

TESIS DOCTORAL

2022

**COMPILACIÓN Y ANÁLISIS DE UN CORPUS
TEXTUAL SOBRE PSICOLOGÍA: HACIA LA
REDACCIÓN ASISTIDA DE TEXTOS
ACADÉMICOS EN ESTE ÁMBITO**

JOANNA MORAZA ERAUSQUIN

**PROGRAMA DE DOCTORADO EN FILOLOGÍA. ESTUDIOS
LINGÜÍSTICOS Y LITERARIOS**

DIRECTORA: DRA. MARÍA ÁNGELES ESCOBAR ÁLVAREZ

Índice General

Tabla de contenido

Capítulo 1. Introducción.....	4
Objetivos	7
Preguntas de Investigación	7
Justificación del Marco	9
Estructura de la Tesis	10
Capítulo 2. Antecedentes: La Complejidad de los Textos en el Campo de la Psicología.....	12
Textos con Fines Específicos	12
Estructura y fases en la redacción de textos con fines específicos	13
Dificultades en la Redacción de Textos en el Ámbito de la Psicología.....	15
Estrategias para Mejorar el Estilo de Redacción en Textos de Especialidad	16
Capítulo 3. Estado del Arte: Sistemas de Redacción Asistida	18
Capítulo 4. Sistema ArText.....	19
Revisión de la <i>Rhetorical Structure Theory</i>	19
La Herramienta de Redacción Semiautomática ArText.....	20
Diseño e implementación del prototipo	21

Capítulo 5. Metodología.....	22
Diseño de la Investigación	22
Descripción de Informantes.....	23
Procedimiento	24
Descripción de cuestionarios	24
Capítulo 6. Análisis de los Resultados.....	25
Exposición de Resultados.....	25
Primera fase.....	26
Segunda fase	27
Tercera fase	29
Capítulo 7. Propuesta de Plantillas en el Prototipo ArText.....	30
Propuesta de Plantillas	30
Artículo de divulgación	30
Artículo de investigación.....	31
TFG.....	32
Discusión de Resultados.....	33
Capítulo 8. Conclusiones.....	35
Referencias	44

Capítulo 1. Introducción

Esta investigación pretende combinar el campo de la lingüística aplicada con el de la psicología mediante la compilación y el análisis de algunos textos académicos y la creación de plantillas que pueden ayudar a la redacción de textos en español como lengua de especialidad. Por tanto, el estudio que se encuentra en esta tesis se basa en algunas investigaciones interdisciplinarias.

El primer objetivo es averiguar cuáles son los géneros textuales que se escriben con mayor frecuencia y, en su caso, los que presentan mayor dificultad en el ámbito de la psicología. También se investigan otros factores que pueden repercutir en la complejidad de este tipo de textos al que se enfrentan los profesionales. Se han considerado tres géneros textuales, a saber: el artículo de divulgación, el artículo de investigación y la tesina de fin de carrera o trabajos de fin de Grado (TFG), en el recién implantado espacio europeo de estudios universitarios. La elección de estos tres géneros se debe a que de acuerdo con nuestra investigación son los que con mayor frecuencia se escriben en el campo de la psicología por parte de los profesionales, investigadores y estudiantes.

La investigación se ha llevado a cabo en tres niveles: nivel textual, donde se han analizado las estructuras de la plantilla y la fraseología, nivel léxico y la extracción de sus datos prototípicos y la reflexión de los datos más relevantes de nivel discursivo. Para cada género en español se seleccionaron 20 textos, lo que hace un total de 60.

Tras el análisis de los textos, se rellenaron una serie de fichas que reflejan los datos recogidos de cada uno de ellos de manera individual. Estas fichas se utilizarán para crear las plantillas que tendrán un formato de muñeca rusa y se mostrarán en el apartado correspondiente.

La presente investigación se enmarca en el proyecto arText que es un sistema automático de redacción de textos especializados. Esta herramienta que ya cuenta con plantillas dirigidas a los ámbitos de la administración pública, la medicina, el turismo y el mundo académico, permite la incorporación de una nueva sección de plantillas en el campo de la psicología. Por tanto, nos centraremos en diseñar plantillas que puedan implementarse en esta herramienta y, así, satisfacer la necesidad de redacción de textos especializados en un nuevo ámbito.

Como estudiante del grado de psicología en la Universidad Nacional de Educación a Distancia (UNED), así como estudiante de doctorado en el programa de filología de la misma universidad, la autora de la presente tesis conoce de primera mano las dificultades que entrañan la redacción de textos específicos, por lo que se decide que, aunando los conocimientos de ambas áreas, se crearán una serie de plantillas destinadas a cubrir el campo del español como lenguaje de especialidad.

Se espera que con esta implementación un gran número de estudiantes de psicología encuentren orientaciones claras a la hora de redactar sus escritos académicos. Además, será muy útil para cualquier profesional o docente que se encuentre en la misma situación.

Tras la lectura de la bibliografía existente sobre el tema, la primera observación es que, dentro de un mismo ámbito, resultaría más adecuado hacer una distinción entre el ámbito académico, por un lado (profesores universitarios), y el profesional y el estudiantil, por otro. Por tanto, se dirigen cuestionarios cualitativos a los profesores y profesionales y a los estudiantes por separado, con preguntas ligeramente variadas a una pequeña muestra.

Una vez recogida la información y analizadas las respuestas obtenidas, se elabora un cuestionario cuantitativo con respuestas de opción múltiple que se envía a una muestra más amplia. Las respuestas recibidas en este segundo cuestionario se utilizan para dilucidar los tipos de textos que se escriben con mayor frecuencia, los que presentan mayores dificultades y cuáles de ellos son más o menos difíciles.

Una vez elegidos los géneros textuales que resultan ser artículos de divulgación e investigación y tesis de fin de carrera, se analizan siguiendo la metodología utilizada en el proyecto arText. En cuanto al nivel textual, el análisis se realiza de forma manual, mientras que en los niveles léxico y discursivo se realiza de forma semiautomática mediante una herramienta para el análisis cuantitativo de rasgos lingüísticos de textos escritos que se ha desarrollado en el marco del equipo arText de la UNED, reflejando toda esta información en los ficheros correspondientes. Una vez completadas las fichas de cada uno de los textos, se eligieron los contenidos para la plantilla final.

Las plantillas se presentan en formato de muñeca rusa y, tras rellenarlas, se han evaluado. Para ello se vuelve a utilizar un cuestionario cuantitativo que ha sido probado de manera previa por un grupo de expertos antes de enviarlo a un mayor número de informantes. Esta evaluación final pretende comprobar la fiabilidad y validez de las plantillas según la experiencia de los informantes.

Una vez obtenidos estos resultados y listas las plantillas, la evaluación final pretende comprobar la fiabilidad y validez de las plantillas según la experiencia de los informantes. Con este fin, se les pidió que tuvieran en cuenta los textos leídos y escritos en cada uno de los géneros textuales estudiados y así poder responder al cuestionario correspondiente. El análisis de estos cuestionarios pretende reflejar si las plantillas son representativas de los géneros elegidos y en qué medida.

Objetivos

El objetivo general de la tesis es la de analizar textos académicos en el campo de la psicología para una implementación futura de plantillas en una herramienta tecnológica y observar si dicha implementación sería beneficiosa para los usuarios, mientras que los objetivos específicos son los siguientes:

1. Analizar el corpus de los géneros textuales elegidos a nivel textual, discursivo y léxico.
2. Elegir el contenido más adecuado, según los criterios de selección, para completar las plantillas finales.
3. Completar la plantilla final para cada género textual elegido.
4. Diseñar un análisis descriptivo de los textos en español de especialidad.
5. Evaluar la adecuación, fiabilidad y representatividad de las plantillas.

Preguntas de Investigación

Para Douglas (2000), el aprendizaje y la enseñanza de lenguas con fines específicos es hoy en día una disciplina cada vez más extendida, aunque no se ha investigado lo suficiente sobre ella hasta hace poco. Si bien el inglés es una de las lenguas más relevantes utilizadas para fines específicos (en especial en ciencia y tecnología) según Kaplan (2011), el presente proyecto se centra en textos específicos en español.

