



REVISTA LATINOAMERICANA DE
METODOLOGÍA DE
LA INVESTIGACIÓN
SOCIAL

**Formas y sustentos en las
búsquedas sociales.
Caminos siempre por
recorrer**

Nº23– AÑO 12

ABRIL - SEPTIEMBRE 2022

PUBLICACIÓN ELECTRÓNICA SEMESTRAL

ISSN 1853-6190

ReLMIS
.com.ar

23

Revista Latinoamericana de Metodología de la Investigación Social



www.relmis.com.ar

Directora:

De Sena, Angélica | CONICET- UNLaM, UBA, CIES, Argentina.

Edición y Coordinación General:

Cervio, Ana Lucía | CONICET-CICLOP-UBA/CIES, Argentina.

Consejo Académico:

Barriga, Omar | *Univ. Concepción, Chile.*

De Sena, Angélica | CIES, UBA, UNLaM, Argentina.

Hamlin, Cynthia | *Univ. Fed. Pernambuco, Brasil.*

Hernández, Marsiela | *Univ. Simón Bolívar, Venezuela.*

Mejía Navarrete, Julio | *Univ. Nac. San Marcos, Perú.*

Padua, Jorge | *COLMEX-CONACYT, México.*

Riella, Alberto | *Univ. De la República, Uruguay.*

Salvia, Agustín | *FSOC-UBA, Argentina.*

Scribano, Adrián | *CONICET-IIGG-UBA/CIES, Argentina.*

Cohen, Néstor | *FSOC-UBA, Argentina.*

Ferreria, Andre | *Univ. Fed. Pernambuco, Brasil.*

Henríquez, Guillermo | *Univ. Concepción, Chile.*

Magallanes, Graciela | *UNVM/CIES, Argentina.*

Mutzenberg, Remo | *Univ. Fed. Pernambuco, Brasil.*

Piovani, Juan | *Univ. Nac. La Plata, Argentina.*

Rivera, Manuel | *Univ. San Carlos, Guatemala.*

Sautu, Ruth | *IIGG-UBA, Argentina.*

Zacarías, Eladio | *Univ. de El Salvador, El Salvador.*

Colaboradores:

Boragnio, Aldana | *CONICET, IIGG- UBA, CIES.*

Chahbenderian, Florencia | *IIGG- UBA, CIES.*

Del Campo, Natalia | *FSOC- UBA.*

D'hers, Victoria | *CONICET- IIGG - UBA, CIES.*

Ferreras, Juan | *CIES.*

Lisdero, Pedro | *CONICET, CIECS- UNC, CIES.*

Diseño de Tapa:

Lucila Salvo

Formas y sustentos en las búsquedas sociales. Caminos siempre por recorrer.

Nº 23. Año 12. Abril – Septiembre 2022.

relmis.com.ar

Publicación electrónica semestral



CENTRO DE INVESTIGACIONES Y ESTUDIOS SOCIOLÓGICOS. Arévalo 2385, piso 8 "D" (CP 1425). Ciudad Autónoma de Buenos Aires, Argentina | E-mail: correo@relmis.com.ar | e-ISSN 1853-6190

Contenido

.Presentación

. Formas y sustentos en las búsquedas sociales.	
Caminos siempre por recorrer	4
<i>Por Angélica De Sena</i>	

. Artículos

. La agencia humana en el decurso de la trayectoria ocupacional	7
<i>Por Ruth Sautu (Argentina)</i>	

. Análise de conteúdo e o uso do Iramuteq	21
<i>Por Rafael Rodrigues Viegas e Natasha Borali (Brasil)</i>	

Procesamiento de Lenguaje Natural aplicado a las ciencias sociales. Detección de tópicos en letras de tango	38
<i>Por Germán Rosati (Argentina)</i>	

. Aplicación de la teoría fundamentada en una investigación sobre travestis y mujeres trans sudamericanas que residen en el AMBA	61
<i>Por Ramiro Nicolás Perez Ripossio (Argentina)</i>	

. Estratégias de divulgação científica nas ciências sociais: escolhas teórico-metodológicas para a criação do website Arquivo Rock Teresina	76
<i>Por Thiago Meneses Alves e Irene Serafino (Brasil)</i>	

. Reseñas de publicaciones

. El método biográfico: estrategias metodológicas para analizar la relación individuo-sociedad en el siglo XXI	89
<i>Por Matías Sbodio (Argentina)</i>	

Presentación:

Formas y sustentos en las búsquedas sociales. Caminos siempre por recorrer

Angélica De Sena

La noción de camino siempre queda atada a la de método y permite colorear a la de metodología. En este caso, aludiremos a las metodologías de la investigación social.

Castillo (1994), al referirse a los modos de enseñar a investigar, incorpora la idea de camino en tanto analogía para dar cuenta de los procedimientos seguidos a lo largo de la investigación social; ello permite evidenciar y poner a prueba los recorridos y las conclusiones a las que se arriba. Ese *camino*, erigido en cada paso que transita espacios de construcción y deconstrucción, obliga a: aprender a preguntar, a observar, a interpretar contextos, relaciones sociales, lenguajes, procesos, realidades sociales diversas en movimiento... Por lo tanto, se impone la necesidad de *reflexionar para investigar e investigar para reflexionar* (Castillo, 1994:21). La noción de camino también le es de utilidad al autor para enfatizar la relevancia de construir y abrir *nuevos caminos* hacia los análisis y explicaciones sociológicas capaces de diseñar mapas, herramientas, gafas y mediaciones para la comprensión de las realidades sociales.

Por otro lado, el término camino fue utilizado como analogía para reflexionar en torno a la metodología de la investigación social desde el abordaje cualitativo, como mapeo de las tierras que se deben transitar desde una actitud que demanda vigilancias epistémicas para recrear sendas ya pisadas, y que permita estar atentos a cambios en conexión con las miradas teóricas. Todo ello, partiendo de que “no hay investigación social sin una reflexiva y relacional tensión entre procedimientos y objetos de indagación (...), por ello la noción de *camino* nos enfrenta con lo que hay de rasgo central en toda indagación: un transitar, un ir de un punto a otro, un pasaje en y a través de un paisaje que se funda y estructura (como toda ciencia) en su complejidad, contingencia e indeterminación” (De Sena, 2015: 13).

Esta presentación de la ReLMIS es una muy buena excusa para reafirmar que el hacer investigación social requiere de policromías, sin atarse a dogmatismos. Siguiendo la perspectiva de Bourdieu et al (2002, 1995), evitar la mera reproducción, con cautela metodológica pero sin obsesiones, siguiendo el imperativo epistemológico (y sociológico) de la reflexividad. Los anclajes y desanclajes del hacer investigación social están atentos a los tiempos/espacios sociales del fenómeno a analizar que dejan espacios para revisar las formas del hacer, narrar, divulgar.

A efectos de detenernos sobre algunos caminos investigativos que toman diferentes posicionamientos teóricos, técnicas y gafas, y que nos invitan a visitar el acto de conocer, el trayecto propuesto en este número 23 inicia con el artículo de Ruth Sautu “*La agencia humana en el curso de la trayectoria ocupacional*”. Aquí la autora parte de considerar la agencia como una expresión de la creatividad y autonomía de las personas para orientar su propia vida. Ello le permite analizar teóricamente el concepto, hurgar cómo se manifiesta a lo largo de la trayectoria ocupacional de dos mujeres y dos varones de clase media residentes en el Área Metropolitana de Buenos Aires, así como observar cómo la agencia humana se filtra en las situaciones de interacción social. La autora da cuenta de los procedimientos metodológicos y concluye que la agencia es intencional, reflexiva, imaginativa-creativa y auto-afirmativa.

El segundo artículo “*Análise de conteúdo e o uso do Iramuteq*”, de Rafael Rodrigues Viegas y Natasha Borali, se entromete con algunos “detalles” de los procedimientos y del análisis de esta estrategia de indagación que mencionan el título, con el objeto de revisar las perspectivas de hacer análisis de contenido desde el abordaje cualitativo como cuantitativo y muestran la versatilidad del mismo.

En tercer lugar, Germán Rosati, con su aporte “*Procesamiento de Lenguaje Natural aplicado a las ciencias sociales. Detección de tópicos en letras de tango*”, muestra que si bien

análisis de letras de tango han sido ampliamente realizados, aún hay nuevos intersticios para observar. Aquí el autor presenta la aplicación de una técnica de Procesamiento de Lenguaje Natural (modelado de tópicos) sobre un *corpus* de letras de tango e introduce un flujo de trabajo posible para el análisis textual computacional y una técnica específica para la detección de tópicos: *Latent Dirichlet Allocation* (LDA). El artículo busca ilustrar las potencialidades que estas técnicas tienen para el análisis de datos textuales en ciencias sociales: su escalabilidad y sus posibilidades de replicabilidad. Se marcan, finalmente, algunas limitaciones de este enfoque.

A continuación, el artículo *“Aplicación de la Teoría Fundamentada en una investigación sobre travestis y mujeres trans sudamericanas que residen en el AMBA”*, de Ramiro Nicolás Perez Ripossio, describe una serie de procedimientos metodológicos aplicados en una investigación empírica cualitativa sobre los proyectos migratorios de travestis y mujeres trans sudamericanas que residen en el Área Metropolitana de Buenos Aires. Se emplea el método de la teoría fundamentada y, de manera auxiliar, el relato de vida. Así, el autor explica el modo en que fue elaborado el muestreo teórico y aplicado el método de la comparación constante. La contribución continúa, junto con otros, el esfuerzo por realizar un aporte respecto al modo de aplicar la Teoría Fundamentada.

En quinto lugar, el trabajo *“Estratégias de divulgação científica nas ciências sociais: escolhas teórico-metodológicas para a criação do website Arquivo Rock Teresina”*, de Thiago Meneses Alves e Irene Serafino, se inmiscuye con la democratización del conocimiento. Para ello, discute las estrategias de divulgación científica en Ciencias Sociales a partir del proceso que culminó en el sitio web Arquivo Roca Teresina. Los autores concluyen respecto a la relevancia de la construcción de instrumentos de divulgación científica para la difusión de los resultados de la investigación a públicos más amplios

Finalmente, Matías Sbodio presenta *“El método biográfico: estrategias metodológicas para analizar la relación individuo-sociedad en el siglo XXI”*, reseña del libro dirigido por Ernesto Meccia, titulado *Biografías y Sociedad. Métodos y Perspectivas*. (2019, EUDEBA); texto que nos convoca a discutir la variedad de formas de llevar adelante investigaciones de corte biográfico.

El recorrido del presente número muestra los diversos caminos de los y las investigadoras, con formas y sustentos para las búsquedas sociales. Y dejan paso a reflexiones para que los nuevos caminos no ocluyan los viejos, y que los viejos no atasquen a los nuevos. Como en un interjuego entre los “viejos” y los “nuevos” parados en los sustentos teóricos y epistémicos que abran nuevas formas y caminos metodológicos a fin de comprender el mundo social del siglo XXI.

Retomando a Adorno (2001), y recordando que al indagar lo social ni su objeto ni su método son unitarios: “Todas las ideas sobre la sociedad entendida como un todo trascienden necesariamente sus hechos dispersos. La construcción de la totalidad tiene como su condición primera un concepto de la cosa en torno al cual se organicen los datos dispares” (Adorno, 2001: 20). Y sin fracturas entre la reflexión teórica y la exigencia empírica. Esto que estamos mencionando como “formas” y “sustentos” quiere también significar que no haya amenazas de “fragmentarse en mera conceptualización formal por un lado, y en recolección de materiales sin concepto, por el otro. En todo el trabajo se trata de crear, entre el elemento de la información y el de la meditación crítica, la relación que una ciencia sociológica como tal debe exigir en no menor medida que la propia conciencia de sus cultores”. (Adorno & Horkheimer, 1969: 7).

Bibliografía

- ADORNO, T. (2001) *Epistemología y Ciencia Sociales*. Fronesis. España: Universidad de Valencia.
- ADORNO, T. & HORKHEIMER, M. (1969) *La Sociedad*. Buenos Aires: Proteo.
- BOURDIEU, P; CHAMBOREDON, J. & PASSERON, J. (2002) *El oficio del sociólogo*. Buenos Aires: Siglo XXI.
- BOURDIEU, P. & WACQUANT, L. (1995) *Respuestas por una antropología reflexiva*. México: Grijalbo.
- CASTILLO, J. (1994) *El trabajo del sociólogo*. Madrid: Editorial Complutense.
- DE SENA, A. (2015) *Caminos cualitativos*. Buenos Aires: CICCUS- Imago-Mundi.

Autora.

Angélica De Sena

Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Universidad Nacional de La Matanza (UNLaM); Universidad de Buenos Aires (UBA); Centro de Investigaciones y Estudios Sociológicos (CIES), Argentina.

Investigadora Independiente de CONICET-UNLaM. Doctora en Ciencias Sociales (UBA), Magister en Metodología de la Investigación Científica (UNLa), Licenciada en Sociología (UBA). Directora de ReLMIS. Profesora en la UNLaM; UBA; USAL. Investigadora del Instituto de Investigaciones Gino Germani (UBA), de la Universidad Nacional de La Matanza (UNLaM) y del Centro de Investigaciones y Estudios Sociológicos (CIES).

E-mail: angelicadesena@gmail.com

Citado.

DE SENA, Angélica (2022). Presentación. Formas y sustentos en las búsquedas sociales: caminos siempre por recorrer. *Revista Latinoamericana de Metodología de la Investigación Social – ReLMIS*, N°23, Año 12, pp. 4-6.



La agencia humana en el decurso de la trayectoria ocupacional

The human agency in the course of the occupational trajectory

Ruth Sautu

Resumen

La agencia es una expresión de la creatividad y autonomía de las personas para orientar su propia vida. El objetivo del artículo es analizar teóricamente el concepto y hurgar cómo se manifiesta a lo largo de la trayectoria ocupacional de dos mujeres y dos varones de clase media residentes en el Área Metropolitana de Buenos Aires (AMBA). Los datos provienen del análisis temático de las transcripciones *verbatim* de entrevistas biográficas; los casos fueron seleccionados entre personas que previamente habían respondido una encuesta aleatoria. La agencia es intencional; se expresa en decisiones y orientaciones de cursos de acción, lo que involucra la movilización reflexiva-creativa de los recursos disponibles y el sentimiento de la propia capacidad de control de la situación. En síntesis, la agencia es intencional, reflexiva, imaginativa-creativa y auto-afirmativa. Aunque el estudio no amerita sacar conclusiones generalizables, es posible intuir que todos los seres humanos poseen, en mayor o menor medida, capacidad de agencia.

Palabras clave: agencia; reflexividad; autonomía; intencional; trayectoria ocupacional.

Abstract

Human agency is an expression of the creativity and autonomy of people to guide their own life. The objective of the article is to theoretically analyse the concept and delve into how it manifests itself throughout the occupational trajectory of two women and two middle-class men residing in the Metropolitan Area of Buenos Aires (AMBA). The data comes from the thematic analysis of the *verbatim* transcripts of biographical interviews; the cases were selected from people who previously responded to a random survey. Agency is intentional; it is expressed in decisions and orientations of courses of action, which involves the reflective-creative mobilization of available resources and the feeling of one's own capacity to control the situation. In synthesis, agency is intentional, reflective, imaginative-creative, and self-affirming. Although the study does not merit drawing generalizable conclusions, it is possible to intuit that all human beings possess a greater or lesser extent of agency capacity.

Keywords: agency; reflexivity; autonomy; intentional; occupational trajectory.

1. Introducción

El análisis de la estructura social y su capacidad explicativa de los comportamientos, relaciones y procesos sociales ha sido siempre una apuesta fuerte en la investigación en ciencias sociales, en particular la sociología, la economía y la educación. Podríamos afirmar, sin temor a equivocarnos, que desde hace más de medio siglo tal hegemonía se ha ido diluyendo para dar paso a enfoques que se apoyan en la interacción social y sus orientaciones socio-psicológicas y culturales e incorporan el tiempo y el espacio en su perspectiva longitudinal.

En esa última perspectiva, el objetivo de este artículo es analizar cómo se manifiesta la agencia humana, en tanto capacidad de autonomía de las personas para orientar su propia vida, en el decurso de sus trayectorias ocupacionales. El marco que subyace al análisis es el de los cursos de vida porque la agencia es dinámica y cambiante en su contexto histórico, económico y social, razón por la cual las entrevistas reconstruyen retrospectivamente las experiencias subjetivas del acceso y desempeño ocupacional.

De una base de encuestas aleatorias se re-entrevistaron 23 personas y se reconstruyeron sus historias de vida. Para este análisis hemos seleccionado cuatro, dos mujeres y dos varones mayores de 55 años, de clase media, residentes en el AMBA. Asumimos que, en términos generales, comparten las experiencias del contexto histórico-político-social; el estado del sistema educativo; las posibilidades y requerimientos de los mercados laborales y, posiblemente, la cultura general de clase, con sus valores, ideas, categorizaciones y modelos de comportamiento. Usaremos las transcripciones *verbatim* de las trayectorias ocupacionales para desentrañar qué es la agencia y cómo se manifiesta en las relaciones sociales; asimismo tendremos en cuenta como las personas reconstruyen e interpretan hoy retrospectivamente su entorno y el contexto socio-histórico. Los cuatro casos son ventanas a través de las cuales, mediante un procedimiento inductivo, intentaremos conocer en qué consiste la agencia humana y cómo se manifiesta en las relaciones a que da lugar el desempeño ocupacional, además de las maneras en que los actores sociales expresan el contexto.

Los alcances del análisis son limitados porque su único objetivo es conceptualizar la agencia y detectar indicios de su presencia en situaciones de la vida laboral. Comenzaremos presentando brevemente la perspectiva general de este análisis: la reconstrucción retrospectiva de experiencias significativas que tuvieron lugar en el curso de sucesos y situaciones de interacción social, y el papel de la agencia humana expresada en ese contexto experiencial. Subsecuentemente, en el análisis de los cuatro relatos identificaremos los temas y cuestiones que los atraviesan, es decir, aquella problemática presente en todos ellos como si se tratara de un texto único. Dentro de estos ejes temáticos trataremos de inferir las maneras en que lo que podríamos identificar como agencia se manifiesta en cada uno de los casos aquí estudiados. En este análisis transcribiremos párrafos *verbatim* en bastardilla. En las conclusiones presentaremos la conceptualización de agencia que, a nuestro entender, puede ser inducida de los relatos. Incluiremos, también, un balance crítico de lo que entendemos son los aportes y limitaciones de este estudio. En el Anexo hemos resumido las trayectorias ocupacionales de nuestros entrevistados.

2. La reconstrucción de experiencias subjetivas de sucesos de la vida aboral

Las personas construyen y cambian la sociedad en que viven (algunos segmentos poblacionales tienen más peso que otros) y el colectivo societal, a su vez, establece condicionamientos y posibilidades que influyen a las personas y familias (Archer, 2007: 10). En esta perspectiva, es de especial interés discutir cómo los actores sociales visualizan y evalúan las condiciones de su entorno, las relaciones sociales en las cuales participan y cómo se conciben a sí mismos y ubican en esas relaciones sociales.

En la narrativa de experiencias de vida, la persona ofrece su interpretación de sí misma, los otros y el contexto (Denzin, 1989: 17). Además del género y la clase social¹, la perspectiva

¹ En nuestro trabajo concebimos a la estructura de clase y las relaciones de clase desde una perspectiva marxista-weberiana que toma en cuenta tanto la inserción en la producción como la posiciones en el

longitudinal-histórica de cursos de vida se apoya en la trilogía edad, generación/cohorte de nacimiento y periodo histórico, los cuales hemos tenido en cuenta en la selección de los cuatro casos. Alwin, citando a Brofenbrenner (1979), los grafica como un conjunto de escenarios/capas incrustados uno en el otro a través de una red de interconexiones. Esos entornos interactúan en los individuos a través de un sistema de roles, relaciones y en un engranaje de actividades que vinculan a los individuos y a la sociedad a lo largo de las trayectorias de vida (Alwin, 1996: 211).

En esa perspectiva relacional, buscamos comprender cómo los actores sociales conciben situaciones o sucesos; muchos, en buena medida, pautados socio-cultural y normativamente. Ellos son como libretos contruidos por la gente y reproducidos y cambiados también por la gente (como cualquier pauta cultural). No obstante, en ese contexto cultural-estructural, es posible ver cómo se expresa la agencia, en tanto autonomía y creatividad humana.

Además del contexto más amplio de la estructura social, económica y cultural, en el análisis de sucesos de la vida cotidiana suele tenerse en cuenta lo que efectivamente pasa y las metas y orientaciones y significados que aparecen activados, así como la naturaleza de la presencia de los actores sociales (Reis, 2012: 13-14). Éstos son rasgos que los actores participantes toman en cuenta, los cuales son movilizados y actuados en el momento de la interacción, ya que ellos saben de alguna manera de qué se trata la situación, así como intuitiva o conscientemente saben de sus propias ideas y sentimientos (Schwarz, 2012: 22).

3. La agencia humana se filtra en las situaciones y procesos sociales

Con la excepción de personas que han perdido el control de sí mismos, la agencia, en un sentido muy amplio, podríamos decir que está presente en todos los comportamientos y relaciones sociales. Según la Enciclopedia Stanford (2019), la concepción estándar de la acción social provee una conceptualización de la agencia en tanto capacidad para actuar intencionalmente, por lo que el ejercicio de la agencia consistiría en el desempeño de acciones intencionales, o vinculadas o derivadas de acciones intencionales; aunque, como lo señala, deben distinguirse clases de agencia según el mayor o menor control y autonomía que denoten.

Una definición ampliamente difundida sostiene que la agencia es, en esencia, una expresión de la creatividad y autonomía de las personas que va marcando sus cursos de vida en un contexto social, estructural y de relaciones sociales (Buss, 2012). En su conceptualización, ya clásica, Emirbayer & Mische (1998) destacan los tres momentos en los cuales:

La agencia humana se filtra en las situaciones de interacción social. La agencia es un proceso subjetivo temporalmente incrustado en las experiencia del pasado (en su aspecto habitual) pero orientado hacia el futuro (como una capacidad de imaginar futuras posibilidades y alternativas) y también hacia el presente (como una capacidad para contextualizar pasados hábitos y futuros proyectos dentro de las contingencias del momento). La dimensión de la agencia de la acción social sólo puede ser capturada en su complejidad completa si es analíticamente situada dentro del flujo del tiempo (Emirbayer & Mische, 1998: 963; la traducción es nuestra).

En síntesis, la agencia es la capacidad de las personas de pensarse a sí mismos y orientar su propia vida dentro de las posibilidades y limitaciones del contexto situacional de los cursos de vida. Los citados autores señalan el carácter repetitivo de las situaciones habituales, cuyo *locus* es la esquematización y aplicación de la experiencia. La agencia se ubica en la imaginación proyectiva de futuras acciones y en “la capacidad de los actores de realizar juicios normativos y prácticos entre posibles alternativas de acción en respuesta a demandas emergentes, dilemas o ambigüedades de la situación actual” (Emirbayer & Mische, 1998: 971, la traducción es nuestra).

La discusión precedente conduce a pensar en dos componentes psicológico-sociales de la agencia. En primer lugar, es un proceso reflexivo del actor social sobre sí mismo. Archer (2007) señala que todas las personas de alguna manera, y con diversa intensidad, *conversa consigo*

mercado de bienes, servicios y trabajo. Asimismo, distinguimos la situación de clase social que establece las chances de vida y los estilos de vida desarrollados en el seno de las clases sociales (Sautu, 2020; 2016).

mismo. En especial, cuando la persona se enfrenta a decisiones sobre cursos de acción o sobre elecciones de todo tipo emergen las experiencias del pasado y las motivaciones e intereses; es una suerte de reactivación imaginaria del pasado que se proyecta y evalúa en el presente (Emirbayer & Mische 1998). Es un razonamiento práctico deliberativo, a veces intuitivo, a veces explícito o solo esbozado.

En segundo lugar, la dimensión psicológica que subyace a la agencia humana es una intrínseca compleja capacidad que permite a las personas actuar sobre su entorno, desplegar su auto-control y desempeñar actividades de diverso tipo en vista a la realización de metas. Sin plantear un enfoque psicológico/individualista, nuestro interés se centra en las personas y familias, y en su capacidad para modelar las circunstancias de su vida, y en ese sentido se asienta en las relaciones sociales. La agencia humana es performativa, e implica un proceso subjetivo interpretativo de la información disponible, y un cierto control sobre el entorno. De allí que lo que Bandura (1997) denomina *self-efficacy* sea un componente a tener en cuenta ya que confronta las actitudes evaluativas de la acción social, sea en términos del azar o de la performance, creatividad y autonomía del *self*².

La concepción individualista de la agencia ha sido criticada por ser sesgada. La concepción neoliberal de agencia presupone en particular un *self* que posee un paquete de capacidades y calificaciones que maneja reflexivamente como si el *self* fuera un negocio (Gershon, 2011). En esta concepción, las motivaciones e intenciones de las personas ganan relevancia, y los análisis tienden a centrarse en el individuo y en las consecuencias de las acciones individuales.

Mientras la concepción individualista de agencia es pertinente en estudios centrados en personas y sus capacidades (por ejemplo, de aprendizaje o en iniciativas empresariales), la perspectiva relacional ubica a las personas en situaciones de “interacción social que son aquellos procesos que comprenden lo que la gente hace cuando ellos se influyen mutuamente” (Turner, 1988: 74).

Una concepción relacional de agencia debe necesariamente tener en cuenta la cultura, sus sistemas de categorías, creencias y valores socialmente compartidos que permiten la interacción social, la cual es influenciada por el entorno inmediato así como también por el contexto macro-social histórico. La interacción social, los círculos o redes de relaciones en las cuales transcurre la vida de las personas no están en el vacío. Las personas portan su individualidad, pero en sus desempeños sus percepciones, juicios y decisiones están en el entorno de las instituciones y cultura socialmente compartida. Biesta, Priestly & Robison (2014) ilustran este punto con referencia al desempeño de los maestros.

Las orientaciones de los actores sociales están modeladas por la circunstancias situacionales ubicadas en el tiempo. “La perspectiva de la agencia en el curso de vida contiene dos aspectos, una forma situada de agencia (el ejercicio de la acción con implicaciones de largo plazo) y la auto-reflejada creencia acerca de la propia capacidad para alcanzar las metas del curso de vida” (Hitlin & Elder, 2007:182; la traducción es nuestra). El análisis de la agencia en su contexto requiere posicionarse en las situaciones específicas de vida, temporal y espacialmente consideradas, y en la interpretación subjetiva de las experiencias de personas con capacidad de agencia (Lvova et. al., 2018).

4. Procedimientos metodológicos

Este artículo es un ejercicio de conceptualización que busca indicios empíricos que permitan analizar la agencia humana. Nuestras preguntas son: ¿Cómo identificar en los relatos las situaciones significativas de la historia laboral? ¿Cuáles son los emergentes textuales que describen su contexto estructural? ¿Cómo se expresa en decisiones y cursos de acción la

² La agencia, en tanto mecanismo psicológico social que se expresa en situaciones de interacción social, comparte con la *self-efficacy* el carácter reflexivo-cognitivo asentado en los sistemas de creencias, valores, normas, disposiciones y orientaciones socialmente compartidas.

capacidad de las personas para orientar su propia vida? ¿Qué estrategias y soluciones emergen en los relatos? ¿Es posible inferir la presencia de la dimensión subyacente de la agencia humana?

En los relatos existe una auto-selección de situaciones significativas para la persona. Ella o él, cuando reconstruyen su historia de vida, marcan aquellas situaciones que son significativas en su propia interpretación del pasado. La tarea de sistematización de las transcripciones *verbatim* de las entrevistas es, probablemente, la que demanda el mayor esfuerzo, ya que sobre el terreno debemos decidir cómo aislar situaciones o sucesos, cómo rastrear en los relatos los rasgos significativos del entorno mostrados por la persona entrevistada y, finalmente, cómo identificar los actores sociales participantes: ¿Todos? ¿Algunos? ¿Cuáles?

El análisis temático completo de todas las transcripciones nos ha permitido aislar aquellos temas que atraviesan los cuatro relatos, dentro de los cuales es posible inferir la presencia de la agencia que se expresa en los modos en los cuales las personas resuelven situaciones de su trayectoria ocupacional. “La cuestión es si puede haber partes de las secuencias de las narrativas (carreras, cursos de vida, etc.) por las que todos tendrían que atravesar” (Abbott, 2001: 19). Es decir, todas las narrativas contienen, de alguna manera, cuestiones que son comunes, como es el caso del ingreso al primer empleo y el de los cambios y secuencias del desempeño laboral. En la reconstrucción retrospectiva de sus experiencias laborales, las cuatro personas entrevistadas destacaron tres ejes temáticos comunes: el papel de los lazos familiares y sociales y del entorno en el acceso ocupacional; la creatividad, resolución de problemas y ejercicio de control en el desempeño ocupacional; y los contenidos simbólicos-evaluativos en su auto-presentación y evaluación.

Teniendo en cuenta esos tres ejes, nuestra principal preocupación ha sido inducir de los relatos los contenidos de la agencia humana y tratar de entender cómo se expresa en situaciones de interacción social. Dentro de los tres ejes temáticos, en el análisis nos hemos detenido en las palabras que las personas eligen para describirse a sí mismas, a los otros y a las situaciones en las cuales les toca vivir, y que ellas/ellos mismos han seleccionado como significativas. En los significados de esas palabras hurgamos indicios sobre su capacidad de agencia.

Si partimos de la idea que la agencia es una disposición/orientación psicosocial a la acción enraizada en esquemas cognitivos culturales imbuidos de emociones, si esta premisa es aceptable, entonces, buscaremos cómo aparece expresada esa disposición/orientación en las auto-presentaciones, en la descripción del entorno, y en las representaciones subjetivamente interpretadas y reconstruidas en los textos. Buscaremos entender sus categorizaciones sociales, que constituyen un mecanismo universal que otorga sentido al mundo social y a sus raíces en los sistemas perceptuales, conceptuales y sociales. En los comportamientos y en la interacción social, los sistemas de categorización permiten identificar y clasificar personas, grupos y objetos materiales, evaluar y explicar conductas, auto-identificarse y establecer cómo la propia identificación se relaciona con los sistemas de representación social (Rhodes & Baron, 2019).

Los sistemas de categorización social son parte constitutiva, como el lenguaje, de la cultura de una sociedad. Es un sistema de ideas y símbolos colectivamente compartidos, generados, sedimentados y re-elaborados históricamente por una sociedad, clase social o población, la que permanece en el tiempo plasmada en productos culturales que adquieren la forma de edificios, objetos, arte, textos, rituales o ceremonias. No es solo un conjunto fijo de normas y pautas o modelos de comportamiento, sino un sistema de significados compartidos que juega un papel central en las disposiciones y comportamientos de las personas (Morling & Kitayama, 2008).

Aun cuando en nuestro análisis identifiquemos categorías cuyo referente son los significados que el propio texto nos lleva a inducir, para comprenderlos debemos insertarlos en las ideas, símbolos y significados comunes prevalentes en la cultura, los que operan como reglas que impregnan comportamientos, desempeños y evaluaciones (Ridgeway & Correl, 2004).

5. Sistematización de los relatos y estrategia de análisis

En términos generales, la agencia es la capacidad de autonomía y creatividad para decidir sobre el curso de nuestras propias vidas en el contexto histórico-social en el cual nos toca vivir. Al

igual que el *self-efficacy*, como parte del *self*, se expresan en modos de pensar y pensarse, emociones, actitudes y comportamientos, los cuales se hallan insertos en situaciones de interacción social.

El análisis de las cuatro historias de vida nos inspiró en nuestros intentos por descubrir cómo la agencia se expresaba y se hacía observable, en qué circunstancias y cuál era su naturaleza: ¿se expresa de manera manifiesta o, por el contrario, aparece sutilmente entretrejida en mayor o menor medida en relatos o comportamientos observables? Reiteradas lecturas de los relatos nos llevaron a pensar que se trata de un proceso psicológico-social continuo que toda persona moviliza en su hacer cotidiano, y que potencialmente se expresa, se hace visible, en el lenguaje, los símbolos y en las interacciones sociales.

En la sistematización de las transcripciones *verbatim* de las entrevistas, comenzamos reconstruyendo la trayectoria educativa y ocupacional de las cuatro personas, además de su lugar de nacimiento y una breve caracterización de su origen de clase. Luego de varias lecturas y análisis de los relatos nos detuvimos y registramos lo siguiente: i.) contenidos verbales, en particular, el uso de verbos y adjetivos, y de metáforas y símbolos; y ii.) descripción de sí mismas/os y de los otros, y de las situaciones de interacción social y su ubicación espacio-temporal. Este análisis fue realizado para cada uno de los relatos separadamente; esto nos permitió revisar cada caso teniendo en cuenta la experiencia de la sistematización de los otros relatos. No buscamos uniformarlos sino incorporar las nuevas ideas que fueron apareciendo en la tarea de sistematización.

Subsecuentemente, mediante el análisis temático de los cuatro casos, identificamos en las secuencias de las narrativas de los cursos de vida laboral los temas que aparecen en todas ellas, los transcribimos en un archivo e identificamos su ubicación en la transcripción completa de cada relato. Esto nos permitió volver a mirar cada tema en el contexto de todos los relatos. Posicionada en cada uno de esos temas, buscamos establecer cuáles eran los indicios por los que se colaba la agencia a través de los relatos. A esta altura retomamos la revisión bibliográfica sobre esos temas, en los cuales se discutía la agencia humana.

La presentación del análisis está ordenada según esos ejes temáticos: el papel de los lazos familiares y sociales y del entorno en el acceso ocupacional; la creatividad, resolución de problemas y ejercicio de control en el desempeño ocupacional; y los contenidos simbólicos-evaluativos en su auto-presentación y evaluación. Apoyaremos nuestro análisis con transcripciones de los relatos que identificamos con letra bastardilla, ligados a nuestros comentarios (es decir, no los transcribimos separados en párrafo aparte).

En el análisis que sigue, no vamos a discutir si existen diferencias entre conjuntos de personas de diferente inserción social. Nuestros datos no ameritan tal análisis. Sí es pertinente citar a Krause (2020)³ quien, en su análisis de la encuesta aleatoria de 1065 personas, señala que: "(...) la atribución de los logros al propio esfuerzo y condiciones personales es generalizada, aunque en menor medida dentro del cuentapropismo informal y los hogares unipersonales y compuestos, donde la suerte adquiere mayor relevancia que en la media" (Krause, 2020:197).

5.1. Clase social y lazos sociales entretrejidos en el acceso al primer empleo

La familia y el barrio, la escuela y los clubes, son las principales fuentes en la cuales se asientan lealtades, confianza y apoyo, en especial, en las edades tempranas de las personas. Constituyen también, a medida que pasa el tiempo, una fuente de identificación social. Todas las personas entrevistadas en este proyecto de historias de vida, incluyendo, por supuesto, las cuatro aquí analizadas, siendo jóvenes, accedieron a su primer trabajo, generalmente informal, a través de familiares o conocidos.

³ Los cuatro casos forman parte de un conjunto de entrevistas en profundidad realizadas a personas que previamente habían respondido la encuesta que analizó Krause.

El negocio de la familia es una manera de comenzar la vida laboral, aunque también en estos casos el vecindario, los conocidos pueden ser para una persona joven fuente de ingresos. Como dice Francisco, antes de ingresar a la universidad, *siempre estuve relacionado al tema del campo...otros trabajos, sí, hice otros trabajos menores, de hecho me había puesto un negocio de revistas en el verano, trabajé de mozo en un restaurant, eh...tareas para tener mi propio dinero.*

El entrenamiento ocupacional temprano de Mariana fue en un negocio de ropa de clase media-alta. En 1982 ingresó al estudio de arquitectura del cual su padrino era el dueño: *habré estado un año y como yo estaba estudiando arquitectura mi papá le pidió a mi padrino que bueno, que me pusiera a trabajar ahí.* También, a través de un familiar, Mariana ingresa a uno de los más importantes canales de televisión abierta; *ahí se abre un mundo para mí. (...) yo no considero un trabajo el de (el negocio de ropa) ni el estudio, cuando empecé realmente a trabajar fue en Canal ZZ que ya yo ahí era una X y tenía que valerme por mí misma y matarme. Por más que mi tío era el gerente y me había hecho entrar, siempre me valí por mí misma en la televisión. Ahí empezó mi carrera y te chupa tanto la televisión y el trabajo que terminé abandonando la carrera de arquitectura.*

También Ana accede a su primer trabajo través de contactos de su papá. *Mi papá tenía mucho contacto en el medio artístico. La oficina tenía tres jefes: uno era representante de jugadores de fútbol, el otro hacía promociones y esas cosas y el otro era un señor que vendía minutos a los canales de todo el interior. Bueno, fue mi primer trabajo. Estuve un año porque era mucho trabajo viste (...) y después mi segundo trabajo fue todo por conocimientos. De un primo mío, él tenía una novia que era la hija de los dueños de una cadena de ropa que en su momento acá era muy importante. Con 20 años le proponen ser encargada porque ella tiene experiencia con números. Y bueno, el jefe de todas las vendedoras me dice vení Ana, vení, ¿Vos te animas? (...) y yo, yo me animo a todo (...) estuve un año ahí bárbaro y después abrieron una sucursal en Santa Fe y Uruguay, y ahí ya era encargada de Lencería y Confección. Trabajaba mucho, ¡era un burro trabajando yo! porque tenía seis vendedoras a cargo, hacer la liquidación de todas las noches, se vendía mucho.*

Los padres de Francisco y de Felipe fueron clase media-alta profesional, propietaria de emprendimientos agropecuarios e industriales. Ambos son graduados universitarios y comenzaron su vida activa participando en los negocios de su familia. Francisco, cuyo padre era europeo, madre argentina, cuando se graduó de médico ingresó a un hospital público y, simultáneamente, ejerció la medicina en clínicas privadas propias y con socios: *en Villa Martelli, con mis socios, teníamos 200 empresas, con eso ya fui dejando después los trabajos en los clubes, los trabajos en otras clínicas.* Felipe, de padre y madre argentinos: *trabajé en la fábrica de mi papá. Ese fue mi primer trabajo. Era una fábrica bastante grande, teníamos más de 100 operarios. Así que inmediatamente empecé a trabajar ahí y pasé por todos los lugares digamos donde se podía estar, la administración, me fue muy bien hasta que las condiciones internacionales hicieron que la fábrica no fuera más viable y hubo que cerrarla en el año 1992.*

Los cuatro relatos nos dicen cómo las personas se vinculan y hacen uso de sus lazos sociales; nos muestran también la presencia de la clase social y de los círculos de relaciones dentro de ellas. Ellas se desempeñan e interactúan en ámbitos específicos, sea en los mercados laborales, en las instituciones o en los ámbitos familiares, de amistad o vecindario. Los lazos existen en esos espacios en los cuales las relaciones y la provisión de información son mutuas basadas en la confianza; que asimismo se influyen mutuamente reforzando la identidad (Espinosa, Rey y Barozet, 2017).

