



**PONTIFICIA
UNIVERSIDAD
CATÓLICA DE
VALPARAÍSO**

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



Locuciones verbales del español chileno: una propuesta de descripción e implantación computacional a partir del marco de la Léxico Gramática

Alumna: Natalia Rivas Folch
Profesor Guía: Dr. Walter Adrián Koza

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

11 de diciembre del 2017

ÍNDICE

I. INTRODUCCIÓN.....	3
II. MARCO TEÓRICO.....	5
2.1 Unidad Fraseológica (UF).....	5
2.2 Léxico-gramática (LG).....	16
2.3 Análisis automático de expresiones multpalabras: la plataforma NooJ	19
2.4 Diccionario en uso del español de Chile (DUECh).....	21
III. METODOLOGÍA.....	23
3.1 Pregunta de investigación.....	23
3.2 Objetivos.....	23
3.3 Tipo y alcance de la investigación.....	24
3.4 Tareas desarrolladas.....	24
IV. RESULTADOS	38
4.1 Tablas léxico-gramáticas.....	38
4.2 Reconocimiento automático.....	43
V. CONCLUSIÓN.....	47
VI. BIBLIOGRAFÍA.....	50
ANEXOS.....	52

I. INTRODUCCIÓN

Uno de los problemas más relevantes que han venido abordando los estudios gramaticales, remite a aquellas estructuras conformadas por más de una palabra (“mover el esqueleto”), pero que pueden actuar como una unidad léxica. Dichas construcciones han recibido diversas denominaciones: locuciones, expresiones idiomáticas, expresiones fijas, etc. (Zuluaga, 1975; Corpas, 1996; Penadés, 2014). Sin embargo, en todos los casos, autores como Zuluaga (1975), Corpas (1996) o Penadés (2014), señalan que estas unidades tienen un comportamiento análogo al de las unidades léxicas, por lo que se pueden clasificar según las mismas categorías gramaticales: nominales (“año de la cocoa”), adjetivales (“de un ala”), preposicionales (“detrás de”), adverbiales (“aquí y en la quebrada del ají”) y verbales (“buscarle la quinta pata al gato”). Esto implica que se trata de unidades predicativas, que proyectan estructura argumental (EA) y, por tanto, seleccionan argumentos.

En este trabajo se utilizarán únicamente aquellas locuciones verbales pertenecientes al español de Chile. A este respecto, se busca realizar una descripción exhaustiva del comportamiento morfosintáctico que poseen estos elementos, tomando como fundamentos teóricos y metodológicos los planteamientos de la Léxico-gramática (LG) propuestos por Maurice Gross (1984). Para ello, se trabajará con el *Diccionario de uso del español de Chile* (DUECh), del que se seleccionaron aquellos lemas que se corresponden con locuciones verbales. Cada una de las entradas se categorizará en diversas clases, según su estructuración y en relación con la exigencia de argumentos que poseen. Posteriormente, se procederá a tabular cada locución en dos tipos de tablas LG: en la primera, se dará cuenta del comportamiento que posee la expresión, en lo relativo a los argumentos que proyecta; y en la segunda, se busca evidenciar las transformaciones que pueden presentar este tipo de unidades. Esto debido a que las locuciones, si bien son estables, admiten variación.

La fase anteriormente descrita es necesaria para llevar a cabo el posterior trabajo computacional. Esta labor se realizará con el software de acceso libre, NooJ, desarrollado por Max Silberztein (2005; 2016), el que permite procesar el lenguaje natural a partir de los diversos niveles: morfológicos, sintácticos y semántico. Por tanto, se efectuará la implantación en máquina de reglas gramaticales que posibilitarán la detección automática

de las locuciones verbales del español de Chile. Además de la confección de diccionarios que sean coherentes con aquellas reglas.

El trabajo se organizará en los siguientes apartados: en la sección II se expone el Marco teórico bajo el cual se sustenta este trabajo; por ello, en esta se definen conceptos fundamentales como: (i) Unidad fraseológica, a partir del cual se entregan las diversas características que poseen este tipo de unidades, para luego dar cuenta de las distintas clasificaciones entre las que se encuentran las locuciones verbales; (ii) Léxico-gramática, la cual entrega los fundamentos teóricos y metodológicos de este trabajo; (iii) Lingüística computacional; (iv) NooJ, software empleado para el trabajo de formalización de reglas morfosintácticas que posibilitarán la detección de las locuciones verbales del español de Chile; entre otros.

La sección III presenta la metodología que se llevó a cabo. Para ello, en primer lugar, se declaran las preguntas de investigación, con los respectivos objetivos que guiarán este trabajo. En segundo lugar, se describen: por un lado, los pasos que se efectuaron para confeccionar las tablas LG que dan cuenta tanto de las propiedades morfosintácticas y argumentales de las locuciones, como las posibles transformaciones que admite cada locución. Por otro lado, el procedimiento de formalización de la estructura de las locuciones para la implantación en máquina de reglas gramaticales que permitan realizar la detección automática de estas unidades con NooJ.

La sección IV da cuenta de los resultados obtenidos tanto en la etapa teórico-descriptiva como en la computacional. Para ello, en el primer caso se adjuntarán las tablas en archivos por separado (ver Anexos del 1 al 11). Cada uno de los archivos está separado por Clase y en cada uno se aprecia la tabla LG y la tabla de transformaciones. En el segundo caso, se presentan los resultados del proceso de detección, en términos de precisión, cobertura y medida F.

Finalmente, en la sección V se presentan las conclusiones de este trabajo. Así se indicará si se logró cumplir con los objetivos propuestos, qué proyecciones se desprenden y las limitaciones que se tuvo para la realización del trabajo.

II. MARCO TEÓRICO

En el presente apartado se definen y describen las bases teóricas y herramientas en las que se enmarca y sustenta el presente trabajo investigativo. Para ello, se precisan conceptos relevantes que permitirán comprender el foco y perspectiva que se posee.

2.1. Unidad fraseológica (UF)

En el sistema de la lengua es posible encontrar, además de combinaciones libres de palabras, ciertas estructuras prefabricadas que poseen mayor estabilidad. Estas, estudiadas desde la fraseología, pueden presentarse, tanto como secuencias memorizadas o frases lexicalizadas (Corpas, 1996). En todos los casos, se refiere a estructuras que, según Corpas (1996:18): “(i) Constan de al menos dos palabras ortográficas; (ii) Presentan cierto grado de lexicalización; (iii) Se caracterizan por la alta frecuencia de coaparición en la lengua”. Por tanto, para este autor, aquellos elementos que contemplan estas tres características son denominados por la autora como *unidades fraseológicas* (UF).

Las unidades fraseológicas han sido arduamente investigadas, por ello es posible encontrar una serie de autores que han clasificado estas estructuras a partir de diversos criterios. En este sentido, Corpas (1996) realiza un recorrido por los distintos aportes teóricos de autores como: Casares (1950), Coseriu (1966), Thun (1978), Zuluaga (1980), Haensch *et al* (1982), Carneado (1985) y Tristá (1979-1980,1985, 1988). A partir de estos, refiere una propuesta de clasificación de las UF, la que se sustenta, principalmente, en la combinación del criterio de enunciado, junto con el de fijación. A este respecto, comprende el enunciado como la unidad mínima de comunicación, pudiendo corresponderse con una oración simple o compuesta, o bien una palabra o sintagma. Con esta primera distinción, establece dos grupos de UF: aquellos que no conforman enunciados completos y aquellos que sí lo hacen.

En el primer grupo incluye a aquellas unidades fraseológicas que no forman parte ni de los enunciados ni de los actos de habla, y que, por tanto, requieren combinarse con otros elementos para convertirse en sintagmas. Esta sección la subdivide en dos: esfera I, en la que pueden encontrarse las colocaciones, como “mano derecha”, esto es, UUFF fijas solo en la norma; y, esfera II, en las que ubica a las locuciones, como “mano de monja”, las que

se constituyen unidades fijas del sistema. En el primer ejemplo, correspondiente a una colocación, se puede apreciar que el nombre “mano” puede seleccionar a cualquier otro adjetivo: “izquierda”, “derecha” o “dañada” que indicarán una cualidad de la “mano”, por tanto, no están prefijadas. En cambio, la locución “mano de monja”, posee un significado fijado como una sola unidad, esto es, no posee un significado literal aludiendo a “la mano de una monja”, sino que remite a habilidades que posee una persona para cocinar o realizar otros quehaceres del hogar (DUECh, 2010). Por tanto, las locuciones se encuentran más consolidadas que las colocaciones.

Y, en el segundo grupo, incluye la esfera III, la que refiere a las unidades fraseológicas que constituyen enunciados fraseológicos. Cabe destacar que para efectos de este trabajo no se considerarán las colocaciones ni los enunciados fraseológicos.

1.1 Locución

Considerando las diversas definiciones existentes para comprender las locuciones, cobra relevancia determinar qué se entenderá por locución. En este trabajo, además de lo referido por Corpas (1996), se contemplará la definición de Casares (1950), reformulada por Penadés (2014). De este modo, locución deja de comprenderse como la “combinación estable de dos o más términos, que funciona como elemento oracional y cuyo sentido unitario consabido no se justifica sin más, como una suma del significado normal de los componentes” (Casares, 1950 citado en Penadés, 2014: 20). Esto debido a que Penadés (2014) realiza una serie de precisiones conceptuales que permiten delimitar el aparente límite difuso entre las locuciones y las demás unidades fraseológicas. Para la autora una locución es una “combinación fija de palabras que funciona como elemento de la oración y cuyo significado no se corresponde con la suma de los significados de sus componentes” (25).

De la anterior definición, se pueden desprender una serie de características fundamentales. En primer lugar, las locuciones deben conformarse como una combinación de palabras. De este modo, tanto “chupársele a alguien”¹ (“verse muy amedrentado por algo y no emprender una acción debido a ello”), como “irse cascando” (abandonar un lugar

¹ Ejemplos extraídos del DUECh.

súbitamente) o bien “echarle pelos a la leche” (causar en manera motivada tensión en una situación), se constituyen como locuciones en tanto conforman una combinación de palabras fijas. En segundo lugar, los tres ejemplos mencionados cumplen una función dentro de la oración, que en este caso sería una función verbal. Y, finalmente, se puede observar que la fijación en las locuciones no es solo estructural o sintáctica, sino también semántica, dado que poseen un significado figurado o no composicional. De este modo, la locución “echarle pelos a la leche” no significa que una persona deja caer pelos en un vaso o jarro de leche, sino que alguien genera que aumente la tensión en una situación determinada.

1.1.1 Características de las locuciones

Existen variados autores que investigan, describen y analizan las locuciones (Corpas, 1996; Zuluaga, F., 2004; Orduña, 2011; Penadés, 2012), expresiones fijas (Zuluaga, 1975) o expresiones idiomáticas (Vietri, 2014). No obstante, pese a que se trabajan desde diversas perspectivas o nomenclaturas, destacan una serie de características en las que concuerdan todos los autores:

1.1.1.1 Fijación

Las locuciones se caracterizan por el grado de estabilidad o petrificación que poseen. Por ello, este tipo de unidades no son producidas por los hablantes, sino más bien reproducidas (Zuluaga, 1975). Esto implica que los usuarios de una lengua, o bien, los aprendices de una L2, deben aprenderlas y emplearlas respetando la estructuración interna que poseen. Esta cualidad es denominada por Corpas (1996) como institucionalización.

En relación con este rasgo, Zuluaga (1975) refiere que se puede manifestar de diversas maneras en las locuciones. Por ende, es posible identificar diversos tipos de fijación:

- a) De orden: pese a que en las locuciones existan estructuras cuya función permite reordenar libremente los elementos internos en ella, al estar fijos los componentes, aquellas estructuras pierden esa cualidad. Por ejemplo: “dar cancha, tiro y lado” (superar con una gran ventaja), pese a ser una enumeración, unida por un nexo coordinante, esta no puede ser reordenada como: *dar lado, cancha y tiro.

- b) De categorías gramaticales: remite a que estas estructuras petrificadas no admitirían cambios flexivos de género y número, pese a que sus componentes sí tengan aquellas propiedades. De este modo, la locución “echar a los leones” (motivar a alguien a que realice algo para lo cual no está preparado), no puede ser flexionada como singular *echar al león. Lo anterior, debido a que al flexionarse deja de constituirse como locución y pasa a ser una combinación libre de palabras.
- c) De inventarios de los componentes: conlleva a que, en general, no se pueden adicionar, eliminar o reemplazar elementos en las locuciones. Por ejemplo, la locución “devolver la mano” (compensar un favor), no admite la inserción de elementos como en: *devolver las dos manos; ni cambiar: *devolver el pie.
- d) Transformativa: remite a las transformaciones que se dan en las combinaciones libres de palabras, como la voz pasiva, pero que no se permitirían en las locuciones (Zuluaga, 1980). Un ejemplo de ello es “embolinar la perdiz” (pretender convencer a alguien utilizando frases confusas o distractores) a *fue embolinada la perdiz.

Ahora bien, pese a que se pueden identificar estos tipos de fijación, Penadés (2014) hace evidente que, para cada caso, se pueden observar excepciones. En la fijación de orden se pueden encontrar locuciones como “ser uña y mugre” (tener una relación muy estrecha con alguien), la cual acepta el desorden de la estructura sintáctica sin tener efectos en el significado en el caso: “ser mugre y uña”. En la de categorías gramaticales, se encuentran locuciones como “nacer parado”, la que se puede flexionar en género (*nació parada*) y número (ellas *nacieron* paradas). En esta construcción se mantiene el sentido de “tener muy buena suerte”, declarado en el DUECh. En la fijación de inventarios de componentes, se pueden vislumbrar locuciones como “meter la mano al bolsillo”, cuyo inventario puede variar, generándose estructuras como “meter la mano *en el* bolsillo”, sin poseer mayor incidencia en el significado idiomático. Finalmente, en relación con las transformaciones, es posible estudiar unidades como “hacer mierda” en casos como: Juan *hizo mierda* las flores, el que a su vez puede modificarse obteniendo la estructura: las flores *fueron hechas mierda* por Juan. En ambos casos implica que alguien hace o afirma algo que se contradice inmediatamente después (DUECh, 2010).

Estas precisiones cobran relevancia dado que permiten comprender que la fijación que poseen las locuciones es relativa, pues existen construcciones que poseen carácter de unidad idiomática, no obstante, no cuenta con todas las cualidades antes descritas. A partir de esto, es posible observar que las unidades fraseológicas aquí investigadas, pueden ser gradadas por el especialista. De este modo, si se piensa en la idea de *continuum*, serán locuciones más periféricas o alejadas del centro, aquellas cuya fijación no es absoluta (Penadés, 2014).

