

## **Reconocimiento de palabras con estructura externa mediante NooJ: una propuesta para estudiantes de español como lengua segunda y extranjera (ELSE)**

Recognition of words with external structure through NooJ: a proposal for students of Spanish as a second and foreign language (ELSE)

**Barrios, Luisina**

Facultad de Humanidades y Artes

Argentina

[luisinabarriosa@gmail.com](mailto:luisinabarriosa@gmail.com)

**Grynberg, Sofía**

Facultad de Humanidades y Artes

Argentina

[s.grynberg474@gmail.com](mailto:s.grynberg474@gmail.com)

### **Resumen**

El presente trabajo tiene como objetivo analizar frases típicas o idiomatismos del español rioplatense, entendidas por la gramática como palabras con estructura externa que son de uso extendido en la variante rioplatense correspondiente a Argentina. Este estudio se realizará mediante el uso de un software de acceso libre que es el sistema *NooJ* para lograr el reconocimiento de dichas construcciones.

Por lo tanto, esta investigación se orienta a reconocer frases idiomáticas del español a partir del sistema *NooJ*, para lo cual se clasificarán las frases del tipo: “clavar el visto”, “pegar un tubazo”, “estirar la pata” y se las etiquetará con el fin de localizarlas en textos de español y contribuir así, a la enseñanza y aprendizaje de este tipo de construcciones por estudiantes de español como lengua segunda y extranjera (ELSE).

Palabras clave: palabras con *estructura externa*, español como lengua segunda y extranjera, lingüística computacional, sistema *NooJ*.

## Abstract

The aim of this paper is to analyze typical phrases or idioms of Rioplatense Spanish, understood by grammar as words with external structure that are widely used in their Argentine Rioplatense variant. This study will be carried out through the use of an open access software, the NooJ system, to achieve the recognition of such constructions.

Therefore, this research is aimed at recognizing idiomatic phrases of Spanish from the NooJ system, for which phrases of the type: "clavar el visto", "pegar un tubazo", "estirar la pata" will be classified and labelled in order to locate them in Spanish texts and thus contribute to the teaching and learning of this type of constructions by students of Spanish as a second and foreign language (ELSE).

Key words: words with external structure, Spanish as a second and foreign language, computational linguistics, *NooJ* system.

## Introducción

Un adulto, hablante promedio de español, aunque se le dé un año para hacerlo, no podrá siquiera acercarse a una explicación certera de cómo se clasifican las frases del tipo “meter la pata” o “hacerse el ganso”. Sin embargo, esta misma persona probablemente nunca, a lo largo de su vida adulta, haya cometido un error en el uso de este tipo de frases, ni las haya usado de un modo inapropiado a sus intenciones o en situaciones en las que podrían no interpretarse adecuadamente.

Son conocidas las frases del tipo “sacate la gorra”, “me clavó el visto”, “pegame un tubazo”, “es re vigilante”, “se hace el banana”, “hizo una burrada”. Este tipo de construcciones, algunas de carácter nominal, otras verbal, han acarreado para los gramáticos y estudiosos de la morfología derivativa numerosos problemas.

Sabemos que “clavar el visto” significa leer un mensaje pero no responderlo y “pegar un tubazo” significa llamar a alguien por teléfono. Sin embargo, los gramáticos dedicados a la morfología derivativa se enfrentan aquí a una dificultad. Por un lado, los adeptos a las teorías lexicalistas establecen que este tipo de frases se forman y se ubican en el lexicon (el de los hablantes de español rioplatense) y allí se almacena su significado. Por otro

lado, para los teóricos sintactistas estas construcciones se forman en la sintaxis para luego almacenarse en el diccionario.

El modelo que Aronoff (1976) propone que los significados idiosincráticos son los que se almacenan en el diccionario. Por ejemplo, en “hizo una burrada” el significado literal que permite la morfología es “unió a un montón de burros”, pero el significado idiosincrático para un hablante de español rioplatense es “hizo algo poco inteligente”. Para Aronoff, este último es el que se almacena en el diccionario de cada hablante. A partir de este planteo, entonces, se nos presenta la necesidad de pensar cómo solventar el conocimiento idiosincrático en un aprendiente de español L2 mediante la utilización de la herramienta *NooJ*.

