

**Hacia un análisis contrastivo de la correferencia
anafórica en español y en francés a partir de dos de
textos narrativos**

Presentado por:

Sara María Galvis Amaya

Asesor:

Prof. Jorge Mauricio Molina Mejía

Dr. en Informática y Ciencias del lenguaje

Facultad de Comunicaciones

Trabajo de grado

para optar al título de:

Filólogo hispanista

Universidad de Antioquia

Facultad de Comunicaciones

Área de Lingüística y Literatura

2019



**UNIVERSIDAD
DE ANTIOQUIA**

1 8 0 3

Copyright 2019, Sara María Galvis Amaya.

Resumen

Título: *Hacia un análisis contrastivo de la correferencia anafórica en español y en francés a partir de dos textos narrativos.*

A partir del presente trabajo, se pretende etiquetar, categorizar y analizar el uso de algunos elementos correferenciales en una muestra de textos escritos originalmente en español y en francés respectivamente, a partir de herramientas ofimáticas para iniciar la construcción de un corpus de textos en español y en francés para futuras referencias de investigaciones y posible creación de herramientas computacionales, en el marco de la enseñanza de dichas lenguas como lenguas extranjeras, ELE y FLE respectivamente.

Palabras clave: Lingüística - referencia - correferencia textual - elementos correferenciales - anáfora - tipología anafórica.

Abstract

Title: *Towards a Contrastive Analysis of the Anaphoric Coreference in Spanish and French Based on two Narrative Texts.*

Based in the present work, we intend to label, categorize and analyze the use of some coreferential elements in a sample of texts originally written in Spanish and in French respectively, using some tools to start the building of a corpus of texts in those languages for future research references and possible creation of computational tools, in the framework of teaching such languages as foreign languages, EFL and FFL respectively.

Key words: Linguistics - reference - textual coreference – coreferential elements – anaphora - anaphoric typology.

Agradecimientos

Al profesor Jorge, por toda su guía y acompañamiento que hicieron posible el desarrollo del proyecto, pero más especialmente por animarme en aquellos momentos difíciles y por su apoyo incondicional, paciencia y comprensión, sin los cuales, tal como este proyecto, mi conclusión del pregrado, no sería una opción. Gracias por esa dedicación desinteresada que brinda a todos sus estudiantes.

A Elsa y a Daniel les agradezco por absolutamente todo.

A todas aquellas personas que sin saberlo me han ayudado incontables veces y, finalmente, al incauto lector que se acerca a este trabajo, espero te sea útil.

Tabla de contenido

Capítulo 1	: Introducción y objetivos de la investigación	12
1.1.	Introducción general	12
1.2.	Justificación	13
1.3.	Pregunta de investigación	13
1.3.	Objetivos	14
1.3.1.	Objetivo general	14
1.3.2.	Objetivos específicos	14
Capítulo 2	: Antecedentes de la investigación	15
2.1.	Lingüística Textual	15
2.1.1.	Correferencia textual y anáforas	16
2.1.2.	Mecanismos de cohesión textual	17
2.2.	Enfoque contrastivo entre lenguas	18
2.3.	Lingüística Computacional y resolución anafórica	18
2.4.	Didáctica de las lenguas y correferencia textual	19
2.5.	Aprendizaje de Lenguas Asistido por Ordenador y correferencia textual	19
2.6.	Conclusión acerca de los antecedentes de investigación	20
Capítulo 3	: Marco teórico	21

3.1. La Lingüística Textual	21
3.1.1. Definición	21
3.1.2. Génesis del campo y objeto de estudio	21
3.2. Cohesión y coherencia textual	24
3.2.1. Referencia y correferencia textual	25
3.2.2. Anáforas y catáforas	27
Capítulo 4 : Marco metodológico	31
4.1. Etapas del marco metodológico	31
4.1.1 Constitución de la muestra	31
4.1.2 Muestra tentativa	33
a. Textos narrativos en español	33
b. Textos narrativos en francés	33
4.1.4. Etiquetado del corpus	34
Capítulo 5 : Análisis y resultados	40
Capítulo 6 Conclusiones y perspectivas	50
Referencias	52
Referencias de la muestra analizada	55
Anexo A	56
Anexo B	57

Índice de figuras

Figura 1. Rastreo preliminar de anáforas.	36
Figura 2. Extracto del análisis en Excel del texto TeNEsp-1.	56
Figura 3. Extracto del análisis en Excel del texto TeNFra-1.	57

Índice de tablas

Tabla 1. Tipología anafórica de Coletta y Karcher (2012), traducción propia.	29
Tabla 2. Tipología anafórica utilizada en el presente trabajo.	35
Tabla 3. Referentes TeNEsp-1	41
Tabla 4. Tipos de referentes TeNEsp-1.	42
Tabla 5. Referentes TeNFra-1.....	42
Tabla 6. <i>Correferentes anafóricos</i>	44
Tabla 7. Relaciones porcentuales TeNEsp-1 y TeNFra-1.....	45
Tabla 8. Relaciones porcentuales Nom-Pron.....	46

Capítulo 1: Introducción y objetivos de la investigación

1.1. Introducción general

En el campo de la enseñanza y aprendizaje de lenguas extranjeras (Español Lengua Extranjera – ELE y Francés Lengua Extranjera – FLE), que hace parte de la lingüística aplicada, es de suma importancia poder desarrollar actividades que formen a los aprendientes de lenguas a partir de la lingüística textual (Adam, 2011; Lundquist, 2014; Molina Mejía, 2015). Por ello, el presente trabajo de investigación tiene como principal finalidad, el estudio de las unidades correferenciales, específicamente las anáforas. Estas unidades serán estudiadas a partir de dos textos narrativos, uno en español y otro en francés. Se enmarca así, la presente investigación, dentro de un enfoque basado en la lingüística de textos y se lleva a cabo a partir de algunas herramientas ofimáticas computacionales, esto con el fin de que el presente trabajo se retome, en el futuro, en el campo del desarrollo de actividades para la enseñanza de ELE y de FLE.

Este trabajo busca, además, indagar acerca de las características generales que adquieren los componentes dentro de la correferencia textual (uso de anáforas) en su relación con las lenguas francesa y española. Con el fin de observar si este fenómeno, de orden lingüístico, se comporta de la misma forma en ambas lenguas, se ha optado por conformar y estudiar una muestra de dos textos literarios en español y en francés; lo que se considera, permitirá observar y analizar el fenómeno antes mencionado.

Es necesario aclarar, no obstante, que, aunque se trata de un fenómeno que previamente ha sido estudiado en cada lengua por separado, su análisis en un estudio comparado a partir de textos literarios hasta el momento no se ha realizado en el contexto en que se circunscribe el presente trabajo. El problema de investigación que aquí se aborda, por lo tanto, tiene que ver con el empleo de correferentes de tipo anafórico, su categorización y posterior empleo en la realización de actividades para el aprendizaje de ambas lenguas.

1.2. Justificación

El presente trabajo nace de un interés propio por la lengua francesa y los análisis contrastivos entre lenguas, además de un previo acercamiento con el tema de la correferencia textual por medio del Semillero de Investigación Corpus Ex Machina de la Facultad de Comunicaciones de la Universidad de Antioquia. Adicionalmente, se tienen en cuenta los diferentes estudios previos consultados sobre el tema para delimitar textos narrativos como objeto de estudio, ya que el tema de la correferencia ha sido poco estudiado a partir de los mismos, desde un enfoque contrastivo de lenguas.

1.3. Pregunta de investigación

Se parte entonces de las siguientes preguntas: ¿Cuál es el posible uso de los elementos correferenciales (anáforas) en textos escritos en español y en francés en el siglo XX? y ¿qué relevancia puede tener el análisis contrastivo de dicho uso en ambas lenguas?

1.3. Objetivos

El objetivo general y los específicos en los cuales se fundamenta el presente trabajo de grado, son los siguientes:

1.3.1. Objetivo general

Analizar las posibles características que adquieren las anáforas en la correferencia textual, dentro de las relaciones inter e intra-oracionales en español y en francés, a partir de una muestra de textos literarios en ambas lenguas. Esto con el fin de obtener una caracterización tipológica que permita, a futuro, la realización de actividades didácticas para la enseñanza y el aprendizaje de ambas lenguas, por parte de estudiantes extranjeros.

1.3.2. Objetivos específicos

- I. Exponer diferentes enfoques teórico-metodológicos concernientes a la correferencia textual.
- II. Iniciar la exploración y construcción de una muestra de textos literarios en español y en francés correspondientes al género narrativo, y originalmente escritos a lo largo del siglo XX.
- III. Enunciar las diferentes tipologías pertinentes con relación a las anáforas.
- IV. Caracterizar el uso de anáforas en español y en francés en un par de cuentos seleccionados, para así contrastar las posibles diferencias o similitudes en dicho uso.
- V. Presentar los resultados con respecto a cada lengua, lo que permitirá evidenciar las características particulares que adquieren los elementos de orden correferencial según las lenguas estudiadas en los textos analizados.

Capítulo 2: Antecedentes de la investigación

2.1. Lingüística Textual

Se tiene, en primera instancia, la investigación de Pardo-Abril (1988), “*Coherencia y cohesión: una aproximación al análisis textual*”, que, ya desde finales de los años ochenta se propuso presentar el papel decisivo de la coherencia y la cohesión como características de la estructura textual en la producción y en la comprensión de textos, y que, además, constituyen un recurso objetivo para un análisis sistemático del lenguaje natural (p.58) y para los métodos de enseñanza-aprendizaje de lenguas. Esto, a partir de las propuestas ya estudiadas desde la textolingüística, y, citando a autores como Longacre (1983); Van Dijk (1977, 1980-1983); Halliday y Hasan (1976); R. De Beaugrande (1980); entre otros, (Pardo-Abril, 1988, p. 25).

