de', 'EMS','prep'].

'un'.
['un', 'TINDF1','indf1a'].

'tratamiento'.
['tratamiento', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'dental'.
 ['dental', 'TADJ','adj2'].

':':
 ['pf', 'EMS','ponc'].

'Fue'.
 ['Fue', mi].

'internado'.
 ['internado', 'EMS','adj', 'GEN','masc', 'NUM','sg', 'TADJ','adj1'].

'de'.
 ['de', 'EMS','prep'].

'urgencia'.
 ['urgencia', 'EMS','nom', 'GEN','fem', 'NUM','sg'].

'el'.
 ['el', 'TDET','art'].

'jueves'.
 ['jueves', 'EMS','nom', 'GEN','masc', 'GEN','_'].

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

'la'.
 ['el', 'TDET','art'].

'noche'.
 ['noche', 'EMS','nom', 'GEN','fem', 'NUM','sg'].

'en'.
 ['en', 'EMS','prep'].

'El Calafate'.
 ['El_Calafate', 'TNOM','npr'].

','
';
['cc', 'EMS', 'coma'].

'donde'.
['donde', mi].

'pensaba'.
['pensaba', mi].

'descansar'.
['descansar', mi].

'durante'.
['durante', mi].

'el'.
['el', 'TDET', 'art'].

'fin'.
['fin', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'sg'].

'de'.
['de', 'EMS', 'prep'].

'semana'.
['semana', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

'largo'.
['largo', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg', 'TADJ', 'adj1'].

','
';
['cc', 'EMS', 'coma'].

'y'.
['y', mi].

'el'.
['el', 'TDET', 'art'].

'viernes'.
['viernes', 'EMS', 'nom', 'GEN', 'masc', 'GEN', '_'].

'fue'.
['fue', mi].

'trasladado'.
['trasladado', mi].

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

'Río Gallegos'.
['Río_Gallegos', 'TNOM', 'npr'].

','.
['cc', 'EMS', 'coma'].

'donde'.
['donde', mi].

'recibió'.
['recibió', mi].

'una'.
['una', 'TINDF2', 'indf2a'].

'transfusión'.
['transfusión', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

','.
['pf', 'EMS', 'ponc'].

'Los'.
['el', 'TDET', 'art'].

'dos'.
['dos', 'TNUM', 'num1'].

'partes'.
['parte', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'pl'].

'médicos'.
['médico', 'TADJ', 'adj2'].
['médico', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'pl'].

'de'.
['de', 'EMS', 'prep'].

'ayer'.
['ayer', 'mi].

'-a'.
['-a', 'mi].

'las'.
['el', 'TDET', 'art'].

'10'.
['10', 'mi].

'y'.
['y', 'mi].

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

'las'.
['el', 'TDET', 'art'].

'19-'.
['19-', 'mi].

'destacaron'.
['destacaron', 'mi].

'su'.
['su', 'TDET', 'pos'].

'evolución'.
['evolución', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

'favorable'.
['favorable', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg', 'TADJ', 'adj1'].

'!'
 ['pf', 'EMS', 'ponc'].

'Allí'.
 ['Allí', mi].

'se'.
 ['se', mi].

'aseguró'.
 ['aseguró', mi].

'que'.
 ['que', mi].

'pasó'.
 ['pasó', mi].

'una'.
 ['una', 'TINDF2', 'indf2a'].

'noche'.
 ['noche', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

'sin'.
 ['sin', 'EMS', 'prep'].

'complicaciones'.
 ['complicación', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'pl'].

'Lo'.
 ['el', 'TDET', 'art'].

'acompañaron'.
 ['acompañaron', mi].

'algunos'.
 ['algunos', 'TINDF2', 'indf2a'].

'amigos'.
 ['amigo', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'pl'].

'porque'.

['porque', mi].

'su'.

['su', 'TDET', 'pos'].

'esposa'.

['esposo', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

','

['cc', 'EMS', 'coma'].

'la'.

['el', 'TDET', 'art'].

'senadora'.

['senador', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

'Cristina'.

['Cristina', 'TNOM', 'npr'].

'Fernández'.

['Fernández', 'TNOM', 'npr'].

'de'.

['de', 'EMS', 'prep'].

'Kirchner'.

['Kirchner', 'TNOM', 'npr'].

','

['cc', 'EMS', 'coma'].

'se'.

['se', mi].

'fue'.

['fue', mi].

'a'.

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

'dormir'.
 ['dormir', mi].

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

'su'.
 ['su', 'TDET', 'pos'].

'casa'.
 ['casa', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

','
 ;:
 ['cc', 'EMS', 'coma'].

'ubicada'.
 ['ubicada', mi].

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

'sólo'.
 ['sólo', 'EMS', 'adv'].

'cuatro'.
 ['cuatro', 'TNUM', 'num1'].

'cuadras'.
 ['cuadra', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'pl'].

'del'.
 ['del', mi].

'hospital'.
 ['hospital', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'sg'].

','
 ;:
 ['pf', 'EMS', 'ponc'].

'Ayer'.
['Ayer', mi].

'lo'.
['el', 'TDET','art'].

'fue'.
['fue', mi].

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

'visitar'.
['visitar', mi].

'después'.
['después', 'EMS','nom'].

'del'.
['del', mi].

'mediodía'.
['mediodía', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'pero'.
['pero', mi].

'entró'.
['entró', mi].

'por'.
['por', 'EMS','prep'].

'una'.
['una', 'TINDF2','indf2a'].

'puerta'.
['puerta', 'EMS','nom', 'GEN','fem', 'NUM','sg'].

'lateral'.
['lateral', 'TADJ','adj2'].

'para'.
['para', 'EMS','prep'].

'eludir'.
['eludir', mi].

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

'los'.
['el', 'TDET','art'].

'periodistas'.
['periodista', 'EMS','nom', 'GEN','_', 'NUM','pl'].

'que'.
['que', mi].

'hacían'.
['hacían', mi].

'guarida'.
['guarida', 'EMS','nom', 'GEN','fem', 'NUM','sg'].

'en'.
['en', 'EMS','prep'].

'la'.
['el', 'TDET','art'].

'acceso'.
['acceso', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'principal'.
['principal', 'EMS','adj', 'NUM','sg', 'TADJ','adj1'].

'.'.
['pf', 'EMS','ponc'].

'El'.
['el', 'TDET', 'art'].

'primero'.
['primero', 'EMS', 'num', 'GEN', 'masc', 'NUM', 'sg', 'TNUM', 'num2'].

'en'.
['en', 'EMS', 'prep'].

'visitarlo'.
['visitarlo', 'mi'].

'ayer'.
['ayer', 'mi'].

'fue'.
['fue', 'mi'].

'el'.
['el', 'TDET', 'art'].

'secretario'.
['secretario', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'sg'].

'de'.
['de', 'EMS', 'prep'].

'Legal'.
['Legal', 'TNOM', 'npr'].

'y'.
['y', 'mi'].

'Técnica'.
['Técnica', 'TNOM', 'npr'].

','
['cc', 'EMS', 'coma'].

'Carlos'.
['Carlos', 'TNOM', 'npr'].

'Zanini'.
 ['Zanini', 'TNOM','npr'].

'::'
 ['pf', 'EMS','ponc'].

'También'.
 ['También', 'EMS','nom'].

'estuvieron'.
 ['estuvieron', mi].

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

'su'.
 ['su', 'TDET','pos'].

'lado'.
 ['lado', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'su'.
 ['su', 'TDET','pos'].

'hijo'.
 ['hijo', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'Máximo'.
 ['Máximo', 'TNOM','npr'].

'su'.
 ['su', 'TDET','pos'].

'hija'.
 ['hijo', 'EMS','nom', 'GEN','fem', 'NUM','sg'].

'Florencia'.
 ['Florencia', 'TNOM','npr'].

