

Dinámicas complejas de artefactos textuales: neurocognición de la escucha y percepción musical

Sección: Artículo
Recibido: 01/05/2026
Aceptado: 06/07/2026

DOI: 10.46530/virtualis.v18i31.496

Complex Dynamics of Textual Artifacts: Neurocognition of Listening and Musical Perception

Juan David Luján-Villar

Secretaría de Educación del Distrito, Bogotá, Colombia

Universidad de las Americas y el Caribe, Colima, México

ORCID: 0000-0001-8622-4774

correo: lujanvillar@gmail.com

Resumen. La literatura científica se establece sobre modelos formales comunicativos, que más allá de ser formatos estandarizados guían la adquisición y despliegue del conocimiento. En este contexto los resúmenes de artículos de investigación posibilitan conocer los alcances y el espacio dialógico de una investigación formal en todos los aspectos básicos. Este trabajo aborda el estudio del corpus de resúmenes del campo de la neurocognición de la escucha y percepción musical (EPM), compuesto por ($n=1430$) documentos y construido a partir de uno de los principales catálogos de información científica actual OpenAlex. Se emplea un modelo de computación lingüística de ponderación a partir de un abordaje que concibe el corpus como un sistema complejo de artefactos textuales. Los resultados ofrecen el vocabulario principal, diversas métricas temporales, estructurales, dinámicas y la evaluación del modelo, desde las perspectivas micro, meso y macroscópicas donde se identifica la consolidación y posterior decrecimiento de la tasa de innovación ($\bar{i}=50,1\%$), una renovación del vocabulario bastante alta ($\tau=94\%$), y una tasa de supervivencia del ($\bar{\varphi}=46,6\%$) en los

Abstract. Scientific literature is built upon communicative formal models that, beyond standardized formats, guide the acquisition and dissemination of knowledge. In this context, research article abstracts make it possible to understand the scope and dialogic space of a formal investigation across all essential aspects. This study examines the abstract corpus from the field of neurocognition of music listening and perception (MLP), comprising ($n=1,430$) documents compiled from OpenAlex, one of the leading current bibliographic databases in science. A weighting-based computational linguistics model is employed, grounded in an approach that frames the corpus as a complex system of textual artifacts. The results yield the core vocabulary, various temporal, structural, and dynamic metrics, and model evaluation from micro-, meso-, and macroscopic perspectives, revealing consolidation followed by mean innovation rate ($\bar{i}=50,1\%$), mean turnover rate ($\tau=94\%$), and mean survival rate ($\bar{\varphi}=46.6\%$) in studies on the neurocognition of MLP. The study concludes by emphasizing the importance and relevance of the proposed design for evaluating scientific data corpora and

estudios sobre la neurocognición de la EPM. Se concluye sobre la importancia y pertinencia del diseño propuesto en la evaluación de un corpus de datos científicos y la aplicación de modelos complejos a través de enfoques interdisciplinares.

Palabras clave: Ciencias complejas, lingüística computacional, música, percepción, escucha

applying complex models through interdisciplinary approaches.

Keywords: Complex sciences, computational linguistics, music, perception, listening

Introducción

En los años recientes la neurocognición de la escucha y percepción musical (EPM) condensa un campo de estudio relevante en el panorama científico mundial. La importancia de esta área reside en la capacidad que posee la música para reconocer y producir estados afectivos en el contexto de prácticas individuales y colectivas de apreciación sonora (Luján-Villar, 2025). Algunos experimentos recientes presentan evidencia de que la música puede fungir un papel trascendental cuando las personas realizan audiciones autoseleccionadas para producir un compromiso musical privado que brinda reparación y sentido de conexión, disminución del sentimiento de soledad y cognición social (Schäfer et al., 2020). Este efecto apela a la música como un subrogado social, una forma de compañía virtual que brinda alivio y apoyo emocional (Groarke et al., 2022).

Taipale et al. (2025) analizan el papel de la escucha musical en el tratamiento de la ansiedad desde el campo de la musicoterapia sobre entrevistas a profesionales experimentados. Las aplicaciones de estos tratamientos revelan que la música en la vida de las personas tiene un papel preponderante a nivel familiar y forma parte del mundo interior de los pacientes. La terapia musical proporciona un lugar de seguridad que es conversacional y promueve acciones efectivas encaminadas a solucionar este problema con una valiosa función de autorregulación anímica. La EPM constituye un campo de investigación consolidado que involucra múltiples disciplinas como la psicología, la etnomusicología, la sociología de la música, las ciencias computacionales y las neurociencias, entre otros. Forma parte de las ciencias comportamentales y cognitivas de manera permanente con un alto potencial metodológico, analítico y presenta un impacto creciente. La principal documentación publicada del campo posibilita comprender la trayectoria de esta área investigativa compuesta en el presente estudio por resúmenes de artículos de investigación (RAI), modelados como un sistema complejo.

Los RAI condensan información que delimita el contexto de una investigación de manera funcional y sintetiza el formalismo y discurso argumentativo científico (Wang y Zhu, 2023). Por lo general los RAI están compuestos por una introducción que establece el contexto, define la temática específica y formula la pregunta investigativa mediante el problema u objetivos delimitadores. Despliega la metodología y el diseño del estudio seguido de la presentación de los resultados principales donde se delinean los hallazgos clave y las conclusiones o implicaciones derivadas, y en ocasiones declaran explícitamente los límites del trabajo. Ermakova et al. (2018) basados en este formato realizan un extenso análisis con 630 investigadores para determinar las partes más importantes de un RAI, e incluyen como indicador el nivel de generosidad o representatividad. Los estudiosos consideran que los aspectos más importantes de una RAI son los siguientes: resultados (32%), objetivos (27%), conclusiones (16%), métodos y diseño (12%), introducción (5%), perspectivas (5%) y limitaciones (3%).

Guerini et al. (2012) realizan un análisis psicolingüístico sobre qué hace viral un RAI en el ecosistema digital. Aplican pruebas de índices de legibilidad y encuentran algunas características estilísticas que incluyen palabras relacionadas con la emocionalidad y la cognición, aspectos fundamentales para establecer claridad, confianza y definir la autoridad investigativa. Suzen et al. (2020) describen el espacio de significado de manera específica en los procesos que conlleva los textos científicos, y abordan el estudio de los RAI como objeto de investigación. En este contexto se destaca en el diseño metodológico el uso de frecuencia inversa de documentos (TF-IDF) con el ánimo de *tokenizar* y encontrar la extracción de características léxicas de los RAI con la finalidad de ponderar la relevancia relativa de esta dimensión (Lopes et al., 2021).

El método TF-IDF es interpretable, computacionalmente eficiente y robusto en corpus pequeños con un vocabulario especializado. Otros paquetes como Word2Vec captura relaciones semánticas distribucionales, los *embeddings* contextuales (BERT, ELMo) modelan la polisemia y algunas dependencias sintácticas complejas, y los LLMs permiten el razonamiento y la generalización en tareas diversas, pero todos exigen recursos computacionales considerables. Debido a que ningún método es universalmente superior se opta por la aplicación del TF-IDF de acuerdo con los recursos disponibles, la tarea de vectorización concreta y la necesidad de interpretabilidad. En este marco de trabajo el TF-IDF puede superar a un LLM en un corpus técnico pequeño con categorías léxicamente bien diferenciadas, así como al Word2Vec sin necesidad de aprender representaciones semánticas densas o grandes volúmenes de texto para poder ejecutar el entrenamiento del modelo por co-ocurrencia.

Esta investigación aplica un modelo de computación lingüística de ponderación aplicado a un conjunto de datos compuesto por RAI (Jurafsky y Martin, 2023; Lopes et al., 2021; Wang y Zhu, 2023), que es concebido como un cuerpo de *artefactos textuales* partícipes de una red que constituye un *sistema complejo*. Tal perspectiva es un esfuerzo interdisciplinar basado en las ciencias sociales computacionales (CSC) (Hovy, 2020) y las ciencias complejas (CC) (Luján-Villar, 2019). En este espacio los RAI son artefactos textuales producto de prácticas discursivas científicas, portadores de marcadores epistémicos y estructuras conceptuales propias del campo de la neurocognición de la EPM. A partir de este marco de trabajo los conceptos recurrentes devienen en nodos de un sistema complejo: elementos interconectados mediante redes semánticas de las cuales emergen propiedades macro como comunidades y centralidades originadas a partir de microinteracciones donde se establecen co-ocurrencias léxicas. Lo anterior posibilita cuantificar y detectar propiedades emergentes, medir la complejidad estructural, determinar las dinámicas del sistema, rastrear la evolución temporal, identificar no-linealidades y evidenciar posibles transiciones.

Objetivos

El objetivo de este estudio es reconocer las dinámicas del conocimiento, emergencia léxica y transiciones de fase conceptual del campo de la EPM y los prepositos neurocognitivos subyacentes, a partir del diseño y análisis de un sistema complejo compuesto por RAI. Los objetivos específicos son los siguientes:

1. Caracterizar la estructura léxico-semántica del corpus sobre vectorización TF-IDF, para identificar el vocabulario temático emergente del sistema.
2. Detectar y caracterizar transiciones de fase en las dinámicas léxicas mediante análisis de puntos de cambio (*change points*), variaciones en entropía de Shannon e indicadores de desaceleración crítica (*critical slowing down, CSD*).
3. Rastrear trayectorias evolutivas conceptuales a través de períodos temporales, identificando tasas de innovación léxica, biografías de términos y flujos de transformación entre comunidades temáticas.