En cambio, para Li y Li (2015) el lenguaje de la ciencia, al igual que otros lenguajes de especialidad, tiene características propias que se corresponden con una determinada terminología que utiliza palabras abstractas, taxonomías técnicas y científicas, y estructuras sintácticas complejas que van desde oraciones pasivas hasta oraciones incrustadas. Los científicos tratan de utilizar la forma más objetiva y eficaz para explicar la aparición, las características y las funciones de los fenómenos (Reeves, 2005), por lo que, al pertenecer la psicología a la rama de las ciencias de la salud, esta condición también se cumple en el caso que nos ocupa.

Como se ha mencionado antes, hay una falta de investigación sobre este tema. Por tanto, la presente investigación está enfocada a resolver un problema práctico, ya que los objetivos que se plantean están orientados primero a desarrollar el prototipo arText y, por tanto, a una posible aplicación y mejora para que la herramienta progrese y llegue a un mayor número de destinatarios.

Se ha mencionado que las características del español como lenguaje de especialidad en campos como la ciencia, el derecho o la propia psicología hacen especialmente difícil la redacción de textos en este contexto. Cumplir con el formato que exigen las revistas o las editoriales para su publicación; escribir de forma clara, concisa y cohesionada; utilizar un léxico o una terminología adecuada a la especificidad del campo o sintetizar de forma correcta son algunas de las dificultades con las que se enfrentan los escritores a la hora de redactar textos de especialidad.

Por otra parte, el tiempo necesario para redactar este tipo de textos es muy elevado, ya que hay que cumplir varias normas (citación, formato o bibliografía) para que sea adecuado al contexto. Por ello, muchos profesionales, profesores y estudiantes no disponen del tiempo ni de las herramientas suficientes para dedicarse a la redacción y posterior publicación, por lo que muchos proyectos, pueden no ver nunca la luz.

A la vista de la falta de investigación y de la falta de herramientas de apoyo a la redacción, y teniendo en cuenta los objetivos principales planteados, se formula la siguiente pregunta de investigación: ¿Resultaría beneficiosa la futura implementación de plantillas como herramienta tecnológica tras la compilación y el análisis de textos específicos en el ámbito de la psicología?

Justificación del Marco

Tal y como se ha mencionado en el apartado anterior, este estudio se enmarca dentro del proyecto arText, cuyo prototipo tiene un acceso gratuito llegando a un gran número de usuarios (<http://sistema-artext.com/>). En consecuencia, y dado que las fichas modelo obtenidas podrían implementarse en esta herramienta como una sección más, sería posible enfocarlo a un nuevo grupo diana del sector de la psicología.

A corto plazo, la aplicación de las plantillas creadas puede ser una herramienta de orientación y guía para todos los interesados en escribir un artículo de divulgación o investigación y un TFG. Con el uso de las plantillas se pretende dar un ejemplo de lo que podría ser este tipo de texto, incluyendo determinados apartados o contenidos recurrentes. A partir de ahí, los usuarios pueden trabajar con unas nociones de formato y fraseología adecuadas para cada tipo de texto.

A medio y largo plazo, se podría plantear la hipótesis de si estas plantillas contribuyen a reducir el tiempo de redacción y si son útiles para la población destinataria. Del mismo modo, se podría analizar si el uso de esta herramienta o de otras similares aumentaría el número de publicaciones con una finalidad concreta. Ambas hipótesis, de comprobarse que son ciertas, beneficiarían a la comunidad científica, ya que un aumento de las publicaciones en áreas especializadas permitiría que un mayor número de personas compartieran sus experimentos, nutriendo y enriqueciendo cada vez más cada campo.

Además, dada la importancia de la investigación hoy en día en casi cualquier campo, los profesores y estudiantes, en especial en el ámbito universitario, están obligados a publicar un número mínimo de trabajos cada cierto tiempo. Esto, junto con otras muchas tareas, puede resultar desmotivador, por lo que dotarles de este tipo de herramientas podría motivarles a publicar con mayor frecuencia.

Al tratarse de una investigación interdisciplinar, el presente proyecto pretende beneficiar a los ámbitos académicos de la lingüística, y más en concreto sobre el español con fines específicos en psicología. Con la futura implementación de las plantillas creadas, la herramienta tecnológica que lleva el sello de la UNED se ampliaría con una nueva sección dedicada al ámbito de la psicología.

Por otro lado, el análisis lingüístico de los 25 textos en inglés se ha realizado de forma manual, incluso a nivel discursivo y léxico, ya que la herramienta original no dispone aún de esta lengua en su base de datos y ha sido necesario analizar los textos uno a uno. Aunque estos datos no se utilizarán para la creación de las fichas modelo finales pueden servir para crear una base de datos y posteriores plantillas en inglés de cara a futuras investigaciones y poder implementarlas en la misma herramienta tecnológica.

Estructura de la Tesis

Los cuatro primeros capítulos están dirigidos hacia la parte más teórica de la tesis. Así, en este primer capítulo se ha introducido el tema del estudio, así como sus objetivos generales y específicos, las preguntas de investigación y el estado del problema, justificando el estudio que se va a hacer. En el segundo capítulo se aborda la complejidad de los textos en el ámbito de la psicología, en concreto de los artículos científicos, los artículos de divulgación y los TFG. También se mencionan las dificultades encontradas en su redacción, así como algunas estrategias para mejorar el estilo de escritura.

En el tercer capítulo se presentan algunos de los sistemas de escritura asistida y gestores bibliográficos de ayuda a la redacción más utilizados, así como algunas de sus ventajas e inconvenientes. Para ello se habla de sus características y particularidades y también se alude a su modo de empleo a la hora de redactar textos de especialidad. Mientras que el capítulo cuatro analiza con más detalle el sistema de redacción asistida que sirve de marco a esta investigación, el sistema arText. Se describen las diferentes fases previas del proyecto y por las que ha pasado el prototipo antes de ser presentado a los usuarios. Algunas de sus características también se plasman en este capítulo.

Los últimos cuatro capítulos tratan la parte más práctica del estudio, por lo que en el quinto capítulo se describe la metodología que incluye el diseño, los informantes, el procedimiento seguido, los cuestionarios utilizados y las preguntas de investigación. A continuación, en el sexto capítulo se recoge el análisis de los resultados obtenidos en cada uno de los géneros textuales y en los niveles textual, discursivo y léxico mostrándose mediante tablas, gráficos y figuras.

Por último, el capítulo séptimo está dedicado a la discusión de los resultados obtenidos descritos en el anterior capítulo y a la presentación y evaluación de plantillas. Además, el capítulo ocho sirve para elaborar las conclusiones finales del estudio, destacando las fortalezas y limitaciones de la investigación, así como los retos para investigaciones futuras.

Capítulo2. Antecedentes: La Complejidad de los Textos en el Campo de la Psicología

Este segundo capítulo empieza desarrollando el estado del arte. En concreto, se tratan los antecedentes de la investigación teniendo en cuenta la complejidad que presentan los textos en el campo de la psicología y las características de los textos de especialidad que son objeto de estudio. Además de las dificultades que entrañan su redacción y las recomendaciones o estrategias a seguir para superar dichos obstáculos y limitaciones.

Textos con Fines Específicos

Los escritos con fines específicos se enmarcan dentro del lenguaje especializado que se adquiere con posterioridad al lenguaje común. En la revisión de la literatura, Parodi (2005) muestra que el discurso de especialidad presenta un mayor grado de abstracción y por ello se debe planificar bien tanto de manera implícita como explícita (Parodi, 2005).

Según Varó et al. (2007) estos textos reflejan los códigos lingüísticos de los diferentes campos, así como los procesos de reflexión de aquellos que se dedican al conocimiento científico. Sirviendo de igual manera para transmitir los resultados mediante un lenguaje más preciso.

Autores como Canales (2010) o Acosta et al. (2008) indican que los lenguajes de especialidad poseen una serie de características como son una amplia tipología textual debido a la inclusión de homónimos y homógrafos, o la comunicación satisfactoria entre especialistas del ámbito satisfaciendo las necesidades dentro del campo en el que trabajan. Por tanto, se emplean todos los recursos lingüísticos utilizados en un ámbito comunicativo que garantiza la comunicación y comprensión entre las personas pertenecientes a él.

Por último, se produce la combinación de diversos códigos para la representación de la realidad dentro del campo específico, que queda sujeta a factores como: el tema y el tipo de interlocutor, su propósito comunicativo en el estilo que emplean, o su longitud u organización. Por tanto, y debido a estas características específicas conviene que el autor conozca de antemano la estructura y las fases en la redacción de este tipo de textos.

