Rodríguez, J., Reguant, M., y Ortega, D. (2024). Un caso práctico de análisis de datos cualitativos con IRaMuTeQ: análisis lexicométrico de narrativas de hombres y mujeres bisexuales. *Revista de Investigación Educativa, 42*(2).
DOI: https://doi.org/10.6018/rie.560411

Un caso práctico de análisis de datos cualitativos con IRaMuTeQ: análisis lexicométrico de narrativas de hombres y mujeres bisexuales

A Practical Case Study of Qualitative Data Analysis With IRaMuTeQ: Lexicometric Analysis of Narratives of Bisexual Men and Women

Julio Rodríguez Rodríguez*[1], Mercedes Reguant Álvarez* y Daniel Ortega Ortigoza**
*Departamento de Métodos de Investigación y Diagnóstico en Educación. Universitat de Barcelona (España)
**Departamento de Pedagogía Aplicada. Universitat Autònoma de Barcelona (España)

Resumen

El análisis de datos cualitativos representa una tarea compleja. Existen distintos tipos de análisis de textos, entre ellos, el análisis lexicométrico, el cual combina la riqueza del análisis puramente inductivo con el rigor de lo cuantitativo. Aunque existen ciertas controversias en la utilización de programas informáticos para el análisis cualitativo de datos, no es el caso de la lexicometría, debido a su propia naturaleza. El objetivo de este artículo metodológico es presentar las posibilidades del análisis lexicométrico con IRaMuTeQ, utilizando como ejemplo el estudio de las narrativas de 80 hombres (15%) y mujeres (85%) bisexuales (edad, M=26,55; DT=6,89). Los resultados muestran la utilidad de IRaMuTeQ en la clasificación del material textual, encontrando seis temáticas que aglutinan el 77,50% de las narrativas estudiadas, que tienen que ver con la ausencia de apoyos, las dificultades en la visibilidad, la diferenciación bisexualidad-homosexualidad, el papel del entorno, el miedo al rechazo, y el papel del colectivo LGTB. En conclusión, IRaMuTeQ se presenta como un software útil para la investigación socioeducativa, permitiendo la obtención de significados latentes en el material textual analizado, mediante un procedimiento de trabajo estructurado y sistematizado como el que se presenta.

Palabras clave: análisis cualitativo; educación social; investigación; lexicometría; metodología.

1 **Correspondencia:** Julio Rodríguez Rodríguez, julio.rodriguez.ro@ub.edu, Universitat de Barcelona, 08035 (España).

Abstract

Qualitative data analysis is a complex matter. There are different forms of text analysis, including lexicometric analysis, which combines the richness of purely inductive analysis with the systematic rigour of quantitative analysis. Although there are controversies in the use of computer programmes for qualitative data analysis, this is not the case with lexicometry, due to its very nature. The aim of this methodological article is to present the potential of lexicometric analysis with IRaMuTeQ, using as an example the study of narratives of 80 bisexual men (15%) and women (85%) of different ages (M=26.55; SD=6.89). The results show the utility of IRaMuTeQ for the categorization of textual material, as it found six themes around which 77.50% of the narratives studied revolved. These include themes such as the absence of support, the difficulties for visibility, the bisexuality-homosexuality differentiation, the role of the environment, the fear of rejection, and the role of the LGTB community. In conclusion, IRaMuTeQ is presented as a useful piece of software for socio-educational research, allowing for the extraction of latent meanings in the textual material analysed using a structured and systematised working procedure such as the one described in this article.

Keywords: qualitative analysis; social education; lexicometry; methodology; research.

Introducción y objetivos

Abordar los fenómenos humanos desde el paradigma cualitativo implica considerar la realidad como una construcción social, y destacar el lenguaje como instrumento de simbolización (García Montejo, 2015). El lenguaje permite representar y comunicar ese mundo compartido desde la subjetividad de las personas, y es la investigación cualitativa la interesada en la perspectiva subjetiva de las personas, y su interpretación y construcción de la realidad (Ander-Egg, 2011; Sandín, 2003; Taylor y Bogdan, 1987). La utilización de la metodología cualitativa es "una de las tendencias más significativas de la investigación educativa a principios del siglo XXI" (Sabariego, 2004, p. 63).

La investigación cualitativa se caracteriza por: ser un proceso iterativo y emergente, utilizado principalmente para explorar, comprender y producir nuevos conocimientos; contener abundante información sobre el fenómeno a estudiar; buscar la producción de conocimiento que permita explicar, y tener en cuenta el contexto para comprender lo investigado (Chandra y Shang, 2019).

Una vez elegido el diseño de investigación, es necesario decidir las estrategias de recogida de información y el tipo de análisis de datos a realizar. Son decisiones que guardan estrecha relación con el paradigma de investigación, la naturaleza del problema y los objetivos. Coincidiendo con lo aportado por Trigueros et al. (2018), la variedad de enfoques cualitativos conlleva la ausencia de modelos claros en el proceso de análisis. Por ello, proponen una perspectiva integradora a la hora de abordar el análisis de los datos, teniendo en cuenta tanto el punto de vista como la interpretación de las narrativas y las teorías implícitas de las personas participantes.

Como estrategias de recogida de información cualitativa tenemos entrevistas, grupos focales, observación participante, documentos escritos (diarios de campo, resúmenes, mensajes de WhatsApp o de Twitter) y registros audiovisuales (audio, fotografía y vídeo). Una vez obtenida esta información, se analiza e interpreta. Esto implica reducir

esa información a unidades con significado para ser analizadas. El análisis cualitativo es una actividad interpretativa de los procesos subjetivos, experiencias personales, creencias, etc. de las personas participantes (Chernobilsky, 2009; Massot et al., 2004). Ander-Egg (2011) señala que "El propósito del análisis es resumir las observaciones llevadas a cabo de forma tal que proporcione respuestas a las interrogantes de la investigación. El objetivo de la interpretación es buscar su significado más amplio a las respuestas mediante su trabajo con otros conocimientos disponibles." (p. 160).

Existen diferentes maneras de afrontar el análisis de datos cualitativos, según las preguntas de investigación, el origen de estos datos, las perspectivas metodológicas, y los propios métodos de análisis (Díaz Herrera, 2018; González Teruel, 2015; Miles y Huberman, 1994; Taylor y Bogdan, 1987): el recuento de palabras clave, el análisis de redes semánticas, el análisis del contenido (cuantitativo y cualitativo), el análisis de las estructuras gramaticales, o las conversaciones y narrativas (Ryan y Bernard, 2000), la teoría fundamentada (Glaser y Strauss, 1967) y el análisis crítico del discurso (van Dijk, 1999), o el análisis temático (Braun y Clarke, 2006). El análisis de datos cualitativos comparte el objetivo de analizar y comprender en profundidad teniendo en cuenta, o no, el contexto en el que se generan esos datos cualitativos. Buscan, en definitiva, acercarse al significado latente de las narrativas a partir del contenido manifiesto. Como señala Abela (2002), "todo contenido de un texto o una imagen pueden ser interpretados de una forma directa y manifiesta o de una forma soterrada de su sentido latente" (p. 2). En cualquier caso, el análisis de datos cualitativos es una tarea compleja, pudiendo complicar el abordaje de esta tarea tanto al alumnado universitario, como investigadoras e investigadores noveles. Un análisis cualitativo exige un procedimiento sistemático y cuidadoso, ya que cabe la posibilidad de perderse en el amplio mundo de los datos cualitativos (Ary et al., 2010; Gibbs, 2012; Kalpokaite y Radivojevic, 2019). El volumen de datos en una investigación cualitativa supera al que se maneja en una investigación cuantitativa (Gibbs, 2012; Taylor y Bogdan, 1987). Hacen falta instrumentos que faciliten el manejo de los datos y la toma de decisiones, y un *software* adecuado puede ser una ayuda necesaria (Huber y Marcelo, 1990; Richards y Richards, 2002; Saldaña, 2013).

En la diversidad de tipos de análisis de datos textuales existentes, el análisis estadístico de textos reúne "aquellos procedimientos que, mediante el cómputo de las ocurrencias de una o varias unidades verbales básicas permiten realizar, a partir de los resultados obtenidos, algún tipo de cálculo" (Jordà et al., 2000, p. 610). Este tipo de análisis, también conocido como lexicometría "agrupan una serie de métodos que permiten reorganizar las unidades presentes en una secuencia textual y operar ciertos análisis estadísticos a partir de la cuantificación del vocabulario resultante de una segmentación operada sobre el texto" (Gil et al., 1994, p. 511). La lexicometría aprovecha las posibilidades del análisis estadístico y la informática, combinando la economía de procesamiento con la precisión que se otorga a los métodos cuantitativos.