Metodología

El enfoque metodológico aplica el diseño de un modelo de computación lingüística de ponderación ejecutado en el lenguaje de programación Python v.3.13.2 (Python Software Foundation, 2024), donde el corpus de RAI o artefactos textuales es un sistema complejo. El primer nivel *micro* aplica el preprocesamiento con el algoritmo de TF-IDF para *tokenizar* el corpus, delinear características léxicas y extraer términos. Esta

técnica se basa en una forma de representación vectorial dónde los RAI se procesan para obtener vectores dispersos (*sparse*) de alta dimensionalidad, una extensión ponderada del modelo de bolsa de palabras (*bag of words*, *BoW*), y parte del procesamiento de lenguaje natural (NLP) (Hovy, 2020). Esta aplicación pondera cada término según la frecuencia de aparición en el documento (TF) e inversamente según la frecuencia en el corpus (IDF), al priorizar términos distintivos (*rareza*) sobre palabras comunes (Suzen et al., 2020). El proceso se verifica de acuerdo con referentes y protocolos principales en el campo de la lingüística de corpus y el discurso académico (Charles, 2006; Hyland, 2005; Thompson y Ye, 1991). El resultado es la cobertura evaluativa de un corpus organizado y cuantificado de manera precisa en una matrix TF-IDF.

En el segundo nivel *meso* se aborda la modelación de redes para consolidar clústeres que devienen en representaciones de espacialidades disciplinares o interdisciplinares (Leydesdorff y Rafols, 2011). Los nodos o conceptos se organizan de manera reticular en términos, documentos y epistemologías para aplicar indicadores como las métricas de centralidad y las asignaciones comunitarias (*community assignments*). En el presente espacio se emplea la red conceptual para identificar nodos centrales, posibles comunidades auto-organizadas y propiedades topológicas del sistema. El tercer nivel *macro* inserta una perspectiva de complejidad formal que incluye clústeres, valores de entropía, CSD y transiciones de fase, análisis de series temporales (*time series analysis*), puntos de cambio y autocorrelación que explican las propiedades y dinámicas relacionales de los RAI concebidos como sistema.

Obtención de la base de datos o corpus (*dataset*)

La construcción del *dataset* se basa en diversas bases de datos producto de una búsqueda focalizada específicamente sobre el campo de la EPM y los procesos neurocognitivos relacionados¹. Este proceso se realiza el día 15 de enero de 2026 a través de la ecuación de búsqueda final en la plataforma y catálogo bibliográfico OpenAlex: ("music perception" OR "musical perception" OR "music listening") OR ("auditory perception" AND music)) AND (cognition OR cognitive OR neuroscience OR neuroeducation OR "music cognition"). El resultado se depura hacia la obtención de los RAI y diseño de la base de datos que incluye un resultado total de ($n=1793$) documentos.

¹ En el siguiente repositorio se puede descargar la base de datos de este estudio: <https://doi.org/10.5281/zenodo.20804771>

La ecuación final es la combinación de múltiples opciones relacionadas con el campo investigativo, producto de la evaluación de la pertinencia del conjunto de objetivos planificados. La necesidad de proponer un enfoque innovador y oportuno funciona como eje motivador y aporte interdisciplinar en la investigación formal en el panorama científico. El resultado inicial de esta búsqueda es un total de ($n=1793$) registros que combinan una serie de formatos habituales en las comunicaciones y resultados investigativos: artículos, libros, capítulos de libros, editoriales, tesis, revisiones y *datasets* para una frecuencia total de ($n=654$) fuentes, donde se destacan revistas y libros, entre otros formatos de publicación. Este conglomerado incluye una gran serie de referencias bibliográficas donde se evidencia información de metadatos especialmente cuando son RAI extendidos, un aspecto que robustece la muestra.

Resultados

Nivel micro: artefactos textuales y computación lingüística

Después de la representación vectorial TF-IDF construida sobre un corpus de RAI, los parámetros estructurales se basan en un total de ($n=1793$) documentos. Los RAI validos son ($n=1522$) registros y después de la limpieza se obtienen ($n=1518$) documentos procesados, los duplicados removidos son ($n=88$) ítems, con lo cual se obtiene un resultado final de ($n=1430$) documentos aptos para la realización del presente análisis. En la reducción del corpus se efectúa la eliminación de ($n=27$) registros sin *abstract* válido o con valores nulos, vacíos o corruptos. Se descartan ($n=4$) documentos que quedaron vacíos tras aplicar la función de limpieza textual: eliminación de DOIs, URLs, números y caracteres especiales.

Este número de trabajos constituye una base representativa solida debido a que posibilita la extracción de clústeres estables, la estimación robusta de distribuciones léxicas y la modelación de redes semánticas. No se detectan evidencias de sobreajuste estructural ni manipulación artificial del modelo durante la normalización textual según los estándares en el NLP científico (Jurafsky y Martin, 2023). El total de la matriz TF-IDF es de ($n=5000$) términos, con una frecuencia mínima de ($n=5$) documentos y frecuencia máxima de exclusión de ($n=0.85$) términos comunes con la remoción de ($n=437$) palabras vacías (*stopwords*). Este proceso es consistente con un análisis semántico medio-alto y el modelado de dominios científicos de clústeres léxicos.

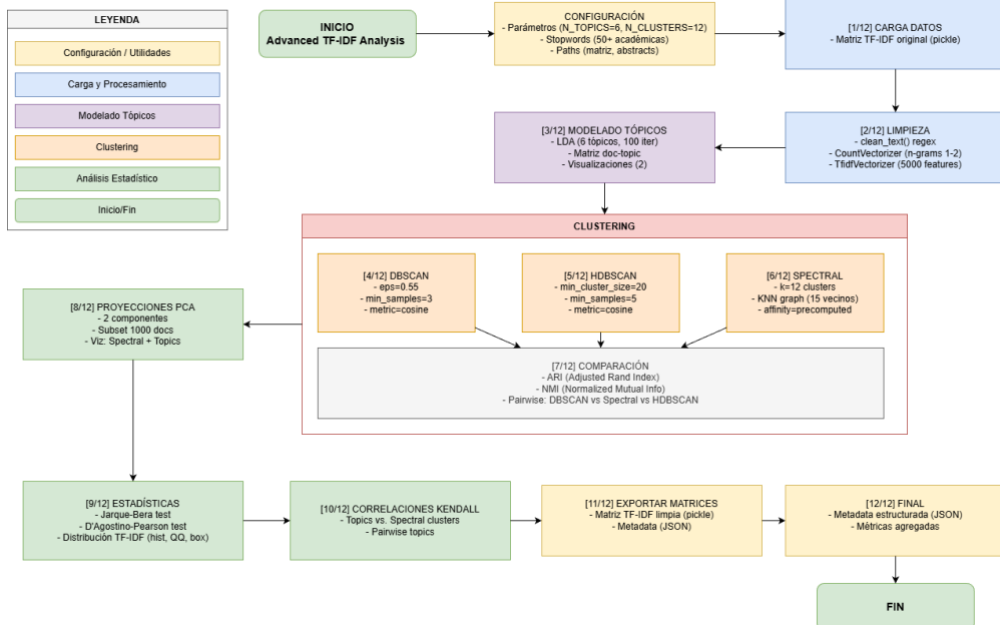
La estructura del programa incluye el uso de las librerías: *numpy*, *pandas*, *scikit-learn*, *scipy*, *matplotlib*, *seaborn*, y *hdbscan* para la vectorización, creación de la matriz dispersa, serialización y articulación de clústeres del corpus debido a la alta efectividad

y las funciones complementarias que estos paquetes aportan al procesamiento, análisis, modelado y visualización de datos lingüísticos. Los hiperparámetros de HDBSCAN configurados en el código son: *min_cluster_size=20*, *min_samples=5*, *metric='cosine'*, *cluster_selection_method='eom'* (*excess of mass*), y *core_dist_n_jobs=-1* (paralelización en todos los núcleos disponibles). Estos parámetros se aplican sobre la matriz TF-IDF previamente normalizada con la norma L2 mediante (*normalize(self.tfidf_clean, norm='l2')*).

En la matriz dispersa TF-IDF cada dimensión corresponde a una palabra del vocabulario, se observada un porcentaje de ($S \sim 97.9\%$ de ceros) lo cual es consistente con literatura estándar en NLP aplicado a textos científicos cortos o RAI (Hovy, 2020; Jurafsky y Martin, 2023). El resultado se encuentra dentro del rango esperado, la reproducibilidad del vocabulario y la estabilidad de los pesos de frecuencia inversa de documento (IDF) demuestra consistencia en transformaciones futuras, y no expresa señales de corrupción del modelo (ver figura 1).

Figura 1.

Seudocódigo del análisis TF-IDF multi-método



En síntesis, el presente modelo emplea niveles o fases analíticas micro, meso y macroscópicas desde la lingüística computacional y de corpus (Hovy, 2020; Lopes et al.,

2021; Wang y Zhu, 2023), incluye la modelación de redes, métricas de complejidad, series temporales y la aplicación del análisis de corpus científicos desde redes de conocimiento bajo una perspectiva teórica de sistemas complejos (Bak, 1996). Esta propuesta investigativa hace parte los estudios de la ciencia de la ciencia o SciSci en los que Fortunato et al. (2018) describen un funcionamiento de la actividad científica sustentado en la formación de comunidades de investigación, redes de colaboración e interacciones entre investigadores, instituciones, publicaciones e ideas que evolucionan y se organizan de manera dinámica. Los nuevos temas investigativos emergen de los vínculos basados en la experiencia propia o entre colegas que forman grupos de autores, donde se demuestra la importancia y pertinencia del abordaje de los títulos investigativos y los RAI para analizar el espacio de las ideas y los territorios conceptuales. A esta perspectiva se suma el enfoque de las ciencias computacionales en el estudio de la complejidad formal, como espacio que viabiliza el análisis de corpus de RAI (Börner et al., 2003).