'está'.
['está', mi].

'en'.
['en', 'EMS','prep'].

'Buenos Aires'.
['Buenos_Aires', 'TNOM','npr'].

'su'.
['su', 'TDET','pos'].

'secretario'.
['secretario', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'privado'.
['privado', 'EMS','adj', 'GEN','masc', 'NUM','sg', 'TADJ','adj1'].

','.
['cc', 'EMS','coma'].

'Daniel'.
['Daniel', 'TNOM','npr'].

'Muñoz'.
['Muñoz', 'TNOM','npr'].

','.
['cc', 'EMS','coma'].

'y'.
['y', mi].

'el'.
['el', 'TDET','art'].

'gobernador'.
['gobernador', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'de'.
 ['de', 'EMS', 'prep'].

'Santa Cruz'.
 ['Santa_Cruz', 'TNOM', 'npr'].

':'.
 ['cc', 'EMS', 'coma'].

'Sergio'.
 ['Sergio', 'TNOM', 'npr'].

'Acevedo'.
 ['Acevedo', 'TNOM', 'npr'].

':'.
 ['pf', 'EMS', 'ponc'].

'No'.
 ['no', 'EMS', 'adv'].

'pudieron'.
 ['pudieron', 'mi'].

'verlo'.
 ['verlo', 'mi'].

'muchos'.
 ['muchos', 'TINDF2', 'indf2a'].

'más'.
 ['más', 'EMS', 'adv'].

'porque'.
 ['porque', 'mi'].

'tiene'.
 ['tiene', 'mi'].

'las'.
 ['el', 'TDET', 'art'].

'visitas'.
['visita', 'EMS','nom', 'GEN','fem', 'NUM','pl'].

'restringidas'.
['restringido', 'EMS','adj', 'GEN','fem', 'NUM','pl', 'TADJ','adj1'].

':'.
['pf', 'EMS','ponc'].

'Más allá de'.
['más_allá_de', 'EMS','prep'].

'la'.
['el', 'TDET','art'].

'guardia'.
['guardia', 'EMS','nom', 'GEN','fem', 'NUM','sg'].

'periodística,'.
['periodística,', mi].

'en'.
['en', 'EMS','prep'].

'la'.
['el', 'TDET','art'].

'puerta'.
['puerta', 'EMS','nom', 'GEN','fem', 'NUM','sg'].

'del'.
['del', mi].

'hospital'.
['hospital', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'no'.
['no', 'EMS','adv'].

'hubo'.
['hubo', mi].

'manifestaciones'.
['manifestación', 'EMS','nom', 'GEN','fem', 'NUM','pl'].

'y'.
['y', mi].

'en'.
['en', 'EMS','prep'].

'Río Gallegos'.
['Río_Gallegos', 'TNOM','npr'].

'se'.
['se', mi].

'respiró'.
['respiró', mi].

'el'.
['el', 'TDET','art'].

'clima'.
['clima', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'de'.
['de', 'EMS','prep'].

'un'.
['un', 'TINDF1','indf1a'].

'sábado'.
['sábado', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'más'.
['más', 'EMS','adv'].

':'.
['pf', 'EMS','ponc'].

'El'.
['el', 'TDET','art'].

'médico'.
['médico', 'TADJ','adj2'].
['médico', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'personal'.
['personal', 'TADJ','adj2'].

'de'.
['de', 'EMS','prep'].

'Kirchner'.
['Kirchner', 'TNOM','npr'].

'señaló'.
['señaló', mi].

'que'.
['que', mi].

'se'.
['se', mi].

'le'.
['le', mi].

'recetó'.
['recetó', mi].

'un'.
['un', 'TINDF1','indf1a'].

'plan'.
['plan', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'de'.
['de', 'EMS','prep'].

'nutrición'.
['nutrición', 'EMS','nom', 'GEN','fem', 'NUM','sg'].

'para'.
['para', 'EMS','prep'].

'este'.
['este', 'TDET','dem'].

'tipo'.
['tipo', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'de'.
['de', 'EMS','prep'].

'patologías'.
['patología', 'EMS','nom', 'GEN','fem', 'NUM','sg'].

'y'.
['y', mi].

'reconoció'.
['reconoció', mi].

'que'.
['que', mi].

'la'.
['el', 'TDET','art'].

'nutricionista'.
['nutricionista', 'EMS','nom', 'GEN','_', 'NUM','sg'].

'del'.
['del', mi].

'hospital'.
['hospital', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'no'.
["'no', mi].

'es'.
['es', mi].

'Petrona C'.
 ['Petrona_C', 'TNOM','npr'].

'
 ;
 ;'
 ['pf', 'EMS','ponc'].

'de'.
 ['de', 'EMS','prep'].

'Gandulfo".'.
 ['Gandulfo".'.', mi].

'El'.
 ['el', 'TDET','art'].

'menú'.
 ['menú', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'permitido'.
 ['permitido', 'EMS','adj', 'GEN','masc', 'NUM','sg', 'TADJ','adj1'].

'incluye'.
 ['incluye', mi].

'nada más que'.
 ['nada_más_que', 'EMS','adv'].

'carnes'.
 ['carne', 'EMS','nom', 'GEN','fem', 'NUM','pl'].

'desgrasadas'.
 ['desgrasado', 'EMS','adj', 'GEN','fem', 'NUM','pl', 'TADJ','adj1'].

'
 ;
 ;'
 ['cc', 'EMS','coma'].

'verduras'.
 ['verdura', 'EMS','nom', 'GEN','fem', 'NUM','pl'].

'hervidas'.

['hervido', 'EMS','adj', 'GEN','fem', 'NUM','pl', 'TADJ','adj1'].

'y'.
['y', mi].

'gelatinas'.
['gelatina', 'EMS','nom', 'GEN','fem', 'NUM','pl'].

':'.
['pf', 'EMS','ponc'].

'Frente a'.

['Frente_a', 'EMS','prep'] .

'una'.
['una', 'TINDF2','indf2a'].

'pregunta'.
['pregunta', 'EMS','nom', 'GEN','fem', 'NUM','sg'].

'puntual'.
['puntual', 'EMS','adj', 'NUM','sg', 'TADJ','adj1'].

':'.
['cc', 'EMS','coma'].

'y'.
['y', mi].

'para'.
['para', 'EMS','prep'].

'que'.
['que', mi].

'no'.
['no', 'EMS','adv'].

'se'.
['se', mi].

'empiecen'.
['empiecen', mi].

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

'hablar'.
['hablar', mi].

'tonterías'.
['tontería', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

'Buonomo'.
['Buonomo', 'TNOM', 'npr'].

'aclaró'.
['aclaró', mi].

'que'.
['que', mi].

'Kirchner'.
['Kirchner', 'TNOM', 'npr'].

'nunca'.
['nunca', mi].

'tuvo'.
['tuvo', mi].

'una'.
['una', 'TINDF2', 'indf2a'].

'afección'.
['afección', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

'estomacal'.
['estomacal', 'TADJ', 'adj2'].

'con'.
['con', 'EMS', 'prep'].

'el'.
['el', 'TDET', 'art'].

'único'.

['único', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg', 'TADJ', 'adj1'].

'antecedente'.

['antecedente', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'sg'].

'de'.

['de', 'EMS', 'prep'].

'un'.

['un', 'TINDF1', 'indf1a'].

'colon'.

['colon', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'sg'].

'irritable".'.

['irritable".'. , mi].

'Siempre'.

['Siempre', mi].

'ejerció'.

['ejerció', mi].

'y'.

['y', mi].

'está'.

['está', mi].

'ejerciando'.

['ejerciando', mi].

'el'.

['el', 'TDET', 'art'].

'mando'.

['mando', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'sg'].

'en'.

['en', 'EMS', 'prep'].