Confianza e identificación se dan en grados, dependen del carácter de las situaciones y quién es quién en la transmisión de la información y recomendaciones, y cuáles son las actividades o intercambios involucrados. El tío puede recomendar a Mariana para que ingrese al Canal de TV; y su primo a Ana para que se desempeñe como encargada de vendedoras. Los padres confían en sus hijos para delegarles responsabilidades que tienen efectos patrimoniales.

Los cuatro casos son personas de clase media que se mueven en espacios sociales diferentes, pero todas, en sus interpretaciones de sus primeras experiencias, muestran un hilo

común⁴. Aunque señalan los lazos sociales que les permitieron acceder a su primer empleo, en todos aparece como una reafirmación de la propia iniciativa; sentirse capaces de desempeñarse en su empleo, ser merecedores de la confianza en sus habilidades que fueron depositadas en ellos/ellas. Esta reafirmación de *self* será una constante a lo largo de toda la entrevista.

5.2. Adaptación con creatividad y ejercicio de control en el desempeño ocupacional

Cuando en 1992 cierra la fábrica de su padre, en la cual había trabajado durante 17 años, Felipe recuerda que *aprovechando que toda la vida yo había montado a caballo y había sido un deportista exitoso en realidad también, pensé que esa era la salida y lo fue. Me dediqué profesionalmente a los caballos; a entrenar caballos, a vender caballos y a los cinco, seis años de eso viendo que por ahí había un poco de inestabilidad con eso, decidí poner un pequeño taller de talabartería que todavía lo tengo*. Reflexivamente, evalúa su situación en el contexto de las reformas económicas neoliberales de los noventa, enumera sus recursos materiales, de información y lazos sociales, y reorienta sus decisiones laborales. La agencia es la capacidad de adaptación creativa a las circunstancias externas, y la resolución de problemas recurriendo a los recursos e información disponible. Además de sus conocimientos, Felipe moviliza también sus lazos sociales dentro del círculo del Club Hípico. Conoce el terreno y usa ese conocimiento para reorientar su vida laboral. Participar de eventos internacionales y vender caballos es su *métie* que se enraíza en su clase social y en los círculos sociales de los que forma parte. Como él mismo lo explica: *voy generalmente a eventos pero como yo transporto caballos, cuando llego a destino la gente que recibe por ahí me tomo unos días, voy a los clubes, visito gente, llevo muestras, (...) se hace toda una relación digamos que me permite vender lo que hago*.

En un círculo social diferente, al igual que Felipe, Ana movilizó los recursos que poseía y reorientó su vida creativamente. Después de dejar su empleo como secretaria ejecutiva (1991 diciembre), se dedicó varios años a la crianza de sus hijos, y en el 2000 comenzó a trabajar terciarizada haciendo pompones de lana para dos empresas. Tuvo conflicto con una de ellas y se desligó mostrando que podía ejercer el control de la situación. *Y un día me cansó y le dije "mira Sigrid vení a buscarte todo porque no te aguanto más", le dije y se llevó todo*. Trabajó por un tiempo con una de las empresas, para eventualmente comenzar a hacer pompones por su cuenta y a vender por Mercado Libre. *Publicaba los pompones de distinta medida a un precio razonable y me empezaron a llamar, a llamar y llamar y tengo del 2000, esperate... Del 2009 ponele al 2013 los pompones que yo he hecho ¡no te puedo explicar! Por Mercado Libre. Fue la primer red que tuve, desde el 2007 que tengo cuenta*.

La capacidad agéntica no funciona en el vacío; consiste en el uso creativo de los recursos disponibles para resolver situaciones nuevas, entre ellas, el desempeño ocupacional. Si efectivamente esto es así, todas las personas entrevistadas poseen en alguna medida esa capacidad. La posee Francisco quien, sin abandonar el ejercicio hospitalario, armó varias clínicas privadas y se organizó para continuar con la administración de sus emprendimientos agropecuarios⁵. También lo mostró Mariana, cuando dejó el primer canal de TV, el ZZ. A través de su relación con un productor fue por un corto periodo a trabajar en un canal público para, eventualmente, acceder a una posición de productora de otro canal privado. En el Canal Público yo estuve *poquito tiempo y esto fue en el 90, en agosto yo me termino yendo porque me llama (menciona el nombre de un gerente-propietario de un Canal)*. La convoca a través de una colega que había trabajado conmigo y le recomendó que me llame a mí, le dijo que yo era muy buena. Por teléfono le dije que no. Le dije *"¿para qué, para ser asistente de producción? no me interesa"*. Ah no y esto fue antes, él me llamó en febrero y yo después me fui con (el gerente del canal público)

⁴ Una concepción marxista-weberiana postula que las clases sociales no son meros sistemas clasificatorios sino entidades ontológicas como son, por ejemplo, la inflación o los mercados. Son construcciones colectivas de existencia real para la gente y, como tales, es posible encontrar diferencias internas.

⁵ El sector salud en Argentina se encuentra altamente fragmentado y con un control creciente de parte de las Obras Sociales y las empresas de Medicina Prepaga. Todo el personal, desde médicos hasta camilleros, con algunas excepciones, desempeñan varios empleos en diversas instituciones públicas y privadas.

como productora y en agosto me vuelve a llamar él y me dice que me quiere hacer una entrevista. Le dije que bueno y entonces fui y tuve una entrevista y era para ser productora de un programa importante. Estos sentimientos de control aparecen en el relato de Mariana cuando se planta frente al gerente-propietario del Canal de TV e informa que solo aceptará el empleo si es designada productora. También lo mostró Ana cuando decide desvincularse de una empresa con la cual estaba en conflicto: *vení no te aguanto más.*

El ejercicio de control y adaptación a las circunstancias del entorno aparece en los relatos como parte de un proceso reflexivo. Cuando recuerda el cierre de la fábrica de su padre, de la cual él era gerente, Felipe reflexiona: *Para mí fue traumático porque yo había nacido con esa fábrica pero también era algo que se veía venir (...) fue terriblemente doloroso pero de repente me encontré con esta "suerte", esta posibilidad de dedicarme a una cosa que me gustaba aún más y así que al poco tiempo sentí una gran liberación y una alegría inmensa.*

El ejercicio de control aparece nuevamente en el relato de Mariana cuando recuerda que en 1996 la nombran productora ejecutiva, y señala la diferencia entre ser productora ejecutiva y su cargo anterior. *Cuando paso a ser Productora Ejecutiva tenía un equipo mucho más grande y ya tengo productores debajo mío a los que mando a hacer cosas pero ellos ya tienen también una independencia y estos le piden a asistentes que hagan el otro trabajo.*

La satisfacción en el ejercicio de control que muestra Mariana aparece también en el relato de Ana, cuando es promovida a una ubicación más importante dentro de las sucursales de la empresa en la cual trabajaba. *Me encantó, porque me sentía re, re importante. Me encantó sí, la verdad que sí, me gustó. Aparte yo siempre digo que me gusta tener responsabilidad. Me encanta viste manejar gente, me encanta.*

Durante la entrevista, Francisco nunca tuvo expresiones que denotaran ejercicio de control. Tal vez porque no fue necesario. Toda su entrevista muestra una persona segura de lo que hace, que llegó a la posición más alta a la que podía esperar en la carrera hospitalaria que es la de director de un gran hospital público. Durante toda la entrevista se mostró seguro de lo que decidía y hacía.

5.3. Los contenidos simbólicos-evaluativos en la representación de sí mismo/a y su entorno

Felipe muestra explícitamente su identidad, y aunque reconoce que su círculo es cerrado, lo reivindica como espacio para toda su familia: *Yo soy directivo en el club y la verdad que brindo mi trabajo ad honorem y me resulta muy grato hacerlo porque la verdad es que el club es como una casa para mí y mis amigos y entonces yo lo hago y eso me devuelve satisfacción en ese sentido (...). Es un poco si querés cerrado, sí, lo es. Pero bueno, la verdad es que se la pasa bien.*

Ana acepta desafíos. Repetidas veces durante la entrevista usa tres frases para describirse a sí misma: aprender, adquirir conocimientos y experiencia, y compromiso con el trabajo duro. Todos los entrevistados muestran compromiso y trabajo duro y competente. Son merecedores de lo que han logrado; sus empleos han llegado naturalmente a ellos. Como lo describe Ana, cuando la designan encargada en la sucursal en Santa Fe y Uruguay: *trabajaba mucho, jera un burro trabajando yo! porque tenía seis vendedoras a cargo, hacer la liquidación de todas las noches, se vendía mucho, ¿viste? y había mucha oferta, mucho esto, mucho el otro.*

Es evidente que tanto Felipe como Francisco trabajan mucho. Sobre todo Francisco que, durante toda su trayectoria, ha desempeñado simultáneamente varios empleos. No obstante, las alocuciones al trabajo duro, a las largas horas en el trabajo, aparecen solo entre las dos mujeres. Trabajar duro, dedicarle muchas horas, comprometerse y aprender forma parte del discurso de Mariana, quien avanzó como asistente, pero en 8 años nunca llegó a productora, aun cuando – como ella nos cuenta – participó como asistente en algunos de los programas más destacados del canal. Fue productora en el Canal público, pero la primera vez que la llamaron para ingresar al segundo canal privado le ofrecieron ser asistente, a lo que ella se negó. Recién después de un tiempo, el canal que se estaba armando le ofreció ser productora. Perseverar, aguantar los contratiempos y seguir adelante es también un componente de la agencia humana.

La posición de director de hospital demanda también compromiso y dedicación de tiempo, aunque Francisco explícitamente no se queja. *Un día de director es...arrancamos antes de las 8 de la mañana, tenemos relación con todo lo que es guardia (...), todo el funcionamiento del tema administrativo del hospital, que trabajan aproximadamente 2.300 personas (...) toda la parte de internados...estamos constantemente en comunicación con todos los jefes de servicio, (...) con los ingenieros, con el tema de la reparación de las cosas que se van rompiendo, se van desgastando, con la parte de obras. Ser director del hospital involucra a todo el hospital. Es importante aquí marcar que Francisco no se queja; tal vez porque se siente seguro de sí mismo, de sus méritos, lo que aparece claramente cuando afirma que a futuro, sea el mismo gobierno o no, me tienen que ratificar o no en el cargo...hay que esperar, creo haber hecho una buena gestión y pienso que me ratificarían, sino volvería a la jefatura de urología que también me gusta mucho. Siempre, a lo largo de la entrevista, Francisco mostró su capacidad de adaptación a los cambios. Como él dice, refiriéndose a sus actividades extra-hospitalarias, uno digamos se va *aggiornando...con los cambios de gobierno (...)* la verdad estoy muy conforme con todo lo que he hecho, y cada etapa me gustó hacerla.*

Felipe, durante toda la entrevista, da la impresión de estar satisfecho con sus logros. Cuando quebró la fábrica de su padre encontró su camino en la venta de caballos y en la talabartería. Parece bastante conforme con su círculo social en el Club Hípico y reconoce: *Yo a veces he sido de clase alta y hoy no. Si bien estoy relacionado mucho con la clase alta (...) hoy soy una clase media-media.* Muy posiblemente esa evaluación esté relacionada con los cambios en su negocio de venta de caballos. *Después del 2008 la cosa cambió y empezó a haber otra prudencia, otra... No tanta alegría para gastar dinero y eso que yo trabajo con gente muy rica, ¿no?* Auto-indulgencia o satisfacción consigo misma/o parecen ser un requisito para el ejercicio de la agencia.

6. A modo de cierre

En la concepción individualista de agencia, el interés se centra en conocer cómo las personas construyen sus espacios de vida psicológicos independientemente de las interacciones interpersonales o lingüísticas. Las situaciones y sucesos son así explicados en términos de los sentidos y valoración de las propias personas actuantes (Ratner, 2000). Esta posición ha sido criticada por su énfasis sesgado en la idea de un *self* de clase media (Skeggs, 2011). Por el contrario, en una concepción relacional, la agencia se expresa en las situaciones de interacción social en el entorno inmediato y en los contextos socio-históricos en los cuales éstas tienen lugar. Las personas procesan, interpretan y responden a las circunstancias situacionales de la interacción en la cual participan (Hitlin & Elder, 2007).

Todas las personas conversan reflexionando consigo mismas, incluso en situaciones habituales. Esta conversación interna es el vínculo con el contexto mediato e inmediato (Archer, 2007). La agencia, como autonomía y creatividad, es parte de esa conversación reflexiva interna cuando se evalúa la situación, se enumeran –a veces intuitivamente– los recursos materiales, de información y lazos sociales, y se reorientan las decisiones laborales.

La agencia es intencional. Se expresa en decisiones u orientaciones de cursos de acción, lo que involucra la movilización creativa de los recursos disponibles y el sentimiento de la propia capacidad de control de la situación, y una cierta confianza en las propias habilidades. En síntesis, la agencia es intencional, reflexiva, imaginativa-creativa, y auto-afirmativa.

Nuestros datos no ameritan sacar conclusiones generalizables. Intuitivamente, diríamos que de alguna manera todos los seres humanos poseemos, en mayor o menor medida, capacidad de agencia. Sin embargo, no sabemos si algunas personas, por clase, género, etnia, generación, etc., están mejor entrenadas en agencia (Appadurai, 2015).

Nuestro estudio ha sido un ejercicio en el cual nuestras/os entrevistadas/os nos muestran, con su reconstrucción selectiva de experiencias laborales pasadas, sus reflexiones hoy acerca de los hilos que sostienen sus auto-interpretaciones de sí mismas/mismos, y de su capacidad para manejar su propio mundo en las circunstancias de su entorno inmediato y de su contexto histórico.

7. Bibliografía

- ABBOTT, A. (2001). *Time Matters. On Theory and Method*. Chicago: The University of Chicago Press.
- APPADURAI, A. (2015). *El futuro como hecho cultural: Ensayos sobre la condición global*. Buenos Aires: Fondo de Cultura Económica.
- ARCHER, M. (2007). *Making our Way through the World*. Cambridge: Cambridge University press.
- ALWIN, D. F. (1996). Taking Time Seriously: Social Change, Social Structure, and Human Lives. En: P. Moen, G.H. Elder Jr. & K. Luscher (Eds.) *Examining Lives in Context. Perspectives on the Ecology of Human Development* (pp. 211–262). Washington DC: American Psychological Association.
- BANDURA, A. (1997) *Self-Efficacy. The Exercise of Control*. New York: Freeman.
- BIESTA, G.; PRIESTLEY, M. & ROBINSON, S. (2015). The role of beliefs in teacher agency, *Teachers and Teaching. Theory and Practice*, 21(6), 624-640.
- BROFENBRENNER, U. (1979). *The ecology of human development*. Cambridge, MA: Harvard University Press.
- BUSS, S. (2012). Autonomous Action: Self-Determination in the passive Mode. *Ethics*, 122 (4), 647-69.
- DENZIN, N.K. (1989). *Interpretative Biography*. Newbury Park: Sage.
- EMIRBAYER, M. & MISCHÉ, A. (1998). What is Agency?. *American Journal of Sociology*, 103 (4), 962-1023.
- ESPINOZA, V.; REY, R.; BAROZET, E. (2017). Incidencia del capital social en la explicación del logro ocupacional. V *Seminario Internacional Desigualdad y Movilidad Social en América Latina*, Ensenada, Argentina.
- GERSHON I. (2011). Neoliberal Agency. *Current Anthropology*, 52 (4), 537-554.
- HITLIN, S. & ELDER Jr G. H. (2007). Time, Self, and the Curiously Abstract Concept of Agency. *Sociological Theory*, 25 (2): 170-191.
- KRAUSE, M. (2020). Auto-interpretaciones de las trayectorias de vida: proyectos y agencia. En: R. Sautu, R.; P. Boniolo, P. Dalle y R. Elbert, R. (Ed.) *Análisis de clases sociales. Pensando la movilidad, la residencia, los lazos sociales, la identidad y la agencia* (pp. 185-214). Buenos Aires: CLACSO-IIGG.
- LVOVA, E. N.; KHAMMATOVA, R. S.; KLEPACH, Y.V.; KAZAKOVA, S. N.; SHUKSHINA, L. V. & DZHAFAROV, N. K. (2018). Enfoques de la investigación de la situación social y su interpretación subjetiva. *Revista espacios*, 39 (20), 1-9.
- MORLING, B. & KITAYAMA, S. (2008). Culture and motivation. In: J.Y. Shah & W. L. Gardner (Eds.) *Handbook of Motivation Science*. New York: The Guilford Press.
- RATNER, C. (2000). Agency and culture. *Journal for the Theory of Social Behaviour*, 30 (4), 413-434.
- REIS, H. T. (2012). Why Researchers should think "Real-World". A conceptual Rationale. In: M.R. Mehl y T. S. Conner (Eds.) *Handbook of Research methods for Studying daily Life* (pp. 3-21). New York: The Guilford Press.
- RHODES, M. & BARON, A. (2019). The Development of Social Categorization. *Annual Review of Development psychology*, 1, 359-386.
- RIDGEWAY, C. L. & CORRELL, S. J. (2004). Unpacking the Gender System: A Theoretical Perspective on Gender Beliefs and Social Relations. *Gender & Society*, 18 (4), 510-531.
- SAUTU, R. (2020). La estructura de clase en el AMBA 2015/2016: Segmentos, fracciones de clase, grupos profesionales y estamentos. En: R. Sautu, R.; P. Boniolo, P. Dalle y R. Elbert, R. (Ed.) *Análisis de clases sociales. Pensando la movilidad, la residencia, los lazos sociales, la identidad y la agencia* (pp. 51-90). Buenos Aires: CLACSO-IIGG.

_____ (2016). *Economía, clases sociales y estilos de vida*. Buenos Aires: Lumiere.

SCHWARZ, N. (2012). Why Researchers Should Think "Real Time". A Cognitive Rationale. In: M.R. Mehl y T. S. Conner (Eds.) *Handbook of Research Methods for Studying Daily Life* (pp. 22-42). New York: The Guilford Press.

SKEGGS, B. (2011). Imagining personhood differently: person value and autonomist working-class value practices. *The Sociological Review*, 59 (3) :496-513.

STANFORD ENCYCLOPAEDIA OF PHILOSOPHY ARCHIVE (2019): Agency. Winter 2019 Edition, First published Mon Aug 10, 2015; substantive revision Mon Oct 28, 2019

TURNER, J. H. (1988). *The Theory of Social Interaction*. Stanford: Stanford University Press.

Anexo: Las trayectorias ocupacionales

La breve síntesis que sigue muestra las fechas y las ocupaciones desempeñadas.

Mariana nació en CABA en 1962. Padres argentinos. Su papá fue vendedor, había trabajado en la tesorería de ISAR (venta de autos, que quebró), en gestión inmobiliaria, y gestión de automóviles. Su mamá, secretaria ejecutiva de un estudio de arquitectos. Terminó la secundaria, hizo cursos de taquigrafía, tenía muy buen francés. 1980, trabajó durante diciembre en un negocio de venta de ropa. 1982, ingresó al estudio de arquitectura del cual su padrino era el dueño. Trabajó allí un año. 1983-1987, canal ZZ como asistente de producción. 1990, trabajó unos meses en un canal público. 1990-2016, Canal NN: productora y Jefa de producción hasta su retiro de la vida activa con jubilación anticipada por razones de salud.

Ana: Nació en CABA en 1957. Padres argentinos. Su papá fue inspector de negocios. Su mamá nació en Entre Ríos y a los 12 o 13 años vino a CABA; educación secundaria completa. En 1982, en el IASE (Instituto Argentino de Secretarías Ejecutiva) hizo el curso de secretaria ejecutiva. 1976, asistente de tres gestores de negocios. 1977- 1980, encargada de vendedoras en la sección lencería y confecciones en sucursales de una gran cadena de tiendas. 1980-1982, secretaria asistente de un abogado laboralista. 1982-1983, oficinista en un sindicato. 1983-1991, secretaria de directorio de un Ente Ejecutivo Represa Embalse (proyecto hidroeléctrico). Son entes interprovinciales que deciden cuestiones generales sobre la construcción y uso de obras públicas hidroeléctricas ubicadas en las provincias. 2000-2007, trabaja terciarizada en la confección de pompones, con dos empresas. 2007-2012, confecciona pompones por su cuenta y vende por Mercado Libre.

Francisco: Nació en CABA en 1959. Su papá nació en Polonia; profesión: médico cirujano de un hospital general de agudos. Su mamá era argentina, hija de italianos, profesora de matemáticas que ejerció unos años, pero después se dedicó a la actividad privada. Sus padres eran dueños de una agencia de autos en una ciudad de la provincia de Buenos Aires. Poseían, además, otros emprendimientos, principalmente agropecuarios. En 1982, ingresó a la facultad de Medicina de La Plata, egresó en 1988. Realizó algunos cursos extra de medicina laboral, auditoría médica, estadística y emergentología. En 1988, ingresó en Urología en un Hospital Público de CABA. Ejerció como médico en clubes, en un supermercado y en una clínica de medicina laboral. Armó con socios varios consultorios privados. En el centro médico ubicado en el conurbano bonaerense, tenían asociadas a 200 empresas. 1997, es jefe de quirófano del Hospital Público. 2005, a la fecha, es asesor de un alto organismo del estado. En 2007, se hizo cargo de un departamento de medicina laboral que constituía una actividad paralela dentro del hospital. En 2014, dejó la jefatura de quirófano y ganó el concurso a la jefatura de urología. En 2015, asume la subdirección del hospital. En 2016, lo designan director que es el cargo que ha ocupado hasta la actualidad.

Felipe: nació en CABA en 1958. Padres argentinos; su papá era contador, dueño de una fábrica de guantes de látex que cerró en 1992; empleaban 100 personas; su mamá era ama de casa. Licenciado en Administración de Empresas por la Universidad de Belgrano, recibido 1982. 1975-1992 trabajó en la gestión de la fábrica de su padre hasta que cerró. 1992 a la fecha: vende y transporta caballos en el exterior, principalmente durante eventos hípicas. Entrena, asimismo, caballos en el club Hípico Argentino. 1997 a la fecha, propietario, diseñador y gestor de una talabartería que ocupa aproximadamente entre 3 y 5 artesanos talabarteros.

Autora.

Ruth Sautu

Instituto de Investigaciones Gino Germani; Facultad de Ciencias Sociales, Universidad de Buenos Aires (UBA), Argentina.

Ph.D. (Economics) Sociology, The London School of Economics and Political Science. Profesora Emérita y Titular Regular de Metodología de la Investigación Social, Carrera de Sociología (UBA). Investigadora del Instituto de Investigaciones GinoGermani, Facultad de Ciencias Sociales (UBA).

E-mail: rsautu@fibertel.com.ar

Citado.

SAUTU, Ruth (2022). La agencia humana en el decurso de la trayectoria ocupacional. *Revista Latinoamericana de Metodología de la Investigación Social – ReLMIS*, N°23, Año 12, pp. 7-20.

Plazos.

Recibido: 13/10/2020. Aceptado: 11/11/2020.



Análise de conteúdo e o uso do Iramuteq

Content analysis and the use of Iramuteq

Rafael Rodrigues Viegas e Natasha Borali

Resumo

A análise de conteúdo é um método bastante utilizado em trabalhos acadêmicos e pesquisas científicas. Contudo, no uso desse método, tem sido evidente a ausência de detalhes suficientes na maioria dos relatórios encontrados na literatura quanto ao procedimento seguido e aos resultados das análises. Neste artigo, examinamos o lugar e a função da análise de conteúdo, buscando superar um pouco as abordagens tradicionais. O objetivo fundamental deste artigo é aumentar o conhecimento, a compreensão e as perspectivas de fazer a análise de conteúdo como método de pesquisa qualitativa e quantitativa, com vistas a promover o seu uso efetivo e criterioso. Para tanto, descrevemos os objetivos, os pontos fortes e fracos da metodologia, os critérios de qualidade, especialmente a triangulação, e, devido ao seu caráter sistemático, ressaltamos a possibilidade do seu emprego em pesquisas apoiadas por computador. Oferecemos dois exemplos específicos do uso de um software livre (Iramuteq) no processo de pesquisa de três decisões do Tribunal de Contas da União e uma do Supremo Tribunal Federal.

Palavras-Chave: Análise de conteúdo; Iramuteq; Tribunal de Contas da União; Supremo Tribunal Federal; Decisão.

Abstract

Content analysis is a method widely used in academic works and scientific research. However, in the use of this method, the absence of sufficient details has been evident in most reports found in the literature, regarding the procedure followed and the results of the analyzes. In this article, we examine the place and function of content analysis, seeking to slightly overcome traditional approaches. The fundamental objective of this article is to increase the knowledge, understanding and perspectives of making content analysis as a qualitative and quantitative research method, with a view to promoting its effective and judicious use. For that, we describe the objectives, the strengths and weaknesses of this methodology, the quality criteria, especially the triangulation, and, due to its systematic character, we emphasize the possibility of its use in computer-supported research. We offer two specific examples of the use of free software (Iramuteq) in the research process of three decisions by the Federal Court of Accounts and one by the Federal Supreme Court.

Keywords: Content analysis; Iramuteq; Federal Audit Court; Federal Court of Justice; Decision.

Apresentação

Um levantamento prévio que fizemos confirmou que a análise de conteúdo (AC) tem sido bastante utilizada em trabalhos acadêmicos e pesquisas científicas. Para se ter ideia, mais de 1.400 dissertações e 290 teses de doutorado, entre 2016 e 2018, foram apresentadas em programas avaliados pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) das áreas de Administração Pública, Ciências Contábeis, Ciência Política, Direito e Economia, com menção à AC. Considerando todas as áreas objeto de avaliação da CAPES no mesmo período, os números saltam para 5.092 dissertações e 1.509 teses que reportam essa metodologia, perfazendo um total de 6.601 resultados (dissertações e teses) concentrados, em sua maioria, nas grandes áreas de ciências humanas (2.245), ciências sociais aplicadas (1.774) e ciências da saúde (1.174). Vale dizer que esse total representa algo em torno de 3% da produção acadêmica (mestrado e doutorado) nacional¹ no período averiguado. No portal de periódicos da CAPES, abordando todas as publicações, foram encontrados 4.570 resultados apenas em português² com referência à AC nesses três anos. Não obstante, identificamos no sistema de busca do *Scielo* mais de 900 artigos publicados que indicam o uso dessa técnica nos seus relatórios de metodologia, nos últimos três anos, no país, em diversas áreas do conhecimento³. Não é pouca coisa, para falarmos apenas dos trabalhos que fizeram constar expressamente a AC.

Contudo, a experiência revela que tem sido evidente a ausência de detalhes suficientes na maioria dos relatórios encontrados na literatura quanto ao procedimento seguido e os resultados das análises de conteúdo. Há indícios, inclusive, de que a análise nem sempre foi utilizada de forma eficaz e criteriosa no processo de pesquisa mesmo por pesquisadores experientes (Bowen, 2009).

Neste artigo, examinamos o lugar e a função da AC, buscando superar um pouco as abordagens tradicionais. O objetivo fundamental deste artigo é aumentar o conhecimento, a compreensão e as perspectivas de fazer AC como método de pesquisa qualitativa e quantitativa, com vistas a promover o seu uso efetivo e criterioso condizente com a realidade atual. É cada vez maior a quantidade de dados à disposição e o seu gerenciamento pode prejudicar a análise e os resultados da própria pesquisa. Para tanto, descrevemos os objetivos, aspectos de categorização, os pontos fortes e fracos da AC e critérios de qualidade que podem ser admitidos, especialmente a triangulação. Devido ao seu caráter sistemático, demonstraremos que a AC qualitativa é também adequada para pesquisas apoiadas por computador. Assim, oferecemos exemplos específicos de uso de um software livre (Iramuteq) no processo de pesquisa de decisões polêmicas do Tribunal de Contas da União (TCU) e do Supremo Tribunal Federal (STF). Com o Iramuteq é possível aplicar diversos testes, mas, dentro dos limites deste trabalho, demonstraremos a nuvem de palavras, a classificação hierárquica descendente (CHD), análise fatorial por correspondência (AFC) e a análise de similitude.

O TCU reprovou a prestação de contas do Executivo Federal (EF) do exercício de 2015. Esse foi o primeiro passo para o julgamento das contas como irregulares pelo Congresso Nacional e, posteriormente, para a instauração do processo de impeachment da Presidente Dilma Rousseff. Desde então, o TCU publicou o parecer do exercício de 2016 referente ao período do governo Dilma Rousseff e ao período do governo Michel Temer. Em 2017, publicou o parecer relativo apenas ao período do governo Michel Temer. Por seu turno, recentemente, em 2019, o plenário do STF reconheceu, por maioria, omissão do Poder Legislativo (Congresso Nacional) por não editar lei que criminalize atos de homofobia. Nessa decisão, o STF criminalizou atos homofóbicos até que o Legislativo edite lei específica. São esses, portanto, os objetos da AC que realizamos com apoio do

¹Catálogos de dissertações e teses da CAPES, disponível em: <https://catalogodeteses.capes.gov.br/catalogoteses/#/>. Filtros de busca aplicados: “mestrado e doutorado”, “2016, 2017, 2018”, “todos”.

² Portal de Periódicos CAPES, disponível em: <https://www.periodicos.capes.gov.br>. Filtros de busca aplicados: “análise de conteúdo”, “2016, 2017, 2018”, “português”, “todos”.

³ Sistema *Scielo*, disponível em: <https://search.scielo.org>. Filtros de busca aplicados: “análise de conteúdo”, “2016, 2017, 2018”, “português”, “todos”.

Iramuteq a título ilustrativo: os três pareceres do TCU (2015, 2016 e 2017) e a decisão do STF que criminalizou atos de homofobia em 2019.

A seguir, o artigo está dividido do seguinte modo: tratamos dos objetivos da AC, história e desenvolvimento; na sequência, lidamos com as abordagens do método, alcances e limitações; elencamos critérios de qualidade que podem ser admitidos enquanto estratégias de controle de qualidade da AC, dando espaço para a triangulação; depois, demonstramos a aplicação do Iramuteq na AC das referidas decisões do TCU e do STF, e; por fim, tecemos considerações finais.

1. Objetivos da análise de conteúdo

O objetivo da AC é o exame sistemático de mensagens de comunicação escritas, verbais ou visuais. Conforme apresentado por Dionísio (2019: 40), esse exame sistemático tem por objetivo:

[...] alcançar uma descrição condensada e ampla do fenômeno, e o resultado da análise são conceitos ou categorias que descrevem o fenômeno. Geralmente, o propósito desses conceitos ou categorias é construir um modelo, um sistema conceitual, um mapa conceitual ou categorias. (Dionísio, 2019: 40, *apud* Cavanagh, 1997).

Por meio da AC, é possível destilar a unidade de análise em categorias relacionadas ao conteúdo (Hsieh; Shannon, 2005). A técnica, que deriva das ciências da comunicação, hoje, no entanto, afirma ser capaz de servir para análise sistemática em uma ampla gama de domínios científicos (Mayring, 2004). Além disso, pode ser usada tanto em bases de dados qualitativos como quantitativos que objetivam descrever e quantificar os fenômenos (Cole, 1988; Downe-Wamboldt, 1992; Elo; Kyngäs, 2007). Os aspectos formais e o conteúdo de significado latente também podem ser objetos de estudo por essa técnica (Mayring, 2004). Uma forma de demonstrar os resultados, o processo de análise e a codificação pode ocorrer por meio de anexos e tabelas, os quais podem ser formulados em softwares (Hsieh; Shannon, 2005; Gerbic; Stacey, 2005). O fundamental, no entanto, é que o material comunicativo, não exclusivamente textual, seja fixado ou registrado de alguma forma (Mayring, 2004).

1.1. História e desenvolvimento

A AC tem uma longa história de uso em comunicação, jornalismo, sociologia, psicologia e negócios. Durante as últimas décadas, seu uso tem mostrado um crescimento constante (Neuendorf, 2017). Sua origem remonta ao século XVIII na Escandinávia, mas sua utilização metodológica começa nos Estados Unidos da América (EUA) no século XX. Inicialmente, a AC foi utilizada como um método qualitativo ou quantitativo, que, posteriormente, se consolidou como um método quantitativo, embora considerada uma técnica simplista pelos críticos do campo quantitativo (Hsieh; Shannon, 2005).

De acordo com Philipp Mayring (2004), o que impulsionou a AC nos EUA foram pesquisas ligadas à comunicação política sobre quantidades de dados textuais da crescente mídia de massa (rádio e jornais). A princípio, segundo o mesmo autor, apenas procedimentos quantitativos foram desenvolvidos, tais como: análises de frequência, nas quais componentes textuais específicos eram contados, p. ex., a frequência com que certos partidos políticos eram mencionados em um jornal; análise de indicadores, onde a frequência de componentes textuais particulares foi definida como um indicador de uma variável superordenada com base em considerações teóricas. Por exemplo, palavras como “deve”, “nunca”, “é” como um indicador do grau dogmatismo de um texto particular; análises de valência e intensidade, que avaliaram o material de acordo com escalas predefinidas, p. ex., quão fortemente os comentários em um determinado jornal expressaram as posições dos partidos atuais no governo; análises de contingência, nas quais foram analisadas as inter-relações entre os diferentes componentes textuais, p. ex., com que frequência, em um jornal, políticos em particular foram mencionados em um contexto direto com atributos positivos (Mayring, 2004: 266-267). Entrementes, nesse contexto, os procedimentos foram limitados ao conteúdo textual em primeiro plano e as estruturas de significado latente foram negligenciadas. A análise ignorou o conteúdo textual que definiu e modificou as unidades textuais específicas. A lógica da análise tinha pouca base linguística e a sistematização e a verificação não podiam ser comprovadas (Mayring, 2004: 267).

Posteriormente, houve repetidas tentativas de desenvolver uma forma qualitativa de AC como alternativa à perspectiva quantitativa, como o procedimento para traçar conteúdos latentes de significados, particularmente conteúdos ideológicos. Foi proposta a técnica para extração e sondagem gradual para a análise de entrevistas abertas e, ao longo do tempo, foi elaborada uma AC qualitativa preparatória para a quantificação. Surgiu a AC estritamente dirigida por hipóteses e a abordagem da codificação de material de entrevistas (Mayring, 2004: 267). De fato, a AC qualitativa atual:

[...] vai além da simples contagem de palavras até o exame intensivo da linguagem, com o objetivo de classificar grandes quantidades de texto em um número eficiente de categorias, que representam significados semelhantes (Hsieh; Shannon, 2005: 1278, tradução livre).

Na sequência, daremos atenção ao emprego dessa metodologia.

1.2. Abordagem do método

A AC é uma técnica versátil e não está limitada por questões ontológicas e epistemológicas que decorrem de diferentes paradigmas. Pelo contrário, o emprego da AC em pesquisas positivistas, pós-positivistas, na teoria crítica e no construtivismo (Guba; Lincoln, 1994; Prior, 2014) amolda-se ao desenho da pesquisa.

Conforme Kaid e Wadsworth (1989), pelo ângulo mais tradicional, podemos resumir a operacionalização do método de AC em sete etapas clássicas: (1) formular as perguntas de pesquisa a serem respondidas; (2) selecionar a amostra a ser analisada; (3) definir as categorias a serem aplicadas; (4) delinear o processo de codificação e o treinamento dos codificadores; (5) implementar o processo de codificação; (6) determinar confiabilidade; (7) analisar os resultados do processo de codificação. Essas etapas são básicas e devem ser assumidas para quem se dispõe a fazer AC, porém, precisam ser adaptadas às diferentes estratégias de pesquisa.

A atividade de coleta de dados, p. ex., muda bastante conforme a estratégia assumida, seja narrativa, fenomenológica, *grounded theory*, etnográfica e estudo de caso (Creswell, 2013). Em relação a qual tipo de informação tipicamente coletada: a estratégia narrativa lida com documentos e material de arquivo, entrevistas abertas, observação participante, conversas casuais; a fenomenológica coleta informações de entrevistas de cinco a 25 pessoas, p. ex. Em relação a como a informação é registrada: as narrativas se valem de notas e protocolos de entrevistas; já estratégias fenomenológicas registram entrevistas, às vezes múltiplas com os mesmos indivíduos; e ainda há outras atividades de coleta e registro para *grounded theory*, abordagens etnográficas e estudos de caso. Em relação a como as informações são tipicamente armazenadas: nas narrativas, os dados são armazenados em pastas de arquivos e arquivos de computador; estratégias fenomenológicas e *grounded theory* armazenam as informações por transcrições e arquivos de computador; nas etnográficas e nos estudos de caso, além das transcrições e dos arquivos de computador, há as notas de campo (Creswell, 2013: 120-121). Todo esse material pode ser objeto da AC, que, conforme dito, precisa ser adaptada às diferentes estratégias de pesquisa.