Vocabulario

El vocabulario resultante del corpus se deriva del total de la suma de todos los valores TF-IDF de cada término a lo largo de todos los documentos de la muestra (*tfidf sum*, W_i). El resultado presenta los principales ($n=10$) conceptos en el grado de importancia de cada término al interior del corpus: *musical* ($W_i=47.3$), *perception*, ($W_i=42.8$), *listening* ($W_i=40.0$), *cognitive* ($W_i=39.0$), *music listening* ($W_i=31.6$), *cognition* ($W_i=30.0$), *music perception* ($W_i=29.4$), *auditory* ($W_i=27.0$), *brain* ($W_i=26.0$), y *processing* ($W_i=23.2$). Estos resultados revelan la representatividad terminológica de un campo decididamente interdisciplinar y especializado en el abordaje de la investigación de los mecanismos neurocognitivos implicados en el procesamiento, la percepción y la comprensión de la música humana en tanto fenómeno acústico complejo.

La cognición musical constituye un proceso psicológico integral que abarca el procesamiento auditivo básico mediante el análisis espectro-temporal y acústico en el córtex auditivo, hasta la percepción superior y las ideas asociadas que involucran específicamente la extracción de altura tonal, el reconocimiento rítmico y la discriminación melódica. Este fenómeno se compone generalmente de la interacción dinámica entre la memoria, la generación de expectativas y el procesamiento estructural para generar una experiencia estética integrada y consciente que es modulada por la atención, las emociones, el contexto cultural y las predisposiciones individuales, sustentado por un sustrato cerebral distribuido en regiones auditivas, motoras, frontales y límbicas que interactúan bidireccionalmente (McFerran y Lotter, 2024). La relevancia de este proceso en las sociedades occidentales contemporáneas es

determinante, según evidencia empírica sirve no solo como una actividad estética también es un sustituto social que brinda compañía, rememora vivencias personales profundas y experiencias significativas (Groarke et al., 2022).

Nivel meso: evolución de redes complejas

En esta parte del trabajo se realiza un análisis longitudinal de los rasgos estructurales de las redes semánticas modeladas en el nivel mesoscópico o de comunidades y macroscópico o de propiedades globales del sistema. Este procedimiento identifica la evolución de las redes semánticas y se basa en la aplicación de diversos indicadores como las métricas por período. Estas medidas incluyen la *densidad* y la *modularidad* en el caso de la identificación de la evolución estructural de las redes, y sobre las *comunidades* se incluye la fragmentación temática, el diámetro y el *clustering* que identifica la cohesión local de la red. Se emplea un número de ($n=200$) términos para ejecutar el modelado de las redes, el umbral de co-ocurrencias mínimas es ($n=2$), la similitud mínima entre documento y documento es ($n=0.3$) con un máximo de documentos en la red de ($n=1000$) nodos. Es importante recordar que los modelos de redes complejas para la síntesis extractiva demuestran un potencial metodológico y explicativo contundente, según la literatura especializada que incluye aplicaciones en el análisis de RAI (Antiqueira et al., 2009) y redes textuales temporales (Vega y Magnani, 2018).

Las métricas resultantes revelan que a nivel estructural la *modularidad* media de la red semántica calculada con el algoritmo Louvain es ($Q=0.175$), es moderado-bajo si se considera un rango teórico de $[-0.5, 1.0]$ para hallar la estructura comunitaria (Newman y Girvan, 2004). Esto se puede comprender como un potencial débil a nivel modular en el caso de los términos y poca capacidad que tienen para formar comunidades fuertemente diferenciadas en sí. Se infiere una ausencia de estructura comunitaria significativa reveladora de una alta densidad conceptual de interconexiones y la ausencia de agrupamientos fuertes en algunos campos aislados. La media de *comunidades* ($\bar{k}=4$) revela la existencia de 4 agrupamientos en promedio de naturaleza indefinida, dónde las fronteras no se trazan de manera definitiva y solo se perciben estas agrupaciones principales. Los diferentes conceptos mantienen conexiones entre sí, lo que puede sugerir la existencia de un vocabulario genérico compartido y un diálogo constante. La *densidad* media ($p=0.4681$) entre un rango teórico $[0, 1]$ evidencia que la red de co-ocurrencias es densa y posee casi la mitad de todas las conexiones posibles entre el total de los ($n=200$) términos principales. Este resultado es un indicio de una posible alta co-ocurrencia entre conceptos, la existencia de un vocabulario técnico interrelacionado y la existencia de un espacio comunitario no definido.

La descripción de los resultados del análisis de la evolución de redes temporales indica que el periodo entre 1985-1996 (períodos 0-2) constituye una red fragmentada con un grado promedio bajo ($k \sim 15-26$), aunque la modularidad ($Q \sim 0.24-0.43$) es alta y tiene como característica la existencia de ($n=5-6$) comunidades separadas compuestas por muchas agrupaciones pequeñas. El periodo que comprende 1997-2004 (períodos 3-4) evidencia un crecimiento que puede ser acelerado con un grado promedio bajo ($k=30-51$) que manifiesta un aumento del 70% en el crecimiento, y mantiene la modularidad decreciente ($Q=0.24-0.17$) lo que sugiere una mayor integración conceptual entre ($n=184-199$) nodos. El último periodo 2005-2024 (períodos 5-9) es temporalmente el mayor con una densificación alta donde el grado promedio bajo ($k=87-194$) aumenta el 124% en el total de la muestra, la densidad ($\rho=0.44-0.98$) pasa a ser un grafo altamente densificado con una modularidad baja ($Q=0.10-0.05$) lo que evidencia una red global unificada donde las comunidades ($n=4-3$) se consolidan de manera definitiva. Cómo hecho notorio la alta transitividad ($\bar{T}=0.63-0.98$) perfila una red bastante cohesiva con un *clustering* promedio de ($\bar{C}=0.72-0.98$), un indicador de la existencia de grupos herméticos en términos conceptuales (ver tabla 1).

Tabla 1.

Evolución de redes

Periodo	Años	N nodos	N vínculos	Densidad (ρ)	Modula. (Q)	N comunida.	Clustering promed.	Grado promed. (k)	Transitividad (T)
0	1985-1988	146	1937	0.182	0.240	6	0.766	26.53	0.472
1	1989-1992	126	1014	0.129	0.307	5	0.825	16.10	0.461
2	1993-1996	165	1308	0.09	0.434	5	0.803	15.85	0.417
3	1997-2000	184	2828	0.168	0.240	4	0.692	30.73	0.437
4	2001-2004	199	5111	0.260	0.174	4	0.652	51.37	0.489
5	2005-2008	200	8680	0.440	0.103	4	0.720	87.00	0.634
6	2009-2012	200	13904	0.699	0.070	3	0.806	13.90	0.785
7	2013-2016	200	16321	0.820	0.065	3	0.871	163.21	0.863
8	2017-2020	200	18184	0.913	0.060	3	0.936	181.84	0.933
9	2021-2024	200	19430	0.980	0.054	3	0.980	194.3	0.980

Comunidades

En el contexto de este estudio una comunidad conceptual se desprende del proceso de creación de redes y se define como un grupo de nodos o términos mayormente conectados entre sí a diferencia del resto de la red. Estos grupos de conceptos detectados co-ocurren frecuentemente y en términos comunitarios funcionan como intermediarios para determinar tópicos o subtemas interpretables como clústeres temáticos del discurso científico, basados en la estructura de la red y no en modelos

probabilísticos debido a que emergen realmente del relacionamiento terminológico a nivel de documentos. Es necesario precisar que el algoritmo de Louvain trabaja sobre la red de relaciones al identificar la modularidad interna, independientemente del azar o de los pesos TF-IDF que son asignados. Para efectos de practicidad analítica en este espacio se presentan los términos en transformación y no los nacimientos, las muertes o las continuaciones conceptuales dentro del campo analizado.

En este marco de trabajo la evolución temporal implica describir los actuales resultados a partir de los mecanismos de transformación e identificar la dirección del campo como un producto de la acumulación del conocimiento a través de la dependencia histórica. Hull (1990) reconoce la ciencia como proceso evolutivo que se articula en este trabajo: existe una *variación* donde nuevos términos emergen, la *selección* mantiene la supervivencia terminológica, la *retención* implica que los términos que sobreviven condicionan el siguiente estado y se produce la *evolución* conceptual sin sustitución o cambio arbitrario. Este enfoque difiere de un *cambio* bidireccional o aleatorio debido a la identificación de una existente dependencia de la trayectoria, un opuesto de posibles cambios aleatorios como puede ver más adelante en la sesión de transiciones de fase donde la memoria del periodo es ($\rho_1=0.87$) (ver tabla 3) y la acumulación incluye herencia selectiva con una tasa de supervivencia del ($\varphi=46,6\%$), lo cual indica que cada *hereda* y *transforma* el anterior.

Entre el periodo 1985 y 1996 aparecen en la dinámica evolutiva del campo los términos: *development, musical, music perception, brain, music listening, importance, music listening, auditory, auditory perception, melody, processing, musical, normal* y cambian hacia temas como *normal, importance, positive, provide, motor* y *perception cognition*. En esta fase se puede reconocer el establecimiento de bases neurocognitivas, el énfasis en el procesamiento auditivo básico, la focalización del desarrollo normal o atípico y la investigación de sustratos cerebrales en las emergencias temáticas. Aspectos que subrayan la importancia o normalización conceptual del procesamiento auditivo, la positividad o valencia afectiva, el sistema auditorio y motor para producir una posible integración sensoriomotora en el campo de la percepción cognitiva. En el periodo de 1993 y 2000 se transforman los conceptos: *positive, music listening, perception cognition, music perception* y *cognition* en las nociones: *different* y *changes*. La unificación entre percepción y cognición enfatiza la variabilidad y diferencia en la sensibilidad a las transformaciones temporales y cambios en la escucha, concebida como proceso cognitivo activo que como se puede observar en los cambios conceptuales, deviene en una emergencia temática producto de la consolidación conceptual de la percepción cognitiva. Este constructo conceptual fundamenta la idea de una dimensión afectiva de

la escucha musical, señala diversos cambios de los procesos o mecanismos subyacentes en la percepción musical basada en una modelización teórica.

Tabla 2.