'contacto'.
 ['contacto', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'permanente'.
 ['permanente', 'EMS','adj', 'GEN','masc', 'NUM','sg', 'TADJ','adj1'].

'con'.
 ['con', 'EMS','prep'].

'el'.
 ['el', 'TDET','art'].

'Gabinete'.
 ['gabinete', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'y'.
 ['y', mi].

'tomando'.
 ['tomando', mi].

'decisiones'.
 ['decisión', 'EMS','nom', 'GEN','fem', 'NUM','sg'].

'dijo'.
 ['dijo', mi].

'el'.
 ['el', 'TDET','art'].

'médico'.
 ['médico', 'TADJ','adj2'].
 ['médico', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'!'.
 ['pf', 'EMS','ponc'].

'Sin embargo'.
 ['sin_embargo', 'EMS','adv'].

'los'.
 ['el', 'TDET','art'].

'funcionarios'.
 ['funcionario', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'pl'].

'que'.
 ['que', 'mi'].

'lo'.
 ['el', 'TDET', 'art'].

'visitaron'.
 ['visitaron', 'mi'].

'ayer'.
 ['ayer', 'mi'].

'afirmaron'.
 ['afirmaron', 'mi'].

'que'.
 ['que', 'mi'].

'lo'.
 ['el', 'TDET', 'art'].

'mantienen'.
 ['mantienen', 'mi'].

'alejado'.
 ['alejado', 'mi'].

'del'.
 ['del', 'mi'].

'teléfono'.
 ['teléfono', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'sg'].

'!'.
 ['pf', 'EMS', 'ponc'].

'Hoy'.
 ['Hoy', 'mi'].

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

'las'.
['el', 'TDET','art'].

'10'.
['10', mi].

'habrá'.
['habrá', mi].

'un'.
['un', 'TINDF1','indf1a'].

'nuevo'.
['nuevo', 'EMS','adj', 'GEN','masc', 'NUM','sg', 'TADJ','adj1'].

'parte'.
['parte', 'EMS','nom', 'GEN','fem', 'NUM','sg'].

'médico'.
['médico', 'TADJ','adj2'].
['médico', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'!'.
['pf', 'EMS','ponc'].

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

'las'.
['el', 'TDET','art'].

'11',.
['11,', mi].

'el'.
['el', 'TDET','art'].

'obispo'.
['obispo', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'de'.
['de', 'EMS','prep'].

'Santa Cruz'.
 ['Santa_Cruz', 'TNOM','npr'].

','
 ;'
 ['cc', 'EMS','coma'].

'Alejandro'.
 ['Alejandro', 'TNOM','npr'].

'Buccolini'.
 ['Buccolini', 'TNOM','npr'].

','
 ;'
 ['cc', 'EMS','coma'].

'dará'.
 ['dará', mi].

'una'.
 ['una', 'TINDF2','indf2a'].

'misa'.
 ['misa', 'EMS','nom', 'GEN','fem', 'NUM','sg'].

'para'.
 ['para', 'EMS','prep'].

'los'.
 ['el', 'TDET','art'].

'internados'.
 ['internado', 'EMS','adj', 'GEN','masc', 'NUM','pl', 'TADJ','adj1'].

'del'.
 ['del', mi].

'hospital'.
 ['hospital', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'':
 ['cc', 'EMS', 'coma'].

'por'.
 ['por', 'EMS', 'prep'].

'la'.
 ['el', 'TDET', 'art'].

'Pascua'.
 ['Pascua.', 'mi'].

'Nadie'.
 ['nadie', 'TPRON', 'prindf'].

'se'.
 ['se', 'mi'].

'animó'.
 ['animó', 'mi'].

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

'anticipar'.
 ['anticipar', 'mi'].

'si'.
 ['si', 'mi'].

'participará'.
 ['participará', 'mi'].

'el'.
 ['el', 'TDET', 'art'].

'Presidente'.
 ['presidente', 'EMS', 'nom', 'GEN', '_', 'NUM', 'sg'].

'':
 ['pf', 'EMS', 'ponc'].

2. Análisis de MPS

Aclaración: Aparecen resaltados en color los snn sujeto que se resaltaron entre corchetes en el Capítulo 5.

'LA SALUD'.
['el salud', 'EMS', 'snn'].

'DEL'.
['DEL' , mi].

'PRESIDENTE:'.
['PRESIDENTE:' , mi].

'LLEVA'.
['LLEVA' , mi].

'TRES DIAS'.
['tres día', 'EMS', 'snn'].

'HOSPITALIZADO'.
['HOSPITALIZADO' , mi].

%fin de la phrase numero : 1

'Kirchner'.
['Kirchner', 'EMS', 'snn'].

'evoluciona'.
['evoluciona', mi].

'bien'.
['bien', 'EMS', 'adv'].

','.
['cc', 'EMS', 'coma'].

'pero'.
['pero' , mi].

'sigue'.
['sigue' , mi].

'internado'.
['internado', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg', 'TADJ', 'adj1'].

'en'.
['en', 'EMS', 'prep'].

'Santa Cruz'.
 ['Santa_Cruz', 'EMS', 'snn'].

En este primer fragmento, debía reconocer el snn sujeto [Kirchner], que para mejor visualización colocamos en color.

'Ya'.
 ['Ya' , 'EMS','adv'].

'se'.
 ['se' , mi].

'le'.
 ['le' , mi].

'retiró'.
 ['retiró' , mi].

'el suero'.
 ['el suero', 'EMS', 'snn'].

'y'.
 ['y' , mi].

'está'.
 ['está' , mi].

'siendo'.
 ['siendo' , mi].

'medicado'.
 ['medicado' , mi].

'por'.
 ['por', 'EMS', 'prep'].

'vía'.
 ['vía', 'EMS', 'snn'].

'oral'.
 ['oral', 'TADJ', 'adj2'].

'.'.
 ['pf', 'EMS', 'ponc'].

%fin de la phrase numero : 2

'Su médico'.
 ['su médico', 'EMS', 'snn'].

'admitió'.
['admitió' , mi].

'que'.
['que' , mi].

'la recuperación'.
['el recuperación', 'EMS', 'snn'].

'llevará'.
['llevará' , mi].

'días'.
['día' , 'EMS', 'snn'].

'o'.
['o' , mi].

'semanas'.
['semana', 'EMS', 'snn'].

'Pero'.
['Pero' , mi].

'no'.
['no', 'EMS', 'adv'].

'aclaró'.
['aclaró' , mi].

'cuándo'.
['cuándo' , mi].

'tendrá'.
['tendrá' , mi].

'el alta'.
['el alta', 'EMS', 'snn'].

'!'.
['pf', 'EMS', 'ponc'].

%fin de la phrase numero : 3

En este segundo fragmento, debía reconocer los snn sujeto: [Su médico] y [la recuperación], los vemos resaltados con color.

'Lanús'.
['Lanús', 'EMS', 'snn'].

'jugó'.
['jugó' , mi].

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

'favor'.
['favor', 'EMS', 'snn'].

'de'.
['de', 'EMS', 'prep'].

'la recuperación'.
['el recuperación', 'EMS', 'snn'].

'del'.
['del' , mi].

'Presidente.'.
['presidente.', 'EMS', 'snn'].

En este tercer fragmento, debía reconocer el snn sujeto [Lanús] que efectivamente está resaltado.

'Lo'.
['el', 'TDET', 'art'].

'dijo'.
['dijo' , mi].

'ayer'.
['ayer' , mi].

'el médico'.
['el médico', 'EMS', 'snn'].

'personal'.
['personal', 'TADJ', 'adj2'].

'de'.
['de', 'EMS', 'prep'].

'Néstor Kirchner'.
['Néstor Kirchner', 'EMS', 'snn'].

'al'.
['al' , mi].

'destacar'.

['destacar' , mi].

'su "evolución'.