1.1.1.2 Idiomática

Remite al grado de especialización semántica que poseen las locuciones y, también, se relaciona con la lexicalización con las que cuentan estas estructuras. Esto implica que el significado que contiene la unidad no se genera gracias a la suma de los componentes, por ello se califican como idiomáticas. Esta característica se relaciona con la fijación, la que, en este caso, se daría por la petrificación semántica, esto es, un significado fijado, que no varía (Corpas, 1996).

Ahora bien, como refiere Zuluaga (1980), las locuciones cuentan con ciertas particularidades que propician la cualidad idiomática que posee. A este respecto, el autor señala como una de ellas el hecho de que existan palabras diacríticas. Este tipo de unidades se caracterizan por aparecer petrificadas en las locuciones, pero que no se emplean en construcciones libres. Un ejemplo de lo anterior se encuentra en la locución verbal de español de Chile: “decir upa” (“se usa para conectar dos hechos de diversas maneras, entre las cuales la más frecuente es la de causa/consecuencia”). En esta se aprecia la palabra “upa”, que no posee un significado propio fuera de la locución e, incluso, su uso se restringe a las locuciones, ya sea como la ejemplificada o, también: “decir upa /chalupa”.

Este mismo autor refiere que la propia estructura interna de la locución provoca que un significado literal sea posible, dado que, si se tomara el de cada elemento de la unidad, la construcción podría carecer completamente de sentido o tener otro completamente diferente de lo que expresa. Esto se puede observar en locuciones como: “dejar la escoba”, cuyo significado idiomático refiere a una situación caótica, confusa o desordenada (DUECh, 2010) (1). Sin embargo, si se considerara su sentido de manera composicional, esto implicaría que una persona dejó la escoba, un instrumento que permite realizar los

quehaceres del hogar, propiciando la limpieza y el orden en algún lugar, como se indica en (2):

(1) Juan *dejó la escoba* al serle infiel a María.

(2) Juan dejó la escoba para que su mamá barra la acera.

La idiomaticidad de las locuciones, en relación con su significado, se puede estudiar, al igual que el caso anterior, a partir de una gradación. Esto se debe a que existen locuciones cuyo significado es más transparente como “echar guata” (“engordar”), cuya alusión a la parte corporal que se abulta o acrecienta, permite deducir el significado. Mientras que existen otras que son más crípticas o cuyo grado de idiomaticidad es mayor, como “emborrachar la perdiz” (pretender convencer a alguien utilizando palabras o frases confusas que distraen).

1.1.1.2.1 Aspectos semánticos

Desde una perspectiva semántica, Corpas (1996) realiza una distinción entre los significados denotativos y connotativos que pueden presentar las locuciones. El denotativo lo subdivide en dos clases: literal y traslaticio o idiomático. En el primer tipo de significado, se encuentran aquellas locuciones que experimentan “cierta delexicalización” (Corpas, 1996: 120) o bien, un mayor grado de composicionalidad. Mientras que, en el segundo, se ubican aquellas unidades cuyo significado es traslaticio o figurado. Estas unidades están institucionalizadas e idiomatizadas, por tanto, el sentido de la locución se desprende a partir de la composición unitaria y no desde cada unidad particular.

Ahora bien, el autor tipifica los diversos significados connotativos y reconoce connotaciones: 1) estilísticas; 2) geográfico-sociales; 3) histórico-culturales; 4) expresivas. De estas, considerando el enfoque de la presente investigación, se profundizará solo la primera tipología, que se puede vincular con la variable diafásica, en la medida en que refiere a un tipo de locución que se selecciona según la situación en la que se encuentre el hablante. Por ello, Corpas (1996) subdivide esta tipología en dos: estilo neutro y estilo bajo. El estilo neutro refiere a un “nivel no marcado” o bien estándar, como “caldo de cultivo ((f) Ambiente propicio’, FEM” (cursiva del autor, 126). En este es posible identificar connotaciones formales, literarias, anticuadas y foráneas. Mientras que el estilo bajo se

emplea en situaciones informales. A tales efectos, admite connotaciones coloquiales, como (“menos que un Candy”) y vulgares o argóticas (“apretada de raja”).

En relación con las connotaciones argóticas, se puede señalar que cobra especial relevancia realizar descripciones profundas de las locuciones que entren en esta categoría, como muchas de las que se trabajarán en esta investigación. Esto debido a que una descripción detallada posibilita la creación de diccionarios que den cuenta tanto de los contextos de uso como de las transformaciones que puede admitir las locuciones, como la voz pasiva en “hacer pebre” en: Juan *hizo pebre* los platos = los platos *fuleron hecho pebre* por Juan. Por lo tanto, realizar una descripción exhaustiva de las propiedades argumentales que poseen las locuciones, así como las transformaciones que admiten estas unidades aporta a los Estudios de español como Lengua Extranjera (ELE). Lo anterior, debido a que se posibilita que los aprendices de una L2 tengan las herramientas necesarias para comprender aquellas expresiones, cuyo significado es demasiado críptico o idiomático (Berri & Boschioli, 2017). Esto se destaca debido a los estudios recientes que se realizan de este tipo de voces en los hispanoamericanos.

1.1.1.3 Variación

Como ya se ha mencionado en los apartados anteriores, la fijación o estabilidad que poseen las locuciones o expresiones idiomáticas es relativa. Por ello, una de las características acuñadas por autores como Zuluaga (1975), Corpas (1996), Penadés (2014), entre otros, es la variación. A este respecto, se concibe esta cualidad “como el fenómeno que se produce cuando una UF, a pesar de estar fijada, se materializa en “formas” distintas sin que dicho cambio conlleve alteraciones en el significado fraseológico” (Ortega & González, 2005: 92). Por lo tanto, y como ya se revisó en el apartado respectivo de la fijación, si bien existen tipos de estabilidad en las locuciones, es posible encontrar algunas excepciones, como aquellas locuciones que admiten cambio de categoría gramatical: el profesor dejó *pegado* a Juan = el profesor dejó *pegada* a María.

No obstante, pese a que la unidad fraseológica puede variar, es importante destacar que este cambio, ya sea morfológico, sintáctico o léxico, no debe alterar el potencial significativo de la locución. En otras palabras, las variantes resultantes del proceso de variación deben continuar siendo locuciones y no constituirse como combinaciones libres de palabras. Por

ello, no todas las UF admiten variaciones. Esto permite referirse al *continuum* entre este tipo de unidades, un *continuum* que se puede gradar a partir de las estructuras que son más estables que otras.

1.1.1.4 Gradación

La gradación es una cualidad transversal a las anteriormente mencionadas, dado que da cuenta de la relativa estabilidad que poseen estas estructuras. De este modo, es posible referir una escala gradual tanto en lo relativo al significado o idiomática que poseen estas unidades, como también respecto a la fijación. En relación con esto, Corpas (1996) propone ciertos criterios, entre los que destaca: 1) grado de fijación sintáctico-estructural y 2) grado de opacidad semántica o idiomática. De este modo, se pueden establecer locuciones, que, al ser más estables o fijas en términos estructurales y semánticos, se pueden ubicar en el centro, si se piensa en términos de centro y periferia, y aquellas que concentren mayor variación, se pueden ubicar en zonas periféricas.

A modo de síntesis, se presenta un cuadro que resume las características de las locuciones abordadas en este apartado, según los autores revisados (Zuluaga, 1975; Corpas, 1980, 1996; Penandés, 2014).

Características	1. Fijación	De orden	Chile es grito y plata *Chile es plata y grito
		De categoría gramatical	Juan hizo perro muerto *María hizo perra muerta
		De inventario de componentes	Juan dejó la zorra en tu casa *Juan dejó unas zorras en tu casa.
		Transformativa	Juan murió pollo *Fue muerto pollo por Juan
	2. Idiomática	Juan pasó por el alambre	
	3. Variación	De orden	Juan es uña y mugre con María Juan es mugre y uña con María
		De categoría gramatical	Juan pilló chanchito a su amigo Juan pilló chanchita a su amiga

		De inventario de componentes	Juan se mandó cambiar ayer Juan se mandó a cambiar ayer
		Transformativa	Juan escribe con la mano y borra con el codo Juan borra con el codo lo que escribe con la mano.
	4. Gradación	Grado de fijación sintáctico estructural	-Juan mandó cortado/a a X +Juan va a todas las paradas
		Grado de opacidad semántica o idiomática	- Juan le pasó la pelota a María + Juan es chanco que no da manteca

Tabla 1. Resumen de características de las locuciones

Una vez abordadas las características de las locuciones, se hace necesario diferenciar estas unidades de aquellas que pertenecen al primer grupo, las colocaciones, siguiendo la clasificación de Corpas (1996). A este respecto, Penadés (2012: 65) refiere una serie de diferencias entre ambos tipos de unidades. En relación con las colocaciones, señala que son “combinaciones de palabras que no pierden el valor semántico que poseen por separado”, esto es, poseen un significado composicional, no idiomático. Por tanto, al combinarse, continúan siendo unidades autónomas. Además, cuentan con un menor grado de fijación, la cual, de existir, es arbitraria y se establece a partir de uso reiterado. Por ello, estas se encuentran más cercanas a los sintagmas libres que a las locuciones, como se apreció al comparar la colocación “mano derecha” con la locución “mano de monja”.

En cambio, según lo señalado anteriormente, las locuciones son unidades, y como tales, su significado no es composicional, debido a la institucionalización e idiomática, antes descrita. Del mismo modo, si se piensa en un *continuum* entre estas unidades y las colocaciones, se puede afirmar que las locuciones poseen un mayor grado de fijación o estabilidad entre sus elementos, los que se traduce tanto en el aspecto semántico, como sintáctico. En relación con lo que se define como límite difuso entre unas unidades y otras, Corpas (1996) refiere una serie de pruebas que se pueden aplicar a las locuciones para diferenciarlas de las demás UUFF: sustitución, eliminación y diferencias transformativas.

Criterios	Colocaciones	Locuciones
Significado	Composicional	Idiomático
Componentes	Unidades autónomas	Unidades dependientes
Grado de fijación	Bajo	Alto

Tabla 2. Cuadro comparativo entre colocación y locución

La primera consiste en cambiar uno de los elementos de la unidad por un sinónimo, hiperónimo o hipónimo. Por ejemplo: *darle al tornillo, en vez de *darle al clavo*. El resultado de la sustitución del complemento indirecto de la locución, por una palabra perteneciente al mismo campo semántico, es una construcción gramatical y posible. Sin embargo, pierde el carácter idiomático de la locución “darle al clavo” (atinar con gran exactitud”, DUECh, 2010).

La segunda conlleva la eliminación de uno de los elementos que conforman la locución. Por ejemplo: frente a la locución “pegarse con una piedra en los dientes” (mostrar gran gratitud), al sustraer uno de los elementos, como *pegarse con piedra en los dientes, al igual que en el caso anterior, se pierde el significado unitario. A este respecto, el autor refiere que tampoco se admite la adición de elementos: *pegarse con una piedra negra en los dientes.

Finalmente, en cuanto a la tercera prueba, la autora refiere que estas unidades fraseológicas no se pueden reordenar sintácticamente, dado que la fijación es tanto sintáctica como semántica. De este modo, la locución “pegarle el palo al gato” (acertar algo que posee consecuencias positivas), no admite la transformación *pegarle al gato el palo. Esta imposibilidad de mover los constituyentes de la locución se fundamenta en que, al ser unidades fijas, no poseen el mismo comportamiento que las combinaciones libres.

Sin embargo, no está de más recordar, como ya se mencionó en los apartados anteriores, que estas pruebas si bien permiten diferenciar las locuciones de otras unidades fraseológicas como las colocaciones, no están exentas de excepciones. Esto debido a que, como ya se ha mencionado, la estabilidad de las estructuras aquí estudiadas es relativa, por lo que admite variación (ver Tabla 1).

1.1.2 Clasificación

Las locuciones, según las diversas definiciones existentes, pueden cumplir una función dentro de la oración. Por ello uno de los criterios de clasificación más comunes se ha definido por el tipo de función que cumple, se encuentra determinada por el núcleo del sintagma. De este modo, se identifican locuciones: nominales, adjetivas, pronominales, adverbiales y verbales (Zuluaga, 1975; Corpas, 1996; Penadés, 2014).

1.1.2.1 Locución verbal

Según Casares (1950, citado en Penadés, 2014) serán locuciones verbales aquellas estructuras compuestas por un verbo que forma un predicado complejo, tras la asimilación de su complemento directo o preposicional. Por ello, este tipo de unidades admiten las modificaciones propias de los verbos: de persona y número, tiempo y modo. Además, al igual que los verbos, estas expresiones pueden designar estados (“andar de maleta”), acciones (“echar la talla”), procesos (“ir a Pelequén”) o propiedades (“llegar con la marraqueta bajo el brazo”).

También, según el tipo de verbo del que se componen, pueden clasificarse en transitivas (“hacer mierda”), intransitivas (“ir al choque”), copulativas o impersonales (“quedar la escoba”). Ahora bien, al analizar estas unidades se debe considerar el comportamiento del verbo como locución y no como palabra libre. A este respecto, se evidencia que locuciones como: “calentar el asiento” (“no hacer nada provechoso en un grupo o en el marco de una actividad”), cuyo núcleo está compuesto por un verbo de naturaleza transitiva, deja de serlo dado que como núcleo de la locución no requiere objeto directo.

Sin embargo, si se emplea esta locución en la oración: *María calentó el asiento* toda la tarde; se obtiene que María es el sujeto agente de la acción “calentar el asiento”, por tanto, si se analiza la locución como una unidad, en este caso, sería de tipo intransitiva, dado que en el uso no presenta objeto directo. Esto producto de que muchas locuciones poseen petrificado algunos de los complementos o preposiciones que requiere (Ruiz, 2001; Penadés, 2014). Esta aclaración resulta relevante para una tarea de reconocimiento automático, como la que se propone en el presente trabajo, pues tanto para la elaboración de las etiquetas como para efectos del análisis, las locuciones se contemplan como

elementos unitarios. Esto implicó contar con un marco teórico que proporcionara herramientas metodológicas para una descripción exhaustiva de las locuciones verbales del español chileno. Por este motivo, se decidió tomar la Léxico-Gramática, propuesta por M. Gross (1984).

2.2 Léxico-gramática (LG)

La Léxico-gramática (*lexique-grammaire*), fundada por Maurice Gross (1984), se constituye como un modelo de descripción gramatical, que pretende construir gramáticas para las diversas lenguas naturales (Mendívil, 1994). Para ello, Gross trabaja con las oraciones simples como unidad mínima de la lengua, dejando fuera palabras y sintagmas. A partir de estas oraciones, busca realizar una descripción de las respectivas propiedades distribucionales y transformacionales que poseen (Gross, 1984).

Por tanto, se pueden destacar tres características fundamentales en este modelo descriptivo del lenguaje: (i) se concibe la sintaxis como un elemento que no puede ser separado del lexicón; (ii) el trabajo con las oraciones simples como unidad mínima de análisis sintáctica y semántica; (iii) puede ser aplicado a cualquier lenguaje como una formalización y descripción metodológica (Elia, Monteleone & Marano, 2011).

