REVISIÓN HERRAMIENTAS DE REPRESENTACIÓN DE CONOCIMIENTO PARA PROCESAMIENTO Y ANÁLISIS TEXTUAL DE TESTIMONIOS DEL CONFLICTO ARMADO COLOMBIANO

LEIDYS YOHANNA GÓMEZ PUENTES

UNIVERSIDAD DE ANTIOQUIA

ESCUELA INTERAMERICANA DE BIBLIOTECOLOGÍA (EIB)

MEDELLÍN

22 DE NOVIEMBRE DE 2018

REVISIÓN HERRAMIENTAS DE REPRESENTACIÓN DE CONOCIMIENTO PARA PROCESAMIENTO Y ANÁLISIS TEXTUAL DE TESTIMONIOS DEL CONFLICTO ARMADO COLOMBIANO

LEIDYS YOHANNA GÓMEZ PUENTES

Informe final trabajo de grado para optar el título de bibliotecólogo

Asesora: Ana María Tangarife Patiño

UNIVERSIDAD DE ANTIOQUIA

ESCUELA INTERAMERICANA DE BIBLIOTECOLOGÍA (EIB)

MEDELLÍN

22 DE NOVIEMBRE DE 2018

CONTENIDO

ÍNDICE TABLAS Y GRÁFICOS 5
RESUMEN 6
PRESENTACIÓN 7
JUSTIFICACIÓN 10
OBJETIVOS 12
OBJETIVO GENERAL 12
OBJETIVOS ESPECÍFICOS 12
MARCO CONCEPTUAL 13
LINGÜÍSTICA COMPUTACIONAL 13
ANALISIS TEXTUAL 14
CORPUS 16
Características y clasificación 16
PAPEL TESTIMONIO DENTRO DEL CONFLICTO ARMADO COLOMBIANO 18
Selección y procesamiento corpus de testimonios 21
Análisis formal y de contenido del corpus 22
Rastreo software para el análisis textual 22
Evaluación y comparación software para el análisis textual 23
Construcción matriz para el análisis de resultados 23
CORPUS 24
PROGRAMAS DE ANÁLISIS TEXTUAL 26
NVivo 26
T-LAB plus 27
Linguakit 28
Stilus 28
AntConc 29
EVALUACIÓN SOFTWARE –EJECUCIÓN 29
Hallazgos ejecución programas 31
EVALUACIÓN SOFTWARE – COMPARACIÓN RESULTADOS EN ANÁLISIS DEL CORPUS 41

CATEGORIZACIÓN DE ELEMENTOS (LUGARES, ACTORES, ACCIONES, FECHAS, TESTIGOS) EN CORPUS A PARTIR DE ANÁLISIS TEXTUAL 47
CONCLUSIONES 50
RECOMENDACIONES 52
BIBLIOGRAFÍA 53
ANEXOS 55

ÍNDICE TABLAS Y GRÁFICOS

Gráficos

Gráfico 1 Visualización NVivo 11 Pro-demo 31
Gráfico 2. Visualización programa T-LAB plus.....34
Gráfico 3. Visualización programa LInguakit.....36
Gráfico 4.Visuliazación programa Stilus 38
Gráfico 5. Visualización programa Antconc 40
Grafico 6.Nube de palabras frecuentes Linguakit. Testimonio 5.....42
Grafico 7.Visulización reconocedor de entidades Linguakit. Testimonio 5 43
Gráfico 8.Visualización palabras frecuentes programa NVivo. Testimonio 5.....43
Gráfico 9. Vista resaltada palabras clave en contexto. Programa Linguakit. Testimonio 5 46
Gráfico 10. Vista resaltada palabras clave en contexto. Programa T-lab. Testimonio 5 46

Tablas

Tabla 1. Listado de temáticas en testimonios corpus 24
Tabla 2. Distribución territorial corpus.....25
Tabla 3. Listado de palabras clave programa T-LAB. Testimonio 5.....44
Tabla 4. Listado palabras clave programa Linguakit. Testimonio 5 45

RESUMEN

El presente proyecto de investigación tiene como propósito revisar las herramientas de representación de conocimiento para el procesamiento y análisis textual de testimonios del conflicto armado colombiano. El cual se encuentra enmarcado dentro del enfoque cualitativo utilizando la técnica de investigación documental. Tiene como fin reconocer si estas herramientas permiten responder a las preguntas ¿quién?, ¿cómo?, ¿cuándo? y ¿dónde? que se plantean dentro del marco de la reconstrucción de la memoria histórica del conflicto armado. Inicialmente se presentan los conceptos teóricos contemplados en la definición de los elementos a evaluar, las características del corpus consolidado, los programas recuperados para el análisis textual y por último los hallazgos de la categorización y evaluación de los resultados al analizar el corpus con los programas de análisis textual.

Palabras claves: análisis textual, corpus, testimonios, lingüística computacional

PRESENTACIÓN

El conflicto armado colombiano presentado en los últimos 60 años ha dejado hasta la fecha alrededor de 262 mil víctimas fatales de acuerdo con registro del Observatorio de Memoria y Conflicto (OMC)[1] a causa de hechos violentos por acciones bélicas y ataques a la población civil en todo el territorio colombiano. A causa de estas acciones se han generado procesos de documentación por parte de las víctimas, entidades gubernamentales y no gubernamentales para brindar información (oral, escrita y audiovisual) que permita reconocer e identificar elementos como: modalidad, momento histórico, región y perfil de la víctima con el fin de brindar recursos para reconstruir los hechos violentos y visibilizar de esta manera las ocurrencias del conflicto dentro de un territorio y como medida para garantizar el derecho a conocer la verdad tanto para las víctimas como para la sociedad en general a partir del esclarecimiento de los hechos, actores armados y víctimas, de voz de todos los actores partícipes de este proceso.

Es por esto que se han desarrollado en las distintas comunidades y poblaciones ejercicios para que las personas den a conocer cómo han vivido y afrontado este proceso a través de testimonios, entrevistas, cartas, manuscritos, noticias de prensa, televisivas y radiales; materiales desarrollados en talleres de memoria, fotografías, cantos, audiovisuales, libros, revistas, piezas comunicativas y copias de expedientes judiciales, compartidos por líderes comunitarios, organizaciones sociales, entidades públicas, investigadores del Centro Nacional de Memoria Histórica y Observatorio de Memoria y Conflicto y ciudadanos con documentos de interés. Como un mecanismo de resistencia y protesta para dar a conocer cómo se ha vivido y afrontado el conflicto armado con miras hacia un camino de verdad y reparación que permita identificar y reconocer los elementos que hacen e hicieron parte de este proceso a nivel nacional.

[1] Observatorio de Memoria y Conflicto (2018). Base de datos. Recuperado de: http://centrodememoriahistorica.gov.co/observatorio/

De ahí el interés del gobierno nacional, organizaciones internacionales, organizaciones religiosas, asociaciones de víctimas, organizaciones sociales, entre otras, de consolidar mecanismos que permitan la conservación y recuperación de esta información, como una apuesta para identificar, reconocer y recuperar todos los contenidos que se han generado a lo largo de los años en relación con el conflicto armado y las acciones de memoria. Es por esto, que en el año 2011 con la Ley 1448 o Ley de Víctimas, se ordenó al Centro Nacional de Memoria Histórica constituir un archivo con documentos originales o copias fidedignas referidos a las graves y manifiestas violaciones de derechos humanos e infracciones al Derecho Internacional Humanitario ocurridas con ocasión del conflicto armado interno, además de archivos sobre construcción de paz e iniciativas de memoria. Todo con el fin de lograr una mayor protección y acceso a la información recopilada, sin alterar los documentos originales y garantizando su perpetuidad y su legitimidad; tener un registro oficial de los periodos de violencia; consolidar la información recolectada y construida por las entidades del gobierno y las iniciativas de la sociedad civil en una base de datos que permita la integración de los registros para evitar la fragmentación de la información y la invisibilización del conflicto y los actores armados.

El testimonio es una de las voces del conflicto, utilizado por las víctimas para brindar información acerca de quién le hizo qué a quién, cuándo, dónde y cómo ocurrieron los hechos violentos de los cuales fueron protagonistas. Dentro de la construcción de memoria del conflicto armado el testimonio cumple con cuatro funciones que señala Nieto (2010) “[…] como fuentes de información, como ilustración para una situación descrita, como estrategias para conocer un evento y como instrumentos para la denuncia” (p.78).

De ahí su importancia, y por la cual se le haya posicionado como una fuente de consulta por investigadores y diferentes organizaciones ya que permite identificar elementos desde un género discursivo que llevan a reconocer cómo percibieron,

sintieron y observaron de manera oral o textual un suceso desde uno de los actores del hecho violento. Además, permite identificar responsabilidades y contrastar información de cada una de las fuentes de un suceso con el fin de llevar a cabo investigaciones que permitan validar los hechos violentos reportados.

Desde el campo bibliotecológico es de interés reconocer el testimonio como una fuente de información, sujeta de análisis formal y textual, con el propósito de identificar fenómenos lingüísticos, problemas asociados a la sistematización, así como su estructura semántica, sintáctica y medios de recolección (orales y textuales) que permitan diseñar un modelo de análisis. Es por esto que este trabajo busca iniciar este proceso desarrollando un estado del arte, el cual permita reconocer las metodologías, técnicas y herramientas de representación de conocimiento para automatizar los procesos de análisis textual, como una medida para contribuir a la construcción de la memoria histórica del conflicto y a la realización de las investigaciones sobre la búsqueda de la reconstrucción de la verdad y la participación de los diferentes actores del conflicto armado.

JUSTIFICACIÓN

De acuerdo con el reporte del Observatorio de memoria y conflicto presentado en el año 2018, se encuentran 353.531 hechos documentados relacionados con el conflicto armado, de los cuales se encuentran analizados e indizados en bases de datos 10.236. Estas cifras dentro de los procesos de análisis y tratamiento llevan al bibliotecólogo dentro de su quehacer a plantear mejoras que faciliten los procesos documentales que se llevan para la recolección, procesamiento, análisis, difusión y evaluación de la información contenida en las diferentes fuentes de memoria del conflicto armado colombiano.

Dada la multiplicidad de formatos e instrumentos que se han utilizado como voz del conflicto, se toma como unidad de análisis el testimonio debido a que este reúne ciertos elementos que le proporcionan complejidad desde su estructura discursiva, semántica y su impacto en la sociedad como instrumento de propuesta hacia los hechos de violencia presentados en el transcurso del conflicto armado de este país. Tal como lo señala Díaz (2012) que citó a García (2003: 19-20):

> El género testimonial es un tipo de discurso característico especialmente de América Latina. Si la literatura tiende a emerger como un reflejo de la realidad social en la que se forma no es extraño, por tanto, que un contexto de profundos conflictos históricos, sociales y políticos, de crisis de todo tipo, de dictaduras, gobiernos militares, y de grupos guerrilleros revolucionarios, segregue todo un caudal de textos encaminados a plasmar y a «dar testimonio» de todos estos movimientos opresores que atentan directamente contra los derechos humanos (p.8)

Es por esto por lo que esta investigación busca identificar modelos de representación de la información que permitan llevar a cabo el análisis de los contenidos de los testimonios de manera ágil y rápida respondiendo a las necesidades de las diferentes instancias encargadas del control de esta

información. Por medio del análisis textual automatizado el cual brinda la posibilidad de analizar textos de gran tamaño y permiten realizar recuento de la frecuencia de cada una de las palabras del texto, cualidades que permiten establecer relaciones entre variables y de contrastar hipótesis. Con el fin de proporcionar herramientas para la identificación de patrones discursivos en los testimonios del conflicto armado que determinen en gran medida cuáles son las variables a tener en cuenta el momento de realizar procesos de organización y recuperación de la información y que permitan al grupo de análisis del Observatorio de Memoria y conflicto responder a las preguntas ¿Quién le hizo qué a quién?, ¿cuándo?, ¿dónde y cómo? en cada hecho de violencia documentado.

