

MÁSTER EN LENGUAS Y TECNOLOGÍA

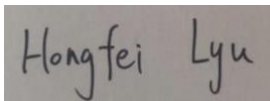
Curso Académico: 2019 / 2020

TÍTULO TRABAJO FIN DE MÁSTER:

Estudio del uso del lenguaje coloquial en la escritura del inglés académico como lengua extranjera

AUTORA: HONGFEI LYU

Declaro que he redactado el Trabajo de Fin de Máster “Estudio del uso del lenguaje coloquial en la escritura del inglés académico como lengua extranjera” para obtener el título de Máster en Lenguas y Tecnología en el curso académico 2019 -2020 de forma autónoma, y con la ayuda de las fuentes consultadas y citadas en la bibliografía (libros, artículos, tesis, etc.). Además, declaro que he indicado claramente la procedencia de todas las partes tomadas de las fuentes mencionadas.



Firmado: Alumna HONGFEI LYU

DIRIGIDO POR: MARÍA LUISA CARRIÓN PASTOR

AGRADECIMIENTOS

En el proceso de mi TFM, he recibido la ayuda, la orientación y el apoyo de muchas personas. Por eso, me gustaría aprovechar esta oportunidad para expresar mi agradecimiento a quienes me han brindado su ayuda y aliento a lo largo de este tiempo.

En primer lugar, quiero agradecerle especialmente a la tutora de mi TFM, María Luisa Carrió Pastor, del Departamento de Lingüística Aplicada de la Universitat Politècnica de València, su ayuda en el proceso de investigación de mi TFM, tanto en su estructura y uso de términos como en la redacción y formación de frases. María me ha dado instrucciones detalladas y correcciones, y me ha ayudado a superar muchas dificultades que me he encontrado en el camino. Por ello, me gustaría expresarle mi agradecimiento.

En segundo lugar, me gustaría agradecerles a todos los profesores del Máster Lenguas y Tecnología su ayuda. Este curso académico ha enriquecido mis conocimientos en lingüística aplicada, traducción y enseñanza de idiomas con asignaturas interesantes. Por eso, quiero darles las gracias por su ayuda desinteresada cuando me encontré con dificultades durante mi estudio.

En tercer lugar, estoy muy agradecida de mis amigos y compañeros de clase. En el proceso de aprendizaje, cada vez que me encontrado con dificultades me han apoyado con sus continuo aliento y apoyo, ayudándome a quitar emociones negativas.

Finalmente, quiero agradecer a mis padres su apoyo y ayuda, con las que he podido superar muchas dificultades y problemas. Desde mi intención de estudiar en España hasta cada pequeña cosa que me he encontrado en la vida, no he podido prescindir de sus ánimos y apoyo en cada etapa de mis estudios. Sé que

siempre que me encuentre en situaciones importantes, ellos me acompañarán y me mostrarán el camino correcto a seguir.

RESUMEN

En los últimos años, a medida que cada vez más personas han ido estudiando de forma académica el inglés, el fenómeno del coloquialismo ha atraído gradualmente la atención de académicos nacionales y extranjeros. La tendencia a un uso del lenguaje coloquial se refiere al uso excesivo de vocabulario y de frases informales. Este fenómeno del aumento en el uso del lenguaje coloquial en la escritura académica afecta seriamente a las intenciones expresivas del autor, debido a una expresión inexacta. En este estudio se analizan las razones de este hecho y se proponen sugerencias a través de los análisis de las características coloquiales de los alumnos ingleses nativos y de los alumnos chinos que utilizan el inglés en la escritura académica.

Esta investigación se ha basado en una investigación existente sobre el coloquialismo en la escritura académica, utilizando una combinación de métodos de investigación cuantitativos y cualitativos, tomando como objetos de investigación a alumnos chinos y a alumnos nativos, y analizando comparativamente dos corpus (BASE y BAWE) que se incluyen en el *Sketch Engine*. Los corpus LM y TFM, comparan y analizan los cuatro corpus con un total de más de 9,27 millones palabras, con el objetivo de analizar los caracteres coloquiales del escritor, su nivel léxico y su nivel sintáctico.