Philipp Mayring (2004) distingue quatro tipos de procedimentos na AC, são eles: redução, indução, explicação e estruturação. O primeiro tipo, redução, busca reduzir os dados, mas preservando os conteúdos essenciais. A redução é levada a efeito se houver interesse apenas em um nível de conteúdo dos dados e quando se faz necessário condensar o conteúdo. Normalmente, na redução um texto menor e mais fácil de trabalhar é produzido. No segundo tipo, indução, a ideia básica é que o procedimento de resumir seja usado tanto para formação gradual de categorias com referência à definição e ao nível de abstração como para subsunção com base em categorias anteriores ou formação de novas categorias. No terceiro tipo, explicação, a operação é oposta à de redução. Aqui, componentes textuais pouco claros (termos, símbolos, sentenças etc.) necessitam de material adicional para torná-los inteligíveis. A ideia básica da explicação é a coleta sistemática e controlada de material que permita distinguir entre uma análise contextual restrita que envolve apenas o ambiente textual direto e uma ampla análise contextual que coleta material adicional além do texto (informações sobre os comunicadores, sujeito, antecedentes

socioculturais, grupo-alvo etc.). Por último, na estruturação, procura-se filtrar aspectos particulares dos dados e fazer um corte transversal do material sob critérios determinados com antecedência ou para avaliar o material de acordo com critérios específicos. Isso envolve procedimentos formais, focados no conteúdo, de tipologização e dimensionamento, dependendo do tipo de dimensões estruturantes que foram desenvolvidas de acordo com alguma teoria, e estas são então subdivididas em categorias individuais. A ideia básica da estruturação é a formulação exata de definições, passagens textuais típicas (“exemplos-chave”) e regras de codificação que resultarão em um guia de codificação que torna a tarefa de estruturação muito precisa (Mayring, 2004: 267-269).

Tabela 1: tipos de procedimentos

Tipos	Objetivo	Usos
Redução	Reduzir os dados preservando o conteúdo essencial	Certo nível de conteúdo dos dados e quando se faz necessário condensar o material em um texto gerenciável curto.
Indução	Definir categorias, fixando o critério de seleção e nível de abstração para formação de categoria	Formação de categorias a partir dos dados com referência à seleção e nível de abstração; subsunção a partir de categorias anteriores ou formação de novas categorias.
Explicação	Tornar certos elementos textuais, confusos ou pouco claros, inteligíveis	Ampliação de informações sobre comunicadores, sujeitos, antecedentes socioculturais, grupo-alvo.
Estruturação	Filtrar aspectos particulares e fazer um corte transversal ou para avaliar o material de acordo com critérios específicos	Formulação exata de definições, que resultarão em um guia de codificação.

Fonte: Elaborada pelos autores. Fonte: Philipp Mayring (2004).

Autores como Hsieh & Shannon (2005) e Elo e Kyngäs (2007), as duas últimas na esteira de Gioia e Chittipeddi (1991), identificam diferentes processos analíticos de aplicação da AC no momento da codificação, que também variam conforme a abordagem do pesquisador. Reunimos cinco processos elencados por esses autores: convencional, direcionada, somativa, indutiva e dedutiva. Na AC convencional, com base em Hsieh & Shannon (2005), as categorias de codificação são obtidas diretamente dos dados de texto e é geralmente usada em estudos cujo objetivo visa descrever um fenômeno, no qual a teoria existente ou literatura de pesquisa são ainda limitadas. A análise convencional se aproxima da análise indutiva, cujas categorias são derivadas da observação dos dados (Elo; Kyngäs, 2007; Gioia; Chittipeddi, 1991). Destarte, com a existência de teoria ou pesquisa prévias sobre um fenômeno, a descrição é incompleta ou pouco detalhada, por isso utilizam-se conceitos como guias para os códigos iniciais da AC direcionada (Hsieh; Shannon, 2005). Essa abordagem se aproxima da AC dedutiva, que é usada quando a construção analítica é operacionalizada com base na literatura existente e o objetivo do estudo é testar a teoria (Elo; Kyngäs, 2007; Gioia; Chittipeddi, 1991).

Já a AC somativa envolve contagem e comparações de palavras, geralmente de palavras-chave ou conteúdo, com o propósito de compreender o uso contextual das palavras ou do próprio conteúdo, seguida da interpretação (Hsieh; Shannon, 2005). A análise somativa, diferentemente das anteriores, utiliza um processo de mensuração e quantificação das palavras, categorias e códigos, contudo, não para uma leitura exclusivamente quantitativa. Parte-se de uma unidade de análise definida com base no problema de pesquisa e, a depender do posicionamento e objetivos do pesquisador, pode seguir vertentes de análise associada a codificação com base nos próprios dados de análise ou com base na teoria. Na somativa, o processo de codificação permite abstrações e inferências e não se limita à contagem da frequência de palavras, categorias ou códigos (Hsieh; Shannon, 2005; Elo; Kyngäs, 2007).

Para cada um dos processos analíticos, aplica-se uma codificação diferente. Na análise convencional, os códigos são definidos durante a análise de dados e derivados dos dados. Na de conteúdo direcionado, os códigos são definidos antes e durante a análise de dados e derivam da teoria ou de resultados de pesquisas relevantes precedentes. Na somativa, as palavras-chave são identificadas antes e durante a análise de dados e derivadas do interesse de pesquisadores ou revisão de literatura. A abordagem indutiva opera uma codificação aberta definida durante a

leitura de dados, cujos códigos são derivados dos dados. Por fim, a dedutiva começa com uma matriz de categorização (estruturada ou não estruturada), que é identificada antes e durante a análise de dados. Na dedutiva, os códigos são baseados em trabalhos anteriores como teorias, modelos, mapas mentais e revisões de literatura.

Tabela 2: Abordagens de codificação na AC

Tipo de AC	Início	Categorização/fonte	Tempo	Autores
Convencional	Observação	Códigos; Dados	Durante a análise de dados	Hsieh e Shannon (2005)
Direcionada	Teoria	Códigos; Teoria ou de resultados de pesquisas relevantes	Antes e durante a análise de dados	Hsieh e Shannon (2005)
Somativa	Palavra-chave	Palavras-chave; Interesse de pesquisadores ou revisão de literatura	Antes e durante a análise de dados	Hsieh e Shannon (2005)
Indutiva	Observação	Codificação aberta; Dados	Definida durante a leitura de dados	Elo e Kyngäs (2007)
Dedutiva	Teoria	Matriz; Teorias, modelos, revisões de literatura	Antes e durante a análise de dados	Elo e Kyngäs (2007)

Fonte: Elaborada pelos autores.

1.3. Critérios e triangulação

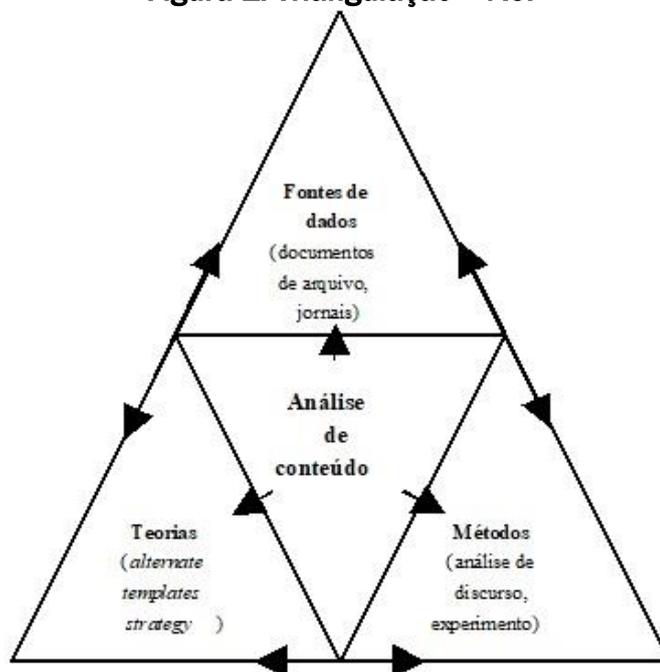
Em qualquer pesquisa científica, a qualidade é garantida pelo uso de critérios apropriados compartilhados pelo campo ou área de conhecimento. Com a AC não é diferente. Referimos a estratégias ou táticas mobilizadas que não representam um conjunto de padrões fixos, mas de condições que ajudam a construir uma pesquisa de qualidade dentro de uma tradição ou paradigma. Pozzebon e Petrini (2013) reúnem critérios tradicionalmente utilizados para avaliar pesquisas positivistas e pós-positivistas, ainda dominantes nas mais diversas áreas de produção de conhecimento. Segundo essas autoras, na tradição positivista, os critérios apontados são: validade interna (o grau no qual os resultados explicam corretamente o fenômeno em questão); validade externa (o grau no qual os resultados podem ser generalizados para outros contextos similares a esse em que o estudo ocorreu); confiabilidade (a medida na qual os resultados podem ser replicados ou reproduzidos por outro investigador); objetividade (a medida na qual os resultados são livres de viés). Na tradição pós-positivista, os critérios são: credibilidade (a “verdade” dos resultados, vista através dos olhos daqueles que estão sendo observados ou entrevistados e dentro do contexto em que a pesquisa é realizada); transferibilidade (a medida na qual os resultados podem ser transferidos para outro ambiente - contextos similares); dependabilidade (a medida na qual a pesquisa poderia produzir resultados semelhantes ou consistentes se realizada como foi descrita); confirmabilidade (pesquisadores precisam fornecer evidências que corroborem os resultados (Pozzebon; Petrini, 2013: 6).

Para efeitos da AC, há uma série de boas diretrizes disponíveis para melhorar os critérios supramencionados, independentemente da tradição de pesquisa (Neuendorf, 2017). Limitamo-nos, aqui, a tratar da triangulação com um pouco mais de afinco.

Existem quatro tipos básicos de triangulação (Denzin, 1978), a saber: por dados (uso de variedade de fontes de dados em um estudo, p. ex.: repositórios de gestão de políticas públicas; entrevistas; documentos de arquivo; periódicos acadêmicos; textos de jornais); por pesquisadores ou avaliadores (o uso de vários pesquisadores ou avaliadores diferentes para equilibrar a influência subjetiva, p. ex.: submeter ao crivo dos pares em eventos acadêmicos e científicos); por teorias (o uso de múltiplas perspectivas teóricas para interpretar um único conjunto de dados, p. ex.: “*alternate templates strategy*” – Langley, 1999: 698), e; por métodos (o uso de múltiplos métodos para demonstrar um único problema ou programa, p. ex.: análise de discurso; experimento) (Patton, 2015: 247; Flick, 2004: 178). Todos esses tipos de triangulação possuem pertinência nos critérios de qualidade acima descritos e são bem-vindos nos relatórios metodológicos dos trabalhos acadêmicos e pesquisas científicas que adotam a AC. Evidente que a adoção de critérios de qualidade pode implicar em maior dispêndio de tempo e de recursos (humanos, financeiros, materiais etc.), portanto cabe ao pesquisador equacioná-los ao campo ou área de conhecimento e

adaptá-los à sua realidade. A seguir, a Figura 1 ilustra um processo de triangulação envolvendo critérios de qualidade que podem ser assumidos na AC.

Figura 1. Triangulação – AC.



Fonte: Elaborada pelos autores.

2. Análise de conteúdo com ajuda do software Iramuteq

Quando as amostras são grandes, é óbvio o benefício da análise por computadores (Pedhazur; Schmelkin, 1991), seja a análise qualitativa ou quantitativa. No campo de ciências sociais e sociais aplicadas, a gama de conteúdo pode assustar, como, p. ex., as centenas de páginas de transcrição de entrevistas, dezenas de pareceres técnicos e decisões judiciais. Isso pode levar o pesquisador a pensar que não conseguirá gerenciar a análise e produzir resultados. Além disso, muitos pontos interessantes e até descobertas surgem quando se analisam os dados (Elo; Kyngäs, 2007), às vezes ampliando ainda mais o material de análise. Nesses casos, programas de computador de análise de dados, como o NUDIST e o NVivo podem e devem ser usados para tornar a AC mais gerenciável e ordenada, permitindo até mesmo novos níveis de análise (Gerbic; Stacey, 2005: 48).

Devido ao seu caráter sistemático, a AC qualitativa é especialmente adequada para pesquisas apoiadas por computador. Não se trata de uma análise automatizada, mas de apoio e documentação dos passos de investigação individuais, bem como de funções de apoio na pesquisa, ordenação e preparação de análises quantitativas. No nosso caso, apresentaremos, a seguir, dois exemplos de uso do Iramuteq para a AC em pesquisas sobre a burocracia. O Iramuteq é um programa de informática livre que viabiliza a análise simultânea de textos pelo cálculo de frequência de palavras⁴. Conforme Camargo e Justo (2013), as possibilidades de análise de dados textuais no Iramuteq são amplas, desde as mais simples, como a lexicográfica básica, até as multivariadas. As análises sobre corpus textuais podem ser estatísticas textuais clássicas (como o cálculo de frequência, média, vocabulário, formas reduzidas, ativas e complementares) e pesquisa de especificidades a partir de segmentação definida do texto (análise de contraste de modalidades

⁴ O programa pode ser baixado no site: <http://www.iramuteq.org>. Existem bons tutoriais disponíveis no mesmo endereço eletrônico, entre eles de Camargo e Justo (2013). Para apresentação dos resultados gerados pelo software (relatório) neste artigo, orientamo-nos pela descrição sugerida por Melo, Cavalcante e Façanha (2019).

de variáveis), como a classificação hierárquica descendente (CHD), a análise de similitude de palavras presentes no texto e a nuvem de palavras (Camargo; Justo 2013). As análises sobre tabelas indivíduos/palavras podem ser CHD conforme algoritmo proposto por Reinert (1987), CHD por matrizes de distância, a análise fatorial por correspondência (AFC), análise de similitude, nuvem de palavras e descrição e X^2 . O corpus textual é produzido pelo próprio pesquisador com base no texto da desgravação de entrevistas, textos de jornal, decisões administrativas e judiciais, códigos legais, jurisprudência, revisão da literatura etc. Pode se referir a um único texto ou a um conjunto desses textos (Camargo; Justo 2013).

Para os fins ilustrativos deste trabalho, demonstraremos a nuvem de palavras, a CHD, AFC e a análise de similitude. Os testes se baseiam na teoria dos grafos (Ratinaud; Marchand, 2012). A nuvem de palavras permite agrupar e organizar e “possibilita rápida identificação das palavras-chave de um corpus” (Camargo; Justo, 2013: 516). A CHD “classifica os segmentos de texto em função dos seus respectivos vocabulários, e o conjunto deles é repartido com base na frequência das formas reduzidas (palavras já lematizadas)” (*idem*). A partir da CHD, o programa possibilita a associação do texto entre palavras, considerando a frequência de incidência de palavras e classes, representando-as no plano cartesiano (AFC) em centro (C) e quatro quadrantes: superior direito (QSD), superior esquerdo (QSE), inferior direito (QID) e inferior esquerdo (QIE). A análise fatorial retoma as frequências de palavras e os valores de correlação chi-quadrado de cada palavra do corpus. A análise de similitude permite “identificar as concorrências entre as palavras e seu resultado traz indicações da conexão entre as palavras, auxiliando na identificação da estrutura de um corpus textual, distinguindo também as partes comuns e as especificidades” (*idem*).

Na sequência, trazemos a AC das referidas decisões do TCU e do STF com suporte do Iramuteq, lembrando, mais uma vez, que o nosso objetivo é meramente exemplificativo das potencialidades do uso desse programa na AC. Em ambos os exemplos, adotamos a técnica de redução de dados (Philipp Mayring; 2004), a análise convencional, pois a codificação inicial tem origem nos dados (Hsieh e Shannon; 2005), e, também, utilizamos a classificação indutiva de Elo e Kyngäs (2007). Destacamos que os gráficos e imagens produzidos pelo software podem ser usados como critérios de triangulação como credibilidade e confirmabilidade (Pozzebon e Petrini; 2013).

2.1. Pareceres do TCU

Temos como premissa que o controle sobre a administração pública é dimensão crucial de uma ordem democrática, tratando-se de uma exigência legal associada ao funcionamento da democracia representativa (Arantes; Loureiro; Couto; Teixeira, 2010: 109). Para suprir a necessidade de controle, frente à sociedade, dos atos e resultados de agente públicos, a Constituição Federal de 1988 (CF-88) contempla uma seção voltada especialmente para o TCU, dentre suas competências está a análise e avaliação da prestação de contas do setor público federal e sua prestação de contas anual.

Entre os processos de auditoria governamental adotadas pelos TC, as auditorias de regularidade são componentes dos controles formais, que “objetivam examinar a legalidade e a legitimidade dos atos de gestão dos responsáveis sujeitos à jurisdição do TC, quanto aos aspectos contábil, financeiro, orçamentário e patrimonial” (TCU, 2011:14). A auditoria de regularidade divide-se em auditoria de conformidade e auditoria contábil. Auditoria contábil é o exame das demonstrações contábeis com o objetivo de expressar uma opinião sobre a adequação desses demonstrativos em relação aos princípios de contabilidade e demais normas (IRB, 2010:13). A auditoria de conformidade ou cumprimento legal é o exame da observância das disposições legais e regulamentares aplicáveis (IRB, 2010:12). O resultado desse processo de auditoria se apresenta ao cidadão na forma de um parecer (opinião) do TCU sobre as contas do EF. Os pareceres contêm aspectos da execução orçamentária e financeira, como a Lei de Responsabilidade Fiscal (LRF, Lei Complementar n. 101/2000), Lei da Contabilidade Pública (Lei n. 4.320/1964), além de limites de gastos impostos pela CF-88 e pela Lei de Licitações e Contratos (Lei n. 8.666/1993) e outros determinantes para a desaprovação das contas. Esse parecer deve se constituir em um julgamento técnico a ser publicado em diário oficial anualmente. Nesse exemplo, analisa-se o conteúdo desses documentos. O TCU reprovou a prestação de contas do EF do exercício de 2015,

conforme Acórdão 1.497/2016-TCU Plenário. Esse foi o primeiro passo para o julgamento das contas como irregulares pelo Congresso Nacional e, posteriormente, para a instauração do processo de impeachment da Presidente Dilma Rousseff. Desde então, o TCU publicou o parecer dos exercícios de 2016 referente a um período do governo Dilma Rousseff e referente ao período do governo Michel Temer. Em 2017 referente apenas ao período do governo Michel Temer.

Submetido os três pareceres do TCU (2015, 2016 e 2017) à AC com ajuda do programa de computador em conjunto, o corpus geral foi constituído por 117 segmentos de texto, com aproveitamento de 62%. Emergiram 4.274 ocorrências, sendo 947 palavras distintas e 426 com uma única ocorrência. A nuvem de palavras nos apresenta frequência de palavras no corpus textual (“União”, “artigo”, “presidência da república”, “lei”, “orçamento” etc.), e serve de indicador da importância das mesmas palavras no conteúdo (Figura 2). Por isso, é a primeira ferramenta do Iramuteq que demonstramos. Observamos que, para fins de categorização, a nuvem permite localizar palavras-chave do corpus e, por conseguinte, dos pareceres em conjunto. Como ponto de partida, a nuvem indica a importância das palavras no conteúdo e sugere ao pesquisador temas centrais.

Realizado o teste CHD, as sete classes obtidas foram divididas em ramificações do corpus em análise, conforme Figura 3 - Dendrograma. Uma primeira codificação foi realizada pelo próprio dendrograma, portanto com uma abordagem indutiva convencional. Outra opção seria uma codificação baseada na teoria ou dedutiva, que categorizaria as classes 1, 2, 3, 4 e 7 - “auditoria de conformidade” (66%), classe 5 e 6 - “auditoria de contábil” (34%). A partir da AFC, realizamos associação entre palavras, considerando a frequência das palavras e as classes (Figura 4). As classes de palavras que reúnem o significado associado ao “parecer prévio” e à organização do “setor público” são encontradas mais no QID, expandindo para C e QSD. A classe que categorizamos como “recomendações e alertas” está no QSE, expandido para C. Já “direitos e garantias” é encontrado em C expandindo para QSE. Não há classe de palavras que se concentra no QIE e pouca presença no limite com QSE.

emitidas 16 recomendações e apenas um alerta, seguido de sete ressalvas sobre o Balanço Geral da União (BGU) e suas demonstrações contábeis, finalizado por dez apontamentos de irregularidade. Dentre os dez apontamentos, seis abordam a mesma temática – a LRF- art. 32, que se aglutina na mesma classe de palavras, mas que poderia ser sintetizado em apenas um apontamento, como feito nos pareceres de 2016 e 2017. Assim, observamos um enfoque demasiado no detalhamento desse conteúdo pelo relator do parecer de 2015, José Múcio Monteiro Filho, presidente do TCU à época.

No exercício de 2016, com relação ao período Presidente Dilma Rousseff, dado o processo de impeachment, não se aplica a auditoria contábil, pois as demonstrações abrangem todo o exercício, assim não apresentam ressalvas sobre balanço, e as contas são aprovadas com quatro ressalvas associadas a execução orçamentária. No mesmo ano de 2016, as contas referentes ao período do governo Michel Temer foram aprovadas com 31 recomendações, o dobro de recomendações do exercício de 2015, mais quatro alertas, sendo apenas um alerta emitido em 2015, e oito ressalvas sobre o BGU e seis ressalvas, incluídas as mesmas ressalvas de execução orçamentária apresentadas em 2016 no mandato da Presidente Dilma. As ressalvas referentes às demonstrações contábeis são atribuídas unicamente à prestação de contas do Presidente Temer, por ser o responsável no momento de encerramento do BGU. As ressalvas apresentadas sobre o BGU de 2016 nos levam a inferir que as demonstrações contábeis apresentadas não estão condizentes com as regras de escrituração contábil da lei de contabilidade pública, pois seu conteúdo não gera insumos para tomada de decisão, apresentando informações superavaliadas com relação ao patrimônio líquido, subavaliadas em termos de ativos e passivos, e, por fim, com classificações indevidas, como apontado o parecer. Novamente, é tênue a distinção entre o conteúdo apresentado como irregularidade e como ressalva. Em 2017, por fim, a prestação de contas do Presidente Temer foi aprovada com 24 recomendações e quatro alertas, novamente uma quantidade superior as de 2015, com sete ressalvas relacionadas ao balanço geral e suas demonstrações contábeis, e cinco ressalvas de execução orçamentária.

A AC nos permitiu identificar que ao pareceres têm um forte viés para a auditoria de conformidade, contudo, se a raiz das escriturações contábeis também são pautadas em leis, uma escrituração indevida que afete o balanço geral e a integridade de uma demonstração também deveria ser entendida como uma irregularidade de auditoria de conformidade, que afeta o erário na medida que impossibilita uma leitura fidedigna da realidade financeira e orçamentária das contas da União, impedindo ou dificultando o controle social. A utilização do Iramuteq como software de auxílio da AC nos mostra que as classes de palavras apresentadas nos pareceres de prestação de contas emitidos pelo TCU são bastante próximas e tendem fortemente à auditoria de conformidade.

2.2. Decisão do STF na ADO n. 26 DF: criminalização da homofobia

Em 13 de junho de 2019, o plenário do STF⁵ reconheceu omissão do Poder Legislativo por não editar lei que criminalize atos de homofobia⁶. Submetido o voto do relator, ministro Celso de Mello, à AC com ajuda do Iramuteq, o corpus geral foi constituído por 104 segmentos de texto, com aproveitamento de 70%. Emergiram 3.599 ocorrências, sendo 1.371 palavras distintas e 960 com uma única ocorrência. A nuvem de palavras obtida nos apresenta a frequência de palavras no corpus textual (“direito”, “como”, “gênero”, “violência”, “lei”, “penal” etc.) (Figura 6). Observamos que, para fins de categorização, a nuvem permite localizar palavras-chave do corpus e, por conseguinte, do voto do relator. Como ponto de partida, a nuvem serve de indicador da importância das mesmas palavras no conteúdo e sugere ao pesquisador temas centrais.

⁵ O plenário do STF é composto pelos 11 ministros do Tribunal. Nesse caso, todos participaram do julgamento. A Ação Direta de Inconstitucionalidade por Omissão (ADO) é uma ação que visa tornar efetiva uma norma constitucional. A falta de norma regulamentadora pode ensejar que dispositivos da CF fiquem sem produzir efeitos.

⁶ “Declarada a inconstitucionalidade por omissão de medida para tornar efetiva norma constitucional, será dada ciência ao Poder competente para a adoção das providências necessárias e, em se tratando de órgão administrativo, para fazê-lo em trinta dias” (CF, 1988).

Figura 6 – ADO 26 - STF - nuvem



Figura 7 – Dendrograma

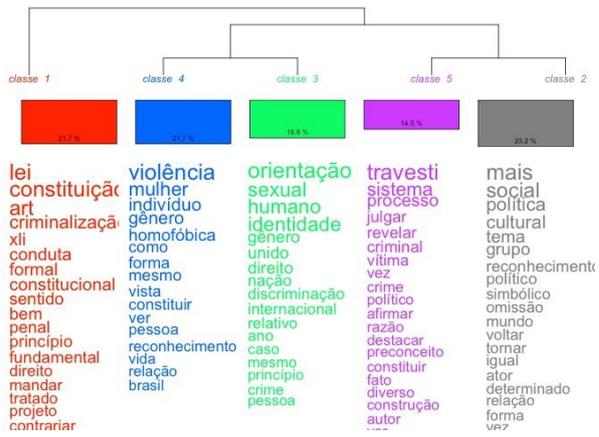


Figura 8 – Plano cartesiano AFC

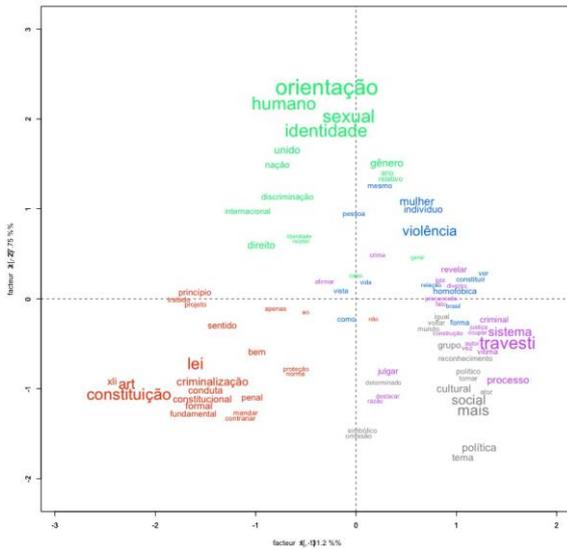
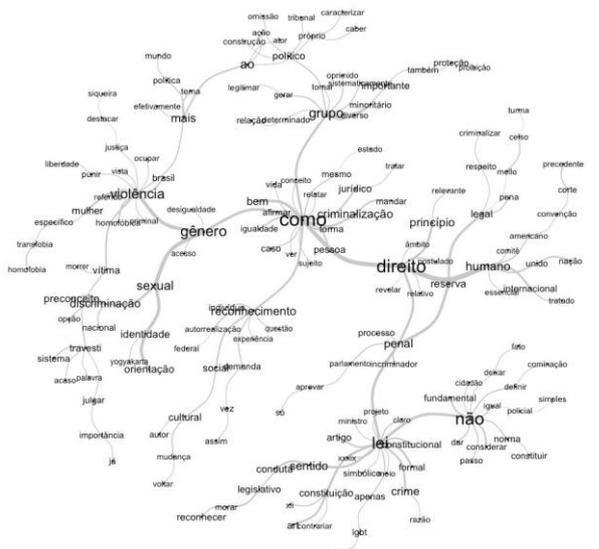


Figura 9 – ADO 26 - STF – similitude



Fonte: Elaboradas pelos autores.

Realizado o teste CHD, obtivemos cinco classes que foram divididas em ramificações do corpus em análise. Tendo em vista que, conforme dito, esse teste classifica os segmentos de texto em função dos seus respectivos vocabulários, sendo que o conjunto deles é repartido com base na frequência das formas reduzidas, conseguimos acesso a um nível razoável do conteúdo da decisão do STF, condensando o material em cinco classes e permitindo a categorização. O conteúdo analisado dessas classes fornecidas pelo Iramuteq foi categorizado por nós em classe 1 – direitos e garantias (21,7%), classe 2 – político e social (23,2%), classe 3 – gênero como direito (18,8%), classe 4 – violência de gênero (21,7%), classe 5 – justiça criminal (14,5%) (Figura 7 – Dendrograma). A classe 1 está representada em apenas uma aresta, enquanto as demais decorrem de uma segunda divisão (4 e 3, de um lado e 5 e 2, do outro).

A partir da AFC, realizamos associação entre palavras, considerando a frequência das palavras e as classes. As classes de palavras que reúnem o significado “político e social” e “justiça criminal” são encontradas mais no QID, expandindo para C e QSD. A classe que categorizamos como “violência de gênero” está no QSE, expandido para C. Já “gênero como direito” é encontrado em C expandindo para QSE. As palavras da classe que representam no plano “direitos e garantias constitucionais” estão concentradas no QIE. Em razão da lógica dos testes CHD e AFC, já dissemos que quanto mais isolada em uma aresta (CHD) e mais distantes nos quadrantes (AFC), menos as

palavras falam da mesma coisa. Ressaltamos que a distância das palavras no plano não significa oposição semântica, contudo, observamos que no corpus em análise não se apresentam em um segmento centralizado (Figura 8), pois expandem para os pontos periféricos, especialmente as palavras da classe 1.

A análise de similitude (Figura 9) com base no mesmo corpus (voto do STF) possibilitou-nos verificar as ocorrências entre as palavras e as indicações de conexidade. Observamos a centralidade da construção conceitual de “gênero”, “como” “direito”, enquanto clusters centrais e irradiadores de conexidade da maior parte das palavras. Existem clusters menores, como o que reúne vocábulos sobre “violência de gênero” à esquerda e o da caracterização da omissão legislativa mais à direita e descendo.

Do ponto de vista estritamente do conteúdo, os resultados dos testes que realizamos permitem a seguinte leitura: a maior frequência de palavras - referentes às classes 2, 3 e 4 - está ligada à conceitualização de “gênero como direito” na teoria social e política, na filosofia e em dados da experiência mencionados no corpus - que o programa localizou em outros quadrantes (C, QSE, QSD e QID). No entanto, é justamente essa conceitualização, afastada do QIE, que serve à configuração da “omissão legislativa” e que autoriza o STF (a si) a intervir na seara do Legislativo.

Ainda, conforme o STF (2019):

[...] ficou assentado que a repressão penal à prática da homofobia não alcança nem limita o exercício da liberdade religiosa, desde que tais manifestações não configurem discurso de ódio. Finalmente, ficou delineado que o conceito de racismo ultrapassa aspectos estritamente biológicos ou fenotípicos e alcança a negação da dignidade e da humanidade de grupos vulneráveis (STF, 2019).

Independentemente do acerto, o STF encampou o voto do relator, que, em termos concretos, inova na seara do Legislativo⁷ - até que este edite lei específica - mediante adequação a preceitos definidos na Lei no 7.716 de 08/01/1989 e no Código Penal. De passagem, vale dizer que o Congresso Nacional não está obrigado a legislar sobre o assunto mesmo depois dessa decisão do STF. A decisão do STF na ADO tem eficácia *erga omnis* e efeito vinculante em relação aos órgãos do Judiciário e da administração pública federal, estadual e municipal, conforme o artigo 28, parágrafo único, da Lei 9868/99 (Tavares, 2019). Todavia, não compele o Legislativo devido ao princípio da separação entre os poderes e à necessidade de abertura do princípio democrático. Na prática, a decisão do STF equivale à edição de lei nova sem prazo de vigência. Assim, vimos mais um capítulo da tensão atual entre a independência dos órgãos do sistema de justiça e accountability democrática no Brasil (em suas diferentes formas) (O'Donnell, 1998; Arantes, Loureiro, Couto e Teixeira, 2010).

3. Considerações finais

A versatilidade da AC e a adaptação a diferentes estratégias de pesquisa, qualitativas e quantitativas, podem passar uma falsa ilusão de que essa técnica independe de rigor analítico e de critérios de qualidade. Vimos que, pelo contrário, a AC compreende etapas e tipos de procedimentos que devem guardar adequação do método ao material e à questão de pesquisa. Não obstante, podem (e devem) ser assumidas estratégias de qualidade da pesquisa com a AC.

Há, ainda, diferentes processos analíticos de aplicação da AC no momento da codificação, que variam conforme a abordagem do pesquisador. Conseguimos sistematizar os tipos de procedimentos e os diferentes processos de codificação em duas tabelas, a fim de facilitar sua aplicação por pesquisadores.

O ponto central foi demonstrar o caráter sistemático da AC e a adequação para pesquisas apoiadas por computador. Trouxemos o Iramuteq como exemplo de programa de computador que auxilia a AC. Assim como a própria AC, o uso desse software está mais condicionado ao material e

⁷ A edição de lei penal foi destinada pela CF-88, exclusivamente, ao Poder Legislativo Federal (Congresso Nacional), por definição do inciso XXXIX do artigo 5º (“não há crime sem lei anterior que o defina, nem pena sem prévia *cominação legal*” - grifamos) e do inciso I do artigo 22 da CF. Vide: CF, 1988.

à questão de pesquisa do que com a viabilidade metodológica. Dentre as limitações do Iramuteq está o formato do material, que, necessariamente, precisa ser fixado no formato textual, seja um ou um conjunto de textos.

De acordo com a teoria sobre AC manejada e o emprego efetivo do programa na análise de decisões do TCU e do STF, o software tem potencial para auxiliar em diferentes tipos de procedimentos da AC. Por exemplo, diante de um volume considerável de dados, podem ser usadas ferramentas do programa para condensar o material em um texto gerenciável curto (redução), manejando as formas ativas e complementares e a sua concordância com o corpus. A partir dos dados, contribui para formação de categorias (indução), por exemplo, se valendo da CHD e da AFC. Ainda, ante eventual complexidade dos elementos textuais, pode ajudar na ampliação de informações sobre comunicadores, sujeitos, antecedentes socioculturais com triangulação de dados oriundos de documentos, entrevistas etc.

Além disso, o Iramuteq pode contribuir para a categorização antes e durante a análise dos dados. Antes, pode assumir cunho exploratório, contribuindo, por exemplo, para a formulação das perguntas de pesquisa a serem respondidas, também ajudando a selecionar a amostra a ser analisada e as categorias a serem aplicadas. Durante a análise dos dados, certamente contribuirá para delinear o processo de codificação, o treinamento dos codificadores e o processo de codificação em si. Possui a vantagem de determinar confiabilidade à pesquisa no sentido de que os testes aplicados podem ser reproduzidos por outros pesquisadores e, conforme já dito, possibilita a triangulação entre dados de diferentes corpus textuais.

Por fim, não está restrito o uso do Iramuteq à AC de um tipo específico, seja convencional, direcionada, somativa, indutiva ou dedutiva. Mas, de fato, pode apresentar mais vantagens nos tipos que partem da observação (convencional e indutiva) e das palavras-chave (somativa). Entre todos os tipos, a somativa, por utilizar um processo de mensuração e quantificação das palavras, categorias e códigos envolvendo a contagem e a comparação de palavras, normalmente palavras-chave ou conteúdo, parece a mais operacional com o programa. Há que se considerar, contudo, a criatividade do pesquisador no processo de adaptação da técnica às diferentes estratégias de pesquisa e às funcionalidades do Iramuteq, desde as mais simples, como a lexicográfica básica, até as multivariadas.

4. Referências

- ARANTES, R.B.; LOUREIRO, M.R.; COUTO, C.; TEIXEIRA, M. A. C. (2010). Controles democráticos sobre a administração pública no Brasil: Legislativo, Judiciário, Tribunais de Contas e Ministério Público. In: LOUREIRO, Maria Rita; ABRUCIO, Fernando; PACHECO, Regina Silvia (orgs.). *Burocracia e Política no Brasil. Desafios para o Estado democrático no século XXI* (pp.109-147). Rio de Janeiro: Editora FGV.
- BOWEN, G. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27-40.
- BRASIL (1988). *Constituição da República Federativa do Brasil de 1988*. Diário Oficial da União, Brasília (05 de outubro).
- BRASIL (1964). *Lei 4320/1964. Lei da contabilidade pública*. Diário Oficial da União, Brasília (17 de março).
- CAMARGO, B.V.; JUSTO, A. M. (2013) Un tutorial en portugais. 9e éd. Universidade Federal de Santa Catarina, Brasil. Disponível em: <http://www.iramuteq.org/documentation>. Acesso em: 11 ago. 2020.
- CAPES. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. Disponível em: <https://www.capes.gov.br>.
- CF. Constituição Federal Brasileira de 1988.
- COLE, F. (1988) Content analysis: process and application. *Clinical Nurse Specialist* 2(1), 53–57.
- CRESWELL, J. (2013) *Qualitative inquiry and research design: choosing among five approaches*. [b] Chapter 7: Data collection (pp. 145-178). Los Angeles: Sage.
- DENZIN, N. K. (1978). *The Research Act: a theoretical introduction to sociological methods*. 2nd ed. Chicago: Aldine. 3rd ed. Englewood Cliffs, NJ: Prentice Hall.
- DIONISIO, G. F. (2019). Análise dos elementos que compõem a cultura de inovação: um estudo teórico. Dissertação (Mestrado em Processos e Gestão de Operações) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos. doi:10.11606/D.18.2020.tde-04022020-151943. Acesso em: 12 ago. 2020.
- DOWNE-WAMBOLDT, B. (1992). Content analysis: method, applications and issues. *Health Care for Women International* 13, 313–321.
- ELO, S.& KYNGÄS, H. (2007). The qualitative content analysis process. *Journal of Advanced Nursing* 62 (1), 107–115.
- FLICK, U. (2004). Triangulation in Qualitative Research. In: FLICK, U., KARDRFF, E. & STEINKE, I (orgs.). *A companion to qualitative research* (pp. 270-275) London: Sage Publications.
- STEINKE, I. (2004). *A companion to qualitative research*. London: Sage Publications, p. 178183.
- GERBIC, Ph; STACEY, E. (2005). A purposive approach to content analysis: Designing analytical frameworks. *The Internet and Higher Education*, 8(1), 45–59.
- GIOIA, D. A. & CHITTIPEDDI, K. (1991). Sensemaking and sensegiving in strategic change initiation. *Strategic Management Journal*, 12, 433–48.
- GUBA, E. G. & LINCOLN, Y.S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105–117). Sage Publications.
- HSIEH; H.F. & SHANNON, S. (2005). Three Approaches to Qualitative Content Analysis. *Qualitative Health Research*, 15(9), 1277–1288.
- Instituto Rui Barbosa (IRB) (2010). *Normas de Auditoria Governamental (NAGS): Aplicáveis ao Controle Externo*. IRB.
- KAID, L.L. & WADSWORTH, A. J. (1989). Content analysis. In: EMMERT, P. & BARKER, L. (eds.). *Measurement of communication behavior* (pp. 197-217). New York: Longman.

