

# Análisis de Discursos Presidenciales con LSA

Marcos Pietto \*, Sergio Romano \*\* and Noelia Rugna \*\*\*

\* CONICET-UNC, \*\* CONICET-UBA, and \*\*\* UBA

Trabajo Práctico del Seminario Introducción a la Neurociencia Computacional

En este trabajo utilizamos “Latent semantic analysis” (LSA) para caracterizar discursos de distintos presidentes de latinoamérica a lo largo de la historia. En ellos buscamos conceptos que nos permitan distinguir elementos típicos de la construcción del “discurso populista” para evaluar la efectividad de LSA como herramienta para el análisis de discursos y para verificar la importancia del eje discursivo en el populismo. Si bien el dialecto de la época de cada discurso, su país de origen y las particularidades de la construcción política de cada contexto hacen difícil su uso efectivo, obtuvimos evidencias que muestran la utilidad del método para analizar conceptos antagónicos típicos en la construcción populista a través de los discursos presidenciales.

lsa | discursos presidenciales | populismo

LSA es una técnica de procesamiento del lenguaje natural para analizar la relación entre un conjunto de documentos y las palabras que lo componen, mediante la creación de una matriz que representa un espacio semántico vectorial en el que los documentos y sus palabras están representados por medio de vectores que contienen sólo la *información principal* para la formación de conceptos [1, 2].

Este trabajo busca experimentar sobre la efectividad de esta técnica para el análisis de discursos presidenciales, con el fin de que pueda utilizarse para obtener métricas cuantitativas sobre distintos conceptos de las ciencias políticas tales como: hegemonía, liberalismo, populismo, etc. En particular este estudio pretende analizar el concepto de populismo.

El populismo es parte constitutiva de la política y de la democracia en América Latina. Durante muchos años, la literatura lo utilizó como un término peyorativo para designar a una vasta variedad de experiencias políticas mundiales, caracterizándolo por su vaguedad, irracionalidad y vacío ideológico [3–7]. Sin embargo, Laclau [8–10] introduce un cambio en esta visión planteando que el populismo no es una ideología de contenido específico, sino un método de construcción política que cohesiona una variedad de demandas y logra unificarlas en torno a una identidad popular común (un significante vacío que condensa las demandas) y a una frontera antagónica que divide al campo social. De aquí, es entonces que al análisis del discurso se vuelve central para entender y caracterizar al populismo.

Por tanto, proponemos una metodología de análisis de contenido a través de LSA sobre los discursos presidenciales dada su importancia en la construcción populista, y examinaremos métricas cuantitativas sobre dimensiones identificadas como propias de un discurso populista para evaluar la efectividad de la técnica de LSA para el ámbito de conocimiento de las ciencias políticas.

Particularmente, nos centraremos en los discursos de apertura de congreso que los líderes de gobierno realizan cada año de su mandato ya que estos suelen ser transcritos y además permiten un análisis comparativo en contextos políticos similares. Debido a que no sería apropiado comparar discursos de diferentes contextos como ser, un discurso de inauguración, de una expropiación, o de una cena con empresarios, pues se tendría un sesgo mayor en el contenido del discurso.

## Método

**Generación de LSA.** Para generar la matriz LSA el primer paso fue descargar el corpus entero de los artículos de la Wikipedia en Español, convertir los artículos a texto plano (quitar el formato de la Wikipedia) y crear el diccionario de palabras y la matriz que contiene la cantidad de apariciones de cada palabra en cada artículo o documento.

El siguiente paso es someter esta matriz a un algoritmo llamado Descomposición del Valor Singular (SVD) para reducir el número de dimensiones sin que se pierda la información sustancial de la matriz original. En nuestro caso, redujimos la matriz en 600, 400 y 200 dimensiones (ya que un valor entre 200 y 500 es el recomendado como un buen estándar [11]).

El código para generar la matriz de LSA se encuentra disponible en *generadorLSA.py* y utiliza la librería *gensim* [12].

**Análisis de textos a partir de LSA.** A partir de la matriz LSA construida sobre el corpus de la Wikipedia en Español, hicimos un primer análisis de cómo funciona para categorizar distintos conceptos / palabras, y luego comparamos en cada corpus de discursos tanto la distancia entre los demás corpus como la distancia a un concepto seleccionado. Para eso, transformamos el concepto y el corpus de los discursos al espacio vectorial de la matriz LSA, y luego medimos el grado de similitud mediante el coseno entre los dos vectores. Valores cercanos a 1 representan una gran similitud, valores cercanos a -1 que son muy disímiles, y valores cercanos a 0 que no hay correlación. Las funciones auxiliares utilizadas en el análisis pueden consultarse en *analizador.py*.

El análisis comparativo entre cada corpus de discursos presidenciales consistió en analizar dos dimensiones pertenecientes al *discurso populista*. La construcción de un significante vacío, que por simpleza y tradición identificamos con el concepto de **pueblo** y la construcción de una frontera antagónica que asociamos al concepto de **enemigo**. Esto nos permitió obtener un índice de acercamiento de los discursos hacia fenómenos que suelen caracterizar al populismo: el nivel *antagónico* del discurso y el hecho de apelar a una identidad popular común.

Para este análisis utilizamos sólo los discursos de líderes de gobierno con mandato presidencial ocurrido en las últimas tres décadas (Raúl Alfonsín, Michelle Bachelet, Hugo Chávez, Rafael Correa, Fernando de la Rúa, Cristina Fernández de Kirchner, Néstor Kirchner, Ricardo Lagos, Mauricio Macri, Nicolás Maduro, Carlos Menem, Evo Morales, Augusto Pinochet, Tabaré Vázquez) y calculamos para cada uno su distancia a los conceptos de **enemigo** y **pueblo**. A fin de obtener un cálculo de la distancia de cada presidente respecto al promedio del grupo entero, convertimos los puntajes medios LSA de cada presidente según condición en puntajes Z.

Una vez efectuado dicho estudio, comparamos el valor LSA medio de los presidentes que superaban 1 desvío estándar respecto a la media ( $z\text{-score} > 1$ ; casos únicos) con la muestra de presidentes por abajo de 1 desvío estándar ( $z\text{-score} < 1$ ; grupo control). Para ello, utilizamos un t-test modificado [13–17] que permite evaluar la significancia a través de la comparación de múltiples puntuaciones individuales con valores normativos derivados de pequeñas muestras (5 sujetos). Este test estadístico resulta más robusto que las distribuciones

no-normales. Este controla efectivamente errores de tipo I y demuestra robustez respecto a otros métodos [15].

Por último realizamos un análisis comparativo entre las distancias obtenidas en los conceptos seleccionados para los diferentes valores de dimensiones de LSA (200, 400 y 600). Para ello comparamos los valores de la distancia de los discursos de cada presidente respecto a *enemigo* y *enemigo* a lo largo de las dimensiones de LSA, a través de la de la prueba no-paramétrica Kruskal-Wallis [18]. Luego, realizamos comparaciones múltiples entre pares de dimensiones utilizando el test no-paramétrico Mann-Whitney [19] y corregimos los p-values obtenidos con el método de Holm-Bonferroni [20].