Todo con el fin de brindar un aporte desde el campo bibliotecológico en el marco del conflicto armado acerca de las herramientas que faciliten el análisis y procesamiento de testimonios para dichas investigaciones que se adelantan y que se brinden recursos que lleven al esclarecimiento de los hechos y las víctimas, para garantizar el derecho a la verdad toda la sociedad colombiana y sobre todo de las víctimas.

OBJETIVOS

OBJETIVO GENERAL

Reconocer las metodologías, técnicas y herramientas de representación de conocimiento para automatizar procesos de análisis textual con el fin de contribuir al diseño de un modelo para análisis de testimonios del conflicto armado colombiano.

OBJETIVOS ESPECÍFICOS

- Caracterizar un corpus de testimonios del conflicto armado colombiano de acuerdo con criterios lingüísticos y de análisis documental.

- Identificar herramientas de representación para el análisis textual de testimonios del conflicto armado colombiano.

- Determinar los requisitos que debe cumplir una herramienta de representación para el análisis textual de testimonios del conflicto armado colombiano.

MARCO CONCEPTUAL

LINGÜÍSTICA COMPUTACIONAL

Guinovart (1998) señala que la lingüística computacional (o lingüística informática) es un campo científico interdisciplinar relativamente reciente de cerca de cincuenta años de investigación y desarrollo cuyo objetivo radica en incorporar en los ordenadores la habilidad en el manejo del lenguaje humano. Llisterri & Moure (1996) definen que la Lingüística computacional engloba una serie de campos diversos que tienen en común la integración de conceptos y procedimientos informáticos en el tratamiento del lenguaje y del habla. Por tanto, se puede decir que la Lingüística computacional es un área interdisciplinaria entre la lingüística y la informática que se ocupa de la construcción de sistemas informáticos capaces de procesar el lenguaje humano por medio del análisis sintáctico, semántico y pragmático del lenguaje escrito y hablado. Entre las aplicaciones y enfoques que presenta, se encuentra:

- Herramientas informáticas para la investigación lingüística: que consiste en la creación de programas que buscan obtener información estadística y distribucional sobre la aparición de unidades lingüísticas y que tienen como objetivo el apoyo de investigaciones lingüísticas.
- El procesamiento del lenguaje natural: tiene como objetivo la realización automática de representaciones de objetos lingüísticos, entre los cuales se encuentra: pasar un texto a una representación con información sobre la categoría gramatical de cada palabra, definición del significado, traducir de una lengua a otra, resumir el contenido del texto, permitir la extracción de información necesaria para recuperarlo después de haberlo introducido en un sistema de archivo.
- Las tecnologías del habla: su interés se centra en transformación del habla en texto de manera automática y a partir de este reconocimiento realizar una

traducción automática de conversaciones, todo esto con el fin de identificar y reconocer patrones que lleven a la creación de sistemas que permitan la comunicación entre hombre - máquina.

- La industria del habla: cubre toda la industria que se encuentra centrada en el desarrollo de productos comerciales destinados a usuarios finales. Estos programas son aquellos que convierten texto al habla y viceversa, los de traducción automática o asistida, los correctores ortográficos y gramaticales que buscan cubrir las necesidades de usuarios no especializados.

ANALISIS TEXTUAL

Consiste en aplicar estos métodos documentales y estadísticos a partir de los datos textuales. El cual tiene como objetivo reproducir la síntesis discursiva, donde se condensa el núcleo semántico, ideológico, informativo, emocional, pragmático que se expresa en las distintas estructuras del texto. Partiendo de la identificación de las relaciones de cohesión y de significado para de esta manera comprender lo que el autor quería emitir.

En relación con los procedimientos documentales y estadísticos, Algranti (1993) define que los procedimientos documentales son aquellos que realizan reorganizaciones de las unidades textuales y se ven representados en los programas como el índice de frecuencias, las concordancias y los segmentos repetidos. Donde el índice de frecuencias informe sobre el número de ocurrencias de una forma en el texto y puede presentarse de forma jerárquica o alfabéticamente. Las concordancias se refieren a la localización del conjunto de concurrencia de una forma en el texto y pueden presentarse de forma alfabética o cronológica. Los segmentos repetidos son aquellos con frecuencia igual o similar a dos, entendiéndose como segmento la sucesión de secuencias no interrumpidas por

caracteres delimitadores de frecuencias como son la coma, el punto y el punto y coma.

Por otro lado, el análisis estadístico se centra en el análisis a través de la comparación. Dado que busca comparar entre sí el discurso de los individuos que han contestado a una encuesta o preguntas pautadas en entrevistas personales, o de grupos de individuos con características comunes. Es por esto por lo que comprende la realización de índices estadísticos y comparaciones entre distintos textos de un corpus. Los productos o funciones que los programas ofertan este tipo de análisis son el análisis factorial de correspondencias y el método de las especificidades.

El *método factorial de correspondencia* requiere antes de su aplicación la elaboración de un sistema de categorías a partir del material verbal. Este permite descubrir afinidades entre dos conjuntos de variables, presentados en forma de tabla de contingencia, tanto de frecuencias como de valores medios. La técnica parte de una tabla de contingencia en la que aparecen en línea las diversas formas gráficas, y en columna las diversas categorías de sujetos; por 10 que cada casilla recoge el número de ocurrencias de una determinada forma gráfica en una categoría de sujetos. Posteriormente, y a partir de esta tabla, se procede a la construcción y análisis de las nubes de puntos generados por dichas variables.

El *método de especificidades* consiste en evaluar si una forma gráfica tiene una presencia significativa más alta en determinadas partes del corpus. El modelo estadístico utilizado para obtener la distribución de probabilidad es la ley hipergeométrica que toma como parámetros el tamaño del corpus, el tamaño de la muestra o parte del corpus considerado, la frecuencia de una determinada forma gráfica en el corpus y la frecuencia de la forma gráfica en la parte del corpus considerada.

CORPUS

Se entiende por corpus como un grupo de documentos susceptibles de análisis, almacenados en algún medio electromagnético que permite la automatización de tareas de búsqueda y recuperación de información (Velásquez & Girón 2010). Es decir, implica la totalidad del material recolectado de acuerdo con unos criterios de selección del cual a partir del análisis se descubren estructuras sistémicas y modelos.

Por tanto, el corpus tiene como propósito servir como referencia para el estudio o el análisis del comportamiento de elementos de interés, siempre y cuando este responda o represente el tema de estudio que será evaluado. De ahí que se requieran establecer criterios para la selección del mismo.

Características y clasificación

Al momento de la construcción de un corpus se deben tener en cuenta algunos criterios que debe cumplir con el fin de garantizar la idoneidad de la información que se selecciona y presentar las siguientes características:

- Representatividad: debe estar conformado por textos reales que permitan representar de manera fiel un código lingüístico determinado, entendiendo que este es una muestra y no la totalidad del conjunto.
- Variedad: considerar cada una de las variables que pueda presentar el tema de estudio que será representado; con el fin de garantizar que abarca la mayor cantidad de variables que pueda presentar el objeto de estudio: tiempo, información personal del informante, localización geográfica, tipo de documento y fuente.

- Equilibrio: en la medida que se permita se debe garantizar un equilibrio entre cada una de las variables en la conformación del corpus con el objeto de garantizar la representatividad de cada elemento.
- Tamaño: aunque esta característica presenta diferentes posturas, sobre cuál es el número ideal de palabas, al momento de la creación se deben tener en cuenta antes que el número de palabras la pertinencia, tiempo disponible para el análisis y facilidad en la adquisición de los recursos, dado que estos determinarán el tamaño adecuado del corpus.

Por otra parte, de acuerdo con los parámetros que se deseen categorizar del corpus, este se puede clasificar por el tipo de documentos, número de lenguas, criterios de recolección, cantidad, distribución, finalidad, tipo de procesamiento y tipo de anotación. De ahí que se presenten corpus orales o textuales, monolingües o multilingües por el número de lenguas involucradas, también estos pueden ser abiertos o cerrados de acuerdo con la posibilidad de desarrollo que estos puedan presentar y a su vez estos también se pueden clasificar por temática y parámetros demográficos. Es por esto, que la clasificación del corpus se debe a la intencionalidad del investigador y de las variables que este tenga en cuenta al representar su objeto de estudio.

MARCO CONTEXTUAL

PAPEL TESTIMONIO DENTRO DEL CONFLICTO ARMADO COLOMBIANO

Al testimonio como género discursivo, se le ha asignado un papel de resistencia y denuncia dentro del contexto colombiano por representar un relato que posibilite dar voz a aquellos que no han podido dar a conocer cómo han afrontado y vivido los hechos de violencia que se han efectuado dentro de su entorno. Cabe señalar que a su vez tal como lo señala Vélez (2003) quien cita a Jelin (2003), el testimonio es un género mediante el cual se sistematiza una memoria autobiográfica y contribuye a la conformación de una memoria social. Con lo cual se puede decir que el testimonio en cierto sentido es un ejercicio de memoria personal que conlleva a la reconstrucción de un relato colectivo y por ende lleva a constituir memoria social que es aquella que todos los ciudadanos de una comunidad reconocen como propia y al enmarcar esta definición con los procesos de memoria, verdad y reparación, se reconoce el testimonio como una herramienta para el esclarecimiento de los hechos lo cual contribuye a avanzar en el camino hacia el post conflicto en el territorio colombiano.

El testimonio es entonces una unidad de análisis donde el narrador documenta el hecho de violencia y proporciona datos para aclarecer el hecho al presentar quién le hizo qué a quién, cuándo, dónde y cómo, donde se presentan los responsables del hecho y los efectos que esto ha tenido en la victima y familiares. Presenta los siguientes elementos que permiten integrar las distintas dimensiones que se requieren para conocer las modalidades de violencia y el perfil de aquellos que fueron afectados:

- Dónde y cuándo: el relator debe describir la fecha de ocurrencia del hecho y el dónde mediante el cual se define el lugar donde ocurrió.
- Cómo: todo lo relacionado con las circunstancias que caracterizaron el hecho, incluyendo los efectos que esto presentó en las personas implicadas.

- Quién lo hace: referencia a quién fue el perpetuador del hecho, desde señalar la identificación general, es decir si pertenece a un grupo armado o si existe una responsabilidad compartida.
- A quién se lo hace: caracterización de la persona o personas sobre las que se infligió el hecho de violencia, además es de importancia conocer el grado de vulnerabilidad en relación con el hecho.

METODOLOGÍA

Esta investigación se enmarca en el enfoque cualitativo. Dada la definición presentada por Baptista, Collado y Sampieri (2010), donde indican que el enfoque cualitativo emplea métodos de recolección no estandarizados y no presenta una medición numérica sobre el tema a investigar. De ahí que esta investigación esté dirigida a profundizar sobre aspectos cualitativos del corpus de testimonios que conforman la población objeto de trabajo, con el propósito de identificar características, atributos o facultades no cuantificables que permitan reconocer las metodologías, técnicas y herramientas de representación de conocimiento para automatizar procesos de análisis textual de los testimonios del conflicto armado colombiano. Es por esto por lo que la unidad de análisis es el corpus seleccionado compuesto por 30 testimonios los cuales responden a las características de representatividad, variedad y equilibrio.