En este tipo de análisis "las personas elaboran representaciones que se configuran en concepciones sobre distintos fenómenos del mundo. La descripción de las categorías a través del método lexicométrico mostraría los aspectos más estables, de estas formas de pensar" (Baccalá y de la Cruz, 2000, p.2). Básicamente, permite identificar "las ideas "nucleares" que se infieren de las palabras características y sus relaciones, es decir, las que se infieren del conjunto de palabras asociadas en cada categoría (Baccalá y de la

Cruz, 2000, p.2). Este análisis se basa en la frecuencia de distintos vocablos y secuencias de términos y su proximidad con otros.

La aproximación estadística a datos textuales permite la obtención del sentido latente del discurso, y proporciona una descripción completa, cuantitativa, y estructurada del texto (Bécue et al., 1992; Peralta et al., 2020). El análisis lexicométrico puede: identificar tendencias o preferencias de términos emergentes en el pensamiento de un individuo, así como el pensamiento colectivo de un grupo; y realizar inferencias en base a diferentes variables inherentes a la persona o personas participantes (Romero–Perez et al., 2018).

Las oportunidades que ofrece la tecnología también se han hecho presentes en los diferentes tipos de análisis de los datos cualitativos. Las siglas CAQDAS (*Computer-assisted Qualitative Data Analysis Software*) o QDA *software* (*Qualitative Data Analysis*), agrupan los programas informáticos de análisis de datos cualitativos (Carvajal, 2002; Farias y Montero, 2005). Trigueros et al. (2018) señalan las posiciones antagónicas de partidarios y detractores del uso de herramientas informáticas para el análisis de datos cualitativos. Como señalan hábilmente, es un debate estéril, puesto que no hay que perder de vista que su utilización no garantiza la calidad del análisis, si bien, la bondad de su uso depende de la postura ideológica, epistemológica y ética de las investigadoras e investigadores. En su opinión, las ventajas en su uso tienen que ver con el ahorro de tiempo, la eficiencia en la gestión de la información, el manejo de gran cantidad de datos, y la mejora de la calidad de la investigación, así como el trabajo en equipo y lo relativo a la gestión del proceso de investigación. Sin embargo, estos mismos autores señalan una serie de inconvenientes: trasladar el diseño cuantitativo a la investigación cualitativa, la pérdida de la globalidad del proceso, y la seducción por la cuantificación de los datos. En definitiva, este tipo de *software* facilita el análisis, si bien no substituye la creatividad e implicación del investigador o la investigadora, ni la utilización acrítica del programa informático (Chandra y Shang, 2019; Gibbs, 2012; Hart y Achterman, 2017; Hernández-Sampieri y Mendoza, 2018). Por tanto, el *software* es una ayuda técnica, un medio, pero no es el fin de la investigación, y creemos que exige igualmente la implicación activa dela persona investigadora en todas las fases del proceso del análisis, como podrá verse a propósito del estudio que se presenta en este artículo, en el cual ejemplificamos el análisis lexicométrico utilizando el *software* IRaMuTeQ.

Entre los programas comerciales se encuentran Atlas-ti, Ethnograph o NVivo, si bien el precio de las licencias excede las posibilidades económicas del alumnado y los equipos noveles. Existen otras opciones basadas en el llamado *software* libre (o código abierto), entre las que se encuentran AQUAD, o Weft QDA, por citar algunos, y que se utilizan en investigaciones académicas (Ávalos et al., 2018; Gonçalves et al., 2021; Marina y Feliz, 2018).

También existen opciones para el análisis lexicométrico, entre los que se encuentran: SPAD-T de CISIA, del Centro International de Estadística e Informática Aplicadas (Francia); LEXICLOUD, desarrollado por el Laboratorio de Lexicometría y Textos Políticos de la Escuela Normal Superior de Fontenay-St. Cloud (Francia); FRECON, desarrollado por la Universitat de les Illes Balears (España); Orange, una suite para la minería de datos, desarrollada por el Laboratorio de Bioinformática, Universidad de

Ljubljana (Eslovenia), e IRaMuTeQ, desarrollado por el laboratorio de investigación LERASS (Francia).

IRaMuTeQ (*Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires*) es un *software* libre y gratuito para el análisis multidimensional de textos, desarrollado por Pierre Ratinaud en la Universidad de Toulouse (Moreno y Ratinaud, 2015). IRaMuTeQ realiza la mayoría de los análisis con una intervención mínima de la persona investigadora.

El objetivo de este artículo metodológico ha sido presentar las posibilidades del análisis lexicométrico mediante el *software* libre IRaMuTeQ, utilizando como ejemplo el estudio de las narrativas de hombres y mujeres bisexuales.

Método: El análisis cualitativo mediante el software libre IRaMuTeQ

A continuación, se detalla un procedimiento para la utilización de IRaMuTeQ que resulte útil para el manejo de un gran volumen de datos cualitativos, y dar un enfoque diferente a los análisis de contenido tradicionales. Estas orientaciones pueden resultar de utilidad para investigadoras e investigadores noveles que quieran apoyarse en programas informáticos para el análisis de datos cualitativos. Camargo y Justo (2013) consideran que IRaMuTeQ tiene mucho que aportar al análisis cualitativo en estudios vinculados a Humanidades o Ciencias Sociales. Estos fundamentos básicos no sustituyen el estudio en profundidad de trabajos como el de Loubère y Ratinaud (2014), Moreno y Ratinaud (2015), Molina-Neira (2017), o Ruiz (2017) para utilizar IRaMuTeQ en todo su potencial.

Antes de empezar, resaltar que IRaMuTeQ tiene algunas particularidades técnicas: previamente hay que instalar el *software* libre y gratuito R y algunas librerías adicionales del mismo (Ruiz, 2017). En caso contrario, IRaMuTeQ no funcionará (consultar cualquiera manual para una descripción más detallada). En este artículo, se ha utilizado R 4.0.0, IRaMuTeQ versión 0.7 alpha 2, y Windows 11 Pro.

IRaMuTeQ plantea diversas opciones de análisis de los datos cualitativos, y aunque la Clasificación Jerárquica Descendente según el método de Reinert (Ratinaud, 2018; Sbalchiero, 2018) es una de las más habituales, hay que explorar otras posibilidades de análisis (Souza et al., 2018). En este sentido, IRaMuTeQ permite desde un análisis lexicográfico básico, la Clasificación Jerárquica Descendente; el análisis factorial de correspondencias, o el análisis de similitudes. Con todos estos análisis, se puede estudiar el contenido simbólico de los datos aportados por las personas participantes (Morales, 2021).

Preparando los datos: premisas básicas

Huber y Marcelo (1990) señalaban que, en la investigación cualitativa, el proporciona una "cantidad abrumadora de ricos y sugerentes datos" (p. 69), de manera que el material cualitativo ha de ser preparado para su análisis posterior: transcribir las entrevistas, o preparar una tabla con las respuestas a preguntas abiertas. Se pueden utilizar procesadores de texto y hojas de cálculo mediante licencia de pago, o utilizar *software* libre como OpenOffice. A diferencia de otros programas, IRaMuTeQ no permite

trabajar a la vez con diferentes archivos de texto, así que tendremos el material textual transcrito en un único archivo.

Antes de iniciar el proceso de análisis, es clave revisar minuciosamente el archivo de texto porque posteriormente IRaMuTeQ no permite su edición. Debe verificarse el formato, la ortografía y demás aspectos formales del documento de texto. Diversos autores (Camargo y Justo, 2016; Gourlay, 2019; Molina-Neira, 2017) sugieren:

- Revisar errores tipográficos y de transcripción.
- Suprimir el material textual de la persona investigadora (preguntas, anotaciones…) o bien conservarlas, marcándolas con un guion bajo (por ejemplo "_PreguntoPorLasDificultades_").
- Evitar el texto justificado a la derecha, y no usar negrita, cursiva o subrayado.
- Si se utilizan siglas, como LGTB o LGTBIQ: siempre las mismas.
- Atender a las palabras sinónimas cuando el objetivo no sea abordar el fenómeno desde la lingüística. Puede interesarnos mantener separadas palabras como "homosexual" y "gay".
- No utilizar caracteres especiales en el texto, como " " & * @ # € $ … ( ) [ ] : ; -
- Las palabras compuestas, como "salir del armario", deben unirse con un guion bajo ("salir_del_armario").
- Todos los textos deben estar escritos en el mismo idioma.
- Los números se mantienen en su formato numérico.

Otra particularidad del documento de texto es que debe guardarse en formato *.txt, utilizando la codificación estándar UTF-8 y LF (Camargo y Justo, 2016). Esto puede hacerse desde OpenOffice.

Por último, tener en cuenta la colocación e identificación del material textual. La disposición e identificación del material podría seguir diferentes opciones (Camargo y Justo, 2016; Loubère y Ratinaud, 2014). Si tenemos textos procedentes de entrevistas semiestructuradas[2], una sería poner el contenido de las entrevistas tal cual, como un único párrafo. En este caso, en la primera línea identificamos a la persona participante, o el número de la entrevista, y las variables o características que queramos asignar (edad, género, etc.). El formato es: un asterisco seguido del nombre de la variable (por ejemplo *sex) y guion bajo a continuación para la modalidad de esa variable (por ejemplo *sex_hombre).

