

Presentación de la Base de Datos del Observatorio lingüístico del español argentino actual

Camila Selene Hernández

UNIVERSIDAD NACIONAL DE RÍO NEGRO

camila97hernandez@gmail.com

Miguel B. de Vanssay

UNIVERSIDAD NACIONAL DE RÍO NEGRO

migueldevanssay@gmail.com

Resumen

El estudio del español argentino, considerado como variedad o conjunto de variedades del español hablado en nuestro país, tiene una larga tradición que se remonta a fines del siglo XIX, la cual cuenta con importantes investigaciones a partir de diversas teorías y ramas de la lingüística. Esta rica tradición no corre el riesgo de agotarse, pues el habla coloquial es dinámica y siempre incorpora nuevos fenómenos, cuyo estudio ofrece una doble recompensa: esclarecer el funcionamiento del sistema lingüístico, así como permitir registrar el habla cotidiana en uso.

A este fin, el actual trabajo se propone como objetivo la presentación de la base de datos elaborada en el marco del Proyecto de Investigación del *Observatorio lingüístico del español argentino actual*, dirigido por el Dr. Pascual José Masullo en la Universidad Nacional de Río Negro, Sede Andina. Caracterizaremos dicha base de datos en términos de los tipos de datos recolectados, los criterios de clasificación de los mismos (no siempre convencionales), la metodología de recopilación y fuentes utilizadas, entre otros.

Asimismo, ilustraremos el valor que esta base de datos posee para las investigaciones lingüísticas en general y, especialmente, para aquellas centradas en el español argentino contemporáneo. En función de este propósito,

destacaremos sus diferencias con otras bases de datos lingüísticos, tales como CREA y PRESEEA.

Por último, ejemplificaremos su utilidad para la investigación desarrollando un breve análisis de algunos fenómenos registrados en la misma, en particular nuevas metáforas conceptuales, así como secuencias de nombre modificado por nombre del tipo *voto castigo*. En ambos casos intentaremos demostrar la relevancia sociocultural de estos fenómenos en nuestra vida cotidiana.

1. Introducción

El objetivo de este trabajo es presentar la Base de Datos, confeccionada en el marco del *Observatorio lingüístico del español argentino actual*, proyecto de investigación de la Universidad Nacional de Río Negro, dirigido por el Dr. Pascual José Masullo, e integrado por diversos estudiantes del Profesorado en Lengua y Literatura y la Licenciatura en Letras. Esta base es el resultado de la observación y el registro sistemático y metódico de los usos que dan cuenta de fenómenos significativos en el seno del español argentino contemporáneo.

El proyecto de investigación se puede anclar en la extensa tradición lexicológica y especialmente lexicográfica que tiene la Argentina. En nuestro país, es y siempre ha sido muy fuerte el interés por todo aquello que define al español argentino en sí, esto es, como una variedad o conjunto de variedades que se diferencian de aquellas otras habladas en el resto de Hispanoamérica (Resnik, 2014). En esta línea, podemos mencionar al *Vocabulario rioplatense razonado*, de Daniel Granada, de 1889; el *Diccionario Argentino* de 1910, de Tobías Garzón, y el *Diccionario de argentinismos de ayer y de hoy*, de Diego Abad de Santillan, de 1976. En tiempos más recientes nos encontramos con obras como el *Nuevo diccionario de argentinismos* de 1993 dirigido por Haensch y Werner; el *Diccionario integral del español de la argentina* del 2009; o el *Diccionario del habla de los argentinos* del 2003 de la Academia Argentina de Letras. Este último fue actualizado en el 2019 en una nueva edición: *Diccionario de la lengua de los argentinos*, presentado en el *VIII Congreso Internacional de la Lengua Española*

realizado en Córdoba en mayo del 2019, y en el *I Congreso Nacional del Español Argentino*, realizado en San Carlos de Bariloche en noviembre de 2019.