Para operar a partir de frases propias del español rioplatense que comportan la cualidad de ser ambiguas en el análisis morfosintáctico, y teniendo como destinatarios a los aprendientes de español como L2, recabaremos la información de series televisivas, publicaciones en revistas, diarios y algunas novelas que focalizan en expresiones coloquiales. Luego, generaremos los ingresos pertinentes en las gramáticas del programa *NooJ* para aclarar las ambigüedades de dichas frases.

Mendivil Giró (2009) establece que estas construcciones se denominan **palabras con estructura externa** (PPEE). El autor explica que las PPEE son unidades que, según ciertas propiedades, pertenecen al eje paradigmático, es decir que son el resultado de una «única elección» en dicho eje, pero que formalmente representan en realidad «trozos» de estructura sintagmática, normalmente sintagmas o frases. En otros términos, una PPEE es un sintagma que funciona, a ciertos efectos semánticos y formales, como una palabra única.

Para analizar construcciones del tipo “ojo de buey”, “cara de piedra”, “batir fruta” “chupar cirios”, “lamer botas” es necesario abordar una definición de locución, para esto nos serviremos de las palabras de Julio Casares:

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 sus componentes (Casares, 1950: 170).

Cuando Casares dice “combinación de dos o más términos” alude al carácter sintagmático de las locuciones. Además, la combinación es de carácter estable. Esta fijación se debe entender como una limitación en la configuración formal de la locución.

Para entender esto pensemos en la expresión “meter la pata” o “ponerse la gorra”: el objeto directo de meter y poner es siempre singular, no se dice “meter las patas” o “ponerse las gorras”. Con esto es necesario decir que la fijación también atañe al potencial combinatorio de los elementos que la integran, pues no se dice “\*introducir la pierna” o “\*ponerse el sombrero”. En el último caso, si se opta por un valor idiomático más culto como “colocarse el sombrero”, la frase pierde su valor peyorativo<sup>1</sup> para volverse positiva. Lo mismo sucede con la inserción de modificadores adjetivales o adverbiales a los núcleos nominales y verbales: \*meter la pata izquierda \*ponerse la gorra lentamente. Esto parece mostrar que la estructura interna de dichas construcciones no es plausible a la configuración de las reglas sintácticas, es una propiedad que atenúa el carácter sintagmático de estas unidades y las aproxima a la noción de palabra.

No obstante, este tipo de locuciones no se comportan exactamente como palabras en lo que respecta a ciertos procesos morfológicos y fonológicos. Por otra parte, no son el producto de una suma del significado normal de sus componentes. Es decir, dicho significado es idiomático y no composicional. Esto se refiere al hecho notable de que el significado de, por ejemplo, las locuciones verbales “tomar el pelo a alguien” ('burlarse de alguien') “batir fruta” ('decir cosas casi sin sentido') o de la locución nominal “talón de Aquiles” ('punto débil') no es una función del significado de sus componentes computada a través de las relaciones formales entre ellos, como sería esperable en un sintagma libre o como sucede en la versión literal de nuestras locuciones ('agarrar el cabello de alguien', 'revolver frutas con un batidor' o 'parte posterior del pie de cierto héroe griego'), sino que es un significado relativamente arbitrario, ya dado, que se aprende de memoria. Esta es la característica más típicamente léxica de las locuciones.