Por su parte, el mismo Pardo-Abril (1988) realiza un análisis textolingüístico del cuento de Juan Rulfo (1953) “*No oyes ladrar los perros*”, en el que propone la demostración de lo explicado anteriormente. Dicho trabajo es pertinente dentro de la presente investigación, pues, no solo retoma bases teóricas de la lingüística textual, sino que, por medio de un análisis textolingüístico realizado a un texto de carácter literario, ejemplifica cómo llevar a la práctica esta teoría.

2.1.1. Correferencia textual y anáforas

En lo que respecta a la correferencia textual y al tema de las anáforas, el investigador R. Escavy (1999) en su artículo “*La anáfora asociativa como mecanismo de cohesión textual*”, señala la importancia de la relación “referente-antecedente” y la de “parte-todo” para el análisis de la anáfora. El autor en cuestión explica que el carácter de la anáfora asociativa, como referente textual, establece diferentes tipos de asociación que propician la cohesión textual, al igual que la anáfora correferencial. Este trabajo aporta a la investigación, información acerca de la tipología de las anáforas y las características que se deben tener en cuenta para un correcto análisis de la anáfora asociativa tomando como base para esto a D. Meiville (1999) y a D. Apotheloz y J. Reichler-Beguel (1999).

Posterior al estudio de Escavy (1999), G. Peña Martínez (2004) publica “*La anáfora lexical y el discurso de divulgación científica: aproximación contrastiva*”, artículo en el que el autor logra una aproximación general a la anáfora lexical y a su funcionamiento discursivo a partir de textos de divulgación científica. Este trabajo resulta particularmente útil para la investigación aquí presentada, ya que, al igual que éste, propone un enfoque contrastivo español / francés (p. 1), por lo que brinda herramientas teóricas y metodológicas.

Similar a este último trabajo, ya no de carácter contrastivo, sino en lengua portuguesa, R. A. Bertucci (2006) propone en su investigación: “*Anáforas Encapsuladoras: Uma Analise em Textos de Opinião*”, analizar la presencia de un tipo específico de anáfora en un tipo específico de texto, siendo esta vez, anáforas encapsuladoras y textos de opinión, respectivamente. El autor presenta las características de este tipo de anáforas, también llamados encapsulamientos, y señala la especial importancia que estas presentan para el texto de opinión ya que ayudan a la “progresión textual y tópica” (p. 218). Por otra parte, “*Análisis*

de la correferencia para su anotación en un corpus en euskara” (2008), es un segundo trabajo que aporta una mirada desde la lingüística computacional y que continúa con el trabajo de K. Ceberio; I. Aduriz; M. A. Díaz De Ilarraza e I. M. García-Azkoaga (2007). Al igual que en la investigación señalada anteriormente, el trabajo de P. M. Martínez, este estudio del 2007 resulta pertinente tanto teórica, como metodológicamente, pues puede ser tomado como referencia para el etiquetado de la anáfora y de la correferencia textual. Este artículo presenta la relación existente entre la correferencia y la anáfora, y el estado actual de los estudios relacionados con el etiquetado sistemático de estas en el ámbito del procesamiento del lenguaje natural (p.3). Tras etiquetar un corpus identificando las redes referenciales, los investigadores esperan que dicho trabajo facilite la creación de un sistema que realice el proceso de etiquetado de la anáfora y la correferencia por sí solo.

2.1.2. Mecanismos de cohesión textual

En “*Mecanismos de Cohesión en un texto de Italo Calvino: propuesta para un análisis*” de Bruni, A. (2004), al igual que en el trabajo de Pardo-Abril (1988), mencionado anteriormente, se propone un análisis textual, basado en el modelo de Raffaele Simone (1990), de un cuento de Italo Calvino y de los diferentes mecanismos de cohesión textual que en él actúan. A partir de dicho trabajo es posible identificar las diferentes cadenas y elementos cohesivos que permiten entender cómo se estructura un texto (p. 21). Al igual que el ya citado trabajo de Pardo-Abril, este trabajo es pertinente con respecto a la presente investigación, pues, no solo retoma diversas bases teóricas de la lingüística textual, sino que, por medio del análisis de un texto literario, aplica, además, diversos conceptos teóricos subyacentes a aquel campo del conocimiento.

2.2. Enfoque contrastivo entre lenguas

Con enfoque contrastivo entre lenguas, además del ya citado trabajo de G. Peña Martínez (2004) se tiene también el trabajo de Cao, Yufei. (2014), esta vez en español / chino, “*Clasificación nominal y anáfora: comparación entre español y chino*” donde compara la clasificación nominal desde la perspectiva de la anáfora en cada lengua. Y explica que, en ambas lenguas, la información clasificatoria puede contribuir a la identificación del referente textual, algunas veces anafóricos (p.3). Este artículo puede aportar a la investigación, en la medida en que no trata un tipo específico de anáfora, sino que muestra el carácter anafórico de la información en la clasificación nominal; además, por su enfoque contrastivo, sirve de ejemplo metodológico.

2.3. Lingüística Computacional y resolución anafórica

Desde la perspectiva de la lingüística computacional y el tema de la anáfora, tenemos el trabajo de P. M. Martínez-Barco (2001), quien en su tesis doctoral “*Resolución computacional de la anáfora en diálogos: estructura del discurso y conocimiento lingüístico*”, desarrolla un sistema de taxonomía de diálogos y de resolución de anáforas en diálogos (ARIADNA). Este estudio tiene la particularidad de proveer información acerca de las anáforas desde una perspectiva computacional, lo que resulta pertinente desde el aspecto teórico además del práctico-metodológico, pues brinda herramientas para el etiquetado automático de las anáforas, hasta el momento poco conseguido por el carácter contextual de las anáforas y la dificultad para su categorización según las diferentes tipologías existentes.

2.4. Didáctica de las lenguas y correferencia textual

L. Lundquist (2014) en “*La didactisation du Français Langue Étrangère: discours, genre et type de texte*”, da continuidad a investigaciones previas realizadas en compañía del lingüista suizo J.-M. Adam. En este trabajo, la autora danesa propone un enfoque del análisis textual que se centra en la enseñanza y el aprendizaje del Francés como Lengua Extranjera (FLE) y que toma el análisis fundamentado en textos académicos de disciplinas especializadas y con fines educativos. Retomando los trabajos previos, explica su propuesta de una tripartición de los textos en: discursos (determinado por un sistema social particular), géneros (definido por el conjunto de características recurrentes que se presentan en situaciones comunicativas idénticas) y tipos de texto (delimitado por sus especificidades lingüísticas). Propuesta de clasificación que deriva de diferentes ejercicios de lectura, y que, según la autora, permite identificar las características de los textos académicos, que crean pistas de coherencia, y para precisar en diferentes estrategias de lectura para aprendientes de FLE.

2.5. Aprendizaje de Lenguas Asistido por Ordenador y correferencia textual

Finalmente, se tiene un trabajo que retoma investigaciones previas, y que reúne elementos de la lingüística textual contrastiva, de FLE, de ALAO (Aprendizaje de Lenguas Asistido por Ordenador), y de la lingüística computacional, es un trabajo que resulta clave para la investigación en curso, y corresponde a la investigación realizada por Molina Mejía (2015): “*ELiTe-[FLE]2 : Un environnement d’ALAO fondé sur la linguistique textuelle, pour la formation linguistique des futurs enseignants de FLE en Colombie*”. En esta tesis, se

desarrolla un dispositivo informático como herramienta para docentes de FLE y se aportan diversas bases teóricas y fuentes bibliográficas sobre los temas anteriormente mencionados. Por otra parte, presenta una tipología completa de las anáforas que se utilizan en el campo de la correferencia textual en francés, lo que resulta de gran utilidad al presente trabajo, sobre todo en lo que respecta al análisis de las mismas en los textos elegidos como objeto de estudio.

2.6. Conclusión acerca de los antecedentes de investigación

Luego del rastreo bibliográfico brevemente descrito en las secciones anteriores, se puede concluir que, no solo ha habido incursiones en las diferentes temáticas y propuestas derivadas de la lingüística textual y del análisis textual desde la década de los ochenta, sino que son temáticas que aún tienen vigencia y que reclaman su importancia ante la atención de los diferentes investigadores contemporáneos. También cabe resaltar que, a pesar de existir numerosas investigaciones con ambiciones similares a las que apunta alcanzar el presente proyecto de investigación, queda por hacer un análisis de la correferencia textual, dentro de las relaciones inter e intra-oracionales en textos literarios y que tenga carácter contrastivo español / francés.

Capítulo 3: Marco teórico

3.1. La Lingüística Textual

3.1.1. Definición

La textolingüística es, para Bernal Leongómez (1986), la tercera y última, pero no menos importante, de las ramas que propone como principales dentro de la lingüística del siglo XX, junto con la lingüística generativista y la estructuralista. El autor aclara que la lingüística no se reduce sólo a esta tríada, y propone el nacimiento de la lingüística textual como una extensión de la semántica generativa, cuyo desarrollo se ha dado, sobre todo, en las lenguas inglesa y francesa. En la introducción de su *“Antología de lingüística textual”*, el autor se dispone a contextualizar las pretensiones que, ya desde sus inicios en las últimas décadas del siglo pasado, se planteaba la lingüística textual.

