



**PONTIFICIA
UNIVERSIDAD
CATÓLICA DE
VALPARAÍSO**

**ILCL
INSTITUTO DE
LITERATURA Y
CIENCIAS DEL
LENGUAJE**



**Análisis de los predicados del dominio médico
desde el marco de la Léxico-Gramática.
Proyecciones para el reconocimiento automático de
textos.**

Alumna: Constanza Suy Álvarez
Profesor Guía: Dr. Walter Adrián Koza

Trabajo Final para optar al grado de Licenciado en Lengua y Literatura Hispánica
Becaria Proyecto Fondecyt 1171033

Viña del mar, enero 2019

ÍNDICE

INTRODUCCIÓN	3
1.- MARCO TEÓRICO	6
1.1.- Léxico-Gramática (LG)	6
1.1.1.- Antecedentes de la Léxico-Gramática	6
1.1.2.- Características fundamentales de la Léxico-Gramática	8
1.2.- Análisis distribucional de la Léxico-Gramática	9
1.2.1.- Clases de objetos.....	11
1.2.2.- Estructura Argumental	12
1.3.- Análisis transformacional de la Léxico-Gramática	13
1.5.- Trabajos investigativos previos	17
2.- MARCO METODOLÓGICO	20
2.1.- Preguntas de investigación	21
2.2.- Objetivos	21
2.2.1.- Objetivo general.....	21
2.2.2.- Objetivos específicos	21
2.3.- Tipos de investigación	21
2.4.- Plataforma NooJ	21
2.5.- Tareas desarrolladas	22
2.5.1.- Selección de verbos	23
2.5.2.- Identificación de clases de objetos.....	24
2.5.3.- Tabla del análisis distribucional de la Léxico-Gramática	26
2.5.4.- Tabla del análisis transformacional de la Léxico-Gramática.....	29
2.5.5.- Gramática de detección automática	31
2.5.6.- Gramática de generación transformativa	45
3.- RESULTADOS	51
3.1.- Clases de objetos	51
3.2.- Estructuras argumentales	54
3.3.- Transformaciones	56
3.4.- Detección automática de EA	58
3.5.- Generación transformacional de oraciones	59

4.- CONCLUSIONES	61
5.- REFERENCIAS BIBLIOGRÁFICAS	64

INTRODUCCIÓN

La presente investigación toma como base de análisis los predicados del dominio médico. La elección para trabajar con dicho dominio no fue arbitraria, además de considerar la medicina como una disciplina fundamental dentro de nuestra sociedad, lo cual se ve reflejado en el aumento progresivo de textos que desean difundir este conocimiento científico a una gran número de beneficiarios (Van Der Stuyft, Soto & Solari, 2016), la principal característica de los dominios de conocimiento recae en que se presentan como parcelas organizadas de la lengua lo que facilita el análisis y posterior clasificación. (Lamiroy, 1991)

Por tanto, este trabajo investigativo se realiza desde el marco de la Léxico-Gramática, modelo descriptivo que fue desarrollado por el lingüista francés Maurice Gross (1975; 1984) que toma como unidad mínima de significado la oración simple. Para ello, se deben elaborar tablas respectivas a un análisis distribucional y otro transformacional que analicen las unidades léxicas de esta oración. Este modelo busca describir de la manera más completa posible las posibilidades combinatorias de las palabras del léxico en el análisis de una oración simple (Lamiroy, 1991)

A partir de este trabajo investigativo, se describirán las estructuras argumentales (EA) de los verbos del área de la medicina considerando la combinación de las clases de objeto que se pudieron determinar mediante el análisis de un corpus especializado. A su vez, se identificarán las posibilidades transformacionales de cada una de las oraciones correspondiente a las estructuras argumentales que fueron categorizadas. Teniendo este sustento teórico, esta investigación se enmarca dentro de la lingüística computacional porque pretende realizar una proyección de detección automática de una estructura argumental y, también, la generación automática de oraciones.

Los apartados posteriores que se presentan en el trabajo son los siguientes: en el apartado 1, marco teórico, se detallan los sustentos teóricos que fundamentaron la investigación como la Léxico-Gramática, sus antecedentes y características principales; el análisis distribucional de este modelo, el concepto de clases de objeto y estructura argumental y el

análisis transformacional de la Léxico-Gramática. Además, se mencionarán ciertos trabajos investigativos que funcionaron como un referente a este.

En el apartado 2, metodología, se indican las preguntas de investigación, el objetivo general y los específicos, el tipo de investigación correspondiente y las tareas que se desarrollaron, tanto las teóricas (selección de verbos, identificación clases de objetos, análisis distribucional y análisis transformacional de la Léxico-Gramática) como las del trabajo computacional, que consiste en la elaboración de una gramática sintáctica a partir de una estructura argumental para su detección automática y generación de oraciones a partir de las transformaciones aceptadas utilizando la plataforma NooJ.

Posteriormente, en el apartado 3 de resultados, se exponen los datos que responden las preguntas de investigación y que se condicen con los objetivos planteados, estos son la determinación de las clases de objetos, las estructuras argumentales y transformaciones de acuerdo a los verbos de medicina seleccionados y los resultados de la proyección que se realizó para la detección y generación automática de una estructura argumental en específico.

Finalmente, en el apartado 4, se dan a conocer las conclusiones de este trabajo tomando en cuenta las preguntas y objetivos de investigación, sus futuras proyecciones y las limitaciones que este trabajo presentó

1.- MARCO TEÓRICO

En este capítulo se presenta el marco teórico que sustenta la propuesta de investigación planteada en la presente investigación. Esta se compone por los siguientes apartados: en primer lugar, se define el marco de la Léxico-Gramática (LG) según los planteamientos de Maurice Gross (1975; 1984), sus antecedentes y características más relevantes; en segundo lugar, se explicará el análisis distribucional de la Léxico-Gramática, a partir de la estructura argumental, y el concepto de clases de objeto, noción que fue desarrollada fundamentalmente por Gaston Gross (2014); en tercer lugar, se detallará el análisis transformacional de la Léxico-Gramática; finalmente, en cuarto lugar, se expondrán algunas investigaciones previas que fueron un referente para este trabajo.

1.1. Léxico-gramática (LG)

La léxico-gramática (*lexique grammair*e) es un modelo descriptivo basado en los trabajos de Zellig Harris (1951; 1954), pero que fue desarrollado por Maurice Gross (1975; 1984) en sus aplicaciones al francés. Este modelo estudia el conjunto de las reglas formales que definen un idioma, o sea, su gramática, y las entradas léxicas a las cuales se aplica el sistema de reglas. (Lamiroy, 1991)

Los fundamentos iniciales de este modelo fueron expuestos por Gross en su obra *Méthodes en syntaxe* (1975) el objetivo principal del autor fue llevar a cabo una descripción exhaustiva del francés con relación a un léxico dado y de las propiedades distribucionales y transformacionales de la oración simple (Garrido, 1999). Por tanto, para el autor una descripción lingüística sistemática es entendida “no solo como la enumeración estructurada de las reglas gramaticales que definen un idioma, sino también la representación exacta de la extensión léxica que caracteriza estas reglas” (Lamiroy, 1991:12).

1.1.1 Antecedentes de la Léxico-Gramática

Para poder entender en profundidad este modelo descriptivo desarrollado por Gross (1975) se hace necesario referirnos a los antecedentes previos, el más relevante recae en los trabajos del teórico del distribucionalismo Zellig Harris (1951), quien constituye el tratado clásico de la lingüística distribucional. Harris, con fin de superar y corregir el

descriptivismo bloomfieldiano —que para determinar las estructuras fonológicas o sintácticas de una lengua utilizaba solo dos criterios: el del sentido y de distribución— decide fundar toda descripción científica únicamente en el inventario de distribución de las unidades con fin de llegar al máximo de objetividad. (Mounin, 1977)

El autor comprende la distribución de una unidad como “la suma de todas sus partes” y señala cuatro fundamentos para considerar las estructuras distribucionales del lenguaje. En primer lugar, indica que los elementos de un lenguaje no son dados arbitrariamente, es decir, cada elemento está en cierta posición debido a que se relaciona con otro elemento que, a su vez, se encuentra en otra posición. En segundo lugar, todos los elementos del lenguaje pueden ser agrupados en clases y, por consiguiente, la restricción distribucional de clases persistirá para todas las ocurrencias, por lo que esta tampoco presenta un carácter arbitrario. En tercer lugar, es posible estimar las ocurrencias de cualquier elemento en relación con otro, sin que se necesite información extra, por ejemplo, se han hallado regularidades de ocurrencias en jergas a partir de la distribución de sus elementos sin la necesidad de recurrir a información histórica sobre la derivación de ciertas palabras. En cuarto lugar, las restricciones relativas a las ocurrencias de cada elemento son más bien descritas por una red de enunciados interrelacionados y no por la simple medida del total de restricciones de cada elemento de forma separada. (Harris, 1954)

M. Gross (1975; 1984) basa su modelo en la noción de transformaciones, que también desarrolla Harris. El transformativismo que este plantea no está orientado hacia la construcción de un modelo hipotético-deductivo que cubra la producción de todas las estructuras del lenguaje, sino que este queda “a ras de texto, tan próximo como es posible a la realidad de enunciados”. (Mounin, 1977: 214). El método que utiliza Harris es cuestionar en qué medida es posible reducir clase de enunciados concretos a fórmulas algebraicas. A partir de esta idea, el autor llega a una serie de conclusiones: considera la gramática como el conjunto de reglas que generan las oraciones de una lengua, elabora la noción de oración nuclear (*kernel*) y *transform* para luego concebir la estructura de una lengua como el conjunto de oraciones nucleares (*kernel-sentences*), más un conjunto de transformaciones (*transforms*). Por ello, las clases sintagmáticas que se juntan para formar núcleos permutan

para realizar operaciones sin cambiar apreciablemente el sentido de los enunciados. (Mounin, 1977). Ejemplo:

(1) 01: María lee un libro. → 02: El libro es leído por María.

También parece oportuno detenerse en el modelo de gramática generativa, que al igual que la Léxico-Gramática fue desarrollado teniendo como base los trabajos de Harris. Lamiroy (1991) señala que:

“Si, para Chomsky, el objetivo principal de la investigación lingüística es la construcción de una sistema muy formalizado llamado *Gramática Universal*, que es a su vez el reflejo de la competencia lingüística mental del ser humano, para Gross la actividad de formalización es subsidiaria de la necesidad imperativa de acumular, de la forma más completa posible, datos empíricos que verifiquen las reglas formales.” (p.12)

Por tanto, el mayor punto de discusión entre los adherentes de la Léxico-Gramática y los seguidores de Chomsky apunta a la importancia que los unos y los otros atribuyen a la comprobación empírica de las teorías formales que ambos lingüistas proponen.

1.1.2 Características fundamentales de la Léxico-Gramática

Para llevar a cabo este modelo, Gross (1975) considera a la oración simple (sujeto – verbo – objeto) como la unidad mínima de significado, por consiguiente, excluye trabajar con palabras y sintagmas. A su vez, señala que son tres los componentes principales para realizar dicho análisis: (i) oraciones libres, que se construyen por medio de los verbos predicativos; (ii) locuciones o expresiones idiomáticas; (iii) verbos soportes. (Gross, 1984). Cabe señalar que en esta investigación solo se analizarán las oraciones libres, quedando para futuros trabajos los estudios con locuciones y verbos soportes.

El autor francés destierra cualquier tipo de formalismo y considera que será el examen sistemático de los datos lo que permitirá establecer principios, para ello, y como fue señalado, se centra en la oración simple y en las transformaciones que esta pueda asumir. En el análisis se evidenciarán elementos operativos y propiedades sintácticas comunes

mediante la comparación de frases emparentadas, por lo que este modelo también alude a los significados obtenidos de las posibles oraciones. (Garrido, 1999)

A partir de lo anterior, es posible resumir las características principales de la Léxico-Gramática en los siguientes tres puntos: (i) se concibe la sintaxis como un elemento que no puede ser separado del lexicón; (ii) la identificación de la oración simples (excluyendo el sintagmas y palabras) como unidad mínima de análisis sintáctico y semántico; (iii) el modelo ofrece una formalización y una descripción metodológica que puede ser aplicable a cualquier idioma. (Elia, Monteleone & Marano, 2011).

En conclusión, la Léxico-Gramática se trata de la descripción más completa posible de las posibilidades combinatorias de las palabras del léxico en el interior de una oración simple. Para ello, se emplea un análisis sintáctico consistente en la descripción de dos tipos de propiedades: las distribucionales y las transformacionales, dicha información será presentada mediante tablas en la que si un elemento léxico admite una propiedad se indicará con un signo de suma (+) o con un signo negativo (-) en caso contrario (Tolone, 2012). Estos dos análisis serán explicados en los siguientes apartados.

1.2 Análisis distribucional de la Léxico-Gramática

El análisis distribucional de la Léxico-Gramática consiste en determinar las propiedades distribucionales según los requerimientos argumentales correspondientes. Desde un punto de vista semántico y teórico, un grupo nominal será un argumento si figura en el ámbito estricto del predicado, por lo que los complementos circunstanciales quedan fuera de esta clasificación. (G. Gross, 2014).

Esta información es presentada por una matriz binaria en la que las filas presentarán los verbos y en las columnas cada oración posible con aquella unidad. Cuando el verbo admite una propiedad se añadirá un signo (+) en la intersección correspondiente, en caso contrario, se indicará que dicho verbo no acepta tal propiedad marcando con un signo (-) (Gross, 1984).

A partir de este análisis se pueden determinar las estructuras argumentales y también los posibles significados que el verbo adquiere mediante la combinación de sus argumentos lo

que permite dar cuenta de la polisemia de las distintas entradas léxicas, tal como se evidencia con el verbo ‘comer’ en siguiente tabla:

Argumento 0		Lema	Argumento 1		Ejemplo
+ANIM	-ANIM		+ALIMENTO	-ALIMENTO	
+	-	comer	+	-	El perro comió un filete.
-	+	comer	-	+	El óxido comió el metal.

Tabla 1. Fragmento de una tabla LG de preferencias de selección para el verbo ‘comer’ (Koza, en prensa)

En la tabla anterior, se distinguen los dos posibles significados para el lema ‘comer’: el primero de ellos refiere a la acción de alimentarse (‘el perro comió un filete’) mientras que el segundo se relaciona con el proceso de corrosión de un metal (‘el óxido comió un metal’). Por tal motivo, las propiedades distribucionales también adquieren un valor clasificatorio en la medida que las restricciones argumentales pueden ser decisivas para atribuir una construcción de un verbo determinado. Por ejemplo, el verbo ‘venir’ dependiendo de sus argumentos puede tratarse de un verbo de movimiento (2) o un verbo auxiliar (3). Por tanto, forman dos clases sintácticas distintas. (Lamiroy, 1991)

(2) La paciente va a la consulta

(3) La operación va a costar 700.000 pesos

Una clase sintáctica agrupa una serie de verbos que tienen por lo menos una propiedad distribucional en común, esta última se considera como el rasgo definitorio de la clase. Que los miembros de una clase no comparten necesariamente todas las propiedades es obvio, por lo que en cada clase existirá una cierta variación que debe ser registrada por la Léxico-Gramática. (Lamiroy, 1991)

1.2.1 Clases de objeto

Las clases de objeto son clases semánticas, es decir, son un conjunto de sustantivos que en una posición argumental específica determinan el significado de un verbo. Estas clases pueden corresponder a hiperclases <humanos> o a subconjuntos <pacientes>, <profesional de la medicina>. Las clases permiten poner en manifiesto los distintos significados, ya que estos no son definidos únicamente por el verbo y las posiciones argumentales, por lo que al pasar de una clase argumental a otra, el sentido del predicado puede cambiar. Para ejemplificar, evidenciamos que tanto la oración (5) como la oración (6) utilizan el mismo verbo ‘arrancar’, sin embargo, con distintos significados.

(5) Juan ha arrancado el coche

(6) El jardinero ha arrancado el arbusto

Si se reemplaza la palabra ‘coche’ por ‘limusina’, ‘moto’, ‘autobús’ el significado del verbo sigue constante, entonces se podría crear la clase de objeto <medios de transporte>. En esta clase se agruparán todos los sustantivos que tengan los mismos rasgos semánticos para que el verbo ‘arrancar’ signifique ‘conducir’. En el caso de la oración (6) la clase de objeto <medios de transporte> no puede ser aceptada, puesto que el verbo adquiere otro valor, significando ‘arrancar de raíz’. Si se sustituye ‘arbusto’ por ‘ficus’, ‘rosal’, ‘geranio’ notamos que el sentido de la oración se mantiene, por lo que es adecuado categorizar la clase de <plantas> para mantener la misma interpretación.