Casares también observa que dicha combinación estable de palabras funciona como elemento oracional. Así, “estirar la pata” parece funcionar como el verbo principal de una oración de sentido intransitivo (Juan estiró la pata ayer) y “ojo de buey” parece ser el núcleo de un sintagma nominal término de preposición en la siguiente construcción: se colaron por los ojos de buey del barco. En este punto debe observarse que formalmente, en ambos casos, estamos ante sintagmas y no ante palabras. En cuanto a la flexión verbal, en el primer caso, y la flexión de número, en el segundo, no van al final

---

<sup>1</sup>El significado peyorativo de la frase “ponerse la gorra” tiene relación directa con la historia política de países que han pasado por procesos dictatoriales.

de la expresión (como sería de esperar si fueran auténticas palabras: \*la-pata-estirar, \*ojo-de-bueyes), sino en sus núcleos formales. Esto es, aunque a efectos semánticos y de ciertos procesos sintácticos, las locuciones se tratan en bloque como si fueran palabras, a efectos de otros procesos morfológicos y fonológicos se tratan como auténticos sintagmas.

Así pues, la afirmación de que las PPEE están a medio camino entre la palabra y el sintagma significa que para ciertos componentes de la gramática mental de los hablantes las locuciones cualifican como palabras, mientras que para otros lo hacen como sintagmas. Determinar si son lo uno o lo otro se convierte entonces, como dice Giró, en una tarea hasta cierto punto arbitraria y que obliga a especificar qué tipo de componentes de la gramática están implicados y qué jerarquía vamos a establecer entre ellos para hacer una elección, tarea que se verá reflejada en la selección clasificatoria que se hará en la programación de *Nooj* en el próximo párrafo.

Entonces, para realizar las entradas al programa resaltaremos que una característica esencial de las locuciones verbales, como ejemplos prototípicos de PPEE, es que tienen totalmente limitado su potencial combinatorio. Así, la diferencia esencial que se puede evidenciar entre un sintagma libre y un sintagma fijado es precisamente que, en este último, los elementos son fijos y no producto de la combinación espontánea por parte del hablante, que tampoco puede alterarlos sin destruir el sentido idiomático.

Las locuciones representan una limitación inesperada y drástica de la libertad combinatoria, en el sentido de que ciertos significados solo se obtienen con determinadas combinaciones fijas. Sin embargo, no son únicamente excepciones que empañan esa teórica libertad combinatoria irrestricta, más bien tendríamos que verlas como casos extremos de toda una continuidad de restricciones combinatorias que afectan, normalmente de manera inconsciente, a la capacidad de elección de elementos paradigmáticos por parte del hablante. Consideramos que esta situación debe ser abordada y trabajada por los docentes de español como L2.

## Propuesta de abordaje de las PPEE con el programa *NooJ*

El programa *NooJ* (Silberztein, 2016) trabaja sobre unidades léxicas individualmente. Si tenemos esto en cuenta, percibimos que resulta imposible especificar las variaciones de las PPEE a través del editor gramatical. Esto se debe a que, generalmente, en frases como “ponerse la gorra” el único ítem que puede derivarse es el verbo (“se puso la gorra”, “ponete la gorra”, “se pondrá la gorra”) y no el resto de sus componentes (“\*se pusieron las gorras”, “\*me pongo el gorro”). Por lo tanto, las derivaciones de estas PPEE no pueden indicarse a través de la extracción de los últimos morfemas que las componen, como sucede, por ejemplo, con los verbos regulares (si extraemos los dos últimos caracteres de “trabajar” y añadimos -o, formaremos la 1ra persona singular del presente)

Para diferenciar las PPEE y facilitar la tarea de la enseñanza del español, propondremos, entonces, establecer etiquetas que indiquen persona, número, tiempo y modo, en el caso de las verbales, y género y número en el caso de las nominales. Además, clasificaremos estas construcciones de acuerdo a su estructura, para que, cuando aparezcan en un texto, el análisis exponga su formación interna (llamaremos PPEE1 a aquellas frases de tipo verbal formadas por V+N y PPEE2 a las que contengan V+DET+N; asimismo, nombraremos PPEE3 a las construcciones con significado nominal).