3.1.2. Génesis del campo y objeto de estudio

Los modelos de lenguaje que se tenían en el siglo XX, propuestos por el estructuralismo y el generativismo, tomaban la oración como unidad de trabajo. Por su parte, la lingüística textual propone el texto y no la oración, como su objeto de estudio. Dentro de dicha disciplina, normalmente existen dos corrientes, una de orden pragmalingüístico y otra de orden semántico e interpretativo; siendo la primera visión la que tiene por objeto su aplicación a la enseñanza y aprendizaje de lenguas. Si bien la lingüística rusa desde 1950 trabajó diferentes cuestiones del texto, Bernal propone la década de 1970, más concretamente en Alemania, el

origen tentativo de la lingüística textual, pues, es en esta década que se desarrollan en Europa los principales postulados de la nueva disciplina, afirma el autor.

La lingüística textual es entonces, para el autor, aquella que “propone el texto como unidad” (Bernal, 1986, p.17) y que, además, lo toma, en su totalidad, como unidad mínima de estudio. Si bien la definición de texto acarrea algunas dificultades, para el autor, también explica que, en términos generales, se entiende como texto aquel que es “resultado de la actividad lingüística del hombre, ha de tener incuestionablemente una específica intención comunicativa, y, por último, ha de explicitarse con suficiencia el contexto en el cual se produce” (p. 17).

Adicionalmente, la textolingüística propone un enfoque interdisciplinar por lo que se plantean como puntos cruciales de la nueva disciplina, el análisis del discurso, la pragmática, la teoría de la acción y el interés en la dualidad anáfora-catáfora que ya había sido trabajada anteriormente por la lingüística transformacional, entre otros temas. En sus comienzos, la textolingüística retoma la dicotomía preexistente competencia/actuación, lengua/habla, reformulándola a la de texto/discurso, dando a su vez paso al nacimiento de disciplinas como la sociolingüística y la pragmática (Bernal, 1986).

Jean-Michel Adam, en su artículo “*L'émergence de la Linguistique Textuelle en France: entre perspective fonctionnelle de la phrase, grammaires et linguistiques du texte et du discours*” (2010), habla sobre la influencia de *L'école de Prague* (el Círculo de Praga) con su sintaxis funcional para la emergencia de la lingüística textual en Francia. Inicialmente revisa 25 años de trabajos de diferentes investigadores, enmarcados entre los años 1975 y 2000, en los que se van introduciendo conceptos relevantes para la presente investigación.

Adam resalta la gran importancia de Denis Slakta en dicho campo, y cita, a través de éste, a Van dijk, “*La différence avec les grammaires de phrase est que les dérivations ne se terminent pas sur des phrases simples ou complexes, mais sur des n-tuples ordonnés de phrases, c’est-à-dire sur des séquences*”, (van Dijk 1973b:19 ; citado y traducido por Slakta 1977:15). Aquí se concibe el texto como la sumatoria de n cantidad de frases.

Como no basta la disociación/integración que permite identificar otras unidades lingüísticas, ya que no se puede aplicar la conmutación de una frase por otra, salvo que estén en un mismo contexto; ya que, diferente a lo que ocurre con la frase, sintagma, signo y morfema, no se pueden descomponer frases en el texto y aplicar los mismos procedimientos.

Esto, explica Adam, conduce a Slakta a querer mostrar cómo se dan las relaciones de dependencia entre las frases, por lo que, debido a su concepción de texto, no existe la frase sin relación con otras frases, lo que implica y hace necesarias las nociones de cohesión y progresión entre aquéllas, es decir, hay una relación intratextual (Adam, 2010).

Se define la frase como «*Unité isolable mais non isolée*»¹ (p.23), además, especifica que para Martin, la frase cumple así, no sólo con exigencias gramaticales y semánticas, sino que también se adaptan al contexto dándole cohesión (p. 28). Y es el componente pragmático, el que asegura el paso de la frase a enunciado y el que permite que haya coherencia.

La importancia de Slakta, para Adam, es principalmente debido a que desarrolla el análisis de las relaciones “transfrásticas” (o transoracionales), por ejemplo, conectores y correferentes. El autor concluye esta primera parte del artículo resaltando la importancia de las investigaciones que se traducen y retoman posteriormente por otros investigadores, como Fuchs en 1985. Y, en la segunda parte del trabajo enlista y resalta una serie de diferentes

¹ Traducción propia realizada a partir de Adam (2010, p.23): “Unidad aislable pero no aislada”.

trabajos que considera importantes en la medida que aportan diferentes criterios para que haya continuidad textual como, por ejemplo, la isotopía, la anáfora y la comunidad proposicional, entre otros, como el trabajo realizado por Martin (1992).

3.2. Cohesión y coherencia textual

En el curso que lleva por nombre: “*Linguistique textuelle et didactique de l’écrit*” de J.-M. Colleta y G. Karcher (2012), se definen varios de los conceptos fundamentales que han sido retomados en el transcurso de la presente investigación, a saber: el de texto, los tipos de texto, la referencia y la correferencia textual y, además, se proponen diferentes formas de abordar los diversos tipos de texto, entre ellos los textos literarios, como herramientas para la didactización del francés como lengua extranjera (FLE).

Si bien, la lingüística textual referencia el texto y la frase como unidades diferentes, también demuestra las sólidas relaciones que se dan entre ambas. Inicialmente, el libro de curso establece que la frase es al texto como la palabra es a la frase, pues se dan relaciones entre los dos tipos de unidades, donde de un conjunto del primer elemento en cuestión en ambos casos se constituye el segundo. Así, de un conjunto de palabras surgen frases y de conjuntos de frases se constituyen textos, convirtiéndose en unidad.

La equivalencia que dice que el texto es aquel tipo de escrito y el discurso es aquel tipo de unidad oral, trae muchas dificultades según los autores, por lo que retoman a Slakta (1975), quien establece que el discurso es el texto más las condiciones de producción; y a su vez, el texto es el discurso sin dichas condiciones. Y, a partir de esta dicotomía, establece una analogía con relación a la distinción frase/enunciado, en la que el primer término se refiere a la teoría y el segundo a la práctica. En otras palabras, es esta diferenciación la que

separa a la lingüística textual del análisis del discurso. Parafraseando a Adam (2011): la lingüística textual no se encarga de estudiar la enunciación. La lingüística textual estudia, por lo tanto, el texto sin tener en cuenta las condiciones de enunciación del mismo.

El curso propone una serie de tipos de texto y, tras examinar varias definiciones de texto, propone entender al texto como un «*objet langagier qui peut se présenter sous deux formes: dialoguée et monologuée, et dans les deux modalités de l’oral et de l’écrit*»² (Coletta; Karcher, 2012, p.19).

3.2.1. Referencia y correferencia textual

Según el modo de enunciación, el texto se divide entre texto dialogado o monologado según sea colectivo o individual respectivamente. En textos dialogados, para los autores, se habla de coherencia pragmática, que reposa en la concatenación de turnos de habla de los participantes, y esta se da dentro de la misma dinámica del diálogo; en los textos monologados se habla de coherencia temática, que reposa en los propósitos del locutor.

Si bien hay diferentes tipologías textuales según las cuales se clasifica al texto bajo diferentes criterios, también se pueden identificar diferentes secuencias textuales específicas que componen la estructura subyacente del texto monologado. Para los autores, “la característica principal de un texto es la coherencia” (Coletta y Karcher, 2012, p.53) y es en este sentido que se hace necesario examinar el contenido del texto para identificar estas secuencias y sub-secuencias al interior del mismo que permiten esa construcción coherente, a saber, los modos de enunciación, tiempos y modos verbales, anáforas, conectores, etc...

² Traducción propia de Coletta y Karcher (2012, p.19): “objeto lingüístico que se puede presentar de dos formas: dialogado y monologado, y bajo dos modalidades de lo oral y de lo escrito.”

Se cita, por otra parte, a Lundquist (1980) quien dice que la coherencia hace parte de la competencia lingüística pues, el hombre es capaz de producir textos a partir de secuencias de frases coherentes y puede decidir si una secuencia de frases es coherente o no, si puede ser considerado como un texto o no (Coletta y Karcher, 2012, p.53). Adicionalmente, los autores relacionan la coherencia con la interpretabilidad y consideran que la cohesión responde a una determinada organización textual, y, es esta última la que le permite ser un objeto organizado y completo. Así, según ellos, la referencia ha tenido un lugar importante dentro de las preocupaciones de muchos pragmáticos y lingüistas al tratarse de una reflexión que concierne tanto a la semiótica, relaciones entre sistemas de signos, como a la pragmática. Retoman a Searle (1982), quien “hizo de la relación lenguaje-mundo un criterio definitorio del acto del lenguaje.” (Coletta y Karcher, 2012, p.63) y señalan el principio semiótico, según el cual “a cada palabra se asocia una cosa en el mundo” (Coletta y Karcher, 2012, p.63).

A partir de conceptos abstractos, la relación referencial, por ejemplo, es interna. Así, el texto debe construir referentes internos a él, a los que pueda retomar posteriormente como anáforas y asegurar la continuidad referencial. De esto se desprende que un referente textual puede ser un elemento o un grupo de elementos que remiten a algo que se conoce del mundo. Generalmente constituye un grupo nominal (GN).