La léxico-gramática de una lengua se construye al analizar las oraciones, concebidas, desde este modelo, como elementales. En su propuesta, M. Gross refiere que son tres los componentes necesarios para poder realizar lo referido: (i) oraciones libres; (ii) expresiones idiomáticas o locuciones; (iii) verbos soportes. De estos componentes, se describe la información lingüística, la cual deberá ser procesada de manera automática a partir del análisis de lenguajes naturales (Gross, 1984).

La LG estudia las propiedades: distribucionales y transformacionales que poseen los diversos elementos a sujetos a descripción. En relación con las primeras, se refiere a los requerimientos argumentales que requieren, por ejemplo, los verbos como “amar”, el cual necesita que un nombre (N0), +Humano, *ame* a un nombre (N1) + Hum o -Hum. En cuanto a las segundas, alude a las posibles transformaciones que puede tener una oración simple sin que esta cambie su significado, como por ejemplo: de voz activa a voz pasiva o la cliticización. Toda esta información es representada por medio de tablas, de las cuales existen

diversos modelos, sin embargo, el descrito por M. Gross, se constituye como una matriz binaria (ver Tabla 3) cuyas filas deberán contener a los verbos y las columnas cada oración posible con aquella unidad (Gross, 1984).

N ₀						N ₁						P.A.														
N ₀ = N hum	N ₀ = N pc	N ₀ = N hum	N ₀ = N nr	N ₀ = V n	N ₀ = N plur obl	N ₀ V	N ₀ est V-ant	N ₀ est V pp	N ₀ pc lui V	N ₀ de N ^o pc	N ₁ = N hum	N ₀ V Prép N pc de N ₁	N ₀ lui V Prép N ₁ pc	N ₀ V Prép N pc de N ₀	N ₀ V Prép N ₁ pc	N ₁ = N hum	N ₁ = le fait que F	Ppv	N ₁ = V n	N ₁ = N plur obl	N ₀ V de N ^o pc Prép N ₁	N hum V sur ce point	N ₀ est V-ant Prép N ₁	il V N ₀ O		
+	-	-	-	-	-	dîner	+	-	-	-	de	+	-	-	-	-	-	-	-	-	-	-	-	-	-	
+	+	+	-	-	-	écarter	+	-	+	+	en	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
+	+	+	-	-	-	écoper	+	-	-	-	de	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
+	+	-	-	-	-	s'emparer	-	-	-	-	de	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
+	-	-	-	-	-	s'emporter	+	-	+	-	contre	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
+	+	+	+	+	-	l'emporter	+	-	-	-	sur	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
+	+	+	+	+	-	évoluer	+	+	+	-	vers	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	+	+	-	-	-	explorer	+	-	+	+	en	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
+	-	-	-	-	-	se fendre	+	-	-	-	de	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
+	-	-	-	-	-	fasciner	+	-	-	-	de	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
+	+	+	+	+	-	finir	+	-	-	-	avec	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
+	+	-	-	-	-	en finir	+	-	-	-	avec	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
+	+	-	-	-	-	fondre	-	-	-	-	en	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
+	+	+	-	-	-	fondre	-	-	-	-	sur	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
+	-	-	-	-	-	fuir	+	-	-	-	devant	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Tabla 3. Representación matriz LG (Gross, 1984)

Como se puede observar en la Tabla 3, el autor establece una lista de los verbos y, en este caso, de preposiciones que lo acompañan. Luego, comienza a establecer una serie de argumentos que acepta (+) o rechaza (-) dicha unidad verbal, a partir de las propiedades predicativas que posee y de las propiedades argumentales que proyecta en la sintaxis.

En relación con las expresiones idiomáticas, este autor refiere que la diferencia entre estas unidades y las libres es que las primeras poseen un segmento idiomatizado o fijado. Esto lo especifica en términos de la distribución de los nombres empleados en cada una de las expresiones. Así, en el primer caso, el cambio de un sustantivo hace variar el sentido de la unidad, pasando de ser una locución a una combinación libre de palabras, por ejemplo:

Juan le *pintó el mono* a María vs Juan le pintó el dibujo a María

De esta manera, y como ya se ha referido en el apartado respectivo a las locuciones, al sustituir una de las unidades de la expresión se obtienen dos expresiones diferentes con significados opuestos. En el ejemplo se aprecia que el sentido no composicional de la locución “pintar el mono” se modifica al cambian el objeto directo *pintar el dibujo, la que posee un significado transparente y deja de comportarse como expresión idiomática.

La situación anteriormente descrita no sucede en las oraciones que poseen sustantivos ubicados libremente, dado que al cambiar o sustituir una unidad se puede producir un cambio en el significado. Sin embargo, esto no afecta el sentido general de lo que se deseaba comunicar. Esto lo ejemplifica a partir de:

El niño *comió* la manzana

Mi hermana *comió* el pie

Este fenómeno es justificado por Gross (1984) a partir de las propiedades que poseen los sujetos y objetos en ambas oraciones. Dado que, en ambos casos, poseen un significado similar y cumplen con los requerimientos argumentales del verbo “comer”.

Ahora bien, pese a que realiza esta distinción, este autor afirma que tanto la combinación libre de palabras como las expresiones fijas poseen el mismo comportamiento sintáctico, motivo por el cual declara que es posible emplear el mismo medio de representación de sus propiedades léxico-gramaticales (Gross, 1984; Mendívil, 1994). Esta afirmación ha sido cuestionada por autores como Mendívil (1994: 256), quien sostiene que:

Si la sintaxis de las formas ‘congeladas’ y las libres es la misma, la única diferencia ha de ser de índole semántica. El problema es que si aceptamos que el comportamiento sintáctico es el mismo (lo cual es muy cuestionable) estamos afirmando que la estructura sintáctica de una oración nada o poco tiene que decir en cuanto a su significado.

La observación de este autor se basa en el criterio descriptivo de las expresiones idiomáticas que emplea M. Gross (1984) es lineal. Esto significa que son abordadas sin contemplarlas como unidades petrificadas o congeladas, sino más bien desde un aspecto composicional. Por ello, Mendívil sostiene que se deben analizar a partir de su estructuración real. Esto implica que, si se tiene la locución “estar lona”, no se debe

describir como una expresión que posee un verbo copulativo y un atributo fijo, pues de este modo no se describe la locución como unidad. Lo indicado sería plantear que un verbo como “estar” al convertirse en el núcleo de una locución como “estar lona”, pasa de ser copulativo a intransitivo.

2. 3 Análisis automático de expresiones multipalabras: la plataforma NooJ

Desde el ámbito de procesamiento del lenguaje natural, las expresiones multi-palabra (*multi-word expressions*), como refiere Vietri (2014), se han convertido en un problema cuya solución se concibe como un desafío relevante dentro de esta área de estudio. Estas expresiones se comprenden como una secuencia de, por los menos, dos palabras fijas, que adquieren un significado especial, cuya estructura da cuenta de propiedades léxicas y morfosintácticas particulares. Ahora bien, este tipo de unidades pueden ser analizadas a partir de softwares como NooJ (Vietri, 2014).

NooJ es un programa de acceso libre desarrollado por Max Silberztein (2005; 2016) que permite analizar el lenguaje natural a partir de aspectos morfológicos, sintácticos y semánticos. Este posee una serie de características, entre las que destaca: ser multilingüe, lo cual amplía el espectro de uso de los usuarios, permitiendo analizar una amplia cantidad de lenguas, tipos de corpus, etc. Además, NooJ cuenta con herramientas que permiten analizar textos, como: gramáticas morfológicas y derivacionales (archivos.nof), productivas (archivo.nom) y sintácticas (archivo.nog).

Las primeras incluyen principalmente lo que respecta a los diversos modelos de flexión y derivación que se deseen incluir para realizar los análisis. Las segundas se constituyen como sistemas gráficos que permiten trabajar a partir de cadena de caracteres, una vez fijadas las propiedades formales. Y, las terceras, al igual que en el caso anterior, incluyen los sistemas regulares que permiten abordar las cadenas léxicas constituidas por una o más unidades (Koza, 2017).

Junto con lo anterior, el software posibilita la definición de las distintas propiedades (archivo.nof) con las que se trabajará y la creación de diccionarios (archivo.dic). Las primeras se entenderán como el conjunto de rasgos utilizados para etiquetar las entradas léxicas en los diccionarios, por ejemplo: número sg. pl. y los diccionarios, por su parte,

contemplarán todas aquellas entradas léxicas que se deseen detectar en los corpus, seguidas con la información gramatical respectiva, lo que permitirá efectuar análisis de mayor profundidad (Bonino, 2015). En este sentido, cabe destacar que este software trabaja los datos de manera veloz y eficiente, pues procesa toda la información lingüística al mismo tiempo.

En el marco de esta investigación, que este programa, según Vietri (2014) puede dar cuenta de las propiedades y características que refiere la léxico-gramática. Esto, dado que en sus diccionarios es posible cargar entradas que concentren la información relativa tanto a las propiedades distribucionales, como a las sintácticas. Por este motivo, para efectos prácticos, se utilizará este software, con el propósito de formalizar las gramáticas que se elaboren en torno a las locuciones verbales del español de Chile.

3.1 Frozen expressions

En el manual de usuario de NooJ (Silberztein, 2003), se describen las unidades denominadas como *frozen expressions*. Estas son definidas como unidades multi-palabra discontinuas, por lo tanto, para su detección Silberztein (2003) plantea dos posibilidades para llevar a cabo la detección: (i) emplear las gramáticas locales que NooJ provee para realizar un reconocimiento de los contextos en los que se puede encontrar aquella unidad; (ii) *linkear* los diccionarios con gramáticas que dan cuenta del comportamiento sintáctico de la expresión. Para llevar a cabo la segunda parte de este trabajo, la que guarda relación con la detección de locuciones verbales, estas unidades serán consideradas como *frozen expressions* y, para su reconocimiento se empleará el segundo método. Esto último será detallado en el capítulo correspondiente con la bibliografía.

3.2 Trabajo de Simonetta Vietri

Contemplando los aspectos referidos sobre la LG y NooJ se debe mencionar el trabajo que realiza Simonetta Vietri, puesto que constituyó el puntapié inicial para esta tesis. La autora en su libro *Idiomatic Constructions in Italian: A Lexicon-grammar Approach* (2014) realiza un análisis de las expresiones idiomáticas del italiano a partir del aparato teórico-metodológico que entrega la léxico-gramática. Por tanto, su investigación busca realizar una descripción exhaustiva de estas unidades, comenzando por su clasificación tomando en

consideración tanto los verbos soportes que posean estas unidades y siguiendo con aquellos que poseen contenido léxico.

Para llevar a cabo sus descripciones, Vietri (2014) parte de la introspección y juicios lingüísticos, para luego corroborar estos dos criterios con información proveniente de Internet. Una vez realizada la descripción de las unidades, la autora da cuenta de las posibilidades que entrega la LG para poder realizar una formalización de las expresiones idiomáticas descritas. Para aquello, emplea el software NooJ y crea gramáticas que le permitan detectar las propiedades de las unidades, con las respectivas variaciones que estas admitan.

En este sentido, alguna de las conclusiones a las que llega la autora son: (i) las expresiones idiomáticas del italiano y las no idiomáticas se rigen por las mismas reglas sintácticas, por tanto, ambas admiten transformaciones, como voz pasiva; (ii) en cuanto al criterio de expresiones descomponibles y no descomponibles, determina que, sintácticamente, es un criterio complejo de reproducir; (iii) la representación semántica de las expresiones idiomáticas (o no idiomáticas) se encuentra directamente ligada con su sintaxis; (iv) el uso de la LG constituye buen marco teórico-metodológico para realizar una descripción exhaustiva de unidades como las estudiadas, considerando la posterior formalización de reglas para implantar en máquina.

El presente trabajo toma como base la propuesta de Vietri (2014) para las locuciones verbales del español chileno. Al igual que esta autora, se realiza una descripción morfosintáctica de estas unidades, incluyendo sus posibilidades combinatorias a través de la estructura argumental que proyectan. Asimismo, se considerará, a grandes rasgos, sus posibilidades transformacionales (nominalización, voz pasiva, etcétera). Y, se propone una formalización y una implantación computacional tanto para la generación como la detección automática en textos de lenguaje natural.

2.4 Diccionario de uso del español de Chile (DUECh)

La lista de locuciones verbales analizadas proviene del *Diccionario de usos del español de Chile* (DUECh, 2010). Este posee dos características fundamentales: (i) como el mismo título lo indica, es un diccionario que da cuenta del uso del español hablado en Chile, por

ello los lemas que en él se pueden encontrar se constituyen como expresiones que solo se emplean de manera estable en el país. Para tener certeza de esta propiedad, el equipo de especialistas encargado de la confección de este diccionario filtró las unidades léxicas según su vigencia (solo se incluyen expresiones producidas entre 1950 y el 2010) y según su aparición en sitios como: blogs, discursos radiales, libros, prensa, entre otros. Además, como el foco está en entregar una descripción de los usos de ciertas unidades léxicas en Chile, no posee un carácter normativo. Únicamente se entregan propiedades que le permite discriminar a los destinatarios el contexto de uso de ciertas unidades, como: “dar la cache”, cuyo lema es descrito como una locución verbal de carácter vulgar.

Además, (ii) el DUECh es un diccionario diferencial, esto implica que se constituye como una obra lexicográfica cuyo propósito es “codificar una variedad de lengua a través de un proceso de estandarización (Koza, Alfaro & Martínez, 2015). Por tanto, las unidades declaradas en este él no pertenecen al español general, pues son unidades cuyo uso se restringe al español chileno. Para que esta cualidad sea efectiva, los encargados de la confección de esta obra debieron verificar aquella “diferencialidad dialectal” (DUECh, 2010: 4) empleando variadas plataformas que contienen unidades del léxico general para comparar cada una de las acepciones contempladas en este.

Por tanto, si bien en el DUECh no se indican con precisión los procedimientos metodológicos que se emplearon para realizar la selección de las unidades léxicas (Sáez, 2011) sigue constituyéndose como una de las obras más actualizadas y exhaustivas que da cuenta de los usos del español de Chile. (Koza et al, 2015). Por este motivo, para realizar la descripción de las locuciones verbales del español de Chile, se utilizó este recurso para la extracción de las unidades que será trabajadas a partir de la LG.

III. METODOLOGÍA

En esta sección se dará cuenta de los lineamientos metodológicos que sustentan esta investigación. Para ello, se detallan: las preguntas de investigación, los objetivos propuestos, los alcances del quehacer investigativo y las diversas tareas que se llevaron a cabo.