Para llevar a cabo los objetivos propuestos en esta investigación, se utiliza la técnica de investigación documental entendiéndose esta como el procedimiento orientado a la aproximación, procesamiento y recuperación de información contenida en documentos, independientemente del soporte documental en que se hallen. Para ello, se utilizaron los siguientes instrumentos: *ficha bibliográfica, ficha de vaciado y matriz analítica de contenido* debido a que estas permiten la adecuada recolección y análisis de la información, que dé cuenta del proceso investigativo de la bibliografía y que favorezca la interpretación de los datos, así como las conclusiones a partir de los mismos.

Además, se desarrollaron actividades que contribuyeron al plan metodológico, las cuales se detallan a continuación:

Selección y procesamiento corpus de testimonios

Se realiza la identificación de fuentes y recuperación de testimonios del conflicto armado. En total se compilan 30 documentos de las siguientes fuentes de información que presentan fondos con colecciones especializadas en temas relacionados con conflicto armado en Colombia:

- Archivo Virtual de los Derechos Humanos y Memoria Histórica del Centro Nacional de Memoria Histórica.
- Banco de testimonios del Centro de documentación Museo Casa de la Memoria.
- Fuentes personales: testimonios producto de investigaciones realizadas por la Escuela Interamericana de Bibliotecología.

En cuanto a los criterios de selección de los testimonios se presentan las siguientes características:

- Narración en primera persona, donde se narren hechos violentos, por víctimas y sobrevivientes del conflicto.
- Hechos ocurridos en diferentes áreas geográficas del país, principalmente en regiones afectadas por el conflicto armado.
- Archivos digitalizados con reconocimiento óptico de caracteres. Esto con el fin de hacer el procesamiento en las herramientas informáticas seleccionadas.

Dado que los documentos recuperados se encuentran en formatos de texto tipo PDF o Word, se realiza el procesamiento de los mismos, el cual consistió en la trasformación del texto original a formato texto plano donde se busca que el documento se encuentre sin formato (negrilla, cursiva), sin imágenes y se encuentre en una codificación UNICODE, el cual es un código diseñado para facilitar el tratamiento informático, transmisión y visualización de textos de múltiples lenguajes compatible con el formato ISO. Este formato garantiza el cumplimiento de los criterios de: universalidad, uniformidad y unicidad. De manera que se garantiza un mejor reconocimiento de caracteres, lo cual facilita la recuperación y la minimización

de ruido o de términos no digitalizados que puedan no ser reconocidos por los programas de análisis textual.

Análisis formal y de contenido del corpus

Se identifican y etiquetan elementos que permitan estructurar la información de los testimonios, con el fin de reconocer patrones o características que se identifiquen en las herramientas de análisis textual para el análisis de testimonios. Se etiquetan los siguientes elementos que son de interés del corpus:

- **Descripción formal:** se describen elementos formales del documento con el fin de facilitar la recuperación del testimonio. Nombre, fecha, autor, lugar de ubicación (fuente), temática y lugar donde se desarrollan los hechos.
- **Descripción de contenido:** se realiza un breve resumen de la información contenida en el testimonio, se identifican género y características físicas del relator (género, parentesco con la víctima en caso de relatar una desaparición o muerte de alguien más, entre otros hechos).

Rastreo software para el análisis textual

Se realiza un rastreo bibliográfico sobre los sistemas de análisis textual para el tratamiento de corpus disponibles online para el uso al público o por descarga de programa ejecutable. Allí se identifican programas de extracción de concordancias, analizadores de sentimientos, extractores de palabras claves, entre otros, que permiten extraer listados de vocabulario e identificar grupos de palabras y comparar la frecuencia del uso, además, de realizar no solo análisis cualitativo sino cuantitativo.

Evaluación y comparación software para el análisis textual

Se evalúan los diferentes programas recuperados para el análisis textual, además se analizan los resultados presentados en los testimonios con el fin de conocer si estos permiten responder a las preguntas quién, cómo, cuándo y dónde. Para llevar a cabo esta evaluación y comparación se utilizan las herramientas tales como el listado de frecuencia de palabras, listado de palabras vacías, nombre de entidades, palabras claves, entre otras.

Construcción matriz para el análisis de resultados

Se lleva a cabo la construcción de una matriz donde se identifican en cada uno de los testimonios los resultados obtenidos en el análisis de los softwares para el análisis textual, en esta se describen los siguientes elementos: actores armados, testigos, hechos, lugares y fechas con el objetivo de conocer si estos responden el sentido y propósito que el testimonio quería señalar de manera que logre responder a las necesidades de representación.

RESULTADOS

En el desarrollo del presente proyecto a partir de la metodología definida, se alcanzaron los siguientes resultados en las actividades establecidas:

CORPUS

Se compilan 30 testimonios los cuales constituyen cerca de 77.000 palabras. Estos reúnen 24 temáticas diferentes de hechos de violencia en 10 departamentos del país. A continuación, se presenta la distribución temática y territorial:

TEMÁTICA - HECHOS DE VIOLENCIA
Desaparición
Asesinato
Despojo de tierras
Abuso de Poder
Masacre
Retención ilegal
Reclamación de tierras
Extorsión
Cultivos ilícitos
Delincuencia común
Secuestro
Desmovilización
Negociaciones Estado-FARC
Posconflicto
Conflicto entre grupos armados
Minas antipersonas
Tortura
Violencia de género
Ataque a la comunidad
Milicias urbanas
Control de territorio
Combates urbanos
Reinserción
Reclutamiento de menores

Tabla 1. Listado de temáticas en testimonios corpus

DEPARTAMENTO	MUNICIPIOS
Sucre	Sincelejo
	Corozal
	Palmitos
	Ovejas
	Los Palmitos
Bolívar	El Carmen de Bolívar
	María la Baja
Antioquia	Puerto Berrío
	Caracolí
	Medellín
	Sonsón
	Nariño
	Urabá
Vichada	Puerto Príncipe
Cesar	Pelaya
Meta	Villavicencio
	San Martín
Santander	Barrancabermeja
Cauca	
Caquetá	
Putumayo	Puerto Asís

Tabla 2. Distribución territorial corpus

Por otra parte, dentro de la caracterización del corpus se identifican los siguientes elementos en la matriz:

- **Descripción formal:** elementos formales para la descripción que facilitan la posterior recuperación del testimonio: nombre, fecha, autor, lugar de ubicación (fuente), temática y lugar donde se desarrollan los hechos.
- **Descripción de contenido:** breve resumen de la información contenida en el testimonio, se identifican género y características sociodemográficas del relator (género, parentesco con la víctima en caso de relatar una desaparición, muerte, entre otros hechos).

- **Características morfológicas y sintácticas**: identificación de la clase o categoría gramatical de las palabras claves dentro del contexto de la narración. Utilizando el análisis textual como herramienta de recuperación de información es importante reconocer las categorías de aquellas palabras claves en el texto. Todo esto con el fin de evaluar si las herramientas de análisis textual reconocen elementos que lleven a la representación del contenido de los testimonios.

PROGRAMAS DE ANÁLISIS TEXTUAL

Como resultado de las búsquedas se identifican programas de extracción de concordancias, analizadores de sentimientos, extractores de palabras claves, entre otros, que permiten extraer listados de vocabulario e identificar grupos de palabras y comparar la frecuencia del uso, además, de realizar no solo análisis cualitativo sino cuantitativo. A continuación, se presentan los programas identificados:

NVivo[2]

Es un software para la investigación de análisis textual el cual cuenta con métodos mixtos. Creado por empresa QRS, la cual se dedica al diseño de sistemas para el análisis de datos cualitativos. Los servicios que prestan están diseñados para organizar, analizar y encontrar perspectivas en datos no estructurados o cualitativos, como: entrevistas, respuestas de encuestas con preguntas abiertas, artículos, contenido de las redes sociales y la web.

Además, tiene tres versiones para sistema operativo Windows: NVivo 11 Starter, NVivo 11 Pro, NVivo 11 Plus. El valor comercial de estas versiones se encuentra 1599.00 USD, con el cual se adquiere una licencia perpetua del programa.

[2] Ver información del producto en https://www.qsrinternational.com/nvivo/products

Cabe destacar que entre los productos que presenta este software es de interés para el análisis del corpus herramientas tales como: consulta de frecuencia de palabras, en diferentes presentaciones (listado en tabla, marca de nube, mapa codificado).

T-LAB plus[3]

Es un software que se encuentra en la versión 4.0.2.6. Es estable y cuenta con respaldo técnico. Presenta versiones para el idioma inglés, italiano, español y francés. Entre sus características principales se encuentra que presenta un conjunto integral de herramientas lingüísticas, estadísticas y gráficas para el análisis textual; además permite el análisis de diversos tipos de textos: transcripciones de discursos, artículos periodísticos, respuestas a preguntas abiertas, mensajes de Twitter, transcripciones de entrevistas y grupos focales, textos legislativos, patentes, documentos de empresas, libros, etc. Por otro lado, admite documentos o datos tabulares en los siguientes formatos: .TXT, DOC, .docx, PDF, .RTF, .HTML, XLS, .XLSX, .CSV, MDB, ACCDB.

Entre las funciones que presenta este programa se encuentran la segmentación de texto, lematización automática o derivación, detección de múltiples palabras y palabras clave, selección de términos clave. Así como el análisis de coocurrencia para medir, explorar y mapear varios tipos de relaciones entre términos clave, ya sea en pares o en grupos, es decir, dentro de todo el corpus o dentro de sus subconjuntos.

El valor comercial de este software es de €1550 para la licencia perpetua con derecho a asistencia técnica.

[3] Ver información del producto en https://tlab.it/es/presentation.php

Linguakit[4]

Es una página web multilingüe, que surge fruto de años de investigación de la empresa Cilenis Language Technology en el campo del Procesamiento del Lenguaje Natural (PLN). Esta plataforma presenta cuatro módulos lingüísticos orientativos: el primero atiende a aspectos más genéricos del lenguaje con módulos como conjugador o traductor; el segundo, para un perfil de usuario más ligado al ámbito educativo, con módulos como el etiquetador morfosintáctico o el analizador sintáctico; el tercer módulo pensado para profesionales de la comunicación y marketing como el analizador de sentimientos o el extractor de palabras clave; y por último, un módulo experimental donde se presentan las nuevas herramientas en proyecto.

Este recurso se encuentra disponible al público y puede ser utilizado sin necesidad de suscripción para realizar 20 consultas diarias a textos no mayores de 5000 palabras. En la página web se encuentra un enlace de planes, pero no se encuentran disponibles.

Stilus[5]

Es una página web que presenta diferentes módulos lingüísticos, entre los cuales se encuentra el conjugador verbal y analizador morfosintáctico. Además, cuenta con cuatro versiones: *usuario no registrado*, donde el usuario puede revisar 5000 palabras/mes totalmente gratuito, con la restricción de encontrarse registrado el usuario; usuario *Stilus gold*, puede revisar 200000 palabras/mes por un pago anual de 29,95 €; *Stilus Platinum* puede revisar 1000000 palabras/mes por un pago anual de 59,95 €; y *Stilus Sprint* pueden revisar 1000000 palabras/mes con un valor de 29,95 €. Luego de tres meses de inactividad la cuenta caduca.