2 Los datos aportados para ejemplificar este artículo metodológico, con una muestra de hombres y mujeres bisexuales, provienen de un estudio más amplio acerca de la homofobia en personas LGTBIQ+. Los datos aquí presentados corresponden a una parte del estudio cualitativo en el que se exploran, mediante una encuesta de preguntas abiertas, las dificultades y las personas de apoyo en el proceso de visibilización como persona LGTBIQ+ de esta muestra de hombres y mujeres bisexuales (n=80, 15% hombres y 85% mujeres; M=26,55; DT=6,89) (España). La encuesta se distribuyó online entre diferentes entidades LGTBIQ+ de Cataluña, mediante la colaboración de personas clave de las diferentes entidades. Se informó por escrito de la finalidad del estudio, y se garantizó el anonimato y confidencialidad de todas las personas que participaron, no recogiéndose ningún dato que permitiera identificarlas. Los requisitos para participar eran: ser persona LGTBIQ+, mayor de edad, y con residencia en Cataluña al menos durante el último año.

**** *sujeto_078 *edad_entre25y35 *sex_h *gen_masc *estadcivil_soltera
Pues nunca me he sentido en un armario, mi proceso ha sido más de apertura que de descubrimiento, así que, igual que no tuve que sentar a mi familia y decirles que era heterosexual, no encontré motivo para hacer lo mismo con mi bisexualidad. Así que se ha ido sabiendo de forma natural y espontanea. Lo que sí que me da más respeto son las amistades del pueblo u otros círculos donde me pueda ver más expuesto a comentarios. Por suerte tengo una familia que lo que quiere es que seamos felices haciendo lo que sintamos y con quien sintamos, solo están para apoyarnos con nuestros proyectos y felicidad.

Otra opción sería identificar las diferentes preguntas o bloques temáticos de la entrevista de cada participante. En ese caso, el procedimiento sería similar al anterior, añadiendo una línea adicional para cada pregunta o temática, con el formato: guion, asterisco, temática, guion bajo y modalidad de esa temática o tema (por ejemplo -*tema_dificultades).

**** *sujeto_225 *edad_menor25 *sex_m *gen_fem * *estadcivil_soltera
-*tema_dificultades
Darme cuenta de que no era heterosexual, salir_del_armario y encontrar parejas a quienes no les importara. La gente piensa que al ser bisexual les pondrás los cuernos con alguien del género contrario.
-*tema_apoyo
Mi mayor apoyo mis amigas, que también son parte del colectivo casi todas.

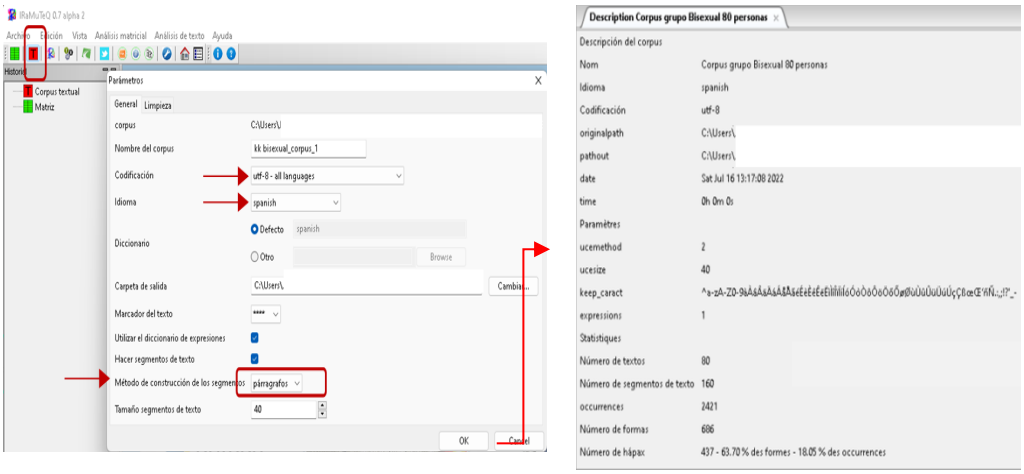
Conviene poner atención al formato del nombre asignado a las variables, ya que el programa diferencia mayúsculas de minúsculas: no sería lo mismo la variable *sex_Hombre que la variable *sex_hombre. Una vez preparado el texto, y guardado en formato *.txt codificación UTF-8, se puede iniciar el análisis lexicométrico con IRaMuTeQ.

Primeros pasos: análisis lexicográfico básico

Cargamos el archivo de texto pulsando el icono "corpus textual". En la pantalla que nos aparece (figura 1) seleccionaremos la opción *utf8-all languages* en la pestaña "Codificación", y el idioma del texto en la pestaña "Idioma". Por defecto, el "Método de construcción de los segmentos" está activa en el método de ocurrencias (se debe cambiar a *párrafos* cuando tenemos respuestas breves a preguntas abiertas de un cuestionario). Esta acción genera una primera información básica sobre el corpus textual que hemos cargado (figura I). En este caso, IRaMuTeQ indica que este corpus textual consta de 80 textos (correspondientes a las 80 personas participantes) y un total de 160 segmentos de texto. Las ocurrencias son 2421 y las formas (palabras genéricas) 686. Las palabras que aparecen con frecuencia igual a 1, denominadas hápax, son en este corpus 437, y representan el 63,70% de las formas y el 18,05% de las ocurrencias.

Figura 1.

Parámetros iniciales y estadísticos descriptivos del corpus textual.

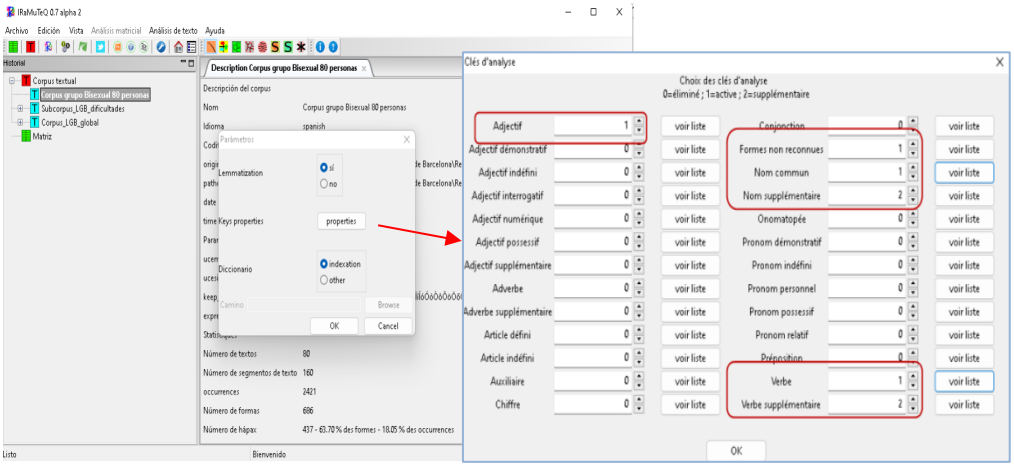


Como señalábamos, IRaMuTeQ permite diferentes análisis de más o menos complejidad. Un primer análisis es el análisis lexicográfico o textométrico que muestra las frecuencias de las palabras, la frecuencia media, y las formas activas y suplementarias. Este primer análisis se ejecuta pulsando el icono "Estadísticas". La opción "propiedades" (figura 2), permite seleccionar el tipo de palabras que queremos incluir en el análisis.

Para corpus textuales vinculados a la Psicología o las Ciencias Sociales, Camargo y Justo (2016) recomiendan activar únicamente las casillas Adjetivos (1), Formas No Comunes (1), Nombre Común (1), Nombre Suplementario (2), Verbo (1) y Verbo suplementario (2), eliminando las "palabras instrumento" (figura II). Esta acción muestra resultados con información importante para continuar análisis posteriores más complejos. La pestaña "Formas activas" (figura III) permite ver el listado total de formas (palabras), su frecuencia y el tipo de palabra, ordenadas según su frecuencia. En nuestro ejemplo (figura IV), los tipos de palabras seleccionados se limitan a los nombres comunes, adjetivos y verbos, como recomiendan Camargo y Justo (2013).

Figura 2.