**Fuentes de datos.** El LSA fue generado a partir del último dump disponible de la Wikipedia en Español al día 2 de junio de 2016. Los discursos presidenciales son todos correspondientes a la apertura de sesiones legislativas con el fin de minimizar que los discursos traten sobre coyunturas particulares de cada país.

## Resultados

**Estructura LSA.** Un primer objetivo es evaluar si la matriz LSA obtenida a partir de los artículos de la Wikipedia y con la reducción de dimensionalidad propia de la técnica, mantiene la información necesaria para encontrar relaciones entre conceptos y, en particular, entre los discursos presidenciales.

Para esto, convertiremos palabras de distintas categorías al espacio LSA y analizaremos las distancias entre ellas, como se pueden ver en las figuras 1, 2, 3, 4, 5 y 6 de escalamiento multidimensional que se encuentran a continuación.

En las figuras 1, 2 y 3 pueden observarse las palabras correspondientes a las categorías de **frutas** y **animales** para las matrices de LSA de dimensiones 200, 300 y 400 respectivamente. En rojo pueden observarse las palabras correspondientes a la categoría **Frutas**: Arándano, Aguacate, Banana, Coco, Ciruela, Cereza, Damasco, Durazno, Frambuesa, Fresa, Grosella, Higo, Kiwi, Limón, Mandarina, Mango, Manzana, Melón, Naranja, Pera, Piña, Pomelo, Sandía, Tomate, Uva. En azul las palabras correspondientes a la categoría **Animales**: Abeja, Burro, Caballo, Elefante, Foca, Gato, Gallo, Jirafa, Kiwi, Koala, Liebre, Mariposa, Mosca, Oso, Oveja, Perro, Paloma, Rana, Serpiente, Sapo, Tiburón, Tortuga, Vaca, Zorro, Zebra.

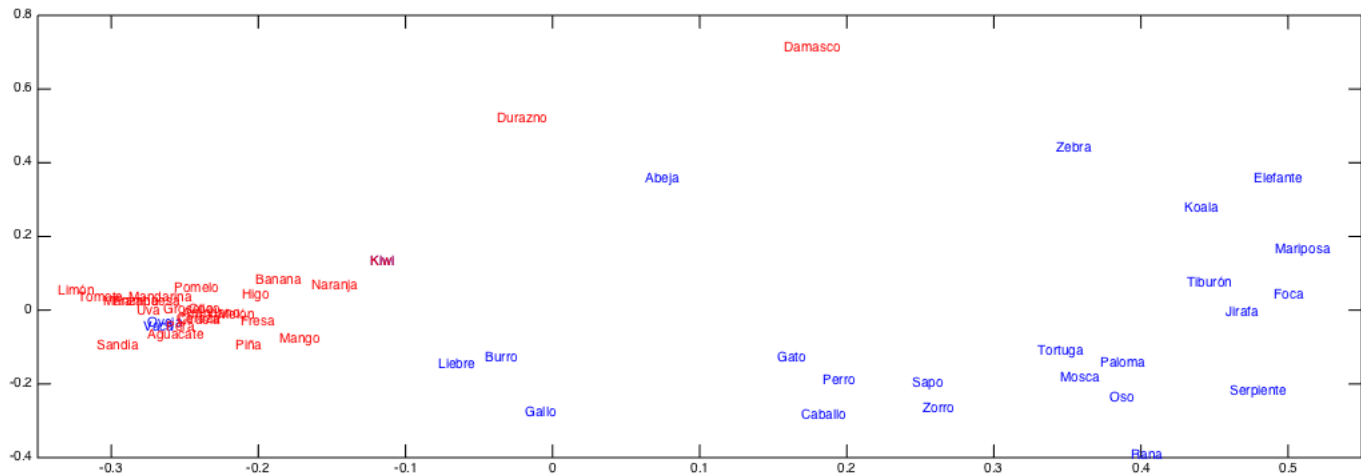


Figura 1: Escalamiento multidimensional de distancia entre conceptos para LSA de 200 dimensiones

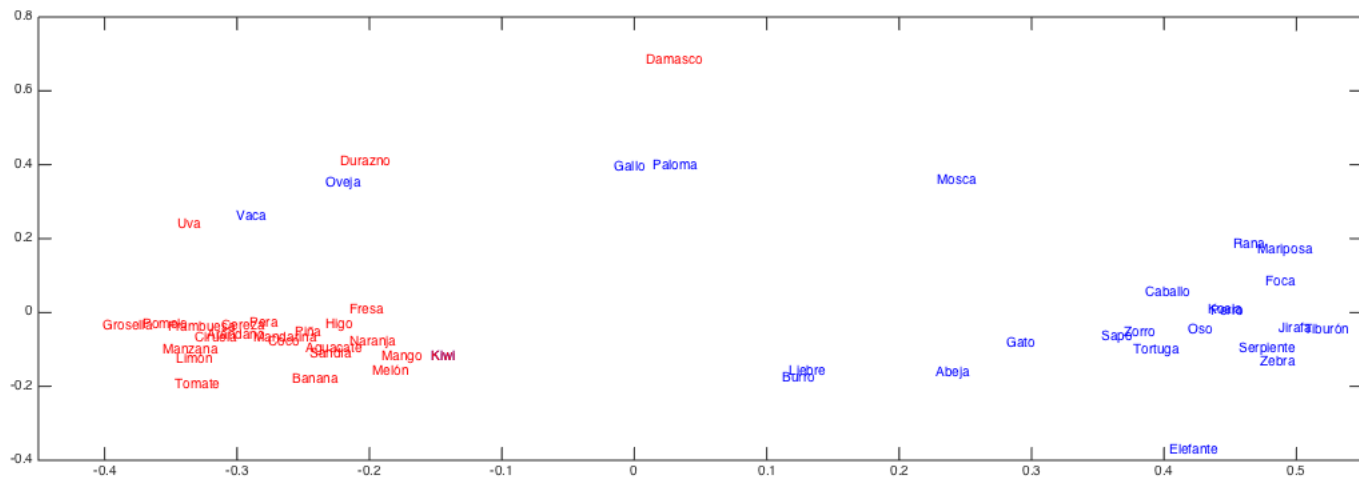


Figura 2: Escalamiento multidimensional de distancia entre conceptos para LSA de 400 dimensiones