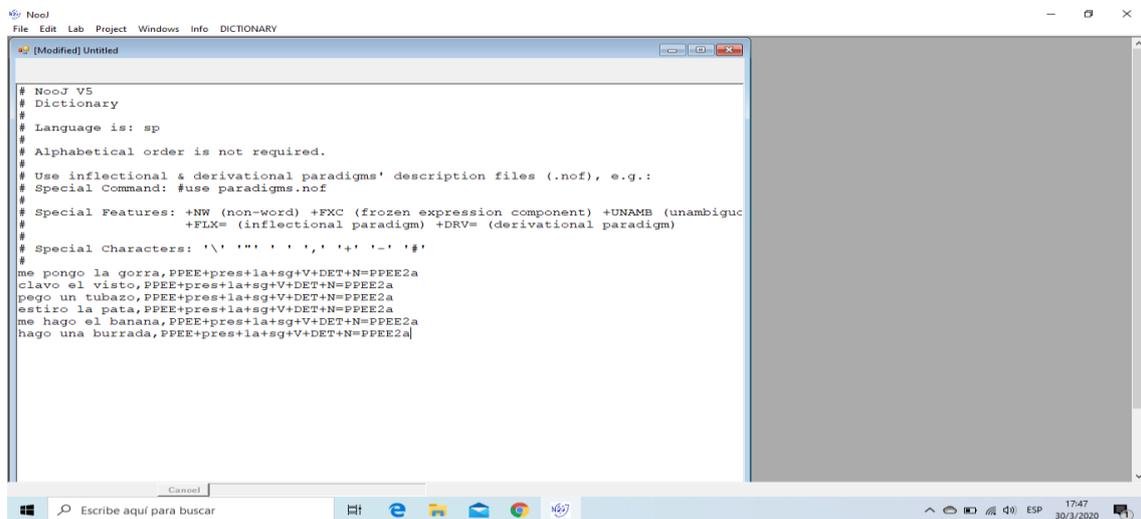
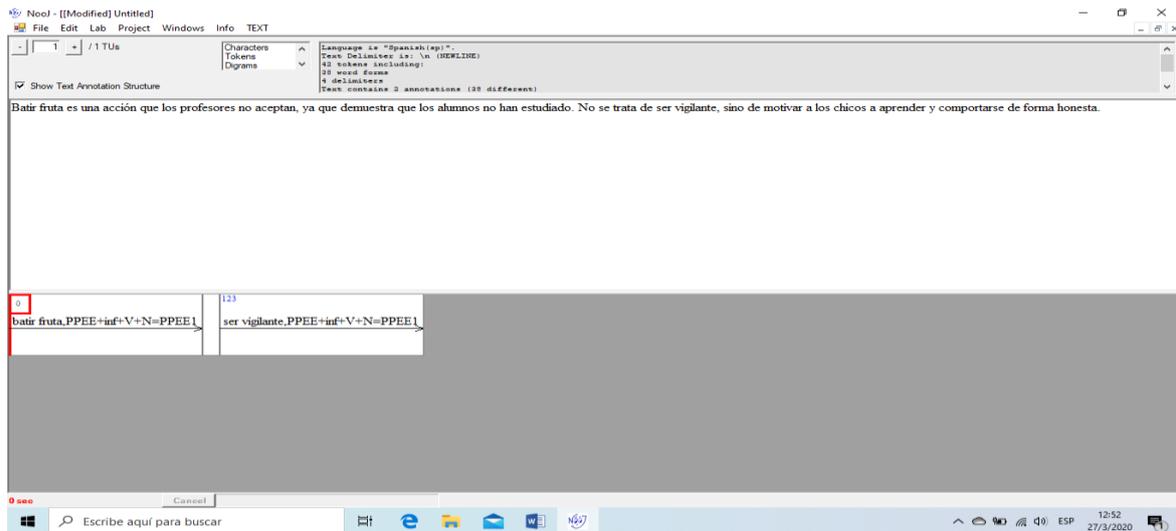
Nuestra propuesta consiste en crear diccionarios con las correspondientes gramáticas de todas las PPEE agrupadas y etiquetadas de acuerdo a sus variaciones morfológicas (PPEE2, por ejemplo, agrupará a todas las construcciones con sentido verbal formadas por V+DET+N en infinitivo (“estirar la pata”, “clavar el visto”); PPEE2a, por su parte, incluirá a todas las que estén flexionadas en 1ra persona singular del presente MI (“estiró la pata”, “clavó el visto”). Cabe aclarar que trabajamos con el módulo español sp (Argentina) (Tramallino, 2013; Solana, Beltrán y Tramallino, 2013) disponible para descargar de forma gratuita<sup>2</sup>.

