

PUBLICACIONES DE LA
ASOCIACIÓN ARGENTINA
DE HUMANIDADES DIGITALES



Voyant Tools en el análisis de letras de canciones. El tango y el rock en Argentina

Voyant Tools in Songs Analysis. Tango and Rock in Argentina

Julieta Yasmín GOLDSMIDT

julieta.gold@gmail.com

RESUMEN

El presente trabajo desarrolla los diferentes usos posibles de la herramienta Voyant Tools para el análisis de letras de canciones. Tal experiencia consistió en la comparación de términos que se encontraban en letras de canciones de tango y rock de Argentina, revelando potenciales y limitaciones de esta aplicación web.

ABSTRACT

The present work develops the different possible uses of Voyant Tools for the analysis of song lyrics. This experience consisted of comparing terms found in the lyrics of tango and rock songs from Argentina, revealing potentials and limitations of this web application.

PALABRAS CLAVE

Humanidades Digitales, Voyant Tools, tango, rock.

KEYWORDS

Digital Humanities, Voyant Tools, Tango, Rock.

PublicAAHD -
2021

Vol. 2

ISSN

2718- 7470



Publicaciones de la Asociación Argentina de Humanidades
Digitales PublicAAHD - 2021 Vol. 2

<https://doi.org/10.24215/27187470e029>

1. DESCRIPCIÓN DEL CORPUS

Se seleccionaron 40 canciones compuestas por músicos argentinos que se ingresaron en la plataforma Voyant Tools¹ para su análisis y comparación. El 50% pertenecían al género *tango*, grabadas entre los años 1924 y 1949². La mitad restante consistía en canciones de género *rock*³, publicadas entre 1970 y 1987.

Las letras de las canciones fueron ingresadas manualmente en archivos de texto plano (.txt), diferenciándose por canción y género. Esto implicó la confección de dos archivos en formato .txt; uno perteneciente a las canciones de tango y otro que contenía las letras de rock. Por otra parte, se cargó de manera individual cada canción de tango en archivos diferenciados, generando 20 archivos adicionales para explorar diversos modos de utilizar la herramienta Voyant.

2. TIPO DE TRABAJO SOBRE LOS DATOS

Para observar las diferentes modalidades de análisis sobre el corpus descrito, se procedió a integrar las composiciones seleccionadas al programa Voyant de diferentes maneras; la tarea se sistematizó en cuatro procedimientos:

1. En una primera instancia, se ingresaron en el sistema las 20 canciones de tango en archivos .txt independientes para observar la modalidad comparativa únicamente en ese género⁴.
2. Luego, se emitió la totalidad de las canciones de tango en un único archivo .txt, a fin de observar diferencias y similitudes con el procedimiento anterior, en cuanto a resultados y experiencia de usuario⁵.
3. En un tercer estadio, se ingresó de manera individual un único archivo .txt con 25 canciones de rock⁶.
4. Finalmente, se cargaron dos archivos .txt: uno que contenía las 20 canciones de tango y otro con las 25 correspondientes al género rock.

Estos diferentes modos de procesar el corpus se generaron a fin de evaluar las diferentes maneras en las que Voyant funciona, de acuerdo a la manipulación e ingreso de los datos.

¹ Accesible desde : <https://voyant-tools.org/>.

² Letras recolectadas de repositorio público. Véase R. García Blaya (2021).