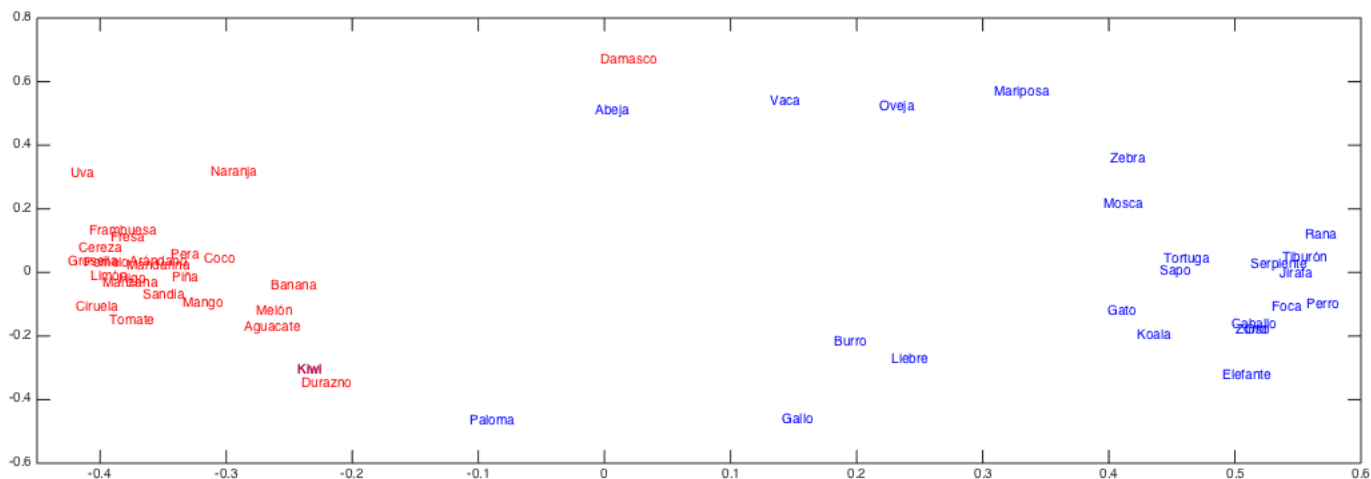


Figura 3: Escalamiento multidimensional de distancia entre conceptos para LSA de 600 dimensiones

Puede observarse en las tres primeras figuras (sobre todo en la figura 1 y 2) que **vaca** y **oveja** se encuentran más cerca de los conceptos de frutas que de animales. Es interesante esto ya que esos animales están también más asociados a la alimen-

tación de los seres humanos que el resto. Algo similar ocurre con Abeja, Gallo y Paloma que sin estar cerca de las frutas necesariamente se encuentran a una distancia intermedia entre estas y el resto de los animales.

En las figuras 4, 5 y 6 presentamos el mismo análisis para categorías más relacionadas con nuestro estudio. En color negro, aparecen las palabras de la categoría **política**: libertad, socialismo, comunismo, populismo, liberalismo, anarquismo, fascismo, igualdad, solidaridad, fraternidad, derechos, gobierno, democracia. En azul, las palabras de **religión**: pueblo, cristianismo, judaísmo, islamismo, budismo, hinduismo, dios, demonio, diablo, fé, religión. En rojo, las palabras de **economía**: mercado, empresa, privado, moneda, bonos, inversión, oferta, demanda.

Nuevamente, el agrupamiento es el esperado y bastante estable entre los tres gráficos con la salvedad de la palabra **pueblo**, que en la figura 6 puede verse cerca de **gobierno** en vez de los conceptos religiosos (lo cual es algo esperable ya que es una idea que aparece mucho en la religión como también en la política y en la conformación de las naciones). También es interesante en la figura 5 como **gobierno** aparece en un punto intermedio entre los conceptos económicos y políticos. Y como **fraternidad** en las tres figuras, aparece bastante cercano a la religión y no sólo cerca de la categoría política ya que es un concepto también muy presente de la hermandad religiosa.

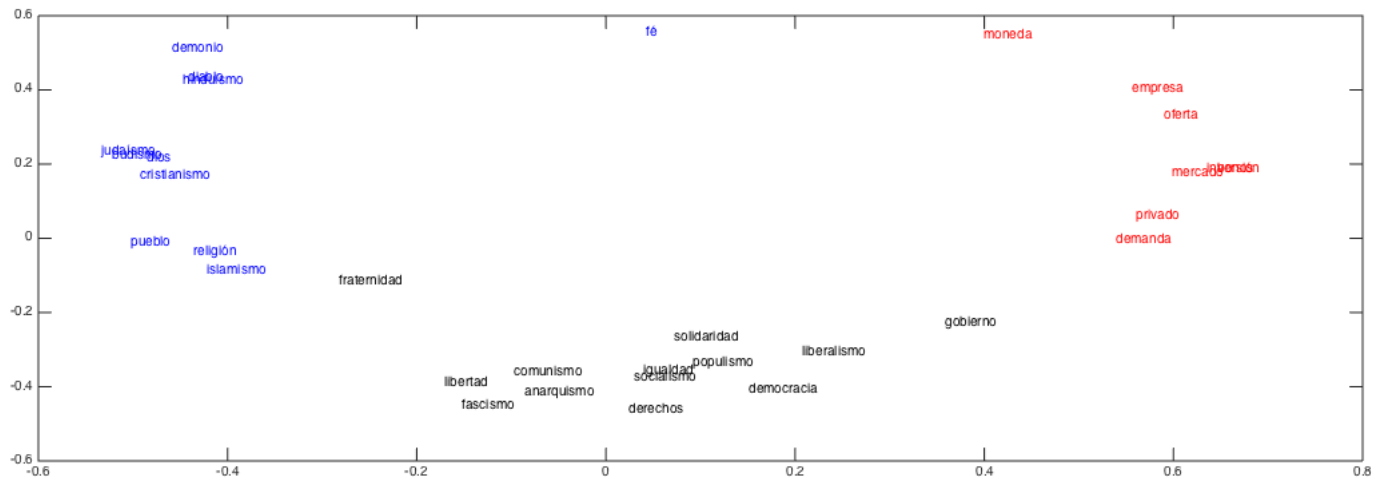


Figura 4: Escalamiento multidimensional de distancia entre conceptos para LSA de 200 dimensiones

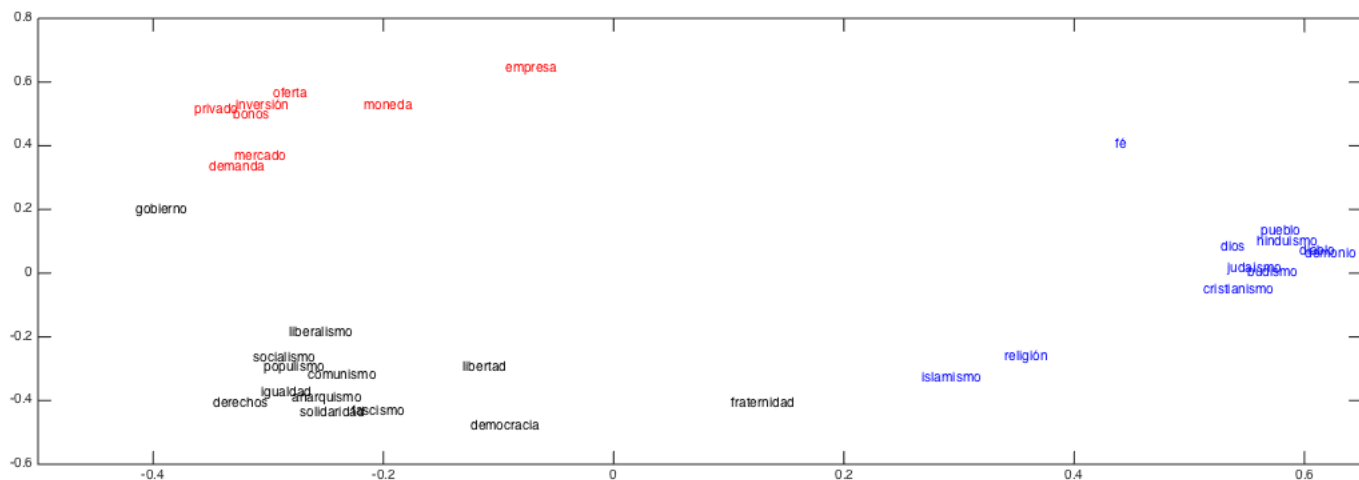


Figura 5: Escalamiento multidimensional de distancia entre conceptos para LSA de 400 dimensiones