Entonces, las clases de objeto son agrupaciones de sustantivos que son semánticamente homogéneos, forman subconjuntos de las hiperclases y seleccionan una interpretación dada de un predicado entre otras posibles. Esta construcción está basada en propiedades gramaticales y no en una clasificación presintáctica como sucede con los árboles semánticos. La noción de clase de objeto se creó para describir con precisión los argumentos. Por ejemplo, señalar que el verbo ‘calzar’ admite el rasgo de ‘concreto’ resulta insuficiente, ya que se generaría una gran cantidad de frases inaceptables; en cambio, si se recurre a la clase <calzado> las posibles oraciones serían adecuadas para mantener un mismo sentido. (G.Gross, 2014)

Ahora bien, también se debe señalar que un mismo sustantivo puede pertenecer a dos clases de objetos distintas a partir del predicado de una estructura argumental determinada. De tal manera, ‘Coca-Cola’ puede ser clasificada como una clase de objeto <líquidos + bebibles> en la oración (7) mientras que en la oración (8) podría formar parte de la clase de objetos de <mercaderías> (Koza, en prensa)

(7) María bebió una Coca-Cola

(8) María vendió una Coca-Cola

1.2.2 Estructura argumental

Antes de definir el concepto de estructura argumental, se hace necesario explicitar la diferencia entre un predicado y un argumento. G. Gross (2014) define ambos conceptos de la siguiente manera:

“un predicado es una palabra que opera una selección determinada entre las restantes palabras del léxico para formar un esquema que sea la base de una aserción. Los argumentos son los elementos léxicos seleccionados por los predicados en el marco de una frase simple. No ejercen por si mismo una selección de una palabra” (26)

Ejemplificaremos dicha diferencia a partir de dos palabras: el sustantivo ‘artículo’ y ‘redactar’. Por una parte, ‘artículo’ puede figurar dentro de un esquema predicativo, pero no posee una estructura sistematizable debido a las posibilidades combinatorias, es decir, un artículo puede ser ‘leído’, ‘apreciado’, ‘escrito’, ‘citado’, etc. Por otro lado, las palabras como ‘redactar’ definen con precisión su entorno: a la izquierda del sustantivo debe presentarse otro sustantivo que designe a un humano (+Humano) y a la derecha un sustantivo que pertenezca a clases de objeto <textos/escritos> ya sea ‘novela’, ‘cuento’, ‘tesis’, etc.

Por tanto, la estructura argumental (EA) refiere al conjunto, ordenado o no, de sus argumentos. En la oración (11) se puede notar que el verbo ‘explicar’, por su significado, denota una acción que requiere la concurrencia de tres participantes: un agente que lleva

acabo la acción (el maestro), una información (la lección) y un destinatario al que esa acción se dirige (los estudiantes) (NGDLE, 2009).

(11) El maestro explicaba la lección a los estudiantes

Todos estos participantes que configuran la noción predicativa se denominan como argumentos, en este caso, este verbo requiere de tres argumentos. N0, el maestro; N1, la lección; y N2, los estudiantes.

Igualmente, los predicados son categorías que designan estados, acciones y propiedades en los que intervienen uno o varios participantes, así también ciertos sustantivos podrían formar un predicado obteniendo una EA (NGDLE, 2009) como también ocurre con las nominalizaciones deverbales que son construcciones que se caracterizan por presentar propiedades propias de los sustantivos, pero que también poseen la facultad de heredar la estructura argumental de los verbos de los que derivan (Perois, 2012). Otra de las grandes diferencias radica en que los predicados pueden tener una doble función de ser predicados y argumentos mientras que un argumento jamás podría ser catalogado como un predicado.

Tomando en cuenta que esta investigación se centra en los verbos que actúan como predicados, se tomará la definición de estructura argumental (EA) como la unión de los argumentos que son exigidos por el verbo en cuestión. Para complementar esta definición, Harris (1951; 1954) indica que una de las facultades que poseen los predicados es la inscribirse en el tiempo, o sea, la posibilidad de ser conjugados, lo que no se puede realizar con los argumentos (G. Gross, 2014). La gran mayoría de los verbos, con escasas excepciones como los auxiliares, tienen una estructura argumental.

1.3 Análisis transformacional de LG

El análisis transformacional de la Léxico-Gramática contempla las posibles transformaciones que admite una oración simple sin alterar su significado. Al igual que el análisis distribucional, esta información se puede presentar mediante matrices binarias que

indican la estructura sintáctica (columnas verticales) en las que cada entrada del léxico (líneas horizontales) puede aparecer¹.

Oración	Voz Pasiva	Nominalización
El perro comió un filete.	+	+
El óxido comió el metal.	+	-

Tabla 2. Ejemplo de Tabla LG de transformaciones. (Koza, en prensa)

En el ejemplo anterior, es posible distinguir los significados de las dos oraciones a partir de las posibilidades de transformación que dicha oración aceptó. La primera oración (11.a) admitió la transformación en voz pasiva mientras que la segunda (11.b) rechazó dicha transformación.

(12) a. 01: El perro comió un filete → 02: La comida del perro fue un filete.

b. 01: El óxido comió el metal → 02: *La comida del óxido fue el metal.

Por lo cual, M. Gross (1984) entiende la noción de transformación tal como lo hizo Harris (1951), puesto que este la considera como un dispositivo experimental destinado a manifestar, mediante frases análogas, las características sintácticas pertinentes de una estructura determinada. Lo anterior, el hecho de someter frases a operaciones transformacionales simples, permite poner de manifiesto los rasgos sintácticos de los verbos (Lamiroy, 1991)

Ejemplificaremos lo anterior con la transformación pasiva. El hecho de prever en la gramática una regla que produce una frase pasiva por transformación se basa en un doble argumento de economía: se puede concluir que la oración declarativa presenta una serie de

¹ La noción de transformación que plantea la Léxico-Gramática de M.Gross (1975) difiere de la GGT (Chomsky, 1957) en que las transformaciones se darían siempre en la estructura superficial.

restricciones y que su variante pasiva está sujeta a las mismas condiciones de gramaticalidad. Por tal motivo, si ambas oraciones fuesen analizadas de manera independiente, las restricciones aparecerían dos veces, lo cual sería redundante.

(13) El médico operó a la paciente.

(14) La paciente fue operada por el médico.

Dicho de otro modo, la solución transformacional consistirá en describir las restricciones de la oración simple, aplicar la operación pasiva y obtener una O2 sin afectar a las restricciones estipuladas. A su vez, se captará también que la transformación pasiva se aplica en principio a cualquier verbo transitivo, sin embargo, para la Léxico-Gramática la extensión de la transformación a través del léxico es considerado como un elemento que debe aparecer explícitamente en la descripción (Lamiroy, 1991).

Si bien es cierto que la base del análisis transformacional se justifica en un argumento de economía en el que las mismas restricciones que posee O1 las tendrá su O2 al realizar la transformación (Palma, 1994) M. Gross (1984) reconoce una serie de excepciones que deben ser tomadas en cuenta. Por ejemplo, la oración (15) pone en manifiesto una excepción de la transformación en pasiva:

(15) O1: La paciente merece mi atención. → O2: *Mi atención fue merecida por la paciente

Esta restricción es descubierta por medio del examen exhaustivo de los verbos transitivos que, a priori, admiten transformación pasiva. Se hace evidente que en este modelo el estudio léxico es tan fundamental como la gramática misma, por lo que, a diferencia de la lingüística tradicional, la descripción conjunta del léxico y la sintaxis es primordial y no son entendidas como disciplinas independientes.

Además de la transformación pasiva, G.Gross (2014) presenta las siguientes opciones transformativas.

1. La nominalización, en la que hay un cambio de categorías morfológicas de los predicados.

(16) 01: Agustín compró un libro. →02: La compra de un libro por Agustín

2. Las interrogaciones que son aplicadas al predicado (17.a) o a un argumento (17.b).

(17) a. 01: Agustín compró un libro. →02: ¿Quién compró un libro?

b. 01: Agustín compró un libro. →02: ¿Qué compró Agustín?

3. La negación que se aplica al predicado o a los argumentos mediante adverbios de negación

(18) 01: Agustín compró un libro. →02: Agustín no compró un libro.

Para esta investigación, se utilizaron las transformaciones señaladas anteriormente y otras que serán indicadas en el apartado de metodología.

2.5.- Antecedentes investigativos

2.5.1.- “Automatic Detection of Negated Findings in Radiological Reports for Spanish Language: Methodology Based on Lexicon-Grammatical Information Processing”

Los trabajos en español que se basan en la Léxico-Gramática son escasos, por tal motivo se hace necesario mencionar algunos antecedentes previos a la presente investigación. En esta ocasión se tomará como referencia el trabajo realizado por Koza, Filippo, Cotik, Stricker, V Muñoz, Godoy, Rivas y Martínez-Gamboa (2018) el cual presenta una metodología para el reconocimiento automático de hallazgos negados en reportes de radiografía, para ello se tomó en consideración información morfológica, semántica y sintáctica. Los autores desarrollaron un diccionario electrónico de terminología médica y gramáticas informáticas, la información fue extraída de la ontología SNOMED CT y la Real Academia Nacional de Medicina (2012); de la misma manera, incluyeron un diccionario general de la lengua. En dicho trabajo, emplearon la Léxico-Gramática como la base de datos que permitió una descripción exhaustiva de la estructura argumental de predicados proyectados por las unidades léxicas. Respecto al trabajo computacional, ellos utilizaron la plataforma NooJ.

El método que emplearon se divide en tres etapas: (i) las unidades léxicas y gramaticales involucradas fueron analizadas para especificar términos y desarrollar un diccionario electrónico, (ii) se formalizaron estructuras sintácticas respecto a la organización de hallazgos negados y (iii) la descripciones se implantaron computacionalmente y la eficacia de la metodología propuesta fue probada en un corpus compuestos por 1120 reportes radiológico obtenidos de la base de datos del Hospital Garrahan (Buenos Aires, Argentina).

Respecto a los resultados, los autores, junto con especialistas del tema, identificaron manualmente un total de 2075 oraciones; 535 de ellas presentan un hallazgo en su estructura argumental y en 404 de ellas hay una negación por adverbio ('no') y otras por el uso de una preposición ('sin'). Para probar la eficacia de esta propuesta metodológica el corpus fue analizado con NooJ, que reconoció automáticamente todas las oraciones que fueron etiquetadas como (SNHall) y otras que también presentaban negación. Luego, el siguiente paso que realizaron fue comparar los resultados de NooJ con los entregados por NegEx, otro software para análisis de textos de lenguaje natural. A partir de ello, el cálculo de los resultados se realizó de la siguiente manera: se consideró como positivos verdaderos (true positives) las oraciones que contenían hallazgos negados. NooJ logró reconocer 384 de 404 oraciones mientras que NegEx detectó 387. Los falsos positivos (false positives) considera a aquellas oraciones que contenían hallazgos no negados, pero que el algoritmo las etiquetaba como negadas. Respecto a esto, NegEx tuvo 5 errores mientras que NooJ no evidenció ningún error. Los verdaderos negativos (true negatives) refieren a las oraciones que no contenían negación y que fueron reconocidas como tal por el algoritmo, NooJ detectó 131 y NegEx 123. Los falsos negativos (false negatives) refieren a las oraciones sin negación, pero que fueron detectadas como negadas, en el que tanto NooJ como NegEx obtuvieron el mismo resultado.

Finalmente, los autores concluyen que tanto NegEx como NooJ mostraron resultados positivos respecto a la detección automática, superando un 90%, lo que beneficia el uso de metodologías basadas en expresiones regulares.

2.5.2.- Clinical Term Recognition: From Local to LOINC Terminology. An application for Italian Language

Otro trabajo que sirve como referente para esta investigación, debido a que utilizó la misma plataforma para el análisis computacional, es el de Parisi (2016) quien en su artículo presenta una metodología para detectar automáticamente el dominio específico de la terminología usada por los laboratorios clínicos y biológicos. La creación de las formas lingüísticas locales se deben mapear con la terminología estándar entregada por LOINC y a los códigos de identificación específicos. Estas construcciones lingüísticas pueden brindar un importante apoyo a los expertos para facilitar el proceso de codificación. La autora construyó una serie de gramáticas mediante las herramientas brindadas por la plataforma NooJ con fin de reconocer formas lingüísticas expresadas en la lengua italiana, por lo que pretende demostrar cómo los elementos lingüísticos que están contenidos en la definición de las pruebas clínicas pueden sugerir valores importante del estándar LOINC, así facilitando el proceso de codificación (por ejemplo, al usar el adjetivo ‘urinario’ en una prueba local sugiere el valor ‘orina’ para el parámetro ‘tipo de muestra’).

La metodología está enfocada en reconocer elementos lingüísticos específicos tales como adjetivos, sufijos, prefijos y acrónimos usados por los laboratorios clínicos que pueden asignar un valor, por ejemplo, el sufijo ‘-emia’ posee un valor en el que las palabras que las contengan pueden ser categorizadas bajo el parámetro de ‘sistema de tipo sanguíneo’ “Sistema de tipo sanguíneo”. El corpus con el que se trabajó se compuso a partir de 5 set de datos de laboratorios clínicos, con fin de identificar las variantes lingüísticas más importantes se consideró 30 tipologías de exámenes clínicos que son las más comunes y usadas por los laboratorios biológicos y clínicos, tales como: ‘glicemia’, ‘colesterol’, ‘hemoglobina’, etc. Para lograrlo, se tuvo que crear un diccionario en NooJ que describiera los acrónimos y los adjetivos que consideraron para la detección automática.

Entonces, el objetivo general del trabajo es crear una herramienta lingüística que permita mejorar el proceso de codificación LOINC. Para eso, se debió reconocer todas las formas lingüísticas locales y la información de parámetros que LOINC pueda sugerir, por lo que las gramáticas sintácticas creadas con NooJ identificaron las características morfológicas de las definiciones clínicas (ejemplo: los sufijos ‘-emia’ y ‘-uria’) y las partes lingüísticas (como ‘urinario’, ‘sierico’, ‘plasmático’, ‘ionizzato’) para deducir otros importantes valores para LOINC. Este reconocimiento es un apoyo importante para el proceso de codificación,

puesto que permite decidir correctamente la asignación de código considerando, particularmente, algunos los sufijos y adjetivos.

Para finalizar, la autora concluye que su trabajo ha evidenciado la importancia en el proceso de codificación, pero también los problemas a los que se enfrentan los expertos clínicos durante este proceso. De tal manera, queda clara la importancia que posee el dominio terminológico de un especialista para asignar un código correcto.

2. METODOLOGÍA

En este apartado se desarrollarán y describirán los pasos metodológicos que se emplearon para llevar a cabo esta investigación. En primer lugar, se presentarán las preguntas de investigación; en segundo lugar, los objetivos generales y específicos; y en tercer lugar, se presentarán las tareas desarrolladas tanto para el trabajo teórico como para el trabajo computacional.

2.1. Preguntas de investigación

1. ¿Cuáles son las clases de objetos que son tomados como argumentos de los verbos del área de medicina? ¿Cómo proyectan dichos verbos su estructura argumental (EA)?
2. ¿Cuáles son las posibles transformaciones de las estructuras argumentales de los verbos del área de medicina?
3. ¿Cómo modelizar las EA y las transformaciones de los verbos del área de medicina con vista a una implantación computacional para la generación y detección automática en textos de lenguaje natural?

2.2 Objetivos

Con fin de responder las interrogantes, se determinaron los siguientes objetivos que guiaron el proceso de esta investigación.

2.2.1 Objetivo general

- Caracterizar la Estructura Argumental proyectada por verbos del área médica y sus posibilidades transformacionales, a partir de la identificación de las clases de objetos que actúan en calidad de argumentos y la estructura argumental que proyectan, con vista a una implantación computacional para el análisis automático de textos del lenguaje natural.

2.2.2 Objetivos específicos

1. Establecer las clases de objeto que son seleccionados como argumentos por verbos del área de la medicina.

2. Describir las estructuras argumentales a partir de la combinación de las clases de objetos de los verbos del área de la medicina.
3. Establecer las posibilidades transformacionales de las estructuras argumentales de los verbos del área de la medicina.
4. Describir formalmente las estructuras argumentales de los verbos del área de medicina y sus transformaciones con vista a una implantación computacional para la generación y detección automática en textos de lenguaje natural

2.3.- Tipo de investigación

La presente investigación descriptiva está fundada, principalmente, sobre los principios planteados por Gross (1975; 1984) y su trabajo de la Léxico-Gramática. Del mismo modo, la investigación pretende determinar y describir las estructuras argumentales, clases de objeto y posibles transformaciones que se dan a partir de los verbos del área de medicina con vista a una implantación computacional, por lo que adquiere un enfoque cualitativo.

2.4. Plataforma NooJ

Para la realización de esta investigación se utilizó el software libre NooJ², creado por Max Silberztein en el año 2002, una herramienta informática útil para el tratamiento de textos de lenguaje natural. Actualmente es utilizada por diversos lingüistas e investigadores que comparten sus trabajos mediante foros y congresos anuales (Silberztein, 2016).

