

Análisis lírico-visual y frecuencias léxicas en la música urbana: una propuesta metodológica para el estudio de narrativas audiovisuales  
*Lexico-visual Analysis and Lexical Frequencies in Urban Music: a Methodological Proposal for the Study of Audiovisual Narratives*

---

**Paz Crisóstomo Flores.** Universidad Finis Terrae (Chile)

Académica e investigadora de la Escuela de Periodismo de la Universidad Finis Terrae y coordinadora de la Línea de Análisis e Investigación. Doctora en Investigación en Medios de Comunicación (UC3M). Ha desarrollado proyectos y publicaciones en revistas líderes. Líneas: metodología de la investigación, calidad periodística, tratamiento informativo, televisión y sucesos imprevistos.

ORCID: <https://orcid.org/0000-0002-9761-2112>

**Nicolás Hernández Pinto.** Universidad Finis Terrae (Chile)

Periodista y licenciado en Comunicación Social con Mención en Comunicación Estratégica en la Universidad Finis Terrae. Ganador del Premio de Excelencia Académica 2022 y del Premio de Periodismo Memoria DDHH del Museo de la Memoria, cubrió los Juegos Panamericanos Santiago 2023 y cuenta con estudios en Pedagogía en Lenguaje. ORCID: <https://orcid.org/0009-0004-7741-4575>

**Vicente Higuera Cortés.** Universidad Finis Terrae (Chile)

Periodista y licenciado en Comunicación Social con Mención en Comunicación Estratégica en la Universidad Finis Terrae. Ganador del Premio de Periodismo Memoria DDHH del Museo de la Memoria, cubrió los Juegos Panamericanos Santiago 2023 y fundador de Engira.cl, medio independiente que cubre la actualidad del género urbano chileno en RRSS. ORCID: <https://orcid.org/0009-0003-3784-0572>

Artículo recibido: 18/10/2025 – Aceptado: 21/11/2025

**Resumen:**

Este artículo presenta un modelo léxico-visual para el análisis multimodal de narrativas audiovisuales, integrado por un diccionario léxico, una matriz visual de codificación y una dimensión relacional basada en análisis de redes. La propuesta combina conteo léxico, codificación visual y evaluación de fiabilidad mediante el coeficiente Kappa de Cohen, permitiendo operacionalizar categorías asociadas a representaciones de estatus, riesgo y transgresión. El modelo se valida a través de una aplicación piloto en un corpus acotado de música urbana chilena, utilizada únicamente como demostración de su operatividad. La triangulación entre dimensiones léxicas y visuales y el cálculo de un índice de convergencia visual/léxico permiten evaluar la coherencia interna del instrumento y su potencial para investigaciones futuras. Se discuten los alcances,

limitaciones y posibilidades de transferencia del modelo a otros contextos audiovisuales.

**Palabras clave:**

Modelo léxico-visual; Análisis multimodal; Análisis de contenido audiovisual; Fiabilidad intercodificador; Música urbana chilena.

**Abstract:**

*This article introduces a lexical-visual model for multimodal analysis of audiovisual narratives, combining a lexical dictionary, a visual coding matrix and a relational component based on network analysis. The proposal integrates lexical counts, visual codification and inter-coder reliability assessment through Cohen's Kappa, enabling the operationalization of categories associated with representations of status, risk and transgression. The model is validated through a pilot application in a small corpus of Chilean urban music, used solely to demonstrate the instrument's operability. Triangulation between lexical and visual dimensions and the computation of a simple lexical/visual convergence index allow for evaluating the model's internal coherence and analytical consistency. The article discusses the methodological scope, limitations and potential transferability of the model to other audiovisual contexts.*

**Keywords:**

*Lexical-visual model; Multimodal analysis; Audiovisual content analysis; Inter-coder reliability; Chilean urban music.*

## 1. Introducción

Desde la irrupción del *trap* chileno en escenarios internacionales, la música urbana ha ocupado un espacio central en el debate público sobre juventud, cultura y seguridad. Su narrativa, que conjuga ritmos globales y relatos locales, incorpora temáticas asociadas al consumo y tráfico de drogas, porte de armas, robo y ostentación material (Díez, 2020). En paralelo, diversos artistas han enfrentado procesos judiciales que han amplificado la visibilidad mediática del fenómeno, reforzando la asociación entre música urbana y delito (Ruiz, 2022). Esta cobertura, centrada a menudo en el espectáculo policial más que en la complejidad social, ha consolidado una lectura simplificada y moralizante del género (Deflem, 2020), que requiere ser abordada con herramientas analíticas más sistemáticas.

En este contexto, los discursos mediáticos sobre el género urbano chileno tienden a reproducir estigmas sobre las juventudes populares, reforzando la idea de que la producción cultural funciona como un espacio de riesgo o desviación. Sin embargo, los videoclips y letras de este repertorio también constituyen narrativas estéticas y políticas que resignifican la desigualdad y las jerarquías sociales. La ostentación de bienes, las armas o las marcas de lujo pueden leerse no solo como símbolos de transgresión, sino como códigos de estatus, pertenencia y resistencia cultural. Precisamente por esa ambivalencia, se hace necesario contar con metodologías que permitan traducir estos repertorios en categorías observables y comparables entre distintos productos audiovisuales.

La literatura latinoamericana sobre narcocultura ofrece un marco conceptual útil para este análisis. Mondaca (2012) la define como un proceso cultural que integra símbolos, normas y visiones del mundo en torno a la economía ilegal, proyectando valores y estilos de vida más allá del narcotráfico. Desde este prisma, la narcoestética se configura como una práctica de ostentación visual y discursiva donde el lujo, la violencia y la masculinidad se articulan en torno a la promesa del éxito (Rincón, 2009; Sandoval, 2020). En Chile, esta estética se recontextualiza en barrios urbanos atravesados por tensiones entre exclusión social y consumo aspiracional (Vásquez, 2017), lo que proporciona un repertorio de signos especialmente fértil para ser operacionalizado en categorías léxicas y visuales dentro de un modelo de análisis multimodal.