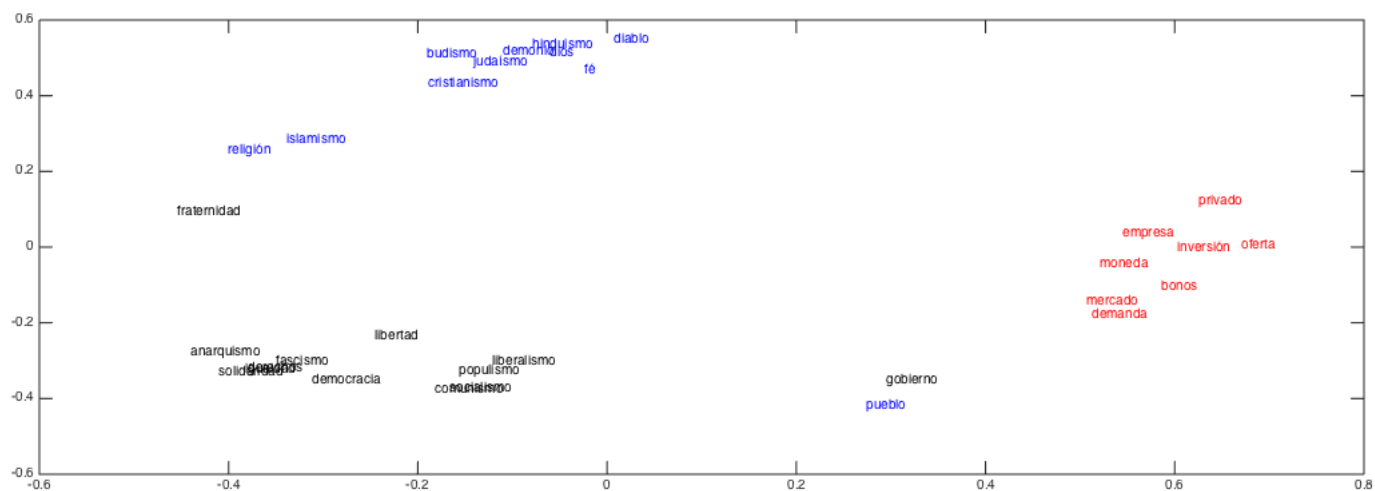


Figura 6: Escalamiento multidimensional de distancia entre conceptos para LSA de 600 dimensiones

**Relación directa entre discursos.** En la figura 7 realizamos una relación entre los discursos de presidencia con los del resto en el espacio vectorial de la matriz LSA de 400 dimensiones.

En la relación entre ellos podemos ver cómo los discursos se relacionan no simplemente por el país al que pertenecen, sino también por época. De esta manera, los discursos de Juan Manuel de Rosas, Justo José de Urquiza, Bartolomé Mitre y Domingo Faustino Sarmiento guardan una mayor relación entre sí, que los de Raúl Alfonsín, Carlos Saul Menem, Fernando De la Rúa, Eduardo Duhalde, Néstor Kirchner y Cristina Fernández de Kirchner, pese a que todos pertenecen a la República Argentina.

Por otro lado, también se destaca claramente en la figura 7 los cluster de presidentes chilenos (Salvador Allende, Augusto Pinochet\*, Eduardo Frei, Ricardo Lagos, Michelle Bachelet y Sebastián Piñera), los de los presidentes uruguayos (Gregorio Álvarez\*, Julio Sanguinetti, Luis Lacalle, Jorge Batlle, Tabaré Vázquez, José Mujica) y los presidentes venezolanos (Hugo Chávez y Nicolás Maduro).

Es interesante destacar el caso de los tres presidentes uruguayos más antiguos del cuerpo de datos: Juan María Bordaberry\*, Óscar Gestido y Aparicio Méndez\*, que pese a las referencias a Uruguay en sus discursos, guardan muy poca relación con sus compatriotas.

Para analizar con mayor precisión esto, veamos en el dendrograma de la figura 8 el resultado del agrupamiento jerárquico, en donde cada par es agrupado de acuerdo al presidente con

discursos más próximos en el espacio de LSA. Otra vez podemos notar por un lado los presidentes argentinos del primer bicentenario, y por otro los presidentes de Latinoamérica correspondientes a los años desde 1970 en adelante. Por otro lado, también puede observarse el agrupamiento por países que mencionábamos, como también la fuerte relación entre Evo Morales, Hugo Chavez y Nicolás Maduro, y la distancia de Gestido (el más antiguo de los presidentes uruguayos, que asumió en 1967) tanto hacia los presidentes argentinos de 1800 como a los presidentes actuales. Para ver qué tan fiel el dendrograma preserva la relación de similitud entre cada par de presidentes, podemos separar los clusters de donde el coeficiente de inconsistencia sea menor o igual a 1.1, y el resultado es el que se detalla a continuación. A su vez, podemos calcular el coeficiente de correlación de cophenetic entre la matriz de distancia y el clustering para analizar qué tan bien se preserva la consistencia entre ambos [21], cuyo valor es 0.71 (valores cercanos a 1 representan una solución de alta calidad).

- Rosas, Urquiza, Mitre, Sarmiento, Roca, Yrigoyen
- Alfonsin, Menem, De la Rúa, Duhalde, Kirchner, C.F.Kirchner, Scioli, Macri
- Lagos, Bachelet, Pinera
- Allende, Pinochet, Frei
- Correa, Alvarez
- Chavez, Maduro, Evo
- Gestido
- Bordaberry, Mendez, Sanguinetti, Lacalle, Batlle, Tabare, Mujica