3.1 Preguntas de investigación

Este trabajo posee dos partes fundamentales: la primera será netamente descriptiva y buscará responder a las siguientes interrogantes:

- ¿Qué propiedades morfosintácticas presentan las locuciones verbales del español chileno?;
- ¿qué estructuras argumentales proyectan las locuciones verbales del español chileno?;
- ¿qué transformaciones admiten las locuciones verbales del español de Chile?

La segunda parte será de aplicación y guardará relación con la formalización de las propiedades morfosintácticas de las locuciones para poder implantarlas en máquina. Por ello, el cuestionamiento que guiará esta sección es:

- ¿Qué propiedades morfosintácticas se deben representar en una gramática computacional para detectar automáticamente locuciones verbales del español chileno?

3.2 Objetivos

Para dar respuesta a las interrogantes antes referidas, en este trabajo se proponen los siguientes propósitos que permitirán guiar el quehacer investigativo:

3.2.1 Objetivos generales

- Describir las propiedades morfosintácticas y la estructura argumental que proyectan las locuciones verbales del español de Chile.
- Detectar locuciones verbales del español de Chile en textos de lenguaje natural.

3.2.2 *Objetivos específicos*

- Identificar las propiedades morfosintácticas de las locuciones verbales del español de chileno.
- Categorizar las locuciones verbales según sus particularidades estructurales.
- Describir las estructuras argumentales de las locuciones verbales del español chileno.
- Describir las transformaciones que admiten las locuciones verbales del español chileno.
- Formalizar las descripciones realizadas con vistas a una implantación en máquina de las locuciones verbales del español chileno.

3.3 Tipo y alcance de la investigación

En relación con el tipo de investigación a la que pertenece este estudio, se puede referir que, como se apreció en el apartado anterior, este trabajo poseerá un enfoque cualitativo. Esto puesto que el principal foco es la descripción de las diversas propiedades que poseen las locuciones verbales del español de Chile. Por ello, el alcance de este trabajo es predominantemente descriptivo.

3.4 Tareas desarrolladas

En el siguiente apartado se realizará una descripción exhaustiva de los procedimientos llevados a cabo para conseguir los objetivos propuestos. Cabe destacar que como esta investigación posee dos partes: una teórico-descriptiva y otra computacional, para otorgar mayor claridad, a continuación, se evidenciarán los pasos que se siguieron en cada caso.

3.4.1 Trabajo descriptivo

Para efectuar el trabajo descriptivo se realizó una serie de tareas que guardan relación con la identificación, categorización y posterior descripción de las locuciones verbales. Para este proceso, un recurso fundamental fue el *Diccionario de uso del español chileno* (DUECh), pues como ya se señaló anteriormente, este texto concentra una serie de entradas léxicas o lemas cuyo uso se restringe al español hablado en Chile. Por tanto, el primer paso que se ejecutó fue la extracción de todas aquellas entradas que pertenecen a la categoría de “locución verbal”.

En este proceso de extracción se recopilaron un total de 842 locuciones verbales. Estas, por fines metodológicos, fueron reunidas en dos grandes grupos: en el primero se concentran expresiones fijas, como: “agachar el moño”, la que significa que alguien cede “ante la voluntad de otra persona” (DUECh), como por ejemplo:

(1) Finalmente, Juan *agachó el moño* y hará lo que dijo María.

Cabe señalar que los elementos incluidos en este grupo admiten emplearse con el adverbio de negación “no” o la preposición “sin”. Por tanto, se pueden encontrar casos como

(2) Juan no *agachó el moño*, por eso María se enojó con él.

En el segundo grupo se encuentran aquellas locuciones cuyas entradas del diccionario se encuentran compuestas por el adverbio de negación “no”, como en el caso: “no ver una”, esto es, el actuar torpe que puede tener una persona (DUECh) y puede emplearse en casos como (3) Juan no *vio una* en el juego de esta noche, por su culpa perdimos. A diferencia de las locuciones del primer grupo, estas no pueden emplearse sin el adverbio de negación, por tanto, expresiones como (4) *Juan *vio una*, resultarán agramaticales.

	Total lemas	Porcentaje
Locuciones Grupo 1	792	95,1 %
Locuciones Grupo 2	41	4,9 %
Locuciones totales (G1 + G2)	833	100 %

Tabla 4. Resumen de locuciones extraídas del *Diccionario de uso del español chileno* (DUECh)

En la Tabla 4 se puede apreciar un resumen del total de locuciones extraídas del DUECh. En el presente trabajo se realizará una descripción de las locuciones del primer grupo. Por ende, tanto la descripción del 4,9% restante, como la implantación en máquina de estas, quedarán para futuras investigaciones.

Posterior al trabajo de recopilación de las locuciones a describir, se realizó una categorización y descripción de estas estructuras según los criterios que plantea la LG (1984). A continuación, se detallarán ambos procedimientos:

3.4.1.1 Categorización

Una vez ejecutada la extracción manual de lemas del diccionario se llevó a cabo el proceso de categorización de las entradas. Para ello, en una primera instancia se tomó en consideración el modo en el que se estructura el lema en el diccionario, esto es, se observaron los elementos constitutivos de la locución. De este modo, expresiones como “dejar la escoba” y “sacarse el pillo” fueron agrupadas en clases diferentes, dado que una requeriría de clítico y la otra no. Sin embargo, al realizar las tablas, se pudo observar que ciertas locuciones poseían requerimientos argumentales diferentes al contemplado inicialmente. Así, aquellas locuciones declaradas en el DUECh con pronombre enclítico, como por ejemplo: “meterse en un queso” podían emplearse sin él, en caso como (5) Juan *metió en un queso* a María. De la misma manera, locuciones cuyo lema aparece sin el pronombre, como “sacar en cara”, su uso exige la presencia de este tipo de elementos: (6) María *te sacó en cara* todo lo que había hecho por ti. Una vez regido el criterio inicial de categorización se obtuvo el siguiente resultado:

Clase	Ejemplo de lema	Ejemplo en uso	Total	%
CI 1: (N0)+V+C1	Echar al agua	María echó al agua a Juan. Lo acusó con su mamá.	404	51,01
CI 2: (N0)+CL+V+C1	Arrancarse con los tarros	Juan se arrancó con los tarros, se adjudicó todo el logro.	283	35,73
CI 3: (N0)+CL+CL+V+C1	Sabérselas por libro	Ella jura que se las sabe por libro.	6	0,75
CI 4: (a)+(N0)+CLdat+V	Crujirle a alguien	A Juan le crujió por fin, tuvo una muy buena idea.	6	0,75
CI 5: (a)+(N0)+CLdat+V+C1	Faltarle a alguien chauchas para el peso.	A María le faltan chauchas para el peso desde que se suicidó su madre.	35	4,37
CI 6: (a)+(N0)+se+CLdat+V+C1	Aconchársele a alguien los meados	A Juan se le aconcharon los meados, por eso no reacciona.	40	4,99
CI 7: (a)+(N0)+se+CLdat+V	Chupársele a alguien	A María se le chupa ir a hablarle	3	0,37
CI 8: (a)+(N0masc)+CLdat+V+C1	Gustarle a un hombre las patitas de chancho	Está claro que a Juan le gustan las patitas de chancho	2	0,25
CI 9: (a)+(N0masc)+se+CLdat+V+C1	Quemársele a un hombre el arroz	María no sabía que a Juan se le quemaba el arroz	7	0,87

Cl 10: (a)+(N0fem)+CL+V+C1	Dejar a una mujer el tren	A María la dejó el tren de tanto esperar a Juan	1	0,12
-----------------------------------	---------------------------	---	---	------

Tabla 5. Resumen de categorización de locuciones pertenecientes al primer grupo.

En la Tabla anterior es posible observar las diversas categorías que emergieron a partir de los diversos requerimientos argumentales que poseen las locuciones. En cada una de las Clases se da cuenta de la organización prototípica que posee la expresión al emplearse. Asimismo, cada categoría contempla una serie de elementos que debe poseer la locución:

Elemento	Significado
(a)	Este elemento representa a la preposición exigida por el pronombre clítico “le”. Por tanto, introduce el complemento indirecto en la locución.
(N0)	Unidad variable en la locución. Su significado se completa en el contexto de uso. En ocasiones la locución exige que el elemento N0 pertenezca a un género determinado, como en las clases 8, 9 y 10. Además, al ser una parte variable, puede prescindirse de su aparición. Por ello, se marca con paréntesis.
le/se	Dan cuenta de los pronombres clíticos determinados o definidos por una locución en particular. En estos casos la locución únicamente admite estas opciones de clíticos.
CL	Indica que en esa parte de la locución se puede insertar un pronombre clítico de cualquier tipo, como dativo o acusativos. Sin embargo, algunas locuciones exigen un tipo particular de pronombre; en esos casos se indica en la Clase como “CLdat”, pues en aquel lugar solo se admiten pronombres dativos.
V	Forma verbal que debe ser conjugada. Se constituye como el núcleo de la locución.
C1	Representa a la sección o parte “invariable” de la locución.

Tabla 6. Descripción de los elementos constitutivos por cada clase.

Posterior al trabajo de categorización, se prosiguió con la confección de las Tablas léxico-gramáticas. Se elaboró una tabla para cada Clase de locución. A continuación, se describirá este procedimiento.

3.4.1.2 Tablas léxico-gramática

3.4.1.2.1 Locuciones verbales del español de Chile

La finalidad de las tablas es describir la estructura argumental de las locuciones verbales que se emplean en Chile. Aquellas tablas, se constituyen como una adaptación de las propuestas por Gross (1984) o Vietri (2014), pues se ciñen a las cualidades que poseen las expresiones descritas. A continuación, se mostrará un ejemplo del procedimiento referido:

N0		PronCl			Verbo	C1	C1						PREP					Ejemplo	
+Anim	-Anim	Se	Dativo	Acusativo			SD	SN	SP	SA	SAd	SV	a	de	con	por	para		N1
(+)	-	-	-	-	hacer	mierda.	-	+	-	-	-	-	-	-	-	-	-	+	Ellos hicieron mierda todo lo que provenía de sus enemigos
-	(+)	-	-	-	hacer	mierda.	-	+	-	-	-	-	-	-	-	-	-	+	La podadora hizo mierda todas las flores
(+)	-	+	-	-	hacer	mierda.	-	+	-	-	-	-	-	-	-	-	-	(+)	Ellos se harán mierda la ropa si siguen peleando
-	(+)	+	-	-	hacer	mierda.	-	+	-	-	-	-	-	-	-	-	-	-	Los platos se hicieron mierda en el suelo
(+)	-	-	+	-	hacer	mierda.	-	+	-	-	-	-	-	-	-	-	-	(+)	Pamela me hará mierda las uñas
-	(+)	-	+	-	hacer	mierda.	-	+	-	-	-	-	-	-	-	-	-	(+)	Las espinas me harán mierda la ropa

Tabla 7. Ejemplo tabla léxico-gramática. Clase 1.

Como se puede apreciar, las tablas LG poseen forma de matriz, en cuyas líneas se define cada una de las locuciones y en sus columnas se precisan las diversas propiedades sintáctico-semánticas. Cuando se realiza el cruce entre ambas secciones de la matriz, se presenta un signo + y uno -, lo que indica si la expresión acepta las cualidades definidas en las columnas. En este caso, se definen ocho propiedades en las columnas, las que tienen como objetivo realizar una descripción exhaustiva y acorde a los fines investigativos propuestos.

En la primera columna se indica el elemento N0, el que da cuenta del primer argumento que exige la expresión para satisfacer los requerimientos del verbo. Esta se subdividió en dos secciones que determinan una cualidad de este elemento, de este modo, puede referir a un objeto +Animado (personas o animales) o -Animado (objetos o elementos abstractos). Esto es útil, pues existen locuciones, como la que se observa en la Tabla 5, que admiten como argumento N0 tanto objetos animados, como es el caso de (7) *Ellos hicieron mierda* todo lo que provenía de sus enemigos; o bien, inanimados: (8) *La podadora hizo mierda* todas las flores. En casos como este, en los que la locución admite más de una cualidad perteneciente al mismo plano, se adiciona otra línea con la información, como se da cuenta en la Tabla.

La segunda columna (PronCl) presenta los pronombres clíticos que puede requerir la locución. En este caso únicamente se realizaron tres subdivisiones: uso de “se”, dativos y acusativos. Esto debido, principalmente, a que se consideró que clasificar cada uno de los clíticos no aporta información valiosa para la posterior formalización computacional de las estructuras.

La tercera (V) y cuarta (C1) columna dan cuenta, por un lado, de la unidad verbal (V) que funciona como núcleo en la locución y, por otro lado, del elemento relativamente invariable

(C1), respectivamente. La información de la quinta columna indica el tipo de sintagma al que pertenece la unidad fija. Como se puede apreciar, se emplean los sintagmas léxicos (nominal, preposicional, adjetival, adverbial y verbal) y, de manera excepcional, un sintagma funcional (determinante). Esto se contempló para poder distinguir a aquellas expresiones que dependen de la presencia o ausencia del determinante para diferenciarlas de las expresiones literales.

Un ejemplo de lo anterior se vislumbra al comparar los elementos que constituyen la locución “hacer gancho” y la oración “hacer el gancho”. Como se observa, mientras que la primera significa, según el DUECh, “Ayudar a alguien a conseguir una relación amorosa con una persona determinada” (2010); la segunda solo indica que alguien o algo construye un “instrumento corvo y por lo común puntiagudo en uno o ambos extremos, que sirve para prender, agarrar o colgar algo” (RAE, en línea). En este sentido, no es lo mismo referir:

(9) Juan *le hizo gancho* a Pedro con María

(10) Juan hizo el gancho que le encargaron los clientes.

En (9) el sintagma nominal al ser parte de la locución no aporta con el significado literal, pues el significado de la expresión no se construye a partir de sus elementos constituyentes de manera aislada, sino que se elabora a partir de un sentido unitario. Mientras que para comprender la oración (10) el significado sí se extrae a partir de sus elementos constituyentes. Por tanto, la presencia o ausencia del determinante, en algunos casos, es fundamental para desambiguar ciertas expresiones.

Las columnas seis (PREP) y siete (N1) dan cuenta de si la locución requiere alguna preposición para introducir un segundo argumento, como la locución “picarle a alguien las manos”, la cual se emplea siempre con la preposición “por”, como: (11) A mi amiga *le pican las manos* por ir a la playa. O bien, de si la expresión requiere un argumento N1, pero sin la necesidad de emplear una preposición que lo introduzca, como es el caso (12) La podadora *hizo mierda* todas las flores. En este caso la locución requiere un segundo argumento que indica aquello que se hizo mierda. Las preposiciones seleccionadas responden a la evidencia extraída durante el proceso de elaboración de la tabla. Finalmente,

la última columna da cuenta de un ejemplo en uso de la locución con sus respectivos requerimientos argumentales.