La presente investigación aborda la apología del delito como fenómeno comunicacional que se expresa simultáneamente en dos planos: el lírico (palabras, frases y campos semánticos que aluden a prácticas delictivas o narcoculturales) y el visual (escenas, objetos, actitudes y símbolos que las representan). Su propósito principal es proponer y validar un modelo metodológico que permita examinar la convergencia entre ambos planos mediante la construcción de un diccionario léxico y una matriz visual de codificación, integrando técnicas de análisis cuantitativo (frecuencias léxicas, concordancia intercodificador) y lectura cualitativa (análisis semiótico y contextual). El foco del artículo reside, por tanto, en la formulación y evaluación de este modelo, más que en la interpretación sustantiva del corpus utilizado como aplicación piloto.

De esta manera, el artículo busca contribuir al campo de los estudios comunicacionales ofreciendo un instrumento replicable que acople la sistematicidad del conteo léxico con la sensibilidad interpretativa de la lectura audiovisual. La propuesta se ilustra mediante una aplicación piloto en un corpus de videoclips y letras de música urbana chilena, utilizado únicamente como escenario de prueba del modelo. El texto se organiza en tres secciones: (i) fundamentos conceptuales y metodológicos para la operacionalización de categorías léxicas y visuales; (ii) diseño del modelo y construcción de los instrumentos; y (iii) validación del modelo a partir de indicadores de fiabilidad y convergencia léxico-visual. Finalmente, se discuten los alcances, limitaciones y posibilidades de transferencia de la metodología a otros contextos y objetos de estudio.

### *1.1. Apología del delito como objeto comunicacional*

El concepto de apología del delito ha sido tradicionalmente definido en el campo jurídico como la exaltación o justificación pública de hechos tipificados como delictivos. En Chile, se encuentra regulado en la Ley N.º 12.927 sobre Seguridad del Estado y en la Ley N.º 20.000 sobre drogas, que penalizan formas de incitación o legitimación del delito. Sin embargo, este marco sancionatorio resulta insuficiente para comprender el fenómeno desde una perspectiva comunicacional y cultural. En los lenguajes mediáticos contemporáneos, la apología no opera como mera incitación directa, sino como narrativización de la transgresión: un repertorio de símbolos, gestos y relatos que expresan experiencias de desigualdad, precariedad y resistencia mediante recursos estéticos y performativos, susceptibles de ser observados y codificados en dimensiones léxicas y visuales.

Como advierte Martín-Barbero (2003), los lenguajes populares articulan formas de conocimiento situadas, donde las culturas juveniles transforman la exclusión en expresividad. Desde una lógica similar, Reguillo (2000) observa que las juventudes urbanas latinoamericanas reescriben el riesgo y el miedo en clave estética, convirtiendo la violencia en una forma de visibilidad y pertenencia. Así, lo que en un registro jurídico se interpreta como apología puede entenderse, en términos comunicacionales, como puesta en escena del límite social: una dramatización simbólica del poder, el deseo y la exclusión que se materializa en palabras, imágenes y performances concretas.

A partir de esta lectura, la música urbana chilena funciona como un dispositivo discursivo en el que la transgresión se estetiza y se codifica. Siguiendo a Hall (1997), toda representación es productiva: los medios no reflejan la realidad, la construyen discursivamente. En este marco, las alusiones a armas, drogas o delitos no implican necesariamente promoción o apología literal, sino configuración de sentido dentro de un campo simbólico donde la marginalidad se reinterpreta como valor cultural y puede ser analizada mediante categorías de contenido específicas.

Este proceso puede vincularse también a la noción foucaultiana de “dispositivo de discurso”, en la que la sociedad produce, regula y recontextualiza aquello que declara prohibido: la visibilidad del delito se convierte en un campo de enunciación donde identidades subalternas reclaman lugar mediante el lenguaje de la transgresión. Autores como Wacquant (2007) y Hartley (2010) aportan una lectura complementaria al mostrar que, en contextos de desigualdad estructural, el delito deja de ser únicamente un acto ilegal para convertirse en forma simbólica de resistencia y reconocimiento en escenarios donde la visibilidad y el consumo operan como nuevas formas de ciudadanía simbólica.

El análisis comunicacional desplaza el foco desde la literalidad del mensaje “peligroso” hacia la función social del discurso: cómo las expresiones artísticas producen significados, disputan estigmas y reconfiguran las fronteras entre lo permitido y lo transgresor. Este giro interpretativo permite operacionalizar la apología del delito en dimensiones observables —léxicas y visuales— que no juzgan el contenido moral, sino que miden su frecuencia, reiteración y función narrativa dentro del ecosistema mediático contemporáneo. Sobre esta base se construyen las categorías del diccionario léxico y de la matriz visual del modelo propuesto, que traducen estas nociones en variables codificables y replicables en el análisis de narrativas audiovisuales.

### *1.2. Narcocultura y narcoestética: entre el valor simbólico y la representación*

La eficacia de la narcocultura, como plantea Mondaca (2012), no reside sólo en relatar el delito, sino en organizar deseos, temores y estatus mediante repertorios estéticos que circulan en música, moda, arquitectura y narrativas audiovisuales. Así, proyecta modelos de éxito anclados en la lógica del riesgo, la ganancia rápida y la visibilidad pública. Dentro de ese entramado, la narcoestética designa la dimensión visual-material del repertorio: una estética del exceso que traduce acumulación (lícita o ilícita) en signos visibles de poder. Rincón (2009) la describe como cultura de la ostentación (“billete, armas, mujeres, música estridente, visaje en autos”), mientras Sandoval (2020) muestra que esta estética disputa los criterios “legítimos” del gusto, proponiendo una

de haber “llegado”.

En Chile, según Vásquez (2017), el país “imprime su sello propio a lo narco” al hibridar desigualdad urbana, consumo aspiracional y pertenencias barriales. Esa especificidad convierte la narcoestética en gramática de visibilidad: una forma de decir “aquí estamos” desde territorios históricamente estigmatizados, donde marcas, joyas, autos y corporalidades funcionan como insignias de reconocimiento. Este fenómeno se potencia en plataformas digitales y videoclips, donde —siguiendo la lógica de la mediatización cultural— el mostrar vale tanto como el decir, y elementos como vehículos, dinero, vestimenta y poses corporales se estabilizan como repertorios reconocibles que pueden ser codificados en categorías visuales.