Estructura y fases en la redacción de los textos con fines específicos

De acuerdo con Clemente Egío y Cremades Montesinos (2016), hay que seguir una estructura clara con el objetivo de que el texto sea lo más adecuado posible en el contexto requerido. Así, se ha de elegir el tema, la audiencia y la motivación (objetivos de la escritura); la organización de la escritura y definición de las metas generales (planificar); contextualizar el escrito; lectura y revisión del texto (revisar) y la producción final del texto.

Para Morales (2003), el lenguaje especializado y sobre todo el científico presenta un esqueleto determinado donde se encuentra presente el vocabulario científico y los compuestos nominales prototípicos del discurso que caracteriza el campo de la ciencia.

Respecto a la expresión escrita o la redacción, cada género textual distingue una serie de estructuras que si bien resultan similares presentan algunas diferencias. Por motivos de extensión y de relevancia se van a plasmar aquellas que se presenten en los tres géneros textuales analizados en la presente tesis dentro del ámbito de la psicología: el artículo científico o de investigación, el artículo de divulgación y el TFG.

La escritura y publicación de un artículo de investigación por lo general responde a la comunicación de los hallazgos y resultados de un estudio o investigación realizados por los propios investigadores. Por tanto, es conveniente conocer el público o la audiencia a la que se va a dirigir el escrito. Argimón et al. (2010), recomiendan que antes de comenzar con la redacción se prepare un esquema o plan de contenido de lo que se pretende escribir que puede estar compuesto por preguntas como ¿qué?, ¿quién?, ¿dónde?, ¿cómo?, etc.

Tras completar cada apartado con la información relevante, se procederá a su revisión y corrección, y se consideran aspectos como la sustitución de palabras comunes por otras de especialidad, la comprobación de ideas o la revisión de los párrafos. En este punto cobra especial relevancia la conectividad de los párrafos y textos por medio de los signos de puntuación y de los conectores discursivos.

En cuanto a los artículos divulgativos que tienen como principal propósito la difusión del conocimiento adquirido, su lenguaje debe ser más accesible ya que están dirigidos a un público más general. En Upegui & Institucional (2003), se propone partir de un titular que incluya un antetítulo, el título y un resumen; a continuación, los datos acerca de la autoría como el nombre, la filiación o el correo electrónico; seguido de una entrada con un comienzo atractivo que logre enganchar al lectorado; después, el cuerpo o desarrollo donde se encuentra la información más importante, y, para terminar, un final atrayente.

Por último los TFG, según Battaner et al. (2009) exponen los conocimientos dentro del ámbito de la educación. La estructura de un TFG consta de varios apartados cuya obligatoriedad varía según los criterios de la propia universidad según da Cunha, (2016). El estilo de redacción, el formato y referencias bibliográficas también están sujetos a las pautas que marque la institución.

En la mayoría de las guías universitarias consultadas se requiere que el TFG cuente con un resumen con palabras clave; una introducción, seguido del método empleado y de los resultados esperados, y para finalizar, la discusión, y las conclusiones junto con las referencias bibliográficas.

Dificultades en la Redacción de Textos en el Ámbito de la Psicología

Aunque existen modelos y guías que ayudan a la redacción de los tres géneros textuales objeto de estudio tanto en el plano formal (orientaciones acerca del formato, el estilo o normas de redacción entre otras) como a la hora de ofrecer consejos y orientaciones previas a la escritura. Sin embargo, los redactores no están libres de encontrarse con dificultades ante este tipo de escritos. A continuación se van a mencionar algunos de los problemas más usuales que se topan los escritores a la hora de enfrentarse a la redacción del artículo de investigación, de divulgación y el TFG en psicología.

El hecho de que el lenguaje de especialidad se adquiera con posterioridad al lenguaje común y debido a las características deben poseer, los redactores no están libres de encontrarse con dificultades ante este tipo de escritos. Ejemplo de ello son la falta aún de estudios en esta área o la falta de pericia de los redactores noveles que tiene como resultado que no se conozcan bien los diversos requisitos exigidos. Por otro lado y durante la revisión de la bibliografía, autores como Pano et al. (2004) en da Cunha y Montané (2019), han encontrado dificultades como las relacionadas con la motivación, la familiaridad con el objeto de estudio, los obstáculos del propio texto o la falta de competencia lingüística.

Por otro lado Fregoso (2008) citado en da Cunha y Montané (2019) y Rodríguez (2013) indican problemas correspondientes al uso de grafemas, formas verbales, vocablos léxicos, estructuras sintácticas y componentes semánticos. En Menéndez et al. (2016), se mencionan problemas en el aspecto formal y metodológico en la redacción de artículos de investigación que se envían sin atender a las normas estipuladas de estructuración y composición. Por último, la falta de precisión, claridad y brevedad del lenguaje así como la deficiente utilización o descontextualización de las palabras clave utilizadas por el autor son otras dificultades halladas en Pellón et al. (2012).

Estrategias para Mejorar el Estilo de Redacción en Textos de Especialidad

Reparar en las dificultades mencionadas anteriormente es el objetivo de estudios e investigaciones. Por eso, sugieren una serie de estrategias para superar los obstáculos y mejorar el estilo de redacción.

Para aumentar la pericia a la hora de escribir textos de investigación y divulgación la autora Chávez (2010) propone la lectura de artículos de esta índole y la identificación de sus diversos componentes, así como un entrenamiento explícito en habilidades lectoras antes de comenzar a redactar.

En López Leyva (2013), se incluyen cuatro etapas a la hora de redactar artículos. En la primera fase se propone definir el plan de escritura y seleccionar la revista a la que irá dirigido el texto. En la segunda fase se comienza a escribir a partir de los datos seleccionados y revisando diferentes modelos de artículos aceptados por la revista escogida. En la tercera fase se establece la manera en la que se va a presentar la información y los datos y por último se redacta un borrador que posteriormente se irá editando.

Como paso previo a la redacción y publicación o difusión del documento se deben conocer las normas y los estilos de citación. Por ello, se recomienda familiarizarse previamente con los requerimientos exigidos por las revistas para no tener problemas a la hora de que se

accepte el escrito. En relación a los estilos de citación resulta conveniente el uso de gestores bibliográficos para referenciar de manera correcta. Por último, conviene releer el texto pasados unos días, y tras un periodo de maduración se repasara la fluidez o la coherencia de las ideas. También se dota de estructura a las oraciones y al texto, y se repasa la gramática, la puntuación, la ortografía, los encabezados y las tablas y gráficos que pudiese haber.

Capítulo 3. Estado del Arte: Sistemas de Redacción Asistida

En este tercer capítulo se sigue profundizando en el estado del arte y para ello se presentan algunos de los sistemas de redacción asistida más relevantes utilizados en la actualidad.

De acuerdo con Heilman y Sagae (2015) estos sistemas y aplicaciones informatizadas resultan muy útiles para redactores tanto noveles como con experiencia. Para su creación y desarrollo, se han efectuado diversos estudios y análisis en los que sus creadores han tenido que revisar la literatura, reunir los datos y llevar a cabo un estudio riguroso para conseguir un producto lo más preciso posible, utilizando los algoritmos adecuados.

El modo de empleo de estos sistemas y sus herramientas son muy similares, además, la mayoría facilitan las búsquedas de datos y cuentan con una barra de herramientas desde la que el usuario puede operar a su antojo. Estos programas también sirven de redes de apoyo y traspaso de información entre usuarios del mismo grupo siendo su principal ventaja el poder exportar la bibliografía creada y tener compatibilidad con otros dispositivos.

Sin embargo, estos sistemas de redacción asistida presentan también una serie de desventajas o limitaciones. En primer lugar, algunos gestores bibliográficos, si bien sirven para administrar y gestionar las referencias, no son compatibles con todas las bases de datos.

El sistema LaTeX y Scrivener también tienen una serie de limitaciones puesto que el proceso de creación y modificación de plantillas no resulta muy intuitivo en estas herramientas. Por último, cabe añadir que no todos estos programas son gratuitos, si bien algunos cuentan con una versión de prueba de un mes y la mayoría de universidades o instituciones permiten el acceso gratuito a investigadores salvando este escollo.