Sin embargo, la Base de Datos del *Observatorio* nace no sólo de un interés lexicográfico, como la gran mayoría de estos diccionarios, sino también lexicológico. Esto significa que, si bien hay un registro y documentación de usos propios del español argentino, también hay una necesidad de ir más allá de la mera compilación para permitirnos a los investigadores adentrarnos en la teoría lingüística propiamente dicha. En otras palabras, la misma se establece como un soporte sobre el cual se busca describir, analizar y explicar cómo estos nuevos usos ilustran el funcionamiento del sistema lingüístico, y cómo evidencian procesos de cambio en el seno del mismo. A este fin, los datos recolectados se confrontan con diferentes aportes teóricos contemporáneos, lo cual da pie a precisar, confirmar, modificar, o proponer otros.

2. Origen y características

La Base de Datos tiene sus orígenes en la década de los noventa, cuando el ahora director de este proyecto, Pascual José Masullo, a partir de un número determinado de intereses teóricos, como las construcciones anticausativas y los sujetos caprichosos, empezó a hacer registros en papel de su uso y aparición en el habla. Estos primeros registros obedecían a la necesidad de confirmar diferentes hipótesis de carácter explicativo.

Cuando se funda la Universidad Nacional de Río Negro en 2009 y Masullo se incorpora al plantel docente, estos registros se reanudaron con más intensidad y nuevas categorías. El número de descriptores también fue aumentando, lo que implica que había cada vez más registros y listas de fenómenos de carácter léxico, morfológico, semántico, pragmático, entre otros.

En el 2010 la tarea de registros se consolida en un proyecto de investigación titulado *Creación de una Base de Datos Audiovisual para la documentación de cambios recientes en el español de la Argentina*, pero es recién a partir de la creación del *Observatorio Lingüístico del Español Argentino Actual* en el 2017

que todos los datos y registros comenzaron a consolidarse en una única base ordenada y clasificada.

Actualmente, la base cuenta con más de diez mil datos. Estos consisten no sólo en ítems léxicos; de hecho, la gran mayoría son frases que evidencian ciertos tipos particulares de construcciones. Los datos están ordenados en alrededor de ciento diez clasificaciones, correspondientes con distintos fenómenos lingüísticos, como son predicados complejos, conectores livianos, preposiciones complejas, desemantización, arcaísmos, verbos ligeros, falsos cognados, expresiones idiomáticas, locuciones preposicionales, nuevas verbalizaciones, entre muchos otros. Estas categorías de fenómenos, a su vez, se ordenan en los siguientes niveles de estudio: léxico, morfología, sintaxis, semántica, y pragmática. Los datos son tomados de todo tipo de fuente oral y escrita, poniendo foco en el habla cotidiana: la calle, el colectivo, la escuela, medios de comunicación, redes sociales, entre otros.

En este sentido, no se puede caracterizar a la base solamente como cuantitativa o cualitativa. Es cuantitativa no en un sentido estadístico, dado que no registra porcentajes de uso, pero sí en el sentido de la extensión, dado que pretende recopilar suficientes datos como para permitir la inferencia de hipótesis. Por otro lado, también es cualitativa ya que tiene como objetivo profundizar en detalle fenómenos concretos.

La clasificación de los datos en diferentes fenómenos, y estos a su vez en niveles, sigue un enfoque generativo. Los niveles (léxico, sintáctico, semántico, pragmático) se entienden en el sentido de componentes lingüísticos con interfaces entre sí, al interior del módulo lingüístico (Anula Rebollo, A. y Fernández Lagunilla, M., 1995). Por este motivo, la Base de Datos obedece a criterios formales en vez de funcionales. Factores contextuales como fecha de recolección, datos del hablante y reiteración de ciertos usos no son tomados en cuenta. Sin embargo, cabe señalar que en la misma sí se consideran y registran fenómenos estudiados por la pragmática, la lingüística cognitiva, teorías del cambio lingüístico, entre otros.

Es por estas características que la Base de Datos se diferencia de otros tipos de corpus, de gran referencia en el ámbito del español en general, como el

PRESEEA, que consiste en muestras de audios de carácter representativo de diferentes sectores demográficos, dando lugar al estudio de distintos lectos de la población de una región; o el CREA, que consiste en muestras representativas de las diferentes variedades del español a partir de conjuntos de textos orales y escritos, y que permite indagar sobre los significados de ciertas palabras, o el registro, contexto o variedad en las cuales son usadas. A diferencia de estos corpus, la Base de Datos del *Observatorio* no tiene un interés sociolingüístico o dialectológico, aunque cabe señalar que sí podría tener valor para estas disciplinas, dado que un investigador puede contrastar lo que esta base revela de las particularidades del español argentino con, por ejemplo, el español colombiano o peninsular.