Ahora bien, no todo es glorificación: la narcoestética convive con ambigüedades (ironía, ficción, estilización) y tensiones (clase, género, raza) que complejizan la lectura. En muchos casos, la violencia se estetiza —se traslada del acto al estilo— y el lujo opera como metonimia de poder más que como evidencia delictiva, lo que obliga a explicitar criterios para decidir qué signos se codifican y cómo se interpreta su función dentro de la narrativa audiovisual.

Para efectos del instrumento, estas nociones se vuelven observables y codificables en dos planos que luego se triangulan: (1) lírico (léxico), mediante campos semánticos asociados a armas, drogas, delitos y lujo, operacionalizados en un diccionario que incorpora lemas, variantes locales, anglicismos y frases compuestas; y (2) visual (imagen), mediante objetos, escenarios, actitudes, vestimenta, vehículos y dinero como indicios materiales de estatus y poder, acompañados de reglas de decisión sobre centralidad narrativa y predominancia (alta, media, baja).

La narcocultura deja de ser una abstracción y se convierte en un sistema medible de signos, apto para el conteo léxico y la codificación semiótica. Esto permite evaluar convergencias (lo que se dice y se muestra) y disonancias (lo que se dice sin mostrarse, o viceversa) no sólo en la música urbana chilena, sino también en otros productos audiovisuales donde circulan imaginarios de riesgo, consumo y poder.

### *1.3. Narrativa lírico-visual: del discurso a la performatividad*

Los productos del género urbano son textos multimodales en los que letra, voz, corporalidad e imagen componen un dispositivo semiótico interdependiente (Kress & Van Leeuwen, 2001). El videoclip no “ilustra” mecánicamente la letra: expande su sentido hacia el terreno del deseo, el cuerpo y la performatividad, generando énfasis, elipsis y redundancias que reconfiguran la narración. En la era de la mediatización, los discursos culturales se organizan bajo lógicas de plataforma y visibilidad: el artista deviene medio de sí mismo, optimizando signos y estilos para circular y conectar con públicos (Hjarvard, 2015), lo que refuerza la necesidad de contar con instrumentos capaces de registrar de manera sistemática esa articulación entre palabra, imagen y performance.

Desde la semiótica cultural, el videoclip es un texto sincrético que articula sistemas verbales, visuales y sonoros en un mismo acto comunicativo (Lotman, 1996). Su

análisis requiere observar la convergencia (refuerzo) y la divergencia (desplazamiento o contradicción) entre lo dicho (léxico) y lo mostrado (imagen). En términos operativos, esto implica distinguir: (1) centralidad narrativa (primeros planos, foco, repetición) versus ambientación (presencias periféricas); (2) índices performativos (gestos, poses, miradas a cámara) frente a indicios objetuales (armas, dinero, autos); y (3) correspondencia diegética (cuando la letra enuncia algo que la imagen realiza) versus desalineación (cuando la imagen estiliza o neutraliza lo enunciado), distinciones que se traducen directamente en variables observables dentro del modelo léxico-visual.

Sobre esa base, la investigación propone una metodología integrada que combina conteo léxico y lectura semiótica para identificar patrones de recurrencia entre campos semánticos (delito, lujo, violencia, drogas) y sus equivalentes visuales (armas, dinero, vehículos, gestos de poder). Operacionalmente: (1) en el plano lírico, el diccionario agrupa lemas y variantes coloquiales por campo semántico y permite medir frecuencia y distribución; y (2) en el plano visual, el libro de códigos registra presencia y predominancia (alta, media, baja) de objetos, escenarios y actitudes, incorporando reglas para casos límite (p. ej., utilería ambigua, símbolos, ironía), de modo que la relación entre letra e imagen pueda ser analizada mediante criterios comparables y replicables.

La doble lectura (léxica-visual) posibilita cuantificar la repetición (frecuencias, proporciones) y cualificar su función narrativa (refuerzo, desplazamiento, contradicción). De este modo, es posible identificar cuándo lo visual potencia lo enunciado en la letra (convergencia) y cuándo lo atenúa o reubica mediante procesos de estetización (disonancia). Más allá del caso específico, el enfoque multimodal propuesto permite pasar de un inventario de “temas” a una cartografía de operaciones discursivas, ofreciendo un procedimiento aplicable a distintos corpus audiovisuales donde se articulan campos semánticos y repertorios visuales bajo lógicas de mediación y performatividad.

#### 1.4. Redes de colaboración y circulación simbólica

El género urbano chileno se sostiene sobre redes de colaboración que operan simultáneamente como estrategia de mercado y régimen de legitimación simbólica. Como señalan Riquelme et al. (2022), estas alianzas producen espacios de cooperación y pertenencia donde el featuring incrementa alcance y consolida comunidades de estilo. En la lógica descrita por Olivares (2023), la circulación digital (*playlists*, *featureings*, sellos colectivos) determina visibilidad y posicionamiento, desplazando el valor desde la obra aislada al entramado de vínculos que la sostiene, lo que hace pertinente incorporar una dimensión relacional en el análisis de estos productos.

En el marco de la narcoestética, la red no solo distribuye música: disemina imaginarios compartidos. Los mismos repertorios de lujo, riesgo o violencia se replican por contagio entre artistas, configurando una ecología simbólica donde signos y poses circulan y se estabilizan como códigos comunes. Este entramado relacional no se limita, por tanto, al plano económico o industrial, sino que también puede ser descrito y analizado a partir de vínculos entre actores y coocurrencias temáticas.

Para efectos de este estudio, el análisis de redes se incorpora como dimensión relacional complementaria al componente léxico-visual:

- (1) Nodos: artistas presentes en el corpus.
- (2) Aristas: colaboraciones acreditadas (A ft. B) en los temas analizados.
- (3) Peso de arista: número de colaboraciones compartidas (o co-ocurrencias en el periodo 2019–2024).

De manera opcional, se puede incorporar una proyección temática mediante una matriz artista por categoría (lujo, armas, drogas, delitos) para estimar afinidades semánticas y observar si las colaboraciones co-varían con determinados campos semánticos.