A partir de los aspectos antes descritos, cada Tabla quedó constituida por las siguientes cantidades de entradas:

Clase	Total locuciones	Total Entradas
Cl 1: (N0)+V+C1	404	798
Cl 2: (N0)+CL+V+C2	283	459
Cl 3: (N0)+CL+CL+V+C3	6	12
Cl 4: (a)+(N0)+CLdat+V	6	6
Cl 5: (a)+(N0)+CLdat+V+C5	35	36
Cl 6: (a)+(N0)+se+CLdat+V+C6	40	40
Cl 7: (a)+(N0)+se+CLdat+V	3	4
Cl 8: (a)+(N0masc)+CLdat+V+C8	2	2
Cl 9: (a)+(N0masc)+se+CLdat+V+C9	7	7
Cl 10: (a)+(N0fem)+CL+V+C10	1	2

Tabla 8. Comparación cantidad de lemas por clase *versus* cantidad de entradas en

Tabla léxico-gramática.

Como se puede apreciar, las primeras Clases poseen una gran diferencia entre la cantidad de lemas y las entradas en el diccionario. Esto porque, si se piensan las clases en un *continuum*, donde la 1 es la más simple y la 10 la más compleja, se puede vislumbrar que a medida que se van complejizando las clases, los usos o las propiedades que admiten para su empleo se restringen, lo cual se traduce en una reducción de la cantidad de entradas en las tablas.

3.4.1.2.2 Tablas transformacionales

Una vez construidas las tablas correspondientes con las posibles estructuras argumentales, se llevó a cabo la tabla de transformaciones. Lo anterior debido a que, como indica la literatura, las locuciones son estructuras relativamente fijas, pues en algunos casos admiten ciertas modificaciones estructurales que no afectan el contenido semántico de la expresión (Harris, 1970), por ello se presentan con un signo de igualdad (=). Para ello, en primer lugar, se debió definir una lista de transformaciones (no exhaustiva) con las que luego se realizará la Tabla. En este caso se trabajó con trece, sin embargo, se tiene como proyección expandir esta lista. A continuación, se resume cada transformación en la siguiente tabla:

Transformación	Significado	Ejemplo
1	Negación con "no"	-dar bola Ella le dio bola = Ella no le dio bola.
2	Negación con "sin"	-hacer bolsa Ella hizo bolsa sus juguetes = Sin que haga bolsa tus juguetes.
3	Voz pasiva	-hacer bolsa María hizo mierda el auto = El auto fue hecho mierda por María.
4	Elisión de C: C por PronCl	-pegar la PLR María le pegó la PLR= María <i>se la</i> pegó.
5	Elisión de objeto diecto: OD reemplazado por PronCL acusativo	-hacer mierda Juan hizo mierda a María = Juan la hizo mierda
6	Pluralización de los clíticos	-dejarle la cagada María le dejó = María les dejó la cagada.
7	Interrogación	-Dejar la escoba Juan dejó la escoba = <i>¿Quién</i> dejó la escoba?
8	Condicional con "si"	-Quedar pegado Juan quedó pegado = <i>si</i> Juan quedó pegado, María también.
9	C con flexión de número	-Estar cagado Él está cagado = Ellos están cagados.
10	C con flexión de género	-hacer huevón Él lo hizo huevón= Él <i>la</i> hizo huevona.
11	C con diminutivo	-hacer tuto Ellos hicieron tuto = Ellos hicieron <i>tutito</i> .
12	C con aumentativos	-caer gordo Él le cayó gordo= Él le cayó <i>gordísimo</i> .
13	Inclusión de un elemento entre V y C o solo en C	-Quedar la cagada Ayer quedó la cagada = Ayer quedó <i>la mansa</i> cagada.

Tabla 9. Transformaciones de locuciones verbales del español de Chile

Una vez elaborado el glosario de las transformaciones a considerar, se prosiguió con la confección de las tablas. Para cada clase de locución se realizó una tabla de transformaciones, como la que se puede observar:

Verbo	C1	Ejemplo	Transformaciones													
			T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	
hacer	mierda.	Ellos hicieron mierda todo lo que provenía de sus enemigos	+	+	+	-	+	-	+	+	-	-	-	-	-	-
hacer	mierda.	La podadora hizo mierda todas las flores	+	+	+	-	+	-	+	+	-	-	-	-	-	-
hacer	mierda.	Ellos se harán mierda la ropa si siguen peleando	+	+	-	-	+	-	+	+	-	-	-	-	-	-
hacer	mierda.	Viene el aluvión y los carteles se harán mierda con la tormenta	+	+	-	-	-	-	+	+	-	-	-	-	-	-
hacer	mierda.	Pamela me hará mierda las uñas	+	+	-	-	+	-	+	+	-	-	-	-	-	-
hacer	mierda.	Las espinas me harán mierda la ropa	+	+	-	-	+	-	+	+	-	-	-	-	-	-

Tabla 10. Ejemplo transformaciones locución Clase 1

Como se puede apreciar en la Tabla anterior, se detalla en primera instancia los componentes nucleares de la locución (V+C) y, posteriormente las transformaciones que admite, lo cual, al igual que en las tablas LG, se va indicando con un signo + si la expresión admite determinada modificación o con un – si no se permite. Tomando como ejemplo el caso de la Tabla 8, se puede señalar que la locución “hacer mierda” admite:

Ejemplo	Ellos hicieron mierda todo lo que provenía de sus enemigos.
T1	Ellos no hicieron mierda todo lo que provenía de sus enemigos
T2	Sin que ellos hicieran mierda todo lo que provenía de sus enemigos.
T3	Todo lo que provenía de sus enemigos fue hecho mierda por ellos.
T5	Ellos lo hicieron mierda
T7	¿ Quiénes hicieron mierda todo lo que provenía de sus enemigos?
T8	Si ellos hicieron mierda todo lo que provenía de sus enemigos...

Tabla 11. Ejemplo transformaciones “hacer mierda”

Como se puede observar la locución descrita no admite todas las transformaciones, por lo que construcciones como (13) *Ellos *los hicieron* (T4) o (14) *Ellos hicieron *mierdas* (T9) resultarán agramaticales.

Una vez realizados estos procedimientos que permiten realizar una exhaustiva descripción de las locuciones verbales del español de Chile, se procedió a realizar la formalización de estas estructuras empleando para ello la plataforma NooJ.

3.4.2 Trabajo computacional

Para llevar a cabo el procedimiento de detección automática de las locuciones verbales del español de Chile, fue necesario formalizar las estructuras que posee cada una de las clases de expresiones. Para ello, se siguieron los siguientes procedimientos: (i) elaboración de gramáticas léxicas; (ii) confección de un diccionario que contemplara todas las propiedades de cada clase de locución.

3.4.2.1 Elaboración de gramáticas y diccionarios

Cada una de las gramáticas elaboradas (ver Anexo 12) presenta como núcleos básicos: (i) el nombre de la clase de locución con la cual se etiquetará el corpus; (ii) la unidad V, que indica el núcleo verbal de la expresión; y (iii) los elementos que componen la parte “fija” de esta. Para aquella sección, se emplean cuatro nodos que admiten diversas combinatorias. De esta manera, la gramática puede detectar locuciones como:

Ejemplo	V	PREPC	CONTRC	DETC	C
Hacer mierda	Hacer	-	-	-	mierda
Dejar la escoba	Dejar	-	-	la	escoba
Atornillar al revés	Atornillar	-	al	-	revés
Caer de cajón	Caer	de	-	-	cajón
Agarrar para el chuleteo	Agarrar	para	-	el	chuleteo

Tabla 12. Ejemplo composición de elementos centrales de una locución verbal del español

El número en el elemento C cambiará según cada clase. Por ejemplo, para la Clase 1, se denomina C1 y en el caso de la Clase 2, se denomina C2 y así sucesivamente.

Las gramáticas irán variando según sus elementos constitutivos, para evidenciar esto se ejemplificará con las gramáticas de Clase 1 y Clase 2. Para comenzar, se mostrará la primera, debido a que estructuralmente contempla los elementos básicos de toda locución:

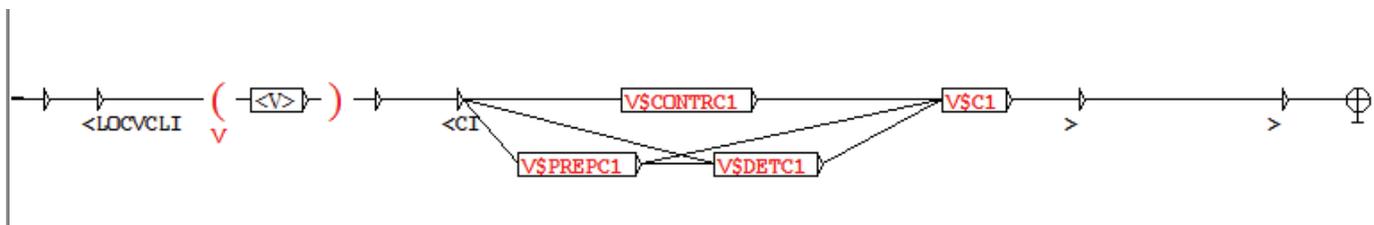


Imagen 1. Gramática léxica para detectar locuciones Clase 1

En la imagen anterior, se pueden observar los elementos antes referidos: al inicio del nodo, el nombre con el que se etiquetará la expresión que posea las características descritas en la gramática. Luego, la variable V, con la que se detectará el verbo nuclear de la locución. Y, posteriormente, se indican las combinatorias de elementos que pueden componer la sección “fija” de la locución. Cada uno de estos elementos está ligado a la variable V, lo cual genera que la gramática únicamente detecte las unidades allí declaradas. Para esto, aquella

información debe indicarse en las entradas del diccionario. De este modo, un ejemplo de entrada sería:

agarrar,V+FLX=AMAR+C1=papa+FXC

agachar,V+FLX=AMAR+DETC1=el+C1=moño+FXC

echar,V+FLX=AMAR+CONTRC1=al+C1=agua+FXC

caer,V+FLX=CAER+PREPC1=de+C1=cajón+FXC

Como se puede evidenciar, las locuciones deben declararse indicando, por un lado, el modelo de flexión que posee el verbo nuclear y luego se deben adicionar los elementos constitutivos de estas unidades. Una vez realizado aquello, se debe indicar el comando +FXC lo cual indica que la información referida en el diccionario pertenece a una “*frozen expresión*” o una unidad invariable. Por tanto, tanto las gramáticas léxicas como las entradas del diccionario trabajan de manera conjunta para llevar a cabo la detección. En la siguiente imagen se puede observar un ejemplo de la detección que realiza NooJ con la información antes referida.

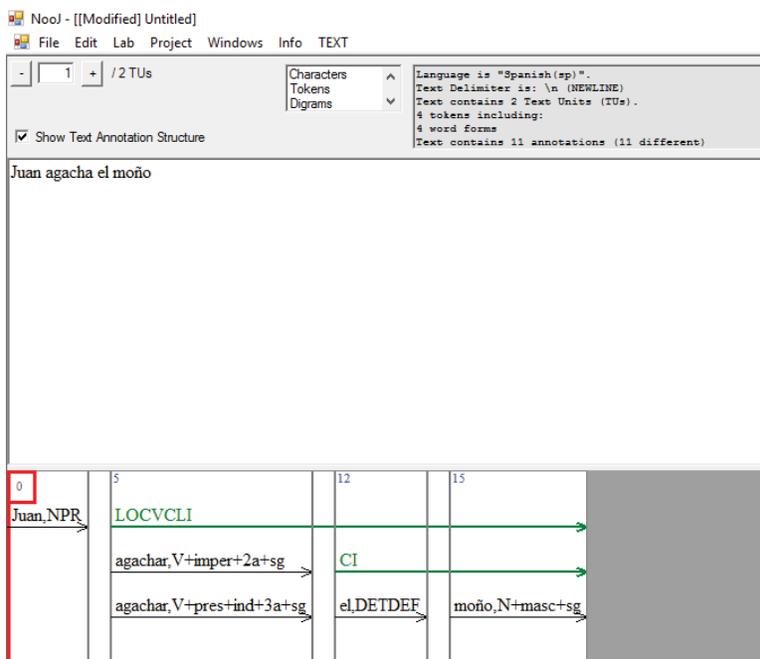


Imagen 2. Ejemplo de detección de locución Clase 1

A continuación, se evidencia la gramática léxica para detectar las locuciones pertenecientes a la Clase 2:

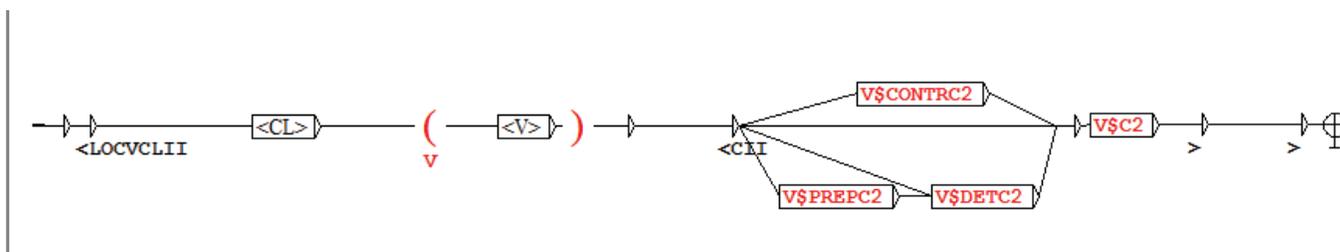


Imagen 3. Gramática léxica para detectar locuciones Clase 2

Como se puede apreciar, la única diferencia de esta gramática con la de Clase 1 es la inclusión del pronombre clítico antepuesto a la unidad verbal. Esta información también debe declararse en las entradas del diccionario. De este modo, al momento de realizar la detección NooJ discrimina entre los elementos que poseen un CL, como en este caso, de los que no, como en el caso anterior. A continuación, se puede apreciar las diferencias entre las entradas.

abrir, V+FLX=PARTIR+CL+C2=cancha+FXC

afirmar, V+FLX=AMAR+CL+DETC2=los+C2=pantalones+FXC

agarrar, V+FLX=AMAR+CL+CONTRC2=del+C2=moño+FXC

caer, V+FLX=CAER+CL+PREPC2=de+C2=maduro+FXC

Del mismo modo, se puede evidenciar en la siguiente imagen un ejemplo de detección. A diferencia del caso anterior, en esta oportunidad se han cargado los diccionarios de Clase 1 y Clase 2, para demostrar que la información declarada posibilita que NooJ discrimine adecuadamente entre un tipo de locución y el otro.