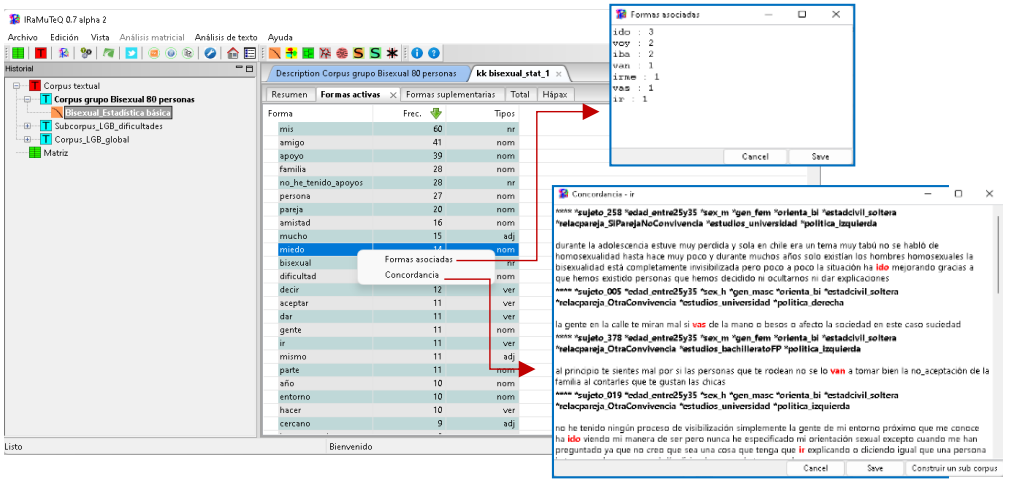
Selección de parámetros y elección de claves de análisis.



Si clicamos con el botón derecho del ratón en cualquiera de las palabras listadas, un submenú permite ver las "formas asociadas" y su frecuencia, y las "concordancias" o fragmentos de texto en los que aparecen esas formas y las asociadas (figura 3). Las formas asociadas son las palabras de la misma familia. Por ejemplo, la forma "ir" tiene como formas asociadas las palabras "ido", "voy", "iba", "irme", "van", "vas" e "ir".

Figura 3.

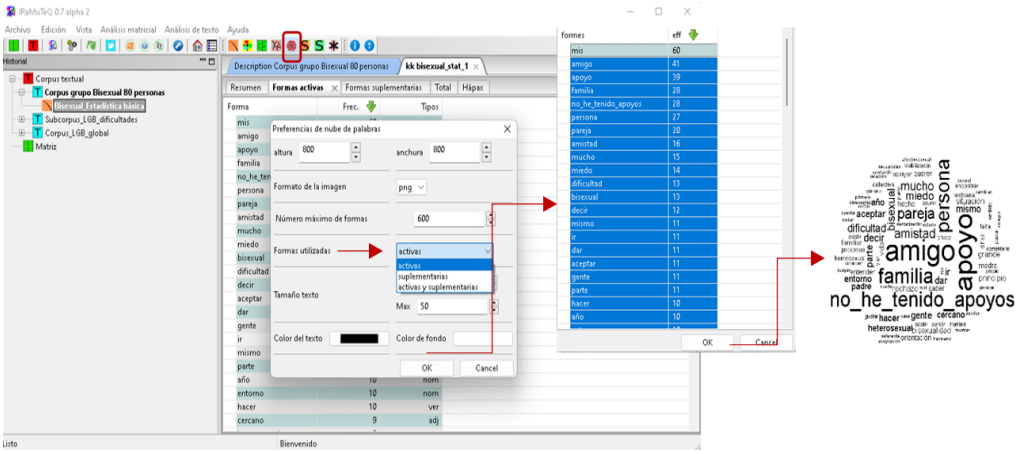
Resumen, formas clave, complementarias, total y hápax.



Otro análisis es la Nube de palabras, una presentación gráfica (figura 4) de las palabras según su frecuencia de aparición en el corpus textual, pudiendo elegir las formas activas, las suplementarias o las formas activas y suplementarias. Se puede seleccionar el color de la nube de palabras y del fondo, el tamaño máximo y mínimo del texto, así como las palabras concretas que queremos que aparezcan dentro de las formas activas, suplementarias, y activas y suplementarias.

Figura 4.

Proceso de generación de la nube de palabras.

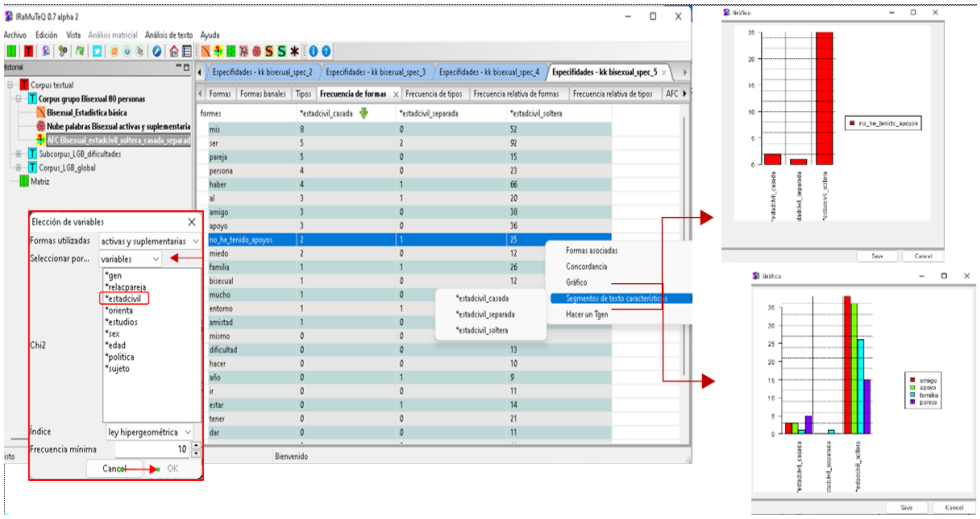


Pasos avanzados (I): el Análisis Factorial de Correspondencias (AFC)

Recordemos que el archivo de texto está organizado en bloques que constan, al menos, de una variable. Lo habitual es que haya más de una (por ejemplo, una que identifica a las personas participantes, otra relacionada con la edad, otra relacionada con el sexo, etc.). Así, una primera variable es en nuestro ejemplo la persona participante, seguida de la edad, el sexo, el género sentido o el estado civil. Para llevar a cabo el análisis factorial de correspondencias (AFC) se elige una variable que tenga al menos tres modalidades (Camargo y Justo, 2016; Ruiz, 2018), como el estado civil, dividida en nuestro ejemplo en: "soltera" (*estadcivil_soltera), "separada/divorciada" (*estadcivil_separada) y "casada" (*estadcivil_casada). Como ilustra la figura 5, la forma "no_he_tenido_apoyos" se encuentra con más frecuencia entre las personas solteras (25), que entre las personas separadas (1) o casadas (2). Este mismo resultado puede representarse gráficamente. Este histograma de frecuencias puede hacerse también por grupos de palabras.

Figura 4.

Análisis Factorial de Correspondencias.



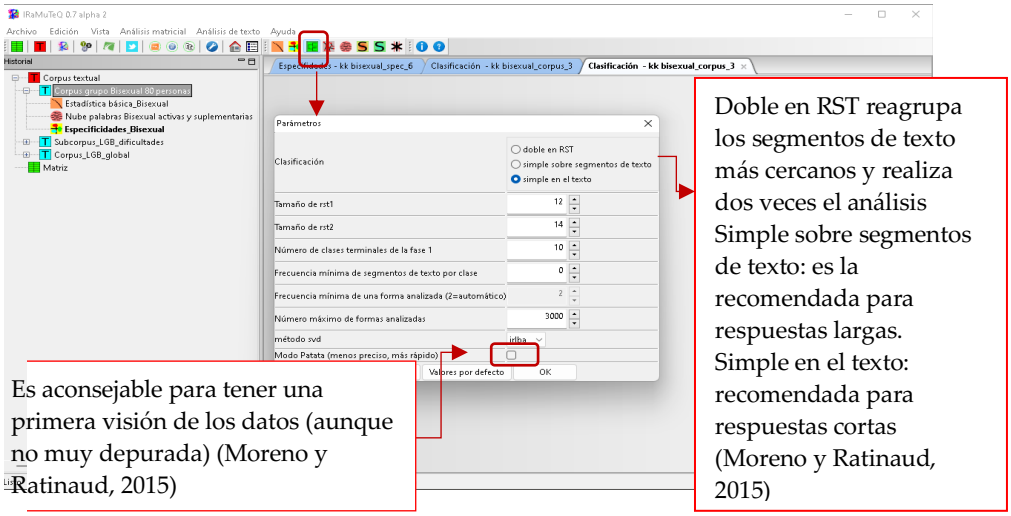
Pasos avanzados (II): la Clasificación Jerárquica Descendente (CHD)

La Clasificación Jerárquica Descendente (CHD) representa la relación entre las formas reducidas y los segmentos del texto. Es decir, IRaMuTeQ agrupa los fragmentos de texto en clases o clústeres de significados, de manera que clasifica las formas en agrupaciones de texto similares en cuanto a su significación. Cada uno de los clústeres tiene una serie de palabras características, ordenadas por su frecuencia de aparición. Ello permite contextualizar el vocabulario típico de cada agrupación.

Es importante, por un lado, seleccionar bien los parámetros de este análisis (figura 5), porque si no, se producen errores técnicos. Así, con fragmentos de respuesta breve, como en un cuestionario de preguntas abiertas, se seleccionará la opción "Simple en el texto" (previamente se habrá seleccionado "párrafos" como método de construcción de los segmentos de texto) (Camargo y Justo, 2016).

Figura 5.