Con esta base se calculan indicadores estructurales que explican la circulación simbólica:

- (1) Centralidad de grado: visibilidad relacional (quién colabora más).
- (2) Intermediación (betweenness): puentes que conectan subescenas y difunden códigos entre clusters.
- (3) Modularidad/comunidades: identificación de bloques estilísticos (micro-escenas) con repertorios visuales/léxicos similares.
- (4) Densidad y diámetro: cohesión de la escena y alcance potencial del contagio simbólico.

Visualmente, el grafo de colaboraciones permite mapear nodos (artistas) y aristas (vínculos), y superponer atributos semánticos (por ejemplo, color por categoría dominante o tamaño por centralidad) para observar cómo determinados repertorios simbólicos se propagan a través de vínculos creativos y comerciales.

Este enfoque amplía la lectura textual hacia una cartografía del poder simbólico: muestra qué actores anclan códigos dominantes, cuáles median entre comunidades y en qué zonas de la red se concentra la convergencia léxico-visual. En suma, la dimensión relacional del modelo ayuda a explicar por qué ciertos signos se consolidan —no sólo qué signos aparecen— y ofrece un marco transferible para estudiar la difusión de imaginarios en distintas escenas audiovisuales, más allá del caso de la música urbana chilena.

## **2. Modelo léxico-visual: propuesta metodológica**

Esta sección presenta el modelo léxico-visual desarrollado para el análisis multimodal de narrativas audiovisuales. Más que describir una metodología aplicada a un caso específico, el propósito es exponer un instrumento replicable compuesto por un diccionario léxico, una matriz visual y un componente relacional basado en análisis de redes. Su aplicación al corpus de música urbana chilena se entiende como una prueba piloto destinada a validar la claridad operativa, la fiabilidad y la coherencia interna del modelo.

## 2.1. Objetivos de la investigación

### 2.1.1. Objetivo general

Disñar un modelo metodológico mixto para el análisis multimodal de narrativas audiovisuales, integrando un diccionario léxico, una matriz visual de codificación y una dimensión relacional orientada al estudio de redes de colaboración.

### 2.1.2. Objetivos específicos

OE1. Construir un diccionario léxico que operacionalice campos semánticos asociados a representaciones de delito, estatus, lujo y violencia.

OE2. Aplicar una matriz visual que permita identificar objetos, escenarios, actitudes y símbolos relevantes en la narrativa audiovisual.

OE3. Analizar la fiabilidad intercodificador del modelo mediante el coeficiente Kappa de Cohen y ajustes sucesivos del libro de códigos.

OE4. Aplicar el modelo a un corpus piloto de música urbana chilena con el fin de demostrar su operatividad y consistencia analítica.

## 2.2. Diseño y enfoque general

Esta investigación se enmarca en un diseño mixto de análisis de contenido multimodal, que articula estrategias cuantitativas —conteo léxico y frecuencias de aparición— con una aproximación cualitativa de tipo semiótico y discursivo. Se adopta la definición clásica de análisis de contenido de Krippendorff (2018) como “una técnica de investigación para formular inferencias replicables y válidas a partir de los textos y de los contextos de su uso”, pero adaptada a entornos audiovisuales, donde texto, imagen y sonido coexisten en un mismo plano narrativo (Rose, 2016).

El modelo propuesto se concibe como un dispositivo metodológico escalonado, capaz de identificar, medir y correlacionar los signos de la apología del delito en el género urbano chileno. De manera específica, el procedimiento integra tres niveles:

- (1) Nivel léxico: detección y cuantificación de unidades de significado (palabras y frases) vinculadas a armas, drogas, delitos y lujo.
- (2) Nivel visual: registro y categorización de objetos, escenarios, actitudes y emblemas de estatus presentes en los videoclips.
- (3) Nivel relacional: análisis de las redes de colaboración entre artistas, entendido como un plano de circulación simbólica donde se difunden y reafirman los mismos imaginarios.

El enfoque multimodal responde al principio de interdependencia de los códigos (Kress & Van Leeuwen, 2001), según el cual las dimensiones verbal y visual no deben estudiarse aisladamente, sino como partes de una misma estructura de sentido. El corpus de música urbana chilena funciona en este artículo como aplicación piloto, cuyo

propósito es verificar la claridad de las definiciones operativas, la pertinencia de las categorías y la fiabilidad del procedimiento.

### *2.3. Enfoque epistemológico*

El estudio asume que los productos culturales no son meros reflejos sociales, sino formas activas de producción de sentido (Hall, 1997). Bajo esta perspectiva, la música urbana no se analiza como registro documental de la realidad delictiva, sino como texto social, una narración estética que traduce desigualdades, aspiraciones y contradicciones de la vida urbana contemporánea.

Desde esta perspectiva, el modelo integra sistematicidad empírica (frecuencias, distribuciones, fiabilidad) con sensibilidad interpretativa (lectura semiótica, análisis contextual), siguiendo las orientaciones de Jensen (2013) sobre metodologías mixtas en comunicación.

### *2.4. Corpus y criterios de selección*

El corpus está compuesto por 12 videoclips y 12 letras de música urbana chilena publicados entre 2019 y 2024. La selección se realizó mediante muestreo intencional orientado a asegurar variedad de estilos, colaboraciones y repertorios visuales.

Este corpus funciona exclusivamente como aplicación piloto del modelo. Su propósito no es producir conclusiones sustantivas sobre la música urbana, sino evaluar la operatividad del diccionario léxico, la matriz visual y la dimensión relacional. Las decisiones de selección responden, por tanto, a criterios metodológicos (diversidad semántica y visual) más que a representatividad estadística.

### *2.5. Instrumentos*

#### *2.5.1. Diccionario léxico de categorías semánticas*

El diccionario léxico estructura los campos semánticos relevantes del fenómeno estudiado a partir de lematización manual y revisión contextual. Incluye lemas base, variaciones locales (Corporación de la Lengua Española, 2023), anglicismos, préstamos del lunfardo y frases compuestas (Rincón, 2009; Sandoval, 2020).