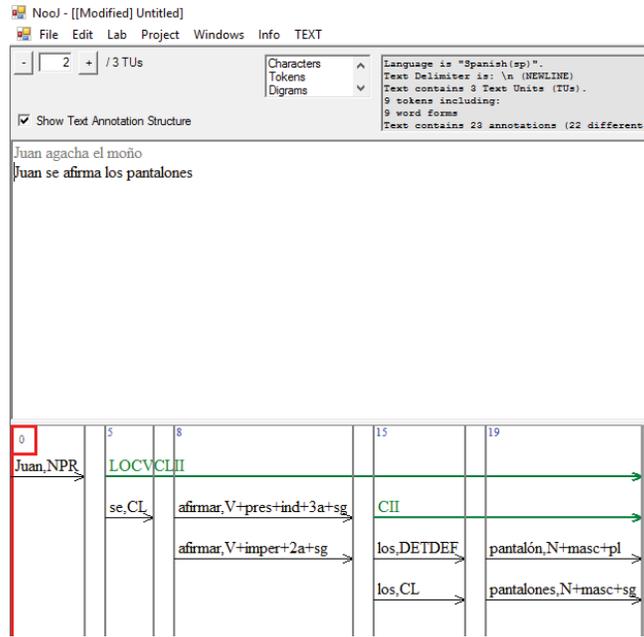


Imagen 4. Ejemplo de detección de locución Clase 2

Incluso si una locución de Clase 1 admite un pronombre clítico antepuesto al verbo, lo que se asemejaría a la gramática de la Clase 2, NooJ realiza la detección de manera adecuada, sin mezclar la información declarada en las gramáticas. Al igual que en el ejemplo anterior, en la siguiente imagen se muestra la detección realizada empleando los diccionarios con las respectivas gramáticas de las Clases 1 y 2.

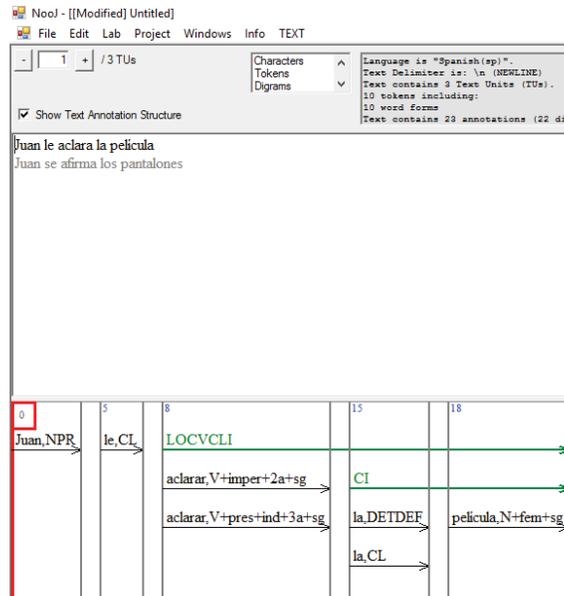


Imagen 5. Ejemplo de detección locución Clase 1 que admite pronombre

3.4.3 *Recopilación de corpus*

3.4.3.1 El corpus

El corpus empleado para poner a prueba la detección automática fue recolectado del periódico nacional La Cuarta. De este medio se extrajeron textos de una de sus secciones denominada “Ventanita sentimental”. En aquel espacio los lectores pueden escribir cartas dirigidas al “Doctor amor”, quien da respuesta a las diversas inquietudes de su público. Se seleccionó este apartado debido a que aquel medio se caracteriza por emplear un registro informal, y por ello se tuvo como supuesto que podrían ser textos ricos en modismos y locuciones verbales como las que se trabajan en esta investigación. En total, el texto está compuesto de 33.094 palabras, lo cual equivale a más de cien cartas con sus respectivas respuestas.

3.4.3.2 Detección manual

Una vez compilado el corpus, para poder comparar los resultados de la detección automática, se realizó una detección manual. Para ello, se hizo lectura del corpus con el propósito de detectar locuciones verbales. Una vez evidenciadas las unidades, se ratificaron con ayuda del DUECh. De este modo, ciertas expresiones que pudieran constituirse como locuciones, quedan fuera de la comparación al no estar consignadas en el diccionario con el que se trabaja. Una vez extraída la lista de locuciones, se lleva a cabo la comparación con el detector automático.

IV. RESULTADOS

4.1 Tablas léxico-gramática

4.1.1 Locuciones verbales del español de Chile

Para dar cuenta de las cualidades y requerimientos argumentales de cada locución se irán evidenciando los resultados obtenidos a partir de la elaboración de la matriz. En primer lugar, se mostrarán los resultados que presenta el N0 en cada Clase.

CLASE	N0					
	PREP		+Anim		-Anim	
	A					
	Cantidad	%	Cantidad	%	Cantidad	%
1	0	0,0%	651	81,6%	68	8,5%
2	0	0,0%	445	96,9%	4	0,9%
3	0	0,0%	12	100,0%	0	0
4	6	100,0%	5	83,3%	1	16,7%
5	36	100,0%	35	97,2%	1	2,8%
6	40	100,0%	40	100,0%	0	0,0%
7	4	100,0%	4	100,0%	0	0,0%
8	2	100,0%	2	100,0%	0	0,0%
9	7	100,0%	7	100,0%	0	0,0%
10	2	100,0%	2	100,0%	0	0,0%

Tabla 13. Resultados N0 por clase

Para obtener los resultados expuestos, en los casos en que las locuciones podrían o no admitir cierta propiedad, indicado con el digno (+), se le otorgó de igual manera el valor de aquellos elementos que exige la locución. Esto es relevante principalmente cuando se describe el primer argumento, pues solo se observó un caso en el que la locución no admitía argumento 1. Este es el caso de la locución Clase 1 “cantar Gardel”. Como se puede apreciar el elemento fijo en la locución es un sintagma nominal. Esta expresión es interesante, debido a que el elemento petrificado guardaría relación con el sujeto de la oración. Por este motivo, el verbo solo admite conjugaciones de 3era persona singular. Esto implica, que solo se tendrá significado unitario expresiones como (15) Aún no *canta Gardel* o (16) Mañana *cantará Gardel*. Si el verbo se conjuga en las otras personas gramaticales la expresión pierde la categoría de locución, en caso como (17) *Nosotros cantamos Gardel.

Otro aspecto a destacar respecto de este primer argumento refiere al uso de la preposición que se antepone al sujeto. Como se puede apreciar, únicamente desde las Clase 4 a la 10 se declara esta información en la matriz. Esto guarda relación al criterio inicial que se empleó para realizar la división por Clases, pues las entradas en el DUECh declaran que las locuciones de las clases indicadas exigen un elemento +Anim, por ejemplo “caerle a alguien el tejo”. En este sentido se decidió dar cuenta de aquella propiedad en la Tabla, ya que solo en dos casos excepcionales las locuciones de este rango de clases no admiten un N0 +Anim. No obstante, en términos generales, se puede apreciar claramente que las Clases requieren un N0 con estas características, aunque perfectamente puede estar elidido.

En relación con los pronombres que admite cada Clase, en la siguiente tabla se da cuenta del detalle:

CLASE	PronCI					
	Se		Dativo		Acusativo	
	Cantidad	%	Cantidad	%	Cantidad	%
1	132	16,5%	232	29,1%	0	0,0%
2	182	39,7%	275	59,9%	13	2,8%
3	6	50,0%	6	50,0%	12	100,0%
4	0	0,0%	6	100,0%	0	0,0%
5	0	0,0%	36	100,0%	0	0,0%
6	40	100,0%	40	100,0%	0	0,0%
7	4	100,0%	4	100,0%	0	0,0%
8	0	0,0%	2	100,0%	0	0,0%
9	7	100,0%	7	100,0%	0	0,0%
10	0	0,0%	1	50,0%	1	50,0%

Tabla 14. Resultados PronCI por Clase

Como se puede observar, los pronombres más recurrentes son los dativos, que dan cuenta de quien recibe determinada acción; luego, el pronombre “se”, sin tomar en consideración su función sintáctica; y por último los acusativos. En el caso de estos, resulta llamativo observar que la única Clase que requiere este tipo de clíticos es la 3, la cual carece de elemento C3. Esto implica que el conocimiento de la lengua de los hablantes del español de Chile, comprenden a qué remite aquel elemento elidido y reemplazado por un pronombre. Un ejemplo de esto se observa en el caso de la locución “sabérselas por libro”, expresión

que significa que “Tener mucha habilidad en algo debido a la experiencia” y se usan en oraciones como (18) Ella cree que se las *sabe por libro*. Aquel pronombre “las” contiene la información de aquello que es sabido, en otras palabras, contiene al argumento N0. Por esto, no es casual que en esta clase, la unidad C3 de la locución está compuesta ya sea por sintagmas adjetivos o preposicionales, los aporta información de aquel elemento elidido.

Para continuar, en la Tabla 13 se da cuenta del tipo de sintagma al que pertenece el elemento “fijo” de la locución:

CLASE	C											
	SD		SN		SP		SA		SAd		SV	
	Cantidad	%	Cantidad	%	Cantidad	%	Cantidad	%	Cantidad	%	Cantidad	%
1	343	43,0%	219	27,4%	152	19,0%	39	4,9%	30	3,8%	16	2,0%
2	128	27,9%	145	31,6%	166	36,2%	15	3,3%	4	0,9%	6	1,3%
3	0	0,0%	0	0,0%	4	33,3%	8	66,7%	0	0,0%	0	0,0%
4	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%
5	17	47,2%	15	41,7%	7	19,4%	1	2,8%	2	5,6%	0	0,0%
6	40	100,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%
7	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%
8	2	100,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%
9	0	0,0%	0	0,0%	7	100,0%	1	14,3%	0	0,0%	0	0,0%
10	2	100,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%

Tabla 15. Resultados C, tipo de sintagma por Clase

A partir de esta información, se desprende que las locuciones verbales del español del Chile tienden a emplear: sintagmas determinantes, nominales o bien preposicionales. Esto hace sentido si se observa el contenido que poseen los elementos contenidos en C, pues en muchos casos la información petrificada contiene uno de los argumentos que exige el verbo, sin embargo, como “agarrar papa”. El verbo agarrar, cuando no se encuentra contenido en una locución puede requerir tres argumentos, como en el caso de (19) Juan agarró los lápices con los pies. Sin embargo, al ser núcleo de una locución pierde uno de sus argumentos, pues este pasa a formar parte de la sección fija de la expresión, por ello, se puede emplear la locución en oraciones como (20) Juan *agarró papa* y le contó todo a María.

Finalmente, se indican los casos en los que las locuciones requiere un argumento N1, ya sea seguido o no por una preposición. Como se observa en la Tabla 14 la preposición más recurrente es “a”, la cual sirve para introducir un N1 ya sea +Animado o -Animado. Como, por ejemplo, la locución “hacer huevón”, en casos como (21) Juan *hizo huevón* a Pedro.

Asimismo, como ya se mencionó anteriormente, en el caso de la expresión “hacer mierda”, la que exige un argumento N1, sin embargo, no requiere preposición introductoria, como en (22) Ellos *hicieron mierda* el auto. No obstante, como se observa en los resultados, tanto el uso de preposiciones para introducir un nuevo argumento, o bien la exigencia de un segundo argumento por parte de la locución, no son cualidades que se puedan generalizar, pues la tendencia es que las locuciones necesitan por lo menos un argumento N0 para tener sentido.

CLASE	PREP										CLASE	N1	
	a		de		con		por		para			Cantidad	%
	Cantidad	%	Cantidad	%	Cantidad	%	Cantidad	%	Cantidad	%			
1	175	21,9%	0	0,0%	8	1,0%	1	0,1%	4	0,5%	1	231	28,9%
2	86	18,7%	0	0,0%	3	0,7%	2	0,4%	0	0,0%	2	107	23,3%
3	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	3	0	0,0%
4	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	4	0	0,0%
5	0	0,0%	1	2,8%	0	0,0%	2	5,6%	0	0,0%	5	3	8,3%
6	0	0,0%	0	0,0%	0	0,0%	1	2,5%	0	0,0%	6	1	2,5%
7	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	7	1	25,0%
8	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	8	0	0,0%
9	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	9	0	0,0%
10	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	10	0	0,0%

Tabla 16. Resultados PREP y N1 por Clase

4.1.2 Transformaciones

Una vez descritas las diversas propiedades de las locuciones, cobra relevancia dar cuenta de las transformaciones que admite cada una de las entradas de la matriz. Para ello, se adjuntan dos Tablas (17 y 18) en las que se puede apreciar la cantidad de transformaciones que admiten las locuciones por cada Clase y también, una Tabla en la que se indica la cantidad promedio de transformaciones que admite cada clase.

CLASE	T1		T2		T3		T4		T5		T6		T7	
	TOTAL	%	TOTAL	%	TOTAL	%	TOTAL	%	TOTAL	%	TOTAL	%	TOTAL	%
1	798	100,0%	796	99,7%	11	1,4%	379	47,5%	118	14,8%	148	18,5%	724	90,7%
2	459	98,7%	459	98,7%	0	0,0%	190	40,9%	30	6,5%	115	24,7%	326	70,1%
3	12	100,0%	12	100,0%	0	0,0%	2	16,7%	2	16,7%	6	50,0%	12	100,0%
4	6	100,0%	6	100,0%	0	0,0%	0	0,0%	0	0,0%	6	100,0%	0	0,0%
5	36	100,0%	36	100,0%	0	0,0%	31	86,1%	0	0,0%	36	100,0%	0	0,0%
6	40	100,0%	40	100,0%	0	0,0%	33	82,5%	0	0,0%	40	100,0%	0	0,0%
7	4	100,0%	4	100,0%	0	0,0%	0	0,0%	0	0,0%	4	100,0%	0	0,0%
8	2	100,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	2	100,0%	0	0,0%
9	7	100,0%	6	85,7%	0	0,0%	0	0,0%	0	0,0%	7	100,0%	0	0,0%
10	2	100,0%	2	100,0%	0	0,0%	0	0,0%	0	0,0%	2	100,0%	0	0,0%

Tabla 17. Resultados transformaciones de la 1 a la 7 por Clase

CLASE	T8		T9		T10		T11		T12		T13	
	TOTAL	%	TOTAL	%	TOTAL	%	TOTAL	%	TOTAL	%	TOTAL	%
1	795	99,6%	30	3,8%	28	3,5%	68	8,5%	5	0,6%	155	19,4%
2	459	98,7%	10	2,2%	20	4,3%	19	4,1%	0	0,0%	110	23,7%
3	12	100,0%	0	0,0%	0	0,0%	6	50,0%	0	0,0%	8	66,7%
4	6	100,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%
5	36	100,0%	0	0,0%	0	0,0%	7	19,4%	0	0,0%	13	36,1%
6	40	100,0%	2	5,0%	0	0,0%	0	0,0%	0	0,0%	15	37,5%
7	4	100,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%
8	2	100,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%
9	7	100,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%
10	2	100,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%

Tabla 18. Resultados transformaciones de la 8 a la 13 por Clase

Como se puede observar, por un lado, existen tres transformaciones que se permiten de manera transversal en todas las Clases (T1, T2 y T8). Esto se traduce a que toda locución admite negación con el adverbio “no”, condicional con “si” y las interrogaciones. Por otro lado, las transformaciones menos comunes son T3 y T12, esto es, voz pasiva e incluir sufijo aumentativo en la sección C de la locución. A este respecto, cobra relevancia que las únicas transformaciones a oración pasiva son admitidas en la Clase 1 en locuciones cuyo verbo nuclear es “hacer”.