[4] Ver información del producto en https://linguakit.com/es/analisis-completo
5 Ver información del producto en http: http://www.mystilus.com/Pagina_de_inicio

Entre sus características principales se encuentra que en la última versión este puede integrar nuevas herramientas como son extracción de palabras y análisis de sentimientos.

AntConc[6]

Es un programa de ayuda para el análisis de corpus. Desarrollado por Laurence Anthony en la Universidad Okayama en Japón. Se encuentra en la versión para Windows 3.2.7 actualmente. Entre las funciones que presenta es la de extraer datos de amplios corpus textuales, entre los cuales se encuentran: frecuencias de palabras, colocaciones, concordancias, búsquedas mediante expresiones regulares. Es de descarga gratuita y liviano por lo que no ocupa mucha cantidad de espacio en la memoria RAM, además, funciona en los programas Windows, Mac y Linux.

EVALUACIÓN SOFTWARE –EJECUCIÓN

Al evaluar los diferentes programas de análisis textual y las funciones que estos presentan, se identifican algunas funciones que permiten recuperar y representar el contenido del testimonio y de esta manera responder a las preguntas quién, cómo, cuándo y dónde. Es por esto por lo que inicialmente se presentan una breve descripción de las funciones que permiten llevar a representar el contenido de los testimonios y luego se presenta la evaluación de cada de los softwares anteriormente descritos de manera que se reconozcan las fortalezas y oportunidades de cada uno. A continuación, se presentan los ítems que se evaluaron:

6 Ver información del producto en http://www.laurenceanthony.net/software.html

Frecuencia de palabras

En esta función se analiza el texto procesado y se extraen las palabras que presentan mayor ocurrencia dentro del recurso. Esta información, de acuerdo con las opciones del programa, se visualiza como mapa ramificado, nube, descarga del listado en XML o TXT. Los resultados de esta consulta permiten conocer la ocurrencia de palabras e identificar cuales palabras se repiten en mayor medida y a partir del análisis de los datos reconocer la temática central del testimonio como también los actores y hechos.

Listado de palabras vacías

Es un recurso que presentan los programas de análisis textual y se utiliza para retirar de los listados de frecuencia aquellas palabras que no presentan un valor semántico en el contenido del documento dado que dentro del relato son palabras que cumplen una función de conectores o que por sí solas, fuera del contexto no permiten conocer la participación pragmática dentro del texto.

Palabras claves

Se realiza de manera automática la identificación de las palabras claves del documento, el cual de acuerdo con el análisis que efectúa el programa, identifica las palabras más importantes de un texto y las clasifica según su grado de relevancia. Esta herramienta es clave en la detección inmediata del tema, lo que puede permitir el análisis de testimonios, el etiquetado automático y la clasificación documental.

Hallazgos ejecución programas

NVivo

Para la ejecución de este programa, debido a que es por suscripción y pago, se descarga la versión *NVivo 11 Pro-demo* la cual se puede utilizar por un periodo de catorce días. Al explorar el programa se encuentran, entre otras, las siguientes herramientas que se privilegian por la posibilidad de análisis exclusivamente textuales:

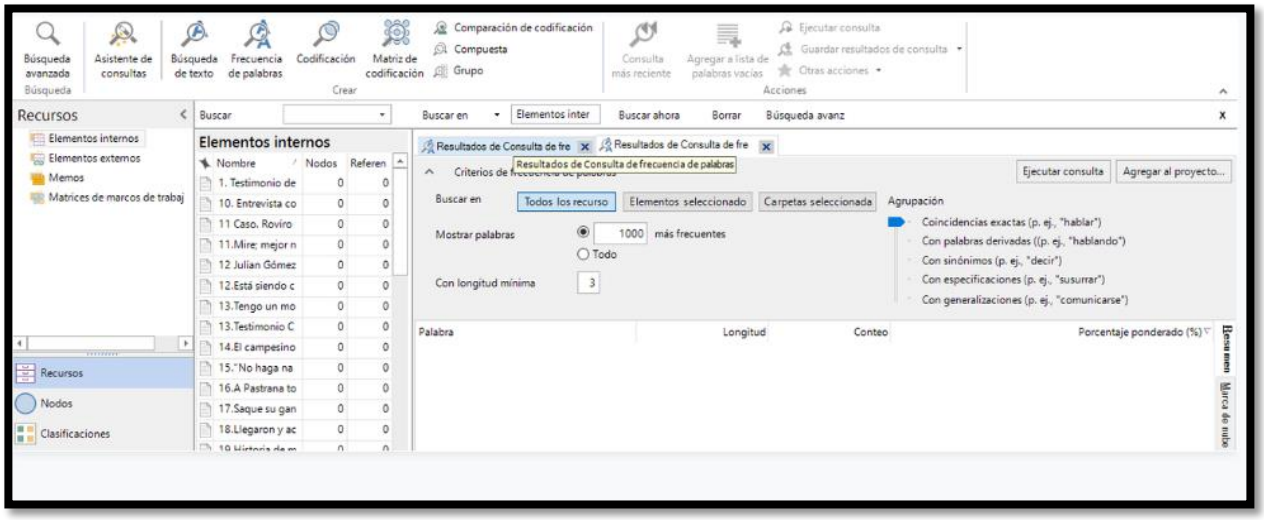

Gráfico 1 Visualización NVivo 11 Pro-demo

Frecuencia de palabras

Este programa permite la extracción de cinco diferentes tipos de listado de frecuencia de palabras, donde lo que se busca es refinar los resultados de acuerdo con las siguientes características:

- Coincidencia exacta: recupera el listado de términos con la palabra exacta, tal como se encuentra en el documento, es decir, que no tiene en cuenta la lematización.
- Palabras derivadas: recoge el listado de palabras a partir de la raíz y relaciona estas palabras de la primera búsqueda con palabras derivadas, tal es el caso de los verbos conjugados, o las variantes de número.
- Palabras con sinónimos: relaciona y agrupa las palabras que presentan sinónimos dentro del texto.
- Palabras con especificaciones: relaciona y agrupa las palabras que presentan términos específicos. Es decir, que derivan una acción específica, tal como se presentan en los verbos.
- Palabras con generalizaciones: relaciona y agrupa las palabras que entre ellas generalizan una acción o nombre.

Listado de palabras vacías

Este programa presenta una opción de lista palabras vacías la cual se va construyendo con las palabras que se etiquetan dentro de esta categoría. Entre las ventajas que presenta este recurso se da el hecho de que a medida que se avanza en el análisis estas palabras dejan de presentarse en los resultados, así se analice un documento diferente en el que se etiquetó.

Hallazgos

Aspectos generales por resaltar

Este es un programa enfocado en el análisis textual mixto, que presenta diferentes opciones de análisis del corpus. Es intuitivo y la interfaz es fácil de manejar. Es un programa de pago y de acuerdo con la edición y las posibilidades de análisis, sin embargo, las últimas versiones cuentan con la posibilidad de analizar textos

provenientes de redes sociales lo que lo hace muy versátil en cuanto a la estructura de la fuente de información.

Fortalezas

- Permite realizar el análisis de la información de manera rápida y depura a medida que avanzan los análisis las palabras vacías identificadas de acuerdo con el contexto del documento.
- Es compatible con diferentes tipos de formato de archivo (doc., .XLM, .TXT, .PDF, entre otros), lo cual facilita el uso dado que no exige que se realice la codificación de los archivos a formatos compatibles.
- Permite la descarga e impresión de los resultados en diferentes formatos entre los que se encuentra el formato Excel el cual tiene una gran utilidad por la facilidad de cuantificar los resultados de los listados de frecuencia.
- Permite la visualización de la información de frecuencia de palabras en marca de nube y mapa ramificado, lo cual facilita la visualización de la información.
- No presenta restricción en el número de palabras y caracteres a analizar, lo que facilita el uso dado que no se fragmenta la información.

Desventajas

- No identifica nombres propios de lugares o personas conformados por dos palabras, es por esto que dificulta reconocer la estructura nombre, apellido o en el caso del nombre de lugares que presentan algunos testimonios.
- Presenta oportunidad en la agrupación de las palabras de acuerdo con las categorías (palabras derivadas, con generalizaciones, con especificaciones). Dado que algunas palabras no se agrupan adecuadamente cuando se selecciona la opción dando como resultados no se asocia palabras que presentan una relaciones gramaticales y semánticas.

T-LAB plus

Se utiliza la versión demo del programa, la cual no tiene límite de tiempo, pero solo permite el análisis simultáneo de tres textos que no superen los 20KB.

Para esta aplicación del análisis de testimonios se utilizan las herramientas de léxico y la lista de palabras clave automática, graficas de frecuencia y coocurrencias.

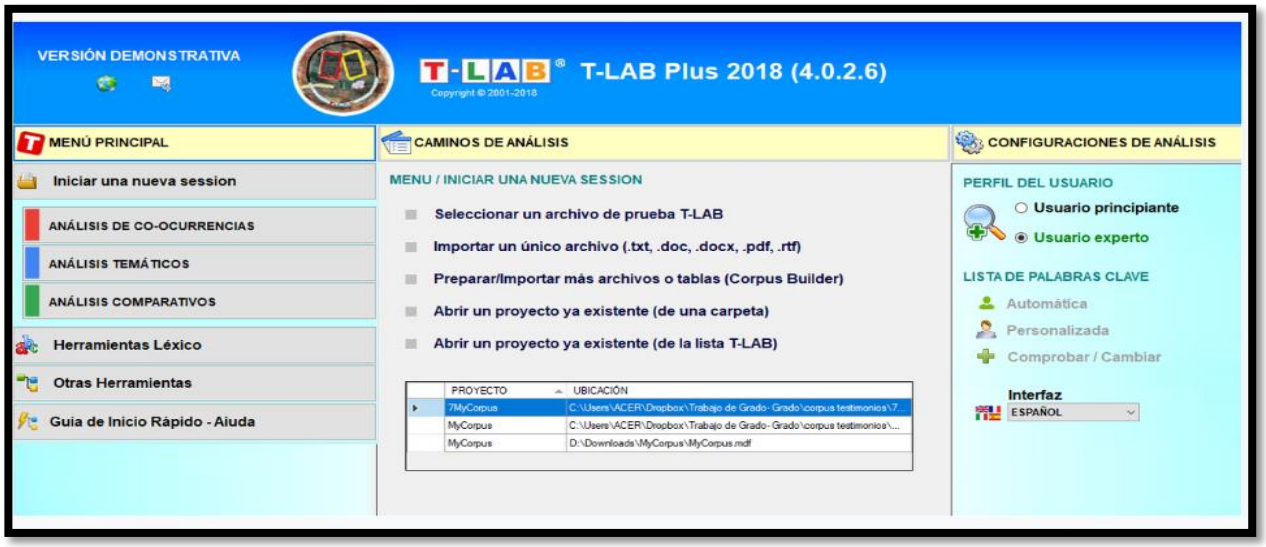

Gráfico 2. Visualización programa T-LAB plus

Frecuencia de palabras

Permite la realización del listado de frecuencia de palabras y de manera automática retira de los resultados de búsqueda las palabras vacías.

Palabras claves

El programa de forma automática identifica las palabras claves y grafica las relaciones que estas palabras presentan entre ellas.

Hallazgos

Aspectos generales por resaltar

Es un programa que presenta un conjunto de herramientas lingüísticas, estadísticas y gráficas para el análisis de los textos. Es de pago. Es de fácil descarga y manejo, dado que presenta una interfaz intuitiva y amigable.