La organización por campos (delito, armas, drogas, lujo/estatus) permite un conteo replicable y la comparación posterior con los patrones visuales. El diccionario fue sometido a una prueba preliminar con dos codificadores para afinar lemas ambiguos o polisémicos. El conteo se efectuó a partir de la lematización manual y validación automática con scripts en RStudio (versión 2024.09), para identificar frecuencias absolutas y relativas. Esta etapa generó una tabla de ocurrencias que permitió detectar los campos semánticos dominantes y su distribución entre artistas.

### 2.5.2. Matriz visual de codificación

La matriz visual opera como libro de códigos que registra presencia, ausencia y predominancia (alta, media, baja) de objetos, escenarios, actitudes y gestos asociados a repertorios de estatus, riesgo y violencia.

Las categorías incluyen: vestimenta y accesorios de lujo; dinero; vehículos; armas; sustancias; corporalidad; y escenografías asociadas al barrio o al consumo. Se incorporaron reglas de decisión para casos límite (utilería ambigua, simbolismos, ironía, ficcionalización).

La matriz se elaboró siguiendo las recomendaciones de Rose (2016) y Ball & Gilligan (2010) para el análisis de imágenes en medios digitales, garantizando consistencia entre unidades visuales (frame-by-frame) y dimensiones narrativas.

### 2.5.3. Plantilla de codificación

Se diseñó una plantilla que integra el registro léxico, el registro visual y las decisiones de codificación. Esta plantilla permite sistematizar cada unidad audiovisual (video y letra) y facilita el cálculo posterior de convergencia léxico-visual y fiabilidad intercodificador.

## 2.6. Procedimiento analítico

El procedimiento se desarrolló en cuatro fases:

- Extracción y depuración de datos: recopilación de letras (Genius) y videoclips (YouTube) en formato textual y audiovisual; normalización ortográfica y segmentación por versos/escenas.
- Codificación léxica: conteo manual y automatizado de términos; agrupación por lema y categoría semántica según el diccionario elaborado.
- Codificación visual: observación sistemática de escenas según el libro de códigos; registro de predominancia visual (1 = baja; 2 = media; 3 = alta) para cada categoría definida.
- Triangulación: comparación entre frecuencias léxicas y visuales; cálculo de convergencia mediante un índice simple de correlación visual/léxico y análisis cualitativo de casos divergentes para ajustar reglas de decisión y categorías.

Este punto constituye uno de los aportes centrales del modelo, al ofrecer un procedimiento replicable para investigaciones futuras.

La información sobre colaboraciones entre artistas se obtuvo mediante la API pública de Spotify, utilizando llamadas autenticadas al *Spotify Web API (endpoints artist, track y album)* para recopilar metadatos de canciones, *featurings* y enlaces entre intérpretes. Los datos fueron procesados en RStudio (versión 2024.09), empleando las librerías *httr*, *jsonlite* y *igraph* para limpieza, normalización y visualización de la red. Esta etapa permitió construir una matriz artista por artista y generar el grafo de colaboraciones empleado en el análisis relacional.

### 3. Validación del modelo léxico-visual: aplicación piloto

#### 3.1. Panorama general de frecuencias léxicas

El análisis léxico de las 12 canciones arrojó un total de 4.238 unidades léxicas pertinentes al diccionario. De ellas, el 38,4% correspondió al campo “lujo”, seguido de “armas” (24,6%), “drogas” (21,7%) y “delitos” (15,3%). Esta distribución muestra que el diccionario léxico es capaz de discriminar de forma consistente entre campos semánticos diferenciados y de concentrar el registro en las categorías teóricamente definidas como relevantes.

**Tabla 1**  
*Frecuencias léxicas por categoría*

Categoría	Frecuencia absoluta	Porcentaje (%)	Desvío estándar
Lujo	1.627	38,4	12,3
Armas	1.042	24,6	9,7
Drogas	920	21,7	8,9
Delitos	649	15,3	6,8

Desde el punto de vista metodológico, estos resultados confirman la operatividad del instrumento para captar la recurrencia de determinados repertorios semánticos en corpus audiovisuales y ofrecen una base sólida para su comparación posterior con la dimensión visual del modelo.

#### 3.2. Distribución visual de categorías

La observación de los videoclips reveló un patrón visual coherente con el diseño del libro de códigos. La categoría vestimenta/estatus estuvo presente en el 100 % de las obras, seguida de dinero (83 %), vehículos (75 %), armas (58 %), drogas (50 %), escenarios (42 %) y actitud (38 %).

**Tabla 2**  
*Presencia visual por categoría*

Categoría visual	Frecuencia de presencia	Predominancia alta (%)	Concordancia con léxico (%)
Vestimenta / estatus	12	91,6	84,3
Dinero	10	75,0	89,2
Vehículos	9	66,6	81,1
Armas	7	58,3	76,4
Drogas	6	50,0	73,0
Escenarios	5	41,6	69,1
Actitud	4	33,3	63,7

Desde la perspectiva del modelo, esta distribución indica que la matriz visual permite registrar de manera sistemática la presencia y predominancia de las categorías definidas, diferenciando entre elementos centrales de la puesta en escena (vestimenta, dinero, vehículos) y otros de aparición más acotada. Ello confirma la aplicabilidad del instrumento a videoclips con alta densidad de signos y su potencial para estudios comparativos en otros contextos audiovisuales.

### 3.3. Convergencia léxico-visual

Para estimar la convergencia léxico-visual (CLV) se definió un indicador que mide la coincidencia proporcional entre la frecuencia relativa de una categoría en la letra y su presencia operativa en el videoclip. En la aplicación piloto, las categorías presentaron valores diferenciados de CLV, con niveles más altos en “lujo” y “armas” y menores en “delitos” y “drogas”.

Metodológicamente, estos resultados muestran que el indicador de CLV es sensible a variaciones entre categorías y permite detectar patrones de alineación y desalineación entre lo verbal y lo visual. La experiencia sugiere que este tipo de medida puede incorporarse en futuros estudios multimodales como complemento al análisis cualitativo, ofreciendo una forma estandarizada de comparar la relación entre letra e imagen en distintos corpus.