---

\* Presidente de facto

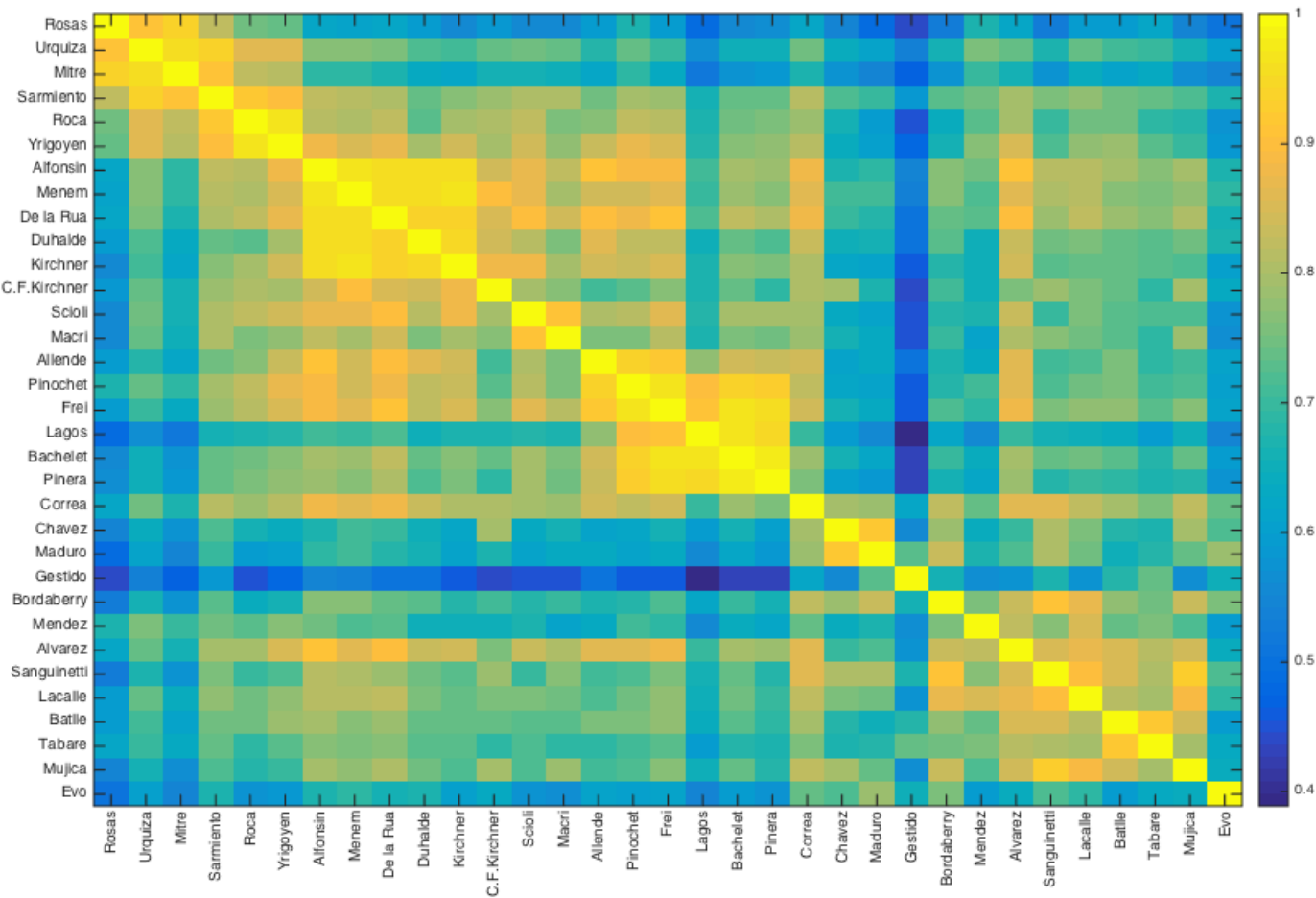


Figura 7: Matriz de relación entre discursos. Dimensión = 400

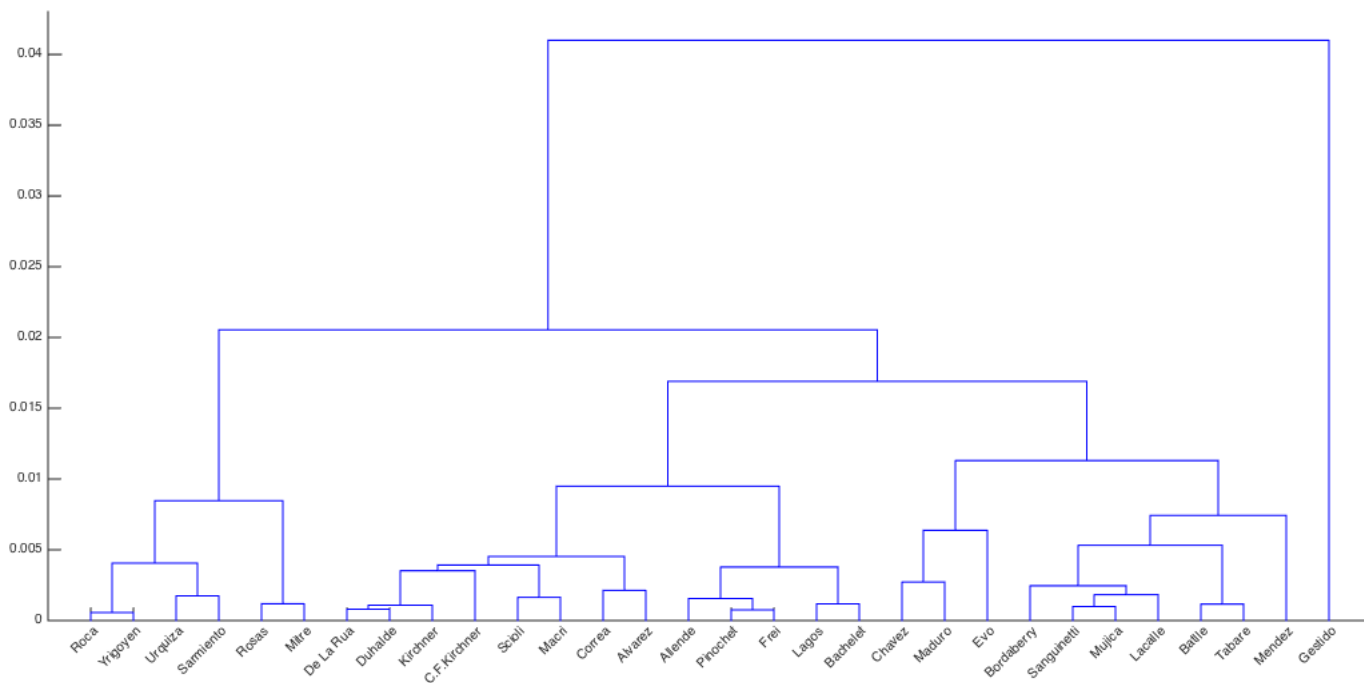


Figura 8: Agrupamiento Jerárquico. Dimensión = 400

**Relación a partir de conceptos.** Respecto a la relación del concepto **enemigo** con los discursos, se puede ver como los discursos de Nicolás Maduro y Hugo Chavez muestran los valores más altos (seguidos de cerca por Rafael Correa y Cristina

Fernández de Kirchner), mientras que los discursos de Fernando de la Rúa y Tabaré Vázquez presentan alta variabilidad (Fig. 9). Cuando consideramos el concepto **pueblo** se destacan Evo Morales y nuevamente Nicolás Maduro con los valores más altos (seguidos por Hugo Chavez y Rafael Correa), y otra vez Tabaré Vázquez muestra alta variabilidad (Fig. 10).