Selección de parámetros de la Clasificación Jerárquica Descendente.

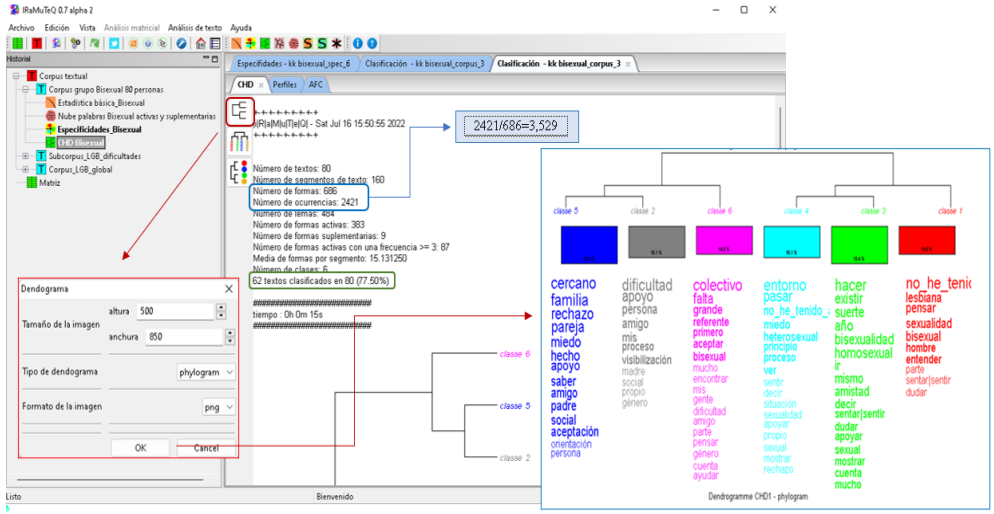


Por otro lado, la CHD requiere una clasificación de los textos mínima del 75% (figura 6). Es decir, que el 75% de los textos quede organizado en las clases que propone IRaMuTeQ. Si el porcentaje es inferior al 70%, Camargo y Justo (2016; 2018) sugieren no tener en cuenta la CHD, y se recurra al análisis de especificidades.

En nuestro ejemplo (figura 5), la clasificación de textos se sitúa en el 77,50% del total, de manera que se puede continuar la CHD. Esto quiere decir que el 77,50% de los textos se clasifican entre las clases que propone el software (restando un 22,50% del material fuera de cualquier clúster). El dendograma resultante (figura VI) muestra como IRaMuTeQ genera estas agrupaciones, seis en nuestro caso. El dendograma resultante (figura 6) muestra como IRaMuTeQ genera estas agrupaciones, seis en nuestro caso. En detalle, muestra una primera división en dos subclases: las clases 5 y 2, junto la clase 6, por un lado, y las clases 4 y 3, y la clase 1, por otro lado. En una segunda división, se separan las clases 5 y 2, y las clases 4 y 3. La división se detiene cuando IRaMuTeQ encuentra, en este ejemplo, seis clústeres estables.

Figura 6.

Resultados preliminares de la CHD y dendograma.

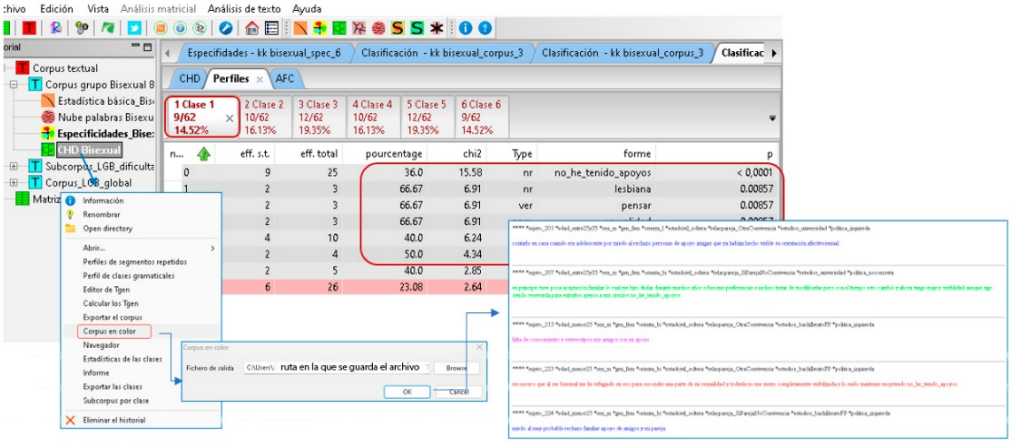


La pestaña "Perfiles" (figura 7) nos da información de cada clúster (Ruiz, 2017): "n" es el número de orden de la forma; "eff. st" es el número de segmentos de texto que contiene la palabra o forma concreta al menos una vez en el corpus; "porcentaje" es el % de ocurrencia de la palabra en los segmentos de texto de ese clúster, en relación con su ocurrencia en el corpus); "*chi2*" representa la fuerza de la asociación de la palabra con su clase; "type" es el tipo de palabra (nombre, verbo...); "Forma", identifica la palabra en cuestión, y "*p*" identifica el nivel de significación de asociación de la palabra con su clase (el nivel de significación asociado a *chi2*) (Moreno y Ratinaud, 2015).

Para un análisis descriptivo del vocabulario de cada clase o clúster, se utilizan simultáneamente dos criterios (Camargo y Justo, 2016): 1) tener en cuenta las palabras con frecuencia mayor que la media del conjunto de palabras de la totalidad del corpus; es decir palabras con una frecuencia superior a núm.ocurrencias/núm.formas (ver figura 6), y 2) tener en cuenta las palabras con $chi2 \geq 3{,}84$ (p<,05).

Figura 7.

Detalles de las clases, sus formas significativas y fragmentos de texto.



La Tabla 1 muestra las diferentes clases, el porcentaje de texto clasificado, las palabras significativas, y su nivel de significación. También se muestra un fragmento significativo del texto que representa esa clase y el nombre atribuido a cada. ¿Qué nombre adquieren estas clases o clústeres? Para ello, las y los investigadores tendrán que revisar en profundidad los fragmentos de texto más representativos de cada clase, atendiendo a la significación estadística (*p*) de cada forma y el valor *chi2* (figura 7). Estas etiquetas se añaden en la pestaña "Perfiles", haciendo doble clic en cada clase.

Tabla 1.

Clases, formas significativas y textos representativos de las clases.

Clase (etiqueta)	**Palabras significativas (*p*)**	**Texto representativo**
1 14,52% "*Ausencia de apoyo en el proceso de visibilización como persona bisexual*"	"no_he_tenido_apoyos" (*p*=,000), "lesbiana" (*p*=,008), "pensar" (*p*=,008), "sexualidad" (*p*=,008), "bisexual" (*p*=,012), "hombre" (*p*=,037)	"Al ser bisexual me he refugiado en eso para esconder una parte de mi sexualidad y todavía no me siento completamente visibilizada y lo suelo mantener en privado. No he tenido apoyos" (s223sexM) "Mi entorno más cercano no entendía cómo podía ser bisexual. No he tenido apoyos" (s246sexM)

Clase (etiqueta)	**Palabras significativas (*p*)**	**Texto representativo**
2 16,13% "*Principales dificultades y personas de apoyo*"	"dificultad" (*p*=,000), "apoyo" (*p*=,000), "persona" (*p*=,003), "amigo" (*p*=,003), "proceso" (*p*=,041)	"Aceptarlo fue difícil ya que quería ser como los demás. Decirlo a mis amigos cercanos y a mis padres fue difícil. Cuando estoy con personas que no conozco mucho intento evitar el tema. Mis amigos cercanos y a mis padres" (s463sexM)
3 19,35% "*Diferenciación de la bisexualidad frente a la homosexualidad*"	"hacer" (*p*=,000), "existir" (*p*=,000), "suerte" (*p*=,000), "año" (*p*=,000), "bisexualidad" (*p*=,000), "homosexual" (*p*=,000), "ir" (*p*=,002), "amistad" (*p*=,003), "decir" (*p*=,007), "apoyar" (*p*=,033), "sexual" (*p*=,033), "sentir/ sentar" (*p*=,033), "mostrar" (*p*=,033), "dudar" (*p*=,033), "cuenta" (p=,033)	"Durante la adolescencia estuve muy perdida y sola en Chile. Era un tema muy tabú, no se habló de homosexualidad hasta hace muy poco, y durante muchos años solo existían los hombres homosexuales. La bisexualidad está completamente invisibilizada pero poco a poco la situación ha ido mejorando gracias a que hemos existido personas que hemos decidido ni ocultarnos ni dar explicaciones." (s258sexM)
4 16,13% "*Papel del entorno en la visibilización*"	"entorno" (*p*=,000), "pasar" (*p*=,000), "no_he_tenido_ apoyos" (*p*=,005), "miedo" (*p*=,007), "heterosexual" (*p*=,017), "principio" (*p*=,041), y "proceso" (*p*=,041)	"Me daba miedo que me juzgaran amigos y me daba aún más miedo que miembros de mi familia me dejaran de hablar. Me sentía incomoda al tener que dar la noticia porque cuando eres heterosexual no la tienes que dar" (s445sexM) "Problema de comprensión al principio" (s441sexM)
5 19,35% "*Miedo al rechazo de los vínculos más próximos*"	"cercano" (*p*=,000), "familia" (*p*=,000), "rechazo" (*p*=,000), "pareja" (*p*=,001), "miedo" (*p*=,002), "hecho" (*p*=,003), "apoyo" (*p*=,009), "saber" (*p*=,016), "amigo" (*p*=,026), "padre" (*p*=,033), "social" (*p*=,033), y "aceptación" (*p*=,033).	"Mis dificultades probablemente he sido yo misma al tardar en asumirlo por miedo a ser jugada (...). Al asumirlo y comunicarlo también encontré una cierta dificultad en el hecho de que a mi madre le chocó, y los primeros días le costó entenderlo, aunque tengo que puntualizar que, aparte de mis padres, mi familia más cercana el resto de mi familia no lo sabe (…)" (s370sexM)
6 14,52% "*Papel del colectivo LGTB*"	"colectivo" (*p*=,000), "falta" (*p*=,000), "grande" (*p*=,002), "referente" (*p*=,008), "primero" (*p*=,008), "aceptar" (*p*=,009) y "bisexual" (*p*=,012).	"Darme cuenta de que no era heterosexual, salir del armario y encontrar parejas a quienes no les importara. La gente piensa que al ser bisexual les pondrás los cuernos con alguien del género contrario. Mi mayor apoyo mis amigas que también son parte del colectivo casi todas" (s225sexM)