La base no pretende hacer un registro integral de diferentes usos, sino uno lo suficientemente extenso o exhaustivo como para permitir la detección de patrones sistemáticos, pudiendo así el investigador inferir posibles hipótesis o teorías descriptivas y explicativas. Para ilustrar lo dicho, se hará un breve análisis de dos fenómenos a la luz de los registros aportados por la base. Estos fenómenos concretos fueron elegidos en función de nuestras líneas de interés e investigación particulares.

3. La Base de Datos en la investigación: El caso de N+N y de la Metáfora Conceptual

Los grupos nominales (N+N), entendidos como nombres que modifican a otros nombres, han sido estudiados hasta ahora como compuestos nominales. Val Álvaro (1999) los distingue en perfectos, cuyos constituyentes están fusionados en una sola palabra fonológica, como en *telaraña* y *aguanieve*; e imperfectos, donde sus componentes no están fonológicamente amalgamados, como en *efecto dominó* o *carta documento*. Según Rainer y Varela (1992), el orden particular en el que se consolidan estos últimos se da porque el segundo nombre es interpretado como cualidad o característica que se integra en el núcleo, o porque lo denota de tal forma que lo subclasifica. Es importante señalar que en muchos casos la relación semántica entre los nombres no está clara y debemos acudir a la pragmática o al conocimiento del mundo para poder establecerla. Es

lo que Reiner y Varela (1992) definen a partir de las nociones de endocéntrico (cuando el significado se deduce fácilmente de sus componentes, como en *papel moneda*) y exocéntrico (cuando es necesario recurrir a conocimiento del mundo, como en *baño maría*).

Sin embargo, y según otra línea teórica, los N+N pueden estudiarse como complejos en vez de como compuestos. Una hipótesis es, según Masullo (en curso), considerarlos como resultado de un proceso de fusión en el componente sintáctico del módulo lingüístico pero sin marcas gramaticales alguna. Se diferencian así de los llamados compuestos sintagmáticos como *ojo de buey* o *casa de comidas*, que, si bien también suceden en el componente sintáctico, se relacionan mediante una preposición de por medio; y de las palabras compuestas de nombres como *noroeste* y *arcoíris*, que ya han quedado fosilizados como listemas en el componente léxico. Si bien esta hipótesis todavía está en sus etapas iniciales de desarrollo, nos permite considerar a los N+N como una demostración de que en español sí es posible que un sustantivo modifique a otro, tal como ocurre en lenguas como el inglés, en grupos como *garden party*. Estos son muy comunes en nuestra lengua, por lo que hay muchos registrados en la Base de Datos, entre los que se encuentran *masa madre*, *modo avión*, y, en relación al contexto económico y sociopolítico actual, por ejemplo, otros como *riesgo país*, *pobreza cero*, *voto castigo*, *dólar futuro*, *dólar rulo*, *fondos buitres*, entre otros.

Otro fenómeno, cuyo estudio ha tenido mucha repercusión en las últimas décadas a partir del campo de la lingüística cognitiva, es el de la metáfora conceptual. Lakoff (1992) define a la metaforización, en tanto proceso cognitivo básico y esencial, como la conceptualización de un dominio mental, generalmente abstracto, en términos de otro más concreto. Esto implica que hay una direccionalidad dada por la proyección de un esquema cognitivo de un dominio fuente a otro destino. En nuestra base tenemos registradas expresiones metafóricas tales como la *tormenta financiera*, *clima de negocios*, y *clima inflacionario*.

Siguiendo la teoría de Lakoff, se puede considerar que la metáfora conceptual que subyace a aquellas expresiones es *economía es clima*. Un concepto tan

abstracto y complejo como el de *economía* carece de base experiencial, es decir, no es un concepto que emerge de nuestra experiencia corpórea, física, perceptible del mundo. Para cognitivistas como Lakoff, el pensamiento y lenguaje humano emergen de los patrones recurrentes de la percepción y acción corpórea (Gibbs, 2005).