### 3.4. Fiabilidad intercodificador

La fiabilidad intercodificador se evaluó mediante el coeficiente Kappa de Cohen, adecuado para variables categóricas nominales. A diferencia del porcentaje simple de acuerdo, Kappa corrige la coincidencia atribuible al azar, proporcionando una medida más rigurosa del grado de consistencia entre codificadores.

El valor obtenido para la codificación visual fue  $\kappa = 0,776$ , lo que según la clasificación de Landis y Koch (1977) corresponde a un nivel de acuerdo “sustancial”. Este resultado indica que las definiciones operativas son suficientemente claras para producir decisiones convergentes entre analistas independientes.

**Tabla 3**  
*Confiabilidad por categoría*

Categoría	$\kappa$	Interpretación	Notas de adjudicación
Armas	0.86	Muy buena	Consistencia visual alta en escenas explícitas.
Drogas	0.74	Sustancial	Dificultades en distinción “decorado vs. narrativa”.
Delitos	0.71	Sustancial	Requiere anclaje contextual (verso/acción).
Lujo	0.83	Muy buena	Alto acuerdo en objetos y marcas visibles.
Escenarios	0.69	Aceptable	Ajuste de criterios para “entorno narrativo”.

Los valores obtenidos confirman la robustez del instrumento y la claridad de las definiciones operativas. Las diferencias más relevantes se concentraron en escenarios y actitud, donde los codificadores tendieron a interpretar de modo divergente la

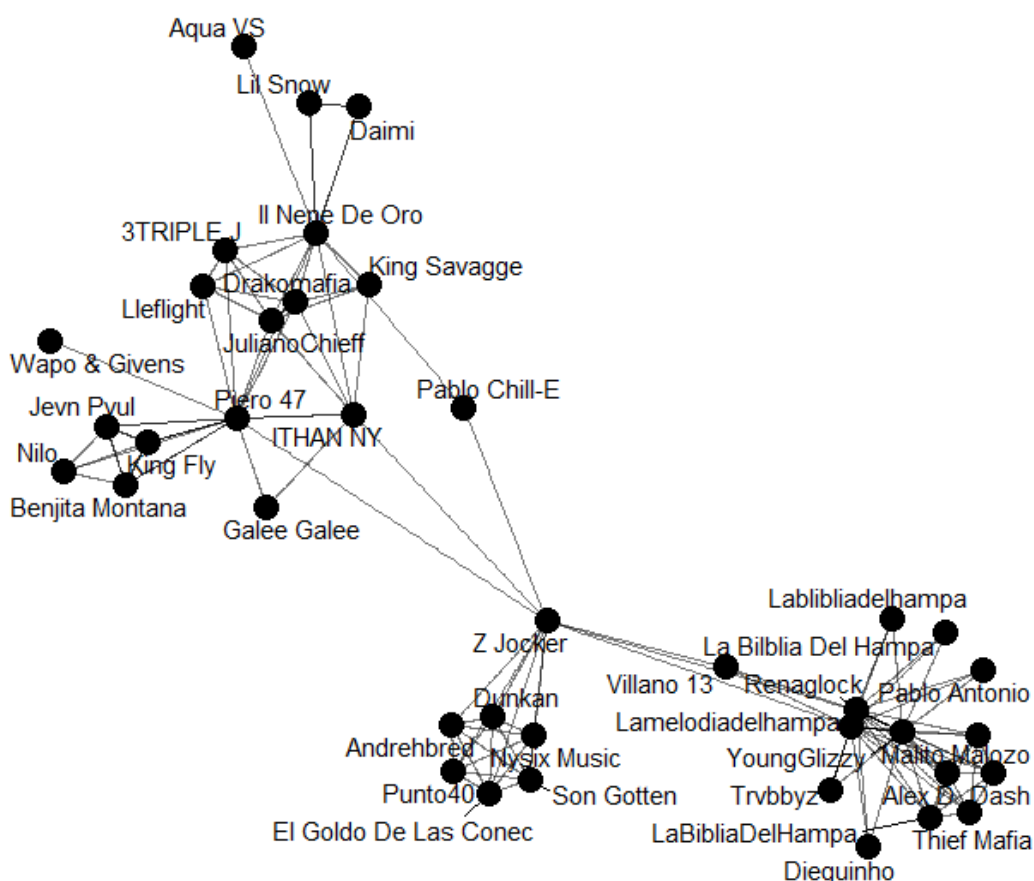
intencionalidad de las posturas o ambientes (p. ej., “barrio”, “fiesta”, “bodega”, “calle”). Tras revisión conjunta, se incorporaron reglas de decisión complementarias (p. ej., “presencia central”, “elemento de apoyo”, “ambientación”).

La validez del modelo se refuerza mediante: (1) la convergencia léxico-visual, que permite comprobar correspondencias entre campos semánticos y repertorios visuales; y (2) la consistencia interna del diccionario, evaluada a través de pruebas preliminares de lematización y selección de variantes.

### 3.5. Análisis relacional: redes de colaboración

El análisis de redes se utilizó en la aplicación piloto para evaluar la dimensión relacional del modelo. A partir de las colaboraciones acreditadas entre artistas se construyó un grafo donde cada nodo representa a un intérprete y cada arista una colaboración registrada en el periodo de estudio.

**Figura 1**  
*Grafo de colaboraciones*



La visualización de la red permitió identificar diferencias en la centralidad de los nodos y superponer información semántica procedente del diccionario léxico (por ejemplo, categorías dominantes asociadas a cada artista), lo que ilustra el potencial del

componente relacional para estudiar la circulación de repertorios simbólicos en escenas audiovisuales. Más que ofrecer conclusiones sustantivas sobre la música urbana chilena, esta etapa muestra que el modelo es compatible con técnicas de análisis de redes y que puede ampliarse hacia enfoques que combinen estructura relacional y patrones multimodales.