A continuación, presentamos la selección de construcciones ingresada al programa, apostando a la progresiva inclusión de todas las ocurrencias del español rioplatense y al perfeccionamiento de la propuesta para obtener resultados más claros.

---

<sup>2</sup> Disponible en





En las PPEE2con estructura V+DET+N

Hacer una burrada > hacer algo poco inteligente

Pegar un tubazo> llamar



## Consideraciones Finales

Este trabajo se inaugura con la afirmación de que casi ningún hablante adulto de español podrá siquiera acercarse a una explicación certera de cómo se clasifican las palabras con estructura externa. Se afirma también, que este mismo hablante las utiliza de forma adecuada.

Tal como se ha intentado dar cuenta en el recorrido de este estudio, las frases con estructura externa son innumerables en el español rioplatense y su desarrollo crece conforme el paso del tiempo, el desarrollo tecnológico y, también, gracias a las nuevas generaciones. Sin embargo, contrariamente a lo que puede pensarse, estas frases tienen estructuras complejas que, en la mayoría de los casos, no permiten cambios de género, número o tiempo.

Si revisamos las discusiones en torno a esta temática que se han expuesto en el primer apartado y si pensamos en los aprendientes de español como lengua segunda y extranjera, las PPEE se vuelven un desafío importante en el aula de ELSE (Tramallino, 2021). Incluso, estas dificultades se han vuelto un espacio de diversión para quienes habitan las redes sociales, ya que los mismos estudiantes generan contenido ironizando sobre el uso de dichas estructuras en situaciones cotidianas.

Por sus características gramaticales, las PPEE también resultan un reto para el trabajo con herramientas y procesadores de textos como el sistema *NooJ*, ya que las posibilidades de clasificación de palabras que permite el programa no son suficientes en este caso. Por lo tanto, se propone la creación de diccionarios específicos dentro de *NooJ* para la clasificación de las PPEE. A partir de la creación de diccionarios en el *software*, podemos concluir que el programa se vuelve una herramienta imprescindible para la interpretación y utilización de las PPEE por parte de estudiantes de español.

## Bibliografía

- Aronoff, M. (1976). *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.
- Casares, J. (1992). *Introducción a la lexicografía moderna*. Consejo Superior de Investigaciones Científicas.
- Fillmore, C. (1970). *The grammar of hitting and breaking* en RA Jacobs y PA Rosenbaum (Eds.), *Readings in English transformational grammar* (págs. 120-133). Waltham, MA: Ginn.
- Mendivil Giró, J. (2009). Palabras con estructura externa en *Panorama de la lexicología* (págs. 83-112). Ariel Letras.

Solana, Z., Beltrán, C., Tramallino, C. (2013). “Análisis automático morfológico con las herramientas Smorph y Nooj”. *Revista de Epistemología y Ciencias Humanas*. Grupo IANUS.

Silberztein, M. (2016). *Formalizing Natural Languages: The NooJ Approach*. ISTE

Tramallino, C. (2013). Análisis morfológico con herramientas informáticas. Reconocimientos de nombres en textos de español con el sistema NooJ. *Revista Lingüística y Literatura*, (63), 33-48.

Tramallino, C. P. (2021). Avances en el tratamiento computacional en corpus de aprendientes de español como lengua segunda y extranjera. *Quintú Quimün. Revista de Lingüística*, 5, 051.