Cabe señalar que las transformaciones con las que se trabajó en este trabajo no son las únicas que podrían admitir las locuciones. Sin embargo, se consideró que pudieran ser las más recurrentes. Sin embargo, cobra relevancia poder dar cuenta de la mayor cantidad de alternativas que admiten la expresión de una locución, de este modo se aporta a la exhaustividad de la descripción. Por ello, en el futuro se buscará, no solo incluir más transformaciones a la lista, sino también, la formalización de estas para poder detectar y

generar a partir de reglas computacionales, de manera automática, cada una de las alternativas.

Finalmente, se muestra un recuadro en el que se da cuenta del promedio de transformaciones que admite cada una de las locuciones según la Clase a la que pertenecen. A este respecto, se observa que la clase que más transformaciones admite es la 3 con una cantidad de seis posibilidades de dar cuenta de una locución y la que menos permite es la 8 con un promedio de tres transformaciones. Estos resultados permiten comprobar que aquella visión de las locuciones como elementos fijos e invariables debe dejarse, pues estas expresiones poseen variaciones que dan cuenta tanto de la riqueza de las propiedades de estas unidades, como de los variados usos que les puede dar un hablante, sin que la locución pierda su significado unitario.

CLASE	PROMEDIO
1	5,07
2	4,79
3	6,00
4	4,00
5	5,42
6	5,25
7	4,00
8	3,00
9	3,86
10	4,00

Tabla 19. Resultados promedio de transformaciones que admite cada Clase

4.2 Reconocimiento automático

Para llevar a cabo la detección automática, se cargaron en NooJ los diccionarios correspondientes a cada una de las Clases de locuciones, con sus respectivas gramáticas. De este procedimiento se obtienen resultados como:

Text	Before	Seq.	After
	ganas locas de salir a	mover el esqueleto	y sacudirse las pulgas con
	con mi señora es al	hacer las tareas	. La casa es chica y
	pesca, y me pide sólo	hacer las tareas	. El cuerpo no me da
	películas, pero para eso debe	tirar toda la carne a la parrilla	y ver si la cosa
	conversar, después nos abrazamos, y	quedó la tendalada	. Nos agarramos a besos, nos
	tengo una polola como para	sacar pica	en mi ciudad natal. Hace
	las ideas del encuentro y	dijo upa	. De ahí a comer, unos
	vamos por calor humano, me	dijo upa	, chahupa. Estoy re feliz, jefe
	es que mis compañeros me	agarraron para el palanqueo	, señalándome que sólo quiero acostarme
	reina, la besaré y le	sacará pica	a sus ex malos compañeros
	considerando que donde estaba le	atomillaban al revés	. 19 Doc: No sé qué hacer
	llamé tipo una, dijo que	venía de vuelta	y nada... Me la encontré
	que de a poco irán	cachando el mote	. Sin embargo, eso no quiere
	Es que le dije que	duraba menos que un candy	, y usted me señaló que
	feliz, pues me dice que	soy una máquina	sexual. Me halaga todo el
	no quiero parar y temo	dejar la escoba	en casa. César Emperador: No
	y, obvio, sin mina joven. ¿	Cachó el mote	? Ya, ahora enmiende el rumbo
	la cabeza al WC y	tiró la cadena	con ganas. No le creo
	el baño y que lo	hagan bolsa	por pobre. Si la jermu
	Quizás usted en muchos años	valió callampa	en el catre o sólo
	a su familia. El fútbol	vale callampa	frente a un problema de
	cada vez que puede me	saca pica	con ella cuando entra a
	Y que su hermano se	haga un lulo	. Así es la diversidad, mi
	casa de la Ana para	hacer las tareas	. En eso estábamos hace un
	y más al momento de	gritar viva Chile	Pero siempre hubo un abejorro
	la veterana es porque le	calienta la sopa	nomás. Pero puro problema. Es
	Pregúntele cuánto le queda para	gritar viva Chile	, respírele en la oreja y
	Doctor Cariño: Ahora sí que	corto las huinchas	. Llevo dos años y medio

Imagen 6. Ejemplo de detección de locuciones Clase 1

Además, como se mencionó en la sección anterior, los resultados de la detección fueron comparados con la realizada manualmente. De este modo, se pueden obtener los porcentajes de precisión, cobertura y medida F, necesarios para medir el trabajo computacional ejecutado. Para dar cuenta de los resultados, se indicarán los valores antes indicados, según la detección general y luego por cada clase. Cabe señalar que para efectos de los resultados se contemplan las locuciones que fueron formalizadas por medio de reglas, por tanto, en este trabajo no se contemplarán aquellas locuciones que aparecen en el DUECh, pero que en el corpus se emplearon con pronombre enclíticos o ciertas modificaciones de género. La detección de este tipo de estructuras se contempla como una proyección de este trabajo.

Resultados	Resumen
Locuciones por clase	103
Detectó bien	95
Detectó mal	3

Detección total	98
Cobertura	92,23
Precisión	96,94
Medida F	94,53

Tabla 20. Resultados detección por total

Como se puede evidenciar en la Tabla en términos generales se obtuvo resultados positivos en cuanto a los porcentajes de precisión y cobertura. Esto permite comprobar que la formalización realizada a partir del aparato metodológico que entrega la LG fue la adecuada. Sin embargo, debido a la limitada cantidad de Clases detectadas, no se pueden validar los recursos elaborados para las Clases no presentes. Ahora bien, compilar un corpus que contenga locuciones de todas las clases se constituye como un desafío, debido al tipo de expresión que se busca detectar.

En relación con los resultados obtenidos por clase, se puede indicar, que la mayor cantidad de locuciones se concentran en la clase 1 y 2. Este resultado no llama la atención considerando que más del 50% de los lemas o entradas en el DUECh pertenece a estas clases, por lo que se constituyen como locuciones diversas y a la vez comunes en el uso del español chileno. En cuando a los porcentajes obtenidos, se puede observar que los resultados más descendidos se encuentran en la cobertura, lo que implica que para próximos trabajos se requerirá realizar gramáticas en las que se incluyan más propiedades de las locuciones.

Resultados	Clase 1	Clase 2	Clase 4	Clase 5
Locuciones por clase	61	47	3	2
Detectó bien	52	39	3	1
Detectó mal	1	2	0	0
Detección total	53	41	3	1
Cobertura	85,25	82,98	100	50
Precisión	98,11	95,12	100	100
Medida F	91,23	88,64	100	66,67

Tabla 20. Resultados detección por Clase

Del mismo modo, se deberá revisar las cualidades declaradas en los diccionarios para detectar aquellas 9 locuciones que quedaron fuera, en el caso de la Clase 1, 8 en relación con la 2 y 1 en la Clase 5. No obstante, se puede señalar que los resultados son

satisfactorios como primera aproximación al fenómeno de las locuciones verbales en el español de Chile.

V. CONCLUSIÓN

Para concluir, cobra relevancia recordar los objetivos generales propuestos. Esto eran dos: (i) describir las propiedades morfosintácticas y la estructura argumental que proyectan las locuciones verbales del español de Chile; y, (ii) detectar locuciones verbales del español de Chile en textos de lenguaje natural. Estos dos grandes propósitos guiaron el quehacer investigativo y a partir de los resultados evidenciados en el apartado anterior, se puede señalar que se consiguieron ambos objetivos. A este respecto, es importante indicar la efectividad del aparato metodológico que aporta la LG en trabajos descriptivos como este, pues por medio Tablas se pueden extraer las cualidades de unidades como las locuciones y a su vez, permiten formalizar las estructuras para poder llevar a cabo la implantación en máquina.

Así, por un lado, fue posible develar que la generalidad de las locuciones requiere un argumento N0 +Animado, pues la tendencia es que “alguien” realice la acción implicada en la locución. También, se observó que, en relación con el uso de pronombres, un amplio porcentaje de locuciones se emplea con pronombres dativos, los que permiten no incluir el argumento N1. En cuando a los elementos “fijos” de las locuciones, se observó que en la mayoría de las clases estas unidades pertenecen a sintagmas determinantes, nominales o preposicionales. Y, por último, se dio cuenta de la poca recurrencia al empleo del argumento N1, ya sea introducido o no por preposiciones.

Esta descripción posibilitó la formalización de estas unidades para realizar la posterior detección computacional, se obtuvieron resultados positivos en cuanto a la detección, pues todos los indicadores, ya sea, precisión, cobertura y medida F, estuvieron por sobre un 80%. Esto, gracias también a las propiedades que posee el software, ya que, la elaboración de diccionario entrega la opción de introducir un comando especial para el tipo de unidades que se desea detectar.

Por otro lado, se observó que, de las trece transformaciones contempladas en este estudio descriptivo, existen tres que son recurrentes de manera transversal a las clases. Asimismo, cobra relevancia indicar el alto promedio de transformaciones que admiten las locuciones de cada clase pues eso da cuenta estas unidades son ricas en cuanto a propiedades

sintáctico-semánticas y morfológicas. Y, por tanto, pese a que tradicionalmente fueron consideradas como invariables, se debe renovar aquella visión.

Ahora bien, a pesar de que, en general, se puede señalar que se lograron los objetivos propuestos, quedaron algunos aspectos fuera del alcance de esta investigación, que deberán ser abordados a futuro. Por ello, quedan como proyecciones, la elaboración de tablas LG para las locuciones perteneciente al segundo grupo, esto es, las que en su entrada del diccionario poseen el adverbio de negación “no”. Esto para posteriormente realizar el procedimiento computacional respectivo.

También, queda como proyección la elaboración de reglas gramaticales que permitan la detección de locuciones que se emplean con uno o dos pronombres enclíticos en la unidad verbal. De esta manera, expresiones como (23) Debes *hacerlos pebre*, puedan ser detectadas de manera correcta por el programa computacional. Del mismo modo, cobra relevancia, como ya se mencionó anteriormente, tanto la inclusión de más transformaciones a la tabla, como la formalización de estas, para complejizar el procedimiento de detección. Para ello, se podría requerir de una mayor descripción de los elementos constitutivos de las locuciones, lo que se traduce a una descripción más detallada de las propiedades que estas unidades.

En síntesis, el trabajo a futuro en relación con el trabajo descriptivo: (i) complejizar aún más la descripción lingüística realizada en las tablas LG, para mejorar detección automática; (ii) extender las posibilidades transformativas que poseen las locuciones verbales del español de Chile; y, (iii) elaborar tablas para locuciones cuyo lema en el DUECh contiene el adverbio de negación “no”. En cuanto al trabajo computacional: (i) complejizar la información lingüística incluida en las gramáticas, para mejorar porcentajes de precisión y cobertura; (ii) elaborar gramáticas que permitan tanto la detección como la generación de las transformaciones que admiten las locuciones; y, (iii) formalizar las cualidades morfosintácticas de las locuciones pertenecientes al segundo grupo para implantación en máquina.

Finalmente, se puede referir que una de las principales limitaciones en este trabajo fue la compilación del corpus. Esto debido a que el tipo de unidades formalizadas no se emplean de manera recurrente en textos escritos, al pertenecer a registros informales que tienden a la

oralidad. Por ello, una de las mayores dificultades fue encontrar un corpus que permitiera probar las gramáticas y diccionarios elaborados.

VI. BIBLIOGRAFÍA

Academia Chilena de la Lengua. (2010). *Diccionario de uso del español de Chile*. Santiago: MN Editorial Ltda.

Berri, M. & Boschioli, V. (2017). "Representación lexicográfica de voces argóticas del español de la Argentina: problemas de marcación, definición y ejemplificación" (pp. 205-218). En Conde, O. (2017). *Argots hispánicos. Analogías y diferencias en las hablas populares iberoamericanas*. Buenos Aires: Universidad Nacional de Lanús.

Corpas, G. (1996). *Manual de fraseología española*. Gredos: Madrid.

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.

Koza, W., Alfaro, P., & Martínez, R. (2015). "Detección automática de chilenismos verbales a partir de reglas morfosintácticas. Resultados preliminares". *Procesamiento del Lenguaje Natural*, (54): 69-76.

Koza, W. (2017). "La estructura de la enumeración. Análisis, descripción y propuesta de detección automática". *Onomázein* 35, 173-194.

Mendívil Giró, J. L. (1994). "Teoría gramatical y sintaxis: el lexique grammair de M. Gross, sus límites y posibilidades". En Corcuera, J., Dijan, M. & Gaspar, A. (1994). *La lingüística francesa. Situación y perspectivas a finales del Siglo XX*, Zaragoza.

Mendívil Giró, J. L. (1999). *Las palabras disgregadas: sintaxis de las expresiones idiomáticas y los predicados complejos*. Zaragoza: Prensas Universitarias de Zaragoza.

Monteleone, M. & Marano, F. (2011). "From the concept of transformation in Harris and Chomsky to the Lexique-grammaire of Maurice Gross" (pp. 76-82). En AA.VV. (2011). *History of Linguistics*. John Benjamins.

Orduña, J. L. (2011). "Estudio gramatical de las locuciones verbales con doble pronombre clítico". *RLA. Revista de lingüística teórica y aplicada*, 49(2), 87-110.

Penandés, I. (2012). *Gramática y semántica de las locuciones*. Alcalá de Henares: Servicio de Publicaciones de la Universidad de Alcalá.

Real Academia de la Lengua. "Gancho". <http://dle.rae.es/?id=IqCcR5P>

Silberztein, M. (2003). *NooJ Manual*. <http://www.NooJ4nlp.net/>

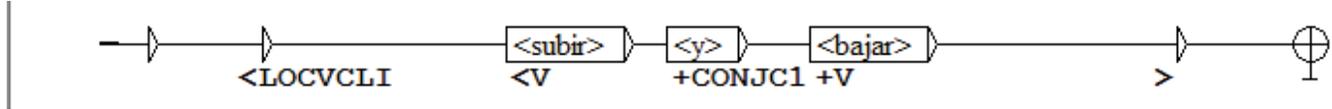
Silberztein, M. (2005). *NooJ: A Linguistic Annotation System For Corpus Processing*. [<http://www.aclweb.org/anthology/H05-2>, fecha de consulta: 20 de diciembre de 2015].