['su "evolución', 'EMS', 'snn'].

'satisfactoria"'.
 'satisfactoria"'.
 ['satisfactoria" , mi].

'de'.

['de', 'EMS', 'prep'].

'la fuerte descompostura'.

['el fuerte descompostura', 'EMS', 'snn'].

'estomacal'.

['estomacal', 'TADJ', 'adj2'].

'por'.

['por', 'EMS', 'prep'].

'la'.

['el', 'TDET', 'art'].

'que'.

['que' , mi].

'debió'.

['debió' , mi].

'ser'.

['ser' , mi].

'internado'.

['internado', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg', 'TADJ', 'adj1'].

'el jueves'.

['el jueves', 'EMS', 'snn'].

'a'.

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

'la noche'.

['el noche', 'EMS', 'snn'].

'::

['pf', 'EMS', 'ponc'].

%fin de la phrase numero : 4

En este cuarto fragmento, debía reconocer el snn sujeto: [el médico] que vemos resaltado.

'De todos modos'.
['de_todos_modos', 'EMS', 'adv'].

':'.
['cc', 'EMS', 'coma'].

'afirmó'.
['afirmó', mi].

'que'.
['que', mi].

'todavía'.
['todavía', mi].

'no'.
['no', 'EMS', 'adv'].

'puede'.
['puede', mi].

'anticipar'.
['anticipar', mi].

'cuándo'.
['cuándo', mi].

'le'.
['le', mi].

'darán'.
['darán', mi].

'el alta'.
['el alta', 'EMS', 'snn'].

':'.
['pf', 'EMS', 'ponc'].

%fin de la phrase numero : 5

'Más allá de'.
['más allá de', 'EMS', 'prep'].

'la broma'.
['el broma', 'EMS', 'snn'].

','
,':
['cc', 'EMS', 'coma'].

'el médico'.
['el médico', 'EMS', 'snn'].

'Luis Buonomo'.
['Luis Buonomo', 'EMS', 'snn'].

'señaló'.
['señaló' , mi].

'que'.
['que' , mi].

'el triunfo'.
['el triunfo', 'EMS', 'snn'].

'del'.
['del' , mi].

'viernes'.
['viernes', 'EMS', 'snn'].

'de'.
['de', 'EMS', 'prep'].

'Racing'.
['Racing', 'EMS', 'snn'].

','
,':
['cc', 'EMS', 'coma'].

'equipo'.
['equipo', 'EMS', 'snn'].

'del'.
['del' , mi].

'que'.
['que' , mi].

'Kirchner'.
['Kirchner', 'EMS', 'snn'].

'es'.
['es' , mi].

'hincha'.
['hincha', 'EMS', 'snn'].

':
';:
['cc', 'EMS', 'coma'].

'sirvió'.
['sirvió' , mi].

'para'.
['para', 'EMS', 'prep'].

'levantarle'.
['levantarle' , mi].

'el ánimo'.
['el ánimo', 'EMS', 'snn'].

'y'.
['y' , mi].

'destacó'.
['destacó' , mi].

'que'.
['que' , mi].

'no'.
['no' , 'EMS', 'adv'].

'ha'.
['ha' , mi].

'evidenciado'.
['evidenciado' , mi].

'ningún tipo'.
['ningún tipo', 'EMS', 'snn'].

'de'.
['de', 'EMS', 'prep'].

'complicaciones'.
['complicación', 'EMS', 'snn'].

'ni'.
['ni' , mi].

'sobresaltos'.
['sobresalto', 'EMS', 'snn'].

':.

['pf', 'EMS', 'ponc'].

%fin de la phrase numero : 6

En este quinto fragmento, debía reconocer los snn sujeto
[el médico] [Luis Buonomo], [el triunfo] que están resaltados.

'Pese a'.

['pese_a', 'EMS', 'prep'].

'esto'.

['esto', 'EMS', 'snn'].

','.

['cc', 'EMS', 'coma'].

'la fecha'.

['el fecha', 'EMS', 'snn'].

'en'.

['en', 'EMS', 'prep'].

'la'.

['el', 'TDET', 'art'].

'que'.

['que', 'mi'].

'Kirchner'.

['Kirchner', 'EMS', 'snn'].

'podrá'.

['podrá', 'mi'].

'dejar'.

['dejar', 'mi'].

'el Hospital'.

['el hospital', 'EMS', 'snn'].

'Regional'.

['Regional', 'EMS', 'snn'].

'de'.

['de', 'EMS', 'prep'].

'Río Gallegos'.

['Río_Gallegos', 'EMS', 'snn'].

'sigue'.

['sigue' , mi].

'siendo'.
 ['siendo' , mi].

'una incógnita'.
 ['una incógnita', 'EMS', 'snn'].

':

['pf', 'EMS', 'ponc'].

n

%fin de la phrase numero : 7

En el fragmento sexto debía reconocer los snn [la fecha] y [Kirchner] oportunamente en color,

'Es'.
 ['Es' , mi].

'un tema'.
 ['un tema', 'EMS', 'snn'].

'delicado'.
 ['delicado', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg', 'TADJ', 'adj1'].

':

['cc', 'EMS', 'coma'].

'porque'.
 ['porque' , mi].

'la salud'.
 ['el salud', 'EMS', 'snn'].

'presidencial'.
 ['presidencial', 'TADJ', 'adj2'].

'es'.
 ['es' , mi].

'una cuestión'.
 ['una cuestión', 'EMS', 'snn'].

'de'.
 ['de', 'EMS', 'prep'].

'Estado'.
 ['Estado', 'EMS', 'snn'].

'pero'.

['pero' , mi].

'también'.

['también', 'EMS', 'snn'].

'a causa de'.

['a_causa_de', 'EMS', 'prep'].

'la personalidad'.

['el personalidad', 'EMS', 'snn'].

'de'.

['de', 'EMS', 'prep'].

'Kirchner'.

['Kirchner', 'EMS', 'snn'].

','

['cc', 'EMS', 'coma'].

'que'.

['que' , mi].

'se'.

['se' , mi].

'resiste'.

['resiste' , mi].

'a'.

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

'aceptar'.

['aceptar' , mi].

'mansamente'.

['mansamente' , mi].

'las indicaciones'.

['el indicación', 'EMS', 'snn'].

'médicas'.

['médico', 'EMS', 'snn'].

','

['pf', 'EMS', 'ponc'].

%fin de la phrase numero : 8

En el séptimo fragmento, debía reconocer el snn [la salud].

'Es'.
['Es' , mi].

'por'.
['por', 'EMS', 'prep'].

'eso'.
['eso', 'EMS', 'snn'].

'que'.
['que' , mi].

'el médico'.
['el médico', 'EMS', 'snn'].

'reiteró'.
['reiteró' , mi].

'ayer'.
['ayer' , mi].

'que'.
['que' , mi].

'si'.
["'si' , mi].

'no'.
['no', 'EMS', 'adv'].

'hay'.
['hay' , mi].

'complicaciones'.
['complicación', 'EMS', 'snn'].

','.
['cc', 'EMS', 'coma'].

'los tiempos'.
['el tiempo', 'EMS', 'snn'].

'de'.
['de', 'EMS', 'prep'].

'recuperación'.
['recuperación', 'EMS', 'snn'].

'para'.
['para', 'EMS', 'prep'].

'un paciente'.
['un paciente', 'EMS', 'snn'].

'común'.
['común', 'EMS', 'adj', 'NUM', 'sg', 'TADJ', 'adj1'].

'son'.
['son', mi].

'entre'.
['entre', 'EMS', 'prep'].

'7'.
['7', mi].

'y'.
['y', mi].

'10'.
['10', mi].

'días'.
['día', 'EMS', 'snn'].

','.
['cc', 'EMS', 'coma'].

'porque'.
['porque', mi].