La deixis, por su parte, es una relación exofórica, es decir, remite a elementos externos a la situación, relaciona la palabra y el referente, que es identificable en el contexto. La anáfora, en cambio, es una relación endofórica, es decir, interna al mismo texto, está ya representada en el contexto por otra palabra. Según Perdicoyanni-Paléologou (2001, p.59), la deixis es un procedimiento espacio-temporal que juega un rol fundamental dentro de la

enunciación, ya que esta permite que los protagonistas de la historia puedan situarse en un lugar y tiempo determinados.

Los autores citan de nuevo a Lita Lundquist (1980), quien afirma que, mientras la referencia relaciona el texto con una realidad, la correferencia establece una relación entre elementos textuales. Así, la referencia es externa y origina el tema, y la correferencia, por su parte, lo hace evolucionar en una estructura temática que puede ser intra o interoracional. Y se aborda, por otra parte, a Martín Rojo y Meeuwis (1993), quienes afirman del referente que: “En español, el referente del sujeto (...) puede tener distintas realizaciones (...) puede realizarse, entre otras maneras, mediante un nombre, mediante una oración de relativo, mediante un pronombre indefinido o demostrativo, mediante un pronombre personal, mediante cero” (1993: p.91).

3.2.2. Anáforas y catáforas

Un correferente textual puede ser un elemento o grupo de elementos que remiten a un referente o a otro correferente. Puede ser un grupo nominal o un pronombre, además, según la dirección en que se encuentre en el texto, puede ser:

- **Anáfora**, si retoma un referente previamente posicionado en el texto.
- **Catáfora**, si anuncia un referente posterior en el texto de una forma más precisa.

Por ejemplo, un nombre después de pronombre o un grupo nominal con más información se usa, generalmente, al inicio de un texto o secuencia textual. Por otra parte, Cuenca (2010) afirma acerca de la catáfora que esta es: “(...) un mecanismo relativamente poco frecuente y muy condicionado por factores que, de algún modo, permiten hacer una

hipótesis o mantener un referente genérico indefinido hasta la aparición de un elemento textual de referencialidad clara y alto grado de definitud” (p.44).

Coletta y Karcher (2012) proponen en su obra dos grandes grupos principales de anáforas que se dan a nivel de la lengua francesa, y estos se distinguen según su función dentro del texto.

1. En el primer grupo reúne la anáfora pronominal total, en opinión de los autores la más recurrente, y la nominal directa. Ambas aseguran la continuidad referencial en el texto creando una red anafórica.
2. En el segundo, se encuentran la anáfora nominal asociativa o nominal indirecta, asociada a la isotopía semántica, ya que permite la construcción de referentes y la progresión de la información de los mismos; la anáfora pronominal parcial y la pronominal integrante, que no sólo aseguran la continuidad referencial en el texto, sino que también participan en la creación de nuevos referentes que se pueden retomar posteriormente. Estas permiten la creación de redes o cadenas referenciales.

Adicionalmente, estos mismos autores subdividen las anáforas nominales directas en cinco subcategorías, a saber: fiel, infiel, determinativa, metafórica y conceptual; también categorizan anáforas verbales, adjetivales, adverbiales y las de segundo orden, que responden más a términos gramaticales que de funcionalidad dentro del texto.

En la tabla 1 podemos apreciar un resumen de la tipología propuesta por ambos autores y la traducción que se ha realizado en el presente trabajo de la terminología en francés.

Tabla 1. Tipología anafórica de Coletta y Karcher (2012), traducción propia.

Primer grupo		Segundo grupo		Otras
<p>Nominal directa Puede ser un nombre o grupo nominal que retoma el referente todo y nada más que el referente.</p> <p>Fiel Retoma idénticamente el referente, con las mismas palabras.</p> <p>Determinativa Se retoma el referente usando artículos determinantes.</p> <p>Infiel Reformula el referente, retoma con sinónimos o palabras diferentes.</p> <p>Metafórica Caso especial de la infiel, reformulación por analogía.</p> <p>Conceptual Sintetiza, resume, nominaliza.</p>	<p>Pronominal total Pronombre que retoma el referente y solo el referente.</p>	<p>Nominal asociativa o indirecta Al igual que la pronominal parcial, también participa en el surgimiento de nuevos referentes. No es una retoma exactamente, funciona bajo una relación de hipo-hiperonimia.</p>	<p>Pronominal parcial Retoma solo una parte del referente.</p> <p>Pronominal integrante Resultado de la fusión de dos referentes diferentes.</p>	<ul style="list-style-type: none"> • Verbal • Adverbial • Adjjetival • Índices de segundo orden (obedecen a reglas de tipo gramatical, como desinencias verbales o concordancias sujeto - adjetivo) <p>(Coletta, Karcher, 2012, pp. 70-72)</p>

Por otro lado, M. J. Cuenca (2010) en *Gramática del texto*, también trabaja la anáfora como un mecanismo fórico de referencia que como, según explica, “(...) casi siempre, el antecedente es anterior al elemento que requiere referencia, cosa del todo lógica si pensamos que establecemos el significado y después lo retomamos” (p.43). Y, además, al interior de dichos mecanismos de referencia, la autora propone tres grupos principales para el español, el de la anáfora gramatical plena, que consta de pronombres y posesivos; el de la anáfora léxica, que se manifiesta por medio de repeticiones y sinónimos discursivos; y el de mecanismos como la repetición, reiteración y la elisión, que alternan con las anáforas gramaticales. Es a la elipsis a la que acuña el nombre de *anáfora cero*, así:

(...) Se puede entender la elipsis como una anáfora cero, es decir, como una pronominalización en la que el pronombre no tiene manifestación fonética, más allá de una pausa en algunos casos. En español, la elisión es casi obligatoria cuando tenemos un referente conocido en posición de sujeto y no hay otro factor (posible ambigüedad, énfasis, etc.) que requiera que aparezca allí un pronombre o un sintagma nominal; pero existen otras elisiones que tienen un carácter más estilístico (Cuenca, 2010, p.45).

Como se pudo apreciar a lo largo del presente capítulo, se han presentado los diferentes conceptos que nutren este trabajo de grado. En el siguiente capítulo se presentará el marco metodológico.

Capítulo 4: Marco metodológico

4.1. Etapas del marco metodológico

Tras exponer y esbozar algunas de las diferentes teorías y enfoques concernientes a la correferencia textual, anteriormente señalados; al igual que las diferentes tipologías existentes con relación a las anáforas, elementos correferenciales a analizar; se pasa, a continuación, a explicar los aspectos metodológicos que soportan el presente trabajo. En este sentido, se pretende seleccionar una muestra de textos escritos en español y en francés con el fin de realizar el análisis de diferentes correferentes textuales anafóricos como los que han sido expuestos en el capítulo anterior.

A continuación, se enuncian las diferentes etapas relacionadas con la metodología a emplear en este trabajo.

4.1.1 Constitución de la muestra

En un primer momento, se hizo uso de herramientas proporcionadas por el campo de la lingüística de corpus como un eje importante en la presente investigación con el fin de establecer una metodología de trabajo, tanto para la construcción del corpus a analizar, como para su posterior análisis. Es, por esta razón, que se propuso una revisión histórica a partir de Bolaños (2015) en *La lingüística de corpus: perspectivas para la investigación lingüística contemporánea*, tras revisar la discusión sobre la definición de *corpus* que se da tanto en Bolaños (2015) como la discusión que se da en *La utilización de la lingüística de corpus en la enseñanza de español* (da Silva, 2006), donde el autor afirma que un corpus ha de cumplir con los siguientes requisitos:

(...) son necesarios algunos criterios para la formación del corpus: I) el corpus debe estar compuesto por textos auténticos en lengua natural, II) el contenido debe ser seleccionado a fin de garantizar que el corpus tenga sus características propias y III) el corpus debe ser representativo (da Silva, 2006, p.221).

Y en vista de que el marco del presente trabajo se consideró un análisis a partir de dos textos, se tienen en cuenta solo los dos primeros requerimientos para la selección de los mismos. En este punto se hizo necesaria la reconsideración del concepto de *corpus* desde una línea diferente a la de la lingüística de corpus, por lo que se retoma la definición propuesta por López Morales (1994): “Llamamos *corpus* al conjunto de textos, sea cual sea su origen, que constituye el material de análisis de una investigación dada. Los textos pueden ser orales o escritos, y estos últimos, impresos o manuscritos” (p.95).

Adicionalmente, para la constitución de la muestra de dos textos de tipo narrativo, se retoma la definición dada por Adam y Lorda (1999) según la cual “Un texto narrativo (...) es una unidad compleja y profundamente heterogénea, compuesta de momentos narrativos, descriptivos y dialogales más o menos puros, que denominamos SECUENCIAS” (sic) (p.35); y que adicionalmente hubiesen sido publicados durante el siglo XX.

En un primer momento de la constitución del corpus, se inició el rastreo de textos que cumplieran dichas condiciones y que, además, tuvieran una extensión aproximada similar, esto con el fin de trabajarlos en su totalidad y no fragmentados. Esta forma de proceder debería permitir que los documentos fueran manipulados lo menos posible, en cuanto al tipo y la cantidad de anáforas en cada texto y en cada lengua. Y, actuando de esta forma, se lograría que el análisis y los respectivos resultados no pudieran ser explicados por variables externas de la constitución de la muestra como podrían serlo: la extensión o el tipo de texto

o la temporalidad del uso de la lengua. Adicionalmente, se tienen en cuenta para el rastreo versiones digitales de los textos por la facilidad que suponen para el manejo de los datos y para evitar posibles modificaciones al texto o errores tipográficos en el proceso de transcripción.