Capítulo 4. Sistema arText

El sistema arText desarrollado en la UNED (sistema-artext.com, 2015), se diseñó dentro del marco teórico de la *RST* de Mann y Thompson (1987) y posteriores revisiones y aplicaciones en los textos de especialidad. Esta teoría ha sido igualmente aplicada a diversos campos de la lingüística computacional, la comparación lingüística de diferentes lenguas, pero sobre todo en el análisis del discurso escrito que es el campo que nos compete.

Revisión de la *Rhetorical Structure Theory*

Este sistema se ha diseñado dentro del marco teórico de la *Rhetorical Structure Theory* (RST) de Mann y Thompson (1987). Resumiendo la extensa bibliografía consultada, la *RST* sirve como referencia para la descripción de la estructura textual en el marco de la lingüística funcional al estar basada en las relaciones discursivas establecidas en el texto.

Los orígenes de la teoría se remontan al año 1983 donde Mann, Thompson y Mathiesen observan que hasta ese momento no se conoce ninguna teoría respecto a la función o estructura del discurso. Debido a ello, el surgimiento de la *RST* tiene como objetivo aportar los datos que sean suficientes para la programación automática de los textos y actualmente es considerada como una de las teorías más relevantes acerca del discurso.

La Herramienta de Redacción Semiautomática ArText

Este editor de textos online gratuito se ha creado y desarrollado en la UNED por el grupo de investigación liderado por la Dra. Da Cunha y sirve para dar respuesta a las necesidades de los redactores en diversos ámbitos específicos. Actualmente cuenta con cuatro géneros textuales de especialidad como son la administración pública, la medicina, el turismo así como algunos textos del ámbito académico.

También tiene integrados tres módulos que incluyen la estructura, el contenido y la fraseología en el primero de ellos. El formato y la corrección de la ortografía en el segundo. Y el tercer módulo contiene las recomendaciones acerca del léxico y el discurso. A lo largo del desarrollo de la herramienta se han dado una serie de pasos en el que el primero resulta ser la selección del corpus que se pretende analizar.

Una vez seleccionados los textos se procede a su análisis lingüístico que sigue las teorías de Van Dijk en la esfera textual, la Teoría Comunicativa de la Terminología de Cabré (2009) en el plano léxico y la ya mencionada *RST* en el discursivo.

Atendiendo a estas teorías, en la esfera textual se extrae la información correspondiente a los apartados, títulos y contenidos. En la esfera léxica se analizan datos como la fraseología, el nivel de especialización, las siglas, las definiciones, el porcentaje de repetición léxica, algunos rasgos morfosintácticos y el uso de marcas de subjetividad. Mientras que en el plano discursivo se tienen en cuenta la longitud de la oración, la segmentación discursiva y el empleo y variación de los conectores (da Cunha et al., 2016).

Por último, toda la información recogida referente a los tres planos se plasma en una ficha de análisis global que finalmente se resumirá en otra ficha específica similar para todos los ámbitos a la que se denomina ficha modelo.

Diseño e implementación del prototipo

En cuanto al diseño e implementación, y tras crear las fichas modelo, se desarrolla en primer lugar una herramienta que servirá para cuantificar los datos en el corpus automáticamente. El prototipo cuenta también con un editor de textos en línea y presenta un módulo de estructura que permite al usuario insertar los apartados prototípicos del documento, añadir títulos o fraseología relacionada (da Cunha 2020). También hay un módulo de corrección ortográfica y formato, y, por último, uno de sugerencias sobre el léxico y discurso.

Para su implementación se ha realizado una evaluación extrínseca en la que se ha preguntado por factores como la accesibilidad al sistema, la estructuración de contenidos, la corrección ortográfica y el formato, la revisión del texto y la valoración final del sistema a usuarios reales. Así como a una evaluación intrínseca donde se emplean textos reales que prueban las reglas y algoritmos del sistema.

Capítulo 5. Metodología

En el capítulo quinto se basa en explicar la metodología empleada a lo largo de la investigación en el que se abordan el tipo de diseño, la descripción de informantes, el procedimiento y la descripción de los cuestionarios.

Diseño de la Investigación

En lo referente al grado de control y de manipulación por parte de la investigadora, el proyecto está enmarcado dentro de un diseño cualitativo. El principal objetivo es identificar una serie de variables que pueden servir para proponer unas plantillas de asistencia a la redacción de textos en el ámbito de la psicología. Para este fin, se han elaborado una serie de cuestionarios que se han distribuido entre profesionales del campo.

Otro de los objetivos principales ha sido el análisis de la información recogida en los cuestionarios, teniendo en cuenta que la manipulación de las variables y su aleatorización ha sido nula, y el grado de control ejercido sobre ellas ha sido mínimo. Al tratarse de un diseño meramente observacional no se pueden crear relaciones causales entre las variables (Sofía et al., 2020), pero las respuestas obtenidas arrojan datos de especial relevancia y utilidad en la construcción del producto final. Por tanto, contaremos con un enfoque cuantitativo y cualitativo presentando un diseño mixto.

Descripción de Informantes

Las distintas muestras con las que se ha contado a lo largo del estudio han estado compuestas por estudiantes, profesionales y profesores de universidad.

El muestreo se ha realizado por medio del método bola de nieve donde los contactos principales han captado a más informantes. De igual forma, se ha contactado con el departamento de psicología de varias universidades. Por otro lado, para completar la muestra del alumnado se han utilizado principalmente los foros y las redes sociales.

La búsqueda y reclutamiento de informantes se dividió en varias fases dependiendo del tipo de información que se buscaba recopilar. Por tanto, no partimos de una única muestra de informantes si no con varios bloques por cada tipo de cuestionario rellenado.

En la fase cualitativa, se contó con una muestra de nueve informantes, mientras que en la fase cuantitativa se recibieron 76 respuestas de parte de la población diana. Mientras que en la fase de evaluación extrínseca realizada a las plantillas, la muestra que validó previamente el cuestionario de evaluación cuantitativo estuvo compuesta por 10 informantes.

Al coincidir los informantes en la extrema longitud del cuestionario decidimos dividir y dirigir este cuestionario a dos muestras diferentes, la primera compuesta por profesionales y profesorado (44 informantes), y la segunda por estudiantes (40 informantes).

Procedimiento

La presente investigación se encuentra dividida en tres fases. Primero, se enviaron una serie de encuestas informativas y experimentales al grupo diana con el fin de seleccionar el corpus de textos de trabajo en el ámbito de la psicología. Una vez seleccionados se procedió al análisis lingüístico de los textos y a la posterior creación de plantillas. Por último, se evaluó la adecuación y validez de estas plantillas a través de otra encuesta de corte cuantitativo.

Descripción de Cuestionarios

Los cuestionarios empleados a lo largo del proyecto, tanto los cualitativos como los cuantitativos, se han creado a través de la aplicación de modelos de cuestionarios *Google Forms*. Esta aplicación se consideró como la más adecuada debido a lo sencillo que resulta confeccionar un cuestionario y almacenar los datos a través de *drive*.

Estos cuestionarios han estado enfocados de manera diferente según la fase de recogida de datos en la que nos encontrábamos. En la primera fase, previa a la elección de los géneros textuales, se confeccionan dos cuestionarios cualitativos de respuesta abierta dirigidos a expertos y alumnado (con 9 preguntas) y docentes (10 preguntas).

También en la primera fase, y gracias a la información recopilada en los cuestionarios cualitativos se redacta un cuestionario de corte cuantitativo con menos preguntas y el objetivo de llegar al máximo número de informantes posible.

En la fase de evaluación extrínseca, se vuelve a confeccionar un cuestionario de corte cuantitativo y se valida previamente con un grupo de expertos. Para la validación, se confecciona otro cuestionario con una única pregunta abierta que se adjunta al cuestionario cuantitativo y se remite a la muestra a fin de conocer su opinión.

Este último cuestionario cuantitativo refleja una serie de preguntas de respuesta múltiple cuyo análisis servirá para evaluar la adecuación de las fichas modelo, y como se ha

mencionado previamente se divide y remite a dos muestras diferentes. Las preguntas relacionadas con los artículos por un lado y las del TFG por otro.

Capítulo 6. Análisis de los Resultados

En este capítulo se dan a conocer los resultados logrados en las tres fases de las que consta el estudio. Estos datos se muestran en cada género textual analizado, así como en el plano textual, léxico y discursivo.