'ese'.
['ese', 'EMS', 'snn'].

'es'.
['es', mi].

'el período'.
['el período', 'EMS', 'snn'].

'en'.
['en', 'EMS', 'prep'].

'el'.
['el', 'TDET', 'art'].

'que'.
['que', mi].

'se'.

['se' , mi].

'puede'.
['puede' , mi].

'producir'.
['producir' , mi].

'un'.
['un', 'TINDF', 'indf1a'].

'nuevo'.
['nuevo', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg', 'TADJ', 'adj1'].

'y'.
['y' , mi].

'potencial'.
['potencial', 'EMS', 'adj', 'NUM', 'sg', 'TADJ', 'adj1'].

'sangrado".'.
['sangrado".' , mi].

'Pero'.
['Pero' , mi].

'reconoció'.
['reconoció' , mi].

'que'.
['que' , mi].

'por'.
['por' , 'EMS', 'prep'].

'la personalidad'.
['el personalidad', 'EMS', 'snn'].

'y'.
['y' , mi].

'responsabilidad'.
['responsabilidad', 'EMS', 'snn'].

'de'.
['de', 'EMS', 'prep'].

'Kirchner'.
['Kirchner', 'EMS', 'snn'].

'existe'.
['existe' , mi].

'una negociación'.
['una', 'EMS', 'snn'].

'El Presidente'.
['el presidente', 'EMS', 'snn'].

'pretende'.
['pretende' , mi].

'dejar'.
['dejar' , mi].

'el hospital'.
['el hospital', 'EMS', 'snn'].

'lo'.
['el', 'TDET', 'art'].

'antes'.
['antes', 'adv'].

'posible'.
['posible', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg', 'TADJ', 'adj1'].

'y'.
['y' , mi].

'el médico'.
['el médico', 'EMS', 'snn'].

'lo único'.
['el único', 'EMS', 'snn'].

'que'.
['que' , mi].

'se'.
['se' , mi].

'animó'.
['animó' , mi].

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

'decir'.
['decir' , mi].

'es'.
['es' , mi].

'que'.
['que' , mi].

'el alta'.
['el alta', 'EMS', 'snn'].

'no'.
['no', 'EMS', 'adv'].

'será'.
['será' , mi].

'hoy.'.
['hoy.' , mi].

""Las'.
[""Las' , mi].

'Pascuas'.
['Pascuas', 'EMS', 'snn'].

'las'.
['el', 'TDET', 'art'].

'va'.
['va' , mi].

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

'pasar'.
['pasar' , mi].

'acá",'.
['acá"', ' , mi].

'dijo.'.
['dijo.' , mi].

En el fragmento octavo, debía reconocer los snn [el médico], [los tiempos], [ese], [una negociación], [El Presidente], [el médico], [el alta] que están efectivamente resaltados. Reconoce todos.

'Para'.
['para', 'EMS', 'prep'].

'referirse'.
['referirse' , mi].

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

'la evolución'.
['el evolución', 'EMS', 'snn'].

'del'.
['del' , mi].

'Presidente'.
['presidente', 'EMS', 'snn'].

','.
['cc', 'EMS', 'coma'].

'el médico'.
['el médico', 'EMS', 'snn'].

'dijo'.
['dijo' , mi].

'que'.
['que' , mi].

'ya'.
['ya' , 'EMS', 'adv'].

'se'.
['se' , mi].

'le'.
['le' , mi].

'retiró'.
['retiró' , mi].

'el suero'.
['el suero', 'EMS', 'snn'].

'y'.
['y' , mi].

'que'.
['que' , mi].

'se'.
['se' , mi].

'lo'.
['el', 'TDET', 'art'].

'está'.
['está', 'mi'].

'medicando'.
['medicando', 'mi'].

'por'.
['por', 'EMS', 'prep'].

'vía'.
['vía', 'EMS', 'snn'].

'oral'.
['oral', 'TADJ', 'adj2'].

'.'.

['pf', 'EMS', 'ponc'].

%fin de la phrase numero : 9

En el fragmento noveno, debía reconocer el snn [el médico] ya resaltado.

'De todos modos'.
['de_todos_modos', 'EMS', 'adv'].

'.'.

['cc', 'EMS', 'coma'].

'agregó'.
['agregó', 'mi'].

'que'.
['que', 'mi'].

'la recuperación'.
['el recuperación', 'EMS', 'snn'].

'completa'.
['completa', 'mi'].

'llevará'.
['llevará', 'mi'].

'días'.
['día', 'EMS', 'snn'].

'o'.
['o' , mi].

'semanas'.
['semana', 'EMS', 'snn'].

'La'.
['el idea', 'EMS', 'snn'].

'es'.
['es' , mi].

'que'.
['que' , mi].

'esté'.
['esté' , mi].

'desestresado'.
['desestresado' , mi].

'y'.
['y' , mi].

'que'.
['que' , mi].

'si'.
['si' , mi].

'vuelve'.
['vuelve' , mi].

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

'trabajar'.
['trabajar' , mi].

'en'.
['en', 'EMS', 'prep'].

'los próximos días'.
['el próximo día', 'EMS', 'snn'].

'lo'.
['el', 'TDET', 'art'].

'haga'.
['haga' , mi].

'con'.
['con', 'EMS', 'prep'].

'la menor cantidad'.
['el menor cantidad', 'EMS', 'snn'].

'de'.
['de', 'EMS', 'prep'].

'presión'.
['presión', 'EMS', 'snn'].

'posible'.
['posible', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg', 'TADJ', 'adj1'].

'.'.

['pf', 'EMS', 'ponc'].

%fin de la phrase numero: 10

En el fragmento décimo, debía reconocer los snn [la recuperación] y [La idea], y efectivamente lo hizo.

'Mi recomendación'.
['mi recomendación', 'EMS', 'snn'].

'serían'.
['serían', mi].

'30'.
['30', mi].

'días'.
['día', 'EMS', 'snn'].

'de'.
['de', 'EMS', 'prep'].

'reposo'.
['reposo', 'EMS', 'snn'].

'psicofísico'.
['psicofísico', 'TADJ', 'adj2'].

'.'.

['cc', 'EMS', 'coma'].

'pero'.
['pero', mi].

'sé'.
['sé' , mi].

'que'.
['que' , mi].

'no'.
['no', 'EMS', 'adv'].

'va'.
['va' , mi].

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

'ser'.
['ser' , mi].

'posible'.
['posible', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg', 'TADJ', 'adj1'].

'afirmó'.
['afirmó' , mi].

'Buonomo'.
['Buonomo', 'EMS', 'snn'].

'resignado.'.
['resignado.' , mi].

En el fragmento once, debía reconocer los snn [Mi recomendación] y [Buonomo] ya resaltados.

'El Presidente'.
['el presidente', 'EMS', 'snn'].

'sufrió'.
['sufrió' , mi].

'una gastroduodenitis'.
['una gastroduodenitis', 'EMS', 'snn'].

'erosiva'.
['erosivo', 'EMS', 'adj', 'GEN', 'fem', 'NUM', 'sg', 'TADJ', 'adj1'].

'aguda'.
['agudo', 'EMS', 'adj', 'GEN', 'fem', 'NUM', 'sg', 'TADJ', 'adj1'].

'con'.

['con', 'EMS', 'prep'].

'hemorragia'.

['hemorragia', 'EMS', 'snn'].

'una lesión'.

['una lesión', 'EMS', 'snn'].

'sangrante'.

['sangrante', mi].

'en'.

['en', 'EMS', 'prep'].

'el duodeno'.

['el duodeno', 'EMS', 'snn'].

'producida'.

['producida', mi].

'por'.

['por', 'EMS', 'prep'].

'un antiinflamatorio'.

['un antiinflamatorio', 'EMS', 'snn'].

'recetado'.

['recetado', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg', 'TADJ', 'adj1'].