NooJ tiene diversas utilidades para el análisis de textos, como son la creación de gramáticas morfológicas, gramáticas productivas y gramáticas sintácticas, además de la creación de diccionarios. Todas estas maquinarias permiten desarrollar operaciones morfológicas mientras se realiza un análisis sintáctico, por ejemplo, la transformación de oraciones en frases pasivas o controlar las concordancias morfosintáctica (Tramallino, 2013) Cabe señalar que NooJ adhiere al concepto de transformación como lo plantea Harris (1951; 1954) y la Léxico-Gramática de Gross (1975), lo cual se corresponde con este trabajo investigativo.

Se explicará algunas de las principales herramientas que presenta el programa:

² <http://www.nooj4nlp.net>

- *Diccionarios (archivos.dic)*: Se elabora un listado de palabras o expresiones en las que se señala un conjunto de información respecto a la categoría a la que pertenecen uno o más paradigmas de inflexiones o derivaciones, por ejemplo, se indica la conjugación de verbos y se determinan las propiedades sintácticas y semánticas, a saber, clases distribucionales como +Humano. (Díaz, Koza, Méndez, Píppolo, Rivero, Rodrigo, Solana & Tramalino, 2014). Por ejemplo, la entrada de ‘cantar’ sería la siguiente: V+FLX=AMAR, lo que indica que ‘cantar’ es un verbo y que se conjuga de la misma manera que ‘amar’, a la vez, AMAR es el modelo que representa todos los verbos regulares de la primera conjugación.

El diccionario con el que se trabajó fue uno compilado en NooJ a partir del proyecto Fondecyt 111130469, el cual está conformado por lemas del Diccionario Esencial de la Lengua Española (RAE, 2006) junto con los nombres y adjetivos que forman parte del dominio médico y que fueron extraídos de los diccionarios Mosby (2005) y Terminología médica (Cárdenas, 2012). (Koza, 2017)

- *Gramáticas*: Son utilizadas para representar una serie de fenómenos lingüísticos, tanto en un nivel morfológico como a un nivel sintáctico. NooJ permite la creación de tres tipos de gramática:
 - Gramática morfológica o derivacionales (archivos.nof): Se elaboran este tipo de gramáticas para representar modelos de inflexiones, a saber, derivaciones o conjugaciones. Se declaran de forma gráfica o mediante reglas.
 - Gramática productivas (archivos.nom). Estas se usan para representar un conjunto de palabras y asociarlas con un tipo de información léxica, son sistemas regulares o gráficos útiles para el tratamiento de cadena de caracteres a partir de una serie de propiedades formales que se determinan. (Bonini, 2015)
 - Gramáticas sintácticas (archivos.nog): Son sistemas regulares o gráficos en función del tratamiento de cadena de caracteres conformadas por unidades léxicas (Koza, 2017). También, se emplean para extraer o desambiguar palabras

filtrando algunos ítems léxicos o anotaciones sintácticas en el texto. (Díaz et al, 2014).

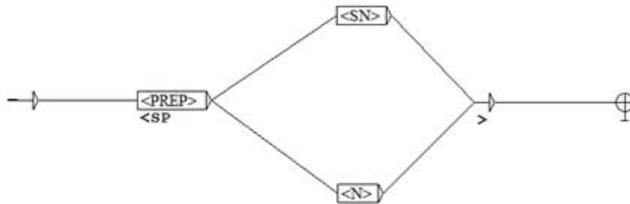


Imagen 1: Ejemplo de gramática sintáctica creada en NooJ (Bonini, Rodrigo, 2013)

La imagen anterior ejemplifica la manera en que se puede estructurar una gramática sintáctica, en esta se reconoce como SP cualquier secuencia que contenga una preposición y un sustantivo (<N> en la rama inferior del gráfico) o un sintagma nominal (<SN> en la superior); esta última rama presupone la aplicación previa de una gramática sintáctica que haya reconocido el SN.

2.5.- Tareas desarrolladas

En este punto se describen las tareas desarrolladas en esta investigación, comenzado por describir las tareas respectivas al trabajo teórico: selección de verbos a trabajar, identificación de clases de objetos y el análisis distribucional y transformacional de la Léxico-Gramática, mediante la elaboración de tablas. Posteriormente, se explican las tareas pertinentes al trabajo computacional, a saber: la creación de una gramática sintáctica que detecte una estructura argumental determinada dentro de un corpus, la que se denominará como *gramática de detección automática*, y otra que genere una serie de oraciones considerando las posibilidades transformativas de una estructura argumental en específico, la cual tendrá como nombre *gramática generativa transformativa*.

2.5.1. Selección de verbos

Para el desarrollo de esta investigación, en primer lugar, se extrajeron automáticamente los verbos del corpus (CCM-2009) el cual está conformado por casos clínicos que fueron

publicados en revistas médicas chilenas indizadas en la plataforma Scielo entre los años 1999 y 2008, fue compilado por Burdiles (2012) sumando un total de 969 textos, sin embargo, para esta investigación, y de acuerdo al FONDECYT 1171033, solo se tomó en consideración los textos especializados en obstetricia y ginecología. (Koza, 2017).

En segundo lugar, los verbos fueron clasificados apelando a un criterio lexicográfico, el cual consistió en considerar verbo del área médica a aquel que estuviera incluido en el diccionario de la especialidad, en este caso, se recurrió a la Real Academia Nacional de Medicina (RANM, 2012).

Finalmente, en tercer lugar, se elaboró una lista de 100 verbos tomando en cuenta la frecuencia dentro del corpus médico. Debido a que esta investigación trabaja con oraciones predicativas³, se excluyeron de la lista los verbos soportes. El resultado de dicho proceso se puede constatar en la siguiente tabla:

1.- Requerir	26.-Explicar	51.-Comentar	76.-Cambiar
2.- Producir	27.-Mantener	52.-Complicar	77.-Carecer
3.- Mostrar	28.-Persistir	53.-Condicionar	78.-Clasificar
4.- Reportar	29.-Proponer	54.-Conocer	79.-Colocar
5.-Describir	30.-Recibir	55.-Controlar	80.-Comprender
6.-Detectar	31.-Contener	56.-Crecer	81.-Comprimir
7.-Generar	32.-Decidir	57.-Cumplir	82.-Comunicar
8.-Incluir	33.-Depender	58.-Desconocer	83.-Conllevar
9.-Observar	34.-Discutir	59.-Desencadenar	84.-Conservar
10.-Tratar	35.-Disminuir	60.-Detener	85.-Consistir
11.-Comprometer	36.-Evitar	61.-Disponer	86.-Consultar
12.-Indicar	37.-Favorecer	62.-Evidenciar	87.-Contribuir
13.-Informar	38.-Invadir	63.-Facilitar	88.-Desear
14.-Plantear	39.-Medir	64.-Impedir	89.-Dificultar
15.-Recuperar	40.-Mejorar	65.-Involucrar	90.-Durar

³ Maurice Gross (1985) utiliza el término de *oraciones libres* para referirse a las oraciones predicativas, por lo que en esta investigación no se hará ninguna distinción entre ambos conceptos.

16.-Causar	41.-Pesar	66.-Manifiestar	91.-Establecer
17.-Comparar	42.-Provocar	67.-Ocasionar	92.-Experimentar
18.-Confirmar	43.-Reducir	68.-Ofrecer	93.-Extraer
19.-Cursar	44.-Registrar	69.-Permanecer	94.-Fallecer
20.-Desarrollar	45.-Sobrevivir	70.-Representar	95.-Fluctuar
21.-Descartar	46.-Sufrir	71.-Responder	96.-Inducir
22.-Determinar	47.-Suspender	72.-Solicitar	97.-Influir
23.-Diagnosticar	48.-Tomar	73.-Asegurar	98.-Intentar
24.-Evaluar	49.-Buscar	74.-Avalar	99.-Interrumpir
25.-Evolucionar	50.-Coincidir	75.-Ayudar	100.- Introducir

Tabla 3. Selección de verbos del área de medicina

2.5.2.- Identificación de clases de objeto

Con fin de establecer las clases de objetos, se optó por localizar cada uno de los cien verbos dentro del corpus para así evidenciar los argumentos. Se utilizó NooJ, para localizar una expresión y las concordancias correspondientes.

Concordance for Text C:\Users\Cona\Documents\NooJ\sp\Projects\corpuscompletoydefinitivoginyobst188milp.not

Reset Display: characters before, and 5 after. Display: Matches Outputs

Text	Before	Seq.	After
	cuidados en salud (6). El diagnóstico mayor dificultad para vaciamiento vesical, máximos de 13 mg/dl, no de los casos esta complicación sin presentar modificaciones cervicales ni	requiere	un alto grado de sospecha de cateterismo intermitente. Se solicitó fototerapia. Se da el alta
	Estudio observacional de casos que despegó de la cabeza fetal, de labios menores periuretrales, que poco dolor postparto, generalmente no	requiere	de un enfrentamiento quirúrgico, este hospitalización, con un buen crecimiento
	incidencia de lesiones fetales que utilizar en los partos que con bolsa de aguas prominente, parto fueron inmediatos, y no	requieren	parto instrumental, entre noviembre de una segunda aplicación e intento sutura. No se realizó episiotomía
	cifras tensionales elevadas, hasta 169/107 mmHg, pacientes sintomáticas como las asintomáticas secundarias a fenómenos hipóxico-isquémicos	requieren	de anestesia, con pequeñas y hospitalización del recién nacido no
	La mayoría de estas pacientes capacidad funcional II, que no es simple, rápida y no	requieren	asistencia instrumental. REV CHIL OBSTET
	en malas condiciones generales, hipotérmica, maneja con Nitroglicerina e Hidralazina. sufrió insuficiencia renal aguda que	requiere	de rotura artificial de membranas de maniobras. Se obtuvo un
		requiere	terapia antihipertensiva biasociada con atenolol una laparotomía para efectuar una
		requiere	un control ultrasonográfico seriado, reservando manejo con altas dosis de
		requiere	hospitalización. Durante su control en de mayor destreza quirúrgica. Creemos
		requiere	ventilación mecánica, apoyo con drogas ventilación mecánica por 5 días y
		requiere	diálisis, y otra, dehiscencia de

Query 142/142

Imagen 2: Ejemplo de localización del verbo <requerir> dentro del corpus

A partir de esta revisión de la plataforma NooJ y las consultas realizadas en el Diccionario de la RANM (2012) se logró conformar 41 clases de objeto. Cabe señalar que se consideraron como clases de objetos nombres deverbales que son producto de una transformación. Para categorizar dichas clases se decidió elaborar una tabla que agrupara toda la información relevante de dicha clase, a modo de ejemplificar esta categorización se presenta la tabla con solo cuatro clases de objeto. En el apartado de resultados se exhibirá la tabla compuesta con todas las clases de objeto que se determinaron en este proceso.

Clases de objeto de los verbos de medicina				
<i>N° de clase</i>	<i>Nombre de CL</i>	<i>Abreviatura</i>	<i>Rasgos</i>	<i>Ejemplo</i>
CL1	Humanos	<Hum>	+Anim, +Hum	Mujer, hombre
CL2	Pacientes	<Pac>	+Anim, +Hum, +Pac	Paciente, enferma
CL3	Profesional de la medicina	<Pdem>	+Anim, +Hum, +Pdem	Médico, ginecólogo
CL4	Parientes	<Parent>	+Anim, +Hum, +Parent	Familiares, parientes
...

Tabla 4. Muestra de la tabla “Clases de objeto de los verbos de medicina”

La primera columna indica el número que obtiene la clase de objeto dentro de nuestra clasificación; en la segunda columna se encuentra el nombre que tiene esa clase de objeto, por ejemplo: <humanos>, <pacientes>, etc; en la tercera columna se presentan los rasgos de la clase de objeto y la última columna presenta un ejemplo de palabras que conforman dicha clase objeto. De tal manera, dentro de una estructura argumental determinada, el sustantivo ‘médico’ formaría parte de la CL3 <Profesional de la medicina>.

2.5.3.- Tabla del análisis distribucional de la Léxico-Gramática

El proceso de elaboración de las tablas léxico-gramaticales fue en paralelo a la de la configuración de clases de objetos. El gran objetivo que estas tienen es determinar las

estructuras argumentales para los cien verbos seleccionados, considerando los argumentos, lema y preposición asociada. Por consiguiente, se creó una tabla compuesta por los siguientes elementos:

BLOQUE 1	BLOQUE 2	BLOQUE 3	BLOQUE 4	BLOQUE 5
Argumento 0 (N0)	Lema	Preposición Asociada	Argumento 1 (N1)	Ejemplo
<i>Clases de objeto</i>			<i>Clases de objeto</i>	

Tabla 5. Estructura de la tabla distribucional de LG.

Para ejemplificar de mejor manera, se presenta una tabla ficticia cuya función es demostrar la manera en que se realizó el análisis, puesto que la tabla original analiza los cien verbos seleccionados y toma a las 41 clases de objeto como posibles argumentos, por ende es mucho más extensa. Los elementos que conforman la tabla serán explicados exhaustivamente posteriormente.

N0	Lema	Prep. Asoc 'a'	N1				Ejemplo
			H U M	H A L L	P A C	E V F A V	
+	Evaluar	+	-	-	+	-	<i>El médico evalúa a la paciente</i>
+	Evaluar	+	+	-	-	-	<i>El médico evalúa al feto</i>
+	Evaluar	+	-	+	-	-	<i>El médico evalúa el tumor</i>
+	Evaluar	-	-	-	-	+	<i>El médico evalúa el alta</i>

Tabla 6. Tabla ejemplo del análisis distribucional de LG

La primera columna de la tabla indica el primer argumento que admite el verbo que, generalmente, es el sujeto de la oración. Estos argumentos corresponden a las clases de objeto anteriormente estipuladas, es decir, hay 41 opciones que pueden funcionar como argumentos y estas se presentan en la tabla a partir de la abreviatura asignada. En caso de

que el verbo utilice dicha clase de objeto como N0 se marca con un signo positivo (+) en la intersección correspondiente o en caso contrario se señala con un signo negativo (-). En el ejemplo anterior, se puede determinar que ‘evaluar’ admite, según lo que se pudo verificar en el corpus, como N0 a <Profesional de la medicina>

Después de la columna de N0 se encuentra la correspondiente al ‘Lema’, debajo de esta se escribe el verbo a analizar. Luego, se presenta el bloque de ‘Prep. Asoc’ en la que se escribe la preposición asociada al lema o un signo negativo (-) en caso de que el verbo no lo requiera. A partir del ejemplo, se puede indicar que la preposición asociada de ‘evaluar’ es ‘a’

Al igual que la columna de N0, esta columna se subdivide en 41 apartados correspondientes a las clases de objeto creadas, en vertical se encuentra la abreviatura de la clase de objeto. En la intersección de esta columna y el verbo se señala con un (+) si utiliza tal clase como argumento 1 o con un negativo (-) en caso de que no corresponda a dicha clase de objeto. En el ejemplo que se presentó, ‘evaluar’ tiene 4 posibles N1. Ahora bien, para los verbos que requieren más de dos argumentos, como ‘comparar’, se añadió la columna de ‘Prep. Asoc’ y otra de ‘N2’ con las mismas clases de objeto señaladas anteriormente.

Luego de estas columnas, la tabla distribucional de la Léxico-Gramática finaliza con una en la que se escribe una oración que ejemplifique la estructura argumental del verbo. Esta EA fue determinada a partir de lo indicado en las columnas anteriores, es decir, mediante el signo positivo (+) en las intersecciones correspondientes y a la preposición asociada pertinente.

Entonces, a partir este análisis distribucional es posible establecer las estructuras argumentales de cada uno de los cien verbos seleccionados mediante la combinación de las clases de objeto. Se ejemplificará la manera en que se categorizaron las estructuras argumentales por cada verbo, para ello se seguirá utilizando el verbo evaluar presente en la tabla 6.

Ejemplo de estructura argumental – Evaluar				
<i>N° EA</i>	<i>N0</i>	<i>LEMA</i>	<i>PREP.ASOC</i>	<i>NI</i>
EA1	CL <Profesional de la medicina>	EvaluarC1	a	CL <Paciente>
EA2	CL <Profesional de la medicina>	EvaluarC2	a	CL <Humano>
EA3	CL <Profesional de la medicina>	EvaluarC3	-	<Hallazgo>
EA4	CL <Profesional de la medicina>	EvaluarC4	-	<Evolución médica favorable>

Tabla 7. Ejemplo de estructura argumental de evaluar

De acuerdo a lo planteado en el modelo de tabla distribucional de LG, se puede determinar que ‘evaluar’ posee cuatro estructuras argumentales (EA). Todas las estructuras argumentales de este verbo tienen como N0 a <Profesional de la medicina>, sin embargo su N1 cambia: para EA1 es <Paciente>; EA2 es <Humano>; EA3 es <Hallazgo>; y EA4 la clase de objeto <Evolución médica favorable>

En el apartado de resultados se explicarán de manera más exhaustiva las estructuras argumentales de los verbos de medicina seleccionados.