En primer lugar, procedemos a la evaluación de estos resultados teniendo en cuenta los objetivos de investigación planteados al inicio de este documento como son el análisis lingüístico en el plano textual, discursivo y léxico del corpus de textos escogidos. A partir de este análisis, se escogen los contenidos relevantes para cada texto, y, así, se completan las plantillas correspondientes.

Tras completar las plantillas, se valora su representatividad para lo que se ha requerido una vez más la colaboración de una amplia muestra de personas relacionadas con la psicología y un cuestionario en el que se refleja esta evaluación.

Exposición de Resultados

Los resultados obtenidos se van revelando a través de las fases seguidas a lo largo del estudio. Primero, se va a resaltar la información cosechada en la fase primera de encuestas para la elección de los géneros textuales. Después se describen los datos arrojados por el análisis lingüístico, y, se concluye con la etapa de evaluación de plantillas.

Primera Fase

Los primeros resultados recabados confirman una distinción entre el campo académico y el campo profesional como se venía viendo en la literatura revisada. También se observa que los cinco géneros textuales específicos más comúnmente redactados no siempre coinciden con los que se consideran que presentan más dificultades.

A este respecto, los datos recogidos de toda la muestra (n=76) que indican que los cinco textos más comunes en el ámbito de la psicología han sido el informe de diagnóstico (40, 8%), el informe clínico (38, 2%), el artículo de investigación (31, 6%) y, por último, el informe de investigación y el texto del profesor (23, 7%, cada uno). Estos resultados muestran que los textos más comunes no son los que ofrecen más dificultades ya que los textos que se consideraron más difíciles de escribir fueron: el artículo de investigación (42, 1%), la tesis doctoral (30, 3%), el informe de diagnóstico y el informe de investigación (23, 7%, cada uno) y el libro especializado (22, 4%).

Mientras que se observa también que conocer y revisar la literatura antes de escribir es el paso más difícil cuando se intenta redactar un texto académico específico (47,4%). A continuación, cumplir con los requisitos de formato (43,3%), resumir de forma correcta (42,1%) y utilizar el vocabulario adecuado (38,2%) fueron los aspectos más difíciles encontrados a la hora de escribir textos académicos.

Estos primeros resultados han servido para escoger el corpus de géneros textuales que han sido los artículos divulgativos y de investigación y los trabajos de fin de grado (TFG). Estos géneros han sido elegidos debido a su disponibilidad, pues existe un extenso corpus de textos en el caso de los artículos disponible para los lectores. Además de que los artículos científicos han sido identificados como los que mayor dificultad presentan en su redacción.

En cada uno de los textos se han analizado los planos textual, léxico y discursivo. El estudio del nivel textual, que se ha desarrollado de manera manual, ha servido para delimitar los apartados, títulos, contenidos y ubicación en la página de estos, así como su fraseología más representativa. En los planos léxico y discursivo el análisis se ha llevado a cabo de manera semiautomática.

Segunda Fase

Una vez seleccionado y analizado el corpus se han reunido los datos prototípicos en los tres géneros textuales analizados concernientes al plano léxico. Comenzando por el nivel de especialización, los artículos de divulgación, según la clasificación propuesta por Cabré (2009) aparecen en el nivel bajo pues están escritos por especialistas para el público en general con el fin de ser divulgados y llegar a un público más extenso a través de revistas de gran alcance.

En el nivel medio nos topamos con los artículos de investigación pues los redacta un especialista a personas más concretas como pueden ser los aprendices de especialista.

En lo respectivo a los TFG y siguiendo con la clasificación de (Cabré, 2009) estos se encontrarían en el nivel avanzado de especialización pues se podría denominar al alumnado como especialista en el medio que compone un escrito para otro especialista (su tutor o tutora). No obstante, y atendiendo al hecho de que un grado es el primer eslabón de una cadena para convertirse en profesional de la psicología, se ha optado por considerar a este tipo de alumnado como semiexperto.

En los datos de cada rasgo lingüístico analizado la media más alta la encontramos en los TFG. Además, en la mayoría de las características examinadas la diferencia respecto a los otros dos géneros textuales es considerable. En cuanto a los rasgos morfosintácticos hallados se observa como el porcentaje de voz activa sobre la voz pasiva es claramente superior en cada uno de los tres géneros textuales. Esto se traduce en que el corpus de textos trabajado está redactado casi al 100% en activa. En lo que concierne al rasgo de persona gramatical, se percibe un porcentaje muy bajo tanto en la primera persona del singular como del plural en todos los géneros.

En el plano discursivo observamos que el total de textos estudiados ha sido 60, 20 por cada género textual, y un total de 402.937. En este plano se revela que son los TFG los que cuentan con una media mayor en el número de palabras y una media de oraciones por texto de 481,6. La media en los segmentos discursivos también es superior en estos escritos, siendo de 753,1. Existen también un gran salto cuantitativo entre la media de frases y de segmentos discursivos en ambos tipos de artículos y los TFG.

Destaca el número de palabras máximo por oración, cabía esperar que la más elevada fuese la de los TFG, no obstante no se encuentra una gran diferencia con respecto a los otros dos géneros. Es más, teniendo en cuenta las cifras podría decirse que las frases resultan ser extremadamente largas tanto en los artículos de divulgación e investigación como en los TFG.

Siguiendo en el nivel discursivo la herramienta ha arrojado información en torno a los conectores discursivos disponibles en la base de datos y que son los que siguen: antítesis, causa, concesión, condición, contraste, propósito, reformulación y resumen.

Es en los TFG donde más conectores se han encontrado y los más empleados son la condición y la antítesis en el caso de los artículos divulgativos, la antítesis seguida de la condición en los de investigación y la causa, la antítesis y la condición en los TFG. En vista de estos datos podría decirse que los conectores mayormente utilizados en este corpus han sido la condición y la antítesis.

Tercera Fase

Tras el análisis de los datos mostrados se ha completado una ficha modelo para cada uno de los géneros textuales analizados. Dichas plantillas se han sometido a una evaluación extrínseca por parte de profesionales, docentes y alumnado del ámbito. El objetivo es observar en qué medida resultan ser representativas cada una de su género textual, si lo fueren, y valorar si son adecuadas o no para ser implantadas en la herramienta.

Esta evaluación ha arrojado una serie de resultados entre los que se encuentran que las réplicas cosechadas señalan la diferenciación entre el campo académico y el profesional. Además, las plantillas relacionadas con el TFG han obtenido un grado de acuerdo mayor en todos los parámetros. Este dato se acentúa en el apartado de la fraseología, donde los porcentajes más altos de acuerdo y total acuerdo se han dado en un 95,7% de las frases en la muestra compuesta por el alumnado.

Las distintas muestras señalan también que en los artículos de divulgación e investigación si bien algunos contenidos no han sido señalados por la mitad de la muestra la mayoría sí que han superado este umbral. En los TFG casi todos los contenidos han logrado superar el 50% de elección entre los informantes.

Capítulo 7: Propuesta de Plantillas en el Prototipo ArText

En este capítulo se proporciona el producto obtenido tras la discusión de los resultados de la investigación llevada a cabo en esta tesis y se discuten los resultados logrados en la sección anterior.

Propuesta de Plantillas

Terminado el proceso de análisis de cada uno de los textos, se procede a rellenar el modelo de ficha adaptado creado por da Cunha y Montané (2020). Esta ficha se completa con los 20 textos de cada uno de los tres géneros textuales a fin de tener una visión global de la información obtenida. Para la creación de estas fichas modelo se emplearon los mismos criterios de selección en cuanto al número de apartados, títulos y contenidos siempre que su aparición fuese el 50% o más de las veces en el corpus analizado.

Artículo de divulgación

El artículo divulgativo consta de los siguientes 10 apartados: cabecera, resumen en español, palabras clave, resumen en inglés (*abstract*), palabras clave en inglés (*key words*), introducción, cuerpo, conclusiones, conflicto de intereses y referencias.

Respecto a los contenidos de cada apartado, en la sección cabecera se encuentran el título y los datos de la autoría y de la institución. El resumen en español incluye los contenidos principales del artículo y su traducción al inglés se encuentra en el *abstract*. Las palabras clave y las *key words* presentan los términos en ambos idiomas que mejor representan el tema.