'para'.

['para', 'EMS', 'prep'].

'calmarle'.

['calmarle', mi].

'un dolor'.

['un dolor', 'EMS', 'snn'].

'de'.

['de', 'EMS', 'prep'].

'un tratamiento'.

['un tratamiento', 'EMS', 'snn'].

'dental'.

['dental', 'TADJ', 'adj2'].

'!':

['pf', 'EMS', 'ponc'].

%fin de la phrase numero : 11

En el fragmento doce, debía reconocer el snn [El Presidente], efectivamente resaltado.

'Fue'.
['Fue' , mi].

'internado'.
['internado', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg', 'TADJ', 'adj1'].

'de'.
['de', 'EMS', 'prep'].

'urgencia'.
['urgencia', 'EMS', 'snn'].

'el jueves'.
['el jueves', 'EMS', 'snn'].

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

'la noche'.
['el noche', 'EMS', 'snn'].

'en'.
['en', 'EMS', 'prep'].

'El Calafate'.
['El_Calafate', 'EMS', 'snn'].

','.
['cc', 'EMS', 'coma'].

'donde'.
['donde' , mi].

'pensaba'.
['pensaba' , mi].

'descansar'.
['descansar' , mi].

'durante'.
['durante' , mi].

'el fin'.
['el fin', 'EMS', 'snn'].

'de'.

['de', 'EMS', 'prep'].

'semana'.

['semana', 'EMS', 'snn'].

'largo'.

['largo', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg', 'TADJ', 'adj1'].

','

['cc', 'EMS', 'coma'].

'y'.

['y', 'mi'].

'el viernes'.

['el viernes', 'EMS', 'snn'].

'fue'.

['fue', 'mi'].

'trasladado'.

['trasladado', 'mi'].

'a'.

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

'Río Gallegos'.

['Río_Gallegos', 'EMS', 'snn'].

','

['cc', 'EMS', 'coma'].

'donde'.

['donde', 'mi'].

'recibió'.

['recibió', 'mi'].

'una transfusión'.

['una transfusión', 'EMS', 'snn'].

':'

['pf', 'EMS', 'ponc'].

%fin de la phrase numero : 12

'Los dos partes'.

['el dos parte', 'EMS', 'snn'].

'médicos'.
['médico', 'EMS', 'snn'].

'de'.
['de', 'EMS', 'prep'].

'ayer'.
['ayer' , mi].

'-a'.
['-a' , mi].

'las'.
['el', 'TDET', 'art'].

'10'.
['10' , mi].

'y'.
['y' , mi].

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

'las'.
['el', 'TDET', 'art'].

'19-'.
['19-' , mi].

'destacaron'.
['destacaron' , mi].

'su evolución'.
['su evolución', 'EMS', 'snn'].

'favorable'.
['favorable', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg', 'TADJ', 'adj1'].

'!'.
['pf', 'EMS', 'ponc'].

%fin de la phrase numero : 13

En el fragmento trece, el snn debía ser [Los dos partes], efectivamente reconocidos.

'Allí'.
['Allí' , mi].

'se'.
['se' , mi].

'aseguró'.
['aseguró' , mi].

'que'.
['que' , mi].

'pasó'.
['pasó' , mi].

'una noche'.
['una noche', 'EMS', 'snn'].

'sin'.
['sin', 'EMS', 'prep'].

'complicaciones'.
['complicación' , 'EMS', 'snn'].

'Lo'.
['el', 'TDET', 'art'].

'acompañaron'.
['acompañaron' , mi].

'algunos amigos'.
['algunos amigo', 'EMS', 'snn'].

'porque'.
['porque' , mi].

'su esposa'.
['su esposo', 'EMS', 'snn'].

','.
['cc', 'EMS', 'coma'].

'la senadora'.
['el senador', 'EMS', 'snn'].

'Cristina Fernández'.
['Cristina Fernández', 'EMS', 'snn'].

'de'.
['de', 'EMS', 'prep'].

'Kirchner'.
['Kirchner', 'EMS', 'snn'].

':
';:
['cc', 'EMS', 'coma'].

'se'.
['se' , mi].

'fue'.
['fue' , mi].

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

'dormir'.
['dormir' , mi].

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

'su casa'.
['su casa', 'EMS', 'snn'].

':
';:
['cc', 'EMS', 'coma'].

'ubicada'.
['ubicada' , mi].

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

'sólo'.
['sólo', 'EMS', 'adv'].

'cuatro cuadras'.
['cuatro cuadra', 'EMS', 'snn'].

'del'.
['del' , mi].

'hospital'.
['hospital', 'EMS', 'snn'].

':
';:
['pf', 'EMS', 'ponc'].

%fin de la phrase numero : 14

'Ayer'.
['Ayer' , mi].

'lo'.
['el', 'TDET', 'art'].

'fue'.
['fue', 'mi'].

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

'visitar'.
['visitar', 'mi'].

'después'.
['después', 'EMS', 'snn'].

'del'.
['del', 'mi'].

'mediodía'.
['mediodía', 'EMS', 'snn'].

'pero'.
['pero', 'mi'].

'entró'.
['entró', 'mi'].

'por'.
['por', 'EMS', 'prep'].

'una puerta'.
['una puerta', 'EMS', 'snn'].

'lateral'.
['lateral', 'TADJ', 'adj2'].

'para'.
['para', 'EMS', 'prep'].

'eludir'.
['eludir', 'mi'].

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

'los periodistas'.
['el periodista', 'EMS', 'snn'].

'que'.

['que' , mi].

'hacían'.
['hacían' , mi].

'guarida'.
['guarida', 'EMS', 'snn'].

'en'.
['en', 'EMS', 'prep'].

'la acceso'.
['el acceso', 'EMS', 'snn'].

'principal'.
['principal', 'EMS', 'adj', 'NUM', 'sg', 'TADJ', 'adj1'].

'::'.
['pf', 'EMS', 'ponc'].

%fin de la phrase numero : 15

'El primero'.
['el primero', 'EMS', 'snn'].

'en'.
['en', 'EMS', 'prep'].

'visitarlo'.
['visitarlo' , mi].

'ayer'.
['ayer' , mi].

'fue'.
['fue' , mi].

'el secretario'.
['el secretario', 'EMS', 'snn'].

'de'.
['de', 'EMS', 'prep'].

'Legal'.
['Legal', 'EMS', 'snn'].

'y'.
['y' , mi].

'Técnica'.

['Técnica', 'EMS', 'snn'].

''

['cc', 'EMS', 'coma'].

'Carlos Zanini'.

['Carlos Zanini', 'EMS', 'snn'].

''

['pf', 'EMS', 'ponc'].

%fin de la phrase numero : 16

En el fragmento catorce, los snn: [algunos amigos], [su esposa], [la senadora], [El primero], resaltados debidamente.

'También'.

['También', 'EMS', 'snn'].

'estuvieron'.

['estuvieron' , mi].

'a'.

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

'su lado'.

['su lado', 'EMS', 'snn'].

'su hijo'.

['su hijo', 'EMS', 'snn'].

'Máximo'.

['Máximo', 'EMS', 'snn'].

'su hija'.

['su hijo' 'EMS', 'snn'].

'Florencia'.

['Florencia', 'EMS', 'snn'].

'está'.

['está' , mi].

'en'.

['en', 'EMS', 'prep'].

'Buenos Aires'.

['Buenos_Aires' , 'EMS', 'snn'].

'su secretario'.

['su secretario', 'EMS', 'snn'].

'privado'.

['privado', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg', 'TADJ', 'adj1'].

':':

['cc', 'EMS', 'coma'].

'Daniel Muñoz'.

['Daniel Muñoz', 'EMS', 'snn'].

':':

['cc', 'EMS', 'coma'].

'y'.

['y', mi].

'el gobernador'.