4.1.2 Muestra tentativa

Teniendo en mente los elementos señalados en el apartado anterior, a continuación, se presentan los textos que se consideraron al inicio del presente trabajo para conformar la muestra a trabajar, dividida en dos apartados, uno para el español (TeNEsp) y otro para el francés (TeNFra):

a. Textos narrativos en español

Código	Título	Autor	Año	Tokens sin etiquetas
TeNEsp-1	<i>Barrabás</i>	Arturo Uslar Pietri	1928	1900
TeNEsp-2	<i>La excavación</i>	Augusto Roa Bastos	1953	2119
TeNEsp-3	<i>El carrusel</i>	Julio Ramón Ribeyro	1977	3449

b. Textos narrativos en francés

Código	Título	Autor	Año	Tokens sin etiquetas
TeNFra-1	<i>Récit de l'étudiant</i>	Jean Lorrain	1900	1730

TeNFra-2	<i>Les deux mendiants</i>	Charles-Louis Philippe	1910	1151
TeNFra-3	<i>Les visites</i>	Victor Du Bled	1910	2431

Finalmente, la conformación de la muestra, objeto de este trabajo, se hizo a partir de los textos TeNFra-1, *Récit de l'étudiant* (Anexo B), y TeNEsp-1, *Barrabás* (Anexo A), ambos correspondientes al género textual narrativo, originalmente publicados en el siglo XX, y escritos por un autor francófono y otro hispanohablante respectivamente, es decir, que no se tuvieron en cuenta textos surgidos de traducciones para la investigación. Si bien, los textos contemplados inicialmente como posibilidades cumplían con los requisitos propuestos, se escogen estos dos textos ya que son los que tienen una extensión similar (de aproximadamente 1900 y 1730 palabras o *tokens*, sin incluir la puntuación) y cumplen con el hecho de que en sus respectivas publicaciones no se hayan muy alejadas la una de la otra, sobre todo, bajo una concepción lineal del tiempo. Esto se hizo con el fin de propiciar la comparación y caracterización del uso de las anáforas en cada lengua de forma equitativa.

4.1.4. Etiquetado del corpus

En el caso de la presente investigación se ha optado por trabajar, a partir de la tipología anafórica propuesta por Coletta y Karcher (2012), seleccionado las subcategorías fiel, infiel y determinativa de la anáfora nominal directa y las subcategorías parcial, total e integrante de las anáforas pronominales. Esto se debe al hecho, como señalan los autores, que se trata de las más comunes y las que permiten dar una continuidad secuencial al texto. Adicionalmente, se propone retomar el concepto de anáfora cero, empleado por Cuenca

(2010), con la finalidad de trabajar la elipsis en el análisis de los textos seleccionados, más específicamente el uso de sujeto tácito; ya que al igual que las anteriores cumple una función importante dentro del texto, y esto es que permite el desarrollo del texto y da continuidad secuencial; y, cuyo uso, como explica la autora, es bastante recurrente en español. A continuación, se pasa a presentar la tipología implementada para la caracterización de los elementos anafóricos encontrados en los textos narrativos estudiados:

Tabla 2. Tipología anafórica utilizada en el presente trabajo.

Nominal Directa	Pronominal	Cero
	Total	
Puede ser un nombre o grupo nominal que retoma el referente todo y nada más que el referente.	Pronombre que retoma el referente y solo el referente.	Elipsis, "proceso de economía discursiva que alterna con la pronominalización según el contexto sintáctico" (Cuenca, 2010, p. 44).
Fiel	Parcial	
Retoma idénticamente el referente, con las mismas palabras.	Retoma solo una parte del referente.	
Determinativa	Integrante	
Se retoma el referente usando artículos determinantes.	Resultado de la fusión de dos referentes diferentes. (Coletta, Karcher, 2012, pp. 70-71)	
Infiel		
Reformula el referente, retoma con sinónimos o palabras diferentes.		

Ya, desde la constitución de la muestra tentativa, se procede con la lectura de los textos y, gracias a esto se realiza un rastreo preliminar de los referentes y correferentes que se encontraban en los mismos; a partir de dicho rastreo se opta por resaltar por colores en un archivo de *Word* las cadenas correferenciales (con sus respectivos referentes y correferentes

anáforicos), como puede observarse a modo de ejemplo a continuación (Figura 1):

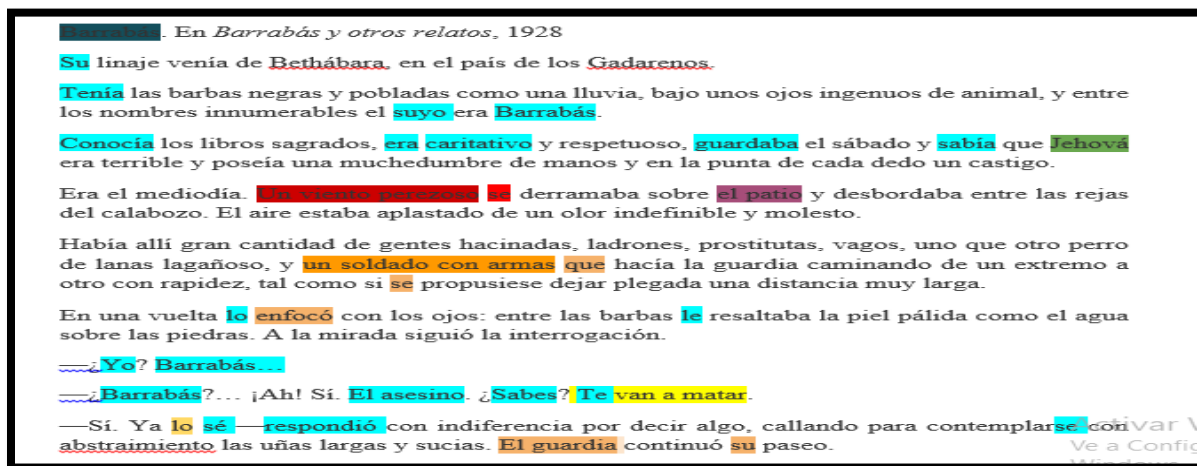


Figura 1. Rastreo preliminar de anáforas.

El etiquetado de anáforas a partir de herramientas computacionales que permitan su correspondiente tratamiento de formas automáticas o semiautomáticas aún es un reto para la lingüística computacional por las diferentes dificultades que representa, como ya ha sido estudiado en investigaciones anteriores por lo que, en esta etapa, se inicia el etiquetado manual de las unidades correferenciales anteriormente mencionadas y presentes en cada texto a partir de herramientas ofimáticas computacionales, como lo son Word y Excel, donde a partir de relecturas se categoriza cada anáfora según la tipología propuesta en el marco metodológico.

En este punto se hizo necesaria la respectiva revisión teórica de los diferentes conceptos y la discusión de qué elementos considerar como constituyentes del texto, como lo son, por ejemplo, el título, que se presenta como el primer referente de un texto, indicando al lector el inicio de, en este caso, una narración y cumpliendo las diferentes funciones explicadas por de la Fuente González en *Funciones de los títulos en la decodificación lectora* (1997). Adicionalmente, en ambos textos trabajados los títulos introducen personajes

principales de cada narración, por lo que se incluyen estos como elementos de análisis, de lo contrario, muchos elementos encontrados hubieran tenido que ser reinterpretados como información nueva en el texto, es decir, como catáforas.

Por otra parte, las unidades encontradas que se prestaran para diferentes interpretaciones, fueran ambiguas, o que a su vez remitieran a un primer elemento no trabajado, no se tomaron en cuenta para el análisis. Tal fue el caso de algunas unidades que, si bien, retomaban información, no concordaban en género y número con el elemento referencial retomado, por lo que podrían referir a diferentes referentes según la interpretación, a manera de ejemplo, tenemos los siguientes casos:

- “*No sé si he soñado, si estoy muerto, o si es mi sangre y mi voz la que le habla.*” Aquí, no hay concordancia de número entre “*mi sangre y mi voz*” y “*la que habla*”, por esta razón no es posible identificar a cuál de los dos elementos se refiere.
- En TeNFra-1 sucede algo similar con el pronombre *on*, el cual según el contexto puede ser interpretado como impersonal (*se*) o como primera persona del plural (*nosotros*), por lo que solo se categoriza cuando se considera que solo hay un posible referente al que puede referir.
- “—¿*Por qué no dijiste eso a los jueces?* —*No me lo preguntaron.*” Donde “*lo*” es correferente de “*eso*”, anáfora conceptual, no trabajada en el presente trabajo, por lo que sería inconsecuente trabajar una cadena correferencial sin su referente principal.

Con el fin de efectuar este proceso, se hicieron necesarias diversas relecturas de los textos y de las tipologías anafóricas, por lo que si bien inicialmente se pretendió asignar a los códigos correspondientes a cada referente y a sus respectivos correferentes una numeración que respondiera al orden de aparición dentro del texto, siendo el primero C1, el segundo C2

y así sucesivamente, la numeración que inicialmente respondía a estas intenciones fue modificada a partir de las posteriores lecturas y revisiones de los aspectos teóricos, pues se decide eliminar algunos elementos marcados anteriormente como referentes o anáforas, o, por el contrario, incorporar algunos otros que inicialmente no se categorizaron, por lo que la numeración responde a esta situación; no siendo en todos los casos numéricamente ordenada según la aparición en los textos.