En la introducción, los contenidos hallados son una presentación y breve descripción del tema del artículo, así como sus características más sobresalientes. La presentación de datos relacionada con la investigación, la presentación de objetivos y su especificación e información sobre trabajos previos relacionados también son contenidos de esta sección.

En el cuerpo se ubican las evidencias y hallazgos de diversos estudios y autores, la recopilación y descripción específica de los datos relacionados con el tema del artículo, los resultados de estudios previos relacionados con el artículo, y por último la presentación de los resultados logrados. Por último en las conclusiones los contenidos principales son la interpretación del autor o autores de los resultados, las necesidades encontradas y retos y futuras líneas de investigación. En cuanto a los contenidos de la bibliografía se encuentra una lista de referencias de todos los estudios e investigaciones citadas a lo largo del artículo.

Artículo de investigación

Respecto a los artículos de investigación, estos se componen de 11 secciones cuyas seis primeras coinciden con los artículos divulgativos y son: la cabecera, el resumen, las palabras clave, *abstract*, *key words* y la introducción. También ocurre en las dos últimas, conflicto de intereses y bibliografía.

Les sigue el apartado método, resultados y discusión. Los contenidos de la cabecera también coinciden con los del artículo divulgativo, sin embargo los de la sección introductoria son los siguientes: breve descripción del tema del artículo, información sobre otros trabajos relacionados con el tema del artículo, presentación de los objetivos y especificación de dichos objetivos.

En el método los contenidos hallados son el procedimiento seguido, las bases de datos utilizadas, los criterios de inclusión y exclusión empleados y la descripción final participantes de la muestra así como el análisis estadístico. Mientras que en los resultados se exponen las tablas de resultados y los resultados obtenidos en la investigación.

En la discusión, se revisan los resultados obtenidos y se plasman las interpretaciones del autor y las limitaciones encontradas en la investigación. Por último, la bibliografía recoge la lista de trabajos y autores citados en el artículo a modo de referencia.

TFG

En la plantilla de los TFG que cuenta con un total de 13 secciones se divide en: portada, índice, resumen, palabras clave, *abstract*, *key words*, introducción, método, resultados, discusión, conclusiones, bibliografía y anexos. Insertos en la portada se encuentran el título del documento, los datos de la autoría y la institución así como el nombre de la persona que dirige el trabajo.

En el índice, se encuentra el índice de contenidos que se presentan a lo largo del trabajo. En las secciones resumen, palabras clave, *abstract* y *key words* se encuentran los contenidos con el mismo nombre. En la sección introductoria los siguientes contenidos resultan prototípicos: la definición y descripción del tema del trabajo, la síntesis de los trabajos de otros autores relacionados con el tema del TFG y los datos de prevalencia. La presentación de objetivos y su especificación, que se dividen en dos contenidos, y el avance de los resultados y las necesidades encontradas se encuentran también en este apartado.

Dentro del método, los contenidos encontrados son la descripción de los criterios de inclusión y exclusión, la descripción de participantes reclutados, la explicación del método y procedimiento llevado a cabo, la descripción del análisis realizado y la descripción de tablas y figuras. Mientras que en los resultados se plasman los resultados obtenidos y la descripción de las tablas y figuras relacionadas con los resultados.

En la discusión se recapitulan los resultados y se interpretan. Otros contenidos versan sobre las necesidades encontradas y las limitaciones. Y por último en la sección conclusiones los contenidos muestran las conclusiones finales del autor. En la bibliografía, y tal y como sucede en los casos anteriores se presenta una lista con las referencias de los autores citados.

Discusión de Resultados

Los datos obtenidos en las distintas fases han mostrado una serie de resultados que en ocasiones ratifican la literatura previamente consultada. En la fase de selección de textos, y del cuestionario cualitativo utilizado se vislumbra que los géneros textuales característicos del ámbito psicológico coinciden con los del campo de la medicina perfilados en el estudio de da Cunha y Montané (2019).

Además, los resultados indican que dentro del dominio de la psicología se presentan dos contextos cuya visión y tratamiento del lenguaje específico es diferente siendo estos dos el lenguaje académico y el profesional. Estos resultados coinciden con los de Trejo et al. (2019) ya que para estos autores el alumnado universitario tiene que escribir sobre diversos géneros destinados a los propósitos y necesidades de cada materia específica, mientras que los escritos profesionales están delimitados por las convenciones de las instituciones públicas y privadas.

Durante la revisión de la literatura se han encontrado en diferentes estudios algunas de las dificultades a la hora de redactar este tipo de textos. En la presente investigación se ha identificado el conocer y revisar la literatura como la mayor dificultad, que coincide con lo señalado por Nieves et al. (2014). Por otro lado, nuestro estudio señala que cumplir con los requisitos de formato adecuados y exigidos por los destinatarios es lo segundo más difícil para los informantes, lo que concuerda con las dificultades señaladas por Perez (2010).

En cuanto a los géneros textuales analizados, las fichas modelo creadas y expuestas coinciden en apartados y algunos contenidos con los formatos mencionados previamente por autores como Perez (2010) o la Asociación para el Avance de la Ciencia Psicológica (AACP) para el artículo científico. Así como la ficha modelo para el TFG coincide en gran medida con lo requerido en las guías de universidades como la UNED, Pontificia de Comillas, Universidad del País Vasco o la Universidad Complutense.

En cuanto al artículo divulgativo, al estar dirigido a un público variado, que por lo general no es especialista en la materia, su estructura cambia con respecto al de investigación (Upegui & Institucional, 2003). La ficha modelo presentada coincide en formato con lo mencionado por Upegui & Institucional (2003) al no contar con apartados tan concretos como el método o los resultados, sino un apartado más general denominado cuerpo.

Capítulo 8: Conclusiones

En este capítulo se trata de hacer un resumen de los aspectos más relevantes de la investigación y de la tesis, así como plasmar las conclusiones principales de la investigación.

En primer lugar, pensamos en una futura implementación de las plantillas, en la que los contenidos que hayan adquirido el mayor porcentaje podrán insertarse como característicos de la sección y los que hayan adquirido el menor podrán mostrarse como una alternativa opcional para el editor. De este modo, se dispondría de una estructura modelo que coincide con la estructura requerida por el medio de publicación, unos contenidos acordes con dicha estructura y otros opcionales que pueden ser utilizados para añadir información relevante.

Teniendo en cuenta lo anterior, se observa que se ha cumplido el objetivo general de analizar textos académicos en el campo de la psicología para una futura implementación de plantillas en una herramienta tecnológica. Asimismo también se logra la observación de si la implementación sería beneficiosa para los usuarios. Mientras que los objetivos específicos, como se observa a continuación, también se han cumplido.

El primer objetivo específico ha sido el de analizar el corpus de los géneros textuales elegidos a nivel textual, discursivo y léxico. Como se ha comentado a lo largo de la investigación, el análisis a nivel textual se ha desarrollado de manera manual, mientras que los otros dos planos se han analizado de manera semiautomática. Tras este análisis se extraen las partes de los que están compuestos cada uno de los textos escogidos así como sus contenidos.

El cumplimiento del primer objetivo da paso al segundo, pues de los contenidos extraídos se elige el más adecuado, siguiendo los criterios de selección y exclusión. Se escoge como representativo del género cada apartado, subapartado y contenido que se repita el 50% o más de las veces, lo cual lleva a completar la plantilla final para cada género textual elegido y a la consecución de los objetivos segundo y tercero. La principal conclusión obtenida tras cumplir estos primeros tres objetivos es que se observan similitudes entre las fichas modelo con los manuales y guías ya existentes para orientar la escritura de artículos y TFG.

Una vez rellenas las fichas modelo se procede a conseguir el objetivo cuarto. Para ello, se lleva a cabo un análisis descriptivo de los textos en español e inglés. Esta última pensando en futuras investigaciones, dado que puede resultar ventajosa una implementación de la herramienta insertando una base de datos en otros idiomas. Se concluye que las partes y contenidos de los documentos analizados en inglés son iguales que sus homólogos en español, pues se trata salvo en el caso de los TFG de los mismos escritos traducidos. Si bien existen diferencias en algunas características lingüísticas que deben ser atendidas sobre todo porque el número de textos analizados no es el mismo.