En orden de transcripción: García, C. (1983). Nos Siguen Pegando Abajo. Los Enanitos Verdes (1986). La Muralla Verde. Virus (1985). Luna de Miel En La Mano. Los Abuelos de la Nada (1984). Lunes Por La Ma-
drugada. Porchetto, R. (1986). Bailando En Las Veredas. Patricio Rey y Sus Redonditos de Ricota (1987). La Bestia Pop. Soda Stereo (1985). Cuando Pase El Temblor. Sui Generis (1972). Canción Para Mi Muerte. García, C. (1983). Los Dinosaurios. Suéter (1984). Amanece En La Ruta. Vox Dei (1970). Presente. Páez, R. (1985). 11 y 6. Pedro y Pablo (1970). La Marcha De La Bronca. Pescado rabioso (1973). Bajan. Los Twist (1983). 25 estrellas de oro. Mateos, M. (1984). Tirá para arriba. Gieco, L. (1978). Sólo le pido a Dios. Sui Generis (1973). Rasguña las piedras. Sumo (1986). Los viejos vinagres. Gieco, L. (1976). La colina de la vida. Carballo, C. (1982). Es la vida que me alcanza. Ross, M. (1982). Puerto Pollensa. Lerner, A. (1983). Todo a pulmón.G.I.T. (1984). En La calle es su lugar. Nebbia, L. (1980). Sólo se trata de vivir.

⁴ Se registraron como palabras vacías: mí, no, me, te, mí, este, tu, qué, más, a, ni, ya.

⁵ Se aplicaron las mismas palabras vacías que en el procedimiento anterior.

⁶ Se registraron como palabras vacías: a, no, uh, me, mí, oh, te, tu, están, dónde, más, este, qué, mis, ni.

3. RESULTADOS

Durante la primera etapa puede observarse que, de un total de 3756 palabras, las más frecuentes fueron: *amor*⁷ (27); *vida* (21); *corazón* (20); *día* (13); *luz* (13) (figura 1). También se registró un promedio de 22.8 palabras por oración.



Figura 1. Nube de palabras más frecuentes generada por Voyant Tools en la fase 1. Elaboración propia.

En la sección *Sumario* puede encontrarse una comparación cuantitativa entre las diferentes canciones (figura 2).



Figura 2. Resultados. Sección "Sumario" del procedimiento 1. Elaboración propia.

Allí puede identificarse aquellas letras de tango más extensas, con mayor densidad o variabilidad en su vocabulario y la cantidad de palabras por oración. La barra inferior llamada *artículos* facilitó aumentar la visibilidad de canciones que formaban parte de los rankings situados por esta herramienta.

Por otra parte, en la sección *Tendencias*, puede observarse el grado de estabilidad en la presencia que puede tener un término en el corpus a partir de las frecuencias relativas (figura 3). De este modo, la palabra *amor* sufre más variabilidad de presencia en las diferentes canciones. Caso contrario, el término *vida* posee una presencia similar en la mayor parte de las composiciones.

⁷ El valor entre paréntesis representa la frecuencia bruta de cada término.