- LANGLEY, A. (1999). Strategies for Theorizing from Process Data. *Academy of Management Review*, 24 (4), 691-711.
- MAYRING, P. (2004). Qualitative Content Analysis. In: FLICK, U., KARDRFF, E. & STEINKE, I.(orgs.). *A companion to qualitative research* (pp. 265-269). London: Sage Publications.
- MELO, C.; CAVALCANTE, A.K.S e FAÇANHA, K.Q. (2019). Invisibilidade do adoecimento psíquico do trabalhador: limites da integralidade na rede de atenção à saúde. *Trabalho, Educação e Saúde*, 17 (2). Disponível em: <https://www.scielo.br/j/tes/a/HPT6S5QXZrFjWRWQgRVXZPQ/?lang=pt>
- NEUENDORF, K.A. (2017). *The content analysis guidebook*. Second edition. Los Angeles: SAGE.
- O'DONNELL, G. (1998). Accountability horizontal e novas poliarquias. In: *Lua Nova* [online], 44, 27-54.
- PEDHAZUR, E. J. & SCHMELKIN, L. P. (1991). *Measurement, design and analysis: an integrated approach*. New York: LEA Publishers.
- PATTON, M. Q. (2015). *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage.
- POZZEBON, M. e PETRINI, M. (2013). Critérios para condução e avaliação de pesquisas qualitativas de natureza crítico-interpretativa. In: TAKAHASHI, A. R. W. (org.). *Pesquisa Qualitativa em Administração: fundamentos, métodos e usos no Brasil* (pp. 51-72). São Paulo: Atlas.
- PRIOR, L. (2014).Content Analysis. In: LEAVY, Patricia (ed.) *The Oxford Handbook of Qualitative Research* (pp. 359-379). Oxford: Oxford University Press.
- RATINAUD, P. & MARCHAND, P. (2012). Application de la méthode ALCESTE à de "gros" corpus et stabilité des "mondes lexicaux": analyse du "Cable-Gate" avec IramuTeQ. In: *Actes des 11eme Journées internationales d'Analyse statistique des Données Textuelles* (pp. 835-844). Liège,Belgique.
- REINERT, M. (1987). Classification descendante hiérarchique et analyse lexicale par contexte: application au corpus des poésies d'Arthur Rimbaud. *Bulletin de méthodologie sociologique*.
- SUPREMO TRIBUNAL FEDERAL (STF) (13 junio 2019). STF enquadra homofobia e transfobia como crimes de racismo ao reconhecer omissão legislativa. Imprensa STF. Disponível em: <http://portal.stf.jus.br/noticias/verNoticiaDetalhe.asp?idConteudo=414010>. Acesso em: 11 ago. 2020.
- TAVARES, A. R. (2019). *Curso de Direito Constitucional*. 17 ed. São Paulo: Saraiva.
- TCU. Tribunal de Contas da União. Portaria-TCU nº 168, de 30 de junho de 2011. Brasília Ano xlv n. 12 5/jul. 2011.

Autor y autora.

Rafael Rodrigues Viegas

Fundação Getulio Vargas (FGV), Brasil.

Doutorando em Administração Pública e Governo pela FGV e mestre em Ciência Política pela Universidade Federal do Paraná (UFPR), Brasil. Dedicar-se à pesquisa sobre: controle e burocracia; recrutamento político; elites políticas e burocráticas; democracia; política brasileira e sistema de justiça; Ministério Público.

E-mail: viegas.r.r@gmail.com

Natasha Borali

Fundação Getulio Vargas (FGV), Brasil.

Doutoranda em Administração Pública e Governo e mestre em Administração Pública e Governo pela FGV, Brasil. Dedicar-se à pesquisa sobre: controle e burocracia; accountability vertical e horizontal; política brasileira e controle interno e externo; Controladoria; Tribunal de Contas.

E-mail: boralinatasha@gmail.com

Citado.

VIEGAS, Rafael Rodrigues e BORALI, Natasha (2022). Análise de conteúdo e o uso do Iramuteq. *Revista Latinoamericana de Metodología de la Investigación Social - ReLMIS*, N°23, Año 12, pp. 21-37.

Plazos.

Recibido: 28/03/2020. Aceptado: 04/11/2020.



Procesamiento de Lenguaje Natural aplicado a las ciencias sociales. Detección de tópicos en letras de tango

Natural Language Processing applied to Social Sciences.
Topic identification in tango lyrics

Germán Rosati

Resumen

En este artículo, se presenta una aplicación de una técnica de Procesamiento de Lenguaje Natural (modelado de tópicos) sobre un *corpus* de letras de tango. Introduce un flujo de trabajo posible para el análisis textual computacional y en una técnica específica para la detección de tópicos: *Latent Dirichlet Allocation* (LDA). Se trabajará sobre un *corpus* de 5.617 letras buscando detectar de forma semiautomática sus temas. Los tópicos detectados abarcan desde imágenes de la ciudad, sobre el tango mismo, sobre emociones negativas y positivas, etc. Se analiza su evolución temporal y se muestra el cambio relativo de los tópicos en las letras de tango. También se valida el modelo analizando la composición de tópicos de algunos tangos canónicos. El trabajo busca ilustrar las potencialidades que estas técnicas tienen para el análisis de datos textuales en ciencias sociales: su escalabilidad y sus posibilidades de replicabilidad. Se marcan, finalmente, algunas limitaciones de este enfoque.

Palabras clave: procesamiento de lenguaje natural; tango; modelado de tópicos; web scraping; minería de Texto.

Abstract

This article presents the application of topic modeling, a natural language processing technique, in the context of tango lyrics analysis. It introduces a general workflow for computational text analysis and a specific topic detection technique: *Latent Dirichlet Allocation* (LDA). A semi-automated topic detection analysis of a 5,617 song *corpus* is described, finding a variety of recurring themes including urban imagery, tango itself, negative and positive emotions, etc. Further temporal analysis evidences the ebb and flow of topics in tango lyrics. The detection model is validated by examining the topic composition found in canonical song examples. This work aims to highlight the potential of these techniques for text data analysis in the social sciences, their scalability and replicability. Some limitations of this approach are also discussed.

Keywords: natural language processing; tango; topic modeling; web scraping; text mining.

1. Introducción

El trabajo empírico en las ciencias sociales se caracteriza por una gran diversidad de fuentes de información utilizadas: desde datos altamente estructurados (cuya forma más clásica son las encuestas) hasta datos de un grado menor de estructuración. En este último grupo, los datos textuales, ya se trate de documentos, noticias, entrevistas, etc.- ocupan un lugar central.

El análisis de textos en sus diversas modalidades ha constituido un ejemplo típico de trabajo sobre datos cualitativos. Los estudios cualitativos se conforman por múltiples líneas de investigación. Una de ellas se caracteriza por tres elementos (Vasilachis, 2006: 6): 1) una posición epistemológica interpretativista, interesada en la comprensión y el análisis de los modos en que el mundo es interpretado por los sujetos; 2) el uso de métodos de recolección de información flexibles, es decir poco estructurados y 3) el empleo de métodos y procesos analíticos centrados en la comprensión del detalle, la complejidad y el contexto. Otros autores (Maxwell, 2004) marcan la importancia que el significado y la interpretación tienen en la investigación cualitativa. Coexisten en esta perspectiva diferentes escuelas sumamente diversas: la etnometodología; el análisis de la conversación, del discurso y de género; el análisis narrativo; la hermenéutica objetiva; la sociología del conocimiento hermenéutica; la fenomenología; el análisis de pequeños mundos de la vida; la etnografía y los estudios culturales (Vasilachis, 2006: 24). Así, las herramientas y técnicas metodológicas de las que se suele hacer uso en este tipo de enfoques se encuentran más vinculadas al análisis literario o del discurso y a métodos que buscan la comprensión profunda de los *corpus*.¹ Dichos enfoques han sido objeto de diversas críticas (Reynoso, 2007). No es objeto del presente trabajo evaluar dichas críticas y solo se marcará un aspecto de carácter metodológico: las posturas interpretativistas tienden a poner el eje en la capacidad subjetiva del investigador para realizar la interpretación o el análisis. Esto hace que parte de las investigaciones puedan presentar una relativa falta de sistematicidad metodológica. Al mismo tiempo, el peligro de la imposibilidad de replicación de sus resultados suele estar latente.

Tales problemas han sido abordados por otros autores y perspectivas que han buscado avanzar hacia un incremento de la replicabilidad de los estudios basados en datos cualitativos. Así, el llamado *Qualitative Comparative Analysis* elaborado en la década del '80 por Charles Ragin (1987) intentó estandarizar aspectos del análisis comparativo en las ciencias sociales. Lo hizo a partir de la sistematización y el ordenamiento de las unidades o casos. El objetivo era identificar diferencias y semejanzas entre las mismas a través del uso de diversos procedimientos inferenciales y comparativos basados en el álgebra booleana.²

El análisis de contenido abarca diferentes enfoques que van desde análisis impresionistas, interpretativos y escasamente estructurados hasta análisis estrictamente textuales (Hsieh & Shannon, 2005). La llamada *Grounded Theory* (Tie, Birks y Francis, 2019) busca la construcción de teoría social mediante procedimientos inductivos generados a partir de los datos producidos en la investigación. Una vez que tales datos han sido recolectados se siguen algunas etapas analíticas básicas: la codificación (que busca identificar los puntos de anclaje conceptuales en las fuentes); la conceptualización (los códigos con contenidos similares son agregados); la categorización (los conceptos similares son agrupados en categorías teóricas relevantes) y la teorización (una teoría surge del agrupamiento de categorías teóricas).

De forma similar, el *Qualitative Content Analysis* (Mayring, 2000) busca sistematizar las etapas y decisiones metodológicas en el proceso de análisis de textos. Procura sistematizar el proceso de investigación (preprocesamiento, análisis e interpretación) y desarrollar procedimientos de inferencia metodológicamente controlados (desarrollo inductivo de categorías, resúmenes, análisis de contexto, aplicación deductiva de categorías).

¹ Uno de los referentes más relevantes de la antropología interpretativa escribe en uno de sus textos más famosos: "*the culture of a people is an ensemble of texts, themselves ensembles, which the anthropologist strains to read over the shoulders of those to whom they properly belong.*" (Geertz, 1974: 452).

² Se ha utilizado como método de análisis de entrevistas en profundidad para diversos propósitos analíticos (Rosati y Chazarreta, 2017).

Este tipo de enfoques suele utilizar softwares como herramientas auxiliares del proceso. Paquetes como *Atlas-Ti* (Murh, 1997) o *Nvivo* (Richards, 1999) permiten estandarizar parte del procesamiento de la información mediante búsquedas en base a diferentes métodos textuales o basadas en métricas de similitud, codificación/anotación del *corpus*, generación de mapas o redes conceptuales y otros procedimientos similares (Casanova Correa y Pavón, 2002).

El uso de procedimientos de control metodológico, ya sean analógicos o mediatizados por algún software, permite avanzar en la estandarización de muchas de las operaciones analíticas mencionadas más arriba. No obstante, el proceso sigue teniendo un carácter fuertemente manual y continúa residiendo en el investigador. Dicha base manual, el uso de interfaces gráficas y la consecuente inexistencia de scripts o sintaxis que codifiquen dichas operaciones muestran que los procesos tienen un grado medio o bajo de estandarización. Las posibilidades de replicar tanto el procesamiento de los datos como los análisis se encuentran atadas a la decisión del investigador de documentar de forma minuciosa cada una de las decisiones tomadas.

A su vez, aparece un problema vinculado a la escala: la transcripción y codificación manual de *corpus* textuales (incluso contando con la mediación de algún software que asista dicha operación) limita fuertemente el tamaño de los *corpus* a analizar. Salvo que se cuente con recursos suficientes, con un equipo altamente capacitado y con disponibilidad de tiempo, el carácter manual de estos procesos puede llegar a limitar la escala de los *corpus* construidos y analizados.

La combinación de técnicas de *web scraping*, *Text Mining* y *Natural Language Processing* (NLP) puede ser de utilidad a las ciencias sociales atendiendo a estos problemas. Por un lado, permiten realizar una sistematización (y, eventualmente, lograr un cierto grado de automatización) de las diversas etapas del proceso de investigación, desde la recolección de datos y construcción del *corpus* hasta el preprocesamiento de un texto y su análisis. Específicamente, las técnicas de NLP habilitan la aplicación de métodos cuantitativos de análisis para una amplia diversidad de tareas (clasificación de textos, detección de temas y tópicos, detección de estructuras semánticas, etc.). También abren la posibilidad de escalar el trabajo de forma eficiente. En lugar de leer cada uno de los textos de un *corpus*, tarea que rápidamente se vuelve imposible, las técnicas de minería de texto permiten analizar de forma automática *corpus* de gran escala.

El presente trabajo tiene como objetivo discutir algunas aproximaciones metodológicas al análisis computacional de textos y presentar algunas de las potencialidades que tienen para las ciencias sociales. Esto se realizará a partir de su aplicación a un caso de estudio concreto: el análisis de los temas en un *corpus* de 5.617 letras de tango. Se concentrará en la discusión del flujo de trabajo y en la aplicación de una técnica específica para la detección de tópicos: el modelo *Latent Dirichlet Allocation* (LDA). Estos aspectos permitirán abordar los dos problemas mencionados: la escalabilidad y replicabilidad. A su vez, se propone realizar una ilustración del tipo de análisis, problemas y preguntas de investigación que tales técnicas habilitan.

2. Antecedentes metodológicos en el análisis de letras de tango

Existe una profusa literatura científica vinculada al análisis discursivo y de contenido de la poética del tango. El presente artículo no pretende abordar dicho campo problemático en la totalidad de sus determinaciones. Es por ello que los documentos reseñados en esta sección son ilustrativos de los aspectos metodológicos mencionados previamente.

Un abordaje común es el rastreo de un tópico o problema particular a lo largo de un *corpus* textual. En el análisis de letras de tango un tópico habitual es la problemática de género: representaciones de la mujer (López, 2010), masculinidades -por ejemplo, la figura del guapo- (Gasparri, 2011) o el sesgo sentimental de las letras (Marchese, 2007). Otros textos tienen un enfoque más amplio e identifican una mayor cantidad de temas. Así, Lucía Willenpart (2011) trabaja algunos temas comunes: el amor, el duelo amoroso, la mujer, la madre, el tango mismo,

etc. Un rasgo común en todos estos textos es el tamaño pequeño de los *corpus*:³ el más grande cuenta con 30 letras (Marchese, 2007).

Por otro lado, Cantón (1972) se pregunta por los objetos y sujetos de los tangos cantados por Carlos Gardel. Este estudio tiene un carácter cuantitativo, por lo cual analiza un corpus significativamente más grande que los anteriores: alrededor de 100 tangos.

De esta forma, en términos metodológicos es posible identificar tres rasgos de los textos reseñados:

1. los criterios para la confección del corpus no aparecen explicitados
2. el tamaño de los corpus es entre pequeño (10 tangos) y mediano (100 tangos)
3. con la excepción del texto de Cantón (1972), los corpus son abordados a partir del rastreo de un conjunto de tópicos o preguntas particulares, privilegiando un análisis interpretativo del contenido.

3. Construcción del corpus

Se presentará un flujo de trabajo clásico aplicado a la detección de tópicos en un corpus de letras de tango. El mismo fue construido a partir de un proceso de *web scraping*. Se descargaron todas las letras de tango, milongas, valeses, etc., disponibles en el sitio Todo Tango.⁴ Además de las letras se descargó información accesorio sobre cada letra particular. Para ello, se confeccionaron dos *web crawlers* (es decir, programas que recorren automáticamente la estructura de un sitio web).⁵

El *scraping* -literalmente, “raspado” o “rascado”- consiste en la descarga y formateo de la información disponible en sitios web, información que generalmente no se encuentra en condiciones de ser trabajada de forma cuantitativa (Mitchell, 2015). El uso de este tipo de técnicas hace posible abordar los dos problemas mencionados más arriba en la etapa de construcción de la información. Por un lado, permite sistematizar y hacer replicable la construcción del corpus. Esto se logra a partir de la ejecución de una rutina (con su correspondiente código) mediante la cual cualquier investigador puede reproducir los procedimientos de descarga de datos y construcción del corpus. Por otro lado, permite incrementar la capacidad de captura de información: el tamaño total del corpus construido en este trabajo es de una escala sensiblemente mayor que los utilizados en los estudios reseñados en la sección anterior.

Ahora bien, además de estas ventajas deben tenerse en cuenta algunas limitaciones. En primer lugar, existe una primera fuente de sesgo: dado que sería notablemente complejo construir un relevamiento completo de letras de tango, el conjunto que conforma el corpus final -pese a su amplitud- ha sido curado y seleccionado por los administradores del sitio. En este sentido, los problemas son similares a los que podrían producirse al utilizar compilaciones realizadas por diversos estudiosos.⁶

A su vez, los metadatos recabados (autor, fecha de composición, etc.) presentan diversos grados de calidad, cobertura e integridad. El atributo correspondiente a la fecha es uno de los que mayores cantidades de no respuesta presentan (alrededor de un 64%). Una vez más, en este caso, prevalece el sesgo presente en la fuente.

³ El tamaño de los *corpus* está referido a la cantidad de textos que citan o mencionan los diferentes autores en cada uno de los trabajos citados. Lógicamente, es razonable asumir que solamente se cita una fracción -desconocida- de los textos analizados.

⁴ <https://www.todotango.com/>

⁵ El código puede ser consultado en uno de los repositorios del proyecto https://github.com/gefero/tango_scrap.

⁶ Véase por ejemplo Russo y Marpegán (2000).

En ambos casos, estos sesgos podrían mitigarse a partir del scrapeo de otras fuentes o sitios.⁷ Dado que el objetivo de este trabajo es mostrar la aplicación de ciertas técnicas analíticas y no la extracción de conclusiones definitivas respecto a la poética del tango se utilizaron estas variables a título ilustrativo.

El *corpus* final consiste en de 5.617 letras de tango, agrupadas en un dataset con la siguiente estructura:

Tabla 1. Ejemplo de base de datos utilizada

Título	Ritmo	Año	Compositor	Autor	Letra
A bailar	Tango	1943	D. Federico	H. Expósito	a bailar a bailar que la orquesta se va...
...
Malena	Tango	1941	L. Demare	H. Manzi	malena canta el tango como ninguna...
Zurdo	Tango	S/D	A. Pontier	F. Silva	era del tiempo lindo que siempre es antes...

Fuente: Elaboración propia.

3.1. Análisis descriptivo de los metadatos

En este apartado se presenta un análisis sintético de los metadatos de las letras analizadas.

Tabla 2. Distribución de las letras de tango por ritmo

Ritmo	f	%
Tango	3964	70.6%
Milonga	477	8.5%
Vals	473	8.4%
Poema lunfardo	273	4.9%
Canción	156	2.8%
Candombe	35	0.6%
Otros	239	4.2%
Total	5617	100%

Fuente: Elaboración propia en base a datos recolectados de todotango.com

El 87.5% se concentran dentro de ritmos del tipo tangos, milongas y vals.

⁷ Se intentó complementar la información de fecha a partir del scrapeo de otro sitio (www.el-recodo.com). Sin embargo, no se lograron mejoras significativas en la no respuesta a la variable fecha. Queda pendiente para sucesivas aproximaciones al problema evaluar otros sitios o eventualmente completar dicho campo mediante la automatización de búsquedas en internet de las fechas faltantes. Todo el código que efectúa el scraping de ambos sitios y los *datasets* en cuestión pueden ser encontrados en el repositorio mencionado en la nota 4.

Tabla 3. Distribución de las letras de tango por autor

Autor	f	%
Enrique Cadícamo	168	3.0%
Homero Manzi	105	1.9%
Celedonio Flores	102	1.8%
Francisco G. Jiménez	80	1.4%
Marta Pizzo	80	1.4%
Horacio Ferrer	79	1.4%
Héctor Negro	77	1.4%
Carlos Bahr	74	1.3%
Cátulo Castillo	73	1.3%
Otros	4779	85.1%
Total	5617	100.0%

Fuente: Elaboración propia en base a datos recolectados de todotango.com

A su vez, se observa que los primeros 9 autores acumulan el 15% de las letras del conjunto de datos. Los principales autores son Cadícamo, Manzi y Flores.

Tabla 4. Distribución de las letras de tango por década

Década	f	%
1900	9	0.1%
1910	24	0.4%
1920	365	6.4%
1930	375	6.7%
1940	398	7.1%
1950	140	2.5%
1960	67	1.2%
1970	65	1.1%
1980	120	2.1%
1990	79	1.4%
2000	179	3.2%
2010	153	2.7%
Sin dato	3643	64.8%
Total	5617	100.0%

Fuente: Elaboración propia en base a datos recolectados de todotango.com

Finalmente, al analizar la distribución de letras por década, se observa la gran cantidad de datos faltantes (casi un 65%), lo cual obligará a tomar con precaución, y a título ilustrativo, análisis posteriores. A su vez, (y como era de esperarse) es el período que va de los años '20 a los '40 el que mayor cantidad de letras encuentra.

4. Preprocesamiento del texto

Al observar en la tabla 1 el campo *Letra* (el núcleo del análisis), puede advertirse que se trata de un caso típico de datos no estructurados: las letras constituyen texto libre y no parece respetarse la estructura tripartita del dato. Las filas sí representan una unidad (los tangos) pero no aparecen atributos, o en todo caso, solamente es visible un atributo, la letra. El primer paso, entonces, es transformar esta representación no estructurada en una estructurada. La misma tendrá como objeto reducir la complejidad del texto, dado que “el lenguaje es complejo. Pero no toda su complejidad es necesaria para analizar un texto de forma efectiva” (Grimmer & Stewart, 2013: 72).

4.1. Modelo BoW (Bag of words)

Para llegar a esa representación estructurada,⁸ será necesario pensar en un formato de datos acorde a las necesidades del análisis. La unidad de análisis serán los tangos individuales, por lo cual, cada fila en la matriz final será un tango. A su vez, cada columna consistirá en un término t del vocabulario general V del corpus C . Cada celda estará constituida por el conteo crudo de ocurrencias de cada palabra (columna) en cada documento (fila). Esta representación es la que se denomina *Bag of Words* o “bolsa de palabras” y se dispone en una Matriz de Frecuencias de Términos (TFM, por sus siglas en inglés).

Por ejemplo, para dos de los tangos analizados esta matriz tomaría la siguiente forma:

Tabla 5. Ejemplo de matriz de frecuencia de términos cruda

Letra	agua	blan da	car tel	cr uel	el	en	era	la	man da	mas	propa ganda	que
Cruel en el cartel la propaganda manda cruel en el cartel	0	0	2	2	2	2	0	1	1	0	1	0
Era más blanda que el agua que el agua blanda	2	2	0	0	2	0	1	0	0	1	0	2

Fuente: Elaboración propia en base a datos recolectados de todotango.com

Puede verse que al construir esta matriz la información sobre el orden de las palabras se ha perdido. El orden de las columnas es ahora arbitrario (en este caso, lexicográfico) y no respeta la estructura secuencial de las palabras en un texto. Esta es una simplificación importante del modelo *BoW*. Esta limitación puede subsanarse parcialmente generando una TFM de bi-gramas (pares de palabras), tri-gramas (tripletras de palabras) o n-gramas. El costo es un crecimiento exponencial en la dimensionalidad de la TFM.

Debe tenerse en cuenta que esta matriz se construye sobre el vocabulario V , o sea el total de términos únicos del corpus C . Así, V incluiría a priori pronombres, preposiciones, diferentes conjugaciones de verbos, sustantivos en singular o plural, etc. Esto hace que el vocabulario “crudo”

⁸ Vale destacar que el flujo de trabajo descrito a continuación es uno de los más comunes pero de ninguna manera el único, ni necesariamente el mejor en términos absolutos. El flujo y las operaciones contenidas en el mismo deberán ser revisadas para cada problema particular (Grimmer & Stewart, 2013).

de C tienda a ser demasiado grande. Es por ello que en la etapa de preprocesamiento del texto se utilizan algunas técnicas para reducir la complejidad y la extensión de V .

Un paso simple es la eliminación de lo que suelen denominarse *stopwords*, básicamente artículos, preposiciones, conectores, etc. La lógica detrás de esta eliminación es que estas palabras aportan poca o nula información acerca del contenido de cada documento.

Tabla 6. Ejemplo de matriz de frecuencia de términos sin stopwords

Letra	agua	blanda	cartel	cruel	era	manda	propaganda
Cruel en el cartel la propaganda manda cruel en el cartel	0	0	2	2	0	1	1
Era más blanda que el agua que el agua blanda	2	2	0	0	1	0	0

Fuente: Elaboración propia en base a datos recolectados de todotango.com

A partir de la eliminación de las *stopwords*⁹ y de los signos de puntuación se obtiene en la tabla 6 una representación más resumida de la información contenida en C . No obstante, ésta no es la única operación disponible para reducir la complejidad de C .

4.2. Normalización de conteos

El último paso¹⁰ supone normalizar los valores de las celdas de la TFM . Al momento de filtrar los *stopwords* se buscaba poder eliminar aquellas palabras con poca información acerca del

⁹ En el caso de los pronombres personales, los mismos fueron eliminados en base a una lista. Sin embargo, sobrevivió a esta eliminación el pronombre “vos”, dado que no se encontraba en la lista original. No obstante, resulta interesante ver cómo en los resultados del ejercicio de detección de tópicos este pronombre aparece en un tópico propio junto con verbos conjugados en segunda persona. En aras de la facilidad de lectura no se explicita de forma detallada el tratamiento realizado en cada etapa del preprocesamiento y, particularmente, el realizado para las diferentes *stopwords*, signos de puntuación, etc. Puede consultarse todo el material de replicación para el análisis de tópicos en el siguiente repositorio https://github.com/gefero/topic_modeling_tango.

¹⁰ Suele ser habitual, a los efectos de reducir la diversidad de términos, realizar algún proceso de normalización de los mismos. Se busca llevar las palabras flexionadas a una forma normal que represente a toda una clase de palabras. Esta forma normal, llamada lema, puede ser pensada típicamente como la palabra utilizada como entrada en los diccionarios de lengua: el infinitivo para las conjugaciones verbales, el masculino singular para adjetivos, etc. Esto suele requerir la detección de la función sintáctica de la palabra, su contexto y la construcción de diccionarios de palabras que mapean formas flexionadas a lemas. También pueden estimarse los lemas a partir de algoritmos y modelos probabilísticos (grafos de asociación, *clustering*, redes neuronales, etc.). A esta operación de reducción se la llama lematización. Existe, además, una segunda forma de reducción *-stemming-* más simple (si bien puede considerarse una variante de la lematización). La misma opera sobre las declinaciones de las palabras. Aquellas palabras que remiten a un mismo concepto básico son reducidas a la misma raíz. Por ejemplo, “familias”, “familia” y “familiar” son reducidas a “familia”. No se entra en mayores especificidades porque dichos procedimientos no fueron utilizados para el presente trabajo. Para mayor detalle puede consultarse (Jurafsky y Martin, 2006).

contenido en todos los textos. Es posible extender este razonamiento para el resto de los términos de V . Así, pueden identificarse dos dimensiones de la frecuencia de dichos términos:

1. un término t es importante si es muy frecuente en un documento d de C
2. a su vez, t es más informativo del contenido de un documento d si está presente en pocos y no en todos los documentos de C .

Es decir, resulta importante analizar la frecuencia de t tanto en el documento d como en el corpus total C . Existen dos métricas para lograr este objetivo. Para la primera dimensión se parte del conteo crudo de t en d : $c(t, d)$, es decir, cada celda de la *TFM* "cruda". Se define una métrica llamada *Term Frequency (TF)* es decir, el conteo crudo normalizado por la extensión del documento (el total de términos en el documento):

$$TF(t, d) = \frac{c(t, d)}{\sum_{t \in d} c(t, d)}$$

En relación a la informatividad de un término a lo largo de C , puede definirse la siguiente métrica, llamada *Document Frequency (DF)*:

$$DF(t) = \log \frac{df(t)}{|C|}$$

donde $df(t)$ es la cantidad de documentos en C que contienen a t ; $|C|$ es el tamaño del corpus.

De esta forma, DF informa acerca de la proporción de documentos que contienen a t . Cuanto mayor es $DF(t)$ menos informativo es t . Es por ello que se usa la inversa de esta métrica:

$$IDF(t) = \log \frac{|C|}{df(t)}$$

$IDF(t)$ entonces, es mayor, cuanto menor es la frecuencia de t en C , es decir, cuanto más informativo es t .

Podemos combinar ambas dimensiones en una métrica resumen, llamada *TF - IDF*, *Term Frequency-Inverse Document Frequency*:

$$TF - IDF(t) = \frac{c(t, d)}{\sum_{t \in d} c(t, d)} \times \log \frac{|C|}{df(t)} = TF(t, d) \times IDF(t)$$

Así, valores altos de $TF(t, d)$ y valores altos de $IDF(t)$ -o sea, valores bajos de $DF(t)$ - arrojan valores altos de $TF - IDF(t)$. O sea, términos t frecuentes en d y poco frecuentes en C .

No obstante, en ciertos casos (como el del corpus que se analiza en este trabajo) suele suceder que el conteo crudo $c(t, c)$ de términos funciona de forma aceptable. En el caso específico

de este trabajo, luego de evaluar ambas alternativas se optó por utilizar $c(t, d)$ como métrica en la *TFM*.¹¹

Para resumir, entonces, el preprocesamiento realizado para este *corpus*:

1. normalización a minúsculas
2. eliminación de *stopwords*
3. eliminación de puntuación y
4. eliminación de caracteres extraños y dígitos
5. construcción de la *TFM* basada en $c(t, c)$

5. Detección de tópicos

Existen muchas técnicas para la detección automática de tópicos en *corpus* textuales. Algunas de ellas están basadas en ciertas formas de descomposición de la *TFM*.¹² Para el presente trabajo se utilizará uno de los métodos más conocidos: *Latent Dirichlet Allocation* o LDA.

La intuición detrás de LDA (Blei, 2012: 79) es que cada d de C puede exhibir varios tópicos, es decir, puede hablar de varios temas simultáneamente. Por ejemplo, al analizar un tango como “Malena” de Homero Manzi, se observa que habla de diferentes temas: del amor, del tango, del barrio, etc. La idea detrás de LDA es poder operacionalizar esta intuición a través de un modelo generativo, es decir, asumiendo la existencia de un “proceso generador de textos”: un proceso aleatorio imaginario por el cual un documento es producido.

Más formalmente,¹³ un tópico se define como una distribución de probabilidad a lo largo de un vocabulario V fijo. Por ejemplo, si existiera un tópico como *sentimientos o emociones* sería esperable que palabras como “amor”, “pena”, “sufrimiento”, tuvieran altas probabilidades de inclusión en el mismo. En cambio, palabras como “ciudad”, “barrio” estarían más asociadas a un tópico que hable acerca de la *ciudad*.

El proceso que genera cada texto del *corpus* funciona bajo los siguientes supuestos. Para cada documento d en el *corpus* C se generan las palabras w que lo componen en un proceso de dos etapas:

1. Se selecciona de forma aleatoria una distribución de tópicos para d
2. Para cada palabra (w) en d
 - i) se selecciona aleatoriamente un tópico de la distribución general de tópicos.
 - ii) se selecciona aleatoriamente una palabra correspondiente a la distribución de todo el vocabulario V .

De esta forma, cada documento d exhibe ciertos tópicos t en diferente proporción (paso 1.), cada palabra w es extraída de uno de los tópicos (paso 2.ii), donde el tópico seleccionado es elegido de la distribución de tópicos de ese documento d particular (paso 2.i).

¹¹ Para una discusión al respecto de estas métricas, puede verse Wiedemann (2016).

¹² Existen métodos que dividen en dos componentes en la *TFM*, una matriz de *documentos x tópicos* y otra de *términos x tópicos*, tales como *Non Negative Matrix Factorization* y *Latent Semantic Analysis* (Hassani, Iranmanesh, & Mansouri, 2019).

¹³ Esta sección se basa en Blei (2012).

Así, el objetivo del modelado de tópicos es descubrir los temas a los que alude un determinado conjunto de documentos. Esa estructura de tópicos puede ser pensada como un set de variables latentes a la *TFM*. Lo único observado es el conjunto de documentos (preprocesado como una *TFM*). La estructura de tópicos (es decir, la composición de tópicos por documento y la asignación de palabras a un documento) puede ser considerada como un conjunto de variables no observadas (justamente lo que se trata de estimar).

Formalizando el razonamiento anterior, es posible ver que se trata de una probabilidad conjunta:

$$p(\beta_{1:K}, \theta_{1:D}, z_{1:D}, w_{1:D}) = \prod_{i=1}^K p(\beta_i) \prod_{d=1}^D p(\theta_d) \left(\prod_{n=1}^N p(z_{d,n} | \theta_d) p(w_{d,n} | \beta_{1:K}, z_{d,n}) \right)$$

Así, los tópicos están indexados en $\beta_{1:K}$, donde cada β_k es una distribución de probabilidad sobre el vocabulario. La proporción de tópicos para el d -ésimo documento están indexadas por θ_d , donde $\theta_{d,K}$ es la proporción del tópico d en el documento k . La asignación de tópicos para el documento d , está dada por z_d , donde $z_{d,n}$ es la asignación de tópico para la n -ésima palabra en el documento d . Las palabras observadas en el documento d , son w_d . Aquí, $w_{d,n}$ es la n -ésima palabra en el documento d , que es un elemento del vocabulario.

Puede verse que existen ciertas dependencias en el modelo: por ejemplo, la asignación de tópicos en un documento ($z_{d,n}$) depende de la proporción de tópicos por documento θ_D . A su vez, la palabra observada $w_{d,n}$ depende tanto de la asignación de tópicos $z_{d,n}$ como de todos los tópicos β_k .

Entonces, el problema es poder estimar la estructura de tópicos a partir de los documentos observados. De esta forma, es posible formular el problema a partir del llamado "posterior":

$$p(\beta_{1:K}, \theta_{1:D}, z_{1:D} | w_{1:D}) = \frac{p(\beta_{1:K}, \theta_{1:D}, z_{1:D}, w_{1:D})}{p(w_{1:D})}$$

El numerador es la distribución conjunta de todas las variables aleatorias del modelo y puede ser estimado de forma relativamente simple. El problema reside en el denominador: la probabilidad marginal de las observaciones, es decir, la probabilidad de observar el *corpus* dado bajo cualquier modelo de tópicos. Si bien, debería (en teoría) poder calcularse la agregación de todas las distribuciones de tópicos para cada una de las posibles estructuras de tópicos, lo cierto es que se trata de un problema computacionalmente intratable. Por ello, al igual que en muchos problemas dentro del marco bayesiano, es necesario recurrir a aproximaciones numéricas.¹⁴

Ahora bien, el método utilizado tiene algunos supuestos que, si bien pueden deducirse de lo expuesto más arriba, es útil hacerlos explícitos:

1. cada documento d se compone de varios tópicos.
2. un tópico, a su vez, se compone de palabras; más precisamente, un tópico es una distribución de probabilidad sobre la totalidad de palabras del vocabulario V .

¹⁴ No es el objetivo de este trabajo desarrollar los métodos de inferencia y aproximación. En general, se basan en métodos de inferencia variacional o en métodos basados en *Markov Chain Monte Carlo*. Para un mayor desarrollo puede verse Asunción, Welling, Smyth, y Teh (2009).

3. los tópicos preexisten a los documentos y la distribución de probabilidad sobre V es constante.
4. dado que se basa en el modelo *BoW*, para la construcción de tópicos se asume que las palabras no tienen orden.
5. el orden de los documentos no es relevante.
6. se asume que existe una cantidad fija de tópicos (y que es un hiperparámetro del modelo). Esto puede ser un problema al analizar *corpus* con documentos de épocas muy diferentes.¹⁵

Utilizando LDA se buscó detectar los tópicos más relevantes en el *corpus* de letras de tango. Ahora bien, como se desprende del apartado anterior, uno de los problemas principales es determinar la cantidad de tópicos a estimar. Este problema es análogo al problema de determinación de la cantidad de clusters al aplicar algoritmos de *clustering* tales como *K-means*. Es más, en la etapa de preprocesamiento, también se tomaron una serie de decisiones: el tipo de *TFM* a utilizar, por ejemplo. Sería posible considerar cada una de estas decisiones como un hiperparámetro a evaluar y testear todas las combinaciones posibles de estos hiperparámetros, junto con el parámetro obligatorio referido a la cantidad de tópicos a detectar en el *corpus*.

Existen diversas métricas que permiten cuantificar el ajuste (es decir, qué tan “bueno” es) del número de tópicos definido en términos cuantitativos (*log-likelihood*, *perplexity*, etc.). En general el uso de estas métricas conduce a modelos que logran buena performance estadística pero que no necesariamente generan tópicos interpretables.