['el gobernador', 'EMS', 'snn'].

'de'.

['de', 'EMS', 'prep'].

'Santa Cruz'.

['Santa_Cruz', 'EMS', 'snn'].

':':

['cc', 'EMS', 'coma'].

'Sergio Acevedo'.

['Sergio Acevedo', 'EMS', 'snn'].

':':

['pf', 'EMS', 'ponc'].

%fin de la phrase numero: 17

En el fragmento quince, debía reconocer los snn: [su hijo], [Máximo], [su hija], [Florencia], [su secretario], [Daniel Muñoz], [el gobernador], [Sergio Acevedo], si bien hay coordinaciones y aposiciones, las reglas nos permiten reconocerlos.

'No'.

['no', 'EMS', 'adv'].

'pudieron'.

['pudieron', mi].

'verlo'.

['verlo', mi].

'muchos'.
['muchos', 'EMS', 'snn'].

'más'.
['más', 'EMS', 'adv'].

'porque'.
['porque', 'mi'].

'tiene'.
['tiene', 'mi'].

'las visitas'.
['el visita', 'EMS', 'snn'].

'restringidas'.
['restringido', 'EMS', 'adj', 'GEN', 'fem', 'NUM', 'pl', 'TADJ', 'adj1'].

'!'.
['pf', 'EMS', 'ponc'].

%fin de la phrase numero : 18

'Más allá de'.
['más_allá_de', 'EMS', 'prep'].

'la guardia'.
['el guardia', 'EMS', 'snn'].

'periodística'.
['periodística', 'mi'].

'en'.
['en', 'EMS', 'prep'].

'la puerta'.
['el puerta', 'EMS', 'snn'].

'del'.
['del', 'mi'].

'hospital'.
['hospital', 'EMS', 'snn'].

'no'.
['no', 'EMS', 'adv'].

'hubo'.
['hubo', 'mi'].

'manifestaciones'.
 ['manifestación', 'EMS', 'snn'].

'y'.
 ['y' , mi].

'en'.
 ['en', 'EMS', 'prep'].

'Río Gallegos'.
 ['Río_Gallegos', 'EMS', 'snn'].

'se'.
 ['se' , mi].

'respiró'.
 ['respiró' , mi].

'el clima'.
 ['el clima', 'EMS', 'snn'].

'de'.
 ['de', 'EMS', 'prep'].

'un sábado'.
 ['un sábado', 'EMS', 'snn'].

'más'.
 ['más', 'EMS', 'adv'].

'!'.
 ['pf', 'EMS', 'ponc'].

%fin de la phrase numero: 19

En el fragmento dieciséis, se detecta el snn: [muchos], resaltado con color.

'El médico'.
 ['el médico', 'EMS', 'snn'].

'personal'.
 ['personal', 'TADJ', 'adj2'].

'de'.
 ['de', 'EMS', 'prep'].

'Kirchner'.
 ['Kirchner', 'EMS', 'snn'].

'señaló'.

['señaló' , mi].

'que'.
['que' , mi].

'se'.
['se' , mi].

'le'.
['le' , mi].

'recetó'.
['recetó' , mi].

'un'.
['un', 'TINDF1', 'indf1a'].

'plan'.
['plan', 'EMS', 'snn'].

'de'.
['de', 'EMS', 'prep'].

'nutrición'.
['nutrición', 'EMS', 'snn'].

'para'.
['para', 'EMS', 'prep'].

'este tipo'.
['este tipo', 'EMS', 'snn'].

'de'.
['de', 'EMS', 'prep'].

'patologías'.
['patología', 'EMS', 'snn'].

'y'.
['y' , mi].

'reconoció'.
['reconoció' , mi].

'que'.
['que' , mi].

'la nutricionista'.
['el nutricionista', 'EMS', 'snn'].

'del'.
['del' , mi].

'hospital'.
['hospital', 'EMS', 'snn'].

""no'.
[""no' , mi].

'es'.
['es' , mi].

'Petrona C'.
['Petrona_C', 'EMS', 'snn'].

'!'.
['pf', 'EMS', 'ponc'].

%fin de la phrase numero : 20

'de'.
['de', 'EMS', 'prep'].

'Gandulfo".'.
['Gandulfo".' , mi].

En el fragmento diecisiete, el snn es: [El médico], [la nutricionista] que aparece resaltado.

'El menú'.
['el menú', 'EMS', 'snn'].

'permitido'.
['permitido', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg', 'TADJ', 'adj1'].

'incluye'.
['incluye' , mi].

'nada más que'.
['nada_más_que', 'EMS', 'adv'].

'carnes'.
['carne', 'EMS', 'snn'].

'desgrasadas'.
['desgrasado', 'EMS', 'adj', 'GEN', 'fem', 'NUM', 'pl', 'TADJ', 'adj1'].

';'.
['cc', 'EMS', 'coma'].

'verduras'.
['verdura', 'EMS', 'snn'].

'hervidas'.
['hervido', 'EMS', 'adj', 'GEN', 'fem', 'NUM', 'pl', 'TADJ', 'adj1'].

'y'.
['y' , mi].

'gelatinas'.
['gelatina', 'EMS', 'snn'].

':'.
['pf', 'EMS', 'ponc'].

%fin de la phrase numero : 21

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

'una pregunta'.
['una pregunta', 'EMS', 'snn'].

'puntual'.
['puntual', 'EMS', 'adj', 'NUM', 'sg', 'TADJ', 'adj1'].

','.
['cc', 'EMS', 'coma'].

'y'.
['y' , mi].

'para'.
["'para' , 'EMS', 'prep'].

'que'.
['que' , mi].

'no'.
['no', 'EMS', 'adv'].

'se'.
['se' , mi].

'empiecen'.
['empiecen' , mi].

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

'hablar'.
['hablar' , mi].

'tonterías'.
['tontería' , 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg'].

En el fragmento dieciocho, el snn es [El menú], resaltado en color.

'Buonomo'.
['Buonomo', 'EMS', 'snn'].

'aclaró'.
['aclaró' , mi].

'que'.
['que' , mi].

'Kirchner'.
['Kirchner', 'EMS', 'snn'].

'nunca'.
['nunca' , mi].

'tuvo'.
['tuvo' , mi].

'una afección'.
['una afección', 'EMS', 'snn'].

'estomacal'.
['estomacal' , 'TADJ', 'adj2'].

'con'.
['con', 'EMS', 'prep'].

'el único antecedente'.
['el único antecedente', 'EMS', 'snn'].

'de'.
['de', 'EMS', 'prep'].

'un colon'.
['un colon', 'EMS', 'snn'].

'irritable'.
['irritable' , mi].

'Siempre'.
['Siempre' , mi].

'ejerció'.
['ejerció' , mi].

'y'.
['y' , mi].

'está'.
['está' , mi].

'ejerciendo'.
['ejerciendo' , mi].

'el mando'.
['el mando', 'EMS', 'snn'].

'en'.
['en', 'EMS', 'prep'].

'contacto'.
['contacto', 'EMS', 'snn'].

'permanente'.
['permanente', 'EMS', 'adj', 'GEN', 'masc', 'NUM', 'sg', 'TADJ', 'adj1'].

'con'.
['con', 'EMS', 'prep'].

'el Gabinete'.
['el gabinete', 'EMS', 'snn'].

'y'.
['y' , mi].

'tomando'.
['tomando' , mi].

'decisiones'.
['decisión' , 'EMS', 'snn'].

'dijo'.
['dijo' , mi].

'el médico'.
['el médico', 'EMS', 'snn'].

'.'.
['pf', 'EMS', 'ponc'].

%fin de la phrase numero: 22

En el fragmento diecinueve, los snn: [Buonomo], [Kirchner] y [el médico]

'Sin embargo'.

['sin_embargo', 'EMS', 'adv'].