- Silberztein, M. (2016). *Formalizing Natural Languages: The NooJ Approach*. Wiley Eds.
- Tercedor Sánchez, M.I. (1999). *La fraseología en el lenguaje biomédico: análisis desde las necesidades del traductor*. PhD Thesis. Universidad de Granada.
- Vietri, S. (2014). *Idiomatic Constructions in Italian. A Lexicon-Grammar Approach*. John Benjamins BV: Amsterdam Netherlands.
- Zuluaga Gómez, F. (2004). “Locuciones, dichos y refranes sobre el lenguaje: unidades fraseológicas fijas e interacción verbal”. *Forma y función*, (18): 250-282.
- Zuluaga, A. (1975). “La fijación fraseológica”. *Thesaurus*, XXX, 235-288.

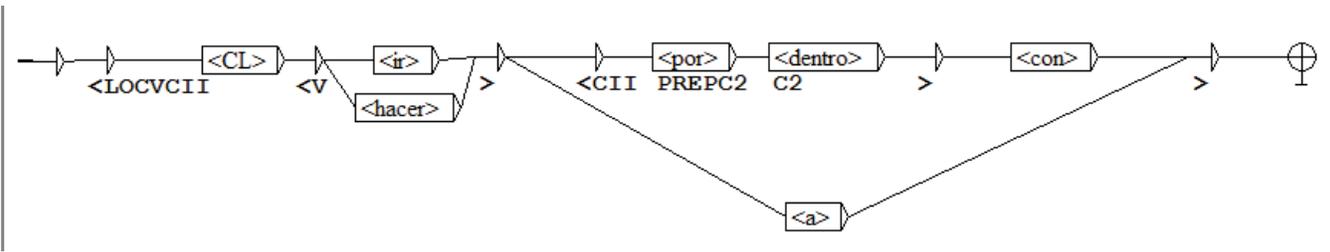
Anexos

Anexo 1: Gramáticas

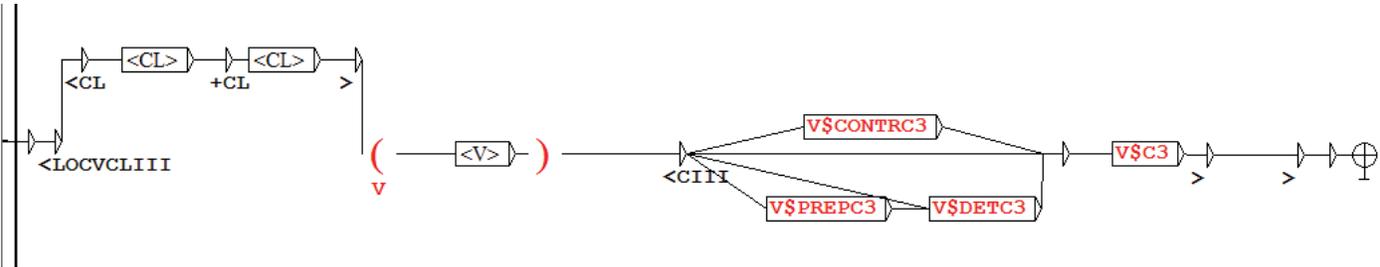
- Clase 1.1



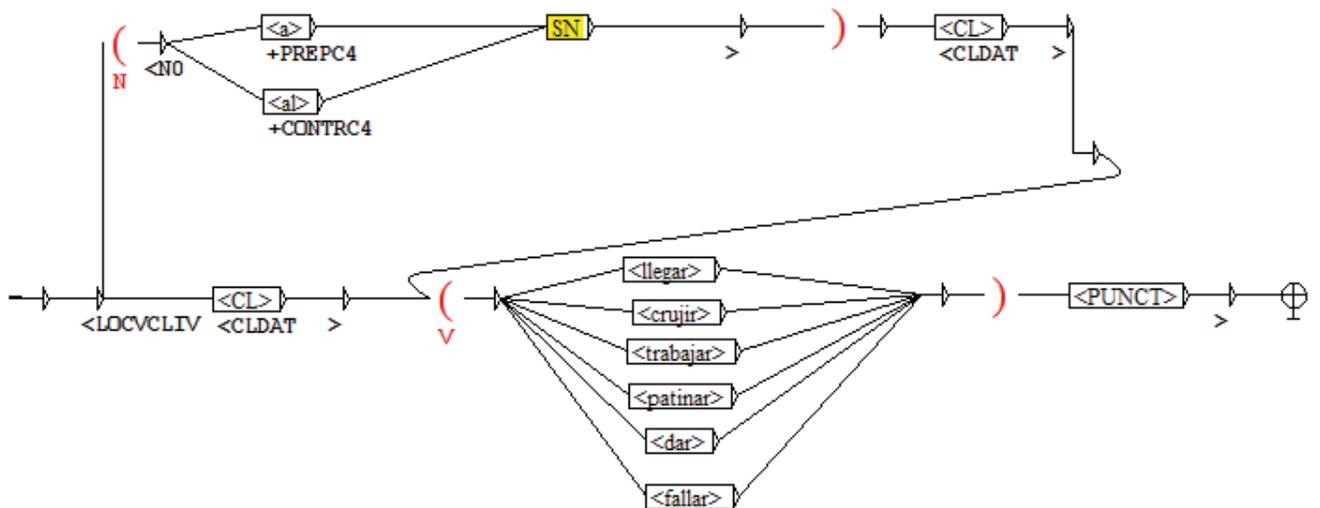
- Clase 2.1



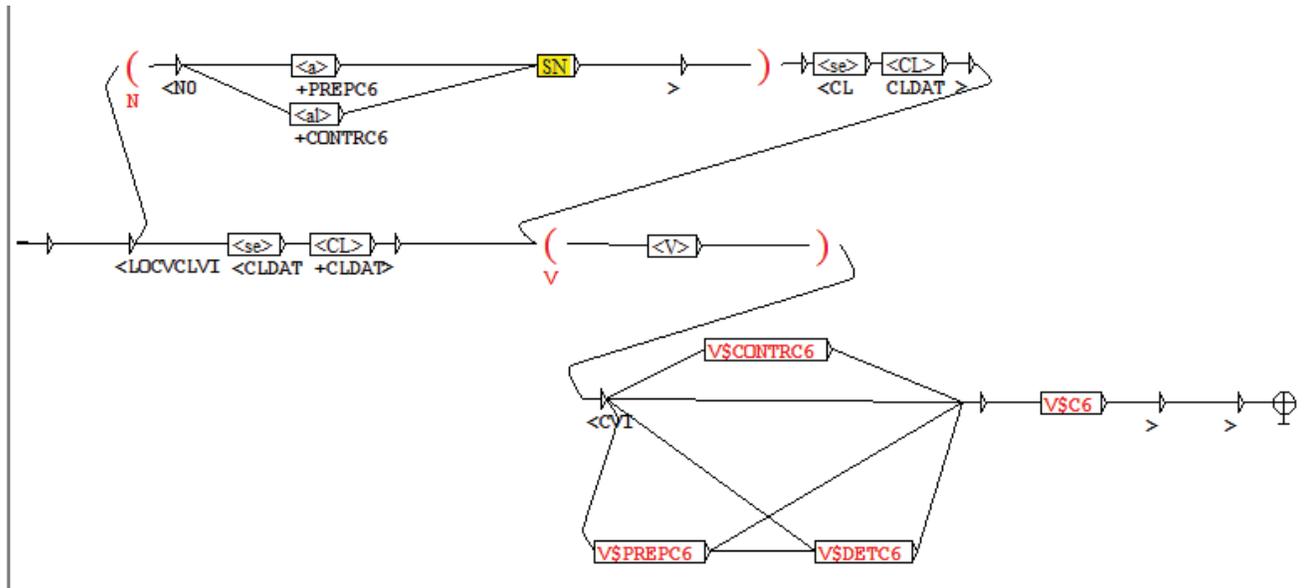
- Clase 3



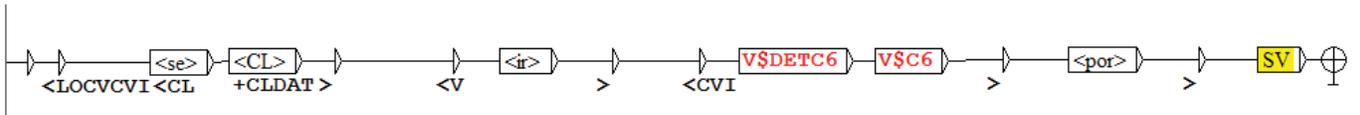
- Clase 4



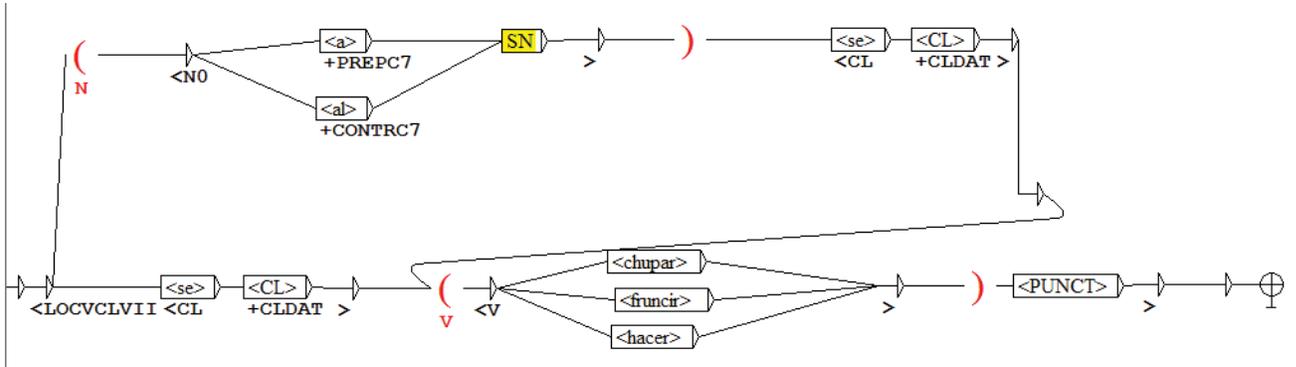
- Clase 6



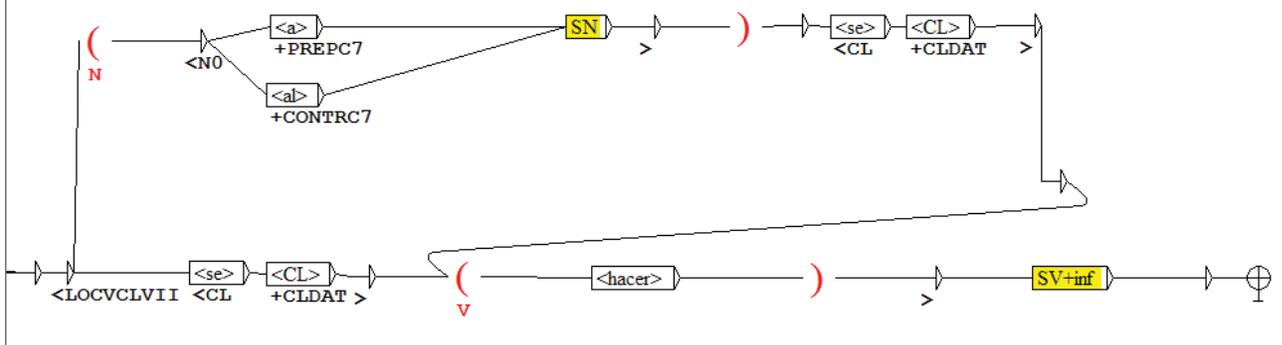
- Clase 6.1



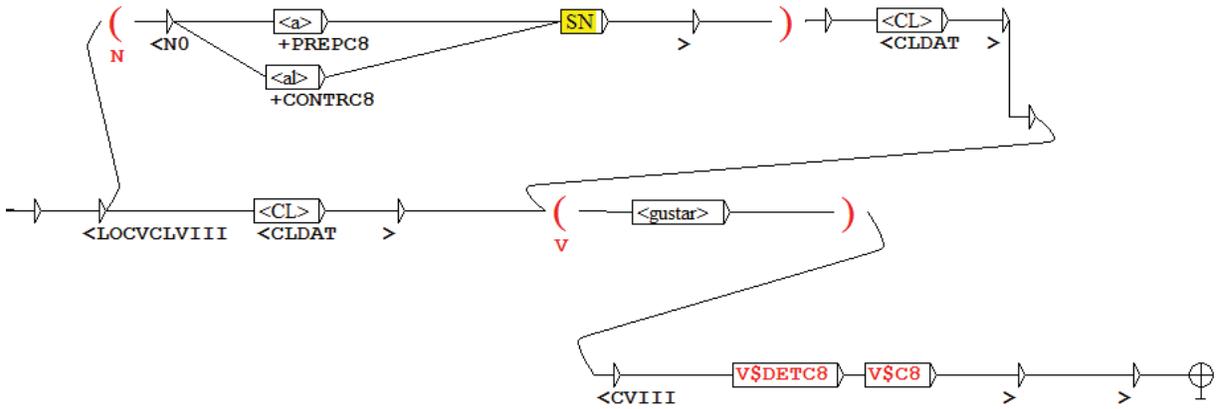
- Clase 7



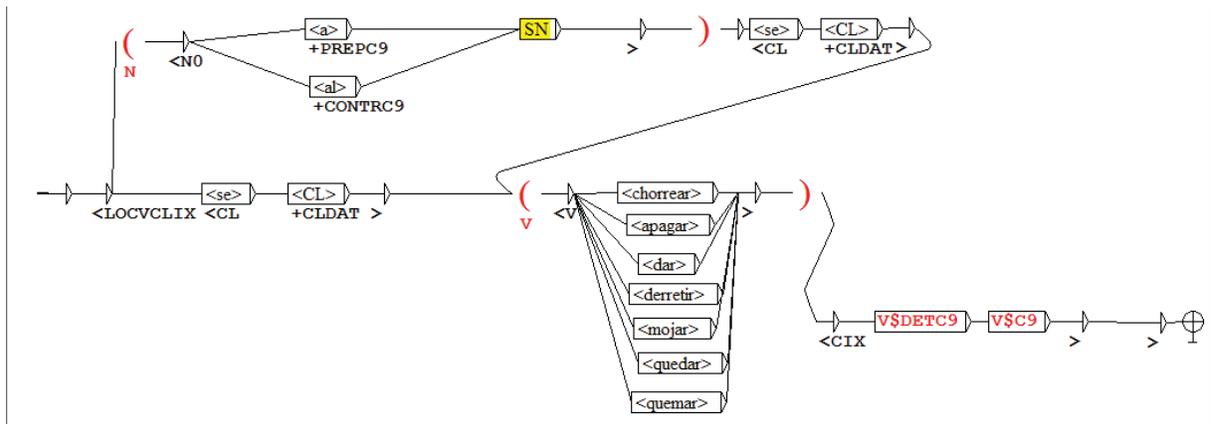
- Clase 7.1



- Clase 8



- Clase 9



- Clase 10