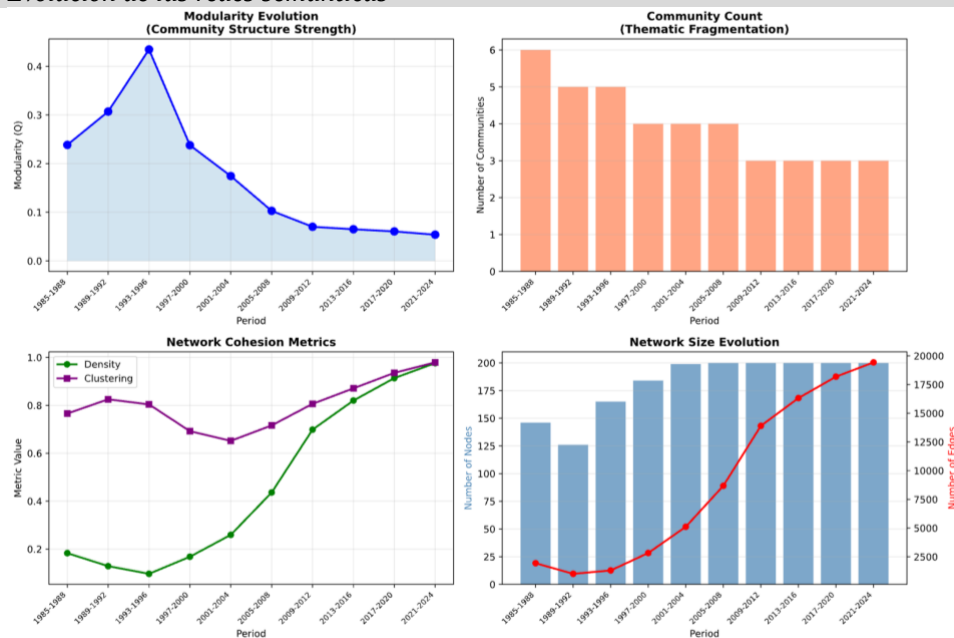
Evolución de comunidades

Periodo ini.	Period fin.	Evento	Similitud	Términos ini.	Términos fin.
1985-1988	1989-1992	Transformación	0.429	development, musical, music perception	processing, musical, normal
1985-1988	1989-1992	Transformación	0.429	brain, music listening, listening	brain, importance, music listening
1989-1992	1993-1996	Transformación	0.25	brain, importance, music listening	positive, music listening, listening
1989-1992	1993-1996	Transformación	0.25	auditory, auditory perception, melody	auditory, provide, motor
1989-1992	1993-1996	Transformación	0.25	processing, musical, normal	perception cognition, musical, perception
1993-1996	1997-2000	Transformación	0.25	positive, music listening, listening	musical, music listening, listening
1993-1996	1997-2000	Transformación	0.25	perception cognition, musical, perception	music perception, perception, cognition
1997-2000	2001-2004	Transformación	0.429	music perception, perception, cognition	perception cognition, music perception, different
1997-2000	2001-2004	Transformación	0.429	musical, music listening, listening	music listening, listening, changes
2001-2004	2005-2008	Transformación	0.429	perception cognition, music perception, different	processing, musical, music perception
2001-2004	2005-2008	Transformación	0.429	music listening, listening, changes	cognitive, listening, emotional
2009-2012	2013-2016	Transformación	0.429	musical, processes, music perception	model, musical, music perception
2013-2016	2017-2020	Transformación	0.25	model, musical, music perception	human, knowledge, processes
2017-2020	2021-2024	Transformación	0.429	provide, clinical, listening	stress, provide, conditions
2017-2020	2021-2024	Transformación	0.429	human, knowledge, processes	knowledge, language, processes

En los años 2001 y 2016 se transforman los conceptos: *perception cognition*, *music perception*, *different*, *music listening*, *changes*, y *processes*, y pasan a ser los términos: *cognitive*, *listening*, *emotional* y *model*. La construcción de modelos explicativos formales de la percepción musical incluye la integración de la dimensión emocional con énfasis en los procesos dinámicos afectivos, y la distinción de las utilidades prácticas de este enfoque. Esto se torna en emergencia temática basada en la aplicación terapéutica, donde la escucha emocional y la salud mental se abordan desde las

condiciones de estrés respecto a los procesos cognitivos, para generalizar teorías y conocimientos disciplinares del campo de la musicoterapia. Como resultado aparecen los primeros modelos de comparación interdominio en el estudio de la música y el lenguaje. Para los años 2013 y 2020 los términos: *model*, *music perception*, *provide*, *clinical*, *listening*, *human*, *knowledge* y *processes* devienen en los conceptos: *human*, *knowledge*, *processes*, *stress*, *provide*, *conditions*, *knowledge*, *language*, y *processes*. Como se puede observar esta terminología está vigente y es central en aplicaciones clínicas y terapéuticas, al abordar de manera relevante la salud mental a partir de una visión de la cognición humana amplia con énfasis en la utilidad práctica de la EPM (ver tabla 2). La figura 2 presenta de manera sistemática las dinámicas de las comunidades y la evolución de las redes derivadas de la muestra analizada.

Figura 2.
Evolución de las redes semánticas



Nivel macro: artefactos textuales y dinámicas de complejidad

Dinámicas entrópicas y tests estadísticos

La entropía de manera general se encarga de medir señales, aumenta con el grado de aleatoriedad, cuantifica la capacidad predictiva de un sistema y cuando los sistemas manifiestan impredecibilidad extrema se dice que la entropía es máxima (Soleymanpour y Arvaneh, 2016). En este contexto la entropía cuantifica la diversidad léxica de un corpus en períodos temporales discretos. Específicamente la entropía media de Shannon [*mean Shannon entropy*] $H(t)$ cuantifica la diversidad léxica o el contenido de información de una fuente de datos, en este caso un corpus de artefactos textuales donde valores altos indica heterogeneidad semántica. La aplicación del presente análisis arroja un total de ($H=9.913$) bits lo que indica una heterogeneidad semántica alta durante un rango temporal de 39 años (1985-2024). Es importante manifestar que el análisis estadístico realizado es riguroso y está atravesado por 12 tests diversos e independientes: Mann-Kendall, Spearman, Kendall Tau, Runs Test y Mann-Whitney U, entre otros.

La regresión lineal demuestra de manera unánime que los 12 tests obtienen $p<0.01$). La entropía de Shannon como el vocabulario exhiben una tendencia creciente significativa de naturaleza *monótona* y de tipo *determinista* en los ($n=10$) periodos establecidos entre 1985 y 2024. La magnitud del cambio a nivel entrópico aumenta a ($H=0.19$) bits cada 4 años ($R^2=0.85$, $p<0.001$), y el vocabulario realiza lo propio con ($n=1,334$) términos cada 4 años ($R^2=0.88$, $p<0.001$). En este nivel la transformación cuantitativa se analiza con la prueba Mann-Whitney ($U=0$), lo que indica una separación completa entre cada fase: temprana (1985-2004) y madura (2005-2024) aspecto que confirma de modo claro un patrón evolutivo cualitativo observable del campo de la EPM a nivel neurocognitivo.

El rango de entropía [*entropy range*] se basa en los indicadores entropía de Shannon, normalización entrópica y la tasa de hápax. La entropía de Shannon presenta un resultado de ($R_H=8.940 - 10.714$) que indica una tendencia general de tipo creciente donde 8.9 pasa a 10.7 y confirma una fase expansiva global con un incremento neto del 18%. Se encuentra una anomalía breve o contracción léxica a partir de una caída del -6.4% en el período 1 (1989-1992), y una aceleración comprendida como cambio mayor en el período 3 a manera de explosión léxica con el 6.2%. Finalmente, el análisis evidencia desaceleración debido a cambios decrecientes en los periodos post-2008 cuando el 2.3% que se reduce y pasa a ser 1.4%, una señal de maduración o estabilización del vocabulario. Se registran señales de crecimiento entrópico leve cuando pasa de 0.9% a 1.4% en el último periodo con un tamaño medio del vocabulario de ($n=5.357$) términos, una muestra de diversidad léxica creciente con el tiempo. Este aspecto sugiere la existencia de una fase exploratoria y diversificada con un vocabulario conceptual en expansión de carácter de dinámico.

La entropía normalizada [*evenness*] describe si la entropía absoluta aumenta –en este caso la diversidad léxica total–, entonces la entropía normalizada disminuye debido a que se presenta una distribución menos uniforme. En este contexto significa *especialización* y como se observa en la tabla 3 decrece al -10% como una posible muestra de concentración en la distribución. La proporción de hápax (*hapax ratio*) pasa de ($HLR=0.456 \approx 0.458$) entre los periodos 8 y 9 con un cambio <0.5%, un marcador de estabilidad y campo en equilibrio dinámico entre innovación y consolidación debido a que la proporción 0.456 revela que el 45% son términos únicos y el 55% restante son términos recurrentes. Este tránsito entre el periodo 8 y 9 exhibe una expansión orgánica dentro de una trayectoria lineal, quizá debido a la madurez del campo que posibilita la absorción de innovación sin fragmentar las dinámicas temporales.