Fortalezas

- Permite mapear las relaciones de coocurrencia entre palabras clave. En estas gráficas se relacionan palabras que acompañan a otras dentro del texto. Esta característica permite visualizar nombre de personas y lugares.
- Descarga de resultados de búsqueda en Excel, facilitando de esta manera el manejo de los datos.
- Es compatible con diferentes formatos, lo cual permite el tratamiento en múltiples formatos de acuerdo con las necesidades específicas.
- Realiza de manera automática el listado de palabras vacías.

Desventajas

Dado que se trabajó en la versión demo, presenta restricción en el análisis de la información ya que tiene en cuenta el número de archivos procesados de acuerdo con el peso.

Linguakit

Se utiliza la versión web y los módulos de análisis lingüístico y analítica textual. En el módulo análisis lingüístico se realiza el análisis de frecuencia de palabras donde se analizan las palabras de cada uno de los testimonios según su frecuencia y tipo:

verbos, sustantivos, adjetivos, además los resultados proporcionan datos genéricos como: el número de frases, de palabras, de lemas y de caracteres con y sin espacios en blanco. Así mismo, en el módulo de analítica textual se utiliza la herramienta palabras claves o 'keyword' en la cual la aplicación determina las palabras semánticamente más relevantes. Además, se utiliza otra herramienta del módulo reconocimiento de entidades donde la aplicación identifica personas, los lugares, organismos, fechas y cantidades, es decir, las entidades que se mencionan.



Gráfico 3. Visualización programa LInguakit

Hallazgos

Aspectos generales por resaltar

Es una aplicación web completa que permite identificar diferentes elementos en los textos procesados, la utilización de este es sin costo, dado que, al revisar los planes disponibles, estos aun no presentan un valor comercial.

Fortalezas

- Presenta un completo conjunto de herramientas para el análisis textual, el cual busca responder a las diferentes necesidades de información.
- Cuenta con tres formas de ingreso de la información, en cuadro de texto, anclado de archivo o página web.
- Presenta un identificador de entidades, donde el programa extrae de un texto las entidades, nombres propios y palabras o datos numéricos relevantes, que clasifica como personas, lugares, organizaciones, organismos, fechas, cantidades. Esta es una herramienta que no se encuentra en otros programas y le proporciona un elemento diferenciador.
- Al recuperar las palabras clave permite recuperar no solo en un listado, sino que esta se encuentra resaltado dentro del contexto del texto.

Desventajas

- Presenta un límite de consultas por día en usuario de 20 consultas lo cual dificulta procesar el corpus.
- Tiene una restricción en el número de palabras máximo para analizar, el cual es de 5000 por consulta, sin embargo, este valor tiene en cuenta los espacios lo cual disminuye el número de palabras a analizar. Por tanto, en los casos cuando se requiere analizar textos con un número mayor de 5000 caracteres se debe fragmentar el texto lo cual ocasiona que la información no responda a la representación del texto completo.

Stilus

Se utiliza la versión web, la cual cuenta con herramientas para corrección de ortografía, gramatical y de estilo. Esta aplicación no cuenta con herramientas para extraer frecuencia de palabras y palabras claves. Las herramientas que presenta son corrección de estilo, conjugador verbal y analizador morfosintáctico.



Gráfico 4.Visuliazación programa Stilus

Aspectos generales por resaltar

Fortalezas

El análisis morfosintáctico permite identificar categorías gramaticales y de forma indirecta identificar actores del conflicto narrados en los testimonios que conforman el corpus.

Desventajas

Presenta restricción en las palabras por analizar, solo permite el análisis de 100 palabras lo que ocasiona que se deba fraccionar demasiado la información.

AntConc

Es un software independiente para el análisis lingüístico de textos, presenta una versión libre disponible para Windows, Mac OS y Linux. Esta permite trabajar en documentos en formato PDF o Word.

Aspectos generales por resaltar

Fortalezas

- Cuenta con herramientas para el análisis de frecuencia de palabras y palabras claves.
- Aunque el programa se encuentra en idioma inglés, permite el análisis de textos en idioma español.
- Permite ingresar listado de palabras vacías para que no tengan en cuenta el analizar frecuencia de palabras.

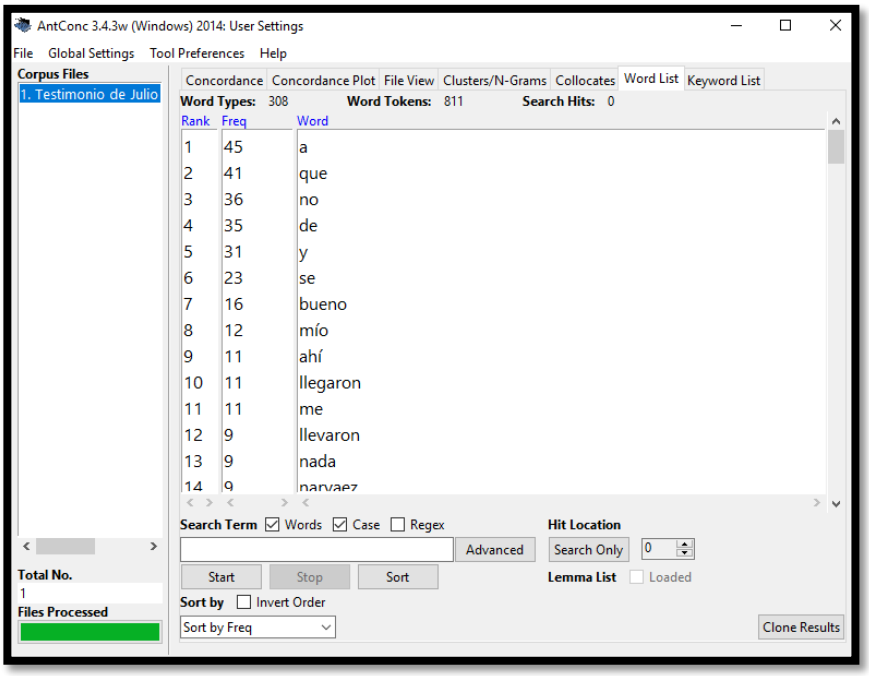

Gráfico 5. Visualización programa Antconc

Desventajas

- A pesar de contar con un listado de palabras vacías, el programa no logra identificarlas todas, lo cual ocasiona que en los resultados de frecuencia de palabras se recuperen palabras que no cumplen con los criterios, afectando los resultados.
- El listado de palabras claves que extrae del texto, aunque se genera automáticamente, se debe ingresar un conjunto de palabras claves como referencia para el programa identifique si estas se encuentran en él. Por tanto, esta herramienta no es de utilidad si se desea utilizar para reconocer la temática del testimonio.

EVALUACIÓN SOFTWARE – COMPARACIÓN RESULTADOS EN ANÁLISIS DEL CORPUS

A partir del análisis realizado al comparar los resultados de los tres principales programas en los cuales se obtuvieron mejores resultados se destacan los siguientes elementos:

Nube de palabras:

Esta aplicación la tienen los programas Nvivo y Linguakit, en los resultados se puede apreciar que al no contar el programa de Linguakit con una herramienta para retirar las palabras vacías al momento de la creación de la nube de palabras frecuentes se observa la presencia de artículos y conectores que al momento de visualizar la representación del contenido no permite distinguir de manera visual elementos repetitivos o concurrentes en el testimonio relevantes en el texto. Sin embargo, al realizar la nube de palabras del resultado de las palabras reconocidas como entidades se aprecia una diferencia en la representación de la información relevante el testimonio.



Grafico 6.Nube de palabras frecuentes Linguakit. Testimonio 5



Grafico 7.Visulización reconocedor de entidades Linguakit.
Testimonio 5



Gráfico 8.Visualización palabras frecuentes programa NVivo. Testimonio 5

Palabras clave:

La identificación automática de palabras claves la presentan los programas Linguakit y T-LAB plus. Al revisar los resultados en un mismo testimonio se identifican diferencias en la determinación del listado de estas palabras.

Al revisar la forma como procesan la información cada uno de los programas se encuentra que el programa T-Lab extrae las palabras a partir de la predisposición de una o más listas de unidades lexicales (palabras, lemas o categorías) a utilizar para crear las tablas de datos a analizar. Es decir que tantas palabras se asocian a estos términos. Por otro lado, el programa Linguakit extrae las palabras más importantes del texto y las clasifica según su grado de relevancia teniendo en cuenta las etiquetas sobre la función de la palabra en el contexto del texto. Es por esto que representa la información de manera al presentar T-Lab, las palabras y el conteo de ellas y Linguakit etiqueta cada palabra de acuerdo a categorías establecidas dado el contexto del texto (localidad, persona, nombre, organización, verbo) que a su vez están relacionadas con las entidades.

Palabra	Conteo
compañero	5
José	5
Tienda	5
Desplazar	4
Gabriel	4

Tabla 3. Listado de palabras clave programa T-LAB. Testimonio 5

Palabra	Etiqueta
Santa Elena	localidad
Padua	localidad
José Gabriel	persona

Sincelejo	localidad
Piñal	localidad
ñame	nombre
machetero	nombre
El Piñal	persona
TESTIMONIO	organización
José Gabriel Jiménez Gómez	persona
Roberto Pérez	persona
Kiosco	organización
José NN	persona
EFECTOS	organización
Fiscalía de Corozal	organización
Tristeza	persona
Margoth	persona
Robiro Causado	persona
pertenecía	verbo

Tabla 4. Listado palabras clave programa Linguakit. Testimonio 5

Vista resaltada palabras clave

Es una función que presentan los programas T-Lab y Linguakit, la cual permite reconocer la asociación de las palabras claves en el contexto del texto. Al comparar los resultados se observa que el programa T-Lab, a diferencia de Linguakit, divide el texto en fragmentos claves y a partir de ellos resalta la palabra; en cambio en el programa Linguakit presenta todo el texto y resalta aquellas palabras claves lo cual al momento de leer lleva al lector a reconocer qué elementos del texto presentan mayor relevancia.

TESTIMONIO CASO: JOSÉ GABRIEL JIMÉNEZ GÓMEZ -Presentado por la compañera Omaira, quien se encuentra desplazada viviendo en el barrio de invasión en Sincelejo. José Gabriel Jiménez Gómez, 31 años de edad, vivía en unión libre, padre de dos hijos. El trabajaba en la Acción Comunal, era secretario de la Junta, allá en Santa Elena, habia comités. El trabajaba en una tienda comunitaria, en ese tiempo se consiguió la luz y otras cosas. Era agricultor, no pertenecia a ninguna organización politica. • HECHO Eso sucedió el 25 de noviembre de 1996, a las 11 :30 pm en el Piñal, él se encontraba en la casa de una prima, Margoth. De ahí lo sacaron. EFECTOS: Tristeza y pasa uno mucho trabajo, los hijos principalmente, todavía no le hemos conseguido nada, ni siguiera colegio. Nosotros después nos tuvimos que desplazar hacia Sincelejo como a los 15 días, llegamos donde una prima de mi papá que vive acá Nos ha tocado pedir ayuda para nuestros hijos, unas veces nos dan y otras no. El hecho fue denunciado en la Fiscalía de Corozal. No he vuelto a saber nada más. Antes de la muerte de José Gabriel, mataron a un sobrino de él, primo mío, Robinson Narvaez, a un compañero de nosotros, Robiro Causado y otro señor en Santa Rosa llamado José NN, no recuerdo el apellido. Nosotros allá estacamos marchando bien, porque trabajabamos, teníamos

Gráfico 9. Vista resaltada palabras clave en contexto. Programa Linguakit. Testimonio 5

TESTIMONIO CASO: JOSÉ GABRIEL JIMÉNEZ GÓMEZ - Presentado por la compañera Omaira , quien se encuentra desplazada viviendo en el barrio de invasión en Sincelejo .