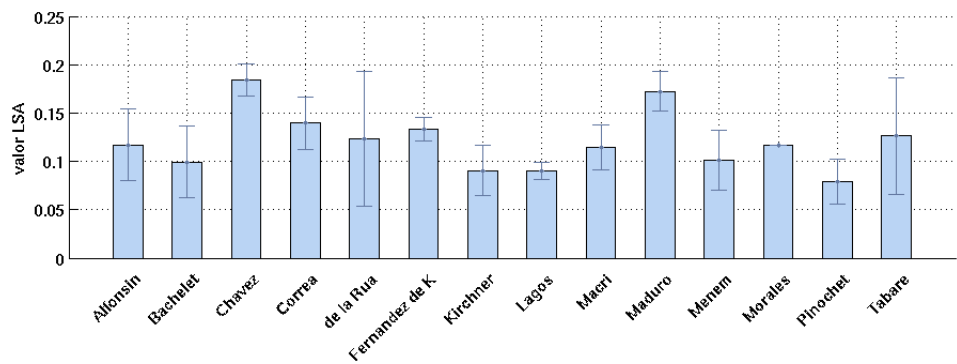


Figura 9: Distribución del valor medio de la distancia del discurso al concepto enemigo en LSA. Dimension = 400

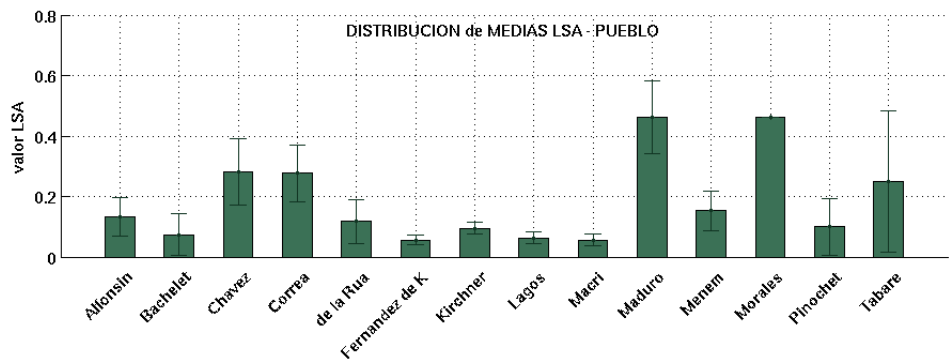


Figura 10: Distribución del valor medio de la distancia del discurso al concepto pueblo en LSA. Dimension = 400

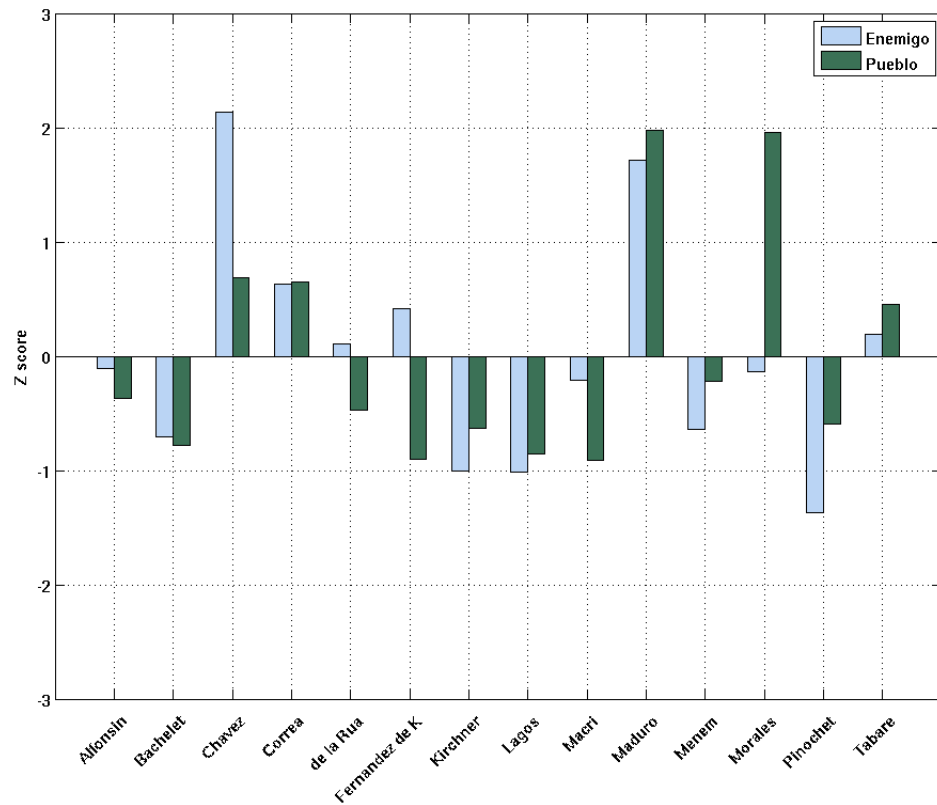


Figura 11: Z Scores de los conceptos enemigo y pueblo por presidente. Dimension = 400

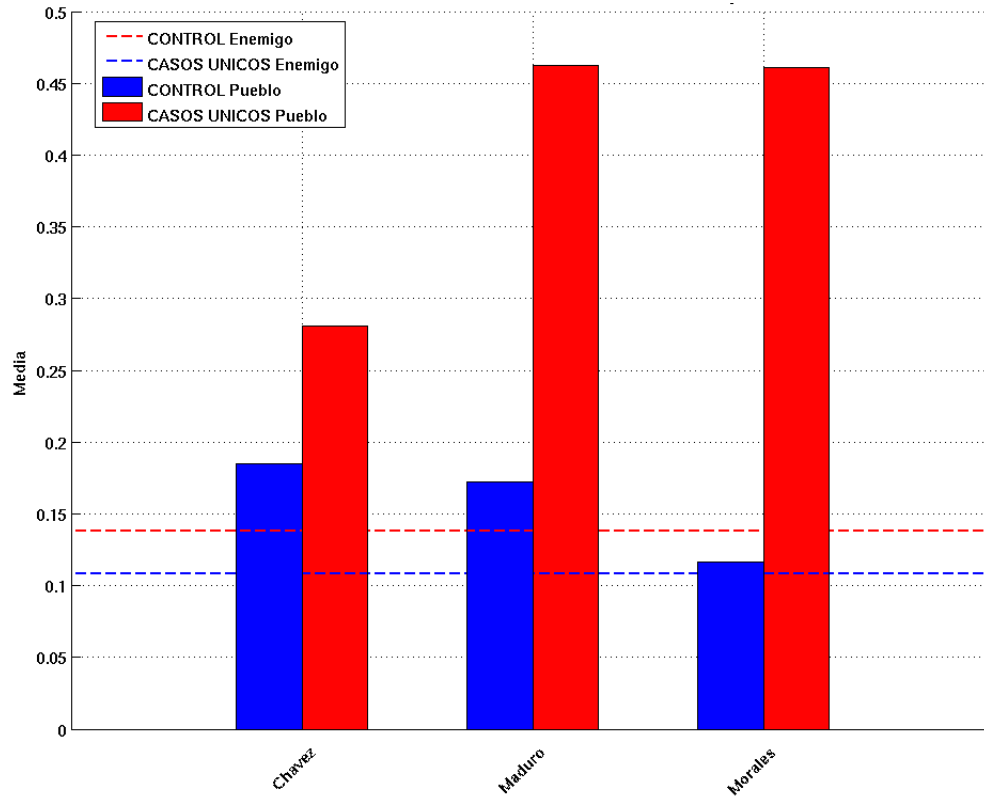


Figura 12: Media grupo control vs presidentes. Dimension = 400

A partir del examen de los Z Scores (Fig. 11) es posible observar que Nicolás Maduro y Hugo Chávez y Augusto Pinochet se destacan fuertemente en el concepto **enemigo** (aunque este último en sentido negativo por su lejanía al concepto) y tanto Nicolás Maduro como Evo Morales para **pueblo** se distancian mayormente del promedio respecto al resto del grupo.