Tabla 3
Dinámicas entrópicas

Años	N docs.	Entrop. de Shannon	Entrop. norma.	Vocabulario (# palabras)	Ratio de términos únicos	Hápax_ legomena ratio	Cambio entróp.	Cambio entróp pct.	Volatilidad entróp.
1985-1988	15	9.547	0.877	1889	0.355	0.600			
1989-1992	13	8.940	0.866	1273	0.370	0.624	-0.607	-6.4	0.429
1993-1996	19	9.023	0.885	1168	0.398	0.617	0.083	0.9	0.329
1997-2000	32	9.581	0.853	2389	0.285	0.544	0.557	6.2	0.348
2001-2004	57	9.941	0.850	3309	0.261	0.525	0.360	3.8	0.462
2005-2008	93	10.061	0.831	4397	0.200	0.494	0.119	1.2	0.250
2009-2012	175	10.297	0.811	6637	0.164	0.482	0.236	2.3	0.181
2013-2016	233	10.464	0.802	8463	0.143	0.488	0.166	1.6	0.202
2017-2020	312	10.556	0.791	10330	0.115	0.458	0.092	0.9	0.131
2021-2024	481	10.714	0.779	13717	0.099	0.456	0.157	1.4	0.126

El vocabulario se expande de 1 889 a 13 717 palabras con un aumento del 33% y una expansión sostenida de 3 387 términos nuevos. Este tipo de distribución concentra términos técnicos dominantes que emergen hacia la especialización sostenida con un vocabulario *core* o técnico en contraposición a un campo semántico periférico, sobre una expansión exponencial de 447% en 35 años. Por último, la volatilidad mínima histórica es de 0.126, -3.82% en contraposición al período 8. Esta visión global sostiene

que el sistema exhibe un atractor estable no caótico, una posible señal de auto-organización hacia el equilibrio o estabilidad paradigmática donde la volatilidad decae espontáneamente si se considera que la escala temporal es larga (39 años), un aspecto que permite ver fases completas de los RAI sobre mecanismos neurocognitivos de la EPM (ver tabla 3).

Transiciones detectadas

La tabla 4 presenta las transiciones de fase detectadas sobre las métricas k y los períodos p , desde un enfoque derivativo a partir de los valores z -scores aplicados durante la modelación de las transiciones. A grandes rasgos el z -score estandariza la magnitud de un valor relativo a la dispersión general, un parámetro idóneo para consolidar transiciones de fase situadas y puede correlacionarse con eventos reales como crisis o sucesos inesperados, por ejemplo, una pandemia. Las transiciones de fase estructurales identifican cambios de punto [*change points*] vía derivadas z -score que se pueden comprender como movimientos cualitativos de cambios del sistema, producto de procesos funcionales de fragmentación disciplinar o la emergencia de nuevas subdisciplinas (ver figuras 3 y 4). El análisis de transiciones de fase revela que el campo experimentó convergencia estructural en el periodo de 1989-1992 donde se evidencia consolidación, quizá debido a una posible convergencia semántica.

Tabla 4.

Periodos de transiciones de fase detectadas

Métrica	Periodo	Periodo_#	Cambio magnitud	Dirección	Interpretación
Entropía	1	1989-1992	2.04	Disminuye	Consolidación, fase de convergencia semántica.
Vocabulario	9	2021-2024	2.98	Disminuye	Aumento significativo del vocabulario.
Modularidad	3	1997-2000	2.30	Disminuye	Integración convergencia interdisciplinaria.
Comunidades	1	1989-1992	2.12	Disminuye	Consolidación del programa de investigación.
Comunidades	3	1997-2000	2.12	Disminuye	Consolidación del programa de investigación.
Comunidades	6	2009-2012	2.12	Disminuye	Consolidación del programa de investigación.

Figura 3.

Transiciones de fase detectadas en el discurso científico

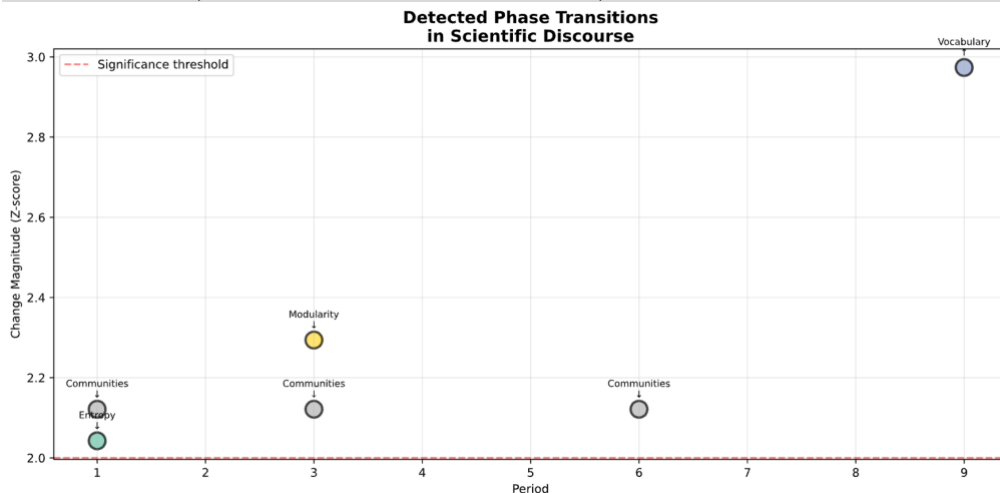
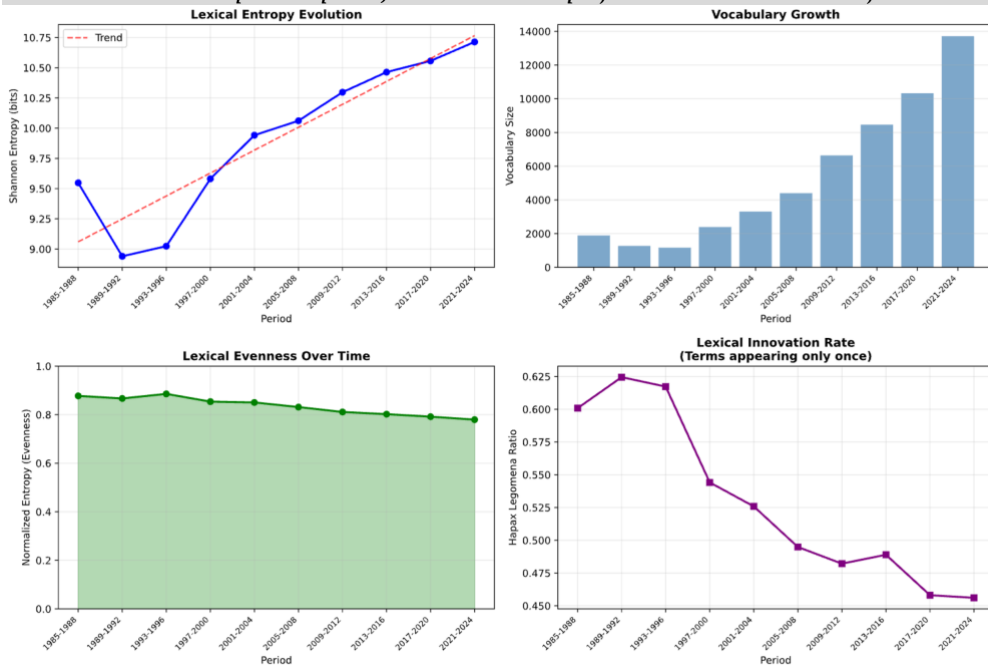


Figura 4.

Dinámicas de entropía temporal, análisis de complejidad del discurso científico



Desaceleración crítica

En este punto del estudio se presenta el análisis de indicadores de desaceleración crítica (CSD), un método que detecta señales de ralentización crítica para medir la estabilidad o comportamiento de un sistema durante series temporales (Ben-Yami et al., 2023). Los lenguajes escritos pueden conceptualizarse como sistemas complejos no-lineales que evolucionan mediante el comportamiento colectivo humano (Amato et al., 2018). Estos sistemas exhiben propiedades características como la sensibilidad a las condiciones iniciales donde pequeños cambios estructurales pueden producir transformaciones drásticas y comportamientos impredecibles en períodos posteriores. Esta dinámica es típica de sistemas en transición cercanos a puntos críticos de bifurcación. Los indicadores de CSD son relevantes para el estudio del cambio lingüístico porque detectan precisamente estos umbrales de transición, permiten identificar cuándo un sistema lingüístico en este caso el inglés y el español se aproxima a cambios sistemáticos significativos, al rastrear las señales de ralentización que preceden a transformaciones profundas.

En el presente sistema compuesto por un corpus de artefactos textuales o RAI los CSD son señales tempranas de transición que pueden tener una mayor autocorrelación y varianza, también conocidas como métricas de señales de alertas tempranas [*early warning signals*] durante la comparación de mitades de series. En este análisis se considera un número de ($n=10$) períodos temporales, con una granularidad de ($n=4$) años por período, una serie robusta de ($n=95 - 250$) documentos en el análisis de redes, y un umbral algorítmico CSD ($n>10$) para el cálculo automático de tendencias. Sin embargo, este número de documentos no permite la realización del análisis de tendencia de autocorrelación y tendencia de varianza para abordar las alertas de CSD. Esta decisión metodológica se mantiene en ($n=10$) documentos para preservar robustez muestral en todos los demás análisis del sistema a nivel temporal como las dinámicas entrópicas, la evolución de redes, las transiciones de fase, el seguimiento de comunidades y las tasas de innovación. En este caso se prioriza la compresión y validez interna sobre la algoritmia puntual de un único análisis.

El resultado presenta evidencia mixta sobre advertencias de desaceleración crítica en tanto sistema dinámico que exhibe un patrón dual, esto señala una posible convergencia semántica de una entropía estable, baja variación y la divergencia estructural de la modularidad volátil y reorganizaciones recurrentes. Como se puede observar no se identifican advertencias CSD formales debido a la limitación técnica del algoritmo con $n \text{ periodos} = 10$, a pesar de esta restricción el patrón dual es robusto y reproducible en ventanas temporales menores. El sistema presenta un estado complejo

caracterizado por estabilidad semántica de entropía local ($CV=6\%$), un vocabulario conceptual que se mantiene relativamente constante y exhibe una trayectoria léxica predecible ($r=0.87$). Por otra parte, se identifica una inestabilidad estructural recurrente definida por una modularidad ($CV=70\%$), esto implica en el sistema reorganizaciones de redes semánticas frecuentes, fusión o división de comunidades temáticas y la existencia de un patrón compatible con criticalidad autoorganizada [*self-organized criticality*, SOC], caracterizada por fluctuaciones sin escala temporal en modularidad ($CV=0.70$) mientras el sistema mantiene convergencia semántica global ($CV=0.06$) (Bak, 1996).