2.5.4.- Tabla del análisis transformacional de la Léxico-Gramática

Para realizar el análisis transformacional de la Léxico-Gramática, se elaboró una tabla en la que se consideraron todas las estructuras argumentales determinadas a partir de análisis distribucional previamente realizado. Se ejemplificó cada EA con una oración, las cuales fueron incluidas en la tabla. A su vez, se creó un glosario con las transformaciones que fueron seleccionadas para el análisis, el criterio utilizado recae en las posibilidades transformacionales indicadas por G. Gross (2014) y la RAE (2010). Cabe señalar que estas son solo algunas de las transformaciones, ya que otras, como la adjetivación y exclamación, no formaron parte de la lista.

			Oración de ejemplo: <u>El médico operó el tumor</u>
T1	[Neg_]	Negación con algún adverbio	<i>El médico no operó el tumor</i>
T2	[Pasiv_se]	Pasiva con "se"	<i>Se operó el tumor</i>
T3	[Pasiva_per]	Pasiva perifrástica	<i>El tumor fue operado por el médico</i>
T4	[Interr]	Interrogación	<i>¿Quién operó tumor?</i>
T5	[Nom]	Nominalización	<i>La operación del tumor</i>
T6	[Clit_OD]	Clitización del OD	<i>El médico lo operó</i>
T7	[Ser_sub]	Verbo ser más subordinada	<i>El tumor que fue operado por el médico</i>

Tabla 8. Glosario de transformaciones

Por tanto, si una oración, correspondiente a una EA en particular, acepta dicha transformación se marca con un signo (+) en la intersección entre la oración y el indicador de la transformación (T1, T2, T3, T4, T5, T6, T7); se señala con un signo negativo (-) si dicha oración no permite tal transformación. La siguiente tabla ejemplifica como se analizaron dos verbos en la presente investigación.

EVALUAR	T1	T2	T3	T4	T5	T6	T7
<i>El cirujano evalúa los ovarios</i>	+	+	+	+	+	+	+
<i>Los trabajos evalúan las alternativas terapéutica</i>	+	+	+	+	+	+	+
<i>El ginecólogo evalúa los gases arteriales</i>	+	+	+	+	+	+	+
<i>El examen evalúa el hígado</i>	+	+	+	+	+	+	+
<i>El urólogo evalúa a la paciente</i>	+	+	+	+	+	+	+
<i>El Doppler color evalúa el flujo sanguíneo</i>	+	+	+	+	+	+	+
<i>El registro FIBROBID evalúa el resultado</i>	+	+	+	+	+	+	+
<i>El médico evalúa la hipoxia</i>	+	+	+	+	+	+	+
<i>El médico evaluó el caso clínico</i>	+	+	+	+	+	+	+
<i>El control ecográfico evaluará la evolución de vitalidad membrana</i>	+	+	+	+	+	+	+

<i>El médico evalúa la cirugía</i>	+	+	+	+	+	+	+
<i>El especialista evaluó el tratamiento</i>	+	+	+	+	+	+	+
EVOLUCIONAR							
<i>La paciente evoluciona</i>	+	-	-	+	+	-	-
<i>El feto evoluciona</i>	+	-	-	+	+	-	-
<i>La enfermedad evoluciona</i>	+	-	-	+	+	-	-
<i>El trastorno de bipolar evolucionó</i>	+	-	-	+	+	-	-
<i>El síndrome TORCH evoluciona</i>	+	-	-	+	+	-	-
<i>El quiste evolucionó</i>	+	-	-	+	+	-	-

Tabla 9. Tabla transformacional de LG

Respecto a la tabla anterior, se concluye que ‘evaluar’ acepta todas las transformaciones para todas sus oraciones mientras que ‘evolucionar’ solo admite cuatro de ellas: la negación con algún adverbio de negación, la interrogación, la nominalización y el verbo ser más subordinada.

2.5.5 Gramática de detección automática

Para poder detectar computacionalmente las estructuras argumentales de los verbos de medicina se deben crear gramáticas que permitan reconocer dicha estructura de manera automática dentro de un texto de lenguaje natural. Se utilizó el software NooJ, puesto que esa plataforma permite generar gramáticas para analizar los textos, estas pueden ser morfológicas y derivacionales (archivos.nog), productivas (archivos.nom) y sintáctica (archivo.nog). Para cumplir los objetivos de esta investigación, se crearon gramáticas sintácticas que se constituyen mediante sistemas gráficos que incluyen las cadenas léxicas constituidas por una o más unidades (Koza, 2017).

Ahora bien, esta investigación pretende realizar una proyección de la detección automática de estructuras argumentales, por lo que solo se considerará un verbo de una clase (EA) específica. Para ello, se ejemplificará la creación de esta gramática con una de las estructuras argumentales del verbo ‘diagnosticar’ que denominaremos como ‘diagnosticarC1’. De acuerdo al análisis distribucional de la Léxico-Gramática, se determinó que dicho verbo posee cuatro EA:

Estructura argumental - Diagnosticar						
<i>N° EA</i>	<i>N0</i>	<i>LEMA</i>	<i>PREP.</i> <i>ASOC</i>	<i>N1</i>	<i>Prep.</i> <i>Asoc</i>	<i>N2</i>
EA1	CL <Profesional de la medicina>	DiagnositcarC1	-	CL <Hallazo>	a	CL <Paciente>
EA2	CL <Profesional de la medicina>	DiagnositcarC2	-	CL <Enfermedad>	a	CL <Paciente>
EA3	CL <Profesional de la medicina>	Diagnosticar C3	-	CL <Síndrome>	a	CL <Paciente>
EA4	CL <Profesional de la medicina>	Diagnosticar C4	-	CL <Trastorno>	a	CL <Paciente>

Tabla 10. Estructuras argumentales de diagnosticar

Como se puede apreciar ‘diagnosticar’ exige que N0 sea un <profesional de la medicina> y que N2 corresponda a <paciente>, la diferencia radica en que N1, el objeto directo, puede ser un <hallazgo>, <enfermedad>, <síndrome> o <trastorno>.

A su vez, de acuerdo a lo indicado por el análisis transformacional de la Léxico-Gramática, todas las oraciones de las estructuras argumentales de ‘diagnosticar’ admiten las siguientes transformaciones:

T1	Negación	NEG	<i>El médico no diagnosticó cáncer a la paciente</i>
T2	Pasiva con se	VP_SEF	<i>Se diagnosticó síndrome Koppler</i>
T3	Pasiva perifrástica	VP_PERIF	<i>El tumor fue diagnosticado por el médico</i>
T4	Interrogación	INTERR	<i>¿Qué diagnosticó el médico?</i>
T5	Nominalización	NOM	<i>El diagnóstico de la paciente</i>
T6	Clitización de OD	CLIT_OD	<i>El médico lo diagnosticó</i>
T7	Verbo ser más subordinada	SER_SUB	<i>El trastorno que fue diagnosticado por el médico</i>

Tabla 11. Transformaciones que admiten las estructuras argumentales para diagnosticar

Por tanto, para poder generar una gramática de detección de una estructura argumental se debe considerar tanto la oración declarativa como las oraciones generadas según las transformaciones aceptadas. En este caso, se elaboró la gramática para EA1 de ‘diagnosticar’. Cabe señalar que NooJ permite el uso y la creación de diccionarios para el análisis de textos (archivos.dic), en estos se incluirán todas las entradas léxicas que se deseen detectar en el corpus junto con la información gramatical correspondiente (Bonini, 2015). Por tanto, se trabajó con el diccionario de la RAE, insumo que fue brindado por el FONDECYT 1171033, pero se modificó la entrada del diccionario para diagnosticar agregándole la clase correspondiente:

diagnosticar,V+C1+FLX=AMAR

A continuación, se indican los elementos principales que componen la gramática sintáctica para la detección automática de ‘diagnosticarC1’. A estos elementos, dependiendo del tipo de oración y transformación, se les añadían otras unidades léxicas como determinantes, proposiciones, adjetivos, pronombres, etc. Posterior a este glosario, se indicará la gramática utilizada para reconocer los sintagmas nominales cuyo núcleo sea una clase de objeto, algo que se denominó como sintagma nominal de la especialidad. Finalmente, se presentará la gramática de ‘diagnosticarC1’ en su totalidad que luego será explicada exhaustivamente.

<p><SN+Subsubtipo=pdem></p>	<p><i>Es el sintagma nominal que tiene como núcleo una palabra perteneciente a la clase de objeto <Profesional de la medicina></i> <i>Ej: El médico del hospital público de Viña del Mar</i></p>
<p><SN+Subsubtipo=pac></p>	<p><i>Es el sintagma nominal que tiene como núcleo una palabra perteneciente a la clase de objeto <Paciente></i> <i>Ej: El paciente de la tercera edad</i></p>
<p><SN+Subtipo=hall></p>	<p><i>Es el sintagma nominal que tiene como núcleo una palabra perteneciente a la clase de objeto <Hallazgo></i> <i>Ej: El tumor benigno</i></p>
<p><V+C1></p>	<p><i>Verbo ‘diagnosticarC1’ correspondiente a su primera estructura argumental (EA1)</i></p>

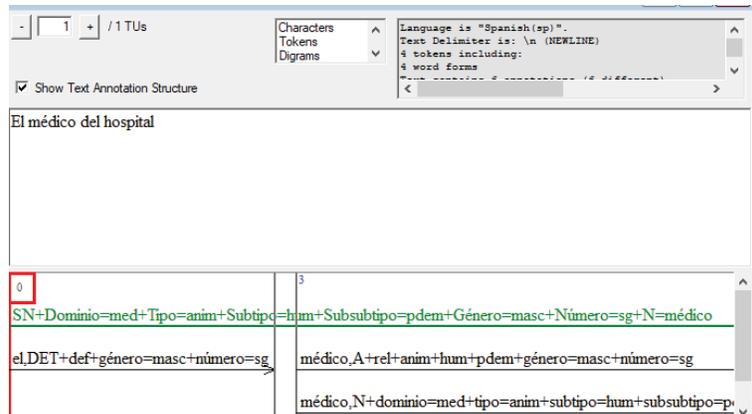


Imagen 4: Detección automática de SN de especialidad

A partir de estos elementos, se elaboró la gramática de reconocimiento para la primera estructura argumental del verbo ‘diagnosticarC1’ denominada como: ‘EAPVC1-DiagnosticarC1’

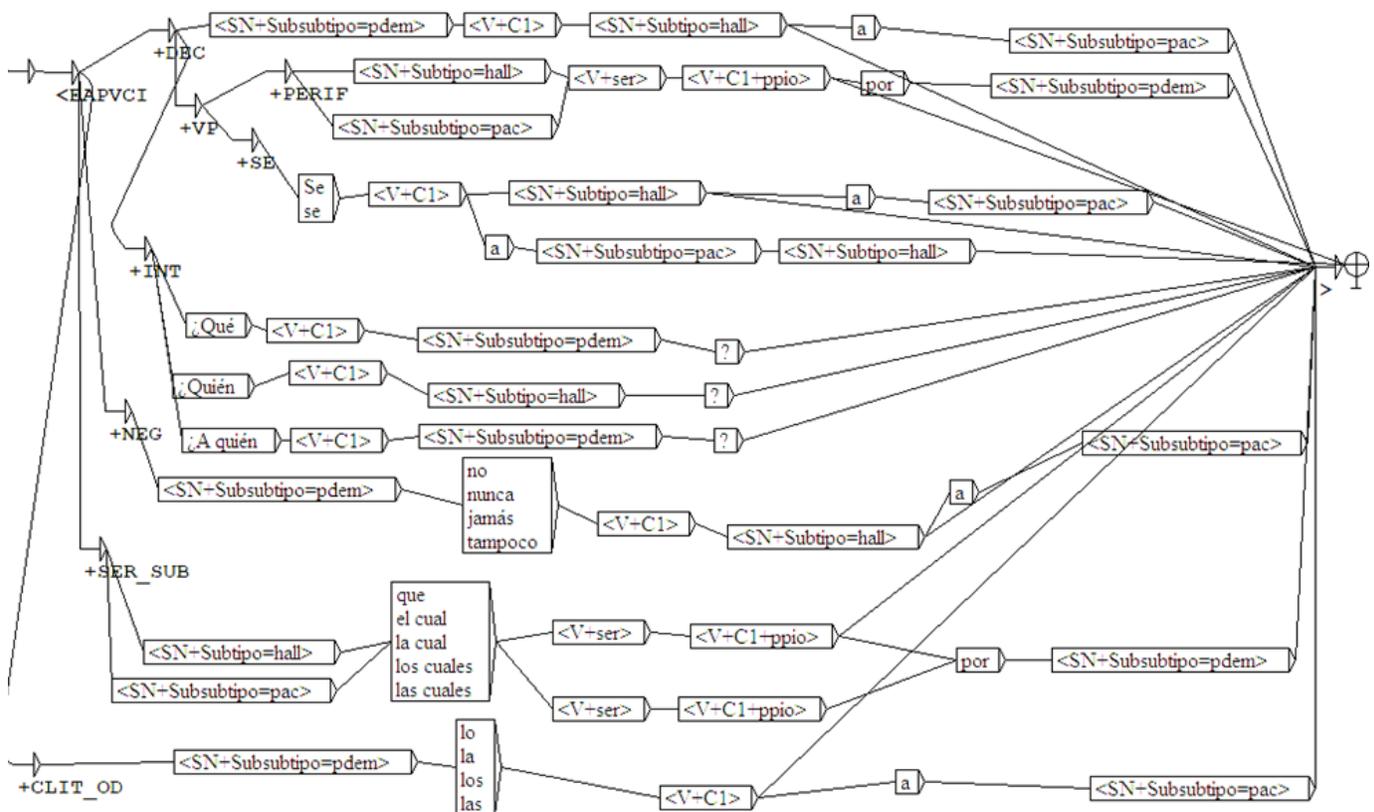


Imagen 5. Gramática de detección para EAPVC1-DiagnosticarC1

En la imagen anterior, se puede observar la estructura argumental para ‘diagnosticarC1’. Al inicio del nodo se encuentra el nombre con el que se etiquetará dicha estructura para su detección dentro del corpus, en este caso será <EAPVC1>. Desde esta etiqueta emergerán las estructuras para la oración declarativa, pasiva perifrástica, pasiva con ‘se’, negación, interrogación, verbo ser más subordinada, y cliticación del objeto directo que serán explicadas a continuación.

- **EAPVC1 + DEC**

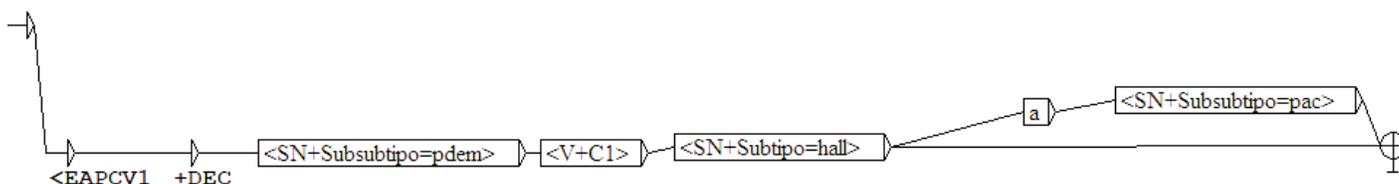


Imagen 6. Gramática de detección para EAPVC1 + DEC

Al inicio, se encuentra la etiqueta de la estructura gramatical de ‘diagnosticarC1’ <EAPVC1> y la de la oración declarativa respectiva +DEC. Los elementos que presentan son los siguientes: se señala que N0 será un sintagma nominal de subsubtipo pdem, es decir, un SN cuyo núcleo sea un nombre perteneciente a la clase de objeto <profesional de la medicina>; el verbo corresponde a ‘diagnosticarC1’, o sea, perteneciente a la primera estructura argumental categorizada; N1, el objeto directo, será un sintagma nominal de subtipo hallazgo, que su núcleo sea un nombre de la clase de objeto <hallazgo>; finalmente, aunque puede no que aparezca en todos los casos, se incluye la preposición ‘a’ para introducir N2 que será un sintagma nominal de subsubtipo paciente, o sea, un núcleo de la clase de objeto <pacientes>. A continuación se presenta un ejemplo de detección con la oración: ‘El medico diagnosticó un tumor’