Para el análisis de comparación utilizamos como casos únicos solamente aquellos presidentes con valores Z mayores a 1 (Fig. 12) ya que nuestro interés era evaluar los discursos que estaban principalmente asociados con los conceptos seleccionados. Los resultados son reportados en la tabla 1, en síntesis estos muestran diferencias significativas entre cada caso único y el grupo de control.

Por último, el análisis de comparación entre los conceptos a lo largo de las dimensiones es reportado en la Tabla 2. Si bien existen diferencias significativas dentro de algunas variables presidenciales dentro *Enemigo* y *Pueblo* entre las dimensiones de LSA, los presidentes que más se destacan en los distintos conceptos resultan ser los mismos en todas las dimensiones (Fig. 11 y 13). Los presidentes Fernandez de K, Pinochet, Lagos y Macri son los que muestran mayor variabilidad en ambos conceptos seleccionados a través de las dimensiones. Por otro lado, el análisis de comparaciones múltiples revela que las diferencias observadas en el concepto *Enemigo* (Tabla 3) surgen principalmente de diferencias entre las tres dimensiones de LSA, mientras que en el concepto *Pueblo* (Tabla 4) tales diferencias se manifiestan entre la dimensión 200 con el resto.

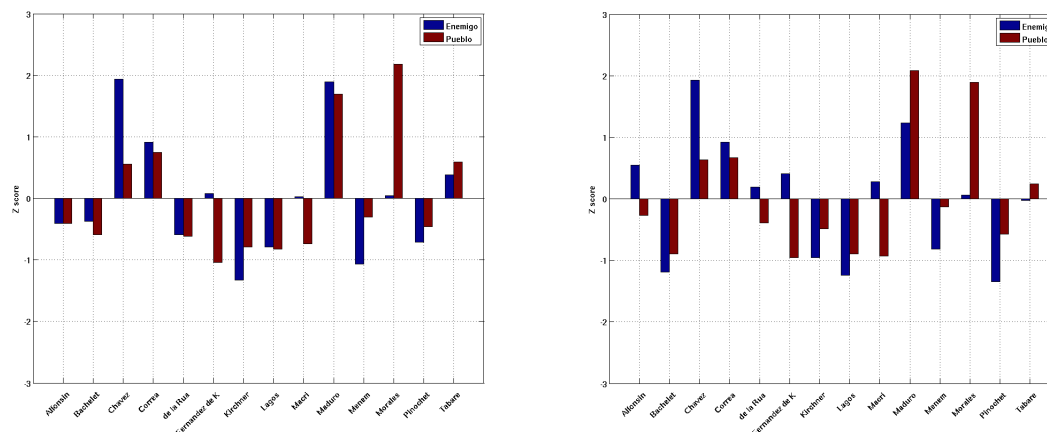


Figura 13: Z Scores de los conceptos enemigo y pueblo por presidente. Dimension = 200 (izquierda). Dimensión = 600 (derecha)

## Discusión

Lo primero que nos gustaría destacar es que si bien la transformación de un discurso al espacio semántico de LSA implica una pérdida de información, igual se logra mantener información relevante para analizar los discursos.

Esto puede notarse en las figuras de análisis de categorías (Fig. 1, 2, 3, 4, 5 y 6) que muestran que la relación entre los conceptos se mantiene bastante estable en las distintas dimensiones. Y más específicamente en el dominio de nuestro problema, en las figuras 11 y 13 donde vemos que los conceptos que pretendemos analizar se mantienen estables para las tres dimensiones en los presidentes que se destacaban sobre el resto (pese a las diferencias que surgen del análisis de la Tabla 2 en la sección de resultados).

Por otro lado, la matriz de relación entre discursos (Fig. 7) y el dendrograma (Fig. 8) muestra un clustering razonable para los presidentes que va más allá del país y sugiere que LSA puede establecer relaciones conceptuales entre los textos más interesantes que la simple aparición de una palabra.

Sin embargo, la poca cantidad de discursos con los que contamos de cada presidente y de cada país (las aperturas de sesiones son una vez por año, no existen mandatos muy largos de un mismo presidente y no todos los discursos de apertura estaban disponibles), hace que el análisis pueda estar fuertemente influenciado por variantes dialécticas de la época y del país del discurso que hagan difícil su uso efectivo. Pese a eso, creemos que los resultados muestran indicios claros de que el análisis de LSA puede ser una herramienta efectiva para el análisis de conceptos de las ciencias políticas en los discursos presidenciales.

Para el caso del populismo hemos dicho que no es esta una categoría teórica estricta de la que exista un consenso en la literatura. Sin embargo, sí existe un consenso en el peso que

los discursos tienen en la práctica populista. Por tal motivo, LSA resulta una herramienta natural para su análisis.

El caso más paradigmático de los resultados fue el de Nicolás Maduro que obtuvo una media alta en sus discursos respecto del resto de los presidentes tanto para el concepto de **enemigo**, como para el concepto de **pueblo** (Fig. 11). Y en menor medida, esto también se ve para el caso de Rafael Correa, Hugo Chavez y Evo Morales. Lo interesante es que, pese a que no siempre suele haber un consenso claro sobre qué presidentes son considerados populistas, sin dudas estos presidentes son caracterizados de esta manera tanto por especialistas como Laclau, como por la mayoría de la literatura.

Por otro lado, presidentes como Néstor Kirchner o Cristina Fernández de Kirchner, no muestran valores altos para los conceptos. Esto no significa que sus gobiernos puedan o no ser caracterizados como populistas sino que permite entender los alcances y limitaciones de esta técnica para el análisis del populismo. Por un lado, el populismo implica la construcción de un significante vacío y una frontera antagónica distinta en cada contexto que adquiere su propio significado en la disputa hegemónica (pueblo-oligarquía, kirchnerismo-corporaciones, revolución ciudadana - pelucones, pueblo-costa ). Distinguir estos conceptos con una única denominación a través de distintos períodos y países, no permite distinguir de manera efectiva la presencia o ausencia de dichos significados como para caracterizar las construcciones políticas.

Nuevamente, los principales inconvenientes resultan ser la escasa cantidad de discursos para poder analizar en cada presidente la presencia o ausencia de estos antagonismos; el poco consenso que existe en torno a la pertenencia de cada líder político en la categoría; la propia práctica populista de la construcción hegemónica que implica cierto juego con los significantes y las cadenas de equivalencia.

Pese a eso, creemos que el trabajo (además del objetivo principal de familiarizarnos con la técnica de LSA), nos dió indicios claros de que el método puede resultar efectivo para analizar la presencia de elementos típicos de la construcción populista en los discursos.