****

José Gabriel Jiménez Gómez, 31 años de edad , vivía en unión libre , padre de dos hijos .

****

El trabajaba en la Acción Comunal , era secretario de la Junta , allá en Santa Elena , habia comités . El trabajaba en una tienda comunitaria , en ese tiempo se consiguió la luz y otras cosas . Era agricultor , no pertenecia a ninguna organización politica .

****

• HECHO Eso sucedió el 25 de noviembre de 1996 , a las 11 : 30 pm en el Piñal , él se encontraba en la casa de una prima , Margoth . De ahí lo sacaron .

****

EFECTOS : Tristeza y pasa uno mucho trabajo , los hijos principalmente , todavía no le hemos conseguido nada , ni siguiera colegio . Nosotros después nos tuvimos que desplazar hacia Sincelejo como a los 15 días , llegamos donde una prima de mi papá que vive acá Nos ha tocado pedir ayuda para nuestros hijos , unas veces nos dan y otras no .

****

Gráfico 10. Vista resaltada palabras clave en contexto. Programa T-lab. Testimonio 5

CATEGORIZACIÓN DE ELEMENTOS (LUGARES, ACTORES, ACCIONES, FECHAS, TESTIGOS) EN CORPUS A PARTIR DE ANÁLISIS TEXTUAL

Los resultados de las herramientas palabras clave, reconocimiento de entidades, frecuencia de palabras y palabras vacías del análisis textual automatizado aplicado en los testimonios que conforman el corpus se valoran las palabras o frases recuperadas y se categorizan dentro de los ítems señalados para dar respuesta a las preguntas ¿Quién? ¿Cómo? ¿Cuándo? ¿Dónde? la cual consiste en identificar aquellas palabras que responden a esos interrogantes teniendo en cuenta la frecuencia y el etiquetado que el programa asigna a las palabras.

Al realizar esta categorización se identifican los siguientes hallazgos que permiten destacar el desarrollo de los programas al aplicar el análisis textual en testimonios del conflicto armado:

- La aplicación para análisis textual Linguakit presenta una estructura en los resultados que facilitan la identificación de los elementos, dado que cuenta con un etiquetado de las palabras que conllevan reconocer la función correspondiente de cada palabra dentro del contexto. Estas etiquetas son: localidad, nombre, fecha, cifra, verbo, organización, adverbios, adjetivos entre otros. Estas facilitan filtrar la información por categorías y depurar los datos. Sin embargo, a pesar de ser automático se requiere de la inversión de trabajo manual por parte del investigador para reconocer estos elementos, dado que la estructura de la representación de la información por sí sola a pesar de ser automática no responde a los interrogantes planteados.

- Dentro del análisis realizado a los testimonios se encuentra de manera repetitiva la aparición de varias ocurrencias por parte de palabras en el texto que no son una palabra clave o que llegan a evidenciar de qué tema se está tratando. Se ha encontrado que las palabras que aportan un sentido para

responder a las preguntas de las categorías de interés, presentan poca frecuencia en algunos casos y que se debe, quizás, a la intención del narrador donde no señala el nombre completo de la víctima o del victimario y recurren a la función o posición jerárquica que presentan dentro de la familia o grupo para nombrarlo, lo cual lleva a que se requiera contar con la interpretación por parte del investigador, dado que la aparición de varias ocurrencias por parte de una palabra en el texto no señalan precisamente que sea una palabra clave o que se trate de la temática; o al contrario, sí presenta un sentido dado que se debe a una intención del narrador. Esto se debe a que listados de palabras frecuentes presentan una gran cantidad de ruido debido a que en muchas oportunidades las palabras con mayor ocurrencias son aquellas que el narrador utiliza con una función ya sea de conector de conjugación o una muletilla ya sea porque no desea señalar de manera puntual la acción o no indicar cual o cuales son las personas que participaron en la ejecución del hecho de violencia dada la intimidación o repercusiones que puedan sucederle por dar el testimonio del hecho violento que afrontaron.

- Se evidencia en los testimonios que el narrador utiliza pronombres para señalar los actores del suceso. Estas palabras en algunas oportunidades el software las cataloga como vacías de manera automática, ya sea porque se filtró por longitud de palabras o porque el software presenta una función como los programas: Antconc, T-Lab, Nvivo, de generar un listado de palabras vacías de forma acumulativa donde el corpus a medida que es analizado va integrando palabras catalogadas como vacías sin tener en cuenta la intención del narrador. Lo que lleva a posibles errores de interpretación dado que una palabra puede cambiar su intencionalidad dependiendo del contexto donde se presenta. Lo que lleva a la tarea manual de explorar cada palabra de acuerdo con el contexto. Es por esto por lo que contar con listado de

palabras vacías automático acumulativo puede llevar a errores o a dejar de lado palabras que dentro de un contexto particular pueden señalar un evento destacado o en otros casos no solo ser ruido.

- Al recuperar datos ya sean de fechas o cifras dentro en los testimonios a partir de los resultados del análisis textual de los softwares no son efectivos dado que estos presentan errores al momento de contextualizar la información. En el caso de las cifras no se reconocen a que se debe y pasan a ser un número vacío sin contexto que no se logra integrar a la información que se desea catalogar. Sin embargo, el programa Linguakit presenta mayor oportunidad y fortalezas en este campo dado que logra integrar la fecha completa y etiquetarla como tal. Por tanto, la fecha no se pierde, no obstante, las cifras pierden el sentido como dato de que no se pueden agrupar a una categoría determinada.

- Dada la forma como se encuentran conformados algunos nombres de personas y lugares en la que están conformadas por dos, tres o cuatro palabras dentro de corpus de testimonios, al recurrir a los resultados del listado de frecuencias de palabras se evidencian problemáticas al momento de reconocer esta información ya que toman cada palabra como una unidad y no realizan conexiones de conjunto palabras. El programa Linguakit, es el que reconoce el nombre de las personas de manera completa reconociendo la estructura de dos o más palabras como un conjunto, pero no lo realiza en el listado de frecuencias sino cuando se le asigna reconocer entidades.

CONCLUSIONES

- La creación de un corpus para la realización de esta investigación fue una labor llena de retos, dado que se debió explorar nuevas áreas del conocimiento e integrar nuevos conceptos que antes no se habían abordado, sin embargo, esta labor enriqueció los conocimientos acerca del área de organización y tratamiento de la información y visualizarla como un área transversal e interdisciplinaria.

- Los testimonios del conflicto armado en Colombia son un relato que permite reconocer cómo la población ha vivido y afrontado en primera persona los hechos de violencia. De ahí el interés de vincular estos registros en la construcción de la memoria social dado que a partir de la autobiográfica se puede llegar a consolidar una memoria que vincule a todos los actores del proceso de violencia en Colombia.

- La función social de la bibliotecología va más allá de su quehacer en el interior de una biblioteca. El buscar nuevos mecanismos para facilitar la recuperación y representación de los contenidos de los testimonios del conflicto armado, responde a la necesidad social de un país para reconstruir su memoria social y conocer la verdad de los hechos de violencia de los últimos 60 años. Lleva al área de tratamiento y organización de la información a buscar respuestas a nuevas necesidades de información por fuera del canon de las fuentes tradicionales. Lo cual permite reconocer que es un área evolutiva y que responde a las necesidades actuales de la sociedad donde se encuentra inmersa.

- La vinculación de otras áreas del conocimiento como es la Lingüística computacional en los procesos de tratamiento y organización de la información dentro del campo bibliotecológico enriquecen de nuevos

conceptos y perspectivas que llevan a explorar nuevos caminos que llevan al bibliotecólogo a plantearse nuevos retos para responder a las necesidades de información que la sociedad del conocimiento día a día requiere.

- Los programas de análisis textuales desarrollados en la actualidad presentan un avance significativo en la identificación de entidades y etiquetado de las palabras. Sin embargo, en la aplicación del análisis textual en los testimonios del conflicto armado se identifica que aún se requiere de mucho trabajo manual y conceptual por parte del investigador. Lo que sugiere que se debe continuar vinculando el área de la lingüística computacional para responder estas necesidades de información y de análisis del texto de este tipo de formatos.

RECOMENDACIONES

Se recomienda continuar con esta investigación dado que esta primera etapa exploratoria solo reconoció ciertos elementos del análisis textual. En una segunda fase se recomienda utilizar el análisis estadístico y evaluar patrones dentro del corpus de testimonios que permitan mejorar la recuperación y representación de estos dentro del campo del tratamiento y organización con miras a que se realice de manera automática y se logre avanzar de manera más ágil en los procesos de recolección de información, procesamiento de datos, análisis de la información, difusión y evaluación que se adelanta en el marco de la reconstrucción de los hechos de violencia que el país ha afrontado.

BIBLIOGRAFÍA

Aguirre, K. F. (2003). Análisis textual: generación y aplicaciones. *Metodología de Encuestas*, *5*(1), 55-66.

Algranti, J. B. (1993). Aproximación al análisis textual informatizado. Anuario de psicología/The UB Journal of psychology, (59), 61-78**.**

Díaz, V. C. (2012). Historias de mujeres. Testimonios de excombatientes del conflicto armado colombiano. *Tonosdigital: Revista electrónica de estudios filológicos*, (22), 1-23

Figuerola, C. G. (2005). Técnicas lingüísticas aplicadas a la búsqueda textual multilingüe: ambigüedad, variación terminológica y multilingüismo (A. Peñas Padilla). *Revista española de Documentación Científica*, *28*(2), 239-240.

Gómez, J. D. V. (2014). Memoria, historias de vida y papel de la escucha en la transformación subjetiva de víctimas/sobrevivientes del conflicto armado colombiano. *Ágora USB*, *14*(1), 37-60.

Guinovart, J. G. (1998). Fundamentos de Lingüística Computacional: bases teóricas, líneas de investigación y aplicaciones. *Bibliodoc: anuari de biblioteconomia, documentació i informació*, 135-146.

i Besora, M. V. (1993). Las disciplinas del discurso: hermenéutica, semiótica y análisis textual. Anuario de psicología/The UB Journal of psychology, (59), 19-60.

Lavid, J. (2005). *Lenguaje y nuevas tecnologías*. Cátedra.

Llisterri, J., & Moure, T. (1996). Lenguaje y nuevas tecnologías: el campo de la lingüística computacional. In *Avances en lingüística aplicada* (pp. 147-228). Universidade de Santiago de Compostela

Satriano, C. R. (2000). Importancia del análisis textual como herramienta para el análisis del discurso. *Cinta de Moebio. Revista de Epistemología de Ciencias Sociales*, (9).

Velásquez, J. M. B., & Girón, W. L. (2010). *Lingüística computacional aplicada*. Universidad Nacional de Colombia, Facultad de Ciencias Humanas, Departamento de Lingüística.

Vélez Rendón, J. C. (2003). Violencia, memoria y literatura testimonial en Colombia. Entre las memorias literales y las memorias ejemplares.

Verdejo Maillo, M. F., Rodríguez Hontoria, H., & Díaz de Ilarraza Sánchez, A. (1998). ITEM: recuperación de información textual en un entorno multilingüe con técnicas de lenguaje natural. *Procesamiento del lenguaje natural, nº 23 (septiembre 1998); p. 189*.