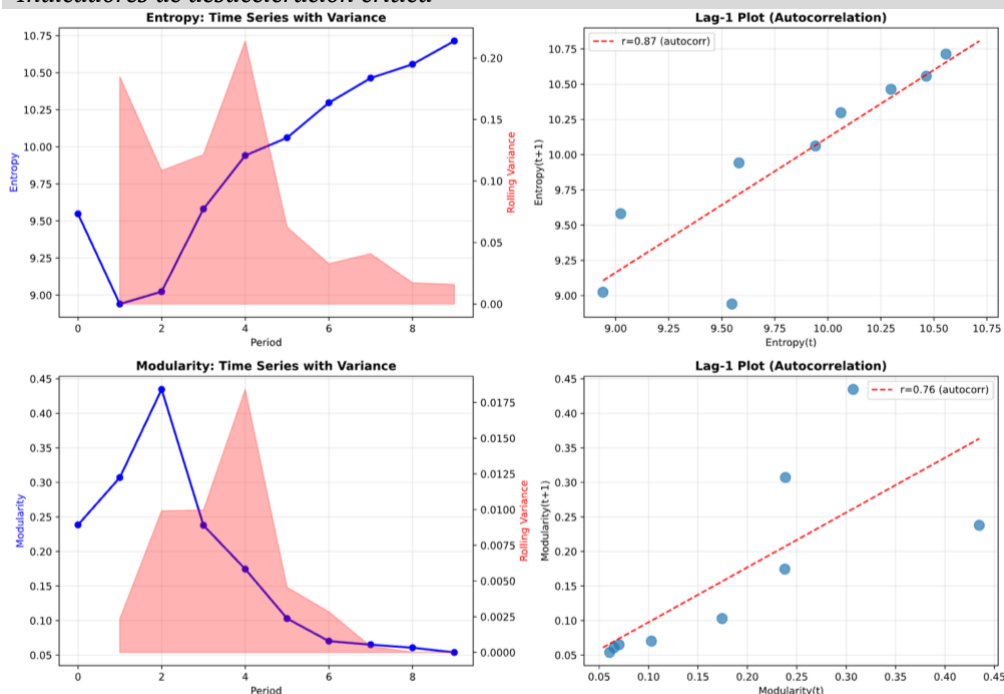
Como se observa en la tabla 5 el sistema presenta un patrón dinámico caracterizado por alta autocorrelación ($lag1=0.874$) para entropía y ($lag1=0.761$) para modularidad, pero exhibe baja varianza estructural para entropía ($CV=0.060$) y modularidad ($CV=0.70$). Este patrón difiere de sistemas complejos cercanos a transiciones críticas que típicamente exhiben tanto autocorrelación aumentada como varianza amplificada (Scheffer et al., 2009; Dakos et al., 2012). En el presente caso la persistencia temporal sin fluctuaciones amplificadas sugiere un estado que se denomina dinámicamente estable con memoria estructural: el campo mantiene coherencia conceptual o alta correlación temporal sin mostrar comportamientos caóticos o bifurcaciones inminentes que son propios de una varianza aumentada. Aunque una validación formal mediante análisis de estabilidad de Lyapunov o mapeo de atractores escapa a las posibilidades técnicas de este estudio, el patrón evidenciado se interpreta como ausencia de alarmas críticas de transición inminente.

Tabla 5.

Indicadores de desaceleración

Métrica	lag1_ autocorr.	lag2_ autocorr.	Varianza	Coefficiente de variación	Tendencia autocorr.	Tendencia variance	Alerta de CSD
Entropía	0.874	0.725	0.351	0.060	Datos insuficientes	Datos insuficientes	Falso
Modularidad	0.761	0.5	0.015	0.70	Datos insuficientes	Datos insuficientes	Falso

Figura 5.
Indicadores de desaceleración crítica



Transiciones de fase

En el método análisis de dependencia de trayectorias (*path dependence análisis, DT*) y las métricas de entropía y modularidad, el retraso (*lag*) indica el número de períodos de tiempo y desplazamiento de una variable respecto a otra. El nivel de correlaciones de Pearson muestra el grado de relación lineal existente entre variables (en el rango de -1 a $+1$), y los valores de p (p value) o pruebas de correlación con un valor bajo $p < 0.05$ sugieren que la relación es estadísticamente real y descarta una correlación de azar en la significancia y posterior interpretación. Esta estructura es fundamental para el análisis de series temporales y la validación de relaciones de causalidad o dependencia temporal (Warren, 2024).

El objetivo de este enfoque es testear si el estado actual del sistema depende de la trayectoria histórica (Brenner y Jeddelloh, 2024). En este modelo el sistema se compone

de artefactos textuales o RAI y la DT revela la existencia de una dependencia significativa de trayectorias: entropía (*lag 1*) ($r=0.874$), entropía (*lag 2*) ($r=0.725$), y modularidad (*lag 1*) ($r=0.762$), lo que se interpreta como una fuerte DT marcada por una continuidad histórica. La tabla 6 exhibe los resultados a nivel general que sugieren considerar el estado actual del campo de la EPM como históricamente contingente. Se puede afirmar que la trayectoria de las ideas dentro del campo depende de los estados pasados, lo cual es un indicador de construcción acumulativa del conocimiento neurocognitivo de la EPM.

Tabla 6.

Transiciones de fase detectadas, indicadores de complejidad con valores de dependencia de trayectorias (DT)

Métrica	Lag (retraso)	Correlación	P_value	Significancia	Interpretación
Entropía	1	0.874	0.002	Verdadero	Fuerte dependencia de trayectorias, continuidad histórica.
Entropía	2	0.725	0.041	Verdadero	Fuerte dependencia de trayectorias, continuidad histórica.
Entropía	3	0.731	0.061	Falso	Sin dependencia de trayectorias, sin memoria.
Entropía	4	0.693	0.126	Falso	Sin dependencia de trayectorias, sin memoria.
Entropía	5	0.407	0.495	Falso	Sin dependencia de trayectorias, sin memoria.
Modularidad	1	0.761	0.017	Verdadero	Fuerte dependencia de trayectorias, continuidad histórica.
Modularidad	2	0.5	0.207	Falso	Sin dependencia de trayectorias, sin memoria.
Modularidad	3	0.419	0.349	Falso	Sin dependencia de trayectorias, sin memoria.
Modularidad	4	0.150	0.775	Falso	Sin dependencia de trayectorias, sin memoria.
Modularidad	5	-0.0026	0.9967	Falso	Sin dependencia de trayectorias, sin memoria.

Dinámicas de innovación

El análisis de la dinámica de innovación léxica del sistema modelado revela una renovación del vocabulario bastante alta (*mean turnover rate*) ($\tau=94\%$) como rasgo distintivo. La mitad del vocabulario activo en cada período compuesto por términos recién introducidos posee una tasa de innovación (*mean innovation rate*) del ($\bar{i}=50,1\%$),

esto indica que la alta tasa de innovación sugiere una frontera conceptual posiblemente en activa expansión. Por otra parte, menos de la mitad del vocabulario del período anterior persiste con una tasa de supervivencia (*mean survival rate*) del ($\bar{\varphi}$ =46,6%). Este resultado en conjunto indica que el 50.1% del vocabulario de cada período son términos que no aparecen con anterioridad en ningún período previo. Una tasa de renovación del 94% subraya entre cada uno de los períodos consecutivos, la existencia de poca superposición léxica una evidencia de renovación semántica constante marcada por un patrón emergente.

La red semántica muestra características estructurales típicas de los campos posiblemente integrados como lo son una baja modularidad ($Q=0,175$) y alta densidad ($\rho=0,47$), características alcanzadas tras múltiples fases de consolidación disciplinar: consolidación en los 90's (entropía baja, comunidades bajan, modularidad baja y se presenta mayor convergencia), estabilidad 2000-2020 sin transiciones detectadas, y expansión reciente 2021-2024 cuando el vocabulario aumenta. Baja modularidad y alta densidad se puede comprender como una posible integración temática que constituiría una conectividad inter-comunitaria y posible integración conceptual en el análisis de redes. Sin embargo, las redes densas pueden producir puntuaciones de modularidad más bajas debido a los límites de resolución de los algoritmos de detección de comunidades (Fortunato, 2010). En este contexto se puede afirmar que el 50.1% del vocabulario actual son términos de primera aparición no registrados en períodos anteriores. La tasa de supervivencia 46.6% es un promedio de periodos del vocabulario previo que persiste, la conectividad alta indica un campo diverso con muchos términos especializados lo que es consistente con la renovación terminológica substancial que el campo experimenta.

Tests estadísticos del modelo

Los tests aplicados posibilitan certificar el modelo y los resultados de su aplicación en términos de tendencias, asociaciones y estabilidad sobre una serie temporal y el conjunto de datos. Buscan determinar la robustez de los hallazgos, algunas pruebas son de carácter no paramétrico como Mann-Whitney ($Z=-3.04$, $p=0.0024$) detectan cambios de nivel. En este caso se identifican cambios donde el valor de este resultado lo que indica una tendencia decreciente monotónica. Spearman ($\rho=-0.92$, $p=0.0002$) correlaciona los rangos de la tasa de innovación (variable) con aquellos rangos del índice temporal (tiempo), el resultado indica la existencia de una alta correlación negativa (-0.92 , cercana al máximo teórico de -1.0), lo cual confirma que los valores más altos están sistemáticamente en períodos tempranos. Kendall Tau ($\tau=-0.78$, $p=0.0009$) confirma estos resultados con mayor robustez estadística, al ratificar la

dirección negativa de modo alternativo: el 78% de predominancia son pares discordantes, es decir, el valor baja a medida que el tiempo sube.