En términos generales, un número de tópicos más grande tiende a arrojar mejores métricas y a permitir una alta resolución de la estructura latente del *corpus*. No obstante, se ha observado que al aumentar el número de tópicos la calidad de los tópicos (en términos de interpretabilidad) tiende a decrecer (Mimno et al., 2011; Chang et al., 2009). De esta forma, al igual que en muchos otros problemas, complejidad del modelo e interpretabilidad tienden a ir en direcciones contrarias.¹⁶

Párrafo aparte merece la problemática de ciertos procedimientos retóricos y estilísticos en textos literarios. Dado que el modelado de tópicos tiende a operar sobre la co-ocurrencia de palabras para la definición de tales tópicos, en algunos casos, recursos como las metáforas pueden generar problemas de identificación. En este sentido, las metáforas no son tematizadas en este trabajo. No obstante, algunos estudios han intentado utilizar el modelado de tópicos (y, particularmente, LDA) como detectores de metáforas.¹⁷

6. Resultados

En el presente ejercicio se intentó buscar un k (número de tópicos) que permitiera identificar temas interpretables. En el anexo se presentan algunos de los principales términos para diferentes k . Un primer rasgo que puede observarse es que, independientemente del número de tópicos que se elija, existen algunos temas que se mantienen:

- Sentimientos y emociones con carácter positivo o negativo
- Imágenes de la noche, oscuridad y sombras asociadas a despedidas
- Imágenes que vinculan al tango y al barrio o al arrabal

¹⁵ Blei (2012) expone varios métodos para flexibilizar este supuesto. Particularmente, los llamados *dynamic topic modelling* son una técnica posible.

¹⁶ Existe otro conjunto de métricas -*coherence*, por ejemplo.- que se centran en la interpretabilidad (Mimno et al., 2011).

¹⁷ Ver, por ejemplo, Navarro-Colorado et al (2011) y Heintz et al (2013).

- Tópico sobre el tango, específicamente

Se han llamado “misceláneos” a aquellos tópicos que presentan una distribución de palabras que no resulta interpretable. Constituyen en este sentido tópicos que no parecen tener un significado atribuible y resultan de carácter residual. Por este motivo no serán objeto de análisis.

6.1. Identificación de tópicos

De esta forma, el modelo seleccionado es el que muestra un $k=12$ tópicos. Es importante recordar que, dado que los tópicos se definen como una distribución de probabilidad sobre las palabras del vocabulario, una misma palabra tiene una cierta probabilidad de pertenencia a todos los tópicos. Las diferencias, entonces, son de carácter relativo. Ciertas palabras pertenecen a ciertos tópicos *con mayor probabilidad que otras*. Esto hace que pueda existir una cierta superposición (*overlapping*) entre términos a lo largo de los tópicos.

En el gráfico siguiente se exponen las primeras 30 palabras de cada uno de los tópicos, es decir, las 30 palabras con mayor probabilidad de pertenencia. A su vez, el tamaño en la nube de cada término es proporcional a dicha probabilidad. Puede verse, entonces, que salvo algunas excepciones como “amor”, las principales palabras tienden a no superponerse.

Gráfico 1. Composición de las 30 palabras de cada tópico ($k=12$)



Fuente: Elaboración propia en base a datos recolectados de todotango.com

Es también importante remarcar que, al igual que en otros métodos estadísticos,¹⁸ la interpretación de los tópicos (es decir, su etiquetado) recae sobre el equipo de investigación. En ese sentido, la identificación y etiquetado se basa en el análisis de sentido que se presentan en las palabras con mayor probabilidad de pertenencia a cada tópico.

Puede verse que el primer tópico detectado tiene palabras como “noche”, “luna”, “cielo”, “sombas”, “viento”. Es decir, nos habla de *imágenes naturales o climáticas*.

A su vez, el segundo tópico capta el tema de la *ciudad y de las imágenes urbanas*: menciona términos como “buenos”, “aires”, (los cuales remiten, claramente a “Buenos Aires”), “ciudad”, “calles”. El tópico 6 habla del *arrabal, pero sobre todo del tango mismo* (“tango”, “barrio”, “arrabal”, “canción”, “milonga”, “bandoneón”). El tópico 7 (“pasado”, “recuerdo”, “tiempo”) menciona palabras vinculadas al paso del *tiempo y a la memoria*.

Los tópicos 4 y 9 contienen palabras vinculadas a las *emociones*. El 4 (“ilusión”, “pasión”, “corazón”, “amor”) con una connotación más positiva y el 9 (“amor”, “dolor”, “pena”, “triste”), aparentemente negativa. Resulta interesante notar cómo un mismo término (“amor”) aparece vinculado tanto a un tópico con signo negativo como a otro con connotación positiva. Es decir, que el modelo parece haber sido capaz de detectar ciertos matices y usos diferentes asociados al término “amor”.

El tópico 5 y el 10 logran evidenciar temas de carácter “étnico”, por decirlo de alguna manera: el 5 con palabras como “china”, “gaucho”, “tierra”, “sangre” capta la cuestión de la *gauchesca y el campo*. El tópico 10, en cambio, (“carnaval”, “negro”, “morena”, “candombe”) habla sobre el *candombe, la cuestión étnica y la negritud*.

Por último, restan cuatro tópicos. Dos de ellos (3 y 8) tienen un carácter residual. Resultan difíciles de interpretar, por ello fueron etiquetados como misceláneos. No obstante, el 11 y el 12, si bien contienen muchos términos que son poco interpretables, puede verse que el 11 (“vieja”, “domingo”, “niños”) contiene algunas palabras vinculadas a la *vida familiar* y el 12, términos en *lunfardo* (“bulín”, “pinta”, “che”, “pibe”). El tópico 12, también parece tener como sus dos palabras más importantes “vos” y “sos”, junto con verbos conjugados en la segunda persona del singular, con lo cual, parece estar captando el hecho de que se habla de *forma directa a un interlocutor*.

Tabla 7. Identificación de los tópicos hallados

Tópico	Etiqueta
01	Imágenes climáticas
02	Ciudad, imágenes urbanas
04	Emociones positivas
05	Campo y gauchesca
06	Tango y arrabal
07	Tiempo, recuerdos
09	Emociones negativas
10	Candombe
11	Misc y familia
12	Misc y lunfardo

Fuente: Elaboración propia en base a datos recolectados de todotango.com

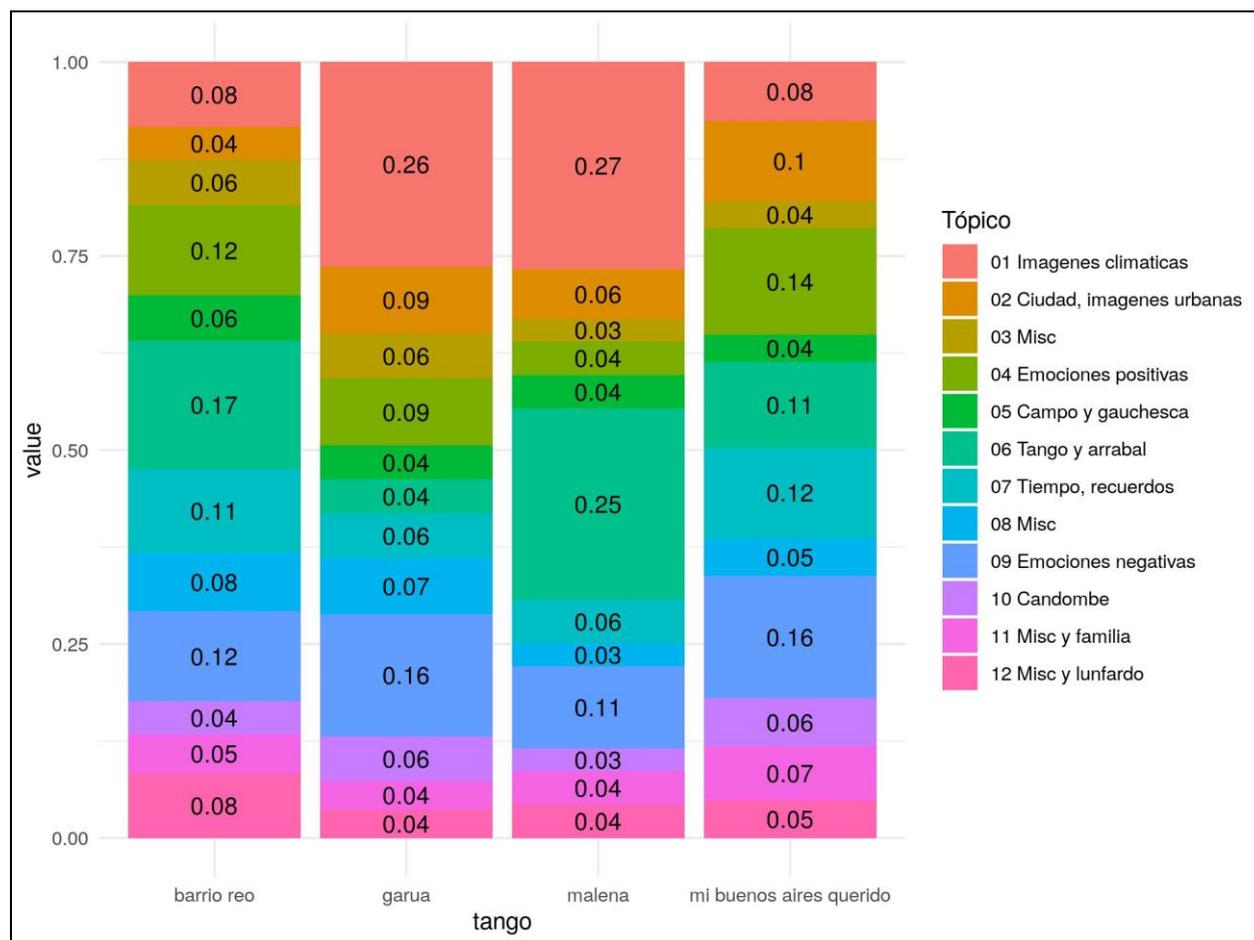
¹⁸ Así, por ejemplo, en otros modelos no supervisados como el análisis de componentes principales (PCA, por sus siglas en inglés) o en las diversas formas de *clustering*, suele resultar necesaria la intervención del investigador para hacer inteligibles algunos de los *outputs* del modelo.

6.2. Algunos análisis posibles

Una vez detectados los tópicos es posible estimar para cada documento d del corpus C en qué proporción presenta cada uno de los tópicos. Es decir, para cada letra de tango será posible calcular qué composición muestra de cada uno de los 12 tópicos detectados. Una forma útil de validación de los resultados es explorar si el análisis manual de los tangos coincide con los tópicos que fueron detectados de forma automática.

A continuación, se exponen, a modo de ejemplo, la composición de tópicos de cuatro tangos de tres décadas diferentes:

Gráfico 2. Composición de tópicos según tangos, 1900-2010



Fuente: Elaboración propia en base a datos recolectados de todotango.com

Un tango como “Barrio reo”, que habla de los gratos recuerdos del cantor al retornar a su barrio y de la tristeza que le produce el encuentro con su deterioro (“Hoy te encuentro envejecido”), muestra valores altos en los tópicos *emociones positivas*, *emociones negativas* y en el que habla sobre el *tango y arrabal*.

El tango “Garúa” tematiza una caminata del narrador bajo la llovizna y pinta un cuadro lúgubre y oscuro (“sobre la calle la hilera de focos/ lustra el asfalto con luz mortecina”) mientras el caminante recuerda a la mujer que, presumiblemente, se fue. Es por eso que las *emociones negativas* y las *imágenes climáticas* aparecen con fuerza en este tango.

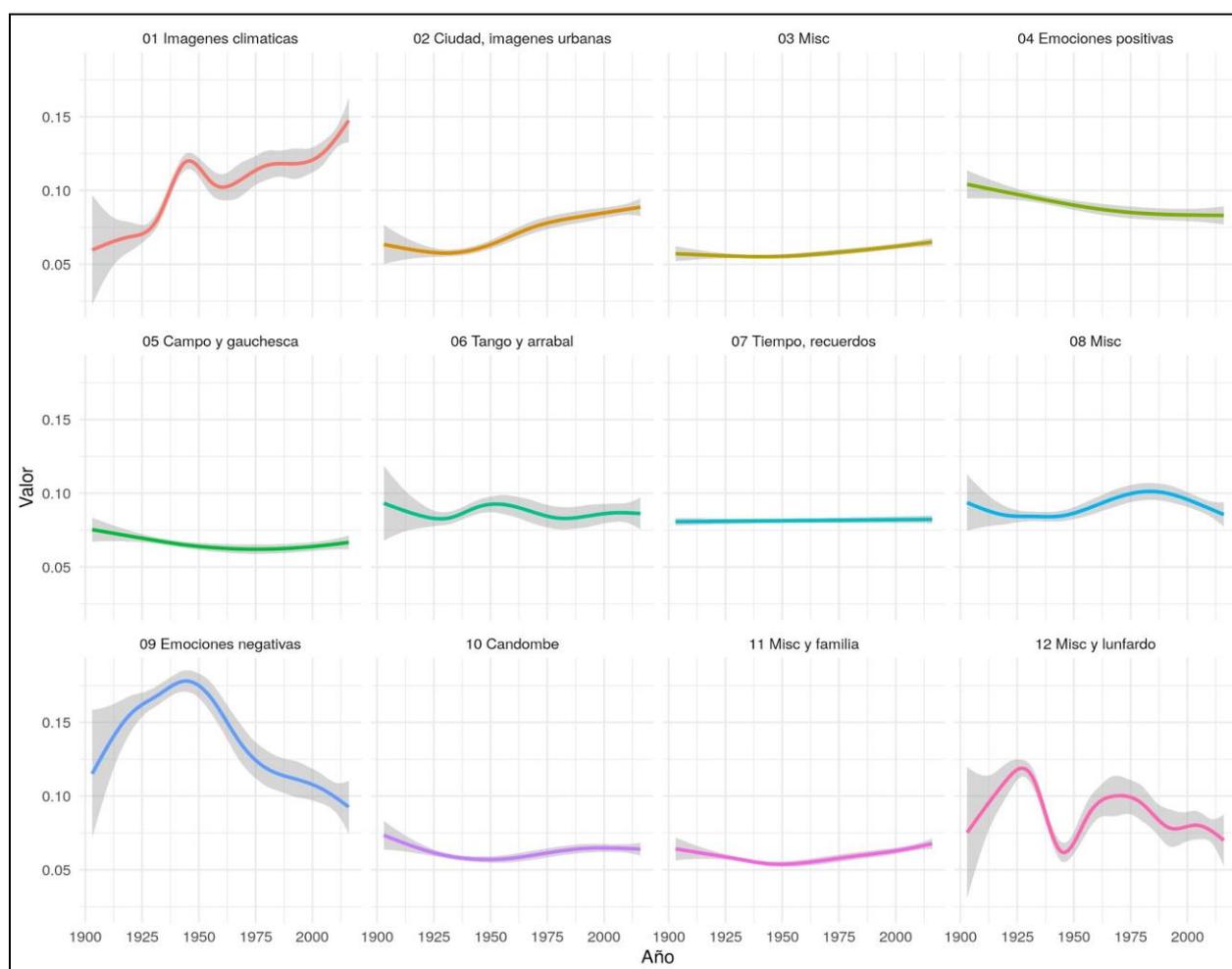
“Malena” habla (en tercera persona) de una cantante de tangos que pasó por desamores, que parece tener cierta predisposición a la bebida y que “canta el tango como ninguna”. Esto se corresponde con la composición de tópicos detectada: *emociones negativas*, *tango y arrabal* e

imágenes climáticas (“tono oscuro”, “el frío del último encuentro”, “tus manos son palomas que sienten frío”).

Por último, en “Mi Buenos Aires querido”, se menciona a la ciudad, la nostalgia del narrador y la esperanza del retorno. Esto se corresponde con los tópicos detectados: *ciudad*, *emociones negativas*, *emociones positivas* y *tiempo y recuerdos*.

Al mismo tiempo, sería posible analizar la evolución de cada uno de los tópicos a lo largo de las diferentes décadas.

Gráfico 3. Evolución de los tópicos, 1900-2010 (suavizado GAM)



Fuente: Elaboración propia en base a datos recolectados de todotango.com

En primer lugar, las imágenes naturales y climáticas ganan predominio de forma sostenida a lo largo del tiempo. En mucha menor medida, el tópico vinculado a la ciudad parece cobrar cierta importancia a partir de la década del 1930. También resultan interesantes las oscilaciones que parece presentar el tema del tango y el arrabal. Parece caer levemente hacia la década del 1920 y vuelve a incrementarse hacia los años '50, mostrando otro valle hacia los años '70.

Pero quizás uno de los cambios más importantes es el que se observa en los tópicos 4 y 9. En efecto, puede verse que la participación de las emociones positivas es relativamente constante a lo largo del tiempo. En cambio, son las emociones negativas las que muestran diferencias más marcadas: se ve una tendencia al crecimiento hasta la década del 1940 y 1950. Posteriormente, tiende a disminuir su importancia.

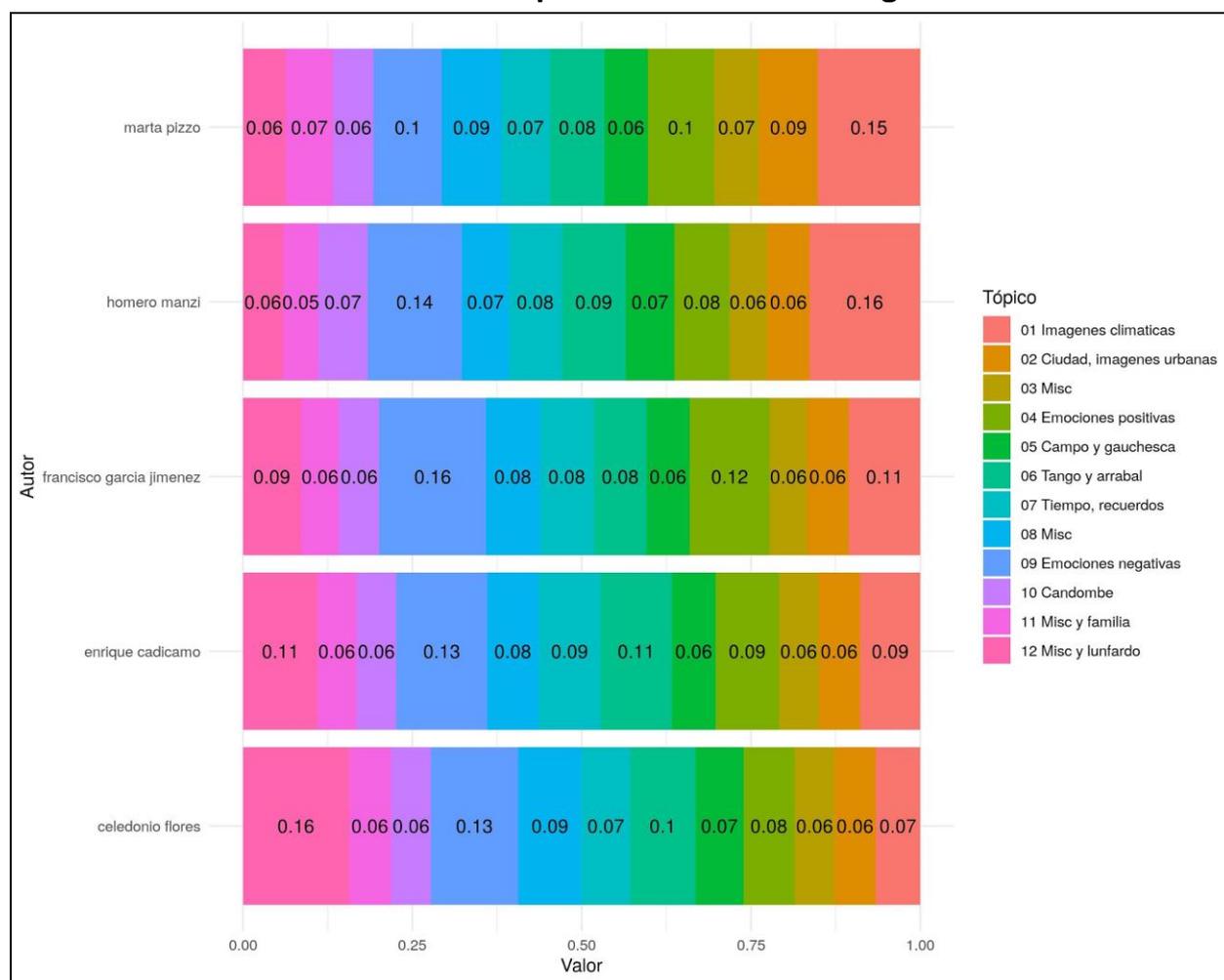
Este punto es interesante porque se vincula con algunas de las discusiones dentro del campo de los estudios del tango. Así, por ejemplo, Borges (2016: 80-81), planteaba hacia la

década del 1960 que “el tango, como hemos visto, empezó, surge de la milonga, y es al principio un baile valeroso y feliz. Y luego, el tango va languideciendo y entristeciéndose...”. De esta forma, puede verse que esta hipótesis parece tener cierto sustento y habilita el estudio más profundo de esta problemática.¹⁹

A partir de las series anteriores podrían plantearse preguntas que interroguen sobre la vinculación existente entre tales cambios en los temas del tango y los procesos de desarrollo y expansión del capitalismo en Argentina y/o a los movimientos de migración rural-urbana.

También, es posible calcular la composición promedio de los diferentes tópicos en cada autor de tango.

Gráfico 4. Composición de tópicos según autores, 1900-2010.
Media de la composición de las letras de tango



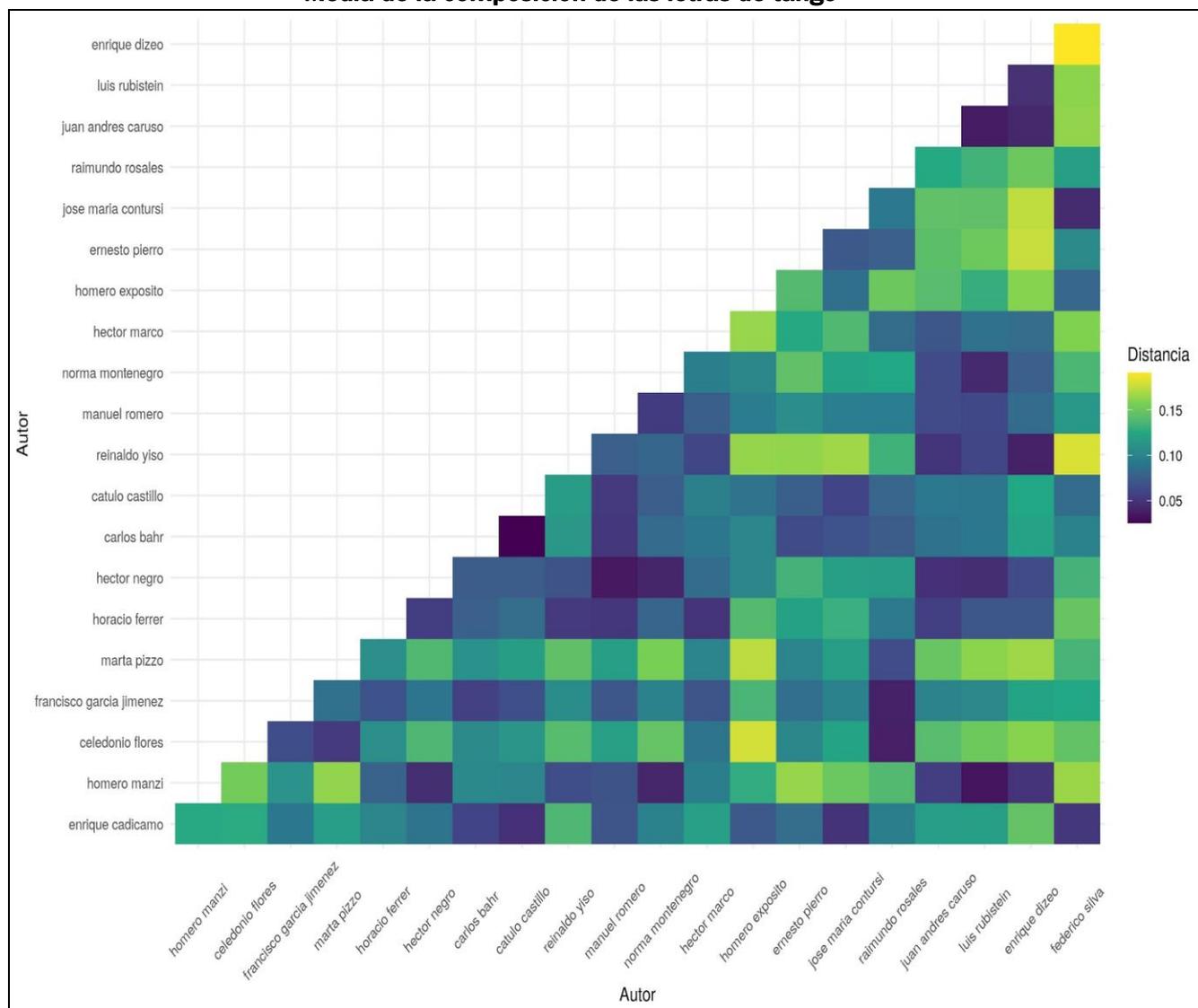
Fuente: Elaboración propia en base a datos recolectados de todotango.com

Si se analizan tales métricas para los cinco autores con mayor cantidad de letras en el corpus, se hace manifiesto que los temas predominantes de Cadícamo parecen ser el lunfardo y las emociones negativas. Esto contrasta con Homero Manzi, quien parece utilizar en mayor medida las imágenes climáticas (y también las emociones negativas).

¹⁹ En el mismo texto (Borges, 2016), discute sobre las causas de dicho cambio: explora hipótesis cuasi-sociológicas sobre la influencia negra, la italiana y otras hipótesis musicológicas, tales como la importancia de la introducción del bandoneón.

Finalmente, puede construirse una matriz de distancias para cada autor en función de la composición promedio de sus tópicos, con el objetivo de encontrar autores que utilizan temas similares.²⁰

Gráfico 5. Distancias en la composición de tópicos según autores, 1900-2010. Media de la composición de las letras de tango



Fuente: Elaboración propia en base a datos recolectados de todotango.com

Celedonio Flores y Homero Manzi parecen ser los que mayores similitudes tienen en relación a los tópicos que utilizan. Algo parecido pasa con Enrique Dizeo y José María Contursi, por un lado y con Ernesto Pierro por el otro.

7. Discusión y comentarios finales

En el presente trabajo se buscó presentar una aproximación metodológica posible para el análisis automático de textos a partir de la aplicación de una técnica de detección de tópicos (LDA) sobre un corpus grande de letras de tango. A su vez, se presentó un flujo de trabajo posible para dicho análisis y se discutieron algunas técnicas para el pre-procesamiento del texto (eliminación de

²⁰ De forma análoga podría construirse una matriz a nivel de tango y buscar los tangos que hablan de tópicos similares.

stopwords y otros signos, normalización de conteos, construcción de una *TFM* mediante el modelo *Bag of Words*).

De esta forma, fue posible identificar los principales temas del tango. El uso de emociones positivas y negativas, imágenes de la ciudad, sobre el tango y el arrabal, sobre el campo y la gauchesca, sobre la temporalidad y la memoria, entre otros, aparecen como los más importantes. Al mismo tiempo, fue posible validar los tópicos a partir de la selección de algunos tangos y del análisis de sus letras evaluando su correspondencia con los tópicos estimados.

Finalmente, fue posible visualizar las diferencias (distancias) entre los tópicos utilizados por diferentes autores.

Ahora bien, más allá de los resultados del ejercicio propuesto (acotados a mostrar un caso de uso de las herramientas), el trabajo busca ilustrar las potencialidades que este tipo de técnicas tiene para la investigación sobre datos textuales en ciencias sociales. Así, con una herramienta que permita analizar con un *corpus* amplio y de forma más sistemática la evolución de los temas del tango (u otros géneros), sería posible vincular analítica y metodológicamente la dimensión cultural con otras esferas de la estructura social.²¹ Quizás una de las posibilidades analíticas más interesantes fue la de poder visualizar la evolución temporal de los tópicos y la eventual posibilidad en trabajos más sistemáticos de plantear hipótesis sobre la vinculación con procesos más generales de la sociedad argentina (los cambios en la estructura económica, movimientos migratorios, etc.).

La detección de tópicos ha sido utilizada en los últimos tiempos para el análisis literario (Jockers & Mimno, 2013), el estudio de comunicados políticos (Grimmer, 2010), de medios (Wiedemann, 2016; DiMaggio, Nag, y Blei, 2013) y de temas en leyes y proyectos (Gerrish & Blei, 2012), por nombrar algunas aplicaciones relevantes.

En efecto, sus principales ventajas radican en su escalabilidad y en sus mayores posibilidades de replicabilidad: utilizando las técnicas de análisis cualitativo de textos “tradicionales”, es posible lograr gran profundidad analítica, pero sobre *corpus* más bien pequeños o medianos y, en muchos casos, con resultados poco o medianamente replicables. Los antecedentes mencionados tenían una escala más bien pequeña: alrededor de 30 letras de tango, con la excepción del texto de Cantón (1972). El proceso de detección de tópicos encarado en el presente trabajo procesó y analizó 5.617 letras de tango.

No obstante, en este tipo de enfoques existen algunas limitaciones que es importante remarcar. En relación a la construcción del *corpus*, la dependencia de la fuente de datos a utilizar resultó una primera desventaja (si bien, no tan diferente al uso de compilaciones de letras de tango realizadas por diferentes autores). En la etapa de análisis, los modelos de detección de tópicos presentan problemas para el tratamiento adecuado de algunas figuras retóricas (como, por ejemplo, las metáforas). También se presentan como problemas potenciales a resolver, la correcta determinación del número de tópicos y el etiquetado de los mismos. La complementación con análisis en profundidad de los documentos del *corpus* resulta, en este punto, de suma utilidad.

Por ello, es importante remarcar que el uso de técnicas de NLP no implica un desplazamiento de los enfoques basados en la interpretación manual de los documentos. Un trabajo (Baumer et al, 2017) compara los resultados obtenidos utilizando dos métodos de análisis sobre un mismo *corpus* de datos: generación de categorías utilizando la metodología de la *Grounded Theory* y detección de tópicos utilizando LDA. Las conclusiones de dicho trabajo sugieren tanto coherencia como complementariedad entre los resultados de ambos métodos.

²¹ Como hemos mencionado más arriba, uno de los supuestos de LDA en su versión básica, es que los tópicos preexisten a los textos y son constantes en el tiempo. Se trata de un supuesto fuerte para un análisis temporal. Es por ello que existen otras versiones de modelado de tópicos que permiten flexibilizar estos supuestos: *dynamic topic modeling*, por ejemplo (Blei, 2012).

8. Bibliografía

- ASUNCIÓN, A., WELLING, M., SMYTH, P., y TEH, Y. W. (2009). On smoothing and inference for topic models. *Proceedings of the twenty-fifth conference on uncertainty in artificial intelligence*, pp-pp. 27-34. Disponible en <http://dl.acm.org/citation.cfm?id=1795114.1795118>. Fecha de consulta: 20/08/2019.
- BAUMER, E., MIMMO, D., GUHA, S., QUAN, E., & GAY, G. (2017) Comparing grounded theory and topic modeling: Extreme divergence or unlikely convergence?" *Journal of the Association for Information Science and Technology* 68 (6). Disponible en <https://mimmo.infosci.cornell.edu/papers/baumer-jasist-2017.pdf>. Fecha de consulta: 12/07/2019.
- BLEI, D. (2012). Probabilistic topic models. *Communications of the ACM* 55 (4), 77-84.
- BORGES, J. L. (2016). *El tango. Cuatro conferencias*. Buenos Aires: Sudamericana.
- CANTÓN, D. (1972). *Gardel, ¿a quién le cantás?* Buenos Aires: De la Flor.
- CASANOVA CORREA, J. y PAVON RABASCO, F. (2002). Nuevas herramientas para el procesamiento de datos cualitativos. *Ágora Digital* 3, 1-13. Disponible en: http://rabida.uhu.es/dspace/bitstream/handle/10272/6616/Nuevas_herramientas.pdf;jsessionid=000A8A1862E6EDF881EA03D14ABCC096?sequence=2
- CHANG, J., BOYD-GRABER, J., WANG, C., GERRISH, S., y BLEI, D. (2009). Reading tea leaves: How humans interpret topic models. *Advances in Neural Information Processing Systems*, 32:288-296. Disponible en: <https://papers.nips.cc/paper/3700-reading-tea-leaves-how-humans-interpret-topic-models>.
- DIMAGGIO, P., NAG, M., y BLEI, D. (2013). Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of U.S. government arts funding. *Poetics* 41(6), 570-606.
- GASPARRI, J. (2011). Che varón, masculinidades en las letras de tango. *Revista Caracol* 2, 175-215.
- GEERTZ, C. (1974). Deep play: Notes on the balinese cockfight. *The interpretation of cultures* (pp-pp. 412-453). New York: Basic Books.
- GERRISH, S., y BLEI, D. (2012). How they vote: Issue-adjusted models of legislative behavior. F. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems* 25, 2753-2761. Disponible en <http://papers.nips.cc/paper/4715-how-they-vote-issue-adjusted-models-of-legislative-behavior.pdf>
- GRIMMER, J. (2010). A bayesian hierarchical topic model for political texts: Measuring expressed agendas in senate press releases. *Political Analysis* 18 (1), 1-35.
- GRIMMER, J., y STEWART, M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis* 21 (3), 267-297.
- HASSANI, A., IRANMANESH, A., y MANSOURI, N. (2019). *Text mining using nonnegative matrix factorization and latent semantic analysis*. Disponible en: <https://arxiv.org/abs/1911.04705>.
- HEINTZ, I., GABBARD, R., SRINIVASAN, M., BARNER, D., BLACK, D., FREEDMAN, M. y WEISCHEDEL, R. (2013). Automatic extraction of linguistic metaphors with lda topic modeling. *Proceedings of the First Workshop on Metaphor in NLP*. Disponible en: <https://www.aclweb.org/anthology/W13-0908.pdf>.
- HSIEH, H.-F., y SHANNON, S. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research* 15 (9), 1277-1288.
- JOCKERS, M., y MIMMO, D. (2013). Significant themes in 19th-century literature. *Poetics*, 41 (6), 750-769.

- JURAFSKY, D. Y MARTÍN, J. (2008). *Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. New Jersey: Prentice Hall.
- LÓPEZ, I. (2010). Morochas, milongueras y percantas. Representaciones de la mujer en las letras de tango. *Espéculo. Revista de Estudios Literarios* 45. Disponible en: <https://webs.ucm.es/info/especulo/numero45/mutango.html>
- MARCHESE, M. (2007). Tango. El lenguaje quebrado del desarraigo". *Revista Latinoamericana de Estudios Del Discurso* 6 (2), 45-60.
- MAYRING, P. (2000). Qualitative Content Analysis. *Forum: Qualitative Social Research* 2(1). Disponible en: <https://www.qualitative-research.net/index.php/fqs/article/view/1089/2385>
- MAXWELL, J. (2004). Reemergent scientism, postmodernism, and dialogue across differences. *Qualitative Inquiry* 10 (1), 35-41.
- MIMNO, D., WALLACH, H., TALLEY, E., LEENDERS, M., y MCCALLUM, A. (2011). Optimizing semantic coherence in topic models. *Empirical methods on natural language processing*. Disponible en: <https://mimno.infosci.cornell.edu/papers/mimno-semantic-emnlp.pdf>
- MITCHELL, R. (2015). *Web scraping with python: Collecting data from the modern web*. California: O'Reilly.
- MURH, T. (1997). *Atlas.Ti-Visual Qualitative Data Analysis-Management-Model Building-Release 4.1*. Berlín. Consultado de <https://personalpages.manchester.ac.uk/staff/andrew.j.howes/manshort.pdf>
- NAVARRO-COLORADO, B., TOMÁS, D., VÁZQUEZ, S., MOREDA, P., IZQUIERDO, R., SAQUETE, E., LLOPIS, F. (2011). Procesamiento automático de metáforas con métodos no supervisados. *Procesamiento del Lenguaje Natural*, 47, 345-346.
- RAGIN, C. (1987). *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. California: University of California Press.
- REYNOSO, C. (2007). El lado oscuro de la descripción densa. *Anthropologika. Revista de Estudio e Investigaciones en Antropología*, 1(1), 136-196.
- RICHARDS, L. (1999). *Using Nvivo In Qualitative Research*. London: Sage.
- ROSATI, G. y CHAZARRETA, A. (2017). El Qualitative Comparative Analysis (QCA) como herramienta analítica. Dos aplicaciones para el análisis de entrevistas. *Revista Latinoamericana de Metodología de las Ciencias Sociales* 7 (1) <https://doi.org/10.24215/18537863e018>
- RUSSO, J. Y MARPEGÁN, D. (2000). *Letras de tango*. Buenos Aires: Basílico.
- TIE, Y., BIRKS, M., y FRANCIS, K. (2019). Grounded theory research: A design framework for novice researchers. *SAGE Open Medicine*, 7. Doi: [10.1177/2050312118822927](https://doi.org/10.1177/2050312118822927)
- VASILACHIS, I. (2006). La investigación cualitativa. En Vasilachis I. (comp). *Estrategias de investigación cualitativa* (pp-pp. 23-6). Barcelona: Gedisa.
- WIEDEMANN, G. (2016). *Text mining for qualitative data analysis in the social sciences. A study on democratic discourse in germany*. Berlín: Springer.
- WILLENPART, L. (2011). El tango: Temas y motivos. *Verba Hispanica*, 9 (1), 219-230.