Por último, y como objetivo específico quinto, se encuentra la evaluación de la adecuación, fiabilidad y representatividad de las plantillas. Tras una evaluación extrínseca por parte del grupo diana, la conclusión más relevante es que existe un alto grado de acuerdo en que los *templates* obtenidos son representativos cada uno de su género textual, si bien existen excepciones como las mencionadas en apartados anteriores.

Teniendo en cuenta lo anterior, al principio de este documento se plantearon algunos objetivos y preguntas de investigación, que se resumirán a continuación. El objetivo general era analizar textos académicos del ámbito de la psicología para su futura implementación en la herramienta tecnológica en el marco del proyecto arText.

Con este objetivo, se ha planteado la siguiente pregunta de investigación, a la que se ha intentado dar respuesta a lo largo de la investigación: ¿Sería beneficiosa para los usuarios la futura implementación de plantillas en una herramienta tecnológica tras el análisis de textos específicos del ámbito de la psicología?

Como se ha observado a lo largo del trabajo, el análisis a nivel textual, léxico y discursivo del corpus de textos recurrentes en psicología se ha desarrollado como se esperaba. Sin embargo, hay que añadir que al realizar el análisis semiautomático de los niveles léxico y discursivo en los textos en español, se encontraron algunos sesgos en la herramienta mencionada anteriormente. Por otro lado, el análisis de los textos en inglés, a pesar de que la herramienta no cuenta con una base de datos en este idioma, se realizó de forma manual en los tres niveles, lo que requirió algo más de tiempo.

Sin embargo, tras el análisis y la creación de las plantillas, ha sido necesario realizar una evaluación de su adecuación y representatividad para ver si la aplicación de la herramienta es posible o no en función de los resultados obtenidos. Si los datos recogidos fueran positivos y las plantillas fueran representativas de los textos recurrentes estudiados, dicha implementación sería posible.

Resumiendo algunas de las conclusiones obtenidas, éstas indican que los usuarios solicitados a través de las encuestas han avalado la representatividad de la mayoría de los contenidos presentados en las diferentes secciones, aunque existen algunas salvedades. Además,

en la valoración de la fraseología o estereotipia de las frases, un alto porcentaje de encuestados ha tenido un alto grado de acuerdo en casi todas las frases referidas a los artículos y un grado de acuerdo total en el caso de las disertaciones de grado. Así pues, de acuerdo con estas conclusiones podría decirse que las tres plantillas redactadas para los géneros textuales de artículos de divulgación e investigación y tesinas reflejan de manera adecuada los apartados y contenidos de los mismos. En definitiva, su aplicación podría ser beneficiosa para los usuarios que necesiten redactar estas tipologías textuales.

En cuanto a los objetivos específicos que se han seguido a lo largo de la investigación, se ha dado cuenta de las características que deben tener las plantillas propuestas. En cuanto al análisis lingüístico a nivel textual, discursivo y léxico del corpus de los géneros textuales elegidos, la elección de los contenidos ha seguido los criterios de selección y se ha podido diseñar la plantilla final para cada género textual.

A lo largo de la consecución de los objetivos y el transcurso del estudio se han encontrado algunas dificultades fruto de la situación propiciada por el COVID-19 y a la imposibilidad de conseguir una serie de géneros textuales debido a su naturaleza y protección de datos. Asimismo, quedan pendientes futuros proyectos e investigaciones como la réplica del mismo estudio con otras lenguas relevantes para seguir profundizando en el campo de la redacción de textos específicos.

Este proyecto tiene una serie de puntos fuertes y débiles que hay que tener en cuenta a la hora de analizar los resultados. Uno de los puntos fuertes de la herramienta es que se ha llevado a cabo un análisis lingüístico de un corpus de textos en un campo en el que antes no existía la herramienta.

La aplicación arText tiene una sección referida a la medicina, pero carece de una sección

de psicología, por lo que uno de los puntos fuertes de esta investigación es haber completado tres plantillas, basadas en el modelo de da Cunha y Montané (2020), en esta última área. Además, estas tres plantillas pueden servir de guía para los investigadores cuyo propósito es publicar tanto trabajos de investigación como de divulgación y para los estudiantes que están escribiendo o están a punto de escribir su tesis.

Para completar las plantillas, se han consultado diversas fuentes, tanto en la literatura sobre el tema como de particulares. La voz de expertos, profesionales y estudiantes ha servido para obtener diferentes perspectivas y extraer información muy valiosa. Primero, para seleccionar los géneros textuales así como para evaluar las plantillas finales, lo que se considera un valor añadido. Del mismo modo, la investigación tiene una estructura clara y ordenada, desde la mencionada selección de textos, pasando por el análisis hasta la evaluación del producto final.

Por último, al revisar los textos en contraste con los datos obtenidos en la herramienta, se han identificado una serie de sesgos como palabras marcadas como unidades subjetivas que no lo eran, errores en la enumeración de siglas, etc. Estos hallazgos pueden servir para corregir esos errores e implementar la herramienta.

Sin embargo, como se ha mencionado anteriormente, el proyecto presenta una serie de puntos débiles que deben ser abordados con el objetivo de tenerlos en cuenta en futuras investigaciones. Las plantillas finales han sufrido cambios e implementaciones con posterioridad a su evaluación por lo que no se ha tenido la oportunidad de preguntar por algunos contenidos añadidos en las revisiones posteriores.

Por otra parte, y debido al tiempo que llevaría preguntar por toda la fraseología analizada y su análisis, esto no se ha evaluado en las encuestas finales. En su lugar, hemos elegido la fraseología de los apartados que resultan ser comunes a ambos tipos de trabajos y disertaciones. En relación con esto, los resultados obtenidos por los encuestados en la fraseología pueden estar

sesgados porque es muy difícil recordar la redacción exacta de todas las frases leídas o escritas a lo largo del tiempo.

Los puntos débiles recién mencionados son el resultado de una serie de dificultades encontradas a lo largo de la investigación, que pueden mejorarse en futuros proyectos. Estas posibles líneas de investigación futuras se comentan con más detalle a continuación, haciendo especial hincapié en la lengua materna de la autora de la tesis, que es el euskara.

En lo concerniente a las limitaciones, en un principio se pensó en analizar tipologías textuales propias del ámbito de estudio, como informes clínicos, informes de diagnóstico o investigaciones sobre casos reales. Sin embargo, dado que este tipo de documentos está protegido por la ley de protección de datos, fue imposible obtener un corpus suficiente para su análisis. Asimismo, fue bastante difícil obtener tesis de grado escritas en el ámbito de la psicología, ya que la mayoría de los repositorios universitarios sólo están disponibles para sus estudiantes. Este problema pudo resolverse en los textos escritos en español, pero sólo se obtuvieron cinco de ellos en inglés.

Por otro lado, las citas previstas cara a cara con profesores, expertos y estudiantes para realizar la primera entrevista cualitativa que sirve de base a la cuantitativa, tuvieron que ser reescritas como cuestionario online debido a la COVID-19. Además, el problema de encontrar una muestra suficientemente representativa también se ha puesto de manifiesto en las diferentes fases, se han enviado un gran número de correos electrónicos y mensajes solicitando colaboración, sin embargo sólo una pequeña parte ha accedido a contribuir.

La falta de una base de datos en otras lenguas en arText ha llevado a pensar en los beneficios que esto aportaría a los editores y usuarios del sistema. Volviendo a Kaplan (2011) y dado que el inglés se considera la lengua de la ciencia, el desarrollo de la herramienta de base de datos y de las secciones en esta lengua serviría para llegar a más personas interesadas en escribir

en inglés. Lo mismo podría ocurrir con otras lenguas, empezando por las cooficiales en España como el euskera y algunas de las más habladas en Europa como el francés, el alemán o el italiano.

Para ello, y como proyecto futuro, esta misma investigación podría replicarse con informantes cuya lengua materna sea o tenga un alto dominio de la lengua de destino, empezando por el inglés. Para ello, la información extraída en este proyecto podría utilizarse y ampliarse para empezar a construir una nueva base de datos.

Dado que la lengua materna de la investigadora es el euskara, para futuras investigaciones también sería interesante realizar el mismo estudio en esta lengua. Según Gonzalez Dios et al., (2021) al estar el euskera todavía en proceso de normalización es mucho más difícil que los alumnos instruidos en esta lengua adquieran las competencias adecuadas para la escritura académica.