Runs Test ($Z=-2.68$, $p=0.0073$) se encarga de la evaluación de la naturaleza aleatoria secuencial de los valores arriba y abajo de la mediana. En este caso hay menos "runs" de los esperados por el azar, esto indica que los valores altos están agrupados en un extremo y los bajos en otro, demuestra que el patrón no es ruido aleatorio y existe una estructura temporal real al interior del modelo. Mann-Whitney U ($U=25$, $p=0.0079$) compara las distribuciones de la primera y segunda mitad de la serie (mediana), el resultado demuestra que las dos mitades provienen de distribuciones distintas y en consecuencia existe un cambio real entre el período temprano y tardío. La regresión lineal ($slope=-0.044$, $R^2=0.52$, $p=0.0185$) realiza el ajuste del modelo lineal y evalúa la significancia de la pendiente, R^2 mide la fuerza o fiabilidad de la tendencia actual donde la pendiente negativa o tendencia decreciente es significativa: el 52% de la varianza explica y predice la innovación a lo largo del tiempo (ver tabla 7 y figura 6).

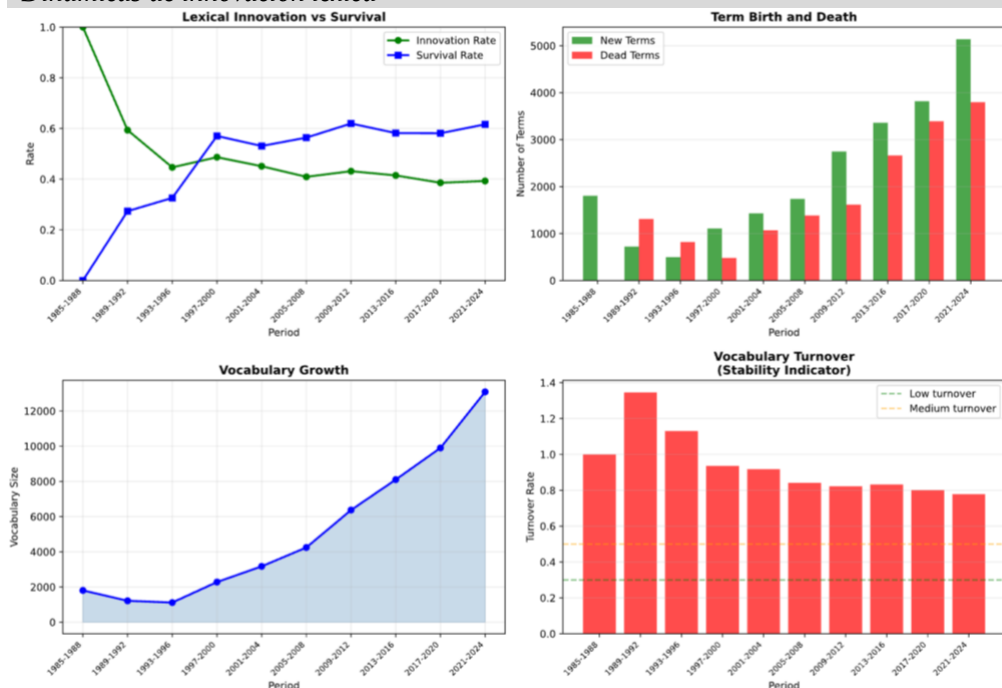
Tabla 7.

Tests estadísticos de rigurosidad

Métrica	Test	Estadístico	P valor	Significancia 005	Interpretación
Tasa innovación	Mann-Kendall	-3,041	0,002358	Verdad	Tendencia decreciente
Tasa innovación	Spearman_vs_Time	-0,915	0,000204	Verdad	Relación monotónica fuerte
Tasa innovación	Kendall_Tau_vs_Time	-0,777	0,000946	Verdad	Concordancia con el tiempo
Tasa innovación	Runs_Test	-2,683	0,00729	Verdad	Patron no aleatorio detectado
Tasa innovación	Mann- Whitney_U_HalfComparison	25	0,007937	Verdad	Cambio significativo: disminución
Tasa innovación	Linear_Regression	-0,044	0,01848	Verdad	$R^2=0.521$, pendiente=- 0.0442

Figura 6.

Dinámicas de innovación léxica



Discusión

Existen diversas tecnologías en las ciencias de la complejidad que se han aplicado al estudio de sistemas complejos (Bak, 1996) en el ámbito de las ciencias sociales y humanas, al emplear presupuestos metodológicos como análisis de redes complejas, dinámica de sistemas, modelado basado en agentes, algoritmos evolutivos (genéticos o culturales), estimaciones fractales, modelos de sistemas multiagentes, sintaxis espacial, sistemas de producción, neurociencias sociales y autómatas celulares (Fortunato et al., 2010; Luján-Villar, 2018). El presente estudio contribuye a plantear un modelo innovador que establece una conexión transdisciplinaria entre investigación sociocultural y modelos originarios del ámbito de las ciencias naturales (física y química), fundamentado en una perspectiva computacional basada en un paradigma cuantitativo, longitudinal y multinivel.

La ciencia contemporánea es un sistema dinámico de empresas impulsado por interacciones complejas entre las estructuras sociales y económicas, las

representaciones del conocimiento científico y los fenómenos del mundo natural. El conocimiento técnico basado en la ciencia está constituido por conceptos y relaciones epistemológicas incorporadas en documentos de investigación que pueden ser libros, reportes, patentes, softwares y otros artefactos académicos, organizados en campos disciplinares y áreas sistemáticas más amplias. Estos elementos sociales, conceptuales y materiales están conectados a través de flujos formales e informales de información, teorías, prácticas investigativas, herramientas de diseño y bases de datos. La ciencia de manera general puede describirse como una enorme red multiescala compleja, auto organizadora y en constante evolución (Fortunato et al., 2018).

El corpus documental reúne los principales RAI comprendidos como artefactos textuales para conformar un sistema dinámico que puede representarse a partir de la elaboración de redes complejas (Ferraz de Arruda et al., 2018; Fortunato et al., 2010). Estas estructuras reticulares son modeladas desde una perspectiva que involucra temporalidad y el análisis de las propiedades internas del sistema desde un nivel detallado, general y global. Aunque la aplicación práctica de esta visión requiere un conocimiento significativo en términos algorítmicos y de programación, es necesario concebir la incursión de modelos computacionales en las ciencias sociales y humanas, en tanto campos del conocimiento verdaderamente interdisciplinares que abordan problemas complejos en el sentido amplio del término. Una gran cantidad de fenómenos requieren tratamientos investigativos de modelación debido a la dificultad que revisten en términos de formalismos y complejidad inherente.

Los sistemas complejos presentan una imprevisibilidad intrínseca (Watts, 2011), pero pueden modelar patrones emergentes, distribuciones, temporalidades, delimitar contextos e identificar trayectorias (Fortunato et al., 2010). La aplicación de este marco de trabajo posibilita identificar los rasgos estructurales y la información relevante de un corpus de lengua escrita, más allá de concebir los registros textuales como simples categorías fijas que develan la naturaleza del lenguaje (Stanisz et al., 2024), en este caso escrito. La idea de concebir los RAI como artefactos textuales implica reconocer la infraestructura científica primaria de un campo científico a nivel representativo, debido a que este formato integra datos, evidencias, enfoques, objetos de modelado e implican una sistematicidad epistemológica compleja.

Conclusiones

Las perspectivas neurocognitivas de la EPM son el resultado final del esfuerzo de miles de investigadores que se interesan por la fascinación que la música despierta en los seres humanos. A nivel interdisciplinar es pertinente aplicar diseños de modelado

complejo a través de la aplicación del potencial hermeneútico propio de los enfoques humanísticos, con el apoyo de diversas algorítmicas computacionales debido a la naturaleza compleja de las dinámicas del mundo social y la producción derivada de las interacciones humanas (Luján-Villar, 2018). La presente investigación integra diversos enfoques en un marco de trabajo multinivel que aborda dimensiones micro, meso y macroscópicas. Esta síntesis metodológica permite de modo innovador caracterizar simultáneamente sistemas complejos que exhiben las propiedades estructurales (redes), dinámicas (entropía temporal) y transiciones críticas (*change points*, CSD) de sistemas semánticos científicos.

La consolidación y posterior decrecimiento de la tasa de innovación ($\bar{i}=50,1\%$), una renovación del vocabulario sumamente alta ($\tau=94\%$) y una tasa de supervivencia media del ($\bar{\varphi}=46,6\%$) en los estudios sobre la neurocognición de la EPM sugieren un fenómeno complejo que no puede reducirse simplemente a una madurez del campo. Tradicionalmente este aspecto suele asociarse a la estabilidad y la replicabilidad, pero aquí coexiste con una alta rotación conceptual que podría interpretarse no como un fin, sino como una transición crítica hacia una fase de integración profunda o reestructuración teórica, donde el campo deja de inventar nuevos principios básicos para comenzar a construir arquitecturas coherentes que integren los conceptos existentes.

La investigación confirma un patrón evolutivo distintivo cualitativo y cuantitativamente evidente en el campo de la neurocognición de EPM, lo que se observa por una entropía que aumenta ($H=0,19$) bits cada cuatro años y una renovación lexical rápida de ($n=1.334$) términos anuales, separándose estadísticamente entre las fases temprana y madura. Esto es consistente con una red semántica posiblemente integrada, caracterizada por baja modularidad ($Q=0,175$) y alta densidad ($\rho=0,47$) lo que sugiere una consolidación disciplinar inicial tras los años 90 y una reciente expansión estructural y temática desde el año 2021.