Anexos:**Tablas con los primeros 8 términos para estimación de tópicos con diferentes k .****Tabla A.1 $k=15$**

topic	V1	V2	V3	V4	V5	V6	V7	V8
Topic 1	noche	sol	tiempo	cielo	voz	dos	luna	luz
Topic 2	vino	hombres	siempre	par	habia	dieron	flor	invierno
Topic 3	dice	mil	muerte	paz	filo	frente	gente	entero
Topic 4	amor	corazon	vida	quiero	dolor	alma	solo	hoy
Topic 5	alguien	final	cada	lejos	dio	niño	pasiones	sabia
Topic 6	siete	historia	volvio	decir	ahora	feliz	ver	viejos
Topic 7	tango	barrio	buenos	aires	milonga	canto	cancion	bandoneon
Topic 8	vamos	quieren	aunque	lleva	siente	oye	presencia	primavera
Topic 9	dias	cuatro	locos	quedo	circo	van	boca	hombres
Topic 10	lleva	gente	fuerte	rueda	sentido	quiero	romance	cuatro
Topic 11	tierra	linda	habia	gaucho	rancho	dijo	china	huella
Topic 12	negro	negra	maria	carnaval	libertad	sangre	niño	hijo
Topic 13	dice	hoy	anoche	puro	rosa	hizo	lagrimas	entro
Topic 14	mano	ahora	dire	paso	saben	traiga	media	casi
Topic 15	vos	sos	bien	hoy	ser	siempre	vida	vas

Fuente: Elaboración propia en base a datos recolectados de todotango.com

Tabla A.2 $k=13$

topic	V1	V2	V3	V4	V5	V6	V7	V8
Topic 1	mira	alli	dia	salio	frente	pelo	vendra	edad
Topic 2	paso	ser	llevo	partida	loco	años	bandera	fuerte
Topic 3	sangre	medio	dice	viene	tumba	vez	cuerpo	quedan
Topic 4	noche	voz	tiempo	cielo	luna	dos	sol	calle
Topic 5	tierra	dijo	don	viejo	tenia	gaucho	grito	rancho
Topic 6	tango	barrio	buenos	aires	milonga	viejo	bandoneon	arrabal
Topic 7	vamos	dias	loco	cuatro	libertad	viva	año	dale
Topic 8	amor	vida	corazon	solo	quiero	dolor	hoy	alma
Topic 9	pues	trabajo	quedo	tras	vida	pies	vio	algun
Topic 10	amor	flor	ojos	corazon	alma	dulce	cancion	ilusion
Topic 11	siempre	ser	vida	mundo	nadie	cosas	hombre	mismo
Topic 12	negro	sangre	negra	maria	niño	carnaval	risa	cuerpo
Topic 13	vos	sos	bien	hoy	vas	tenes	gran	pobre

Fuente: Elaboración propia en base a datos recolectados de todotango.com

Tabla A.3 k=10

topic	V1	V2	V3	V4	V5	V6	V7	V8
Topic 1	amor	corazon	alma	flor	ilusion	ojos	pasion	dulce
Topic 2	cada	buenos	aires	calle	ciudad	esquina	calles	van
Topic 3	noche	voz	cielo	dos	adios	luna	sol	manos
Topic 4	tierra	gran	don	gloria	alli	huella	grito	gaucho
Topic 5	tango	barrio	milonga	canto	bandoneon	arrabal	cancion	cantar
Topic 6	pobre	triste	día	noche	aquella	madre	pena	dio
Topic 7	amor	vida	corazon	quiero	solo	dolor	nunca	alma
Topic 8	ser	siempre	mundo	nadie	hombre	voy	aqui	mejor
Topic 9	hoy	tiempo	ayer	vida	viejo	años	recuerdo	vez
Topic 10	vos	sos	bien	vas	tenes	gran	hace	che

Fuente: Elaboración propia en base a datos recolectados de todotango.com

Tabla A.4 k=5

topic	V1	V2	V3	V4	V5	V6	V7	V8
Topic 1	triste	vida	ojos	día	pobre	alma	pena	noche
Topic 2	dos	noche	voz	tiempo	cielo	sol	adios	cada
Topic 3	tango	viejo	barrio	buenos	aires	milonga	canto	cancion
Topic 4	amor	corazon	vida	solo	quiero	dolor	hoy	nunca
Topic 5	vos	bien	sos	ser	vas	siempre	hombre	hoy

Fuente: Elaboración propia en base a datos recolectados de todotango.com

Autor.

Germán Rosati

Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET); Universidad Nacional de San Martín (UNSAM); Instituto de Altos Estudios Sociales (IDAES); Programa de Investigaciones sobre el Movimiento de la Sociedad Argentina (PIMSA), Argentina

Investigador Asistente del CONICET. Doctor en Ciencias Sociales por la Universidad de Buenos Aires (UBA). Magíster en Generación y Análisis de Información Estadística por la Universidad Nacional de Tres de Febrero (UNTREF). Licenciado en Sociología por la Universidad de Buenos Aires (UBA). Docente en la carrera de Sociología y Antropología (IDAES-UNSAM).

E-mail: german.rosati@gmail.com

Citado.

ROSATI, Germán (2022). "Procesamiento de Lenguaje Natural aplicado a las ciencias sociales. Detección de tópicos en letras de tango". *Revista Latinoamericana de Metodología de la Investigación Social - ReLMIS*. N°23, Año 12, pp. 38-60.

Plazos.

Recibido: 23/12/2019. Aceptado: 20/10/2020.



Aplicación de la Teoría Fundamentada en una investigación sobre travestis y mujeres trans sudamericanas que residen en el AMBA

Application of the grounded theory based on an investigation on transvestites and South American trans women residing in the AMBA

Ramiro Nicolás Perez Ripossio

Resumen

El artículo describe una serie de procedimientos metodológicos aplicados en una investigación empírica cualitativa sobre los proyectos migratorios de las travestis y mujeres trans sudamericanas que residen en el Área Metropolitana de Buenos Aires. Se emplea el método de la teoría fundamentada y de manera auxiliar el relato de vida. Así, se explica el modo en que fue elaborado el muestreo teórico y aplicado el método de la comparación constante. Estos procesos llevan a considerar el procesamiento de la información, el uso del software atlas ti y la codificación elaborada (abierta, axial y selectiva), teniendo en cuenta que el método propuesto por la teoría fundamentada establece que el investigador se encuentra en condiciones de delimitar y escribir la teoría emergente sobre un determinado fenómeno, cuando se alcanza la saturación teórica. Para ello se especifica en qué consiste esta tarea y cuáles son los criterios para determinarla. El artículo aporta una manera de aplicar la Teoría Fundamentada que resulta útil para los abordajes pedagógicos didácticos que se propongan la enseñanza del quehacer investigativo.

Palabras clave: Teoría fundamentada; métodos cualitativos; muestreo teórico; travestis/trans; migraciones.

Abstract

The article describes a series of methodological procedures applied in a qualitative empirical research on the migratory projects of transvestites and South American trans women residing in the Metropolitan Area of Buenos Aires. The method of grounded theory is used and, in an auxiliary way, the life story. Thus, the way in which the theoretical sampling was developed and the constant comparison method applied is explained. These processes lead to consider the information processing, the use of the atlas ti software and the elaborate coding (open, axial and selective), taking into account that the method proposed by the grounded theory establishes that the researcher is in a position to delimit and write the emerging theory about a certain phenomenon, when the theoretical saturation is reached. For this, it is specified what this task consists of and tasks are the criteria to determine it. The article provides a way of applying Grounded Theory that is useful for didactic pedagogical approaches that are proposed to teach research.

Keywords: Grounded theory; qualitative methods; theoretical sampling; transvestites/trans; migrations.

1. Introducción¹

El artículo describe algunos procedimientos y herramientas metodológicas aplicadas en una investigación durante el período 2017-2019 sobre los proyectos migratorios de las travestis y mujeres trans sudamericanas que actualmente residen en el Área Metropolitana de Buenos Aires. Para ello se expone una aplicación del método propuesto por la teoría fundamentada de Strauss y Corbin (2002) y de manera auxiliar y en combinación, se exponen algunos elementos del relato de vida.

Las travestis y mujeres trans migran hacia el AMBA de acuerdo a una pluralidad de motivos, pero, por lo general, se debe a la hostilidad que experimentan en las sociedades emisoras. Los proyectos migratorios, de este universo de estudio, son pendulares, polimórficos e inestables. Sin embargo, pueden diferenciarse algunas fases, entre ellas: los motivos de la migración, los modos de migrar que se constituyen mediante cadenas, la asimilación en la sociedad receptora, la transitoriedad hacia destinos europeos y las experiencias de regreso hacia las sociedades emisoras. En el AMBA, experimentan mayores libertades, acceden a derechos y entablan vínculos sexoafectivos, aunque también vivencian una reconfiguración en las discriminaciones producto de que la condición étnica nacional se presenta como una característica que es jerarquizada y en conjunción con sus identidades de género.

A pesar de que en los últimos veinte años se ha constituido un campo de investigaciones sobre procesos migratorios y diversidad sexual (Stang, 2018), las investigaciones que han tomado como objeto la migración de travestis/trans sudamericanas son aún escasas. A nivel internacional, las tradiciones anglosajonas e ibéricas, han realizado aportes permitiendo aseverar que la migración *queer* (Manalansan, 2006; García y Oñate, 2008, entre otras) es un hecho social de relevancia mundial. Estas investigaciones se centran en poblaciones ubicadas en el colectivo LGTTTBIQ+ (lesbianas, gay, travestis, transexuales, transgéneros, bisexuales, intersexuales y *queer*) y permiten entender los motivos de la migración, la discriminación, la violencia y la readecuación de la corporalidad mediante diversas intervenciones.

En Latinoamérica, de manera puntual, se han elaborado investigaciones etnográficas sobre travestis y trans brasileñas que migraron hacia Europa (Pelúcio, 2009 y Vartabedian, 2012, entre otras). Estas han demostrado que la migración se produce como consecuencia de la discriminación y que aspectos asociados al *glamour* y al encanto de la brasileidad son identificados en los mercados sexuales europeos como atractivos, por lo que ejercer la prostitución o el trabajo sexual representa una estrategia de supervivencia común ante las dificultades para acceder al trabajo formal. En México (López Fernández, 2018; Zarco Ortiz y Chacón Reynosa, 2020), se han problematizado las tácticas y estrategias para atravesar las fronteras mediadas por dispositivos de seguridad que obligan a *performar* la identidad de género de manera binaria y; por lo tanto, enmascarar la expresión de género travesti/trans. Por último, en el ámbito local, de manera indirecta, Fernández (2004), Berkins (2007) y Boy (2017) han realizado aportes para reflexionar sobre las migraciones internas de personas travestis y trans que desde distintas provincias migran a centros urbanos con el fin de buscar el anonimato de la gran ciudad y experimentar mayores libertades.

El artículo se divide en cuatro apartados. En primer lugar, se mencionan algunas premisas del enfoque cualitativo y se describen los métodos y técnicas empleadas. En segundo lugar, el modo en que fue aplicado el muestreo teórico de acuerdo con los criterios para seleccionar los casos. Luego, explica los procedimientos vinculados al procesamiento, y, por último, los procedimientos de análisis de la información, es decir, la aplicación del método de la comparación constante.

¹ El artículo forma parte de una tesis doctoral en Sociología en la Universidad de Buenos Aires ya concluida. Agradezco especialmente a Miury Cirone por las discusiones y reflexiones presentadas en las XIII Jornadas de Sociología de la Universidad de Buenos Aires sobre este tema.

2. Enfoque métodos y técnicas

Como fue mencionado en la introducción, el enfoque empleado en esta investigación fue cualitativo y sin intención de definirlo de manera acabada, es posible determinar que este se propone comprender los sentidos y significados que las personas le otorgan a sus acciones observando sus perspectivas (Scribano, 2008).

Por otra parte, los métodos refieren a los instrumentos y las concepciones sobre cómo alcanzar la cientificidad. Según Marradi, “Lo esencial del concepto de método está en lo siguiente: en la elección de las técnicas a aplicar, en la capacidad de modificar técnicas existentes, adaptándolas a los propios problemas, y en imaginar técnicas nuevas” (2002: 119). El principal método empleado por esta investigación fue la Teoría Fundamentada, aunque de manera auxiliar, se utilizaron la etnosociología mediante el relato de vida. Con fundamento en De Sena, Lisdero y Scribano (2015), un diseño multimetódico se puede construir en función de tres posibilidades: triangulación, complementación y combinación. En este caso, se empleó una combinación en simultáneo, acorde con los métodos mencionados y al interior del enfoque cualitativo estableciendo sus respectivas jerarquías. Sin desconocer las críticas a la triangulación ni cayendo en un realismo ingenuo, esta tarea se orientó a comprender en profundidad el fenómeno de estudio paliando los déficits de cada uno de los métodos y técnicas empleadas. En otras palabras, “no existe el método para dar cuenta de un fenómeno, contrariamente, existen métodos o combinación de métodos que mejor posibilitan el tratamiento de la base empírica a partir de conceptos” (Cohen y Gómez Rojas, 2019: 22).

2.1. La Teoría Fundamentada

La Teoría Fundamentada es un método de investigación cuyas bases se exponen en este artículo en función de la versión e interpretación presentada por Anselm Strauss y Juliet Corbin en el libro titulado: *Bases de la investigación cualitativa: técnicas y procedimientos para desarrollar la teoría fundamentada*, aunque esto no significa que no puedan contemplarse otras referencias que resulten pertinentes. Se trata de una estrategia de producción y de análisis cualitativo que consiste en una forma de pensar y mirar el mundo. La interpretación sobre la Teoría Fundamentada que ha prevalecido en el ámbito académico ha sido considerarla un método inductivo que parte de la relación del investigador con los datos y, a partir de esa instancia es que se constituye la teoría sustantiva sobre un fenómeno. La teoría es generada valiéndose del muestreo teórico, además de un proceso de recolección de información mediante el cual el analista reúne, codifica y analiza los datos conjuntamente y toma decisiones sobre qué información debe recopilarse en lo sucesivo y dónde encontrarla.

Glaser y Strauss (cofundadores del método) (1999 [1967]) dejan en claro en los distintos párrafos de la obra *The Discovery of Grounded Theory Strategies for Qualitative Research* e Glaser y Strauss que los investigadores *no* se encuentran despojados de teorías ni conceptos. En el siguiente párrafo afirman:

(...) la teoría sustantiva es un vínculo estratégico en la formulación y generación de teoría formal fundamentada. Creemos que aunque la teoría formal se puede generar directamente a partir de los datos, es más deseable, y generalmente necesario, comenzar la teoría formal desde una sustantiva. (Glaser y Strauss, 1999 [1967]: 79).

De esta manera, queda claro que la teoría se construye a partir de los datos cuando la investigación carece de una teoría aplicable al problema de investigación, pero si otras teorías sustantivas han abordado procesos emparentados con el propio, entonces esas teorías se formalizan y representan un marco para luego, en sucesivas investigaciones, verificarlas, ampliarlas y modificarlas, tal como sucedió en la investigación expuesta en este artículo. Lo que se pretende señalar es que “... muchos investigadores piensan que es difícil entrar al campo sin ningún concepto sobre lo que se va a estudiar” (Strauss y Corbin, 2002: 223), por lo que es apropiado dejar de lado presupuestos ingenuos como la ausencia de teoría al investigar un determinado fenómeno. En todo caso, en la mayoría de las investigaciones se cuenta con un marco conceptual flexible que puede modificarse, contradecirse o, incluso, robustecerse conforme avanza la investigación mediante descubrimientos y hallazgos.

El proceso de producción de datos, en este método, es controlado por la teoría emergente y, más allá de las decisiones relativas a una recolección inicial, es prudente complementar esta elaboración una vez que asoman los primeros atisbos teóricos que se convertirán luego en una teoría general (Glaser y Holton, 2004). Cabe destacar que Strauss y Corbin (2002) señalaron que el propósito de la Teoría Fundamentada también podría consistir en ampliar y perfeccionar la teoría existente. Por eso, este artículo se diferencia de posiciones que consideran que los investigadores pueden aproximarse al campo despojados de nociones teóricas y construcciones hipotéticas.

Por otra parte, con base en Charmaz (2006) la teoría fundamentada es el resultado de la confluencia entre la escuela de Chicago y la escuela de Columbia, ya que sus fundadores pertenecieron a estas instituciones. Barney Glaser, por ejemplo, se encontraba influido por los criterios de codificación aplicados por Paul Lazarsfeld en investigaciones cuantitativas. Por eso, algunos de los procedimientos de codificación se asemejan a la operacionalización de variables. Hasta podría afirmarse que poseen considerables similitudes. No obstante, características como la exhaustividad y la mutua exclusión no se consideran criterios decisivos en la elaboración de las categorías de análisis. Además, la Teoría Fundamentada sienta sus bases teóricas en el interaccionismo simbólico de Herbert Mead (Páramo Morales, 2015), con lo cual se conjugan procesos rigurosos de codificación con procedimientos de análisis interpretativos y minuciosos del material empírico. Profundidad, sistematicidad y rigurosidad expresan las principales bases de este método.

La base del método de la Teoría Fundamentada establece que la teoría es emergente. Por lo tanto, la investigación no comienza con un marco teórico o un diseño rígido, preestablecido e invariable, sino con un contexto nutrido por conceptos sensibilizadores. Es preciso dejar espacio para que el diseño —al igual que los conceptos— emerja de los datos sin caer en un ingenuo empirismo. Si se utiliza de manera adecuada, la Teoría Fundamentada permite obtener un enfoque claro sobre aquello que está sucediendo sin sacrificar el detalle. “Como una cámara con muchos lentes, primero se observa una amplia barrida del paisaje. Posteriormente, al cambiar sus lentes tantas veces como sea necesario es posible acercarse cada vez más a las escenas a la vista” (Charmaz, 2006: 14). Las categorías iniciales, con sus propiedades y dimensiones, son generadas mediante una minuciosa indagación conocida como *microanálisis* y *análisis línea por línea*. Cabe destacar que las categorías refieren a conceptos que representan fenómenos; las propiedades son atributos de las categorías que les otorgan significado y, mediante su desarrollo, permiten que se constituyan como teóricamente densas, y las dimensiones remiten al intervalo mediante el cual las propiedades varían.

2.2. El método biográfico

Las biografías de las personas pueden tener significado social y la Sociología ha entendido esto desde sus orígenes. Así, la etnosociología es definida por Bertaux (2005) del siguiente modo:

La hipótesis central de la perspectiva etnosociológica es que las lógicas que rigen el conjunto de un mundo social o mesocosmos se dan igualmente en cada uno de los microcosmos que lo componen: observando con atención uno solo, o mejor varios de estos microcosmos, y por poco que se logre identificar las lógicas de acción, los mecanismos sociales, los procesos de reproducción y de transformación, se deberían poder captar al menos algunas de las lógicas sociales del mesocosmos mismo (18).

Con el fundamento en la hipótesis determinada por Bertaux, cuya premisa es que las particularidades poseen significado social, las investigaciones que emplean el método biográfico utilizan una diversidad de términos *espacio biográfico*, *biografías*, *relatos de vida* y *trayectorias biográficas*. Tales dispersiones pueden conducir a equívocos. Para este artículo, el método biográfico puede presentarse bajo diversas formas; por ende, una de sus modalidades es el relato de vida que se concretiza, en este caso, mediante la aplicación de entrevistas cualitativas individuales y semiestructuradas a migrantes sudamericanas travestis/trans. A su vez, existen otras posibilidades de investigar amparándose en este método, como el análisis de los diarios personales y la historia de vida. Los relatos de vida, en la investigación elaborada, consideraron los proyectos migratorios de las travestis/trans desde el momento en que las entrevistadas recuerdan

haber percibido las primeras manifestaciones de sus identidades de género y se proyectan hasta el presente, haciendo un recorrido por los distintos momentos de la migración.

Según Mallimacci y Giménez (2006), el relato de vida permite, desde una perspectiva sociológica, comprender un proceso social a partir de la vida de una persona o de un grupo de personas. De esta manera, los sucesos, vivencias y representaciones que los sujetos construyen a lo largo de su historia personal reflejan aspectos estructurales complejos. Por lo tanto, “cada acto individual es la totalización de un sistema social. El análisis sociológico nos lleva a la hermenéutica de una interacción donde cada relato biográfico nos da la imagen totalizadora de un sistema social y una totalización en marcha” (Mallimacci y Giménez, 2006: 3).

En esta línea de interpretación, Arfuch (2002) sostiene que el espacio biográfico supone una multiplicidad de modalidades que poseen en común contar experiencias de vida que se materializan mediante narrativas. Esto es el resultado de un tercer tiempo, que representa el resultado de articulación de un tiempo cronológico, simbólico y estático; y uno lingüístico, enunciativo que se constituye de manera intersubjetiva construyendo una identidad narrativa que simboliza el entrecruzamiento de lo fáctico con lo ficcional (Ricoeur, 2003). En definitiva, las narrativas biográficas poseen en común que describen aspectos de las biografías atravesadas por la temporalidad y estas permiten aproximarse a procesos sociales de mayor densidad y complejidad. Pero este procedimiento metodológico se ve atravesado por una cláusula de verosimilitud porque el narrador, en un tiempo presente, ficcionaliza aspectos de su vida, los construye (Meccia, 2016).

Respecto de las técnicas de investigación empleadas, se utilizó la entrevista cualitativa con un guión semi estructurado de pautas y la observación participante. En las entrevistas la intención consistió en reconstruir los relatos de vida de las entrevistadas y mediante lo que Kornblit (2004), retomando a Bertaux, denomina *virajes o puntos de inflexión* en las trayectorias biográficas y que desde la perspectiva de la Teoría Fundamentada pueden denominarse incidentes. Las experiencias biográficas de las entrevistadas permitieron identificar indicadores de procesos estructurales sin perder las particularidades de cada una de ellas. La observación participante se empleó como técnica complementaria a las entrevistas e implicó el registro de verbalizaciones y conductas en los espacios en los que se realizó trabajo de campo, a saber: bares, discotecas y talleres dictados por organizaciones políticas en el AMBA.

Para concluir, se afirma que, habiéndose tomado algunos elementos del método biográfico, el marco metodológico principal ha sido diseñado desde los pilares de la Teoría Fundamentada. Luego de describir algunas de las premisas de los métodos de investigación empleados, a continuación, se presenta el modo en que fueron seleccionados los casos.

3. Muestreo teórico y criterios de selección de casos

Dado que entrevistar al conjunto de la población migrante travesti/trans sudamericana es imposible por cuestiones de recursos y tiempos, las investigaciones cuentan con un subconjunto de unidades de análisis que se denominan *muestras*. El muestreo cualitativo contempla una serie de posibilidades relacionadas con la selección de casos, escenarios y acontecimientos. En esta investigación se aplicaron distintos tipos de muestreo, a saber: teórico, “bola de nieve” y por oportunidad o muestreo de voluntarios (Guber, 2004).

En primer lugar, el muestreo teórico, propuesto por la Teoría Fundamentada según Glaser y Strauss, “es el proceso de recolección de datos para la generación de códigos, por el que el analista recoge conjuntamente la teoría y analiza sus datos y decide qué datos recoger en adelante y dónde encontrarlos, para desarrollar su teoría que emerge de los mismos” (1999 [1967]: 45). De esta manera, “el muestreo teórico es acumulativo. Cada acontecimiento que entra en el muestreo se suma al análisis y a la recolección de datos hecha antes y los aumenta” (Strauss y Corbin, 2002: 221).

La muestra, elaborada en el contexto de esta investigación, quedó integrada por 44 casos, de los cuales 3 fueron informantes clave (dos travestis/trans nativas y una mujer cisgénero). El resto se trató de 41 migrantes travestis y mujeres trans sudamericanas con diferentes

características sociodemográficas. Además, se tuvo contacto con alrededor de 130 travestis/trans entre las nativas y las sudamericanas que involucró tanto a las entrevistadas como aquellas con las que se generaron conversaciones en el contexto de la observación participante. Hubo dos entrevistas que fueron interrumpidas y, a falta de registro, no se consideran como parte de la muestra.

Los criterios para seleccionar a las entrevistadas fueron los siguientes: nacionalidad, edad, tiempo de permanencia en el país, nivel instructivo alcanzado y participación o no en organizaciones políticas. En una primera instancia, durante la fase de muestreo abierto, se intentó realizar los primeros contactos con el fin de poder delimitar con mayor precisión los criterios de selección y su justificación. En este momento del muestreo, se incorporaron los casos a los que pudo accederse y se aplicó una codificación abierta, procedimiento que se elaboró en las primeras ocho entrevistas. El principal criterio empleado fue la conveniencia (Patton, 2002), ya que se trató de casos a los que se accedió con relativa facilidad.

Si bien las entrevistadas poseen una variedad de aspectos, algunos de ellos devinieron en criterios en la medida en que fue necesario contemplarlos a la hora de constituirlos como casos. De este modo, se fijó que el tiempo de permanencia en el país debía ser de, al menos, dos semanas. Con respecto a la edad de las entrevistadas, estas debían tener 18 años o más. La nacionalidad estaba contemplada dentro de los países de Sudamérica y, con respecto a la participación política, la muestra contempló casos de participación efectiva. La construcción de la muestra intentó contener la mayor variabilidad posible de acuerdo con los criterios establecidos. Estos criterios fueron constituidos y delimitados porque poseen una mayor incidencia en la conformación de los proyectos migratorios. En definitiva, otorgarle a la muestra la máxima variación (Patton, 2002) implicó reunir casos cuyos criterios tuvieran considerables diferencias, así podrían observarse casos negativos, diversidades, matices y similitudes que permitirían desarrollar una teoría sustantiva con profundidad y densidad teórica.

El intervalo de edad de las entrevistadas fue entre los 19 y los 57 años, con una media de 27 años. Las nacionalidades fueron las siguientes: peruana, colombiana, ecuatoriana, brasileña y paraguaya. El tiempo de permanencia en el AMBA osciló entre dos semanas y veinte años y, en algunos casos, participaban de organizaciones políticas o habían participado alguna vez. Algunas de ellas poseen nivel instructivo secundario completo, otras no han finalizado esa instancia y hay quienes poseen estudios universitarios/terciarios incompletos y casi la totalidad de la muestra ha ejercido la prostitución/trabajo sexual. Respecto de la nacionalidad, se trató de un criterio para seleccionar casos, ya que, la mayoría de las contactadas eran peruanas, con lo cual dar con otras nacionalidades fue clave para poder establecer comparaciones y determinar diferencias según nacionalidad.

Luego, en la instancia de muestreo emparentado con la codificación axial, la mayoría de los criterios para seleccionar a las entrevistadas ya se encontraban delimitados. En virtud de ello, se le otorgó a la muestra variación en cuanto a la nacionalidad, edad y tiempo de permanencia en el país. Por último, algunos casos se eligieron de manera selectiva con el fin de reconstruir experiencias de regreso y dotar de sentido la caracterización de los proyectos migratorios, en calidad de categoría medular. Se trató de casos en los que la entrevista se focalizó en detectar los momentos del proyecto migratorio seleccionando unidades de análisis de manera específica, con el fin de saturar la muestra. De manera particular, se seleccionaron cuatro migrantes travestis/trans sudamericanas que, en conversaciones informales previas, se había podido constatar que poseían trayectorias de retorno temporal y migración en tránsito. Las experiencias de regreso y migración en tránsito forman parte de algunos los proyectos migratorios y han sido aspectos escasamente abordados por los antecedentes.

Como se mencionó, existen distintos tipos de muestreo cualitativo. A diferencia de las muestras cuantitativas, donde la clave radica en la representatividad del universo y el tamaño de la muestra, lo fundamental radica en la calidad muestral que se mide de acuerdo con los aportes que realizan los casos incorporados a la construcción de la teoría emergente. Por eso, los criterios para seleccionar los casos se encuentran guiados e intencionados por propósitos teóricos. En otras palabras, el desarrollo de la muestra, cuyo fin último es la generación de teoría sustantiva, se logra mediante la selección de casos de relevancia teórica conforme a criterios que se estipulan de manera anticipada y durante el desarrollo de la investigación.

De acuerdo con lo que señalan Strauss y Corbin:

(...) el investigador se interesa por generar el mayor número posible de categorías; de ahí que recopile datos en una gran variedad de áreas pertinentes. Una vez que el analista tiene algunas categorías, el muestreo apunta a desarrollarlas, a hacerlas más densas y a saturarlas (2002: 221).

Este procedimiento, que consiste en hacer robustas las categorías de análisis, puede involucrar a personas, acontecimientos y escenarios, y se vincula con el muestreo teórico definido como:

(...) el proceso de recolección de datos para la generación de códigos, por el que el analista recoge conjuntamente la teoría y analiza sus datos, y decide qué datos recoger en adelante y dónde encontrarlos, para desarrollar su teoría que emerge de los mismos (Strauss y Corbin, 2002: 45).

Con respecto a la saturación teórica, los autores afirman: (...) “ninguna información adicional por medio de la cual el sociólogo pueda desarrollar propiedades de la categoría, ya que, al ver instancias similares una y otra vez, el investigador adquiere confianza empírica de que una categoría está saturada” (Strauss y Corbin, 2002: 84). De este modo, la investigación presentada aquí alcanzó la saturación teórica. Tal definición se asienta en que la incorporación de nuevos casos, luego de la entrevista número 40, no proporcionaron información adicional que permitiera robustecer las categorías de análisis. Sin embargo, Strauss y Corbin (2002) no suministran elementos claros para determinar cómo se logra tal proceso. Alcanzar la saturación teórica implica que el investigador se encuentra en la fase de delimitar la teoría emergente. Se considera que, para poder determinar que se ha alcanzado la saturación teórica, es necesario tener en cuenta algunos factores que se enumeran a continuación:

- 1) Si bien suele remitirse a la calidad de los datos producidos, es importante remarcar que la cantidad de casos es un factor para tener en cuenta porque un volumen importante de entrevistas, observaciones y documentos otorgan mayores posibilidades de robustecer las categorías de análisis y desarrollarlas.
- 2) Cuando se observa en el material empírico producido una reiteración significativa de las propiedades de las categorías propuestas. En resumen, no se saturan la totalidad de las categorías de la investigación, solo aquellas que se consideran de mayor relevancia para el análisis.
- 3) Los factores técnicos como el tiempo y los recursos deben tenerse en cuenta. La permanencia prolongada en el campo contribuye a la saturación teórica, pero debe recordarse que las investigaciones cuentan con tiempos y recursos estipulados; por lo tanto, estos factores generan que, en una determinada instancia, el trabajo de campo se dé por finalizado.
- 4) Los investigadores son quienes mejor conocen sus investigaciones. En consecuencia, poseen la sensibilidad teórica necesaria para poder determinar que la incorporación de nuevos casos no generará nuevas categorías y propiedades.

No puede determinarse una receta única para conocer en qué momento se alcanza la saturación teórica. De todas maneras, en esta investigación, luego de la realización de 44 entrevistas y observaciones participantes en múltiples escenarios, los criterios antes mencionados se cumplieron. Las categorías que remitieron al contexto de la transición y motivos de la migración fueron las primeras en saturarse, mientras que, para el modo de migrar, tipo de discriminación, la migración en tránsito y las experiencias de retorno temporal, fue necesario superar las treinta entrevistas para alcanzar ese objetivo.

Por otra parte, un tipo de muestreo que fue menos utilizado en la investigación es el denominado “bola de nieve” y por “casos voluntarios”. En el muestreo por “bola de nieve”, los esfuerzos destinados a la incorporación de nuevos casos a la muestra se depositan en otras personas distintas al investigador. Dicho de otro modo, algún entrevistado, informante o miembro de la comunidad se desempeña como nexo para acceder a otras unidades de análisis, lo que permite el acrecentamiento de la muestra (Blanco y Castro, 2007) mediante un proceso de

encadenamiento. El muestreo por “oportunidad” o por “casos voluntarios” ocurre cuando los entrevistados se presentan con espontaneidad y desean participar en la investigación aportando información. Se trata de alternativas de muestreo que, ante la falta de mejores opciones o debido a las dificultades de acceso, permiten comenzar a familiarizarse y construir los datos necesarios para delinear la incipiente teoría emergente.

Estas modalidades se presentaron de manera escasa. Si bien se consideró que el muestreo por bola de nieve podía ser una herramienta útil para incorporar otros casos, las migrantes travestis y trans que tenían la posibilidad de facilitar contactos con el fin de entrevistar a sus amigas o conocidas rara vez terminaban de concretar ese contacto. Con respecto al muestreo por oportunidad, solo se presentó un caso de manera espontánea en uno de los espacios en los que se realizó el trabajo de campo, ya que la migrante poseía conocimiento por medio de otras personas sobre la labor que se estaba realizando.

4. Procesamiento de la información

El procesamiento de la información suele ser uno de los aspectos al que menos atención se le presta en las investigaciones cualitativas. Según Valles (2005) los investigadores cualitativos se han preocupado por el análisis y la presentación de sus resultados, pero han subestimado la manera de procesar e intervenir los datos. En consecuencia, la escasa atención sobre estos procedimientos puede afectar la calidad y rigor de la investigación en su conjunto. Si bien la utilización de un *software* no es necesario para realizar estos procedimientos, ya que el procesamiento puede realizarse de manera artesanal o emplear algún tipo de programa del paquete Office como Excel o Word, es pertinente establecer una aclaración: el procesamiento de la información es un procedimiento metodológico necesario para reducir, sistematizar y analizar los datos. Por ende, puede excluirse, para su elaboración, el uso de herramientas informáticas, como los *software*, pero el procedimiento metodológico como tal representa una condición en el proceso de investigación, con lo cual debe estar presente. Este proceso supone que el investigador transforma la información en datos, les inscribe un sentido, una marca teórica que permite hacerlos abordables de acuerdo con las metas de la investigación. Por eso, estos procedimientos deben emplearse con rigurosidad teniendo en cuenta que las técnicas empleadas proporcionan una variedad y una complejidad en la información con la que luego se estructura el análisis. Cuando se carece de procesamiento, o se asume con laxitud, se presupone que los datos se explican por sí mismos asumiendo la realidad como algo dado.

En algunos casos, estos procedimientos se asumen con debilidad y la fase de procesamiento de la información suele ser inexistente, lo que genera modos de análisis que suponen una suerte de “codificación directa”, es decir, los fragmentos de entrevistas o las notas de campo se insertan en el texto para luego ser analizados y el proceso de reducción mediante una mirada teórica se encuentra ausente. Señalar esto es pertinente porque, a la tradición cualitativa, se le ha criticado su falta de sistematicidad, cuando en verdad esta falencia tiene que ver con el modo en que se aplican estos procedimientos. La investigación cualitativa es compleja porque debe constituir descripciones densas a partir de la perspectiva del actor y, a la vez, sostener interpretaciones ajustadas a criterios de validez y de confiabilidad, con la traducción correspondiente.

En esta investigación, se empleó, en primer lugar, el programa Excel 2016, en el que se volcaron las categorías de análisis iniciales y los fragmentos de entrevistas cuyo contenido expresaba, en sus distintas variaciones, su significado. Las primeras ocho entrevistas se procesaron desde esta modalidad, pero el crecimiento del material empírico implicó la necesidad de incorporar un *software* que permitiera facilitar el análisis y proveyera los recursos necesarios para lograr la mayor sistematicidad y rigurosidad. De esta manera, se incorporó el *software* Atlas ti versión 8.3. Es pertinente señalar que, en esta instancia, la investigación realizó análisis preliminares en la medida en que el grillado elaborado en Excel supuso una instancia de codificación y de análisis de las principales categorías a partir de la identificación de palabras y oraciones que se observaron en las entrevistas y que resultaron relevantes desde el punto de vista teórico.

Como las investigaciones cualitativas suelen carecer del empleo de *software* para realizar un procesamiento de los datos producidos, los recursos suelen quedar limitados a algún programa del paquete Office. En el caso de emplear métodos cuantitativos, el quehacer artesanal se vuelve imposible debido a la cantidad del material empírico con el que se trabaja, que suele tornarse inabarcable. Los paquetes de procesamientos de datos cualitativos permiten integrar diferentes tipos de informaciones y llevar a cabo procedimientos de análisis como, por ejemplo, la coocurrencia de códigos, que ilustra la vinculación y codependencia de las categorías de análisis, lo que permite también verificar la consistencia de la teoría sustantiva. Es decir, estos representan vectores para otorgarle al análisis calidad y sistematicidad.

En el caso de esta investigación, como se mencionó, se empleó la versión 8.3 del Atlas ti y los principales recursos utilizados fueron la asignación de códigos², las familias de códigos, las notas, los memos, los diagramas, el diseño de tabla de coocurrencia de códigos y de código-documento. Los memos³ son dispositivos analíticos en los que se constituyen interpretaciones sobre el material empírico. Los códigos representan etiquetas que le dan nombre a las categorías y reducen los datos, las notas son elaboraciones escritas de primera mano que permiten realizar interpretaciones iniciales y las familias de códigos permiten llevar a cabo un ordenamiento y una jerarquización de las categorías con la meta de otorgarle estructura a la teoría emergente. Respecto de los memos se construyeron descriptivos, analíticos, teóricos y metodológicos, lo que permitió integrar los procedimientos de análisis sobre la información provista por las entrevistas y la observación participante.

Luego, las variaciones construidas de acuerdo con las propiedades y dimensiones de las respectivas categorías se volcaron en una matriz elaborada en el programa Word 2016 y esta se presenta a modo de tablas que sintetizan esa información. El *software* Atlas ti resulta útil cuando se trabaja con la Teoría Fundamentada porque las funciones que posee se articulan de manera congruente con el método. Una de sus principales limitaciones radica en que los documentos cargados no pueden ser editables y, por lo tanto, los errores de tipeo no pueden corregirse, factor que, si bien está ligado a la operatividad, es para tener en cuenta.

Más allá de la operatividad y de los recursos técnicos que utilizó la investigación, el análisis de los datos cualitativos requiere de un minucioso proceso en el que el analista debe sumergirse en las complejidades y en la trama de significados que le aportan las entrevistadas. Para ello, es necesario describir los procedimientos de análisis empleados.