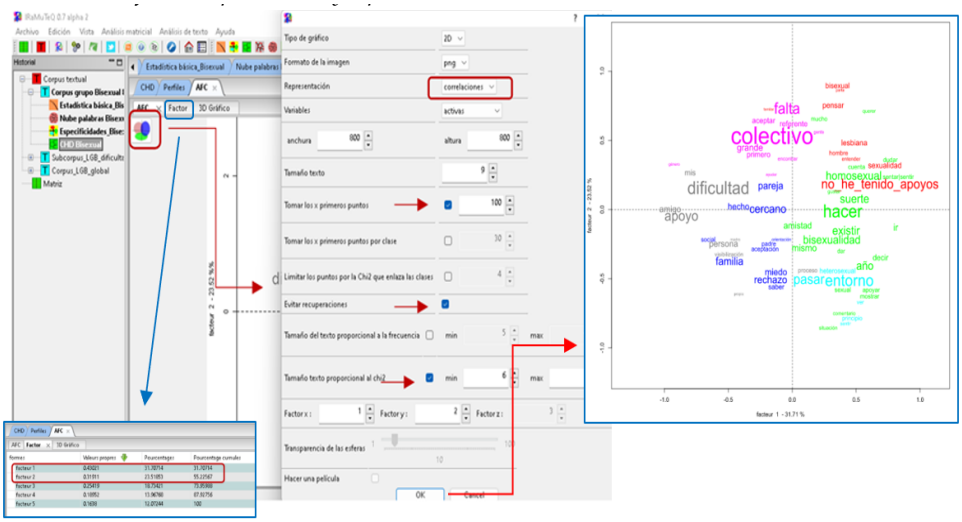
Podemos utilizar el procedimiento de la CHD como paso previo a la categorización manual (un procedimiento más inductivo). O bien, codificar manualmente el material textual y, posteriormente, utilizar el programa para complementar la categorización manual. Para Santos et al. (2020), las diferencias entre la codificación manual y la automática tienen que ver con: a) la primera requiere mayor intervención por parte del equipo investigador, y b) el estilo más intuitivo en la codificación del material que el procedimiento analítico de IRaMuTeQ. En cualquier caso, la codificación automática no elimina la participación activa del equipo investigador en la comprensión y significación de los clústeres.

Además, la CHD permite una segunda parte, que es la extracción de los factores, que son la combinación significativa de variables (figura 8). La proximidad en el espacio significa la correspondencia entre esas formas. El número de factores es igual al número de clases de la CHD menos 1 (Camargo y Justo, 2016). Estos ejes muestran el porcentaje de varianza que recogen. En nuestro ejemplo, el plano factorial 1x2 recoge el 55,22% de la varianza, y en él se representan las variables o las formas.

Para la representación en el eje de coordenadas, Camargo y Justo (2018) sugieren algunas opciones del menú "AFC" (figura 8): en "Representación" marcar la opción correlaciones; "Tomar los primeros puntos" marcando 100; marcar la opción "Evitar repeticiones", y "Tamaño del texto proporcional al chi2".

Figura 8.

Extracción de factores, parámetros y representación cartesiana.



El resultado (figura 8) es una representación de las formas activas en un eje de coordenadas (-1 y +1) en relación con su correlación (*chi2*). La clase 1 tiene una relación positiva con los factores 1 y 2; las clases 2 y 5 tienen una relación negativa con

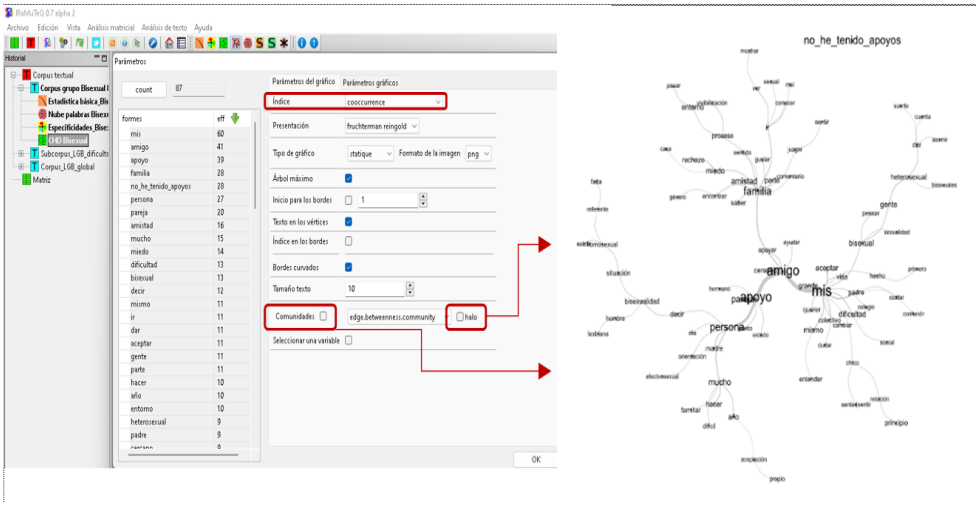
los factores 1 y 2; las clases 3 y 4 tienen una relación positiva con el factor 1 y negativa con el factor 2, y la clase 6 tiene una relación negativa con el factor 1 y positiva con el factor 2. El tamaño de las formas activas se relaciona con su frecuencia, y la proximidad con la correlación entre estas. Los colores de las formas son asignados en función del clúster al que pertenecen. Además, IRaMuTeQ genera por defecto otra representación solamente con las etiquetas de las clases (con el mismo color que las palabras que aglutinan), permitiendo ver como la posición de las clases en relación con los factores.

Pasos avanzados (III): el Análisis de Similitudes

El árbol de similitudes (figura 9) se basa en la teoría de los grafos, y representa gráficamente la estructura de un corpus textual, diferenciando las partes comunes y las específicas de las variables codificadas (Marchand y Ratinaud, 2012).

Figura 9.

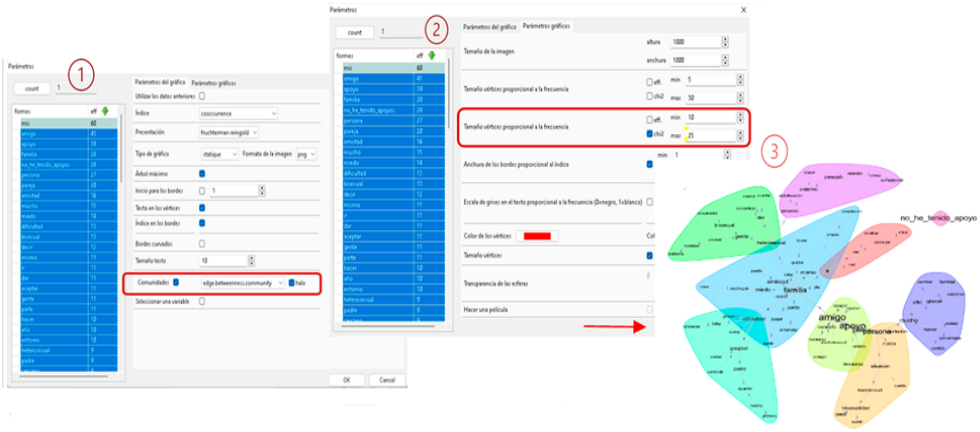
Análisis de similitudes sin configuración.