Finalmente, la categorización se llevó a cabo en un archivo de Excel, en el que se importan los fragmentos de los textos que se clasifican en referentes, tipos de referentes, o anáforas y se categorizan según lo anteriormente expuesto (Anexo A y B). Dichos fragmentos, se extraen con el fin de tener presente el contexto para la clasificación, y fueron únicamente modificados al agregarles paréntesis. Esto permitió, en la medida de lo posible, que se pudieran indicar los elementos que no fueron considerados como parte del correferente a analizar como, por ejemplo, las preposiciones de los elementos nominales o pronominales, verbos o complementos circunstanciales; también se indica entre paréntesis los dos elementos que hacen parte de los respectivos títulos.

Por otra parte, el etiquetado de los referentes también hizo necesaria la revisión teórica del tratamiento que diferentes autores le dan al referente como unidad, como la de Martín Rojo y Meeuwis (1993) en la que contemplan, al igual que Cuenca (2010), no solo la anáfora cero sino también los referentes tácitos. Los autores explican que, en español y en las lenguas de sujeto tácito, como las denominan, hay una alternancia de uso entre el sujeto pronominal y la elipsis o *anáfora cero*, esto se puede evidenciar en TeNEsp-1, donde se encuentran dos elementos, que a partir de lo explicado por Martín Rojo y Meeuwis, serían referentes tácitos.

Tras haberse etiquetado y caracterizado los diferentes usos anafóricos en cada texto, se procedió al análisis contrastivo de las posibles diferencias o similitudes en dicho uso y, a continuación, se presentan los diferentes resultados con respecto a cada lengua, lo que permitirá evidenciar las características particulares que adquieren los elementos de orden anafórico según las lenguas estudiadas.

Capítulo 5: Análisis y resultados

Una vez que los textos han sido etiquetados y revisados se procede a su respectivo análisis y conteo; esto con el fin de realizar un estudio estadístico del empleo de anáforas en la muestra seleccionada. Como bien se ha expresado en los capítulos anteriores, en el presente trabajo se propuso categorizar las unidades anafóricas encontradas en dos textos narrativos; motivo por el cual se hizo necesario la identificación y categorización de los respectivos tipos de referentes correferenciales de dichas unidades.

Cabe resaltar que la parte cualitativa del presente análisis en gran parte consistió en la identificación y categorización de cada uno de los 809 elementos correferenciales trabajados en total en ambos textos entre referentes y anáforas de la forma que se observa en los anexos A y B y se explica a continuación. La muestra de TeNEsp-1 fue constituida por 500 elementos, 47 de los cuales fueron referentes; a su vez, TeNFra-1 contó con 309 elementos de los cuales 43 fueron referentes posteriormente retomados total o parcialmente.

En primer lugar, en TeNEsp-1, se encontraron 47 referentes que posteriormente son retomados en el texto y que así dan inicio a diversas cadenas correferenciales. Como puede observarse en la Tabla 3, 40 referentes son de tipo nominal (Nom), tres nombres comunes, tres nombres propios y 34 grupos nominales (GN), siendo el grupo más representativo. En los demás hallazgos se trataba de casos aislados, siendo estos: tres referentes pronominales, dos referentes tácitos, uno adjetival y uno verbal. En este punto, como ya se mencionó anteriormente, se hizo necesaria la revisión del concepto de referente tácito, ya que en el texto se presenta la caracterización de un personaje colectivo que no es introducido de forma directa en la narración.

Tabla 3. Referentes TeNEsp-1

Referentes TeNEsp-1			
Ref Nom	40	Nom	3
		NP	3
		GN	34
Ref Pron	3		
Ref Adj	1		
Ref Vbal	1		
Ref Tácito	2		
Total	47		

También, al igual que otros autores en investigaciones anteriores, como es el caso de Muñoz (2017), se encontraron cuatro tipos de referentes en los textos, a saber:

a) Referente mencionado: solo se introduce el referente y no es retomado en la narración; b) Referente mencionado y referido: se introduce el referente y se retoma a lo largo del relato con distintos elementos correferenciales; c) Referente solo referido: no se introduce el referente en forma explícita, pero sí se señala a través de elementos correferenciales; d) Referente mencionado y en elipsis: se introduce el referente pero se retoma a través de la marcación morfológica del verbo, tomando forma de sujeto tácito. (Muñoz, 2017, p.33).

De esta lista, el *referente mencionado* no se toma en cuenta en la presente investigación, pues no es pertinente si tenemos en cuenta los objetivos de la misma; esto debido al hecho de que un referente que no es posteriormente retomado, no participa en ninguna relación correferencial. Por otra parte, se pueden observar en la Tabla 4, los valores porcentuales correspondientes a cada una de las tres categorías de los que cabe aclarar que, no necesariamente los referentes mencionados y retomados en elipsis son únicamente y exclusivamente retomados en dicho fenómeno textual, sino que además pueden ser empleados de forma anafórica, por medio de otros mecanismos.

Tabla 4. Tipos de referentes TeNEsp-1.

Referentes TeNEsp-1		
Mencionado y referido	26	55,31%
Mencionado y en elipsis	19	40,42%
Solo referido	2	4,25%
Total	47	100%

En segundo lugar, en TeNFra-1 como aparece cuantificado en la Tabla 5 se encontraron un total de 43 referentes, de los cuales 37 son nominales (todos son de tipo GN), un pequeño grupo de cinco referentes pronominales y un referente adverbial. Llama la atención que, si bien en ambos textos los dos grupos predominantes son el de referentes nominales y el de referentes pronominales en segundo lugar, en ambos casos, se presentan referentes únicos de diferentes clases gramaticales que no se vuelven a presentar, como el caso del referente adjetival y verbal en TeNEsp-1 y el referente adverbial en TeNFra-1. Y, en este último, a diferencia del primero, el 100% de referentes que se presentan son de tipo mencionado y referido, ya que no se presentan casos ni de referentes tácitos, ni de anáforas cero, que serían las que determinarían la categoría de solo referido y de mencionado y referido en elipsis respectivamente.

Tabla 5. Referentes TeNFra-1.

Referentes TeNFra-1			
Ref Nom	37	GN	37
Ref Pron	5		
Ref Adv	1		
Total	43		

Por otra parte, en TeNEsp-1, como se muestra en la Tabla 6, se encontraron un total de 452 anáforas, divididas en tres grandes grupos: anáforas nominales, pronominales y

anáforas cero. Si se considera que, según la clasificación propuesta por Cuenca (2010) y esbozada anteriormente, en términos generales las anáforas nominales corresponden a las que la autora denomina como lexicales, y las pronominales a su vez equivaldrían a las anáforas gramaticales, entonces se tiene que, en este texto hay una predominancia de empleo de la anáfora pronominal con 208 elementos, seguida de la anáfora cero con 141; es decir, una predominancia de la anáfora gramatical sobre la lexical, tal como se muestra en la Tabla 7. Tanto la pronominal como la nominal, se subdividieron en esta categorización en tres pequeños subgrupos (véase Marco Metodológico).

Siendo para las unidades nominales el menor grupo el de las anáforas infieles, es decir, aquellas que recurren a sinónimos o reformulación para la retoma de la información, las anáforas fieles y las determinantes guardan una equivalencia cuantitativa. Mientras que, dentro del grupo de las anáforas pronominales sí existe un margen de diferencia cuantitativa más amplio entre los tres subgrupos. El mayor de los cuales, de anáforas totales, es decir, que retoman la totalidad del referente y solo el referente, cuenta con 168 elementos; el segundo subgrupo cuenta con 39 anáforas parciales, que retoman solo una parte del referente y que a su vez ayudan a crear nuevos referentes, tales son los casos de aquellas que retoman atributos de los personajes, por ejemplo; finalmente el subgrupo de anáforas integrantes solo cuenta con una anáfora que retome dos referentes al tiempo.

Para el grupo de las anáforas cero, que es el segundo más relevante cuantitativamente para esta categorización de TeNEsp-1, se encontraron 141 elementos, constatando la afirmación de autores como Muñoz (2017), según la cual, el español puede considerarse una lengua de sujeto tácito, por el frecuente uso que hace del recurso de la elipsis. Dicho grupo,

se ve influido de forma directa por el uso de imperativos en el texto y en los diálogos entre los personajes, propiciados por el narrador, quien les cede la voz.

Tabla 6. *Correferentes anafóricos.*

Anáforas TeNEsp-1			Anáforas TeNFra-1		
Nominal Directa	Pronominal	Cero	Nominal Directa	Pronominal	Cero
103	208		73	192	
Fiel	Total	141	Fiel	Total	0
39	168		13	139	
Determinativa	Parcial		Determinativa	Parcial	
39	39		30	50	
Infiel	Integrante		Infiel	Integrante	
25	1		30	3	
Total			Total		
452			265		

Así, se confirma el planteamiento de Coletta & Karcher (2012), según el cual el tipo de anáfora más recurrente en los textos escritos es la pronominal total; en este caso con un 37,16% del total de los correferentes anafóricos encontrados en TeNEsp-1 y, según la afirmación de Cuenca (2010):

La anáfora gramatical y la léxica, generalmente, alternan en el discurso. Se suele preferir la anáfora gramatical si el antecedente es próximo y claramente identificable y se mantiene el mismo tema o tópico discursivo, mientras que la anáfora o cohesión léxica se suele utilizar con antecedentes menos accesibles, bien porque están más lejanos textualmente y/o porque, hay posibles antecedentes disponibles y se puede crear ambigüedad interpretativa. (Cuenca, 2010, p. 41).

Además, teniendo en cuenta que la autora propone como anáfora gramatical, las acuñadas anáforas cero, sumadas a aquellas que según la tipología seguida de Coletta y Karcher (2012) serían las pronominales. Se podría concluir, por lo tanto, que, si se toman en

cuenta las cifras propuestas en la Tabla 7, que en TeNEsp-1 la mayor parte de referentes son claramente identificables debido a su proximidad con el respectivo correferente.