5. Procedimientos de análisis

Dado que en las investigaciones cualitativas suele haber una superposición de fases, la escisión entre procesamiento y análisis se relacionan con la claridad en la presentación, puesto que "... en la investigación cualitativa no es posible independizar la etapa del procesamiento de la del análisis: el desarrollo de la primera conlleva la segunda" (Cohen y Gómez Rojas, 2014: 13). Dicho de otro modo, constituir una muestra, seleccionar los casos que proveen de información relevante y establecer procedimientos de codificación y análisis mediante la comparación del material empírico representan instancias involucradas en un mismo proceso.

Una de las herramientas propuestas por la Teoría Fundamentada, y que conduce al análisis de la información, consiste en distintos tipos de codificación que se denominan abierta, axial y selectiva, y dan lugar a la construcción de categorías, subcategorías, propiedades y dimensiones. Para lograr esta meta es necesario emplear el método de análisis denominado *comparación constante* en el que deben detectarse incidentes y compararlos según grupos o personas mediante las variaciones máximas y mínimas para que las categorías se construyan de manera robusta (Seid, 2017). Este procedimiento consiste, en primer lugar, en reducir las diferencias del

² Si bien puede construirse un libro de códigos que remitan a las categorías y a sus variaciones, en este caso, se asimilaron los códigos a las categorías y sus variaciones fueron construidas a partir de una matriz en Word.

³ Cabe destacar que Strauss y Corbin (2002) se refieren a memorandos, pero se prefiere utilizar *memos* debido a que el Atlas ti le asigna este nombre a esa herramienta de análisis inicial y complementario.

material empírico para diseñar las categorías y luego se acentuarlas para encontrar variaciones con respecto a las propiedades y dimensiones. Esta instancia supone poder delimitar las principales categorías relevantes para el análisis y luego fortalecerlas mediante la incorporación de propiedades que muestren una importante variación. Fue mediante este procedimiento que la investigación apostó por construir categorías robustas y densas mostrando la mayor variabilidad y riqueza posible. Desde el punto de vista de los relatos de vida, los incidentes implicaron virajes biográficos (Meccia, 2016), es decir, puntos de ruptura que impulsaron nuevos rumbos en las vidas de las entrevistadas.

El microanálisis inicial, o análisis “línea por línea” incluye la codificación abierta y axial y permite en un primer momento descubrir categorías y comenzar a develar las relaciones entre los conceptos. El proceso analítico que permite identificar conceptos y descubrir en los datos sus propiedades y dimensiones se conoce como *codificación abierta*. Partiendo del registro de campo, se comienza a realizar comparaciones que buscan identificar similitudes en el contenido. Estos incidentes que pueden ser objetos, acciones, acontecimientos, ideas o interacciones y que comparten un común denominador de acuerdo con la perspectiva teórica del investigador son identificados con un código común. De esta forma, son ubicados en categorías existentes o emergentes al momento de su descubrimiento y la comparación constante permite generar, de forma dinámica, propiedades derivadas de las categorías. Los nombres de las categorías pueden ser establecidos por el investigador de acuerdo con las evocaciones que producen desde un contexto teórico, o bien, nominarse con códigos *in vivo*, es decir, utilizando palabras propias del lenguaje de las personas entrevistadas. En este caso, uno de los códigos *in vivo* de mayor importancia fue *madrinas*, que remite a travestis/trans de edad avanzada que facilitan la asimilación en la sociedad receptora.

La codificación axial implica que el investigador ya posee las principales categorías de análisis establecidas y el procedimiento de análisis inicial se finalizó, con lo cual las categorías comienzan a integrarse, refinarse y perfeccionarse. Esta codificación permite ligar las categorías a sus subcategorías para alcanzar explicaciones más minuciosas sobre los fenómenos. De esta forma, *axial* hace referencia al procedimiento que, tomando como eje una categoría, enlaza sus propiedades y dimensiones, y le da, de esta forma, profundidad y estructuración. Se construye, en este punto, teoría. La finalidad de la codificación axial es construir, de modo sistemático, categorías densas relacionándolas entre sí, ejercicio que resulta, sin dudas, complejo. La comparación en este punto se centra en observar las categorías frente a los datos a los fines de refinarlas y fortalecerlas.

De este modo, algunas de las categorías principales elaboradas por la investigación, fueron los motivos de la migración, el modo de migrar, la edad al asumir la identidad de género, el tipo de representaciones sociales de la República Argentina como destino migratorio, modo de migrar, tipo de discriminación, tipo de vínculo con las familias a la distancia, tipo de vínculos entre las travestis/trans y tipo de experiencias de retorno temporal. Algunas de ellas se presentan en el anexo del artículo a modo de ejemplo. También, algunas categorías secundarias, pero asociadas a las principales fueron las siguientes: modo de ejercer la prostitución/trabajo sexual, tipo de prácticas sexuales, tipo de vínculos afectivos, uso de drogas, tipo de intervenciones corporales, readecuación genital, obtención de DNI y tipo de vínculo entre travestis y trans, entre otras.

Por otra parte, diversos autores, algunos de ellos ya mencionados, que se han detenido en el análisis de los procedimientos de codificación propuestos por la teoría fundamentada han determinado que este proceso no se encuentra atravesado por etapas temporales que pueden construirse de acuerdo con una linealidad concreta. En otras palabras, puede ser necesario volver a codificar de manera abierta en la medida en que la generación de nuevas entrevistas elucide categorías relevantes que no se habían observado o no se habían tenido en cuenta al principio de la investigación (Cohen y Seid, 2019). Este aspecto es pertinente, puesto que en esta investigación fue necesario volver atrás y recodificar entrevistas al emerger categorías de análisis no contempladas al inicio.

De este modo, en un principio, se ignoraron las experiencias de retorno temporal a las sociedades emisoras de las migrantes travestis/trans sudamericanas que ocurrían luego de haber consolidado el proyecto migratorio en el AMBA. En algunas entrevistas, la cuestión del regreso había aparecido de manera solapada, pero no se profundizaba en su significado. El retorno

migratorio representa uno de los aspectos más importantes en los procesos de desplazamiento poblacional y, en el caso de las migraciones de las travestis y trans sudamericanas, ha sido prácticamente inexplorado. Por eso, los antecedentes consultados no ahondaban en las experiencias de regreso. Solo caracterizaban el abandono de los países de origen y los conflictos que poseían con sus familias de origen en relación con la identidad de género en las sociedades emisoras. Sin embargo, la construcción identitaria y el proyecto migratorio también involucran el retorno porque, en esa instancia, las unidades de análisis atraviesan una prueba ante sus familias de origen mostrando que el proceso de expresar socialmente sus identidades de género es genuino y esto impacta en los vínculos que mantienen con ellas que, luego de haber transitado la migración, tienden a ser de mayor aceptación. Entonces, fue necesario volver a la instancia anterior, codificar las entrevistas en las que el retorno migratorio había emergido y luego, en la realización de nuevas entrevistas, incorporar en la guía de pautas interrogantes que hicieran foco en esa categoría. De ahí el importante volumen de las entrevistas realizadas y los obstáculos para alcanzar la saturación teórica.

Por otra parte, la codificación selectiva implica el descubrimiento de una o de dos categorías que poseen poder analítico. Para este artículo, *proyecto migratorio* e *identidad de género* son las categorías medulares, ya que atraviesan al conjunto de las categorías construidas. Poseen poder analítico y las definiciones del marco teórico dan cuenta de su densidad teórica. Esto no quiere decir que otras categorías de análisis no sean centrales. De hecho, una tercera categoría medular que se consideró en instancias finales del proceso de investigación fue la desviación (Becker, 2009), debido a que esta condensa y anida la identidad de género y la condición migratoria de las entrevistadas tanto en las sociedades emisoras como en la receptora. La desviación es una categoría que atraviesa las trayectorias biográficas de las entrevistadas desde que comienzan a expresar las primeras manifestaciones de sus identidades de género y continúan durante la totalidad del proceso migratorio.

Por último, las redes analíticas o diagramas, también generados con el software Atlas ti permitieron graficar las relaciones entre categorías. Estos recursos representan dispositivos visuales analíticos que permiten aproximarse a la creación de teoría sustantiva. Las redes vinculan los códigos que hacen referencia a las categorías y permiten integrar el análisis mediante la conexión jerárquica de los códigos construidos. En términos de Strauss y Corbin: “al final, es importante tener una descripción clara y gráfica de la teoría, que sintetice los principales conceptos y sus conexiones” (2002: 258).

6. Conclusiones

El artículo consistió en describir algunos procedimientos metodológicos en el contexto de una investigación empírica sobre los proyectos migratorios de las travestis y mujeres trans que en la actualidad residen en el AMBA. De este modo, se expuso la manera en que fue aplicada la teoría fundamentada de manera auxiliar algunos aspectos del relato de vida y la etnografía que se efectivizaron mediante técnicas como la entrevista semi-estructurada y la observación participante. Se apuesta a contribuir a la metodología de la investigación social tanto para ejemplificar la aplicación de la Teoría Fundamentada como para favorecer diferentes estrategias pedagógicas didácticas que apunten al empleo de la metodología cualitativa.

Con respecto al muestreo teórico se describieron los diferentes criterios para seleccionar los casos conforme avanzaba la investigación y se formalizaron algunos criterios para alcanzar la saturación teórica. El muestro teórico permite seleccionar casos de acuerdo a criterios que se encuentran ligados a los análisis de la información. En relación con el procesamiento y análisis de la información, se halló que los procesos de codificación no suponen etapas rígidas; por el contrario, el investigador puede revisar los procedimientos de codificación y regresar a instancias anteriores. Por último, puede concluirse que el software Atlas ti versión 8.3 proporciona herramientas de análisis útiles cuando se emplea la teoría fundamentada, ya que sus funciones se adaptan al método y facilitan el análisis de la información.

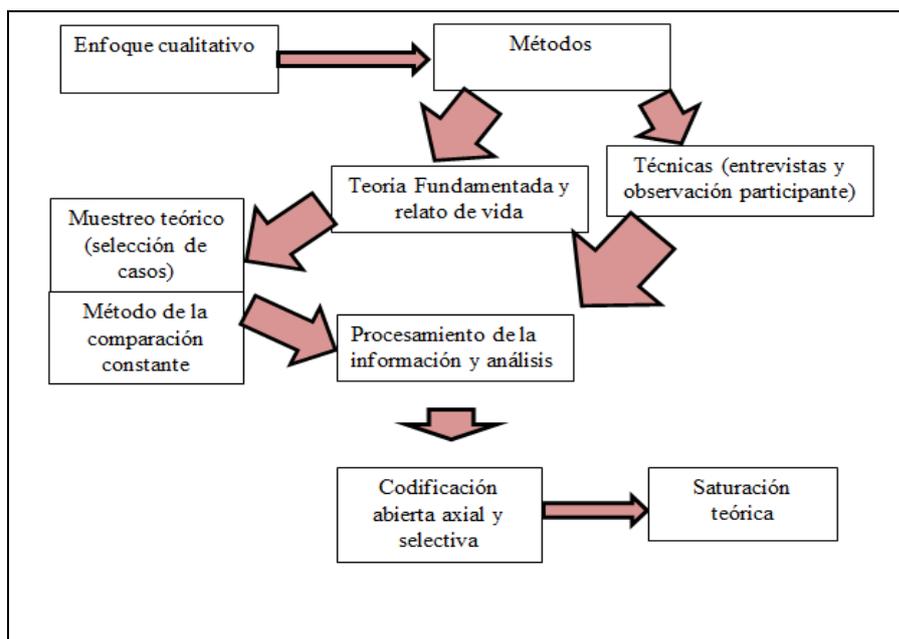
7. Bibliografía

- ARFUCH, L. (2002). *El espacio biográfico: dilemas de la subjetividad contemporánea*. México: Fondo de Cultura Económica.
- BECKER, H. (2009). *Outsiders. Hacia una Sociología de la desviación*. Buenos Aires: Siglo XXI.
- BERKINS, L. (2007). *Informe Nacional sobre la situación de las travestis, transexuales y transgéneros. Cumbia, copeteo y lágrimas*. Buenos Aires: Asociación de Lucha por la Identidad Travesti-Transexual.
- BERTAUX, D. (2005). *Los relatos de vida: perspectiva etnosociológica*. Barcelona: Bellaterra.
- BLANCO, M. C. M. C. y Castro, A. B. S. (2007). El muestreo en la investigación cualitativa. *Nure investigación*, 27(4). Recuperado de: <https://www.nureinvestigacion.es/OJS/index.php/nure/article/view/340>.
- BOY, M. (2017). Cuerpos e identidades extranjerizados: vecinos/as y travestis en disputa El caso de la zona roja de Palermo, 1996-200. En *Fronteras en la Ciudad (re)producción de desigualdades y conflictos urbanos*. Buenos Aires: Teseo.
- CHARMAZ, K. (2006) *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis*. Los Angeles, CA: SAGE Publications.
- COHEN, N. y GÓMEZ ROJAS, G. (2014). Esa cosa llamada datos. *Revista latinoamericana de metodología de la investigación social*, 4(8), 10-18. Recuperado de: <http://relmis.com.ar/ojs/index.php/relmis/article/view/132/198>.
- _____ (2019). *Metodología de la investigación, ¿para qué? La producción de datos y los diseños*. Buenos Aires: Teseo, Clacso.
- COHEN, N. y SEID, G. (2019) Producción y datos de análisis cualitativos. En Cohen, N. y Gómez Rojas, G. (2019) *Metodología de la investigación, ¿para qué? La producción de datos y los diseños*. (pp. 203-228). Buenos Aires: Teseo, Clacso.
- DE SENA, A., LISDERO, P. y SCRIBANO, A. (2015). *Caminos cualitativos: aportes para la investigación en Ciencias Sociales*. Buenos Aires: CICCUS.
- FERNÁNDEZ, J. (2004): *Cuerpos desobedientes. Travestismo e identidad de género*. Buenos Aires: Edhasa.
- GARCÍA, A. y OÑATE, S. (2008). Transexuales ecuatorianas: el viaje y el cuerpo, en Herrera, G. y Ramírez, J. (eds.). *América Latina migrante: Estado, familia, identidades* (pp. 343-360). Quito: FLACSO – Ministerio de Cultura.
- GLASER, B. G.Y HOLTON, J. (2004). *Remodeling Grounded Theory*. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, 5(2), Art. 4, <http://nbn-resolving.de/urn:nbn:de:0114-fqs040245>.
- GLASER, B. G., & STRAUSS, A. L. (1999 [1967]). *Discovery of grounded theory: Strategies for qualitative research*. Routledge.
- GUBER, R. (2004). *El salvaje metropolitano: reconstrucción del conocimiento social en el trabajo de campo*. Buenos Aires: Paidós.
- KORNBLIT, A. L. (Coord.) (2004). *Metodologías cualitativas en ciencias sociales*. Buenos Aires: Editorial Bilblos.
- LÓPEZ FERNÁNDEZ, V. (2018). Diásporas trans, fronteras corporeizadas y tránsito (s) migratorios en México. *Cuicuilco. Revista de ciencias antropológicas*, 25(71), 9-34.
- MALLIMACI, F. y GIMÉNEZ, V. (2006). Historias de vida y método biográfico. En *Estrategias de investigación cualitativa* por I. Vasilachis (Coord.) (pp. 175-212). Madrid: Gedisa.
- MANALASAN IV, M. F. (2006). Queer intersections: Sexuality and gender in migration studies. *International migration review*, 40(1), 224-249.
- MARRADI, A. (2002). Método como arte. *Revista de Sociología*, 67, 107 -127.

- MECCIA, E. (2016). *El tiempo no para. Los últimos homosexuales cuentan la historia*. Santa Fe, Argentina: Eudeba y Ediciones UNL.
- PATTON, M. Q. (2002). *Qualitative Evaluation and Research Methods* (3a ed.). Londres: Sage.
- PELÚCIO, L. (2009). "Sin papeles" pero con glamur. Migración de travestis brasileñas a España (Reflexiones iniciales). *VIBRANT-Vibrant Virtual Brazilian Anthropology*, 6(1), 170-197. Recuperado de: <https://www.redalyc.org/articulo.oa?id=4069/406941907008>.
- RICOEUR, P. (2003). *La memoria, la historia, el olvido*. Madrid: Trott.
- SCRIBANO, A. O. (2008). *El proceso de investigación social cualitativo*. Buenos Aires: Prometeo.
- SEID, G. (2017). La pluralidad de procedimientos para alcanzar validez en las investigaciones cualitativas. *Revista Latinoamericana de Metodología de la Investigación Social: ReLMIS*, (12), 41-55. Recuperado de: <http://www.relmis.com.ar/ojs/index.php/relmis/article/view/165>.
- STANG, F. (2018). *Fronteras, sexualidades procesos de subjetivación. Migrantes LGTBQ colombianos y peruanos en Santiago de Chile*. (Tesis de doctorado no publicada). Córdoba: Universidad Nacional de Córdoba.
- STRAUSS, A. y CORBIN, J. (2002). *Bases de la investigación cualitativa: técnicas y procedimientos para desarrollar la teoría fundamentada*. Colombia: Universidad de Antioquia.
- VALLES, M. (2005). El reto de la calidad en la investigación social cualitativa: de la retórica a los planteamientos de fondo y las propuestas técnicas. *Revista Española de Investigaciones Sociológicas (REIS)*, 110(1), 91-114.
- VARTABEDIAN, J. (2012). *Geografía travesti: cuerpos, sexualidad y migraciones de travestis brasileñas (Río de Janeiro-Barcelona)*. (Tesis de doctorado). Barcelona: Universitat de Barcelona. Recuperado de: <http://hdl.handle.net/2445/35232>.
- ZARCO ORTIZ, E., & Chacón Reynosa, K. (2020). Dispositivos de seguridad y sexualidad en la frontera sur de México: biopolíticas en mujeres transgénero centroamericanas. *Tabula rasa*, (33), 137 - 163. doi:10.25058/20112742.n33.06.

Anexos:

Diagrama N° 1: Diferentes procedimientos aplicados de acuerdo con la Teoría Fundamentalada



Fuente: Elaboración propia.

Ejemplos de categorías construidas en la investigación mediante la Teoría fundamentada

N°1: Categoría modo de migrar según subcategorías, propiedades y dimensiones

Categoría	Subcategorías	Propiedades	Dimensiones
Modo de migrar	<i>Cadenas migratorias horizontales</i>	Organizaciones políticas	Garantías mínimas de supervivencia (alimentos, indumentaria y vivienda)
		Amistades	
		Familiares	Asumir costos vinculados al viaje Respeto y protección
	<i>Cadenas migratorias verticales</i>	Parejas (<i>maridos, chongos y garrones</i>)	Garantías mínimas de supervivencia (alimentos, indumentaria y vivienda) Asumir costos vinculados al viaje Tareas domésticas reproductivas
		Madrinas	Socialización diferenciada y asimilación Protección ligada a la prostitución/trabajo sexual
			Asumir costos vinculados con el viaje Consejos sobre cómo llevar a cabo actuaciones de género Conductas abusivas Dominación Respeto, lealtad y liderazgo

Fuente: Elaboración propia.

N° 2: Categoría tipo de vínculos con las familias a la distancia según propiedades y dimensiones

Categorías	Subcategorías	Propiedades	Dimensiones
<i>Tipo de vínculos entre travestis/trans</i>	<i>Cooperación/solidaridad</i>	Ayuda económica, compartir informaciones, protección, exigir el cumplimiento de derechos, defenderse ante situaciones de violencia	Intensidades (fuerte, moderada y leve)
	<i>Indiferencia</i>	Falta de comunicación, decidia y apatía	
	<i>Competencia/hostilidad</i>	Agresiones verbales y físicas, acciones destinadas al control del espacio, celos vinculados a lo afectivo	Frecuencia (alta, media y baja)

Fuente: Elaboración propia.

Autor.

Ramiro Nicolás Perez Ripossio

Universidad de Buenos Aires (UBA); Instituto de investigaciones Gino Germani (IIGG); Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina.

Doctor en Ciencias Sociales (UBA). Docente de metodología de la investigación en la carrera de Sociología de la Universidad de Buenos Aires (UBA).

E-mail: ramiro7242@hotmail.com

Citado.

PEREZ RIPOSSIO, Ramiro Nicolás (2022). Aplicación de la Teoría Fundamentada en una investigación sobre travestis y mujeres trans sudamericanas que residen en el AMBA. *Revista Latinoamericana de Metodología de la Investigación Social - ReLMIS*, N°23, Año 12, pp. 61-75.

Plazos.

Recibido: 16/09/2020. Aceptado: 02/04/2021.



Estratégias de divulgação científica nas ciências sociais: escolhas teórico-metodológicas para a criação do website Arquivo Rock Teresina

Scientific divulgation strategies in Social Sciences: theoretical-methodological choices for the creation of the Arquivo Rock Teresina website

Thiago Meneses Alves e Irene Serafino

Resumo

O artigo objetiva discutir estratégias de divulgação científica em Ciências Sociais a partir do processo que culminou no website Arquivo Rock Teresina. Este surgiu da necessidade de disponibilizar de forma acessível os resultados de uma tese de doutorado em Sociologia sobre a produção de rock autoral independente da capital do Piauí. São explicadas a formatação do estudo de caso, os critérios e técnicas de recolha de dados, o planejamento e formulação de um instrumento de divulgação científica e a estrutura e o conteúdo do site. Levando em consideração o debate sobre a democratização do conhecimento, argumenta-se que a construção de instrumentos de divulgação científica é fundamental para a disseminação dos resultados das pesquisas para audiências mais amplas, respondendo a objetivos educativos, cívicos e de mobilização popular.

Palavras-Chave: Divulgação científica; democratização do conhecimento; música independente brasileira; arquivo em website; investigação social.

Abstract

The article aims to discuss strategies of scientific divulgation in the Social Sciences from the process that has culminated in the Arquivo Rock Teresina website. This site was created to provide in an accessible way the results of a sociology PhD thesis about authorial independent rock from Teresina, capital of Piauí State. Here are explained the construction of the case study, criteria and techniques to collect data, planning and formatting of an instrument for scientific divulgation and the structure and content of the website. Taking into consideration the debate about knowledge democratization, it is argued that the construction of instruments of scientific divulgation is essential for the dissemination of the researcher's results for bigger audiences, answering also educational, civil and popular mobilization goals.

Keywords: Scientific divulgation; knowledge democratization; independent Brazilian Music; Website archive; social research.

1. Introdução

Este artigo surge das reflexões teóricas e metodológicas sobre uma experiência concreta de construção de um instrumento de divulgação científica de uma tese de doutorado em Sociologia sobre a produção de rock autoral de Teresina, capital do Piauí, nos primeiros anos do século XXI. O contexto analisado na pesquisa está situado na vasta e, para muitos, anônima, produção musical gerada fora das grandes estruturas da indústria cultural e longe do tradicional eixo Rio-São Paulo.

Há uma bibliografia acadêmica e jornalística consistente sobre essa produção, com boa representatividade, inclusive, do rock gestado longe dos principais centros econômicos, políticos e midiáticos brasileiros (Benevides, 2019: 161). Contudo, o mesmo não pode ser afirmado sobre iniciativas de divulgação científica dos resultados de tais investigações para um público não acadêmico.

Na medida em que foram se desenrolando as etapas de construção de um instrumento de diálogo com um público fora do alcance restrito da academia, vários pontos de reflexão sobre os modos de disseminar esse tipo de informação foram surgindo. Estes podem ser resumidos nas seguintes questões: 1) de que forma pode ser agilizada uma divulgação eficiente nas Ciências Sociais e Humanas, aumentando as possibilidades de diálogo entre acadêmicos e o público em geral? 2) como afinar as estratégias de divulgação com o potencial oferecido pelo ambiente digital? 3) como, por outro lado, lidar com as dificuldades de um contexto contemporâneo marcado pelo excesso informacional? 4) de que formas estes questionamentos são importantes para as pesquisas sobre os fenômenos musicais em territórios longe dos grandes centros políticos, econômicos e midiáticos do Brasil?

Após a finalização do website Arquivo Rock Teresina (Alves & Serafino, 2018), no ar em caráter experimental, começaram a surgir respostas preliminares para estas questões. Estas sugerem estratégias para a construção de instrumentos que contribuam para o incremento de uma “alfabetização científica” (Bueno, 2010: 1). Em termos mais específicos, para a familiarização, para uma faixa maior de público, sobre as pesquisas veiculadas no meio científico, com a consequente abertura de possibilidades de inserção deste mesmo público nos debates em torno destas pautas.

No que diz respeito à eficiência na divulgação de material científico no âmbito das Ciências Sociais (ponto um), verificou-se a necessidade de formatar um conteúdo de fácil disseminação, substancialmente diverso do produto acadêmico tradicional. Um procedimento fundamental neste sentido foi a escolha por um estilo de escrita simples, que evite a prolixidade, endossado por alguns intelectuais renomados que se dedicaram à temática (Becker, 2015; Eco, 2007; Schopenhauer, 2005).

A escolha pelo formato website foi a estratégia geral para inserir de forma mais efetiva os resultados da pesquisa no ambiente digital (ponto dois). O auxílio das redes sociais da internet como *Facebook* para impulsionar o alcance do conteúdo foi outro passo importante. Além disso, buscou-se uma divulgação mais personalizada a partir da apresentação do material para um conjunto de agentes e instâncias com potencial interesse pela temática. Espera-se, assim, a replicação espontânea do material por especialistas que possuem elevado capital social no âmbito da música independente, seja na esfera acadêmica, seja na jornalística.

Além da formatação de um texto mais direto do que aquele convencionalmente verificado no ambiente acadêmico, a utilização de *hiperlinks* e a inserção de fotografias e material audiovisual foram as opções adotadas para um formato que tivesse apelo junto a uma audiência mais diversificada (ponto três). Buscou-se esta formatação tendo em vista o ambiente veloz e sobrecarregado, do ponto de vista do volume de informações, que caracteriza a internet. Neste ambiente, o texto longo e linear acaba por não surtir o mesmo efeito do hipertexto enxuto (Ferrari, 2010: 44), – que possibilita uma leitura mais personalizada e potencializa a interatividade com esse público.

Por último, concluiu-se também que um arquivo nos moldes do que é aqui apresentado contribui para o debate sobre novos modos de abordagem para segmentos musicais independentes (ponto quatro). De modo mais específico, sobre a importância de, nas análises,

serem incluídas e posteriormente divulgadas informações mais substanciais sobre a história, assim como o volume de obras produzidas nestes territórios de produção musical situados longe de grandes centros econômicos, políticos e midiáticos brasileiros. Como discutido adiante, ainda que exista já um bom número de trabalhos sobre a independência musical no Brasil (Giorgis, 2017), ainda são poucas as pesquisas que priorizam as obras produzidas nestes territórios, organizando estas informações em arquivos e disponibilizando-as para um público mais amplo (Alves, 2018; Quintela & Guerra, 2017).

2. Divulgação científica e a democratização do conhecimento

Em linhas gerais, a divulgação científica consiste na utilização de certos recursos, técnicas, processos e produtos para disseminar informação de natureza científica para o público leigo, não familiarizado com os ambientes de discussão acadêmica (Bueno, 2010: 2). A divulgação científica é diferente quer da comunicação científica (conteúdo especializado que visa o público acadêmico), quer da difusão científica (inclui todos os modos de difundir este tipo de informação, especializados ou não) (Albagli, 1996: 397).

O primeiro grande objetivo da divulgação científica é de teor educacional: aumentar o acesso ao conhecimento do público não acadêmico sobre os procedimentos e lógicas que regem a ciência. Este tipo de procedimento visa ainda estimular a curiosidade da população sobre o que é produzido em âmbito científico. O segundo grande objetivo é de teor cívico. Visa esclarecer sobre os impactos da prática científica nas sociedades em que se desenvolvem, formando uma opinião pública consciente, que possa ter voz ativa nas tomadas de decisões. O terceiro grande objetivo visa a mobilização popular. Trata-se de utilizar a ciência como instrumento para a construção de indicadores que permitam que estas tomadas de decisões coletivas estejam mais afinadas com as necessidades dos territórios onde as pesquisas se desenvolvem (Albagli, 1996: 397).

Embora divulgação e comunicação científica tratem de disseminar o conteúdo produzido no meio científico, há estratégias específicas para cada uma destas modalidades. Estas estratégias devem levar em consideração, em primeiro lugar, as funções de cada modalidade de disseminação: informação mais especializada entre os pares (comunicação científica) e democratização do acesso ao conhecimento científico a partir da formatação de um conteúdo mais acessível (divulgação científica). Além deste pressuposto geral, é necessário levar em conta, também, o perfil do público-alvo, o nível do discurso a ser utilizado e a natureza dos canais e veículos de comunicação (Bueno, 2010: 2).

Na comunicação científica o público-alvo é formado por especialistas que compreendem a linguagem técnica utilizada. Já na divulgação científica é formado por indivíduos que, *a priori*, não dominam tal linguagem. É neste sentido que comunicação e divulgação científica apresentam níveis diferenciados de discurso. Enquanto no primeiro caso os termos técnicos, assim como uma menor preocupação com o didatismo, são comuns, no segundo caso, há uma preocupação com a promoção de acessibilidade no tocante ao entendimento vasto e irrestrito do conteúdo. A tipologia de canais de promoção do conteúdo também é diversa nas duas modalidades. Enquanto na comunicação científica estes são em número restrito (eventos e periódicos técnico-científicos), a divulgação científica pode e deve ser feita num leque bem mais vasto – veículos de imprensa diversos (rádio, televisão, jornais, revistas, Internet), livros didáticos, palestras, jogos, entre outros.

A adoção de posturas que tenham em conta pressupostos como a disponibilização mais democrática dos resultados das pesquisas científicas contribui para uma missão ética fundamental: o retorno da ciência à sociedade.

Antes da apresentação da estrutura e conteúdo do Arquivo Rock Teresina, são explicados os modos como foi delimitado o estudo de caso do rock independente autoral da capital do Piauí e os critérios utilizados para a construção do arquivo que compõe o site.

3. Rock independente de Teresina: a delimitação de um estudo sociológico

Em linhas gerais, a produção musical independente consiste no empreendimento artístico que nasce fora do raio de atuação das estruturas ligadas à grande indústria cultural. Esta relação tem sido historicamente problematizada no âmbito da indústria fonográfica¹ – ou seja, entre as grandes gravadoras, geralmente partes de conglomerados multinacionais, e pequenos e médios selos mais especializados (Shuker, 1998: 144-145).

Estes selos, assim como as rádios vinculadas à população negra estadunidense, exerceram um importante papel no nascimento e consolidação do rock enquanto um dos principais gêneros de música jovem no século XX (Friedlander, 2002: 39). Apesar das polarizações que costumam ser feitas na caracterização destas instâncias, a complementaridade entre mundos musicais *mainstream* (das grandes indústrias culturais) e independentes ocorrem com alguma frequência. Apesar de complexos, estes arranjos entre pequenos selos especializados e grandes gravadoras não necessariamente configuram uma destruição dos pressupostos estéticos e das lógicas de trabalho das pequenas empresas por parte da grande estrutura capitalista (Hesmondhalgh, 1999: 34).² O independente, de vários modos, também impregna com suas lógicas o *mainstream* musical. E vice-versa. A história do rock é prova viva neste sentido.

Nos últimos anos, há um incremento razoável das pesquisas sobre a música independente brasileira, com um número razoável de trabalhos a nível de pós-graduação³ (Giorgis, 2017: 22-27). Em contraste às iniciativas que fornecem indicadores para um diagnóstico mais geral deste heterogêneo e robusto segmento musical do país, o caráter mais especializado dos recortes de pesquisa é a tônica. Isso significa a ênfase em estudos de caso que podem focar em um território, artista, espaço de fruição, gênero musical, etc. A pesquisa que originou o Arquivo Rock Teresina inclui-se neste conjunto de investigações mais especializadas, cujo foco foi a produção de rock autoral da cidade nestes primeiros anos do século XXI.

No caso desta pesquisa, a delimitação de um estudo de caso em termos territoriais corresponde ao perímetro urbano de Teresina, que conta com uma área de 1.391km² e cerca de 814 mil habitantes (IBGE, 2010). Em termos temporais, o recorte abrange os primeiros anos do século XXI (2001-2017) uma vez que é neste período que ocorrem mudanças substanciais nos modos de produzir, circular e consumir música, ligadas principalmente à emergência do ambiente digital.

Entre outros impactos, o acesso facilitado às tecnologias da informação e comunicação foi um dos principais fenômenos responsáveis pela desestruturação parcial do modelo hegemônico de indústria fonográfica que vigorou no decorrer do século XX. A maior parte das receitas deste modelo vinha principalmente da venda de fonogramas em suportes materiais (Herschmann, 2010: 62). Passados alguns anos após a euforia inicial, que previa a diminuição do poder das grandes estruturas capitalistas na intermediação entre artistas e público, o quadro atual aponta para uma reorganização do grande *business* fonográfico. Uma reorganização menos libertária do que supunham os prognósticos mais otimistas.

Assim, o novo panorama é caracterizado por uma perda de protagonismo da estrutura industrial e verticalizada da grande gravadora que passa a dar lugar gradativamente às empresas digitais (De Marchi, 2011: 147-152). O controle exercido por estas novas empresas na fruição de música via *streaming* (*Spotify*, *Youtube*, entre outras) é uma das principais características desta

¹ A partir do começo do século XXI esta histórica divisão não abarca mais a complexidade das relações travadas no universo da música popular. A intermediação entre público e artistas sai do controle da grande gravadora industrial para empresas digitais como *Spotify* ou *Youtube*. Para detalhes conferir De Marchi (2011: 147-152).

² Para uma discussão mais detalhada sobre os contornos específicos da trajetória da independência musical no Brasil conferir Alves (2017: 98-106).

³ Uma análise que permita uma visualização detalhada da pesquisa sobre a música independente no Brasil ainda está por ser feita. No âmbito de uma dissertação de mestrado, Belisa Giorgis (2017: 22-27) fez um levantamento inicial neste sentido. O predomínio da área da Comunicação nestas abordagens é notório.

nova fase da intermediação fonográfica, com sérias implicações para a diversidade musical do país (Vicente, Kischinhevsky & De Marchi, 2018: 26).

Não obstante as dificuldades para a produção musical independente, em particular para o rock nacional,⁴ nos dias de hoje, o acesso recente às ferramentas de registro e circulação musical permitiu a contextos musicais como o da produção de rock de Teresina superar um velho entrave: a materialização em obras (produtos fonográficos) de um repertório, em qualidade e quantidade, representativo da diversidade artística que já caracteriza a capital do Piauí há um bom tempo (Medeiros, 2013: 41). Assim, se existiam várias complexidades para o registro fonográfico até meados dos anos 2000, nos dias de hoje essa prática é recorrente, generalizada e acessível.

É devido a estas mudanças recentes no mundo da música que não se abriu mão da contextualização histórica na delimitação do rock autoral de Teresina enquanto estudo de caso. Assim, procedeu-se a uma reconstituição de trajetória, sempre tomando o contexto nacional mais amplo como padrão comparativo.⁵ Com início em 1966 em torno do movimento artístico da Jovem Guarda e com o aparecimento da primeira banda local, a trajetória deste segmento atravessou as décadas, sendo marcado por diversas fases que viram o surgimento e a consolidação do rock autoral de Teresina como o conhecemos hoje.⁶

A contextualização do rock independente de Teresina nestes termos foi fundamentada a partir dos pressupostos elencados por Pierre Bourdieu nas suas análises sobre os segmentos artísticos especializados. Mais especificamente, no postulado de que o estudo da dimensão histórica é fundamental para a compreensão dos mecanismos de funcionamentos da atitude estética e da própria obra de arte (Bourdieu, 2007: 283).

Após a reconstituição histórica, buscou-se verificar a existência de certos atributos inerentes aos campos artísticos. Estes são: (I) um conjunto de obras produzidas que contenham marcas específicas do contexto social em que foram geradas (Bourdieu, 2003: 123); (II) a existência de um corpo de críticos e curadores que atuam, via de regra, em instâncias específicas de legitimação; (III) a existência de instâncias específicas de reprodução e difusão das obras (Bourdieu, 1996: 326).

As obras são verificadas em duas dimensões: uma fonográfica, com mais de 200 produtos catalogados até o momento de escrita deste artigo, distribuídos em diversos formatos, seja em suportes materiais, seja no ambiente digital;⁷ uma videográfica, com mais de 90 vídeos catalogados. Os críticos e demais divulgadores/ legitimadores atuam seja no âmbito da imprensa tradicional, seja na imprensa alternativa. Estas instâncias de reprodução e divulgação foram verificadas a partir da catalogação de mais de 50 entidades promotoras de eventos e mais de 70 espaços de fruição (2014-2016), distribuídos no âmbito público, privado e do terceiro setor.⁸

Em resumo, foram delimitadas quatro dimensões de análise para investigar o rock independente de Teresina: (I) a sua genealogia, com a reconstituição desde o marco-zero até o período de desenvolvimento da pesquisa (1966-2016); (II) a morfologia, conjunto de agentes e instâncias em interação, analisada a partir de um trabalho de levantamento dos projetos musicais, promotoras de eventos, espaços de fruição musical e instâncias de divulgação e

⁴ Uma pesquisa sobre o papel do rock no panorama musical macro brasileiro neste início de século XXI ainda está por ser feito. O rock deixa de figurar na lista das 100 músicas mais executadas nas rádios em 2017.

⁵ Para isso foi feita também uma reconstituição histórica do rock no Brasil, desde as primeiras movimentações na década de 1950 que desencadeariam na Jovem Guarda até estes primeiros anos do século XXI, período marcado pela abundância de artistas atuantes, segmentação em nichos de consumo e ostracismo no *mainstream* cultural.

⁶ Para mais detalhes sobre a contextualização histórica conferir Alves (2017: 158-222).