ANEXOS

CORPUS DE TESTIMONIOS

Testimonio 1

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Caso: Santa Elena Testimonio de: julio Rafael Narváez Gómez - padre de Robinson	Casa campesina Sincelejo	Archivo virtual de los derechos humanos y memoria histórica	1211	6510	5314	7 de septiembre de 1996	Municipio de Los Palmitos. Corregimiento de El Piñal. Sucre	Desaparición, asesinato	El día 7 de septiembre de 1996 llegaron 17 uniformados y 4 de civil a la vivienda ubicada en Vereda Santa Elena, del municipio de Los Palmitos. Corregimiento de El Piñal. Estos se llevaron a Robiro Causado Márquez y Robinson Manuel Narváez Arco; este último hijo de Julio Rafael Narváez Gómez. Testimonio de Julio Narváez Gómez.

Testimonio 2

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Caso: Félix Manuel baldes chamorro	Casa campesina Sincelejo	Archivo virtual de los derechos humanos y memoria histórica	6656	35034	28467	25 de febrero de 1990	Corregimiento de Macayepo, municipio de El Carmen de Bolívar	Desaparición, abuso de poder	Testimonio de la esposa del desaparecido Félix Manuel Baldes Chamorro, hombre de 42 años y padre de cuatro hijos. Narra los hechos acerca de la desaparición y los sucesos que acontecieron luego de este. Como fue el proceso de búsqueda y como este hecho afecto a la familia y a la comunidad.

Testimonio 3

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Caso: Jairo Antonio Chamorro Romero	Casa campesina Sincelejo	Archivo virtual de los derechos humanos y memoria histórica	569	3278	569	29 de Julio de 1986	Municipio de Corozal Sucre	Asesinato, despojo de tierras	Jairo Antonio Chamorro Romero, de 34 años de edad, dedicado a la agricultura, activista campesino de la ANUC. Es asesinado en conflictos por la tenencia de la tierra por sicarios contratados por el terrateniente de la Hacienda Don Alonso

Testimonio 4

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Caso: Edilberto Acosta Ortega	Casa campesina Sincelejo	Archivo virtual de los derechos humanos y memoria histórica	914	5135	4219	25 de noviembre de 1996	Municipio de Corozal Sucre	Asesinato, despojo de tierras	Compañera sentimental relata los hechos y efectos del asesinato de Edilberto Acosta Ortega el 25 de noviembre de 1996 a las 1 am por 7 personas vestidos de policía y 2 encapuchados que ingresaron a su vivienda.

Testimonio 5

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Caso José Gabriel Jiménez Gómez	Casa Campesina Sincelejo	Archivo Virtual de los Derechos Humanos y Memoria Histórica	469	2822	469	25 de noviembre de 1996	Piñal. Sucre	Desaparición, asesinato	Compañera sentimental relata los hechos acerca de José Gabriel Jiménez Gómez, hombre de 31 años desaparecido el 25 de noviembre de 1996

Testimonio 6

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Caso familia contreras triple homicidio.	Casa campesina Sincelejo	Archivo virtual de los derechos humanos y memoria histórica	1519	8606	7016	30 de agosto de 1991	Vereda El Desvelo, del municipio de Ovejas, (Sucre)	Asesinato, masacre	Relato acerca del triple homicidio presentado en la familia Contreras por parte de desconocidos. Las víctimas son: Julio Cesar Contreras Morales - 46 Años – Agricultor; Huber Contreras Hernández -22 Años-Agricultor Orlando José Contreras Morales -30 Años

Testimonio 7

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Navidad y terror	Dioselina Pérez / Programa de Atención a Víctimas del Conflicto Armado. Alcaldía de Medellín	Jamás olvidaré tu nombre	1827	9854	8224	25 de diciembre	Puerto Berrío, entre Virginia y Caracolí, a orillas del Río San José del Nus.	Retención ilegal, masacre	Dioselina Pérez relata la retención ilegal que su padre, Ángel Antonio, sufrió por parte de un grupo paramilitar donde se le acusa a él y a otras personas de ser colaboradores de la guerrilla, esta retención se da cuando se desplazaba. Los hechos generan los efectos en el estado de salud. Por otra parte durante la retención ingresan a varias fincas, entre las que se encuentra la de una señora y sus hijos los cuales debieron entregar los

									pocos recursos para el sustento, al mismo tiempo ellos fueron víctimas dado que fueron maltratados y amenazados por estos hombres armados.

Testimonio 8

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Uno no está ahí nada más porque "yo llegué y aquí estoy"	Observatorio de derechos humanos y derecho internacional humanitario de FUNDAGÁN	Acabar con el olvido. Segundo informe	1098	6178	5158	1989-1990	María la Baja. Bolívar	Despojo de tierras, reclamación de tierras	Se relata cómo se vivió la llega de grupos armados a la zona y como estos utilizaban el secuestro como una herramienta para adquirir recursos económicos. Además, los efectos del desplazamiento de personas de la zona y el desalojo de tierras.

Testimonio 9

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Bájense que aquí se va a morir alguno	Observatorio de derechos humanos y derecho internacional humanitario de FUNDAGÁN	Acabar con el olvido. Segundo informe	1659	8690	7150	1997	Pelaya. Cesar	Retención ilegal, secuestro	Relato acerca de cómo es la convivencia con grupos armado (ELN) y como se vivió en el territorio la llegada de otros grupos armados (PARAMILITARES), además de cómo se vivió el secuestro por parte de los paramilitares.

Testimonio 10

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Entrevista con Rodrigo Marulanda	Sandra Arenas	EIB	8189	41572	33519	1973;1981;1992	Medellín	Asesinato, masacre	Entrevista realizada dentro del marco de la investigación de altares ubicados en las comunas de Medellín a Rodrigo la persona encargada de mantener un altar en el barrio la Milagrosa. Este altar fue creado por la comunidad a causa de una masacre ocurrida en el barrio donde asesinaron entre 7 a 8 personas entre las cuales se encontraba dos mujeres en una esquina del barrio por presunto personal de la

									policía. Anqué en esta masacre no murió ningún familiar cercano a Rodrigo, el relata el asesinato de su padre y hermano y las consecuencias que este hecho tuvo en su familia. Además, cuenta apartes de la vida cotidiana el barrio y como es la convivencia con grupos al margen de la ley.

Testimonio 11

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Caso: Roviro José Causado - Testimonio del padre Pedro Causado	Casa Campesina Sincelejo	Archivo Virtual de los Derechos Humanos y Memoria Histórica	776	4468	3692	7 de septiembre de 1996	Municipio de Los Palmitos. Corregimiento de El Piñal. Sucre	Desaparición, asesinato	El día 7 de septiembre de 1996 llegaron 17 uniformados y 4 de civil a la vivienda ubicada en Vereda Santa Elena, del municipio de Los Palmitos. corregimiento de El Piñal. Estos se llevaron a Roviro Causado Márquez y Robinson Manuel Narváez Arco; este último hijo de Julio Rafael Narváez Gómez. Testimonio de padre Pedro Causado.

Testimonio 12

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Testimonio Carmen	Sandra Arenas	EIB	15940	83109	67216		Medellín	Asesinato	Entrevista realizada a Carmen en varias secciones donde relata porque realizo un altar (Calvario)a su hijo fallecido con los restos de sangre que quedaron en el lugar donde fue asesinado Robin su hijo mayor a los 30 años por hombres de color negro que vestían de jean, y gabán negro con taches (se presume que esta vestimenta era para ocultar las armas) que hace poco habían llegado al barrio y se encontraban extorsionando a

									las habitantes del barrio. Además, narra cómo era la vida de Robin y sus otros hijos, y como ha sido la vida después de su muerte y los efectos que esto a tenido en sus nietos y en su vida misma.

Testimonio 13

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Tengo un montón de papeles que dicen que fui víctima	Observatorio de derechos humanos y derecho internacional humanitario de FUNDAGÁN	Acabar con el olvido. Segundo informe	1356	7453	6.093	1963 1968 1967-2015	Meta	Extorsión, secuestro, asesinato, cultivos ilícitos, desplazamiento forzado, delincuencia común.	Ganadero relata cómo llego al Meta y como fue el proceso de extorsión y secuestro por parte de las FARC y el EPL. Además de cómo se vivió el proceso de distinción y como ejercía el poder las FARC en el territorio. Por último, dice cuál es su percepción acerca del proceso de paz y como los mecanismos de control del país para la entrega de recursos no presentan una validación adecuada de los papeles dado

									que siendo él un desplazado no ha recibido recursos y otras personas que no han afrontado hechos de violencia están recibiendo beneficios.

Testimonio 14

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
"No haga nada porque lo matamos, Se tiene que ir con nosotros"	Observatorio de derechos humanos y derecho internacional humanitario de FUNDAGÁN	Acabar con el olvido. Segundo informe	1361	7912	6658	Sin fecha	Magdalena Medio	Extorsión, Secuestro	Hombre relata cómo fue su vida antes y después de la llegada de la guerrilla al paramo, como se vivió la extorsión en los pobladores. Como afrontaron el secuestro del papá y sufren por no conocer su paradero, dado que la guerrilla nunca reconoció el hecho. Por otra parte, relata cómo vivió su secuestro y como logro recobrar su libertad luego de que su familia pagara la extorsión. Por último, relata como el gremio ganadero no ha

									encontrado en los gobiernos un respaldo para afrontar y superar los hechos de violencia.

Testimonio 15

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
A Pastrana todo mundo le reclama el fracaso del Caguán, Pero yo pienso que esa era una pela que había que darse	Observatorio de derechos humanos y derecho internacional humanitario de FUNDAGÁN	Acabar con el olvido. Segundo informe	2166	12721	10725	Sin fecha relacionada	Cauca	Desmovilización, despojo de tierras, secuestros, asesinato, extorsión, posconflicto, negociaciones Estado-FARC	Hombre narra su percepción acerca del proceso de paz realizado en el Caguan por el expresidente Pastrana, Inicialmente narra cómo se presentó la lucha indígena sobre la tierra ancestral y como el M-19 colaboro con los indígenas para que estos se apropiaran de las tierras. Por otra parte porque fue tan exitoso el plan Colombia y la seguridad democrática llevada a cabo en el gobierno de Uribe. Donde señala que el

									Cauca ha sido un territorio marcado por el secuestro y la extorsión y el proceso de paz que se adelanta con las FARC no es sinónimo de paz y tranquilidad solo, es un cambio en estructura armada donde se conformaran nuevos grupos de milicias los cuales serán más fáciles de combatir por parte del estado.

Testimonio 16

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Saque su ganadito o venda la finca, porque si vuelve lo Bajamos o le quitamos todo	Observatorio de derechos humanos y derecho internacional humanitario de FUNDAGÁN	Acabar con el olvido. Segundo informe	1585	8318	7412	1980-2000	Caquetá	desplazamiento, secuestro, conflicto entre grupos armados, extorsión	Hombre relata cómo su familia fue víctima de la violencia al ser su padre y el secuestrados por la guerrilla, además de como la guerrilla M-19 y FARC, tenían el control del territorio lo que generaba que debieran contribuir a la guerrilla con el pago de vacunas y algunas tierras se comenzaron a utilizar para la siembra de cultivos ilícitos. Luego de esto describe como se vivió la llegada de los paramilitares y como

									comenzaron hacer barbaries y a solicitar el pago para el uso de la gasolina, cemento entre otros enseres. Por último señala que espera que el proceso de negociación con las FARC genere buenos resultados.