## 4. Discusión y Conclusiones

### 4.1. Alcances del modelo léxico-visual

La aplicación piloto del modelo léxico-visual permite reubicar el debate sobre fenómenos como la llamada apología del delito en un terreno primordialmente metodológico. Más que ofrecer una interpretación cerrada del corpus analizado, el estudio muestra que es posible traducir discusiones complejas —delito, narcoestética, desigualdad, estatus— en dimensiones observables que pueden ser codificadas y comparadas de forma sistemática. Al trabajar simultáneamente con campos semánticos y repertorios visuales, el modelo ofrece un marco operativo para analizar cómo se articulan los lenguajes de la transgresión en productos audiovisuales contemporáneos.

Los resultados de la aplicación piloto muestran que el diccionario léxico y la matriz visual son capaces de captar regularidades y matices en la representación de la transgresión, el lujo y el riesgo. La frecuencia diferencial de los campos semánticos y la distribución de las categorías visuales no se presentan aquí como “prueba” de una tesis sustantiva sobre la música urbana chilena, sino como evidencia de que el modelo puede detectar patrones de recurrencia, jerarquías de temas y diferencias en el grado de explicitud entre lo verbal y lo visual. En ese sentido, la contribución principal no reside en lo que el corpus “dice” sobre la apología del delito, sino en cómo el modelo permite observar, medir y comparar tales repertorios.

### 4.2. Convergencias, disonancias y lectura multimodal

La introducción de un indicador de convergencia léxico-visual (CLV) constituye uno de los aportes específicos del modelo. La posibilidad de estimar de manera estandarizada el grado de alineación entre lo que se enuncia en la letra y lo que se representa en la imagen abre un espacio fértil para el análisis multimodal. Las diferencias de CLV entre categorías muestran que el modelo es sensible a la coexistencia de convergencias (cuando la imagen refuerza el campo semántico dominante) y disonancias (cuando lo verbal y lo visual se desajustan o se desplazan).

Este comportamiento no se interpreta aquí como confirmación de una hipótesis sobre la “esencia” del género urbano, sino como demostración del potencial del modelo para mapear operaciones discursivas complejas: intensificación, estilización, atenuación o desplazamiento de determinados contenidos. De este modo, la propuesta dialoga con la tradición multimodal que concibe el videoclip como texto sincrético (Kress & Van Leeuwen, 2001; Lotman, 1996), aportando un procedimiento concreto para cuantificar y cualificar la relación entre los distintos modos de significación.

La doble lectura léxica-visual permite, además, articular de manera más precisa el giro planteado por autores como Hall (1997) o Martín-Barbero (2003): la representación no se limita a “reflejar” la realidad, sino que la construye mediante códigos, géneros y estilos. El modelo léxico-visual ofrece una forma de operacionalizar esta premisa, identificando cómo los signos se organizan en patrones observables y cómo esos patrones pueden compararse entre escenas, géneros o periodos distintos.

#### 4.3. *Articulación con la literatura metodológica previa*

El modelo propuesto se inscribe en la tradición del análisis de contenido (Krippendorff, 2018; Neuendorf, 2017) y de las metodologías visuales (Rose, 2016; Ball & Gilligan, 2010), pero introduce dos desplazamientos relevantes para el campo metodológico. En primer lugar, integra de manera explícita la dimensión multimodal, evitando tratar texto e imagen como objetos separados. La articulación entre diccionario léxico y matriz visual responde al llamado de Jensen (2013) y otros autores a desarrollar metodologías mixtas que reconcilien medición y comprensión en los estudios de comunicación.

En segundo lugar, la incorporación de una dimensión relacional mediante análisis de redes amplía los alcances tradicionales del análisis de contenido. Al trabajar con nodos (artistas u otros actores), aristas (colaboraciones) e indicadores estructurales (centralidad, comunidades, densidad), el modelo permite vincular los repertorios semánticos y visuales con patrones de circulación simbólica. Más que ofrecer un análisis de industria musical, esta dimensión muestra que es posible integrar herramientas de la teoría de redes con el análisis multimodal, generando un campo de diálogo entre enfoques que habitualmente operan por separado.

De esta manera, el modelo no busca sustituir las metodologías existentes, sino articular elementos de varias de ellas en un dispositivo coherente: el rigor del análisis de contenido, la sensibilidad interpretativa de la semiótica cultural y la perspectiva estructural del análisis de redes.

#### 4.4. *Limitaciones del modelo*

La aplicación piloto también deja en evidencia un conjunto de limitaciones que deben ser consideradas. En primer lugar, la codificación visual sigue dependiendo de la interpretación situada de los codificadores. Aunque el coeficiente Kappa de Cohen indica un acuerdo sustancial, las categorías “escenarios” y “actitud” mostraron mayores dificultades de consenso, lo que sugiere la necesidad de seguir refinando las definiciones operativas, incorporar más ejemplos y, eventualmente, desarrollar protocolos de entrenamiento más extensos.

En segundo lugar, el diccionario léxico, aun cuando fue sometido a pruebas preliminares, se construye a partir de un corpus específico y de un contexto lingüístico determinado. Su transferencia a otros contextos culturales o idiomáticos requiere ajustes, revisión de modismos y, probablemente, la incorporación de nuevas categorías. El modelo es replicable, pero no universal: demanda adaptaciones cuidadosas en función de las características de cada escena.

En tercer lugar, la dimensión relacional depende de la disponibilidad y calidad de metadatos en plataformas digitales. Cambios en los sistemas de recomendación, etiquetado o visibilidad podrían afectar la posibilidad de reconstruir redes de colaboración con el mismo nivel de detalle. Asimismo, la integración de la información relacional con la dimensión multimodal requiere decisiones adicionales sobre qué vínculos son relevantes para cada investigación.

Por último, la validación se realizó sobre un corpus acotado, que funciona como prueba de concepto. Futuras investigaciones deberán aplicar el modelo a muestras más amplias y variadas, y contrastar sus resultados con otros enfoques metodológicos para evaluar convergencias, discrepancias y puntos ciegos.