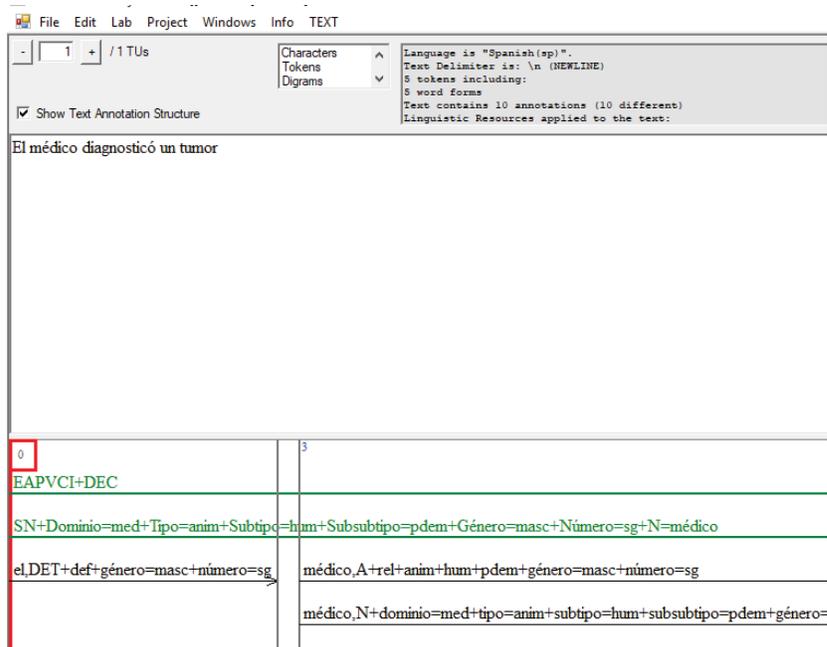


Imagen 7. Ejemplo de detección de EAPVC1+DEC

- **EAPCV1 + DEC +VP +PERIF**

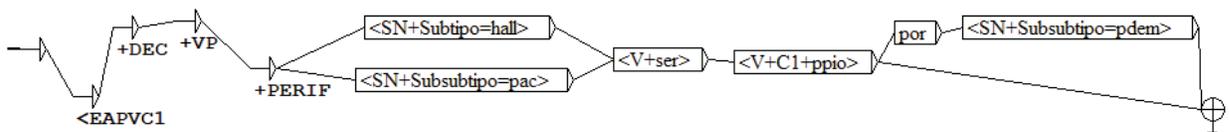


Imagen 8. Gramática de detección para EAPVC1 + DEC + VP + PERIF

Como en todas las gramáticas que se presentarán, el nodo comienza con la etiqueta de la estructura argumental de ‘diagnosticarC1’ <EAPVC1>, luego la etiqueta de oración declarativa +DEC, en voz pasiva +VP y el tipo de pasiva, en este caso, perifrástica +PERIF. Hay dos opciones para N0, esta puede tratarse de un sintagma nominal de subtipo hallazgo o un sintagma nominal de subsubtipo paciente, seguido de estos se presenta el verbo ser (V+ser) y el verbo ‘diagnosticar’ en participio (V+C1+ppio), también es posible incluir la preposición ‘por’ y finalizar con el sintagma nominal de subsubtipo pdem <profesional de

la medicina>. Ejemplo de detección con la oración: ‘El quiste fue diagnosticado por el médico’

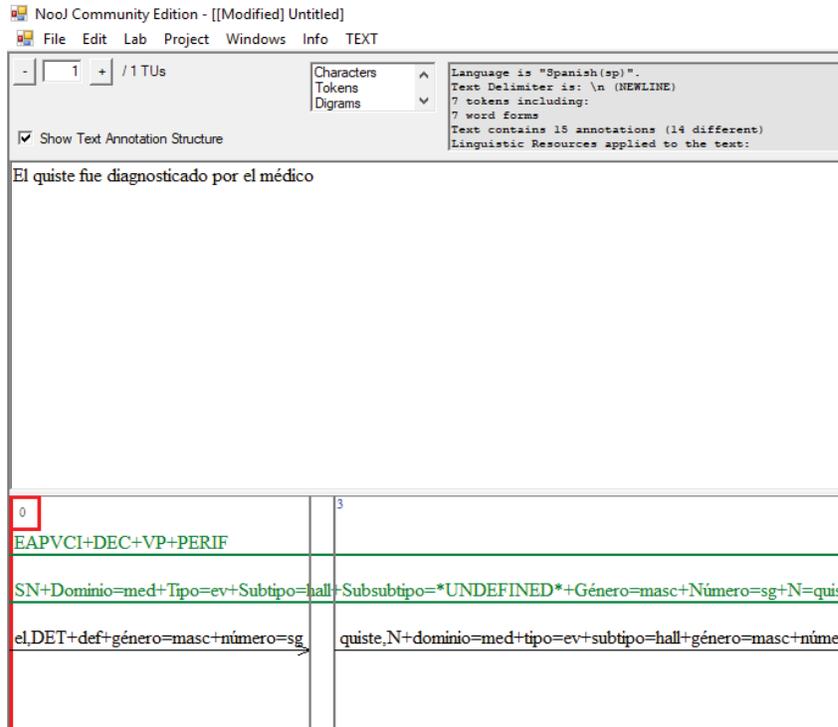


Imagen 9. Ejemplo de detección de EAPVC1 + DEC + VP + PERIF

- EAPVC1 + DEC + VP + SE

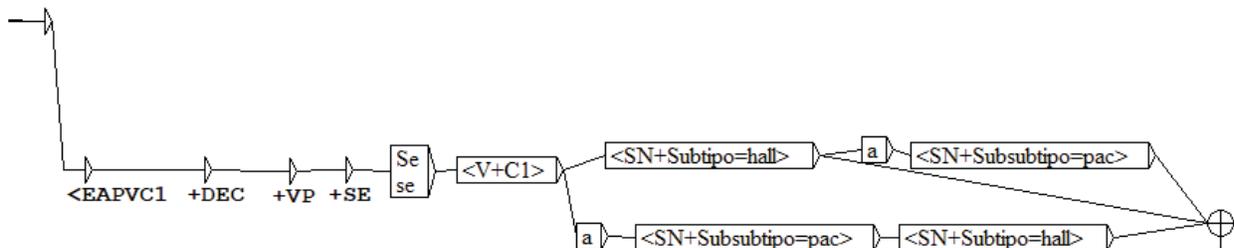


Imagen 10. Gramática de detección para EAPVC1 + DEC + VP + SE

En el inicio están las etiquetas de la estructura argumental, oración declarativa, voz pasiva y el tipo de pasiva, que en esta gramática, es la pasiva con ‘se’ +SE. Por consiguiente, el

primer elemento que se presenta es el pronombre ‘se’ seguido por ‘V+C1’ acompañado por el sintagma nominal de subtipo hallazgo, clase de objeto <hallazgo>, y la posibilidad de incluir objeto indirecto con la preposición ‘a’ y el sintagma nominal de subsubtipo paciente. Otra opción de estructura sintáctica radica en incluir la preposición ‘a’ seguido de un sintagma nominal de subsubtipo paciente para concluir con el sintagma nominal de hallazgos. Ejemplo de detección con la oración: ‘Se diagnosticó un tumor’

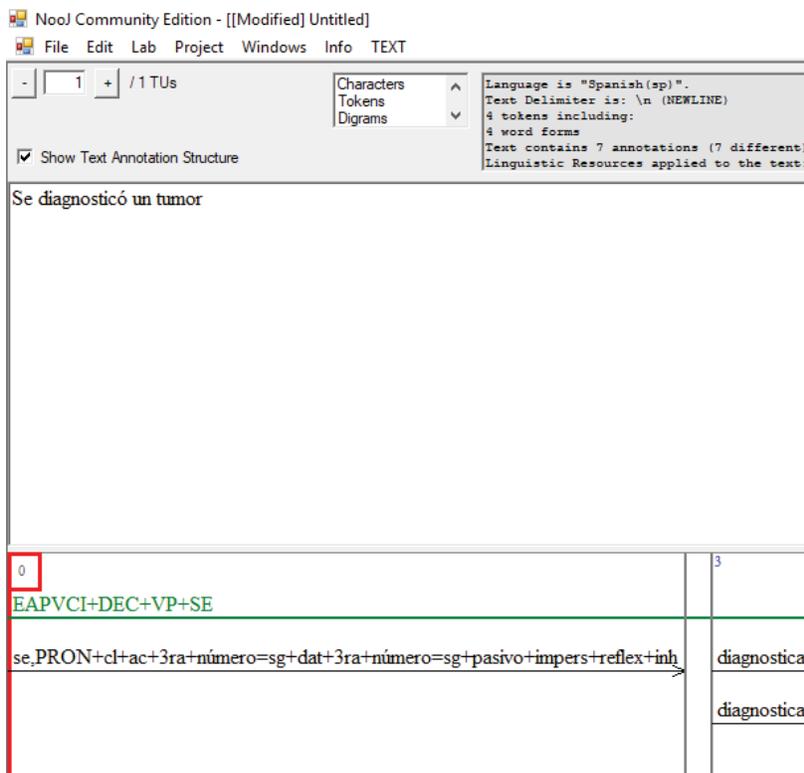


Imagen 11. Ejemplo de detección de EAPVC1 + DEC + VP + SE

- EAPVC1 + INT

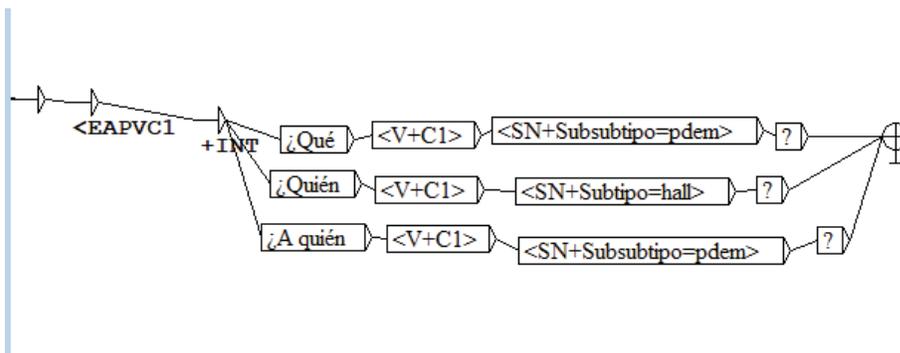


Imagen 12. Gramática de detección para EAPVC1 + INT

La gramática comienza con la etiqueta de la estructura argumental seguido por la etiqueta de la transformación interrogativa +INT, que admite tres posibilidades:

a) comenzar por el pronombre interrogativo ‘qué’, seguido por el verbo ‘diagnosticarC1’ y finalizar con el sintagma nominal de subsubtipo pdem correspondiente a la clase de objeto que designamos como <profesional de la medicina>. Ejemplificaremos con la oración: ‘¿Qué diagnosticó el médico?’

Start	End	Annotation
0	17	EAPVC1+INT
0	5	qué,INTERR
5	17	diagnosticar,V+C1+tiempo=pps+modo=ind+persona=3a+número=sg
17		SN+Dominio=med+Tipo=ani
5	17	diagnosticar,V+tiempo=pps+modo=ind+persona=3a+número=sg
17		el,DET+def+género=masc+n

Imagen 13. Ejemplo de detección de EAPVC1+INT (a)

b) Comenzar con el pronombre ‘quién’, seguido por el verbo ‘diagnosticarC1’ (‘V+C1’) y finalizar con el sintagma nominal de subtipo hallazgo, que en nuestra clasificación de clases de objeto corresponde a un nombre de tipo <hallazgos>. A continuación se muestra la detección para la oración: ‘¿Quién diagnosticó la hernia?’

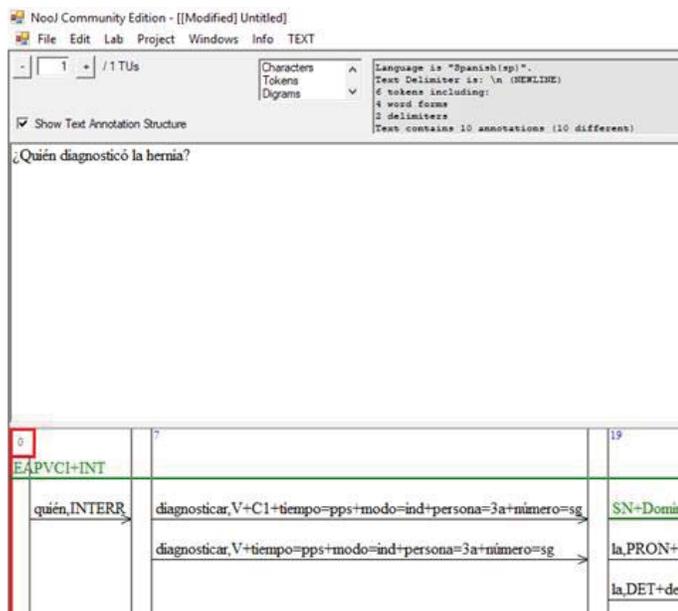


Imagen 14. Ejemplo de detección de EAPVCI+INT (b)

c) Iniciar con la preposición ‘a’ junto con el pronombre interrogativo ‘quién’, seguido del verbo ‘diagnosticarC1’ y terminar con el sintagma nominal de subsubtipo pdem correspondiente a <profesional de la medicina>. Ejemplificaremos la detección con la siguiente oración: ‘¿A quién diagnosticó el cirujano?’

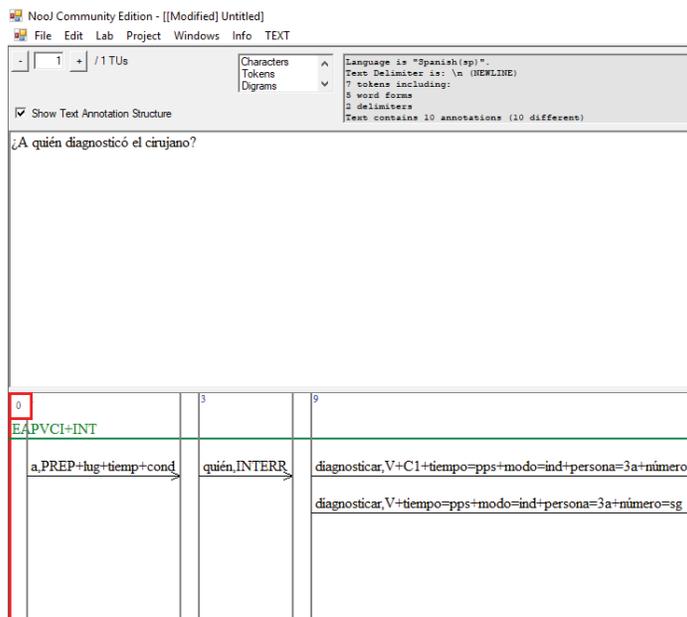


Imagen 15. Ejemplo de detección de EAPVCI+INT (c)

- **EAPVC1 + NEG**

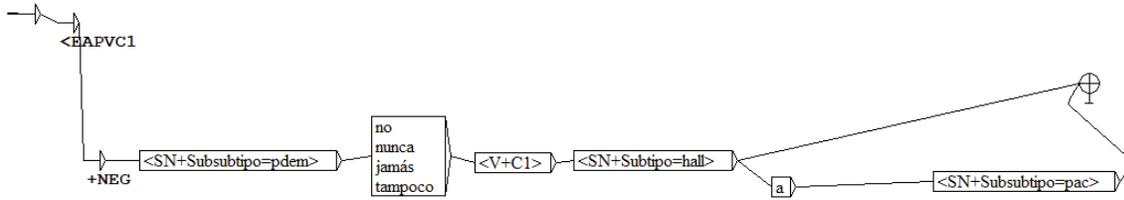


Imagen 16. Gramática de detección para EAPVC1 +NEG

Se inicia con la etiqueta de la estructura argumental y la etiqueta de negación +NEG. El primer elemento que aparece es el N0 correspondiente sintagma nominal subsubtipo pmed, <profesional de la medicina>, seguido por alguno de estos adverbios de negación: ‘no’, ‘nunca’, ‘jamás’, ‘tampoco’. Luego, el verbo ‘diagnosticarC1’ y posteriormente N1, que es el sintagma nominal subtipo hallazgo con la posibilidad de introducir N2 mediante la preposición ‘a’ y el sintagma nominal subsubtipo paciente. Ejemplificaremos con la oración ‘El médico no diagnosticó tumor’

0	3
EAPVC1+NEG	
SN+Dominio=med+Tipo=anim+Subtipo=hum+Subsubtipo=pdem+Género=masc+Número=sg+N=médico	
el,DET+def+género=masc+número=sg	médico,A+rel+anim+hum+pdem+género=masc+número=sg
	médico,N+dominio=med+tipo=anim+subtipo=hum+subsubtipo=pdem+género=masc+núm

Imagen 17. Ejemplo de detección de EAPVC1+NEG

- EAPVC1 + SER_SUB

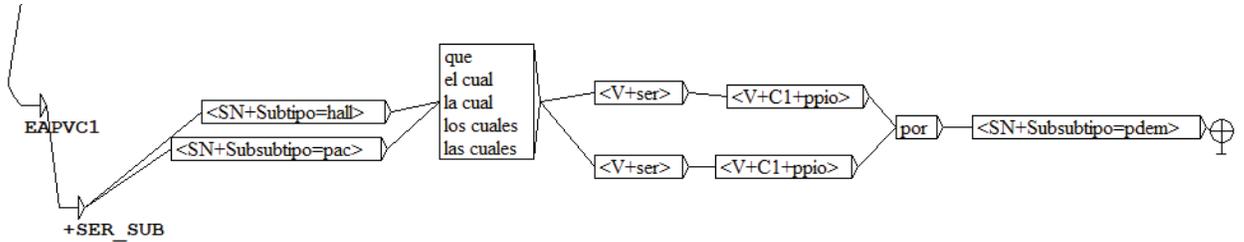


Imagen 18. Gramática de detección para EAPVC1 + SER_SUB

Al igual que en las gramáticas anteriores, la estructura comienza indicando la etiqueta de la estructura argumental para ‘diagnosticarC1’ seguido por la etiqueta creada para identificar la transformación de la oración en la que se añade un verbo ‘ser’ y una subordinada +SER_SUB. Está también presenta dos opciones para el primer argumento, ya que puede ser un sintagma nominal de subtipo hallazgo o un sintagma nominal de subsubtipo paciente, seguido de estos se incluyen los pronombres relativos que introducen una subordinada: ‘que’, ‘el cual’, ‘la cual’, ‘los cuales’; estos van acompañado del verbo ‘ser’ y la forma participial de ‘diagnosticar’, luego la preposición ‘por’ y finaliza con el sintagma de subsubtipo pdem. Ejemplificaremos con la oración ‘El tumor que fue diagnosticado por el médico’.