Las formas/palabras son los vértices del gráfico y las aristas representan las coocurrencias entre ellas, de manera que el tamaño de las palabras aumenta con su frecuencia, y los enlaces entre las palabras son más gruesos en la medida en que son más coocurrentes (Baril y Garnier, 2015).

Figura 10.

Árbol máximo de similitudes con configuración.



Este análisis complementa la información aportada por el CHD, o la sustituye cuando no se dan las condiciones para el segundo (figura 6). Se puede pulir este primer producto seleccionando, a la derecha, las palabras que queremos recoger (por defecto se incluyen todas). Además, se puede clicar la opción "Comunidades" y la opción "halo". Finalmente, para mejorar más la representación, se hacen cambios en "Ajustes gráficos" (figura 10).

Pasos avanzados (IV): subcorpus por temática y subcorpus por metadatos

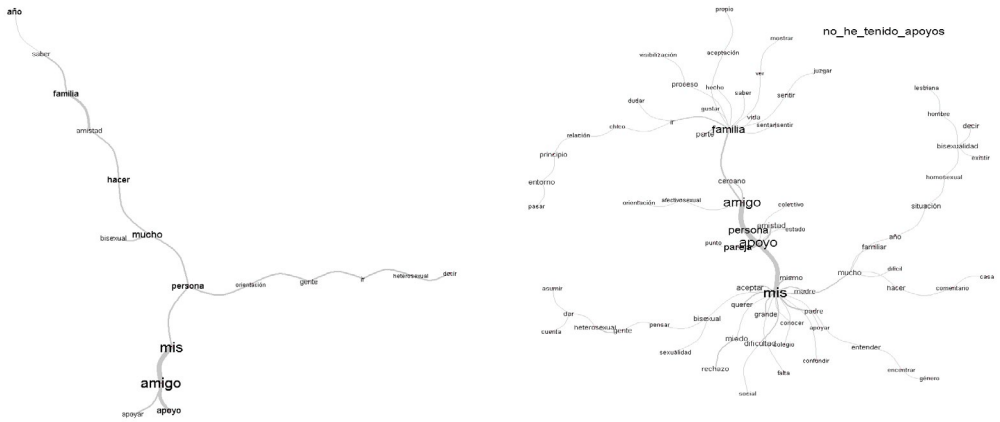
Si queremos trabajar exclusivamente con una temática de los datos, seleccionaremos la opción "Subcorpus por temática", donde aparecerán las temáticas en las que hayamos dividido el corpus textual, dos en nuestro ejemplo: "dificultades" y "apoyo". Si seleccionamos una, podemos realizar los análisis solo para esa temática.

Sin embargo, cuanto más segmentamos el corpus textual, menor es el número de textos a analizar, y esto puede ser un problema. Como estándar, IRaMuTeQ requiere entre 20 y 30 textos, cuando se trata de entrevistas. En el caso de respuestas breves (de tres o cuatro líneas) a un cuestionario, se necesita un número mayor de textos para que el análisis sea apto (Molina-Neira, 2017; Moreno y Ratinaud, 2015). En nuestro ejemplo, en el subcorpus para la temática "dificultades" pueden llevarse a cabo todos los análisis. Sin embargo, el subcorpus para la temática "apoyo", solo permite el análisis de especificidades y el de similitudes.

Finalmente, puede interesarnos analizar por metadatos, es decir, concretando un tipo de variable. Se puede seleccionar más de una modalidad de la misma variable. En nuestro ejemplo, "*sex_h" (sexo hombre), al tener un número reducido de participantes (12 personas), IRaMuTeQ solo permite realizar estadísticos básicos y el análisis de similitudes.

Figura 11.

Dificultades y apoyos en hombres (izq.) y mujeres bisexuales (de.).



En cambio, para "*sex_m" (sexo mujer), se puede realizar el análisis de especificidades y el de similitudes. Este procedimiento ha de permitir comparar la representación de hombres y mujeres sobre sus dificultades y apoyos en el proceso de visibilización como personas bisexuales (figura XI).

Conclusiones

El análisis de datos cualitativos puede abordarse desde distintas perspectivas. La opción de la lexicometría emerge como una forma de obtención de los significados latentes en un texto, prescindiendo del marco referencial del equipo investigador que analiza los datos textuales. Gracias al análisis lexicométrico, y en base a las palabras utilizadas y la forma de expresión adoptada por las personas participantes, emerge el sentido del discurso, bien de una persona o un colectivo.

Entre los diferentes programas informáticos existentes para el análisis de datos textuales, la utilización de programas de código abierto o *software* libre es una opción interesante, especialmente para las personas o los equipos de investigación que quieran utilizar *software* de acceso libre. Entre éstos, destaca IRaMuTeQ, que tiene unas características que lo hacen atractivo para su uso en investigaciones sociales y educativas, aunque hay que tener en cuenta algunas particularidades en su instalación, que pueden provocar el recelo inicial en usuarias y usuarios poco experimentados en la instalación y utilización de *software* libre.

El objetivo de este artículo metodológico ha sido presentar las posibilidades del análisis lexicométrico mediante el *software* libre IRaMuTeQ, utilizando como ejemplo el estudio de las narrativas de hombres y mujeres bisexuales. Como se muestra, IRaMuTeQ ha identificado de forma fiable seis categorías o temáticas de clasificación del corpus

textual. Estas categorías o clústeres son: "*Ausencia de apoyo en el proceso de visibilización como persona bisexual*", "*Principales dificultades y personas de apoyo*", "*Diferenciación de la bisexualidad frente a la homosexualidad*", "*Papel del entorno en la visibilización*", "*Miedo al rechazo de los vínculos más próximos*", y "*Papel del colectivo LGTB*".

El procedimiento de análisis se ha presentado de forma ordenada y sistematizada para aprovechar el potencial del programa informático, teniendo en cuenta que su utilización no substituye la participación y la creatividad de las personas investigadoras a la hora de obtener resultados. En este sentido, IRaMuTeQ organiza el material textual, pero es el equipo investigador quien ha de preparar minuciosamente el material textual, así como decidir si hacer los diferentes análisis disponibles por áreas temáticas previas, o bien con el material textual como un todo. Además, la agrupación del material narrativo (los clústeres) se dotan de significado a partir del ejercicio que hace el equipo investigador de dotarlos de nombre, otorgando así el sentido al contenido manifiesto (las palabras) implicadas en ese clúster. El programa, aunque realiza un análisis automático, no sustituye la creatividad del investigador o la investigadora, dado que son ellas y ellos los que han de dar sentido al material agrupado por IRaMuTeQ.

Referencias

Ander-Egg, E. (2011). *Aprender a investigar. Nociones básicas para la investigación social*. Editorial Brujas.

Ary, D., Jacobs, L. C., y Sorensen, C. (2010). *Introduction to Research in Education*. Wadsworth Publishing

Ávalos, M. A., Merma, G., Hernández-Amorós, M. J., Urrea-Solano, M. E., y Aparicio, M. P. (2018). Rediseño del Plan de Acción Tutorial a partir del grupo focal. El caso de la Facultad de Educación. En: R. Roig-Vila (Ed.), *El compromiso académico y social a través de la investigación e innovación educativas en la Enseñanza Superior* (pp. 911-921). Octaedro.

Baril, E., y Garnier, B. (2015). *IRaMuTeQ 0.7 alpha 2*. http://www.iramuteq.org/documentation/fichiers/Pas%20a%20Pas%20IRAMUTEQ_0.7alpha2.pdf_

Baccalá, N., y de la Cruz, M. (2000). La importancia de la estadística textual aplicada al estudio de las concepciones de enseñanza. En: M. Rajman. y J. C. Chappelier (eds.) 5es Journées Internationales d'Analyse Statistique des Données Textuelles (pp. 519-524). Suiza. https://lexicometrica.univ-paris3.fr/jadt/jadt2000/pdf/39/39.pdf

Becue, M., Lebart, L., y Rajadell, N. (1992). El análisis estadístico de datos textuales. La lectura según los escolares de enseñanza primaria. *Anuario de Psicología*, 55, 7-22. https://www.raco.cat/index.php/AnuarioPsicologia/article/view/61168

Braun, V., y Clarke, V. (2006). *Using thematic analysis in Psychology. Qualitative Research in Psychology, 3(*2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Camargo, B. V., y Justo, A. M. (2013). IRAMUTEQ: Um Software Gratuito para Análise de Dados Textuais. *Temas em Psicologia, 21*(2), 513-518. http://www.redalyc.org/articulo.oa?id=513751532016

Camargo, B. V., y Justo, A. M. (2016). *Tutorial para uso do sotfware IRaMuTeQ*. UFSC Brasil.

Carvajal, D. (2002). Las herramientas de la artesana. Aspectos Críticos en la Enseñanza y Aprendizaje de los CAQDAS. *Forum: Qualitative Social Research*, *3*(2). https://doi.org/10.17169/fqs-3.2.853

Chandra, Y., y Shang, L. (2019). *Qualitative Research Using R: A Systematic Approach*. Springer.