⁷ Mais de 220 produtos fonográficos foram catalogados até o momento. Estes estão distribuídos em suportes materiais (CD, LP e K7), em arquivos digitais (álbuns, *extended plays*-EP, demos, *singles*, *bootlegs*, compilações) e outros tipos de fonogramas. Para mais detalhes conferir Alves (2017: 351-389).

⁸ Para mais detalhes conferir Alves (2017: 223-324).

legitimação na imprensa; (III) as dinâmicas de interação entre estes agentes e instâncias a partir de dois pontos inerentes aos campos – as alianças que permitem alguma coesão interna e os conflitos em busca de hegemonia; (IV) produtos, que consistiu na catalogação e análise das obras geradas. Todas estas dimensões estão presentes no website Arquivo Rock Teresina.

4. A importância dos arquivos para as ciências sociais e sua aplicação sobre os contextos musicais periféricos e independentes brasileiros

O arquivo foi o formato adotado para a organização do material a ser divulgado via Arquivo Rock Teresina. Em linhas gerais, “o arquivo é o local para armazenamento de documentos e registros” (Featherstone, 2006: 591). O termo tem sido utilizado majoritariamente para definir o ato de reunir informações no âmbito das políticas de inteligência dos estados modernos. Com a disseminação das tecnologias digitais na vida cotidiana, os conteúdos que podem potencialmente ser organizados desta maneira aumentam consideravelmente. Já existem, inclusive, arquivos derivados de pesquisas sobre o rock no âmbito das Ciências Sociais em língua portuguesa (Quintela & Guerra, 2017).

Além das razões elencadas anteriormente, a opção por construir este arquivo também está relacionada com as necessidades, inerentes a qualquer trabalho de investigação científica nas Ciências Sociais: o de encontrar meios próprios para melhor responder às indagações que estruturam os problemas de pesquisa. Ainda que isso não anule a necessidade óbvia de obedecer diretrizes gerais inerentes à disciplina em que se está inscrito e ao próprio fazer científico, “o sociólogo ativo não somente pode como deve improvisar as soluções que funcionam onde ele está e resolve os problemas que ele quer resolver” (Becker, 1993: 13).

O questionamento geral desta investigação consistiu em analisar a forma como estava estruturado o conjunto de agentes e instâncias em interação cujas atividades possibilitavam a permanência de um segmento de rock autoral em Teresina no início do século XXI. Portanto, foi a necessidade de um levantamento meticuloso que fornecesse uma panorâmica, qualitativa e quantitativa, do fenômeno investigado: o pontapé inicial que culminou no lançamento do Arquivo Rock Teresina.

As duas características básicas para uma proposta artística figurar no arquivo foram: a formatação em uma estética roqueira; a disponibilização de registros fonográficos que permitissem comprovar as atividades autorais.

Na construção deste tipo de arquivo é fundamental conciliar múltiplas técnicas, cruzando dados advindos de diferentes fontes. É neste sentido que o surgimento deste catálogo foi possível graças à articulação da recolha documental ao trabalho no terreno. Em ambas as situações foram feitos procedimentos de natureza *on* e *off-line*.

Em investigações como estas, que abordam objetos de estudo associados à produção e circulação de obras artísticas, a internet tem um papel-chave. Mesmo tratando-se de uma área territorial limitada, correspondente à cidade de Teresina, o estudo das fontes *on-line* tornou-se fundamental para alcançar os objetivos de exaustividade e representatividade da amostra.

No caso dos documentos, buscou-se uma diversidade de fontes que possibilitassem cobrir com eficácia o recorte temporal delimitado para a análise (2001-2016). Em primeiro lugar, foram vasculhadas as plataformas digitais onde os artistas depositam os fonogramas. Algumas das principais foram: *Spotify*, *Deezer*, *Bandcamp*, *Palco MP3*, *Toque no Brasil*, *Soundcloud* e *Last FM*.

O acompanhamento sistemático da programação no período em que ocorria a pesquisa foi outro procedimento importante. Assim, a partir da coleta das informações contidas nos cartazes de eventos foi possível não apenas enriquecer o catálogo de artistas. Mas, também, construir uma lista dos principais espaços de fruição dedicados ao rock em Teresina entre os anos de 2014-2016.⁹

⁹ Coleta feita em duas frentes: acompanhamento de sites focados na programação cultural de Teresina (Agenda THE); acompanhamento dos perfis nos *social mídia* (bandas, entidades promotoras de eventos,

Contudo, um problema surgiu: como catalogar de forma eficaz os projetos artísticos que desenvolveram atividades no início dos anos 2000, período em que as plataformas e a divulgação dos eventos musicais não convergiam tão intensamente para o ambiente digital? Foi no intuito de preencher esta lacuna que não se abriu mão do trabalho de terreno. Este consistiu na realização de 32 entrevistas, também nas modalidades *on* e *off-line*, com um conjunto variado de agentes que exercem ou exerceram atividades no rock independente de Teresina no decorrer deste início de século XXI.

Muitos destes agentes contribuíram para o incremento do catálogo de dois modos: a partir da disponibilização de material dos seus acervos pessoais; a partir da interlocução continuada com o pesquisador após a entrevista, fornecendo informações sobre novos fatos, eventos, programações e produtos referentes ao segmento analisado.

Portanto, devem ser levadas em consideração duas questões principais que não podem prescindir de um trabalho de catalogação nos moldes do que é aqui apresentado:

- 1) O estudo de documentos *on-line* acerca de um objeto que envolve múltiplas instâncias e agentes não pode direcionar-se a um único tipo de material. É por essa razão que, além das plataformas digitais onde os artistas comumente disponibilizam as canções, foram examinadas também as programações e outras fontes documentais, sobretudo no âmbito da imprensa.
- 2) O estudo aprofundado de documentos não diminui a importância de um trabalho no terreno cuidadoso e substancial. A catalogação de artistas ausentes ou com presença irrisória no ambiente digital foi possível graças a este trabalho.

Em suma, uma catalogação representativa do contexto analisado somente se faz a partir da multiplicidade de técnicas de recolhidas e de fontes materiais. Caso contrário, é muito provável a permanência de lacunas.

Mesmo diante de todos estes procedimentos, importa sublinhar a improbabilidade de uma coleta total na construção deste tipo de catálogo. Em primeiro lugar, pela rapidez com que surgem e desaparecem artistas com os atributos potenciais para figurar neste arquivo, muitas vezes deixando pouquíssimas pistas de existência. Este tipo de situação ocorre principalmente nos primeiros anos 2000, época em que a informação convergia em menor intensidade para o ambiente digital.

Outro caso bastante comum é o de artistas que afirmam posturas autorais das quais, porém, não foram encontrados produtos fonográficos que as comprovassem. Há ainda o caso de artistas que estavam registrando os seus primeiros trabalhos na altura do fechamento parcial para a escrita deste artigo. Nestes dois últimos casos, estes nomes podem vir a ser incluídos numa próxima atualização na medida em que forem verificados fonogramas que comprovem a atividade autoral.

Portanto, o objeto deste arquivo é dinâmico, em contínua evolução. Não se propõe aqui uma catalogação completa. Isso não deve significar a impossibilidade de considerações científicas. Importa sublinhar que, apesar da consciência sobre a improbabilidade de construir uma amostra total, o catálogo tem por objetivo ser o mais representativo possível. Vide a triangulação de fontes e a coleta contínua e prolongada no tempo. A adoção destes dois pressupostos tem permitido não só a construção de um catálogo que forneça uma panorâmica do estado atual do rock independente de Teresina. Mas, também, a possibilidade de refletir de modo embasado sobre potenciais mudanças, continuações e tendências futuras.

É neste sentido que a catalogação tem continuado mesmo após a finalização da tese. Sobre este aspecto, cabe ressaltar sobre as possibilidades de *continuar* uma tese mesmo depois de cumprido o rito acadêmico da defesa. O gosto por esta continuação deriva na vontade de aprofundar pontos que surgiram no decorrer do trabalho e que, por várias razões, podem (e devem) ser mais desenvolvidos (Eco, 2007: 234).

entre outros). No ano de 2015, por exemplo, foram catalogados mais de 180 eventos de rock independente em Teresina.

5. A estrutura e o conteúdo do Arquivo Rock Teresina

O website está dividido em quatro grandes áreas: a lista de bandas e artistas em carreiras-solo que atuam ou atuaram no rock independente de Teresina nos primeiros anos do século XXI (Artistas); a lista de produtos fonográficos recolhidos *on-line* e no terreno nos mais diversos formatos (Registos fonográficos); a lista de uma parte importante¹⁰ da grande produção videográfica (Videoclipes); as entrevistas com agentes de destaque, registradas em formato vídeo, no decorrer de 2015 (Entrevistas).

Na parte superior da tela, estão as informações complementares. Na secção (Sobre) há uma apresentação breve do site e da equipe que o criou. Na secção (Curiosidades) o intuito é apresentar alguns acontecimentos interessantes da história do rock independente de Teresina descobertos no decorrer da pesquisa.

Uma lista da produção bibliográfica, com a disponibilização de *links*, é o foco de (Destques). A ideia para esta secção surge de outra constatação verificada na pesquisa: a necessidade de expandir quer a produção acadêmica sobre a temática analisada, quer as conexões entre os pesquisadores.

(Contatos) é a última secção da página *web*. É aqui onde são disponibilizados os endereços de *social mídia* (Facebook) e e-mail. Uma vez que o conceito do site pressupõe a construção colaborativa, este espaço busca a interlocução com o público visando principalmente dois pontos: o *feedback* mais geral a respeito do conteúdo com críticas e sugestões. E o recebimento de informações e materiais novos que, após avaliação dos administradores, podem vir a compor o acervo.

Para uma visão panorâmica, segue uma fotografia da sessão Artistas.

Figura 1. O Arquivo Rock Teresina: Artistas



Fonte: Alves & Serafino (2018), Artistas.

¹⁰ A produção audiovisual é vasta. Além dos videoclipes, há uma série variada de registros (*lyric vídeos*, filmagens amadoras do público, documentários, entre outros) referentes ao rock independente de Teresina.

Tendo em vista a discussão feita até o momento, esta última parte objetiva ilustrar como está estruturado o conteúdo do site. Isto é feito a partir da apresentação de uma secção: Artistas. Até o momento de escrita deste artigo constavam no catálogo 151 nomes. Estes estão distribuídos em 5 subgrupos: Rock Pop, Indie Rock, Híbridos, Rock Vintage e Outros. Para mais detalhes sobre os critérios que nortearam a catalogação conferir Alves (2017: 223-228).

Depois de inserido no respectivo subgrupo, a etapa final foi uma descrição a mais direta possível de cada projeto musical. Portanto, não apenas a pura e simples listagem. Mas um breve comentário analítico que situe imediatamente o leitor na proposta catalogada. Além do nome do artista e de uma breve descrição, existe também no site o *link* para a página oficial de cada projeto musical. Como explicado, alguns artistas foram catalogados a partir do diálogo com agentes, que forneceram material fonográfico que comprovava a produção autoral. De todo modo, o número de artistas que disponibiliza material na *web* é maioritário.

Buscou-se respeitar critérios de clareza e diretividade anteriormente assinalados. Sem reproduzir o conteúdo da secção, que pode ser consultado diretamente no site, abaixo segue uma explicação resumida de cada subgrupo.

- 1) Rock Pop: Alguns dos artistas mais relevantes em termos de visibilidade do rock independente de Teresina estão situados neste subgrupo. O termo configura uma categoria que serve para ilustrar a ambivalência que tem caracterizado o rock. Esta é verificada na sua relação íntima com as indústrias culturais apesar do seu discurso contestador (Friedlander, 2002: 12; Regev, 2013: 10). No Brasil, o Brock (Dapieve, 1995), referente aos grupos que atingiram o *mainstream* musical na década de 1980, foi a principal manifestação do gênero. O Rock Pop teresinense, quer em termos estéticos (sonoridade mais leve, letras em português), quer em termos de visibilidade na música autoral da capital do Piauí, está alinhado com o Brock. É composto até o momento de escrita deste artigo por 42 nomes, representando 27% do conjunto geral.
- 2) Indie Rock: A dimensão mais alternativa do rock independente de Teresina. Enquanto um gênero de música popular, o *indie* tem origens no contexto britânico da segunda metade da década de 1980 (Hesmondhalgh, 1999: 38). Até o momento de escrita deste artigo foram catalogados 37 nomes no Indie Rock em Teresina, correspondendo a 25% do conjunto geral.
- 3) Híbridos: A hibridação cultural (Canclini, 2003) aqui é baseada no diálogo entre rock e elementos da cultura brasileira e nordestina. Consiste nas inúmeras possibilidades de fusão entre um gênero musical nascido em um contexto anglo-saxão e manifestações culturais locais (Bennett, 2001: 2; Regev, 2013: 11). No paralelo com o panorama nacional mais amplo, está alinhada com a produção musical da década de 1990, fortemente influenciada por estes procedimentos, e que ficou conhecida por Música Pop Brasileira (MpopB) (Abramo, 1996: 1; Alexandre, 2013: 31). Alguns dos artistas mais importantes em termos de visibilidade também estão neste subgrupo. Foram catalogados até o momento de escrita do artigo 32 nomes, correspondendo a 21% do conjunto geral.
- 4) Rock Vintage: Manifestação, no rock independente de Teresina, de uma tendência na cultura *pop* atual: a revisitação do seu passado (Reynolds, 2011). No caso específico do rock isso é verificado principalmente nos artistas que formatam suas músicas a partir dos cânones do gênero, reunidos sob o termo Classic Rock. Estes constituem um grupo heterogêneo, cronologicamente situado entre a primeira metade da década de 1960 até meados dos anos de 1970 (Shuker, 1998: 235). Foram catalogados até o momento da escrita do artigo 19 nomes, correspondendo a 13% do conjunto geral.
- 5) Outros: conjuntos de artistas não-enquadráveis em nenhum dos subgrupos anteriormente citados, mas que, por diferentes razões, cravaram seus nomes na história do rock independente de Teresina. O critério para a inserção de cada artista está na descrição de cada projeto musical proposto no site. Foram catalogados 21 nomes até o momento da escrita do artigo, correspondendo a 14% do conjunto geral.

6. Considerações finais

A importância de divulgar de forma clara os processos metodológicos e os resultados que subjazem as considerações cientificamente orientadas foram as principais motivações para a escrita do artigo. Sublinha-se, aqui, que a validação em pesquisas qualitativas nas Ciências Sociais passa necessariamente pela coerência, mas, também, pela transparência dos procedimentos adotados (Serafino, 2019: 65).

É neste sentido que se buscou, complementarmente, à divulgação dos resultados, a apresentação clara dos métodos e técnicas de recolha de dados que permitiram a formação do Arquivo Rock Teresina, além dos critérios utilizados para a delimitação do estudo de caso objeto de análise. Em suma, além dos resultados em si, buscou-se explicitar os modos como se chegaram às conclusões cientificamente orientadas.

A escolha por uma postura científica nestes termos está ligada a implicações éticas. Em primeiro lugar, na premissa de que o retorno dos resultados da pesquisa sociológica para uma fatia mais ampla da sociedade é um imperativo. Este retorno à sociedade passa necessariamente por uma divulgação em termos claros, não herméticos, amparada em escrita rigorosa, porém, simples. E que, quando necessário, sejam criados instrumentos específicos de divulgação fora da seara acadêmica.

A divulgação científica nestes termos permite o debate mais democrático não apenas dos resultados de um determinado estudo, mas também dos processos e lógicas do fazer acadêmico nas Ciências Sociais. Além disso, fomenta mais intensamente a continuação e/ou comparação de pesquisas sobre uma dada temática.

Espera-se, finalmente, que um desenho de pesquisa seguido de uma divulgação nos termos aqui apresentados contribua para a formulação de indicadores que sirvam para a discussão de estratégias voltadas à promoção de democracia cultural. Isso significa promover e fomentar manifestações culturais que não estão associadas, por um lado, à cultura erudita, e, por outro lado, à oferta abundante do grande mercado de bens culturais (Botelho, 2007: 173).

Por apresentar números de uma faceta substancial da música autoral da cidade, uma iniciativa como a do Arquivo Rock Teresina pode também contribuir discussões que justifiquem, na pauta governamental, políticas culturais que levem em consideração o segmento analisado, a partir de indicadores produzidos de modo científico. Uma vez que o site fornece a possibilidade de visualização detalhada do volume de artistas e produtos do rock autoral da capital do Piauí, pode vir a ser utilizado para subsidiar esse tipo de discussão, uma vez que prova a existência, no tempo e no (*ciber*)espaço, de um segmento artístico-musical já bem sedimentado. Porém, ainda pouco contemplado no seu potencial pelas políticas culturais, e ainda não devidamente reconhecido, na sua dimensão e diversidade, por boa parte da população de Teresina.

É dessa forma que se podem proferir previsões mais embasadas sobre as mudanças, permanências e tendências futuras da produção de música rock autoral de Teresina. A posse deste tipo de indicador aumenta consideravelmente as possibilidades de ações estratégicas mais assertivas, ancoradas em pressupostos ligados não apenas à divulgação científica, mas também à democracia cultural.

7. Referências

- ABRAMO, B. (1996). Rock Made in Brasil. *Revista Teoria e Debate*.
- ALBAGLI, S. (1996). Divulgação Científica. Informação científica para a cidadania?. *Revista IBICT*, 3, 369-404. Recuperado de: <http://revista.ibict.br/ciinf/article/view/639/643>
- ALEXANDRE, R. (2013). *Cheguei bem a tempo do palco desabar: 50 casos e memórias do rock brasileiro (1993-2008)*. Porto Alegre: Arquipélago.
- ALVES, T. (2017). *Genealogia, morfologia, dinâmicas e produtos do rock independente de Teresina no início do século XXI*. 566f. Tese (Doutorado em Sociologia). Universidade do Porto, Porto Alegre. Recuperado de: <https://hdl.handle.net/10216/106637>
- _____ (2018). Outros territórios da música independente brasileira no início do século XXI: uma discussão a partir do catálogo de artistas do rock independente de Teresina (2001-2017). *Revista Hodie*, 18 (2), 254-269. Recuperado de: <https://doi.org/10.5216/mh.v18i2.52374>
- ALVES, T. & SERAFINO, I. (2018). *Arquivo Rock Teresina: Artistas, Teresina, 2018*. Recuperado de: <https://rockteresina.wixsite.com/arquivo>. Acesso em: 04/06/2020.
- BECKER, H. S. (1993). *Métodos de pesquisa em ciências sociais*. São Paulo: Hucitec.
- _____ (2015). *Truques da escrita: para começar e terminar livros, teses e artigos*. Rio de Janeiro: Zahar.
- BENEVIDES, M. F. (2019). *Música além da música: uma etnocartografia dos afetos e das coletividades na cena Rock de Fortaleza-CE*. 336f. Tese (Doutorado em Sociologia). Universidade Federal do Ceará, Fortaleza. Recuperado de: <http://www.repositorio.ufc.br/handle/riufc/45664>
- BENNETT, A. (2001). *Cultures of Popular Music*. Maidenhead: Open University Press.
- BOURDIEU, P. (1996). *As Regras da Arte: gênese e estrutura do campo literário*. São Paulo: Companhia das Letras.
- _____ (2003). *Questões de sociologia*. Lisboa: Fim de Século.
- _____ (2007). *A Distinção: crítica social do julgamento*. Porto Alegre: Zouk.
- BOTELHO, I. (2007). Políticas Culturais: discutindo pressupostos. En G. Marchiori Nussbaumer (Org.), *Teorias e políticas da cultura: visões multidisciplinares* (pp. 171-180). Salvador: EDUFBA.
- BUENO, W. C. (2010). Comunicação científica e divulgação científica: aproximações e rupturas conceituais. *Revista Informação & Informação*, 15 (1), 1-12. Recuperado de: <http://dx.doi.org/10.5433/1981-8920.2010v15n1espp1>
- CANCLINI, N. (2003). *Culturas Híbridas: estratégias para entrar e sair da modernidade*. São Paulo: Editora da Universidade de São Paulo.
- DAPIEVE, A. (1995). *BROCK. O rock brasileiro dos anos 80*. São Paulo: Editora 34.
- DE MARCHI, L. (2011). Discutindo o papel da produção independente brasileira no mercado fonográfico em rede. En M. Herschmann (Org.), *Nas bordas e fora do mainstream musical: novas tendências da música independente no início do século XXI* (pp. 145-163). São Paulo: Estação das Letras e das Cores.
- ECO, U. (2007). *Como se faz uma tese em ciências humanas*. Barcarena: Editorial Presença.
- FEATHERSTONE, M. (2006). Archive. *Revista Theory, Culture & Society*, 2-3, 591-596. Recuperado de: <https://doi.org/10.1177/0263276406023002106>
- FERRARI, P. (2010). *Jornalismo digital*. São Paulo: Editora Contexto.
- FRIEDLANDER, P. (2002). *Rock and Roll: uma história social*. Rio de Janeiro, São Paulo: Record.
- GIORGIS, B. Z. (2017). *Arqueologia da mídia da Apanhador Só: produção de presença na cibercultura*. 152f. Dissertação (Mestrado em Processos e Manifestações Culturais) – Programa de Pós-Graduação em Processos e Manifestações Culturais da Universidade Feevale. Recuperado de: <https://biblioteca.feevale.br/Dissertacao/DissertacaoBelisaZoehlerGiorgis.pdf>

- HERSCHMANN, M. (2010). *Indústria da música em transição*. São Paulo: Estação das Letras e Cores.
- HESMONDHALGH, D. (1999). Indie: the institutional politics and aesthetics of a popular music genre. *Revista Cultural Studies*, 1, 34-61. Recuperado de: <https://doi.org/10.1080/095023899335365>
- IBGE – Instituto Brasileiro de Geografia e Estatística. (2010). *Sinopse do Censo Demográfico 2010*. IBGE, Rio de Janeiro, Disponível em: <https://censo2010.ibge.gov.br/sinopse/index.php?dados=6&uf=00>. Acesso em: 10/05/2020.
- MEDEIROS, H. (2013). *Acordes na cidade: música popular em Teresina nos anos 1980*. 131f. Dissertação (Mestrado em História) – Programa de Pós-Graduação em História do Brasil da Universidade Federal do Piauí. Recuperado de: <https://docplayer.com.br/10269122-Acordes-na-cidade-musica-popular-em-teresina-nos-anos-1980.html>
- QUINTELA, P. & GUERRA, P. (2017). Ciências sociais, arquivos e memórias: considerações a propósito das culturas musicais urbanas contemporâneas. *Revista Sociologia*, 33, 155-181. Recuperado de: <http://ojs.letras.up.pt/index.php/Sociologia/article/view/2825>
- REGEV, M. (2013). *Pop-rock music: aesthetic cosmopolitanism in late modernity*. Cambridge: Polity Press.
- REYNOLDS, S. (2011). *Retromania: Pop Culture's Addiction to its Own Past*. London: Faber & Faber.
- SCHOPENHAUER, A. (2005). *A Arte de Escrever*. Porto Alegre: L&PM Pocket.
- SERAFINO, I. (2019). Processos de criação artística comunitária: questões metodológicas. *Revista Sociologia*, 28, 52-70. Recuperado de: <https://doi.org/10.21747/08723419/soc38a3>
- SHUKER, R. (1998). *Popular music: the key concepts*. Manchester, Nova Iorque: Routledge Taylor & Francis Group.
- VICENTE, E.; KISCHINHEVSKY, M. & DE MARCHI, L. (2018). A consolidação dos serviços de streaming e os desafios à diversidade musical no Brasil. *Revista EPTIC*, 20 (1), 25-42. Recuperado de: <https://seer.ufs.br/index.php/eptic/article/view/8578>

Autor y autora.

Thiago Meneses Alves

Departamento de Comunicação Social (Jornalismo) de la Universidade Federal do Piauí, Brasil.

Doctorante en Sociología pela Faculdade de Letras de la Universidade do Porto, Portugal. Profesor del Departamento de Comunicação Social (Jornalismo) de la Universidade Federal do Piauí.

E-mail: thiagomeneses85@gmail.com

Irene Serafino

Universidade Federal do Piauí, Brasil.

Doctorante en Sociología pela Faculdade de Letras de la Universidade do Porto, Portugal. Investigadora independente. Colaboradora del Núcleo de Pesquisas Sobre Crianças, Adolescentes e Jovens de la Universidade Federal do Piauí, Brasil.

E-mail: irene.serafino85gmail.com

Citado.

MENESES ALVES, Thiago e SERAFINO, Irene (2022). Estratégias de divulgação científica nas ciências sociais: escolhas teórico-metodológicas para a criação do website Arquivo Rock Teresina. *Revista Latinoamericana de Metodología de la Investigación Social - ReLMIS*, N°23, Año 12, pp. 76-88.

Plazos.

Recibido: 14/07/2020. Aceptado: 09/12/2020.

Reseña bibliográfica:

El método biográfico: estrategias metodológicas para analizar la relación individuo-sociedad en el siglo XXI

Reseña del libro:

MECCIA, Ernesto (Dir.) (2019). *Biografías y Sociedad. Métodos y Perspectivas*.

Santa Fe: Ediciones UNL – EUDEBA

Matías Sbodio

Biografías y sociedad. Métodos y Perspectivas es un extenso volumen metodológico dirigido por Ernesto Meccia; tal vez el primero en presentar en el campo académico argentino tanta variedad de formas de llevar adelante investigaciones de corte biográfico.

Se advierten tres características singulares. Primero, es una obra interinstitucional en la que se dan cita investigadorxs provenientes del Instituto de Investigaciones Gino Germani, de la Universidad de Buenos Aires, y del Departamento de Sociología de la Universidad Nacional del Litoral. Segundo, es intergeneracional: incluye tanto a investigadorxs de extensa trayectoria como a estudiantxs y jóvenes recientemente egresados de sus carreras de grado. Tercero, el estilo pedagógico-didáctico: el libro está pensado como una caja de herramientas de la que pueden servirse docentes y estudiantes en contextos áulicos, especialmente en talleres de tesis. No presenta prescripciones metodológicas en abstracto, sino un abanico de alternativas para encarar trabajos donde, si bien la biografía es central, los intereses teóricos que llevan al uso de la misma son distintos, es decir, procura poner en íntima relación las posturas teóricas con la elección de las estrategias metodológicas.

Previamente, en el capítulo introductorio, Meccia trabaja para demostrar que el incremento de las inquietudes académicas por el método biográfico en las Ciencias Sociales está vinculado con las transformaciones sociales de las últimas décadas. Plantea que las sociedades atraviesan un proceso de *biografización* (Delory-Momberger, 2009). Vivimos en sociedades profundamente reflexivas (Giddens, 1995) producto, entre otras variables, de la incertidumbre generada por la desintegración de la sociedad salarial, la caída de los macro-relatos modernos y el concomitante surgimiento de movimientos sociales que traen a las arenas públicas cuestiones relacionadas con las vidas realmente vividas, las vidas que deberían vivirse y las vidas que (nos) deberían importar.

Según Meccia, nuestro tiempo se caracteriza por un clima subjetivo particular que posiciona tanto al individuo como a la narración de su vida en el centro de la escena. Como si fuera poco, esta información sobre la vida de la *gente común* goza de un alto grado de publicidad, un fenómeno que Sibila (2008) denomina *extimidad*. Sin lugar a dudas, la revolución que suponen Internet y el desarrollo de las redes sociales ocupan un lugar central en este proceso. La civilización digital, cuyos indicadores podemos mirar en las *historias* de Instragram, o leer en la *biografía* de Facebook, viene materializando lo que Meccia denomina *show del yo*, una característica de la subjetividad biográfica contemporánea.

Como se adelantó, las Ciencias Sociales no son ajenas a este proceso. Por el contrario, las transformaciones sociales impactan sobre las formas de hacer ciencia, y sus efectos van

aterrizando en las producciones teóricas y en las reflexiones metodológicas (Plummer, 2001; Merrill & West, 2009).

Según prologa Piovani en esta obra, las instituciones dedicadas a la formación en Ciencias Sociales no han puesto el foco precisamente sobre el método biográfico, al menos en Argentina. Piovani advierte que el escaso desarrollo de la formación en investigación biográfica quizás se deba a que la misma pone el ojo sobre los individuos y su relación con lo social. Esto se explica recordando que el individuo constituye un *problema no resuelto* en varias perspectivas analíticas sociales (Rustin, 2005). Para Piovani, esta desatención representa una vacancia de la que *Biografías y Sociedad* se hace cargo.

Quien se sumerja en esta lectura, apreciará que se encuentra frente a un tratado metodológico sobre la producción de datos empíricos a partir de las biografías de las personas. El libro presenta un alegato de la figura del individuo y de sus relatos como fuentes de datos empíricos. Enraíza así el giro de las Ciencias Sociales hacia el interés por las interpretaciones de primer grado (Schutz, 1974), como son los datos biográficos en general, y en particular las narrativas del yo.

En cuanto a sus contenidos específicos, el libro propone diferentes estrategias de aplicación de lo que define como *método biográfico*, sin desentenderse de las reflexiones filosóficas, epistemológicas y teóricas que implica el estudio de las biografías y, en particular, la referida a la intrigante relación entre individuo y sociedad cuyo estudio reclamaba hace más de sesenta años Charles Wright Mills en su clásico "La imaginación sociológica".

¿Qué es el método biográfico y para qué sirve? El "método biográfico" designa un amplio conjunto de procedimientos para la producción de datos empíricos relativos al estudio de la vida de los individuos. Los procedimientos pueden enmarcarse en la metodología cualitativa y/o cuantitativa y están destinados a dar cuenta de un transcurso, de un devenir, es decir: son datos que deben informar sobre los impactos del paso del tiempo en las biografías. El tiempo comprendido por las investigaciones puede coincidir con el de toda una vida, o, más generalmente, con algunos de sus momentos o transiciones (sin importar que sean largas o breves). El método biográfico estudia las biografías de dos maneras que no son excluyentes: por un lado, puede reconstruir sucesiones de "hechos" biográficos o, por otro, reconstruir las "experiencias" de la vida. Cada una de estas maneras representan ventanas de observación de los fenómenos biográficos que requieren procedimientos (y convocan conceptos) distintos: los "hechos" refieren a lo que efectivamente pasó, a cuestiones fácticas que (se) sucedieron; las "experiencias", en cambio, a las formas que tiene la gente de significar esos hechos por intermedio de su propia memoria biográfica. A menudo los hechos se reconstruyen estadísticamente (las nociones de "curso de vida" o "trayectoria" ayudan a dar una idea concreta); las experiencias, casi sin variación, se reconstruyen cualitativamente (las nociones de "relato de vida" o "narrativa" hacen lo propio).

Partiendo de una amplia revisión de investigaciones empíricas, Meccia identifica diferentes focos analíticos, y los utiliza como criterios para presentar las cuatro estrategias. Precisamente la pregunta que estructura la obra es: ¿para qué sirven las biografías en una investigación social? ¿Qué se puede hacer a través de ellas? ¿Qué se puede observar desde ellas? Las biografías, sostiene Meccia, son una ventana para ver zonas del mundo social, raramente interesan por sí mismas.

La primera estrategia consiste en utilizar las biografías *para ilustrar fenómenos socioestructurales*. Esta estrategia observa cómo las trayectorias de vida de las personas se adaptan a tendencias históricas y también hacen esa tendencia, habilitando un viaje de lo microsocio a lo macrosocio, y viceversa. En esta clave aparecen, por ejemplo, el capítulo de Mercedes Di Virgilio y Mercedes Najman sobre trayectorias residenciales de familias de sectores populares y sectores medios que muestra la desigual intensidad de la movilidad espacial de los individuos a lo largo de su vida; algo que las autoras vinculan con factores estructurales tales como la clase social en su relación con el mercado inmobiliario y las disputas del territorio. En la misma línea, el capítulo de Pablo Dalle estudia procesos de movilidad social de distintas generaciones de familias de origen popular. El autor sostiene que, observando trayectorias de las familias, es posible dar cuenta de transformaciones en diversas dimensiones de la estructura social. Esta primera estrategia permite observar relaciones entre biografías individuales y amplios procesos de

cambio social. Por ejemplo, cómo trayectorias biográficas de descenso social son condicionadas por determinados eventos sociales y espacios residenciales (capítulo de Paula Boniolo). O cómo trayectorias de ingreso universitario se relacionan con desigualdades sociales, pero también con procesos contemporáneos como la individualización (capítulo de Virginia Trevignani). Para las investigaciones que parten de esta estrategia, los hechos sociales, lo fácticamente acontecido, constituyen un punto nodal.

La segunda estrategia propone utilizar las biografías *para realizar historia oral o micro-historia*, es decir, para acceder a historias no oficiales de acontecimientos históricos. Estas investigaciones exploran hechos históricos y a la gente hundida en ellos, tratando de rescatar su propia perspectiva histórica para ayudar a dar ciudadanía a las memorias no oficiales. Las historias profanas pueden ser puestas en relación, por ejemplo, con estadísticas o registros de diversa índole. No obstante, Meccia expone que la intención nunca es verificacionista, ya que el valor radica en la subjetividad con que los individuos tiñen los hechos vividos con sus versiones individuales o colectivas. Justamente en el versionado de quien vivió esos acontecimientos está el dato biográfico más interesante. El estudio del discurso y la memoria aparece entonces como un rasgo característico.

La tercera estrategia procura utilizar las biografías *para dar cuenta de procesos de des- y re-socialización* experimentados por personas luego de un suceso particular o *turning point*, como suele decirse. En efecto, caídas en el consumo abusivo de drogas (véase el capítulo de Esteban Grippaldi) o en el mundo de los juegos de azar (capítulo de Astor Borotto), decisiones de asumir públicamente identidades no heterosexuales, o el acaecimiento de una enfermedad crónica (capítulo de Daniel Jones y Juan Pedro Alonso), suponen para los sujetos un arduo trabajo de reconstrucción biográfica que, no pocas veces, alienta la búsqueda de una redefinición del yo en contextos grupales (en sentido amplio). Otro ejemplo es el capítulo en el que Yamila Gómez describe procesos de re-construcción identitaria de un conjunto de judíos ortodoxos gays mediante la interacción dentro de la asociación Judíos Argentinos Gays (JAG). Allí puede verse el trabajo biográfico de ensamble de dos identidades sociales en principio no compatibles. Este estilo también puede verse en el capítulo de Lucía Pussetto, quien reconstruye calendarios de sufrimiento explorando los sentidos atribuidos al sufrimiento en los relatos de un grupo mujeres bajo tratamiento psiquiátrico sin redes de contención familiar.

La cuarta estrategia plantea utilizar las biografías *para revelar las marcas narrativas de los individuos*. Las narrativas son pensadas como formas discursivas que, si bien no son un espejo de los hechos vividos, sirven para dotarlos de sentido. Meccia explica que: "los relatos son construcciones que no informan primariamente las *verdades fácticas* de una historia de vida sino las *verdades narrativas* que maneja el narrador" (Meccia, 2020: 54). Los relatos del yo construyen imágenes públicas de ese yo, y es claro que el valor de esas imágenes solo se puede entender si lo ponemos en relación con ciertas regulaciones sociales (hegemónicas o alternativas) relativas a cómo deberíamos ser las personas (Plummer, 1995). He aquí el sentido profundo –dialógico y referido– de poner los relatos del yo al servicio de una reflexión sobre la identidad social y personal.

El libro es presentado como el primer volumen, y se ha dedicado a presentar dos de las cuatro estrategias propuestas: la que utiliza biografías *para la reconstrucción de entidades socioestructurales* y la que hace lo propio *para la reconstrucción del yo en contextos grupales*. Será interesante contar con el segundo volumen para seguir trabajando en los talleres de tesis la cuestión biográfica la cual, acorde a nuestros tiempos, se ha convertido en uno de los temas predilectos de las nuevas generaciones de investigadorxs.

Bibliografía

DELORY-MOMBERGER, Ch. (2009). *Biografía y educación. Figuras del individuo-proyecto*. Buenos Aires: Facultad Latinoamericana de Ciencias Sociales-Facultad de Filosofía y Letras, Universidad de Buenos Aires.

GIDDENS, A. (1995). *Modernidad e identidad del yo. El yo y la sociedad en la época contemporánea*. Barcelona: Península.

MECCIA, E. (director) (2019). *Biografías y sociedad. Métodos y perspectivas*. Santa Fe: Ediciones UNL. Buenos Aires: Eudeba.

MERRILL, B. & West, L. (2009). *Using Biographical Methods in Social Research*. London: Sage Publications.

MILLS, C. W. (1974). *La imaginación sociológica*. México: Fondo de Cultura Económica.

PLUMMER, K. (1995). *Telling Sexual Stories. Power, Change and Social Worlds*. London: Routledge.

—— (2001). *Documents of Life 2. An Invitation to A Critical Humanism*. London: Sage Publications.

RUSTIN, M. (2005). Reflections on the Biographical Turn in Social Science. In: P. Chamberlayne, J. Bornat & T. Wengraf (Ed.) *The Turn to Biographical Methods in Social Science. Comparative Issues and Examples*. London y New York: Routledge.

SCHUTZ, A., (1974). *El problema de la realidad social*. Buenos Aires: Amorrortu.

SIBILA, P. (2008). *La intimidad como espectáculo*. Buenos Aires: Fondo de Cultura Económica.

Autor.

Matías Sbodio.

Universidad Nacional del Litoral (UNL); Instituto de Humanidades y Ciencias Sociales del Litoral (IHuCSO - UNL - CONICET), Santa Fe, Argentina.

Licenciado en Ciencia Política por la Universidad Nacional del Litoral (UNL). Doctorando en Estudios Sociales (UNL).

E-mail: matias.sb@live.com.ar

Citado.

SBODIO, Matías (2022). El método biográfico: estrategias metodológicas para analizar la relación individuo-sociedad en el siglo XXI. *Revista Latinoamericana de Metodología de la Investigación Social - ReLMIS*, N°23, Año 12, pp. 89-92.

Plazos.

Recibido: 17/06/2021. Aceptado: 02/08/2021.