#### *4.5. Potenciales aplicaciones y líneas futuras de investigación*

A pesar de estas limitaciones, el modelo léxico-visual presenta un alto potencial de aplicación en diversos campos de la comunicación. Su estructura permite adaptarlo al análisis de videoclips de otros géneros musicales, campañas publicitarias audiovisuales, productos de ficción seriada, noticias televisivas sobre seguridad o migración, y contenidos nativos de plataformas digitales (reels, shorts, TikTok). En todos estos casos, la combinación de diccionario léxico, matriz visual y dimensión relacional ofrece un conjunto de herramientas para examinar cómo circulan y se reconfiguran imaginarios sobre riesgo, desigualdad, consumo, identidad o ciudadanía.

En el plano técnico, una línea prometedora consiste en explorar la integración del modelo con herramientas automatizadas de análisis de imagen y video (computer vision), así como con técnicas avanzadas de procesamiento del lenguaje natural. Estas incorporaciones podrían agilizar el conteo léxico y el registro visual, manteniendo un núcleo interpretativo cualitativo que supervise y contextualice los resultados.

Asimismo, el modelo podría vincularse con estudios de recepción y participación de audiencias, incorporando métricas de engagement o interacciones en redes sociales. Ello permitiría relacionar los patrones multimodales detectados con la forma en que los públicos interpretan, discuten o reapropian los contenidos audiovisuales, ampliando el enfoque desde la producción textual hacia los procesos de circulación y resignificación.

En síntesis, la validación realizada muestra que el modelo léxico-visual constituye un marco robusto y flexible para el análisis de narrativas audiovisuales. Su principal aporte no radica en resolver definitivamente el debate sobre la apología del delito, la narcoestética o la música urbana, sino en ofrecer un conjunto de herramientas metodológicas que permiten abordar estos y otros fenómenos de manera más sistemática, transparente y comparativa. Al combinar conteo léxico, codificación visual, evaluación de fiabilidad y análisis relacional, la propuesta contribuye a consolidar un campo metodológico capaz de responder a los desafíos que plantean las formas contemporáneas de representación multimodal.

## 5. Referencias bibliográficas

- Ball, M., & Gilligan, C. (2010). *Visual methodologies for digital culture*. SAGE.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46.  
<https://doi.org/10.1177/001316446002000104>
- Corporación de la Lengua Española. (2023). *Diccionario de chilenismos y expresiones coloquiales*. Santiago: CLE.
- Deflem, M. (2020). Popular culture and social control: The moral panic on music labeling. *American Journal of Criminal Justice*, 45(1), 2–24.  
<https://doi.org/10.1007/s12103-019-09495-3>
- Díez, S. N. (2020). *Sobre rap, trap y calle: imágenes y fenómenos*. Kamchatka. Revista de análisis cultural, 16, 93-128. <https://doi.org/10.7203/KAM.16.16627>
- Hall, S. (1997). *Representation: Cultural representations and signifying practices*. SAGE.
- Hartley, J. (2010). Silly citizenship. *Critical Discourse Studies*, 7(4), 233–248.
- Hjarvard, S. (2015). *The mediatization of culture and society*. Routledge.
- Jensen, K. B. (2013). *A handbook of media and communication research: Qualitative and quantitative methodologies* (2nd ed.). Routledge.
- Kress, G., & van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. London: Arnold.
- Krippendorff, K. (2018). *Content analysis: An introduction to its methodology* (4th ed.). SAGE.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. <https://doi.org/10.2307/2529310>
- Ley que Sustituye la Ley N.º 19.366, que Sanciona el Tráfico Ilícito de Estupefacientes y Sustancias Sicotrópicas y Deroga Otras Disposiciones Legales, N.º 20.000, Diario Oficial de la República de Chile, 16 de febrero de 2005.
- Ley sobre Seguridad del Estado, N.º 12.927, Diario Oficial de la República de Chile, 19 de junio de 1958.
- Lotman, Y. (1996). *La semiosfera I: Semiótica de la cultura y del texto*. Cátedra.

- Martín-Barbero, J. (2003). *De los medios a las mediaciones: Comunicación, cultura y hegemonía* (3ª ed.). Anthropos.
- Mondaca, A. (2012). Narcocultura y poder simbólico: entre la violencia y la estética. *Revista de Estudios Culturales Latinoamericanos*, 9(1), 55–72.
- Neuendorf, K. (2017). *The content analysis guidebook* (2nd ed.). SAGE.
- Olivares, P. (2023). Estrategias de visibilidad en la escena urbana chilena: Plataformas, colaboraciones y capital simbólico digital. *Revista Comunicación Contemporánea*, 8(1), 22–39.
- Reguillo, R. (2000). *Emergencia de culturas juveniles: Estrategias del desencanto*. Norma.
- Rincón, O. (2009). *Los narcos también cuentan: Narcocultura y narrativas mediáticas*. Tercer Mundo Editores.
- Riquelme, S., Sule, V., Castillo, V., & Núñez, V. (2022). La periferia urbana a través de los lentes del trap chileno: narrativas juveniles. *Última Década*, 30(59), 140–172. <https://doi.org/10.4067/S0718-22362022000200140>
- Rose, G. (2016). *Visual methodologies: An introduction to researching with visual materials* (4th ed.). SAGE.
- Ruiz, O. (2022). The influence of narcoculture on popular music: A critical look at reggaeton's narco-messages and narco-representations. *Journal of the Royal Musical Association*, 147(1), 77–104. <https://doi.org/10.1017/rma.2022.13>
- Sandoval, M. (2020). La estética del exceso: Cultura visual y narcocapitalismo. *Cuadernos de Comunicación*, 5(2), 89–110.
- Vásquez, M. (2017). *Narcoestética y desigualdad urbana: Crónicas visuales de la periferia chilena*. Ediciones UDP.
- Wacquant, L. (2007). *Los condenados de la ciudad: Gueto, periferias y Estado*. Siglo XXI Editores.

**Financiación:** no se ha recibido financiación.

**Conflicto de intereses:** no existen.

**Traducción al inglés:** aportada por los autores.

#### **HOW TO CITE (APA 7ª)**

Crisóstomo, P., Hernández, N., & Higuera, V. (2025). Análisis lírico-visual y frecuencias léxicas en la música urbana: una propuesta metodológica para el estudio de narrativas audiovisuales. *Comunicación y Métodos | Communication & Methods*, 7(2), 54-72. <https://doi.org/10.35951/v7i2.257>