Los resultados que presenta este estudio exponen el nivel de detalle necesario para resolver los objetivos propuestos. Brindan evidencias de la dinámicas de un sistema complejo lo que implica indagar la temporalidad estructurada, la integración temática y la renovación léxica del campo de la neurocognición de la EPM (Luján-Villar, 2025). La dinámica temporal identificada no es aleatoria, como lo demuestra la ejecución de la prueba $p=0.007$ que evidencia un patrón sistemático, coherente y ratifica el marco referencial temporal del corpus. Los seis tests o métricas de evaluación automática aplicadas sobre la tasa de innovación, convergen sobre la tendencia lingüística, y evidencian dependencia histórica. El método de análisis de dependencia de trayectorias

evaluado con correlaciones y retrasos (*lags*) indica que la tasa de innovación del ($\bar{i}=0.51\%$) depende de los cambios temporales como un mecanismo regulador de la tasa de generación de conocimiento.

El análisis revela una desincronización epistémica característica de campos epistemológicos en fase de transición estructural. La consolidación de la red semántica sobre núcleos temáticos compartidos: baja modularidad ($Q=0.175$) y alta densidad ($\rho=0.47$) coexisten con una tasa de renovación léxica del 94%. Esta disociación funcional donde el vocabulario técnico se reconfigura con una velocidad superior a la estabilización conceptual subyacente no es meramente lingüística, sino reflejo de un posible *lag* cognitivo donde nuevos términos actúan como marcadores de innovación paradigmática antes de ser asimilados al marco teórico central. Los términos no se estabilizan a la misma velocidad que las ideas subyacentes, lo que genera riesgo de fragmentación epistémica si no existen mecanismos rigurosos de estandarización terminológica. La proliferación de jerga no coordinada puede erosionar la capacidad de comunicación intersubjetiva necesaria para integración disciplinar. Esta volatilidad léxica puede ser sensible a cambios de paradigmas científicos, movimientos ideológicos y políticas públicas, factores que pueden amplificar el efecto corrosivo sobre la coherencia del conocimiento acumulado.

Obras citadas

- Amato, R., Lacasa, L., Díaz-Guilera, A., y Baronchelli, A. (2018). The dynamics of norm change in the cultural evolution of language. *Proceedings of the National Academy of Sciences*, 115(33), 8260-8265. <https://doi.org/10.1073/pnas.1721059115>
- Antiqueira, L., Oliveira Jr, O.N., da Fontoura Costa, L., y Nunes, M.D.G.V. (2009). A complex network approach to text summarization. *Information Sciences*, 179(5), 584-599. <https://doi:10.1016/j.ins.2008.10.032>
- Bak, P. (1996). *How nature works: The science of self-organised criticality*. Copernicus. https://doi.org/10.1007/978-1-4757-5426-1_2
- Ben-Yami, M., Skiba, V., Bathiany, S., y Boers, N. (2023). Uncertainties in critical slowing down indicators of observation-based fingerprints of the Atlantic Overturning Circulation. *Nature Communications*, 14(1), 1-11, 8344. <https://doi.org/10.1038/s41467-023-44046-9>
- Börner, K., Chen, C., y Boyack, K.W. (2003). Visualizing knowledge domains. *Annual Review of Information Science and Technology*, 37(1), 179-255. <https://doi.org/10.1002/aris.1440370106>
- Brenner, T., y Jeddelloh, S. (2024). Path dependence in an evolving system: a modeling perspective. *Cliometrica*, 18(1), 1-36. <https://doi.org/10.1007/s11698-023-00266-z>
- Charles, M. (2006). The construction of stance in reporting clauses: A cross-disciplinary study of theses. *Applied Linguistics*, 27(3), 492-518. <https://doi.org/10.1093/applin/aml021>
- Dakos, V., Carpenter, S.R., Brock, W.A., Ellison, A.M., Guttal, V., Ives, A.R., ... y Scheffer, M. (2012). Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. *PloS One*, 7(7), e41010. <https://doi.org/10.1371/journal.pone.0041010>
- Ermakova, L., Bordignon, F., Turenne, N., y Noel, M. (2018) Is the abstract a mere teaser? Evaluating generosity of article abstracts in the environmental sciences. *Front. Res. Metr. Anal.*, 3(16), 1-17. <https://doi.org/10.3389/frma.2018.00016>
- Ferraz de Arruda, H., Nascimento Silva, F., Queiroz Marinho, V., Raphael Amancio, D., y da Fontoura Costa, L. (2018). Representation of texts as complex networks: a mesoscopic approach. *Journal of Complex Networks*, 6(1), 125-144. <https://doi.org/10.1093/comnet/cnx023>
- Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, 486(3-5), 75-174. <https://doi.org/10.1016/j.physrep.2009.11.002>

- Fortunato, S., Bergstrom, C.T., Börner, K., Evans, J.A., Helbing, D., Milojević, S., ... y Barabási, A.L. (2018). Science of science. *Science*, 359(6379), eaa0185. <https://doi.org/10.1126/science.aao0185>
- Groarke, J.M., MacCormac, N., McKenna-Plumley, P.E., y Graham-Wisener, L. (2022). Music listening was an emotional resource and social surrogate for older adults during the COVID-19 pandemic: A qualitative study. *Behaviour Change*, 39(3), 168–179. <https://doi.org/10.1017/bec.2022.10>
- Guerini, M., Pepe, A., y Lepri, B. (2021). Do linguistic style and readability of scientific abstracts affect their virality?. *Proceedings of the International AAAI Conference on Web and Social Media*, 6(1), 475-478. <https://doi.org/10.1609/icwsm.v6i1.14305>
- Hovy, D. (2020). *Text analysis in Python for social scientists: Discovery and exploration*. Cambridge University Press. <https://doi.org/10.1017/9781108873352>
- Hull, D.L. (1990). *Science as a process: An evolutionary account of the social and conceptual development of science*. University of Chicago Press. <https://doi.org/10.7208/chicago/9780226360492.001.0001>
- Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies*, 7(2), 173-192. <https://doi.org/10.1177/146144560505036>
- Jurafsky, D., y Martin, J.H. (2023). *Speech and language processing (3rd ed. draft)*. Stanford University.
- Leydesdorff, L., y Rafols, I. (2011). Indicators of the interdisciplinarity of journals: Diversity, centrality, and citations. *Journal of Informetrics*, 5(1), 87–100. <https://doi.org/10.1016/j.joi.2010.09.002>
- Lopes, L.G., Vieira, T., y Lira, W.W. (2021). Automatic evaluation of scientific abstracts through natural language processing. *arXiv preprint. arXiv:2112.01842*. <https://doi.org/10.48550/arXiv.2112.01842>
- Luján-Villar, J.D. (2018). Ciencias sociales y sostenibilidad: tecnologías de investigación social aplicadas a lo urbano y lo rural. *Revista de Antropología y Sociología: Virajes*, 20(1), 61–81. <https://doi.org/10.17151/rasv.2018.20.1.4>
- Luján-Villar, J.D. (2019). Evolución de la red musical afronorteamericana. *Aposta. Revista de Ciencias Sociales*, 83, 8-29, <http://apostadigital.com/revistav3/hemeroteca/lujanvillar.pdf>
- Luján-Villar, J.D. (2025). ¿Por qué Escuchamos Música? Perspectivas Científicas sobre la Percepción Musical. *Miradas*, 20(2). <https://doi.org/10.22517/25393812.25795>
- McFerran, K.S., y Lotter, C.B. (2024). *The psychology of music listening for health and wellbeing professionals*. Springer Nature. <https://doi.org/10.1007/978-3-031-55373-8>

- Newman, M.E.J., y Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E*, 69(2), 1-15, 026113. <https://doi.org/10.1103/PhysRevE.69.026113>
- Python Software Foundation (2024). *Python* (Versión 3.13) [Software de computación]. <https://www.python.org/>
- Schäfer, K., Saarikallio, S., y Eerola, T. (2020). Music may reduce loneliness and act as social surrogate for a friend: Evidence from an experimental listening study. *Music & Science*, 3, 1-16. <https://doi.org/10.1177/205920432093570>
- Scheffer, M., Bascompte, J., Brock, W.A., Brovkin, V., Carpenter, S.R., Dakos, V., ... y Sugihara, G. (2009). Early-warning signals for critical transitions. *Nature*, 461(7260), 53-59. <https://doi.org/10.1038/nature08227>
- Soleymanpour, R., y Arvaneh, M. (2016). Entropy-based EEG time interval selection for improving motor imagery classification. En *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 004034-004039). IEEE. <https://doi.org/10.1109/SMC.2016.7844864>
- Stanisz, T., Drożdż, S., y Kwapień, J. (2024). Complex systems approach to natural language. *Physics Reports*, 1053, 1-84. <https://doi.org/10.1016/j.physrep.2023.12.002>
- Suzen, N., Mirkes, E.M., y Gorban, A.N. (2020). Informational space of meaning for scientific texts. *arXiv preprint*. [arXiv:2004.13717](https://doi.org/10.48550/arXiv.2004.13717). <https://doi.org/10.48550/arXiv.2004.13717>
- Taipale, M., Carlson, E., y Saarikallio, S. (2025). The role of music listening in the music therapy process for people with anxiety disorders: The therapist's perspective. *Nordic Journal of Music Therapy*, 34(2), 86-104. <https://doi.org/10.1080/08098131.2024.2377974>
- Thompson, G., y Ye, Y. (1991). Evaluation in the reporting verbs used in academic papers. *Applied Linguistics*, 12(4), 365-382. <https://doi.org/10.1093/applin/12.4.365>
- Vega, D., y Magnani, M. (2018). Foundations of temporal text networks. *Applied Network Science*, 3(25), 1-26. <https://doi.org/10.1007/s41109-018-0082-3>
- Wang, M., y Zhu, L. (2023). Linguistic analysis of verb tense usage patterns in computer science paper abstracts. *Journal of Advanced Computing Systems (JACS)*, 3(6), 27-38. <https://doi.org/10.69987/JACS.2023.30603>
- Warren, J. (2024). A path-dependence analysis of Italian social and community cooperatives. *JEOD*, 13(2), 62-96. <http://dx.doi.org/10.2139/ssrn.4739502>
- Watts, D.J. (2011). *Everything is obvious: Once you know the answer*. Crown Currency.