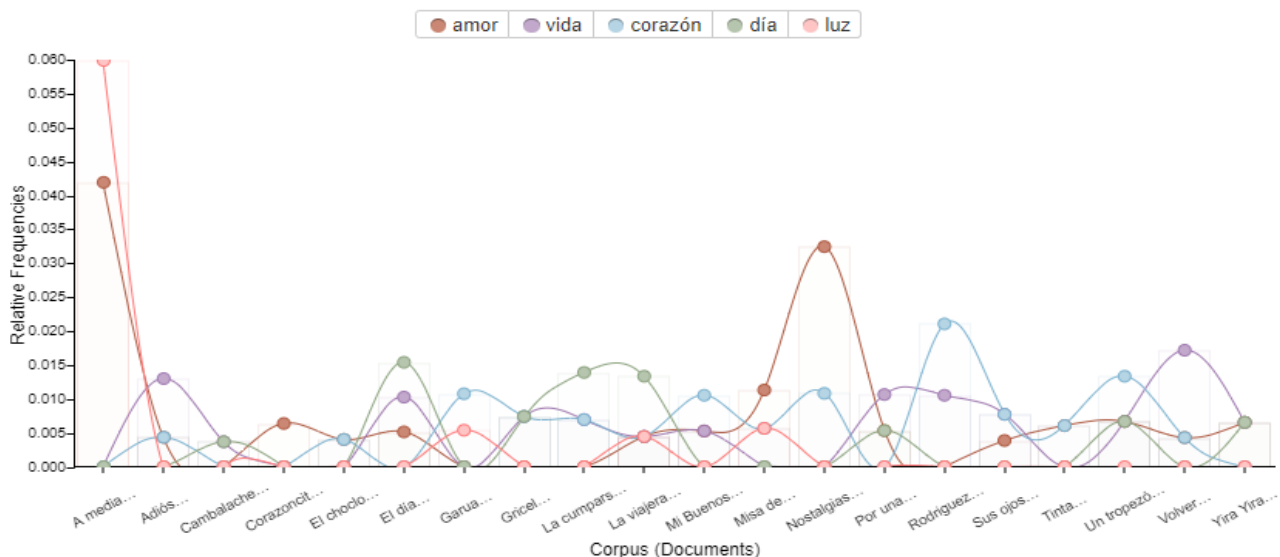


Figura 3. Resultados. Sección Tendencias del procedimiento 1. Elaboración propia.

En la segunda fase de la experiencia, al ingresar el archivo de texto con las canciones de tango unificadas, los resultados globales son exactamente los mismos. No obstante, la diferencia más clara puede situarse en el apartado Sumario, que lógicamente ya no exhibe la comparación entre canciones por estar colocadas en un solo archivo.

Asimismo, en el apartado Tendencias, el corpus se segmenta en 10 bloques seriados y puede observarse que se reduce la precisión con relación al primer procedimiento, dado que este último exhibía la frecuencia relativa en cada documento (figura 4).

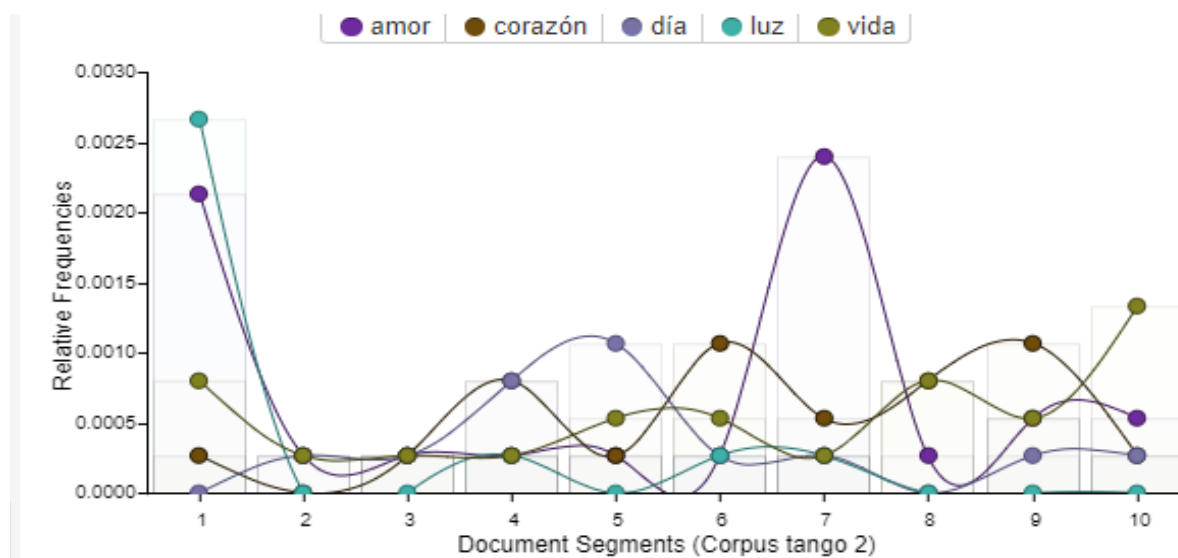


Figura 4. Resultados. Sección Tendencias del procedimiento 2. Elaboración propia.

Prosiguiendo con el tercer procedimiento, se ingresó en Voyant el corpus donde solo se encuentran las canciones de rock. Se registraron un total de 4539 palabras, siendo las más frecuentes: amor (51); desaparecer (37); noche (29); bronca (20); vida (20); brillar (17) (figura 5).