Acorde con esta perspectiva, la forma en que conceptualizamos abstracciones como *economía* o *sistema económico* es mediante procesos de extensión metafórica, a partir de conceptos emergentes de la experiencia, como *clima*. La noción de *clima* emerge de un conjunto de percepciones y consideraciones que realizamos diariamente, las cuales se han refinado y extendido con saberes o información transmitida socialmente. Se lo conceptualiza como estados atmosféricos sucesivos. Los cambios de estado son generalmente regulares, lo que posibilita el pronóstico o predicción. Sin embargo, sabemos que también hay eventos climáticos extremos que causan desastres. Por último, el clima no es agentivo, por ende, carece de intencionalidad o volición: lo que sucede, sucede por causas naturales.

A través del proceso de metaforización, este esquema de nociones es usado para conceptualizar la economía, la cual entendemos como estados del mercado, y sujeta a cambios generalmente regulares. Por ejemplo: la inflación se predice, y algunos economistas, cuales meteorólogos, pronostican crisis. Facilita, además, concebir la economía como una suerte de fuerza natural, lo cual ciertamente invita a reflexiones desde el prisma de Análisis del Discurso, dado que desdibuja responsabilidades e invisibiliza intereses.

Esta conceptualización de la economía en términos climáticos también está detrás de expresiones metafóricas más concretas que las ya citadas, como *lluvia de inversiones/dólares/pesos*, *sequía de inversiones*, *turbulencia económica*, entre otras. Siguiendo a Lakoff y Johnson (1986: 92) podemos entender a estas metáforas como más específicas, haciendo foco en aspectos más concretos del dominio origen *clima*, tales como la lluvia, la sequía, la turbulencia, etcétera. Sin embargo, continúan participando del sistema metafórico más amplio señalado anteriormente: *economía es clima*.

4. Conclusión

El valor de la Base de Datos del Observatorio radica en que permite no sólo estudiar los aspectos más intrínsecos del funcionamiento del sistema lingüístico, sino también caracterizar el español argentino atendiendo a su existencia cotidiana y concreta. La diversidad de posibilidades de estudio y de líneas de investigación es enorme.

Dado que dispone de los datos ya ordenados, clasificados, y procesados, se conforma como una herramienta fácilmente predispuesta a los fines de cualquier investigación lingüística, ya sea para el estudio específico de fenómenos morfosintácticos (como es el caso de N+N), pragmática, semántica, etc.

Uno de los actuales objetivos del proyecto de investigación es disponer de la Base de Datos en un soporte on-line, interactivo y de fácil acceso para que cualquier investigador o interesado en el tema pueda acceder, buscar, y además contribuir con datos que disponga y que no estén previamente consignados. En suma, que la base esté disponible en una plataforma virtual de libre acceso. Se espera que esto sea posible en el 2020.

5. Bibliografía

Anula Rebollo, A. y Fernández Lagunilla, M. 1995. *Sintaxis y cognición. Introducción al conocimiento, el procesamiento y los déficits sintácticos*. Madrid: Editorial Síntesis.

Gibbs, R. 2005. *Embodiment and Cognitive Science*. Cambridge: Cambridge University Press.

Lakoff, G., y Johnson, M. (1986). *Metáforas de la vida cotidiana*. Madrid: Ediciones Cátedra.

Lakoff, G. (1992). The Contemporary Theory of Metaphor. En Ortony, A. (Ed.), *Metaphor and Thought*, Cambridge: Cambridge University Press.

Rainer, F., y Varela, S. (1992). Compounding in Spanish. *Rivista di Linguistica*, 4 (1), 117-142.

Resnik, G. (2014). Norma y variación lingüística en los diccionarios del español de la Argentina. En Kornfeld, L. (Comp.), *De lenguas, ficciones y patrias*. Los Polvorines: Universidad Nacional de General Sarmiento

Val Álvaro, J. F. (1999). La Composición. En Bosque, I. y Demonte, V. (Eds.), *Gramática descriptiva de la lengua española*. Madrid, Espasa-Calpe.