Tabla 7. Relaciones porcentuales TeNEsp-1 y TeNFra-1.

TeNEsp-1			TeNFra-1		
Nominal	Pronominal	Cero	Nominal	Pronominal	Cero
22,78%	46,02%	31,20%	27,55%	72,45%	0%
Lexical	Gramatical		Lexical	Gramatical	
22,78%	77,22%		27,55%	72,45%	
Total	100,00%		Total	100,00%	

Adicionalmente, en TeNFra-1, como se muestra de igual forma en las Tablas 6 y 7, se categorizaron un total de 265 anáforas, subdivididas en dos grupos principales: el de anáforas nominales y el de las pronominales, ya que en este texto no se encontraron elipsis relativas al sujeto tácito, o anáforas cero. Estos dos grupos principales cuentan con 73 unidades nominales y 192 pronominales, es decir, con un 27,55% de anáforas lexicales y un 72,45% de anáforas gramaticales. De estas últimas, llama la atención que, a pesar de no contar con anáforas cero, siga siendo un porcentaje bastante alto respecto al total de la muestra.

De los dos grupos principales de TeNFra-1 también puede observarse una división en las tres subcategorías de subtipo total, parcial integrante para las pronominales y de subtipos fiel, determinativa e infiel para las nominales. En este texto también se encontró una equivalencia cuantitativa entre las anáforas nominales, pero esta vez entre las determinativas e infieles, categorías con 30 elementos cada una. Al contrario de TeNEsp-1 el menor grupo de las nominales es el de las anáforas fieles, con solo trece elementos. Mientras que, para las pronominales de TeNFra-1 sí se tienen relaciones cuantitativas similares a las de TeNEsp-1, pues de un total de 192, el grupo mucho más pequeño de anáforas integrantes tiene 3 casos,

un segundo grupo de parciales cuenta con 50 elementos y el mayor de los grupos, el de las pronominales totales, posee 139 elementos, es decir, un 72,39% del total de las pronominales (véase la Tabla 8).

Tabla 8. Relaciones porcentuales Nom-Pron.

TeNEsp-1		TeNFra-1	
Nominal Directa	Pronominal	Nominal Directa	Pronominal
Fiel 37,86%	Total 80,76%	Fiel 17,80%	Total 72,39%
Determinativa 37,86%	Parcial 18,75%	Determinativa 41,10%	Parcial 26,05%
Infiel 24,28%	Integrante 0,49%	Infiel 41,10%	Integrante 1,56%
Total			
100,00%	100,00%	100,00%	100,00%

Adicionalmente, a continuación, se enlistan algunos ejemplos de las categorías encontradas en los textos para mejor comprensión de la tipología de anáforas trabajadas.

TeNEsp-1:

- **Nominal directa fiel:** A partir del título, que es el referente que da origen a la cadena correferencial C1, “*Barrabás*”, se producen quince anáforas fieles, es decir, quince más “*Barrabás*” dentro del texto.
- **Nominal directa determinativa:** Llama la atención la anáfora “*las dos*” que, a pesar de que la tipología previamente expuesta de Coletta y Karcher no contemple una categoría de nominales integrantes, es, además de determinativa, de carácter integrante, pues retoma dos referentes diferentes (C8 y C10).
- **Nominal directa infiel:** Del referente C10 (*Jahel*) surgen diferentes anáforas infieles como “*una sombra*” y “*una mujer desgreñada y temblorosa*”.

- **Pronominal total:** Se trata de aquellos casos tales como “*él*”, “*yo*”, “*mí*”, “*me*”, “*lo*”, “*le*”, “*el que dormía*”, “*aquella mujer*”, etc.
- **Pronominal parcial:** Este tipo hace referencia, por lo general, a atributos de un personaje, por ejemplo, como “*su inteligencia*”, “*mi crimen*”, “*sus voces*”. Un fenómeno presente en ambos textos fue la correlación entre personajes mediante enunciados de los que se desprendían dos anáforas: una pronominal parcial, una nominal directa, por ejemplo, “*su hijo*” o “*mi hombre*” y siendo las parciales “*su*” y “*mi*” respectivamente.
- **Pronominal integrante:** El único caso de anáfora integrante encontrada en este texto “*les*”, remite a Jahel y a su hijo (C8 y C10).
- **Anáfora cero:** Verbos cuyas desinencias permiten identificar los sujetos tácitos a que refieren como “*Tenía*”, “*Conocía*”, “*preguntaron*”. De esta categoría llama la atención la posible influencia de la polifonía de voces que introduce el narrador al ceder la voz a los personajes en los diferentes diálogos, posibilitando formas imperativas y preguntas tales como “*¿Sabes?*”, “*¡Oye!*” y “*cállate*”.

TeNFra-1:

- **Nominal directa fiel:** En concordancia con lo comentado anteriormente acerca de la característica de las anáforas pronominales parciales, según la cual participan en la creación de nuevos referentes, se tiene, por ejemplo: “*sa chambre*” como referente posteriormente retomado como “*une chambre*”.
- **Nominal directa determinativa:** En las cadenas correferenciales C43 y C18 se tienen “*une armoire*” y “*quelque voleuse*” como los respectivos referentes, a partir de

los cuales se retoman las siguientes anáforas: “*l’armoire*” y “*la voleuse*”, respectivamente.

- **Nominal directa infiel:** Del referente extraído del título del texto, “(*Récit de l’étudiant*)”, surge la anáfora directa infiel “(*qu*)*un pauvre étudiant*”, y del referente “*une cliente*” surgen “*une locataire*”, “(*sa*) *cliente*” y “*Mme de Prack*” que retoman los referentes no de forma exactamente igual.³
- **Pronominal total:** Pronombres tales como “*que*”, “*ce*”, “*y*”, “*ça*”, “*on*” retomados en contextos que no hubieran dado lugar a ambigüedades respecto a su referente, “*je*”, “*se*”, “*toi*”, etc.
- **Pronominal parcial:** De igual forma que en TeNEsp-1, en francés estos se presentan para referirse a atributos como “*mon attention*”, “*son menton pointu*”, “*ses yeux*” o para establecer relaciones entre diferentes personajes, es decir, diferentes referentes; dándose casos en que de un mismo enunciado parten dos correferentes diferentes, el pronominal parcial y el nominal que refiere al segundo referente, por ejemplo, “*leur amie*” (pronominal que refiere a C12) y la nominal “(*leur*) *amie*” que refiere a C3, o “*ma voisine*” (parcial que refiere a C1) y “(*ma*) *voisine*”, nominal que refiere a C3, etc.⁴
- **Pronominal integrante:** Dos de los tres casos de pronominales integrantes presentados en TeNFra-1 fueron el pronombre “*on*”, el tercer caso fue “*Nous*”, correferente, este último, del que cabe resaltar un carácter catafórico a la vez que anafórico, ya que retoma información previa y, a su vez, anuncia la aparición de nueva

³ Paréntesis propios y sus contenidos no hacen parte de los elementos correferenciales.

⁴ Paréntesis propios y sus contenidos no hacen parte de los elementos correferenciales.

información (en este caso retoma y anuncia la aparición de personajes) en el texto, que posteriormente participa en la creación del referente “*un pensionnaire et moi*”.

Finalmente, cabe aclarar que, si bien en la muestra se presentaron casos cuyas categorizaciones representaron un reto, puesto que cumplían con las condiciones para ser catalogadas bajo diferentes subgrupos ya que, como se puede observar en los ejemplos previamente enunciados, tienen características de diferentes categorías tipológicas, no obstante, las decisiones de estas catalogaciones fueron discutidas y tomadas según la información dada por el contexto del respectivo texto.

Capítulo 6 Conclusiones y perspectivas

Esta investigación se desarrolló siguiendo los parámetros necesarios para dar cumplimiento a los objetivos planteados desde el inicio de la misma. Se considera que dichos objetivos han sido alcanzados y que, sobre todo, el presente trabajo invita a la continuación investigativa en el área de la lingüística textual, y de la correferencia anafórica. También evidencia la importancia de continuar desarrollando herramientas desde la lingüística computacional que faciliten, en la medida de lo posible, el tratamiento de este tipo de datos, para así continuar ahondando el tema desde miradas comparativas entre lenguas.

A partir de la muestra de textos analizada se pudo comprobar, que hay un mayor uso de anáforas de tipo gramatical respecto a las de tipo lexical, por lo que se observa que: la mayoría de referentes registrados en los textos están próximos a sus respectivos correferentes. En este sentido, existe un mayor uso de anáforas de tipo pronominal respecto a las de tipo nominal. De igual forma, se comprueba que, en la totalidad de la muestra, hay un comportamiento o uso similar de las subcategorías rastreadas en la medida que: entre los tres subgrupos de anáforas pronominales se dan usos similares, siendo las anáforas pronominales integrantes las menos empleadas, y las más utilizadas las anáforas pronominales totales; entre los tres subgrupos de las anáforas nominales se dan equivalencias cuantitativas de las determinativas respecto de alguno de los otros. De dicha equivalencia, llama la atención que esta sea diametralmente opuesta en cada texto, sin embargo, se observa que dichas relaciones pueden verse afectadas por los juegos narrativos o las escogencias estilísticas de un determinado autor.