Figura 5. Nube de palabras generada por Voyant Tools en la fase 3. Elaboración propia.

El término *desaparecer* presenta una frecuencia bruta destacable, lo cual se asocia con el período histórico que se ha seleccionado para las canciones de rock (1970-1987). Este abarca la fase de terrorismo de Estado ocurrido en Argentina desde el año 1976 hasta 1983.

Por otra parte, se presenta un promedio de palabras por oración de 90.8, lo cual resulta un número poco esperable para la estructura de una canción. Esto lleva a inferir algún conflicto en la elaboración de los datos con relación a su espaciado⁸.

A su vez, en la sección Tendencias puede encontrarse una llamativa presencia de los tres términos más frecuentes en una parte específica del corpus segmentado en el gráfico, lo cual lleva a cuestionar la posible relevancia global de estas palabras. También es destacable que las tres muestran su alta frecuencia relativa en el mismo sector del documento.

Cabe resaltar que, en las estructuras de las canciones, resulta esperable encontrar frases repetitivas, a modo de estribillo, que podrían alterar la frecuencia relativa de este tipo de corpus (figura 6).

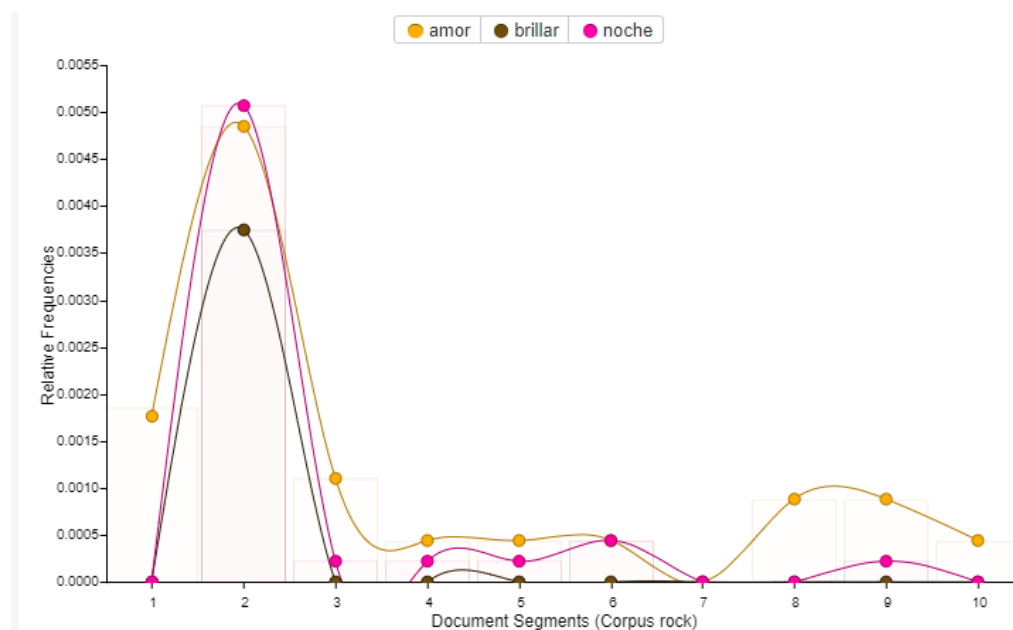


Figura 6. Resultados. Sección Tendencias del procedimiento 3. Elaboración propia.

⁸ Esta apreciación llevaría a situar un testeo de las diferentes maneras en las que Voyant detecta las oraciones y los conflictos con relación al espaciado.

Al continuar con el cuarto procedimiento, se pueden observar los resultados comparados entre ambos géneros. Se ingresaron en Voyant los dos archivos que almacenaban cada uno el total de las canciones, diferenciándose por género: Corpus Tango.txt y Corpus Rock.txt.