'los funcionarios'.

['el_funcionario', 'EMS', 'snn'].

'que'.

['que' , mi].

'lo'.

['el', 'TDET', 'art'].

'visitaron'.

['visitaron' , mi].

'ayer'.

['ayer' , mi].

'afirmaron'.

['afirmaron' , mi].

'que'.

['que' , mi].

'lo'.

['el', 'TDET', 'art'].

'mantienen'.

['mantienen' , mi].

'alejado'.

['alejado' , mi].

'del'.

['del' , mi].

'teléfono'.

['teléfono', 'EMS', 'snn'].

':':

['pf', 'EMS', 'ponc'].

%fin de la phrase numero: 23

'Hoy'.

['Hoy' , mi].

'a'.

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

'las'.

['el', 'TDET', 'art'].

'10'.

['10', 'mi'].

'habrá'.

['habrá', 'mi'].

'un nuevo parte'.

['un nuevo parte', 'EMS', 'snn'].

'médico'.

['médico', 'EMS', 'snn'].

':':

['pf', 'EMS', 'ponc'].

%fin de la phrase numero : 24

En el fragmento veinte el snn es [los funcionarios].

'A'.

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

'las'.

['el', 'TDET', 'art'].

'11,'.

['11,', 'mi'].

'el obispo'.

['el obispo', 'EMS', 'snn'].

'de'.

['de', 'EMS', 'prep'].

'Santa Cruz'.

['Santa_Cruz', 'EMS', 'snn'].

':':

['cc', 'EMS', 'coma'].

'Alejandro Buccolini'.

['Alejandro Buccolini', 'EMS', 'snn'].

':':

['cc', 'EMS', 'coma'].

'dará'.
['dará' , mi].

'una misa'.
['una misa', 'EMS', 'snn'].

'para'.
['para', 'EMS', 'prep'].

'los internados'.
['el internado', 'EMS', 'snn'].

'del'.
['del' , mi].

'hospital'.
['hospital', 'EMS', 'snn'].

','.
['cc', 'EMS', 'coma'].

'por'.
['por', 'EMS', 'prep'].

'la'.
['el', 'TDET', 'art'].

'Pascua.'.
['Pascua.' , mi].

En el fragmento veintiuno, los snn son: [el obispo] y [Alejandro Buccolini]

'Nadie'.
['nadie' , 'EMS', 'snn'].

'se'.
['se' , mi].

'animó'.
['animó' , mi].

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

'anticipar'.
['anticipar' , mi].

'si'.
['si' , mi].

'participará'.
['participará' , mi].

'el Presidente'.
['el presidente', 'EMS', 'snn'].

':':
['pf', 'EMS', 'ponc'].

%fin de la phrase numero: 25

En el fragmento veintidós debía reconocer dos snn: [nadie] y [el Presidente], efectivamente resaltados.

Bibliografía

Abbaci, F. (1999) *Développement du Module Post-Smorph*. Mémoire de DEA de Linguistique et Informatique. GRIL, Université Blaise-Pascal, Clermont-Fd.

Abney, Steven (1991). *Parsing by Chunks*. En Berwick et al. (1991).

Aït-Mokhtar S. (1998): *L'analyse presyntaxique en une seule étape*. Tesis de Doctorado, GRIL, Univ. Blaise Pascal.

Beesley K. R., Karttunen L. (2003): *Finite State Morphology*, CSLI Publications.

Berwick, R. , Abney, S. y Tenny C. (eds.).(1991) *Principle-Base Parsing*. Kluwer Academic Publishers, Dordrecht.

Bès G. G. (1999): *La phrase verbale noyau en français*. En: Recherches sur le français parlé, Volume 15.

Bès G. G., Hagège C., Coheur L. (1999): “*De la description des propriétés linguistiques à l'analyse d'une langue*”. En Proceedings of VEXTAL '99, Venezia.

Bès, G.G. Hagège, C. (2001) *Properties in 5 P. Technical report, Groupe de Recherche dans les Industries de la Langue* (GRIL), URL.

Bès G. G. (2002): “*La linguistique entre science et ingénierie*”. En: TAL Vol. 41 n. 3.

Bès G. G. ,Dahl V. (2003): “*Balanced parentheses in NL texts: a useful cue in the syntax / semantics interface*” . En Nancy Workshop: Prospects and Advances in the Syntax / Semantics Interface.

Bès G. G. , Dahl V., Guillot D., Lamadon L., Milutinovici I., Paulo J.(2003): “*A parsing system for balanced parentheses in NL texts*” . En Nancy Workshop

Bès G. G., Solana Z. (2004/a) *Los clíticos en español*. Congreso Internacional Políticas de la Cultura, UBA Buenos Aires.

Bès G.G., Solana Z. (2004/b) “*Análisis morfológico y gramáticas locales: introducción y una aplicación concreta*” En: Jornadas Argentinas de Lingüística Informática Modelización e Ingeniería, JALIMI 2004.

Bès G.G., Solana Z. (2004/c) "*Extracción del sintagma verbal núcleo y resolución de ambigüedades en la asignación categorial*". En: Revista de Letras Nº 9. Facultad de Humanidades y Artes. Laborde Editor.

Beltrán- Bès (2005) "*La obtención de límites de oraciones*" en Desarrollo, implementación y utilización de modelos para el procesamiento automático de textos, compilado por Víctor Castel.

Bochensky, I. M. (1958). *Los métodos actuales del pensamiento*. 2ª Edición. Madrid. Rialp, Traducción al español de Raimundo Drudis Baldrich.

Coheur L., Mamede N., Bès, G. G. (2003): "*AsdeCopas: a syntactic-semantic interface*". En EPIA03 Workshop on Natural Language and Text Retrieval, Evora, Portugal, 2003.

Gala-Pavia N. (1999): *Using the Incremental Finite-State Architecture to Create a Spanish Shallow Parser*, SEPLN.

Gazdar, G. , Klein, E. , Pullum, G. , Sag, I. (1985) *Generalized Phrase Structure Grammar*. Oxford, Blackwell.

Gelbukh and Bolshakov (1999), *Avances en Análisis Automático de Textos- Proc. Foro: Computación, de la Teoría a la Práctica*. IPN, México City. May 26-28, 1999.

Hagège C. (2000): *Analyse syntaxique automatique du portugais*. Tesis de Doctorado GRIL, Univ. Blaise Pascal.

Hagège, Meireles, Trindade, Diogo, Leite (2002) *Taggers para o Português-Workshop*. Analisador morfosintáctico do ILTEC-

Hagège, Carolina *SMORPH: Um analisador/Girador morológico para O português* Taggers para o Português- Workshop.

Caroline Hagège, Nuno Nemedé & Joana Paulo, *Pasmo- pós análise morfológica*, Technical report, INESC-ID, Lisboa.

Hagège C., Bès G. G. (2002): *Encoding and reusing linguistic information expressed by linguistic Properties*. En COLING, Proceedings of the Workshop on Grammar Engineering and Evaluation, Taiwan.

Jrvinen T., Tapanainen P. (1998): «*Towards an Implementable Dependency Grammar*». En Proceedings of the COLING-ACL98

Manning C. D., Schuetze H. (1999): *Statistical Natural Language Processing*, MIT Press, Massachusetts.

Rodrigo Mateos, J. L., Bès, G.G. (2004): *Análisis e implementación de clíticos en una herramienta declarativa de tratamiento automático de corpus*. VI Congreso de Lingüística General, Santiago de Compostela, España, 2004.

Solana Z., Rodrigo A. (2005) "*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. ISBN 987-575-019-0 del soporte *Internet*.

Solana Z., et al. (2006) *Morfología del verbo español*, Colección Infosur, Centro de Estudios de Adquisición del Lenguaje, Facultad de Humanidades y Artes, UNR, Ediciones Juglaría, Rosario.