Los principales hallazgos de esta investigación han sido: 1). En términos de visibilidad del escritor, tanto los alumnos nativos como los alumnos chinos muestran un uso excesivo de los pronombres personales de primera y segunda persona. 2). En términos de nivel léxico, los alumnos chinos muestran una mayor tendencia coloquial y utilizan mucho vocabulario coloquial. 3). En términos de nivel sintáctico, los alumnos nativos tienden a usar frases más largas y complejas, mientras que los alumnos chinos hacen lo contrario. El resultado final muestra

que los alumnos nativos y los alumnos chinos tienen diferentes grados de uso de la lengua coloquial en su escritura académica, pero con la mejoría del dominio del inglés, esta tendencia coloquial tiene una tendencia significativamente a la baja. Por ejemplo, en comparación con los alumnos universitarios, los alumnos de inglés son más profesionales, ya que utilizan menos lenguaje coloquial en sus escritos académicos.

Finalmente, a través de los resultados de la presente investigación, se han presentado sugerencias con tal de ayudar a los alumnos nativos y chinos con su escritura académica en inglés.

Palabras clave: Escritura académica, tendencia coloquial, alumnos chinos

RESUM

En els últims anys, a mesura que cada vegada més persones han anat estudiant de forma acadèmica l'anglès, el fenomen de l'col·loquialisme ha atret gradualment l'atenció d'acadèmics nacionals i estrangers. La tendència a un ús de l' llenguatge col·loquial es refereix a l'ús excessiu de vocabulari i de frases informals. Aquest fenomen de l'augment en l'ús de l' llenguatge col·loquial en l'escriptura acadèmica afecta seriosament a les intencions expressives de l'autor, a causa d'una expressió inexacta. En aquest estudi s'analitzen les raons d'aquest fet i es proposen suggeriments a través de les anàlisis de les característiques col·loquials dels alumnes anglesos nadius i dels alumnes xinesos que utilitzen l'anglès en l'escriptura acadèmica.

Aquesta investigació s'ha basat en una investigació existent sobre el col·loquialisme en l'escriptura acadèmica, utilitzant una combinació de mètodes d'investigació quantitativus i qualitativus, prenent com a objectes d'investigació a alumnes xinesos i alumnes nadius, i analitzant comparativament 2 corpus (BASE i BAWE) que s'inclouen en el Sketch cerca. Els corpus LM i TFM, comparen i analitzen els quatre corpus amb un total de més de 9.270.000 paraules, amb l'objectiu d'analitzar els caràcters col·loquials de l'escriptor, el seu nivell lèxic i el seu nivell sintàctic.

Els principals troballes d'aquesta investigació han estat: 1). En termes de visibilitat de l'escriptor, tant els alumnes nadius com els alumnes xinesos mostren un ús excessiu dels pronoms personals de primera i segona persona. 2). En termes de nivell lèxic, els alumnes xinesos mostren una major tendència col·loquial i utilitzen molt vocabulari col·loquial. 3). En termes de nivell sintàctic, els alumnes nadius tendeixen a fer servir frases més llargues i complexes, mentre que els alumnes xinesos fan el contrari. El resultat final mostra que els alumnes nadius i els alumnes xinesos tenen diferents graus d'ús de la llengua

col·loquial en la seva escriptura acadèmica, però amb la millora de l' domini de l'anglès, aquesta tendència col·loquial té una tendència significativament a la baixa. Per exemple, en comparació amb els alumnes universitaris, els alumnes d'anglès són més professionals, ja que utilitzen menys llenguatge col·loquial en els seus escrits acadèmics.

Finalment, a través dels resultats de la present investigació, s'han presentat suggeriments per tal d'ajudar els alumnes nadius i xinesos amb la seva escriptura acadèmica en anglès.

Paraules clau: Escriitura acadèmica, tendència col·loquial, alumnes xinesos

ABSTRACT

In recent years, as more and more people have studied English academically, the phenomenon of colloquialism has gradually attracted the attention of national and foreign scholars. The tendency to use colloquial language refers to the excessive use of vocabulary and informal phrases. This phenomenon of the increase in the use of colloquial language in academic writing seriously affects the expressive intentions of the author, due to an inaccurate expression. In this study, the reasons for this fact are analyzed and suggestions are made through the analysis of the colloquial characteristics of native English students and Chinese students who use English in academic writing.

This research has been based on existing research on colloquialism in academic writing, using a combination of quantitative and qualitative research methods, taking Chinese students and native students as research objects, and comparatively analyzing two corpus (BASE and BAWE) that are included in the Sketch Engine. The LM and TFM corpus compare and analyze the four corpus with a total of more than 9.27 million words, in order to analyze the writer's colloquial characters, their lexical level and their syntactic level.