Por otra parte, y aunque existen portales de internet y empresas dedicadas a la traducción especializada en el lenguaje científico o jurídico, sólo hay algunas plataformas dedicadas a la redacción de textos específicos en euskera como la herramienta Analhitza desarrollada por el Grupo Ixa en la EHU. Por tanto, futuras investigaciones podrían estar enfocadas a la creación de este tipo de interfaces que asistan la redacción de textos específicos.

Otra posible línea de investigación e hipótesis a plantear es si estas plantillas contribuyen a reducir el tiempo de redacción y si son útiles para la población destinataria. Estos estudios con el objetivo de analizar estos parámetros se podrían llevar a cabo con dos grupos, uno de control que no utilice la herramienta y el otro que sí la emplee, y observar cuántas publicaciones recoge cada grupo al finalizar el proyecto. Del mismo modo, se podría analizar si el uso de esta herramienta o de otras similares aumentaría el número de publicaciones con una finalidad concreta.

Todas estas futuras líneas de investigación contribuirían a aumentar el conocimiento que se tiene acerca del lenguaje específico en general, y de la redacción de textos con fines concretos en particular. Todo ello resultaría en una contribución sustancial al lenguaje de la ciencia y a los científicos, investigadores y académicos inmersos en este campo.

Referencias

- Acosta, R. G., Castellví, M. T. C., & Lidon, J. M. C. (2008). Clasificación de textos especializados a partir de su terminología. *Íkala, Revista De Lenguaje Y Cultura*, 13(19), 15-39.
- Argimon, J., Jiménez, J., Zurro, A. M., & Vilardell, M. (2010). Publicación Científica Biomédica: cómo escribir y publicar un artículo de investigación. *Actividad Dietética*, 14(1), 47.
- Asociación para el Avance de la Ciencia Psicológica. (15 de diciembre del 2011). *Guía Introductoria de Redacción Científica*. https://cife.edu.mx/recursos/wp-content/uploads/2019/01/AACP_Guia_de_Redaccion_Cientifica.pdf
- Battaner, M. P., Atienza Cerezo, E., López Ferrero, C., & Pujol Llop, M. (2009). Características lingüísticas y discursivas del texto académico. *Textos: Didáctica De La Lengua Y La Literatura*.2009; 50: 47-67,
- Cabré, M. T. (2009). La Teoría Comunicativa de la Terminología, una aproximación lingüística a los términos. *Regué Française De Linguistique Appliquée*, 14(2), 9-15.
- Canales, A. B. (2010). Características lingüísticas, pragmáticas y funcionales de las lenguas de especialidad. Implicaciones didácticas y metodológicas. *RAEL: Revista Electrónica De Lingüística Aplicada*, (9), 72-85.

Chávez, V. P. (2010). Clasificación de párrafos y desarrollo de habilidades escritoras en estudiantes de psicología. *Revista Mexicana De Psicología*, 27(2), 257-268.

Clemente Egío, V., & Cremades Montesinos, A. (2016). Estrategias para la escritura académica: ajustes formales y composición final.

da Cunha, I. (2016). El trabajo de fin de grado y de máster: Redacción, defensa y publicación. *El Trabajo De Fin De Grado Y De Máster*, , 1-206.

da Cunha, I. (2020). Una herramienta TIC para la redacción del Trabajo de Fin de Grado (TFG).

da Cunha Fanego, I., Montané March, M. A., & Escobar Álvarez, M.A. (15 de diciembre de 2016). *ARTEXT. Un sistema automático de ayuda a la redacción de textos especializados*. [Sesión de seminario]. Seminario Departamento Filologías Extranjeras y sus Lingüísticas, UNED, España.

da Cunha, I., & Montané, M. A. (2019). Géneros textuales y dificultades de redacción en ámbitos especializados. *Revista Signos*, 52(99), 4-30.

da Cunha, I., & Montané, M. A. (2020). A corpus-based analysis of textual genres in the administration domain. *Discourse Studies*, 22(1), 3-31.

de Ves Cuenca. (2003). E.Un pequeno manual de LATEX.

Douglas, D. (2000). *Assessing language for specific purposes*. Cambridge University Press.

- Gonzalez Dios, I., Iñurrieta, U., & Zabala, I. (2021). General and specialised corpora to raise linguistic awareness in a language undergoing the normalisation process: academic writing in Basque [Innovation and digital technologies in Languages for Specific Purposes]. Paper presented at the *Multilingual Academic and Professional Communication in a Networked World. Proceedings of AELFE-TAPP 2021 (19th AELFE Conference, 2nd TAPP Conference)*. Vilanova i La Geltrú (Barcelona), 7-9 July 2021.
- Heilman, M., & Sagae, K. (2015). Fast rhetorical structure theory discourse parsing. *arXiv Preprint arXiv:1505.02425*,
- Kaplan, R. B. (2011). *English—the accidental language of science*. De Gruyter Mouton.
- Li, X., & Li, L. (2015). Characteristics of English for Science and Technology. Paper presented at the *Proceedings of the 2015 International Conference on Humanities and Social Science Research*, 161-165.
- López Leyva, S. (2013). El proceso de escritura y publicación de un artículo científico. *Revista Electrónica Educare*, 17(1), 5.
- Mann, W. C., & Thompson, S. A. (1987). *Rhetorical structure theory: A theory of text organization*. University of Southern California, Information Sciences Institute Los Angeles.
- Menéndez, M. R., Cantalapiedra, Z. G., & Polo, M. G. (2016). Problemas frecuentes en la redacción de artículos científicos. *EduSol*, 16(57), 137-147.
- Morales, S. M. (2003). Inglés para fines específicos: el discurso escrito en las ciencias de la salud. *Revista De Lenguas Para Fines Específicos*, 10, 151-168.

Nieves, O. V., Rivera, M. M., & Rodríguez, C. A. W. (2014). Guía para la redacción de una propuesta de investigación. Universidad Autónoma de Querétaro. [Archivo PDF].

https://www.academia.edu/Guia_para_la_redaccion...-with-cover-page-v2.pdf

Parodi, G. (2005). Discurso Especializado Y Linguística de Corpus: Hacia el desarrollo de una competencia PSICOLINGÜÍSTICA. *Boletín De Linguística*, 17(23), 61-88.

Pellón, R. T., Miranda, D. R., & Esquijarosa, D. C. (2012). Sugerencias para mejorar el estilo de redacción de un artículo científico en las ciencias de la salud. *Revista Cubana De Educación Médica Superior*, 26(1), 112-120.

Pérez, M. (2010). Guía para la escritura de artículos científicos en psicología. Universidad CES. Colombia. [Archivo PDF].

[http://Biblioteca.Uniminuto.Edu/Documents/968618/1883027/Guia_Escritura_De_Articulos_Cient% C3% ADficos En Psicologia.Pdf/0154b49e-0f6c-485cba20-0602b678b725.](http://Biblioteca.Uniminuto.Edu/Documents/968618/1883027/Guia_Escritura_De_Articulos_Cient%20C3%ADficos_En_Psicologia.Pdf/0154b49e-0f6c-485cba20-0602b678b725)

Reeves, C. (2005). *The language of science*. Routledge.

Rodríguez, M. N. (2013). La redacción de artículos de investigación, desde la construcción de Tesis Doctorales. *Visión Educativa IUNAES*, 7(15), 8-20.

Scrivener (2021). *User Manual*. <https://www.literatureandlatte.com/learn-and-support/user-guides>.

Sistema-artext.com. (2015). *arText - un sistema de ayuda a la redacción de textos en español de ámbitos especializados*. <http://sistema-artext.com>

Sofía, F. D. G., Carmen, G. G., Laura, Q. C., Raquel, R. F., & Encarnación, S. S. (2020). *Fundamentos de investigación en psicología*. Editorial UNED.

- Trejo, R. G., Chacón, G. M., Suárez, G., & de Rojas, M. F. (2019). Vínculos y diferencias entre la escritura académica y la escritura profesional en una carrera técnica universitaria. *Cuaderno De Pedagogía Universitaria*, 16(32), 35-47.
- Upegui, C. A. S., & Institucional, C. (2003). Elementos de escritura académica. *Revista Virtual Universidad Católica Del Norte*, (11).
- Varó, E. A., Martínez, J. M., & Ramos, F. Y. (2007). *Las lenguas profesionales y académicas*. Ariel.