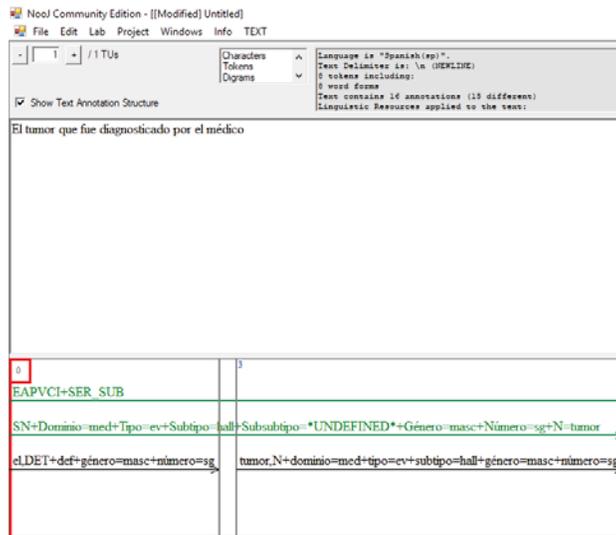


Imagen 19: Ejemplo de detección de EAPVC1+ SER_SUB

- **EAPVC1 + CLIT_OD**

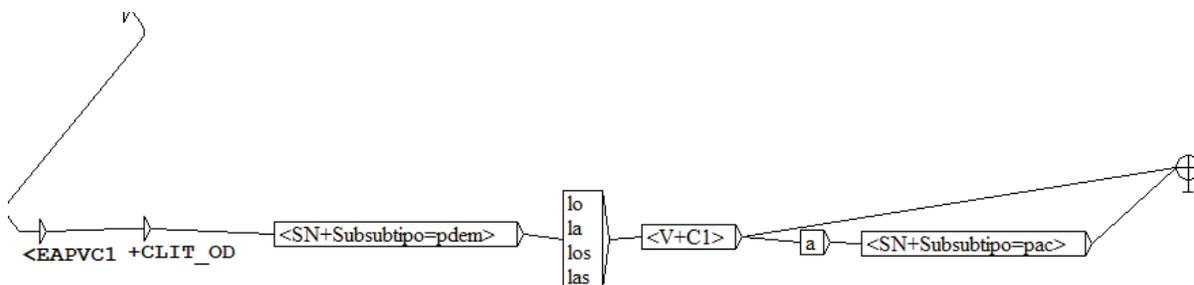


Imagen 20: Gramática de detección para EAPVC1 + CLIT_OD

Se comienza con la etiqueta de la estructura argumental de ‘diagnosticarC1’ y la etiqueta de la transformación correspondiente a la cliticación del objeto directo +CLIT_OD. El primer elemento que se presenta es el sintagma nominal subsubtipo pdem, <profesional de la medicina>, seguido de las palabras que cliticizan el objeto directo (‘lo’, ‘la’, ‘los’, ‘las’) y se presenta el verbo más la clase correspondiente (‘V+C1’). Esta estructura también puede complementarse añadiendo una preposición ‘a’ y el sintagma nominal subsubtipo paciente, equivalente a un núcleo de la clase de objeto creada como <pacientes>. Ejemplificaremos la detección con la oración ‘El médico lo diagnosticó’

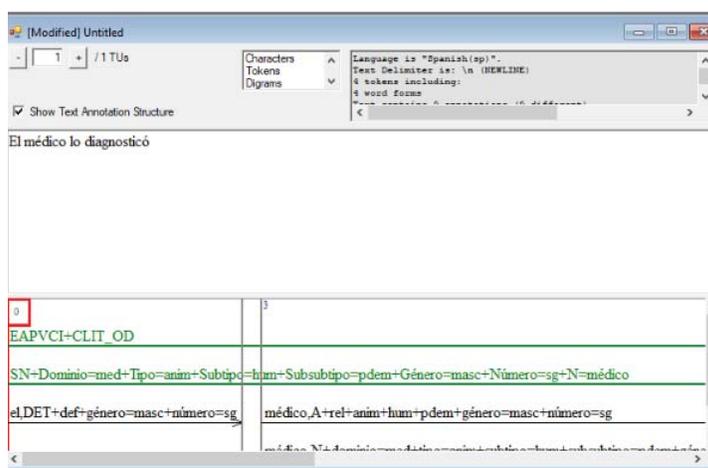


Imagen 21: Ejemplo de detección de EAPVC1+ CLIT_OD

2.5.6 Gramática generativa transformativa

Se elaboró una gramática sintáctica que tiene como finalidad generar todas las posibles oraciones a partir de las transformaciones anteriormente señaladas mediante los elementos que componen la estructura y las palabras que se seleccionaron para ser consideradas en la generación de oraciones. En primer lugar, se describirán los componentes de la estructura para luego presentar esta en su totalidad. Ahora bien, al igual que en la gramática anterior, es necesario señalar que se trata de una gramática simplificada que contiene solo algunos elementos. En trabajos futuros se pretende incluir otras complejidades, como variación de género y número de los argumentos, puesto que esta gramática sirve para ‘la paciente fue diagnosticada’, pero no para ‘el paciente fue diagnosticado’. Un trabajo que consideró todas estas variantes fue el de Silberztein (2016) quien logró 3.000.000 de transformaciones para la oración ‘Joes loves Lea’.

- **ARG0CLI:** Esta etiqueta refiere al argumento 0 (N0) de la estructura y, por ende, a la clase de objeto <Profesional de la medicina>. En esta estructura se deben indicar ciertos sustantivos que conforman dicha clase, en este caso se seleccionaron cuatro SN simples: ‘el médico’, ‘el especialista’, ‘el ginecólogo’ y ‘la matrona’.

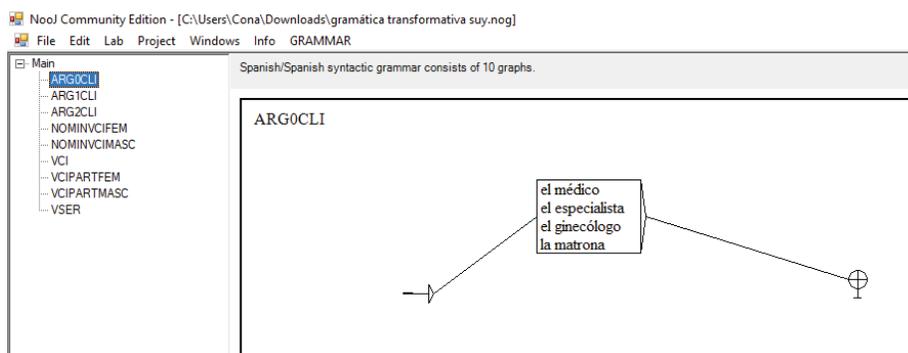


Imagen 22. Gráfico de elementos de ARG0CLI

- **ARG1CLI:** En esta categoría se indican ciertos nombres que conforman el argumento 1 (N1) de la estructura. En este caso, se optó por incluir cuatro SN simples junto con el determinante correspondiente: ‘un tumor’, ‘un quiste’, ‘el tumor’, ‘el quiste’.

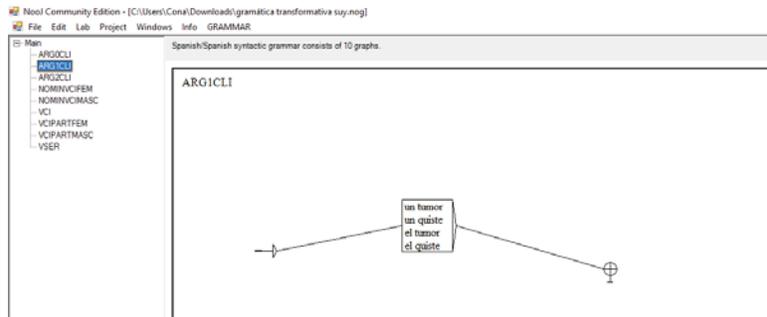


Imagen 23. Gráfico de elementos de ARG1CLI

- **AR2CLI:** Esta etiqueta corresponde al argumento 2 (N2) de la estructura gramatical, los SN simples que se incluyeron para la generación de oraciones son los siguientes: ‘la paciente’, ‘la enferma’ y ‘la paciente embarazada’. Debido a que se realiza una proyección de generación automática, en esta ocasión no se tomó en su variante en masculino.

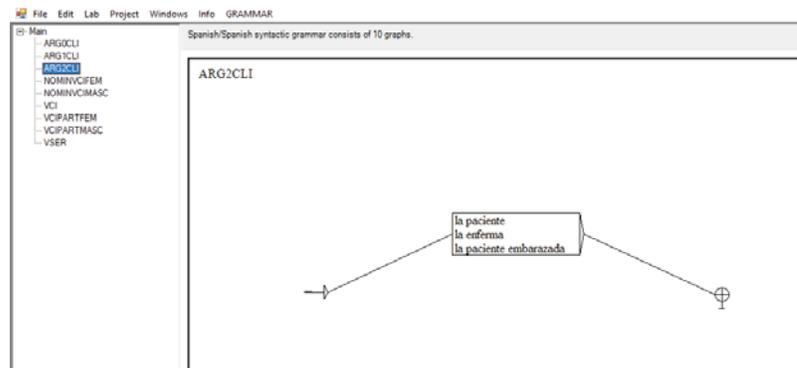


Imagen 24. Gráfico de elementos de ARG2CLI

- **VC1:** A esta etiqueta se le añadieron dos verbos: ‘diagnosticar’ y ‘evaluar’. Se añadió el verbo ‘evaluar’ puesto que a partir del análisis distribucional se evidenció que posee la misma estructura argumental que ‘diagnosticarC1’ y un significado similar a este.

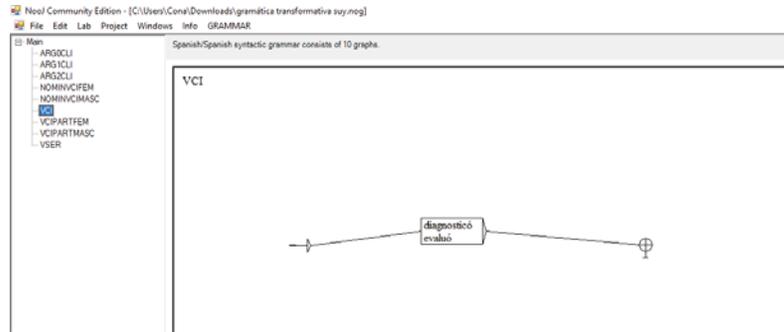


Imagen 25. Gráfico de elementos de VC1

- **VC1PARTFEM:** Esta etiqueta agrupa la forma participial de los verbos anteriormente señalados y en género femenino, por tanto, se indican las palabras ‘diagnosticada’ y ‘evaluada’.

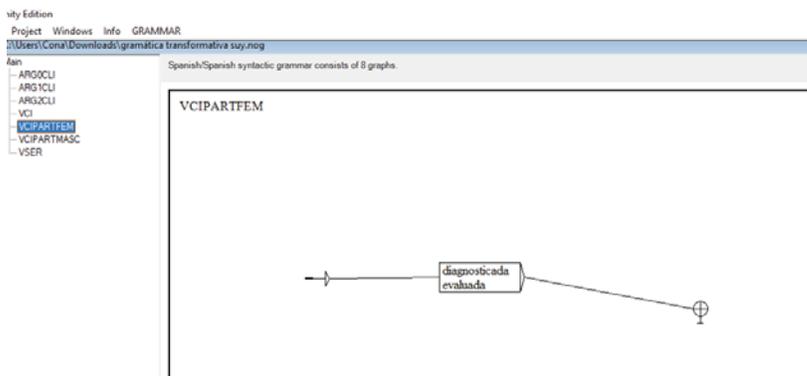


Imagen 26. Gráfico de elementos de VC1

- **VC1PARTMASC:** Esta etiqueta agrupa la forma participial de los verbos anteriormente señalados y en género masculino, por tanto, se indican las palabras ‘diagnosticado’ y ‘evaluado’.

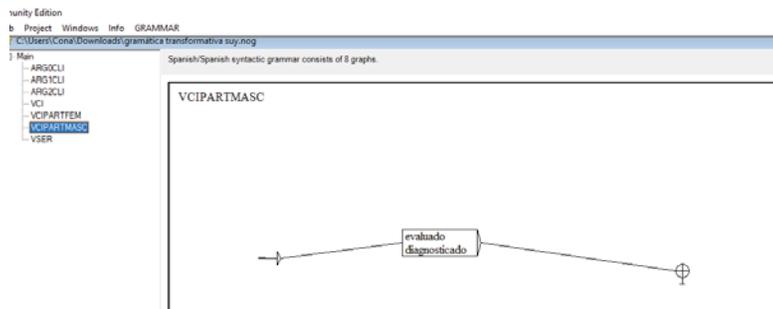


Imagen 27. Gráfico de elementos de VC1

- **VSER:** En este apartado se añaden las palabras referentes a posibles conjugaciones del verbo ser, las seleccionadas fueron: ‘es’, ‘fue’, ‘será’, ‘había sido’ y ‘serán’.

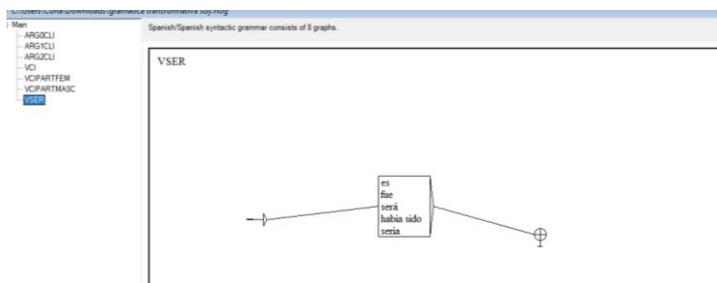


Imagen 28. Gráfico de elementos de VC1

- **NOMINVC1FEM:** En esta etiqueta se incluye la palabra ‘evaluación’, puesto que es la nominalización del verbo ‘evaluar’ que fue incluida para la generación de oraciones. Por tal motivo, se señala que el género es femenino.

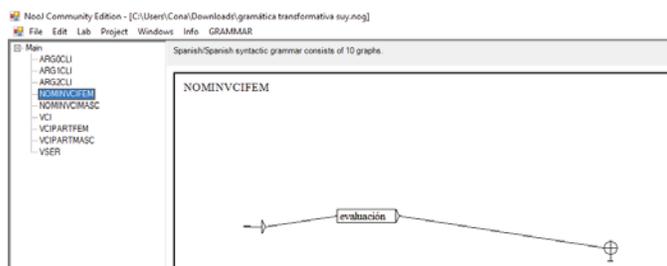


Imagen 29. Gráfico de elementos de VC1

- **NOMINVC1MASC:** En este caso, se señala la nominalización en género masculino correspondiente a ‘diagnostico’ proveniente de la transformación del verbo ‘diagnosticar’ utilizado para esta estructura gramatical.