En la sección de sumario puede encontrarse la comparación por género que revela una mayor densidad de vocabulario en el tango (figura 7). Las palabras más frecuentes como *amor* (78), *noche* (41), *vida* (41) y *corazón* (26) resultan de interés. En el caso del término *desaparecer* (37), debe considerarse propio del género rock, lo cual se sitúa en el listado de palabras diferenciadas.

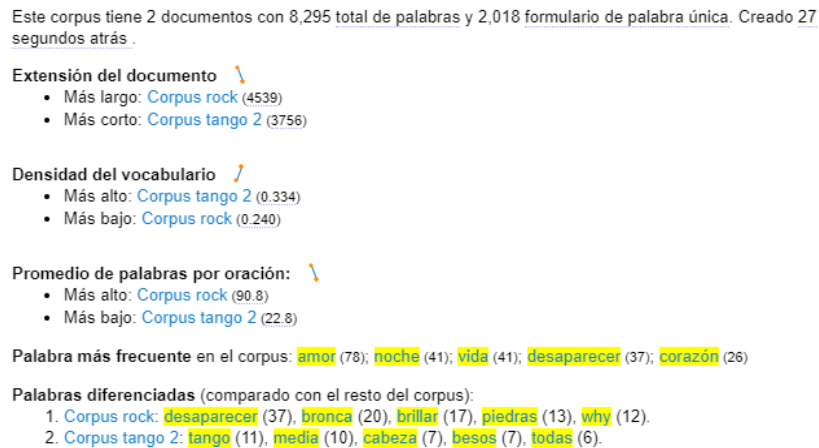


Figura 7. Resultados. Sección Sumario del procedimiento 4. Elaboración propia.

También se observa el mismo detalle en la sección Tendencias (figura 8).

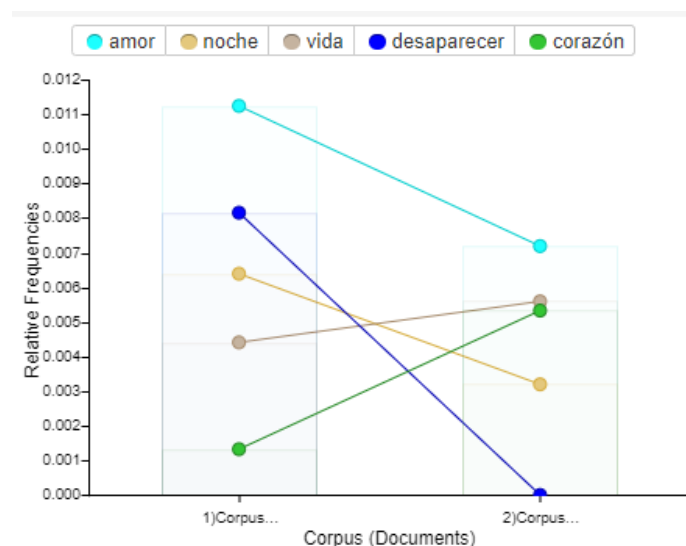


Figura 8. Resultados. Sección Tendencias del procedimiento 4. Elaboración propia.

4. ALCANCES, DESAFÍOS Y LIMITACIONES

A lo largo de esta experiencia ha podido observarse cómo ciertos valores cuantitativos permiten generar nuevos análisis cualitativos respecto de una lectura más cercana del corpus. Estos resultados han llevado a la producción de nuevos campos exploratorios para pensar cómo afectó el contexto de dictadura en las letras del rock de las décadas del 70 y 80. También lleva a situar el *amor* como un término protagonista tanto en el rock como en el tango, lo que podría expandir el

análisis a otros géneros musicales. Asimismo, resultaría de interés indagar la densidad de vocabulario en los diferentes géneros y épocas de la música.

De este modo, Voyant Tools funciona como una herramienta que puede ofrecer resultados cuantitativos interesantes, no pretendiendo generar una hermenéutica ontológica de la lectura distante, sino promoviendo otras maneras de explorar las lecturas cercanas y cualitativas.