Chernobilsky, L. B. (2009). El uso de la computadora como auxiliar en el análisis de datos cualitativos. En: I. Vasilachis (ed.), *Estrategias de investigación cualitativa* (pp. 239-273). Gedisa.

Díaz Herrera C. (2018). Investigación cualitativa y análisis de contenido temático. Orientación intelectual de revista Universum. *Revista General de Información y Documentación*, *28*(1), 119-142. https://doi.org/10.5209/RGID.60813

Farías, L., y Montero, M. (2005). De la transcripción y otros aspectos artesanales de la investigación cualitativa. *International Journal of Qualitative Methods, 4*(1). https://sites.ualberta.ca/~iiqm/backissues/4_1/pdf/fariasmontero.pdf

García Montejo, S. (2015). Aspectos metodológicos de la investigación cualitativa. En L. Abero, L. Berardi, A. Capocasales, S. Garcia Montejo y R. Rojas, *Investigación educativa. Abriendo puertas al conocimiento* (pp. 101-118). CLACSO.

Gibbs, G. R. (2012). *El análisis de datos cualitativos en investigación cualitativa*. Morata.

Gil, J., García, E., Rodríguez, G., y Corrales, A. (1994) Análisis estadístico de datos cualitativos textuales: enfoque lexicométrico. *Revista Investigación Educativa*. 23, 510-514. http://hdl.handle.net/11162/185311

Glaser, B. G., y Strauss, A. L. (1967). *The Discovery of Grounded Theory. Strategies for Qualitative Research*. Aldine.

Gonçalves, J., Sales, J. P., Moreno, M. M., y Rolim, M. L. (2021). The Impacts of SARS-CoV-2 Pandemic on Suicide: A Lexical Analysis. *Frontiers in Psychiatry*, 12, 593918. https://doi.org/10.3389/fpsyt.2021.593918

González-Teruel, A. (2015). Estrategias metodológicas para la investigación del usuario en los medios sociales: análisis de contenido, teoría fundamentada y análisis del discurso. *Profesional de la Información*, *24*(3), 321–328. https://doi.org/10.3145/epi.2015.may.12

Gourlay, S. (2019). *Preparing IRaMuTeQ input files*. Kingston University.

Hart, T., y Achterman, P. (2017). Qualitative Analysis Software (ATLAS.ti/Ethnograph/MAXQDA/NVivo). En: C. S. Davis y R. F. Potter (Eds.), *The International Encyclopedia of Communication Research Methods*. John Wiley y Sons.

Hernández-Sampieri, R., y Mendoza, C. P. (2018). *Metodología de la investigación. Las rutas cuantitativa, cualitativa y mixta*. McGraw-Hill.

Huber, G., y Marcelo, C. (1990): Algo más que recuperar palabras y contar frecuencias: la ayuda del ordenador en el análisis de datos cualitativos. *Enseñanza*, 8, 69-84.

Jordà, G., Pàmies, C., y Vicens, C. (2000) El análisis lexicométrico en literatura. Propuesta metodológica y aplicaciones. En M. L. Casal, G. Conde, J. Lago, L. Pino y N. Rodriguez (eds.), *La lingüística francesa en España camino del siglo XXI* (pp. 610-617). Arrecife.

Kalpokaite, N., y Radivojevi, I. (2019). Demystifying Qualitative Data Analysis for Novice Qualitative Researchers. *The Qualitative Report, 24*(13), 44-57. https://nsuworks.nova.edu/tqr/vol24/iss13/5

Loubère, L., y Ratinaud, P. (2014). *Documentation IRaMuTeQ 0.6 alpha 3 version 0.1*. http://www.iramuteq.org/documentation/fichiers/documentation_19_02_2014.pdf

Marchand, P. y Ratinaud, P. (2012). L'analyse de similitude appliquée aux corpus textuels: les primaries socialistes pour l'élection présidentielle française (septembre-octobre 2011). Actas del *XI Journées Internationales d'Analyse statistique des Données Textuelles* (pp. 687-699). Liège. Bélgica.

Marina, J., y Feliz, T. (2018). Percepciones en la búsqueda de información y educación para la salud en entornos virtuales en español. *Revista Española de Salud Pública*, *92*, https://tinyurl.com/26jqwe9s

Massot, I., Dorio, I., y Sabariego, M. (2004). Estrategias de recogida y análisis de la información. En: R. Bisquerra (coord.), *Metodología de la investigación educativa* (pp. 329-368). La Muralla.

Miles, M. B., y Huberman, A. M. (1994). *Qualitative Data Analysis. An Expanded Sourcebook*. Sage Publications.

Molina-Neira, J. (2017). *Tutorial para el análisis de textos con el software IRaMuTeQ*. Universitat de Barcelona.

Morales, C. (2021). Uso de software IRaMuTeQ. En: M. A. Casillas, J. Dorantes y C. Ortiz-Blanco (coord.), *Representaciones sociales, educación y análisis cualitativo con IRaMuTeQ* (pp. 23-38). Universidad Veracruzana.

Moreno, M., y Ratinaud, P. (2015). *Manual uso de IRaMuTeQ*. http://www.iramuteq.org/documentation/fichiers/guia-iramuteq

Peralta, N., Castellaro, M., y Santibáñez, C. (2019) Analysing Textual Data As a Methodology to Approach Argumentation: A Research Work Among Undergraduate Students from Chilean Universities**.** *Íkala, Revista de Lenguaje y Cultura, 25*(1), 209-227**.** https://doi.org/10.17533/udea.ikala.v25n01a02

Ratinaud, P. (2018). Amélioration de la précision et de la vitesse de l'algorithme de classification de la méthode Reinert dans IRaMuTeQ. En D. F. Iezzi, L. Celardo, & M. Misuraca (Eds.), *JADT' 2018, Proceedings of the 14th international conference on statistical analysis of textual data* (pp. 616–625). Universitalia. http://lexicometrica.univ-paris3.fr/jadt/JADT2018/actes-jadt18.pdf

Richards, L., y Richards, T. (2002). From filing cabinet to computer. En: A. Bryman y R. G. Burgess, (Eds.), *Analyzing qualitative data* (pp. 146-172). Routledge.

Romero-Pérez, I., Alarcón-Vásquez, Y., y García-Jiménez, R. (2018). Lexicometría: enfoque aplicado a la redefinición de conceptos e identificación de unidades temáticas. *Biblios*, 71, 68-80. https://doi.org/10.5195/biblios.2018.466

Ruiz, A. (2017). *Trabajar con IRaMuTeQ: Pautas*. Universitat de Barcelona. http://hdl.handle.net/2445/113063

Ryan, G. W., y Bernard, H. R. (2000). Data management and analysis methods. En: N. K. Denzin y Y. S. Lincoln (Eds.), *Handbook of Qualitative Research* (pp. 769-802). Sage Publications.

Sabariego, M. (2004). El proceso de investigación (parte 2). En: B. Rafael (coord.), *Metodología de la investigación educativa* (pp. 127-163). La Muralla.

Saldaña, J. (2013). *The Coding Manual for Qualitative Researchers*. SAGE.

Sandín, M. P. (2003). *Investigación Cualitativa en Educación. Fundamentos y tradiciones*. McGraw-Hill.

Santos, I. C., Rosário, V. M., Amaral-Rosa, M. P., Moreira, L., y Ramos, M. G. (2020). Handcrafted and Software-Assisted Procedures for Discursive Textual Analysis: Analytical Convergences or Divergences? En: A. P. Costa, L. P Reis y A. Moreira (Eds.), *Computer Supported Qualitative Research* (pp. 189-205). Springer.

Sbalchiero, S. (2018). Topic Detection: A Statistical Model and a Quali-Quantitative Method. En: A. Tuzzi (ed.), *Tracing the Life Cycle of Ideas in the Humanities and Social Sciences* (pp. 189-209). Springer.

Souza, M. A. R., Wall, M. L., Thuler, A. C., Lowen, I. M. V., y Peres, A. M. (2018). The use of IRAMUTEQ software for data analysis in qualitative research. *Revista Da Escola de Enfermagem Da USP*; 52, e03353. http://dx.doi.org/10.1590/S1980-220X2017015003353

Taylor, S. J., y Bogdan, T. (1987). *Introducción a los métodos cualitativos de investigación*. Paidós.

Trigueros-Cervantes, C., Rivera-García, E., & Rivera-Trigueros, I. (2018). *Técnicas conversacionales y narrativas. Investigación cualitativa con Software NVivo*. Universidad de Granada. Escuela Andaluza de salud Pública. https://www.easp.es/wp-content/uploads/publicaciones/UGR-EASP_Libro_Cualitativa-NVivo_12.pdf

van Dijk, T. A. (1999). El análisis crítico del discurso. *ANTHROPOS*, 186, 23-36.

Fecha de recepción: 10 de marzo de 2023.
Fecha de revisión: 28 de mayo de 2023.
Fecha de aceptación: 9 de octubre de 2023.