Testimonio 17

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Llegaron y acabaron hasta con el nido de la perra	Observatorio de derechos humanos y derecho internacional humanitario de FUNDAGÁN	Acabar con el olvido. Segundo informe	988	5375	4459	Sin fecha relacionada	Puerto Asís	cultivos ilícitos, masacres, enfrentamientos entre grupos armados, posconflicto	Relato acerca de cómo el municipio de Puerto Asís comenzaron con el cultico de coca, como se comercializaba y el papel de los grupos armados en la negociación. Por otro lado como se vivió la llegada de las autodefensas donde dejo más de 100 muertos en enfrentamientos con las FARC por el control territorial y las consecuencias en la población (represarías). Por último, las exceptivas en el

									proceso de paz donde indica que se espera que avancen y que se logre un poco de tranquilidad en el territorio.

Testimonio 18

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Historia de mi accidente	John Ferney Giraldo. / Programa de Atención a Víctimas del Conflicto Armado. Alcaldía de Medellín	Jamás olvidaré tu nombre	1091	5958	4988	26 de mayo de 2004	La Quiebra Río Verde de los Montes, del municipio de Sonsón.	Minas antipersonas, desplazamiento forzado	Relato de John Ferney Giraldo, quien con su familia debieron salir de la fincan done Vivian debido a que el grupo las FARC le solicita a su familia salir del territorio, en este desplazamiento alrededor de 26 familias salen de sus casas. En el recorrido hacia el municipio de Argelia Jhon Ferney, cae en una mina antipersonas lo que ocasiona que pierda un pie. La recuperación la realiza en el municipio de Rionegro, hoy en

									día tiene una prótesis y trata de llevar una vida normal a pesar de las cicatrices.

Testimonio 19

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Urabá manchada de sangre	Yeraldín Zapata Osorno / Programa de Atención a Víctimas del Conflicto Armado Alcaldía de Medellín	Jamás olvidaré tu nombre	984	5257	4380	Sin fecha	Vereda Caucheras. Urabá Antioquia	Desplazamiento, tortura, masacre, violencia a la mujer	Yeraldín narra cómo vivieron la llegada de las autodefensas, como estas torturaron a su mamá cuando esta se encentraba en gestación de 8 meses, además cono su padre vio morir a sus compañeros de trabajo luego de los paramilitares llegaran y asesinaran a todo aquel que se encontraba en una lista .Como resultados de estos hechos y a amenazas recibidas la familia decide desplazarse a Medellín donde

									deben pasar por muchas dificultades económicas y de vivienda debido a que no cuentan con un lugar donde vivir. Por último señala que su familia tiene la esperanza de algún día volver a su tierra natal.

Testimonio 20

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
La Guerra	Lesmin Yuliana Pérez Gómez/ Programa de Atención a Víctimas del Conflicto Armado Alcaldía de Medellín	Jamás olvidaré tu nombre	731	3838	3185		Nariño, Medellín. Antioquia	desplazamiento, ataque a la comunidad	Narración del desplazamiento que debió realizar la familia de Les min Yuliana Pérez Gómez, luego que el pueblo cerca de la vereda donde vivían fuera atacada por un grupo armado, destruyendo la alcaldía y otras edificaciones. Al llegar a Medellín se enfrentaron a no tener recursos y un lugar donde vivir dado que no tenía trabajo fijo lo cual llevo a que uno de los hermanos debiera pedir limosna y trabajar además

									vivieron en un terreno de invasión donde poco a poco pudieron construir su vivienda.

Testimonio 21

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Los barrenderos	Marlín Yuliana Benítez Mosquera / programa de atención a víctimas del conflicto armado alcaldía de Medellín	Jamás olvidaré tu nombre	839	4365	3619	Sin fecha	Medellín. Antioquia	milicias urbanas, asesinato, control de territorio	Marlín Yuliana Benítez Mosquera, relata cómo se vivía en su barrio el conflicto armado. Donde había presencia de un grupo llamado los bandoleros. En esta época de violencia dos de sus primos que eran parte de la banda son asesinados luego de descubrir por parte de los jefes de la banda que uno de ellos era un infiltrado. La familia se entera de la participación en esta banda en el sepelio al conocer el motivo de su muerte.

									Hoy en día su tía tiene 5 nietos dos de ellos de uno de sus hijos asesinado.

Testimonio 22

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Mi amigo Mello	Jesús Eduvier Correa Echavarría / Producción de historia: Roberto Sanabria	Jamás olvidaré tu nombre	839	4365	3619		Medellín. Antioquia	milicias urbanas, asesinato, combates urbanos	Jesús Eduvier Correa Echavarría, narra cómo conoció a Francisco alias Mello, quien luego de ser su hermano (drogas) y su tía (colaboración) fueron asesinados debido a limpiezas realizadas por las milicias urbanas tomo de decisión de alejarse de los estudios y ser parte de uno de estos grupos. Luego de un tiempo este fue capturado y después de su regreso decide salir de esa vida, sin embargo, pierde la vida al ser asesinado

									cerca de la terminal de transporte dejando una esposa, hijo y madre atrás.

Testimonio 23

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Testimonio N°3	Centro de memoria y conflicto	Testimonios San Carlos	3231	16511	13280		San Carlos Antioquia	Desaparición, reconciliación, desplazamiento	Entrevista donde la participante describe cómo vivió la desaparición de su hija, como luego en el proceso de reconciliación realizado por el grupo armado, fue indicado el lugar donde se encontraba el cuerpo sin vida, además como vivió el desplazamiento y los efectos que esto ha tenido en su vida.

Testimonio 24

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Testimonio Laura Acevedo	Sandra Arenas Grisales	EIB	1340	6161	7475	15/08/2012	Medellín. Antioquia	Resistencias sociales, movimientos sociales	Entrevista con Laura, quien participo en la construcción de un mural en la UdeA y el UN como un homenaje para reconocer el movimiento estudiantil y los hechos de violencia que se vivieron dentro de la institución.

Testimonio 25

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Testimonio Eliana Sánchez	Sandra Arenas Grisales	EIB	17034	71882	89066		Medellín. Antioquia	Movimientos sociales, acciones bélicas, resistencias sociales	Eliana Sánchez en una entrevista narra los hechos ocurridos en la Universidad de Antioquia donde luego de disturbios una estudiante fallece (P), otro compañero debe ser llevado al hospital debido a problemas en un oído y como fue todo el proceso dado que al llegar al hospital no dijeron que eran estudiantes debido a temor por parte de la fuerza pública. Además, relata cómo fue su relación con las víctimas, familia y comunidad universitaria; los sucesos de días anteriores y consecuencias de estos hechos de violencia.

Testimonio 26

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Mi Historia	Fernando Iván Méndez	Narrativas Visibles 2010. La historia detrás de estos ojos.	236	1382	1184	2 de noviembre de 1987	Doradal (Antioquia)	Asesinato, Enfrentamiento	Fernando Iván Méndez, detective del DAS narra lo sucedido luego de disfrutar de una visita a la Hacienda Nápoles en Doradal Antioquia con su familia y compañeros, donde unos hombres armados iniciaron una riña luego de afectar el carro que ellos conducían. Este enfrentamiento con algunos paramilitares de la zona ocasiono la muerte del detective Oswaldo Buitrago Huertas y la incapacidad del relator de historia. Esta incapacidad ocasiona que pierda a su esposa y se enfrente día a día a

									la incertidumbre sobre que pasara en un futuro en su vida cuando llegue la vejez y no pueda movilizarse adecuadamente.

Testimonio 27

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Historia de José Albertino Gutiérrez	Inés Gutiérrez	Narrativas Visibles 2010. La historia detrás de estos ojos.	1460	7360	6180	11 y 14 de noviembre de 1991	San Martín (Meta)	Desaparición	Inés Gutiérrez, madre de José Albertino Gutiérrez narra cómo fue su vida y como fueron los primeros años de vida de su hijo dentro de su familia. Él fue una persona que no le gusto estudiar y por tanto se dedicó a trabajar, dentro de la comunidad donde vivía presento problemas con el orden público ya que fue señalado como delincuente, sin embargo, nunca fue detenido. Este fue visto por última vez luego de visitar a su familia en unas fiestas patronales 11 y 14 de noviembre de 1991. Ya han pasado 20 años y aún desconocen cuál es su paradero. Este relato presenta testimonios de los hermanos donde indican como lo llamaban en casa y cuanto lo extrañan.

Testimonio 28

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Historia de Cristian, Rodrigo y Rubiela Wilches	Ana Elvia Hurtado	Narrativas Visibles 2010. La historia detrás de estos ojos.	870	4472	3740	28 de julio de 1989/ 6 de noviembre de 1993 / 14 de enero del 2002	Meta	Asesinato, desplazamiento	Ana Elvia Hurtado, relata cómo fue asesinada su familia. El primer miembro fue su esposo quien era comerciante y fue asesinado sin conocer el motivo. El segundo fue el hijo quien estaba trabajando en una Caño Jabón y fue asesinado, tampoco se conoció el motivo. Por ultimo su hija quien luego de tanto insistir fue contratada para trabajar en el DAS y tenía el sueño de comprar una casa fue asesinada en la casa luego de llegar a la casa tres hombres armados y uno de ellos llamado como el comandante la asesino de dos disparos. Desde

									entonces Ana Elvira es una desplazada en la ciudad de Villavicencio en espera de apoyo por el estado.

Testimonio 29

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Relatos de Memoria y Dignidad	Ciro Galindo	Narrativas Visibles 2010. La historia detrás de estos ojos.	746	3952	3334	22 de abril de 2003	Villavicencio (Meta)	Asesinato, reinserción, reclutamiento de menores	Padre relata cómo su hijo fue reclutado por parte de las FARC a la edad de 13 años, como este luego se retira del grupo por el programa de reinserción y es dejado en custodia por el ICBF. Luego de poder salir del amparo del programa su hijo es sometido a constantes amenazas y retenciones por parte de hombres de grupos paramilitares. Después de varias denuncias no reciben ayuda por parte de ningún organismo. Un

									día se reciben una llamada y su hijo es llevado por unos hombres dos días después lo encuentran en la morgue asesinado. El asesinado ocurre en Villavicencio. Luego de ese hecho la familia recibe amenazas lo que los obliga a salir de la cuidad donde luego de dos años regresan en espera que la situación mejore.

Testimonio 30

TÍTULO	AUTOR	FUENTE	CARACTERES			FECHA	LUGAR	TÉMATICA	Descripción de los hechos narrados
			PALABRAS	CON ESPACIO	SIN ESPACIO				
Biografía de Naydi Heraque Noguera	Dain Antonio Heraque	Narrativas Visibles 2010. La historia detrás de estos ojos.	963	5267	4462	22 de noviembre de 1997	Puerto Príncipe (Vichada)	Desaparición, asesinato	Hermano relata apartes de la vida de Naydi Heraque Noguera, quien un día salió de la casa en búsqueda de una oportunidad laboral para cumplir con el sueño de comprar casa propia, luego de partir llamo a su familia e informo que se encuentra retenida por un grupo armado de las FARC y que le tenían los papeles, esa fue la última vez que hablaron con ella, las personas que salieron con ella al trabajo volvieron pero ella nunca volvía, después de un

									tiempo se conoció que fue asesinada por directriz del Negro Acacio quien la acuso de ser paramilitar. La familia aun llora su partida dado que perdieron a un miembro importante y muy querido.