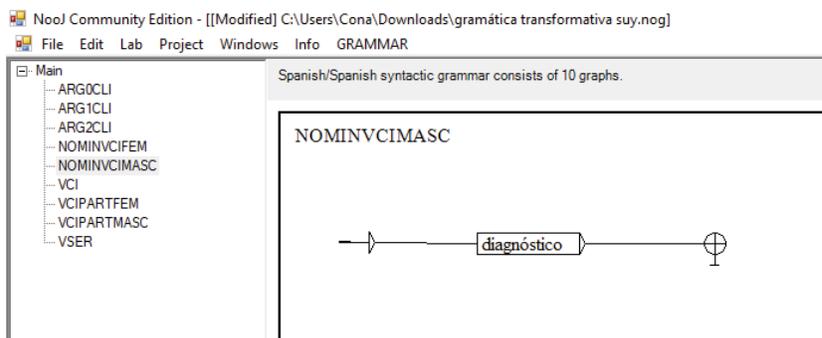


Imagen 30. Gráfico de elementos de VCI

Por consiguiente, a continuación se presenta la gramática generativa para todas las posibilidades oracionales a partir de la combinación de los argumentos y elementos mencionados anteriormente:

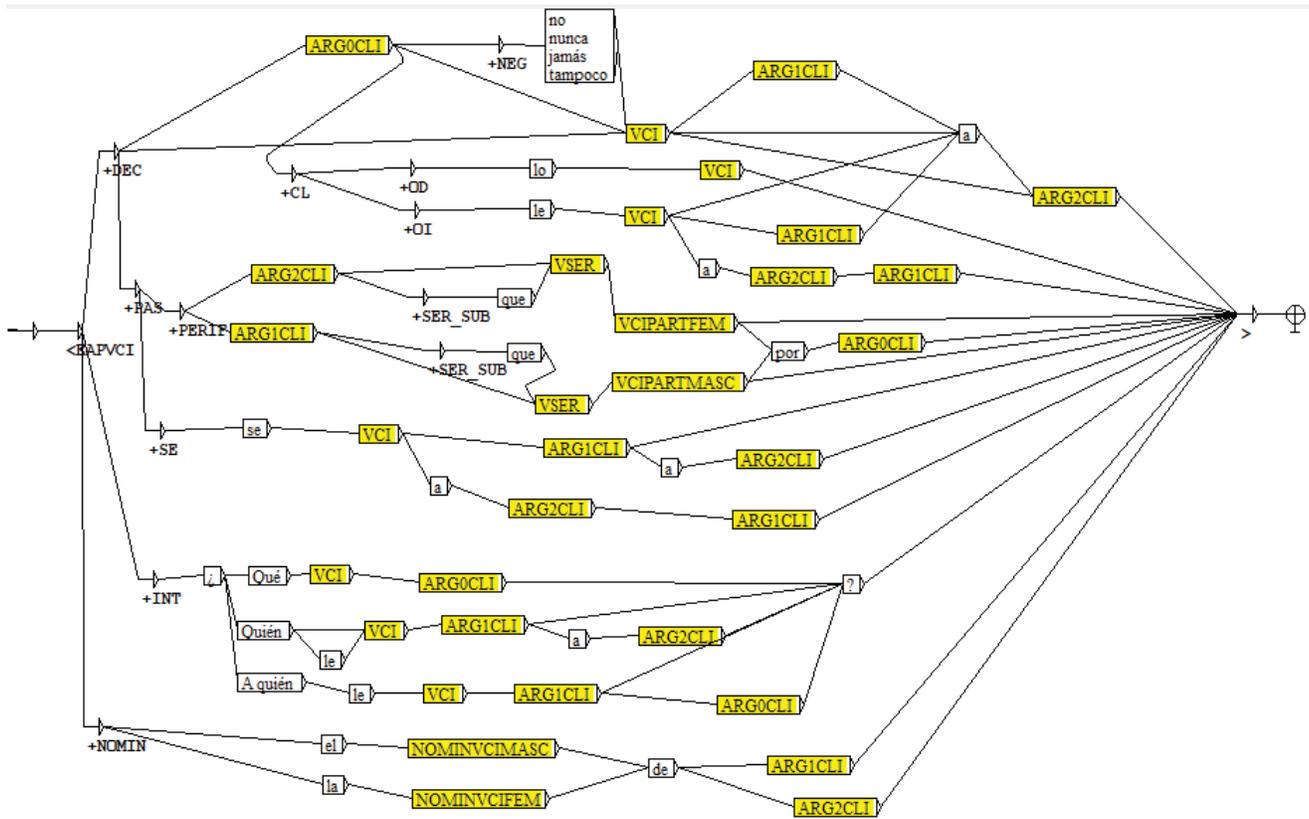


Imagen 31. Gramática transformativa de EAPVCI

Al igual que la gramática señalada anteriormente, en el inicio del nodo se encuentra la etiqueta correspondiente al nombre de la estructura <EAPVCI> seguido de las etiquetas que indican la estructura sintáctica para una oración declarativa (+DEC), pasiva con ‘se’ (+DEC +PAS +SE), pasiva perifrástica (+DEC +PAS +PERIF), verbo ‘ser’ más subordinada (+SER_SUB), interrogativa (+INT), cliticación del objeto directo (+CLIT

+OD), también se añadió la cliticación del objeto indirecto (+CLIT +OI), la nominalización de ‘evaluar’ (+NOMINVC1FEM) y la nominalización de ‘diagnosticar’ (+NOMINVC1MASC)

Ejemplificaremos con una oración generada por cada clasificación.

EAPVC1 +INT	<i>¿ A quién le evaluó el tumor el especialista ?</i>
EAPVC1+ DEC	<i>El médico diagnosticó un quiste</i>
EAPVC1+ DEC+ SER_SUB	<i>la paciente que es diagnosticada por el ginecólogo</i>
EAPVC1+ DEC+ PAS + SE	<i>Se diagnosticó un tumor a la paciente</i>
EAPVC1+DEC+ PAS + PERIF	<i>la paciente embarazada será diagnosticada por el médico</i>
EAPVC1+ DEC+ CL + OD	<i>la matrona lo evaluó</i>
EAPVC1+DEC+ CL + OI	<i>La matrona le evaluó el quiste</i>
EAPVC1+DEC+NEG	<i>El especialista no evalúa los tumores</i>
EAPVC1+ NOMIN	<i>El diagnóstico de la paciente</i>

Tabla 13. Ejemplo de oraciones generadas a partir de la gramática transformativa

3. RESULTADOS

En el presente apartado se evidenciarán los resultados de esta investigación, los cuáles estuvieron abocados en determinar las clases de objetos junto con la estructura argumental que poseen los verbos de medicina y la cantidad de concordancias halladas dentro del corpus mediante la gramática de detección de una EA y las oraciones que fueron generadas a partir de la gramática transformativa.

3.1 Determinación de clases de objeto

Una de las preguntas de investigación aludía a determinar cuáles eran las clases de objeto de los verbos de medicina. El resultado de aquello se encuentra en la siguiente tabla:

Clases de objeto				
<i>N° de clase</i>	<i>Nombre de CL</i>	<i>Abreviatura</i>	<i>Rasgos</i>	<i>Ejemplo</i>
CL1	Humanos	<Hum>	+Anim, +Hum	Mujer, hombre
CL2	Pacientes	<Pac>	+Anim, +Hum, +Pac	Paciente, enferma
CL3	Profesional de la medicina	<Pdem>	+Anim, +Hum, +Pdem	Médico, ginecólogo
CL4	Parientes	<Parent>	+Anim, +Hum, +Parent	Familiares, parientes
CL5	Evento	<Ev>	+Ev	Existencia, presencia
CL6	Proceso fisiológico	<Pf>	+Ev, +Pf	Parto, contracciones
CL7	Proceso médico	<Pm>	+Ev, +Pm	Operación, cirugía
CL8	Estudios	<Estud>	+Ev, +Estud	ecotomografía, reportes
CL9	Tratamientos	<Trat>	+Ev, +Estud	Quimioterapia
CL10	Técnicas	<Téc>	+Ev, +Tec	Técnica de Manchester, técnica loporoscópica
CL11	Accidentes	<Acc>	+Ev, +Acc	Caída

CL12	Estados	<Est>	+Ev, +Est	Embarazo
CL13	Síntomas	<Sin>	+Sin	Dolor
CL14	Signos	<Sig>	+Sig	Fiebre, vómitos
CL15	Hallazgos	<Hall>	+Ev, +Hall	Tumor, hernia
CL16	Enfermedad	<Enf>	+Ev, +Hall	Cáncer, diabetes
CL17	Enfermedad Social	<Enfsoc>	+Ev, +Hall, +Enfsoc	Tabaquismo, alcoholismo
CL18	Síndromes	<Sind>	+Ev, +Hall, +Sind	Síndrome FORCH, síndrome hemo
CL19	Trastornos	<Trast>	+Ev, +Hall, +Trast	Trastorno bipolar, trastorno de ansiedad
CL20	Evolución médica favorable	<Evfav>	+Ev, +Hall, +Ev.Fav	Mejoría.
CL21	Evolución médica no favorable	<Evnofav>	+Ev, +Hall, +NoFav	Empeoramiento, deterioro
CL22	Medida	<Medida>	+Medida	Dosis, etapas
CL23	Medida longitudinal	<Long>	+Medida, +Long	Centímetros, metros
CL24	Medida peso	<Peso>	+Medida, +Peso	Kilogramos, gramos
CL25	Medida tiempo	<Tiempo>	+Medida, Tiempo	Minutos, días
CL26	Instrumentos Quirúrgicos	<Quir>	+Quir, +Instr	Bisturí
CL27	Implantes	<Impl>	+Quir, Impl	Pierna ortopédica,
CL28	Sustancia biológica	<Bio>	+Sust, +Bio	células
CL29	Sustancia molecular	<Mol>	+Sust, +Mol	Estrógeno, carbono
CL30	Fármacos	<Farm>	+Sust, +Farm	Aspirina, remedios

CL31	Institución	<Inst>	+ <i>Inst</i>	Hospital, consultorio
CL32	Complicaciones	<Problem>	+ <i>Problem</i>	Riesgos, problemas
CL33	Resultados	<Result>	+ <i>Result</i>	Resultados, datos
CL34	Diagnósticos	<Diagnost>	+ <i>Diagnost</i>	Diagnósticos
CL35	Elementos epistemológicos	<Teorimed>	+ <i>Teorimed</i>	Términos
CL36	Supuestos médicos	<Supuest>	+ <i>Supuest</i>	Sospecha, Pronósticos
CL37	Zona anatómica	<Zonant>	+ <i>Zonant</i>	Pared vaginal,
CL38	Estructura corporal	<Estcorp>	+ <i>Estcorp</i>	Vagina, ovario
CL39	Sentidos	<Sent>	+ <i>Sent</i>	Olfato, visión
CL40	Disciplinas	<Discp>	+ <i>Discp</i>	Obstetricia, ginecología
CL41	Control	<Controlmed>	+ <i>Controlmed</i>	Seguimiento, cuidados, manejo

Tabla 14. Clases de objeto de los verbos de medicina

A partir del análisis realizado, se concluyó que dentro de los 100 verbos seleccionados se utilizaron 41 clases de objeto.

Además, tal como se puede apreciar, hay clases de objetos que admiten más de una rasgo semántico, el motivo de aquello recae en la rigurosidad de la clasificación propuesta. Explicaremos lo anterior a partir del siguiente ejemplo: la RANM (2012) define hallazgo como “dato, idea o cosa encontrada, hallada o descubierta mediante observación o investigación” (805) por lo que en esta clasificación podrían incluirse tanto tumores, síndromes y/o enfermedades. Sin embargo, se observó que aquello sería inexacto para algunos verbos, como ocurre en el caso de ‘pesar’ en el que la oración (19) es aceptable mientras que la (20) no lo es.

(19) El tumor pesó 2,5 Kg = N0 (+Hall) + V + N1 (+Medida, +Peso)

(20) *El cáncer pesó 2,5 kg ≠ N0 (+Hall) + V + (N1) (+Medida, +Peso)

Por tanto, para evitar dicha imprecisión se optó por generar clases de objetos para <hallazgos> ('tumor', 'quistes', 'hernias, etc), <enfermedad> ('cáncer', 'diabetes', etc), <síndromes> ('síndrome hemagofocítico', 'síndrome de TORCH', etc) y <trastornos> ('trastorno bipolar', 'trastorno de ansiedad', etc).

3.2.- Estructura argumental

A partir del análisis distribucional de la Léxico-Gramática, se pudo concretizar los resultados respecto a las estructuras argumentales de los cien verbos seleccionados, lo que generó un total de 604 EA. Debido a la gran cantidad de datos, en este apartado se ejemplificarán solo con algunos resultados, no obstante, en el anexo 1 se detalla cada verbo con su estructura argumental.

1. Tratar

Dentro del corpus, se detectó que este verbo posee cuatro estructuras argumentales, por ende, cuatro clases.

Estructura argumental - Tratar				
N° EA	N0	LEMA	PREP.ASOC	N1
EA1	CL <Estudios>	TratarC1	de	CL <Hallazgos>
EA2	CL <Profesional de la medicina>	TratarC2	-	CL <Estructura corporal>
EA3	CL <Estudios>	TratarC3	de	CL <Estado>
EA4	CL <Profesional de la medicina>	TratarC4	-	CL <Enf>

Tabla 15. Estructuras argumentales del verbo 'tratar'

Este verbo admitió dos posibilidades como N0, la clase de objeto de <Profesional de la medicina> y la correspondiente a <Estudios>, es en este último caso donde se agrega una

preposición asociada ‘de’. A priori, se pueden distinguir dos significados, pero la determinación de significados excede los alcances de este trabajo.

1. Evolucionar

A partir de la revisión, se determinó que el verbo evolucionar posee seis estructuras argumentales:

Estructura argumental - Evolucionar				
<i>N° EA</i>	<i>N0</i>	<i>LEMA</i>	<i>PREP.ASOC</i>	<i>NI</i>
EA1	CL <Paciente>	EvolucionarC1	-	-
EA2	CL <Humanos>	EvolucionarC2	-	-
EA3	CL <Enfermedad>	EvolucionarC3	-	-
EA4	CL <Trastorno>	EvolucionarC4	-	-
EA5	CL <Síndrome>	EvolucionarC5	-	-
EA6	CL <Hallazgo>	EvolucionarC6	-	-

Tabla 16. Estructuras argumentales del verbo ‘evolucionar’

Se concluye que este verbo requiere de un solo argumento (N0), los cuales pueden ser, por ejemplo: ‘la paciente’ (EA1), ‘el hombre’ (EA2), ‘el cáncer’ (EA3), ‘el trastorno’ (EA4), ‘el síndrome de HORCH’ (EA5) y ‘el tumor’ (EA6). Ahora bien, también se puede señalar que las distintas EA determinan un distinto significado para ‘evolucionar’. Por ejemplo, la oración (21) puede significar una mejoría mientras que en la oración (22) podría adquirir otro significado.

(21) La paciente evolucionó

(22) El humano evolucionó

Si bien, algunas clases de objeto dan como resultado estructuras argumentales similares, se optó por respetar los principios clasificadores de la Léxico-Gramática y las listas de estructura argumentales según las clases de objeto involucradas.

2. Pesar

En el caso de este verbo, se concluyó que posee cinco estructuras argumentales a partir del análisis distribucional realizado:

Estructura argumental - Pesar				
<i>N° EA</i>	<i>N0</i>	<i>LEMA</i>	<i>PREP.ASOC</i>	<i>N1</i>
EA1	CL <Humanos>	PesarC1	-	CL <Peso>
EA2	CL <Estructura corporal>	PesarC2	-	CL <Peso>
EA3	CL <Hallazgo >	PesarC3	-	CL <Peso>
EA4	CL <Profesional de la medicina>	PesarC4	-	CL <Hallazgo >
EA5	CL <Paciente>	PesarC5	-	CL <Peso>
EA6	CL <Profesional de la medicina>	PesarC6	a	CL <Paciente>

Tabla 17. Estructuras argumentales para verbo ‘pesar’

La mayoría de las EA de pesar (EA1, EA2, EA4, EA5) tiene como N1 la clase de objeto <Peso> (*60 kilogramos, 100 gramos, etc*) mientras que EA4 es la única que admite otra clase de objeto. Esto indica que el verbo puede ser transitivo o intransitivo según los argumentos que los acompañan. A su vez, se puede apreciar que en la oración (23) y (24) le dan un mismo significado al verbo pese a que sean estructuras argumentales distintas. Este ejemplo será retomado en el siguiente apartado.

(23) El médico pesó un tumor

(24) El médico peso a la embarazada.

3.3 Análisis transformacional de la Léxico-Gramática

Otra de las preguntas de investigación apuntaba a las posibilidades transformacionales que aceptaban los verbos de medicina según su estructura argumental, para ello, y como fue señalado en metodología, las transformaciones que fueron tomadas en consideración fueron las siguientes:

T1	<i>Negación con adverbios de negación</i>
T2	<i>Pasiva con se</i>
T3	<i>Pasiva perifrástica</i>
T4	<i>Interrogación</i>
T5	<i>Nominalización</i>
T6	<i>Clitización del objeto directo</i>
T7	<i>Verbo ser más subordinada</i>

Tabla 18. Glosario de transformaciones

El requisito para aseverar que una oración determinada admite dicha transformación radica en que el significado de esta se debe mantener. A partir de esto, los resultados fueron los siguientes:

Verbos	T1	T2	T3	T4	T5	T6	T7
<i>N° de oraciones que admiten la transformación</i>	604	519	519	604	486	519	519
<i>N° de oraciones que no admiten la transformación</i>	0	85	85	0	118	85	85

Tabla 19. Cantidad de oraciones que admiten la transformación

En la tabla anterior se puede evidenciar que hay dos transformaciones que todas las oraciones podían admitir, a saber: la negación con algún adverbio (T1) y la posibilidad de ser transformadas en oraciones interrogativas (T4). Ahora bien, con T2, T3, T6 y T7, 519 oraciones pueden aceptar la transformación y 85 la rechazan. El motivo por el cual esas oraciones rechazan las transformaciones señaladas se debe a que el verbo actuaba como intransitivo, generalmente, esto ocurría cuando el verbo requería solo un argumento ('evolucionar', 'persistir', 'sobrevivir', etc.). A su vez, se determinó que ciertos verbos pueden actuar como transitivo o intransitivo dependiendo de las clases de objeto que actuaban como argumentos, estos verbos fueron 'medir', 'cambiar' y 'cumplir'; por tanto, cuando estos verbos actuaban como transitivos dentro de su estructura argumental presentaban un N1 correspondiente al objeto directo, por lo que era posible realizar una

clitización del objeto directo (25.a). En caso contrario, como ocurre con la oración (16. b), el verbo se tornaba intransitivo y no aceptaba la transformación en voz pasiva.