Finalmente, la principal diferencia hallada en el empleo correferencial entre los dos textos seleccionados se da con relación a las anáforas cero, o sea, de aquellas que hacen uso del mecanismo de la elipsis para referirse al sujeto previamente identificado, presentes únicamente en TeNEsp-1, y cuantitativamente relevantes en dicho texto. Por otra parte, llama la atención la aparición de los referentes de tipo tácito, o *solo referidos*, en TeNEsp-1 y se propone continuar investigando en posteriores trabajos la posible relación con la lengua o el género narrativo.

Referencias

- Adam, J.-M. (2010). L'émergence de la linguistique textuelle en France: entre perspective fonctionnelle de la phrase, grammaires et linguistiques du texte et du discours. *Revista Investigações*, 23(2), 11–47.
- Adam, J.-M. (2011). *La linguistique textuelle: introduction à l'analyse textuelle des discours*. Armand Colin: Paris.
- Adam, J.-M. & Lorda, C.-U. (1999). *Linguística de los textos narrativos*. Ariel: Madrid.
- Aduriz, I., Ceberio, K., de Ilarraza, A. D. y García, I. (2008). Análisis de la correferencia para su anotación en un corpus en euskera. *Actas de Congreso: VIII Congreso de Lingüística General*.
- Bertucci, R. A. (2006). Anáforas encapsuladoras: uma análise em textos de opinião. *Revista letras*, 70.
- Bolaños Cuéllar, S. (2015). La lingüística de corpus: perspectivas para la investigación lingüística contemporánea. *Forma y función*, 28(1), 31–54.
- Bruni, A. (2004). Mecanismos de cohesión en un texto de Italo Calvino: Propuesta para un análisis. *Núcleo*, 21, 10–23.
- Colletta, J. y Karcher, G. (2012). *Linguistique textuelle et didactique de l'écrit*. Cned Université Stendhal-Grenoble 3: Grenoble.
- Cuenca, M. J. (2010). *Gramática del texto. Cuadernos de lengua española*. Arco Libros: Madrid.

- Da Silva, I. M. (2006). La utilización de la lingüística del corpus en la enseñanza del español: usos de pero y sino. *Actas del III simposio internacional José Carlos Lisboa de didáctica del Español como lengua extranjera*, pp. 257.
- De la Fuente González, M. Á. (1997). Las funciones de los títulos en la descodificación lectora. *Tabanque: Revista pedagógica*, (12), 185–202.
- Escavy Zamora, R. (1999). La anáfora asociativa como mecanismo de cohesión textual. *Revista de investigación lingüística*, Vol 2, no 2 (1999).
- Leongómez, J. B. (1986). *Antología de lingüística textual*. Vol. 26, Instituto Caro y Cuervo: Bogotá.
- López Morales, H. L. (1994). *Métodos de investigación lingüística*. Ed. Colegio de España: Salamanca.
- Lundquist, L. (1980). *La cohérence textuelle : syntaxe, sémantique, pragmatique*. Nyt Nordisk Forlag: Copenhague.
- Lundquist, L. (2014). La didactisation du français langue étrangère: Discours, genre et type de texte. classification mise a jour. *Genres Et Textes*, Presses universitaires de Lyon, pp. 287–296.
- Martínez-Barco, P. (2001). *Resolución computacional de la anáfora en diálogos: estructura del discurso y conocimiento lingüístico*. PhD thesis. Universitat d'Alacant-Universidad de Alicante: Alicante.

- Martínez, G. P. (2004). La anáfora lexical y el discurso de divulgación científica: aproximación contrastiva. *Le français face aux défis actuels: histoire, langue et culture*, Editorial Universidad de Granada, pp. 783–794.
- Martín Rojo, L. y Meeuwis, M. (1993). Referentes del sujeto pronominales y tácitos en la conversación en español: un enfoque pragmático. *Diálogos hispánicos*, (12), 87–118.
- Molina Mejía, J. M. (2015). *ELiTe-[FLE] 2: Un Environnement d'ALAO fondé sur la Linguistique Textuelle, pour la Formation Linguistique des futurs Enseignants de FLE en Colombie*. PhD thesis. Grenoble Alpes: Grenoble.
- Muñoz Olavarría, B. (2017). *La construcción de la correferencia en la narración de niños con TEL*, ILCL, Pontificia Universidad Católica de Valparaíso: Valparaíso.
- Pardo-Abril, N. G. (1988). Coherencia y cohesión: una aproximación al análisis textual. *Litterae*, (1), 25–57.
- Perdicoyanni-Paléologou, H. (2001). Le concept danaphore, de cataphore et de déixis en linguistique française. *Revue québécoise de linguistique*, 29(2), 55–77.
- Yufei, C. (2014). Clasificación nominal y anáfora: comparación entre español y chino. *Círculo de lingüística aplicada a la comunicación*, (59), 3–15.

Referencias de la muestra analizada

[TeNEsp-1] Úslar Pietri, A. (1928). Barrabás. *Ciudad Seva* [En línea], URL: <https://ciudadseva.com/texto/barrabas/>

[TeNFra-1] Lorrain, J. (1900). Récit de l'étudiant. Bibliothèque Municipale de Lisieux [En línea], URL: <http://www.bmlisieux.com/litterature/lorrain/etudiant.htm>

Anexo A

	Contexto	Refere	Ref tip	Correferent	Clase corre	Subtipo corre	Cadena
2	Barrabás (título)	Ref Nom	NP				C1
3	Su linaje			Anaf Pron	Parcial		C1
4	Tenía			Anaf Cero			C1
5	las barbas negras	Ref Nom	GN				C31
6	los nombres innumerables	Ref Nom	GN				C2
7	el suyo			Anaf Nom	Directa	Determinativa	C2
8	suyo			Anaf Pron	Parcial		C1
9	Barrabás			Anaf Nom	Directa	Fiel	C1
10	Conocía			Anaf Cero			C1
11	era			Anaf Cero			C1
12	guardaba			Anaf Cero			C1
13	sabía			Anaf Cero			C1
14	Jehová	Ref Nom	NP				C3
15	Un viento perezoso	Ref Nom	GN				C4
16	se			Anaf Cero			C4
17	el patio	Ref Nom	GN				C5
18	(d)el calabozo	Ref Nom	GN				C32
19	gentes hacinadas	Ref Nom	GN				C7
20	un soldado con armas	Ref Nom	GN				C6
21	que			Anaf Pron	Total		C6
22	se			Anaf Pron	Total		C6
23	(propusie)se			Anaf Pron	Total	Clítico	C6
24	lo			Anaf Pron	Total		C1
25	enfocó			Anaf Cero			C6
26	las barbas			Anaf Nom	Directa	Infiel	C31
27	le			Anaf Pron	Total		C1
28	¿Yo?			Anaf Pron	Total		C1
29	Barrabás			Anaf Nom	Directa	Fiel	C1
30	¿Barrabás?			Anaf Nom	Directa	Fiel	C1
31	El asesino			Anaf Nom	Directa	Determinativa	C1
32	¿Sabes?			Anaf Cero			C1
33	Te			Anaf Pron	Total		C1
34	van (a matar)			Anaf Cero			C11
35	Te van a matar	Ref Nom	GN				C47
36	lo			Anaf Pron	Total		C47
37	sé			Anaf Cero			C1
38	respondió			Anaf Cero			C1

Figura 2. Extracto del análisis en Excel del texto TeNEsp-1.

Anexo B

	Refere	Ref tí	Correfer	Clase cc	Subtipo corre	Cadena
1 Contexto						
2 (Récit de) l'étudiant (título)	Ref Nom	GN				C1
3 l'hôtel garni	Ref Nom	GN				C2
4 que			Anaf Pron	Total		C2
5 j'(habitais)			Anaf Pron	Total		C1
6 j'(avais)			Anaf Pron	Total		C1
7 une cliente	Ref Nom	GN				C3
8 Je			Anaf Pron	Total		C1
9 (qu')un pauvre étudiant			Anaf Nom	Directa	Infíel	C1
10 cette femme			Anaf Pron	Total		C3
11 mon attention			Anaf Pron	Parcial		C1
12 (qu')elle			Anaf Pron	Total		C3
13 des autres pensionnaires de l'hôtel	Ref Nom	GN				C37
14 (de) l'hôtel			Anaf Nom	Directa	Determinativa	C2
15 C'(était)			Anaf Pron	Total		C3
16 une locataire			Anaf Nom	Directa	Infíel	C3
17 (dirai-)je			Anaf Pron	Total		C1
18 (qu')elle			Anaf Pron	Total		C3
19 sa chambre			Anaf Pron	Parcial		C3
20 sa chambre	Ref Pron					C4
21 elle			Anaf Pron	Total		C3
22 (n')y			Anaf Pron	Total		C4
23 (qu')elle			Anaf Pron	Total		C3
24 s'(y)			Anaf Pron	Total		C3
25 (s')y			Anaf Pron	Total		C4
26 Elle			Anaf Pron	Total		C3
27 un homme	Ref Nom	GN				C10
28 plusieurs femmes	Ref Nom	GN				C5
29 des amies			Anaf Nom	Directa	Infíel	C5
30 on			Anaf Pron	Integrante		C3, C5
31 (l')on			Anaf Pron	Integrante		C3, C5
32 l'hôtel			Anaf Nom	Directa	Determinativa	C2
33 on			Anaf Pron	Total		C37
34 elle			Anaf Pron	Total		C3
35 le tenancier et sa femme	Ref Nom	GN				C7
36 le tenancier	Ref Nom	GN				C8
37 sa femme			Anaf Pron	Parcial		C8

Figura 3. Extracto del análisis en Excel del texto TeNFra-1.