Presidentes	Enemigo (Media; Desviación estándar (máx y mín))				Pueblo (Media; Desviación estándar (máx y mín))			
	M = 0.1083; DE = 0.0172 (0.0796-0.1329)				M,= 0.1378; DE = 0.0846 (0.0556-0.2808)			
	<b>LSA</b>	<b>t</b>	<b>p</b>	<b>z-cc</b>	<b>LSA</b>	<b>t</b>	<b>p</b>	<b>z-cc</b>
<b>Chavez</b>	0.1846	4.2491	<0.01	4.4380	-	-	-	-
<b>Maduro</b>	0.1721	3.5537	<0.01	3.7117	0.4625	3.6860	<0.01	3.8365
<b>Morales</b>	-	-	-	-	0.4607	3.6652	<0.01	3.8149

Tab. 1. Comparaciones (Crawford) entre CTL y presidentes > 1 desvío estándar

PRESIDENTES	ENEMIGO						PUEBLO					
	CHI-Q	P value	N	LSA			CHI-Q	P value	N	LSA		
				200	400	600				200	400	600
Alfoinsín	1.4	0.5	7	13.1	10.6	9.3	0.68	0.71	7	12.3	9.6	11.1
Bachelet	5.36	0.07	5	10.8	8.8	4.4	5.18	0.08	5	11.4	7.6	5
Chavez	23	<0.01*	10	24.4	16.5	5.6	0.34	0.84	10	15.1	16.8	14.6
Correa	7.28	0.03*	8	17	13	7.5	0.18	0.91	8	13.3	12.5	11.8
De la Rúa	1.14	0.56	2	3.5	4.5	2.5	0	1	2	3.5	3.5	3.5
Fernández de K	16.62	<0.01*	8	19	13.8	4.8	13.07	<0.01*	8	19.1	12	6.4
Kichner	1.26	0.53	5	8.6	9.2	6.2	1.82	0.4	5	8.2	6	9.8
Lagos	13.93	<0.01*	6	15.3	9.3	3.8	8.92	0.01*	6	14.7	8	5.8
Macri	11.08	<0.01*	9	20.7	13	8.3	16.66	<0.01*	9	22.8	10.3	8.9
Maduro	8.77	0.01*	4	10	7	2.5	1.88	0.39	4	4.5	7.3	7.8
Menem	3.97	0.14	10	17.3	18.2	11	0.19	0.91	10	16.4	14.7	15.4
Morales	2	0.37	1	3	2	1	2	0.37	1	2	3	1
Pinochet	18.07	<0.01*	16	35.4	23.8	14.4	12.56	<0.01*	16	34.6	19.2	19.7
Tabaré	0.86	0.65	2	4	4	2.5	1.14	0.56	2	4.5	3.5	2.5

Tab. 2. Análisis comparativo de las dimensiones de LSA

	Chavez	Correa	F. de Kirchner	Lagos	Macri	Maduro	Pinochet
[200; 400]	0.0036	0.1949	0.0379	0.0087	0.0434	0.1143	0.0112
[200; 600]	0.0005	0.0443	0.0005	0.0065	0.0107	0.0857	0.0005
[400; 600]	0.0005	0.166	0.0012	0.0087	0.1333	0.0857	0.0205

Tab. 3. p-values corregidos del análisis de comparaciones múltiples para el concepto de enemigo

	Fernandez de K	Lagos	Macri	Pinochet
[200; 400]	0.0413	0.0303	0.0017	0.0064
[200; 600]	0.0009	0.026	0.0017	0.0064
[400; 600]	0.065	0.3939	0.5962	0.8653

Tab. 4. p-values corregidos del análisis de comparaciones múltiples para el concepto de pueblo

**Código fuente.** El código utilizado para el trabajo se encuentra disponible en <https://github.com/sromano/lsatp>

1. Thomas K Landauer, Peter W Foltz, and Darrell Laham. An introduction to latent semantic analysis. *Discourse processes*, 25(2-3):259–284, 1998.
2. Thomas K Landauer and Susan T Dumais. A solution to plato’s problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological review*, 104(2):211, 1997.
3. Aníbal Viguera. “populismoz” neopopulismo.<sup>en</sup> américa latina. *Revista Mexicana de Sociología*, pages 49–66, 1993.
4. Carlos Vilas. Entre la democracia y el neoliberalismo: los caudillos electorales de la posmodernidad. *Socialismo y participación*, 69:31–43, 1995.
5. Carlos María Vilas. La democratización fundamental: el populismo en América Latina. Consejo Nacional para la Cultura y las Artes, 1995.
6. Aboy Carlés. Repensando el populismo. *Política y gestión*, 4, 2003.
7. Gerardo Aboy Carlés. Populismo y democracia en la argentina contemporánea. entre el hegemonismo y la refundación. *Estudios sociales*, 28(1):125–149, 2005.
8. Ernesto Laclau. Populismo:¿ qué nos dice el nombre? In *El populismo como espejo de la democracia*, pages 51–70. Fondo de Cultura Económica, 2009.
9. Ernesto Laclau. *La razón populista*. Fondo de cultura Económica, 2012.
10. Ernesto Laclau and Chantal Mouffe. *Hegemonía y estrategia socialista. Siglo Veintiuno de España ed.*, 1987.
11. Roger B. Bradford. An empirical study of required dimensionality for large-scale latent semantic indexing applications. In *Proceedings of the 17th ACM Conference on Information and Knowledge Management, CIKM ’08*, pages 153–162, New York, NY, USA, 2008. ACM.
12. Radim Řehůřek and Petr Sojka. Software Framework for Topic Modelling with Large Corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, pages 45–50, Valletta, Malta, May 2010. ELRA. .
13. John R Crawford and David C Howell. Comparing an individual’s test score against norms derived from small samples. *The Clinical Neuropsychologist*, 12(4):482–486, 1998.
14. John R Crawford, Paul H Garthwaite, and Sara Porter. Point and interval estimates of effect sizes for the case-controls design in neuropsychology: Rationale, methods, implementations, and proposed reporting standards. *Cognitive Neuropsychology*, 27(3):245–260, 2010.
15. John R Crawford and Paul H Garthwaite. Single-case research in neuropsychology: a comparison of five forms of t-test for comparing a case to controls. *Cortex*, 48(8):1009–1016, 2012.
16. John R Crawford, Paul H Garthwaite, and David C Howell. On comparing a single case with a control sample: An alternative perspective. *Neuropsychologia*, 47(13):2690–2695, 2009.
17. John R Crawford, Paul H Garthwaite, and Kevin Ryan. Comparing a single case to a control sample: testing for neuropsychological deficits and dissociations in the presence of covariates. *Cortex*, 47(10):1166–1178, 2011.
18. WW Daniel. Kruskal-wallis one-way analysis of variance by ranks. *Applied Nonparametric Statistics*,, pages 226–230, 1990.
19. Henry B Mann and Donald R Whitney. On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, pages 50–60, 1947.
20. Sture Holm. A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics*, pages 65–70, 1979.
21. Robert R Sokal and F James Rohlf. The comparison of dendrograms by objective methods. *Taxon*, pages 33–40, 1962.