(25) (a) O1: Miguel pesó la manzana → O2: Miguel la pesó

(b) O1: Miguel pesa 85 kilos → O2: *85 kilos fueron pesados por Miguel.

A partir de esto, y de los resultados evidenciados, se pudo generar un patrón en el que si una oración aceptaba la transformación en pasiva perifrástica (T2) tampoco lo haría para la pasiva con ‘se’ (T3), clitización del OD (T6) y el verbo ser más subordinada (T7). En el caso de T5, la nominalización, 486 oraciones admiten la transformación manteniendo el significado, por lo que las 118 la rechazan.

3.4.- Detección automática de una EA

En la presente investigación, se formalizó la estructura argumental de un verbo de la clase 1, ‘diagnosticar’, junto con sus posibles transformaciones, a esta gramática sintáctica se le denominó como ‘EAPVC1-diagnosticarC1’. Si bien en el apartado anterior se hizo una prueba satisfactoria en la que se logró detectar las estructuras en distintas oraciones, los resultados están abocados al reconocimiento de dichas estructuras dentro del corpus, lo que se puede apreciar en la siguiente imagen.

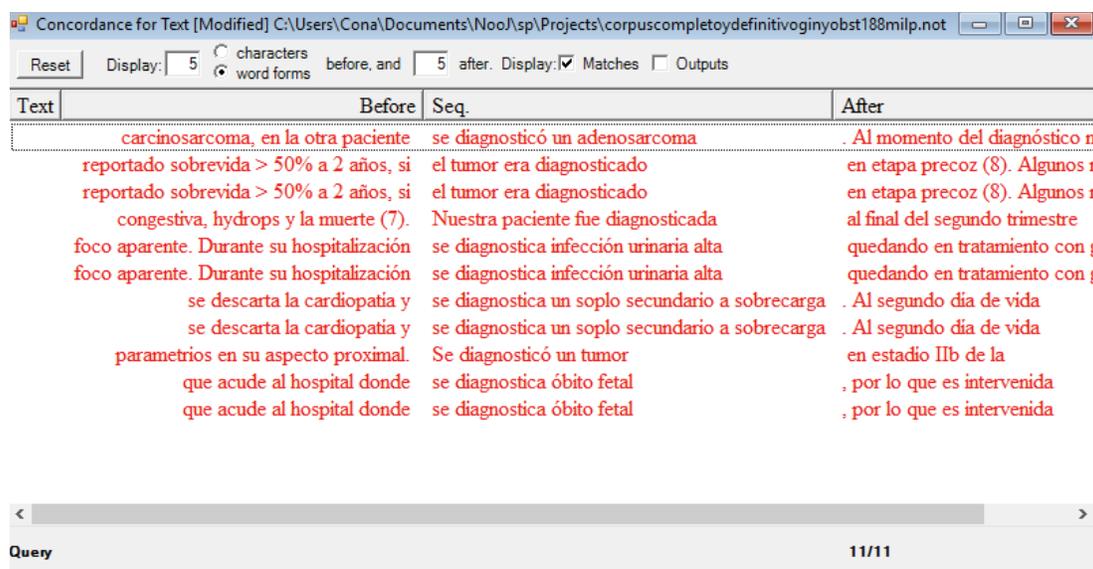


Imagen 32. Resultados de detección automática para EAPVC1 – DiagnosticarC1

El programa NooJ detectó 11 concordancias dentro del corpus para ‘EAPVCI-diagnosticarC1’, todas ellas fueron en transformación de voz pasiva, 3 de ellas pasiva perifrástica y las restantes corresponden a pasiva con ‘se’. Uno de los posibles motivos se debe a la naturaleza del corpus, ya que este está compuesto por textos académicos que tienden a priorizar el uso de voz pasiva por sobre otras estructuras.

3.5. Generación de oraciones transformacionales

Finalmente, los últimos datos que forman parte de los resultados de esta investigación recaen en las posibilidades transformacionales que adquiriría la estructura argumental para ‘EAPVCI – diagnosticarC1’, a la gramática se le agregó el verbo ‘evaluar’ ya que se evidenció comparten sus argumentos y significado en la primera de sus estructuras argumentales (EA1). Por tanto, a partir de la gramática generativa transformativa, que fue descrita en el apartado anterior, NooJ generó un total de 1862 entradas que consideraban las posibles oraciones que se podrían formar:

```

NooJ Community Edition - [C:\Users\Cona\Downloads\gramática transformativa suy_gen.dic]
File Edit Lab Project Windows Info DICTIONARY
1862 entries.
¿ Quién le diagnosticó un tumor a la enferma ?,EAPVCI#+INT
¿ Quién le diagnosticó un tumor a la paciente ?,EAPVCI#+INT
¿ Quién le diagnosticó un tumor a la paciente embarazada ?,EAPVCI#+INT
¿ Qué evaluó la matrona ?,EAPVCI#+INT
¿ Qué evaluó el ginecólogo ?,EAPVCI#+INT
¿ Qué evaluó el especialista ?,EAPVCI#+INT
¿ Qué evaluó el médico ?,EAPVCI#+INT
¿ Qué diagnosticó la matrona ?,EAPVCI#+INT
¿ Qué diagnosticó el ginecólogo ?,EAPVCI#+INT
¿ Qué diagnosticó el especialista ?,EAPVCI#+INT
¿ Qué diagnosticó el médico ?,EAPVCI#+INT
la matrona evaluó la enferma,EAPVCI#+DEC
la matrona evaluó la paciente,EAPVCI#+DEC
la matrona evaluó la paciente embarazada,EAPVCI#+DEC
la matrona evaluó el quiste a la enferma,EAPVCI#+DEC
la matrona evaluó el quiste a la paciente,EAPVCI#+DEC
la matrona evaluó el quiste a la paciente embarazada,EAPVCI#+DEC
la matrona evaluó el tumor a la enferma,EAPVCI#+DEC
la matrona evaluó el tumor a la paciente,EAPVCI#+DEC
la matrona evaluó el tumor a la paciente embarazada,EAPVCI#+DEC
la matrona evaluó un quiste a la enferma,EAPVCI#+DEC
la matrona evaluó un quiste a la paciente,EAPVCI#+DEC
la matrona evaluó un quiste a la paciente embarazada,EAPVCI#+DEC
la matrona evaluó un tumor a la enferma,EAPVCI#+DEC
la matrona evaluó un tumor a la paciente,EAPVCI#+DEC
la matrona evaluó un tumor a la paciente embarazada,EAPVCI#+DEC
la matrona evaluó a la enferma,EAPVCI#+DEC
la matrona evaluó a la paciente,EAPVCI#+DEC
la matrona evaluó a la paciente embarazada,EAPVCI#+DEC
la matrona diagnosticó la enferma,EAPVCI#+DEC
la matrona diagnosticó la paciente,EAPVCI#+DEC
la matrona diagnosticó la paciente embarazada,EAPVCI#+DEC
la matrona diagnosticó el quiste a la enferma,EAPVCI#+DEC
la matrona diagnosticó el quiste a la paciente,EAPVCI#+DEC
la matrona diagnosticó el quiste a la paciente embarazada,EAPVCI#+DEC

```

Imagen 33. Resultados de la gramática transformativa (fragmento)

Tal como se aprecia en la imagen anterior, el software abre una nueva ventana con cada una de las 1862 oraciones. Al costado de la oración se encuentra la clasificación pertinente a la estructura 'EAPVC1' y al tipo de oración correspondiente.

4. CONCLUSIONES

El modelo de la Léxico-Gramática desarrollado por Maurice Gross (1975; 1984) permite realizar una descripción exhaustiva mediante un análisis distribucional de las unidades léxicas a analizar, en este caso la unidad de análisis que consideramos para realizar esta presente investigación fueron los verbos predicativos propios del área de la medicina en el que, a partir de la elaboración de tablas, se pudo realizar un análisis distribucional para determinar las estructuras argumentales. De manera conjunta, se realizó un análisis transformacional que permitió determinar las restricciones que poseían los verbos de acuerdo con sus posibilidades transformativas. A partir de un análisis previamente teórico, se pudo formalizar la EA1 del verbo ‘diagnosticar’ para su detección automática considerando tanto su oración declarativa y las transformaciones que este admitía. Del mismo modo, de acuerdo al análisis teórico se pudo llevar a cabo la elaboración de una gramática con los argumentos correspondiente a una de las estructura argumental de ‘evaluar’ y ‘diagnosticar’, en el que se generó un total 1862 posibles oraciones.

Uno de los pasos fundamentales para alcanzar los resultados refiere a determinar cuáles eran las clases de objeto que estos cien verbos empleaban dentro de su estructura argumental. Mediante una revisión del corpus utilizando la plataforma NooJ se concluyó que hay 41 clases de objetos que actuaban en calidad de argumentos, los cuales tenían distintas posibilidades combinatorias, lo que generó un total de 604 oraciones a partir de las EA descritas. Estos datos apuntan a la primera pregunta de investigación respondiéndola satisfactoriamente.

Respecto al análisis transformacional, se consideraron un total de 7 posibles transformaciones en las que T1 (negación con adverbio) y T2 (interrogación al predicado o sus argumentos) fueron aceptadas por todas las oraciones. Este análisis evidenció, también, que un mismo verbo, de acuerdo a su estructura argumental, podría no aceptar ciertas transformaciones como ocurrió con: medir, pesar y cumplir, puesto que para alguna EA sí permitía su transformación en pasiva (01: ‘El médico midió los cambio’ → 02: ‘Los cambios fueron medidos por el médico’) mientras que en otras la rechazaba (01: ‘El tumor midió 10 cm → 02: *10 cm fueron medidos por el tumor’). También, se pudo evidenciar un patrón en que si el verbo no aceptaba la transformación pasiva tampoco lo haría para la

pasiva con ‘se’, cliticación del objeto directo y el verbo ser más subordinada. Puede que dicha afirmación sea una obviedad, sin embargo, no hay que desconocer que ciertos verbos transitivos no pueden admitir la transformación pasiva. Lo que respecta a la nominalización, estas no eran aceptadas puesto que cambiaba el significado de la oración. Este análisis alude a la segunda pregunta de investigación contestándola significativamente.

Luego del análisis teórico, se realizó un trabajo computacional en el que se elaboró, mediante la plataforma NooJ, una gramática sintáctica de la estructura argumental 1 para ‘diagnosticar’ (‘EAPVC1 – DiagnosticarC1’) con el objetivo de poder detectar automáticamente dicha estructura dentro del corpus. Se registraron un total de 11 concordancias, todas ellas fueron en voz pasiva lo que condice con el género del corpus. A su vez, se creó una gramática transformativa compuesta por las transformaciones que aceptaba EA1 de ‘diagnosticar’ y de ‘evaluar’, el resultado total fue de 1862 entradas.

De acuerdo a lo mencionado anteriormente, es posible dictaminar que se cumplieron los objetivos de esta investigación y que las preguntas fueron respondidas de manera adecuada. Sin embargo, hay diversos elementos que no fueron considerados en esta tesis, pero que pueden ser retomados en futuros proyectos. Uno de estas temáticas que no fue considerada en determinar los significados que adquirirían estos cien verbos a partir de su estructura argumental, puesto que, a priori, el significado podría variar dependiendo de los argumentos que acompañaban al verbo, por lo que sería pertinente realizar un trabajo teórico en el que se cuantifique la cantidad de significados y se determine cada uno de ellos. De la misma manera, esta investigación no consideró los posibles significados de las clases de objeto, o sea, no se determinaron los sustantivos que podrían pertenecer a más de una clase de objeto dependiendo el verbo que lo acompaña. Por tanto, la determinación de las EA de los verbos de medicina se torna relevante y sirve para sentar las bases a otros trabajos investigativos, puesto que pueden ser retomadas y analizadas de diversas maneras.

Para futuros trabajos se podría ampliar la selección de transformaciones, como la adjetivación, y también elaborar gramáticas sintácticas para otras estructuras argumentales de los verbos de medicina. Un ejemplo de dicha transformación sería la siguiente:

O1: El médico operó a la paciente. → O2: La paciente operada por el médico.

Respecto a las limitaciones de esta investigación, se puede señalar que determinar las clases de objeto de los verbos de medicina fue un proceso complejo, puesto que muchos sustantivos eran difíciles de comprender por su especificidad de terminología médica o se ponía en duda la eventividad de estas unidades léxicas, lo que ocasionaba problemas al momento de la clasificación. Para poder subsanar esta problemática, se consultaba recurrentemente a diccionario de RANM (2012), no obstante, en algunas ocasiones se debió contar con la asesoría de expertos del dominio. De la misma manera, se puede considerar que el corpus con el que se trabajó fue un tanto reducido, sin embargo, sigue siendo una muestra adecuada para realizar una proyección de detección y generación automática mediante la elaboración de gramáticas sintácticas. Lo anterior, tiene estrecha relación con el tiempo que fue asignado para realizar la presente investigación, ya que fue escaso si se toma en cuenta que se llevó a cabo un extenso trabajo teórico y, además, un trabajo de aplicación computacional.

REFERENCIAS BIBLIOGRÁFICAS

- Bonino R: Una propuesta para el tratamiento de los enclíticos en NooJ. Infosur Revista 7:31–40, 2015
- Elia, A., Monteleone, M. & Marano, F: From the concept of transformation in Harris and Chomsky to the Lexique-Grammaire of Maurice Gross. En Kasevich, V., Kleiner, Y. & Sériot, P. (Eds.) History of linguistics, 76-82, Amsterdam: John Benjamins, 2011
- Gallardo, S. (2005). Los médicos recomiendan. Un estudio de las notas periodísticas sobre salud. Buenos Aires: Editorial Universitaria de Buenos Aires.
- Gross, M. (1975). *Méthodes en syntaxe*. París: Hermann.
- Gross, M. (1984). “Lexicon-grammar and the syntactic analysis of French”. En *Proceedings of the 10th International Conference on Computational Linguistics and 22nd annual meeting on Association for Computational Linguistics* (pp. 275-282). Association for Computational Linguistics.
- Gross, G. (2014). *Manual de análisis lingüístico. Aproximación sintáctico-semántica al léxico*. Barcelona: Editorial UOC.
- Harris, Zellig S. 1951. *Methods in Structural Linguistics*. The University of Chicago Press, Chicago.
- Harris, Zellig S. 1954. Distributional structure. *Word*, 10:146-162
- Koza, W. Traducción automática del conocimiento médico. Propuesta basada en explicitación y reglas transformacionales. Proyecto FONDECYT Regular 1171033. Comisión Nacional de Investigación Científica y Tecnológica (Conicyt), Chile.
- Koza, W. (2017). La estructura de la enumeración. Análisis, descripción y propuesta de detección automática. *Onomázein*, (35).
- Koza, W. Filippo, D. Cotik, V. Stricker, V. Muñoz, M. Godoy, N... Martínez-Gamboa, R. (2018). Automatic Detection of Negated Findings in Radiological Reports for Spanish

Language: Methodology Based on Lexicon-Grammatical Information Processing. Journal of Digital Imaging doi.org/10.1007/s10278-018-0113-8

Lamiroy, B. (1991). *Léxico y gramática del español. Estructuras verbales de espacio y tiempo*. Barcelona: Editorial Anthropos.

Mounin, G. (1979) Historia de la lingüística (desde los orígenes al siglo xx). *Trad. de Felisa Marcos*. Gredos, Madrid.

Palma, S. (1994). Hacia un enfoque semántico de las expresiones idiomáticas. En Coursera, J., Djian, M., Gaspar, A. (Eds.) *La lingüística francesa. Situación y perspectivas a finales del siglo XX* (pp. 313-321). Zaragoza: Prensas de la Universidad de Zaragoza.

RANM: Diccionario de términos médicos. Editorial Médica Panamericana: Buenos Aires, 2012

Real Academia de la Lengua Española. (2014). Diccionario de la lengua española. Madrid: Espasa.

Sandoval, A. M. (1998). Lingüística computacional. Teide.

Tolone, E. (2009). Les tables du Lexique-Grammaire au format TAL. Ponencia presentada en MajecsTIC 2009, Avignon, Francia

Van Der Stuyft, N., Soto, A. & Solari, L. (2016). Traducción del conocimiento: pautas básicas para los profesionales de la salud pública. *Revista peruana de medicina experimental y salud pública*, 33(3), 1-7. ISSN: 1726-4642