No obstante, debe resaltarse la dificultad cuando se manipulan estructuras lingüísticas diferentes a los textos literarios, como ha sido el caso de esta experiencia. La estructura repetitiva de las canciones puede alterar ciertos rasgos del análisis cuantitativo, por lo que deben considerarse los límites de esta herramienta al momento de interpretar los resultados.

Respecto de la experiencia e interfaz de usuario, Voyant Tools muestra un diseño intuitivo y accesible que permite el acercamiento a sus funciones de una manera rápida y simple. De todos modos, podría considerarse una mejora en la traducción al español de alguno de los elementos de la interfaz.

En cuanto a los posibles nuevos desafíos que se presentan, resultaría de interés vincular el nexo de este tipo de recursos con el abordaje de la inteligencia artificial y el análisis del discurso. Esto permitiría generar un mejor filtrado de las palabras vacías y asegurar la expansión del potencial que esta aplicación web presenta.

REFERENCIAS BIBLIOGRÁFICAS

- Abuelos de la Nada, Los (1984). Lunes Por La Madrugada [Canción]. En *Himno de mi corazón*. DG Discos.
- Carballo, C. (1982). Es la vida que me alcanza [Canción]. En *Me vuelvo cada día más loca*. DG Discos.
- Enanitos Verdes, Los (1986). La Muralla Verde [Canción]. En *Contrarreloj*. Sony.
- García, C. (1983). Nos Siguen Pegando Abajo [Canción]. En *Clics modernos*. Ariola.
- García, C. (1983). Los Dinosaurios [Canción]. En *Clics modernos*. Ariola.
- García Blaya, R. (Dir.). (2021). *Todo Tango*. <https://www.todotango.com/>
- Gieco, L. (1976). La colina de la vida [Canción]. En *PorSuiGieco*. Music Hall.
- Gieco, L. (1978). Sólo le pido a Dios [Canción]. En *León Gieco. 7 años*. D&D Música.
- G.I.T. (1984). En La calle es su lugar [Canción]. En *G.I.T.* DG (Interdisc/Universal).
- Lerner, A. (1983). Todo a pulmón [Canción]. En *Todo a pulmón*. Interdisc.
- Mateos, M. (1984). Tirá para arriba [Canción]. En *Tengo que parar*. Music Hall.
- Nebbia, L. (1980). Sólo se trata de vivir [Canción]. En *Creer*. Mellopea Records.
- Páez, R. (1985). 11 y 6 [Canción]. En *Giros*. EMI.
- Patricio Rey y Sus Redonditos de Ricota (1987). La Bestia Pop [Canción]. En *Gulp! Wormo Del Cielito* Records.
- Pedro y Pablo (1970). *La Marcha De La Bronca* [EP]. CBS.
- Pescado rabioso (1973). Baján [Canción]. En *Artaud*. Talent/Microfon.
- Porchetto, R. (1986). Bailando En Las Veredas [Canción]. En *Noche y día*. RCA/Ariola.

- Ross, M. (1982). Puerto Pollensa [Canción]. En *Puerto Pollensa*. MICSA.
- Soda Stereo (1985). Cuando Pase El Temblor [Canción]. En *Nada Personal*. CBS.
- Suéter (1984). Amanece En La Ruta [Canción]. En *Lluvia de gallinas*. DG Discos.
- Sui Generis (1972). Canción Para Mi Muerte [Canción]. En *Vida*. Talent/Microfon.
- Sui Generis (1973). Rasguña las piedras [Canción]. En *Confesiones de Invierno*. Microfon.
- Sumo (1986). Los viejos vinagres [Canción]. En *Llegando los monos*. CBS.
- Los Twist (1983). 25 estrellas de oro [Canción]. En *La dicha en movimiento*. EMI.
- Virus (1985). Luna de Miel En La Mano [Canción]. En *Locura*. CBS.
- Vox Dei (1970). Presente [Canción]. En *Caliente*. Mandioca.