The main findings of this research have been: 1). In terms of writer visibility, both native and Chinese students show excessive use of first and second person personal pronouns. 2). In terms of lexical level, Chinese students show a greater tendency to colloquial and use a lot of colloquial vocabulary. 3). In terms of syntactic level, native speakers tend to use longer and more complex sentences, while Chinese students do the opposite. The final result shows that native-born students and Chinese students have different degrees of use of the colloquial language in their academic writing, but with the improvement of English proficiency, this colloquial trend has a significantly downward trend. For example, compared to university students, English students are more professional as they

use less colloquial language in their academic writing.

Finally, through the results of the present investigation, suggestions have been presented in order to help native and Chinese students with their academic writing in English.

Keywords: Academic writing, colloquial tendency, Chinese students

AGRADECIMIENTOS	2
RESUMEN	4
RESUM	6
ABSTRACT	8
1. INTRODUCCIÓN.....	17
1.1. Antecedentes	17
1.2. Objetivos	19
1.3. Preguntas de investigación	19
1.4. Estructura del trabajo.....	20
2. FUNDAMENTACIÓN TEÓRICA.....	23
2.1. La escritura académica.....	23
2.1.1. Definición de la escritura académica	23
2.1.2. El artículo académico	25
2.2. El inglés como lengua extranjera	27
2.2.1. El inglés como lengua franca.....	27
2.2.2. El inglés en China	29
2.3. La escritura académica en una lengua extranjera.....	32
2.3.1. Discusión teórica sobre la escritura académica en las lenguas extranjeras	32
2.3.2. La escritura académica del inglés	34
2.4. El lenguaje coloquial.....	35
2.4.1. Definición del lenguaje coloquial	36
2.4.2. Las características del lenguaje coloquial.....	36
2.5. La escritura académica del inglés en China	39
2.5.1. La situación actual de la escritura académica del inglés en China	39
2.5.2. La tendencia coloquial de la escritura académica de los alumnos chinos.....	40

3. METODOLOGÍA	42
3.1. Métodos.....	42
3.1.1. El método basado en el análisis de corpus.....	42
3.1.2. El método cuantitativo y cualitativo.....	43
3.2. Corpus.....	44
3.2.1. <i>British Academic Spoken English Corpus</i> (BASE).....	44
3.2.2. <i>British Academic Written English Corpus</i> (BAWE).....	45
3.2.3. Corpus de pregrado en la lengua materna (LM)	46
3.2.4. Corpus de trabajos fin de Máster (TFM).....	47
3.3. Diseño experimental	48
3.3.1. Parámetros utilizados en el estudio.....	48
3.3.1.1. La visibilidad del escritor	48
3.3.1.2. El nivel léxico	49
3.3.1.2.1. Diversidad léxica (TTR).....	49
3.3.1.2.2. Densidad léxica	50
3.3.1.2.3. Vocabulario coloquial	51
3.3.1.2.4. Abreviaturas y contracciones	52
3.3.1.3. El nivel sintáctico	52
3.3.1.3.1. Longitud promedio de la frase	53
3.3.1.3.2. Estructura " <i>There be</i> "	53
3.3.1.3.3. Verbos perifrásticos.....	54
3.3.2. Herramientas utilizadas en el estudio.....	54
3.3.2.1. <i>Sketch Engine</i>	54
3.3.2.2. <i>AntConc</i>	56
3.3.2.3. <i>WordSmith 4.0</i>	56
4. RESULTADOS Y DISCUSIÓN	58
4.1. Resultados de los coloquialismos de visibilidad del escritor en el inglés académico	58
4.1.1. Análisis de BASE.....	60
4.1.2. Análisis de BAWE.....	62

4.1.3. Análisis de LM	63
4.1.4. Análisis de TFM	64
4.2. Resultados de los coloquialismos léxicos en el inglés académico	68
4.2.1. Diversidad léxica (TTR).....	68
4.2.2. Densidad léxica	70
4.2.3. Frecuencia de vocabulario	71
4.2.3.1. Frecuencia de vocabulario coloquial.....	72
4.2.3.2. Frecuencia de abreviaturas y contracciones	78
4.3. Resultados de los coloquialismos sintácticos en el inglés académico	81
4.3.1. Longitud promedio de la frase	82
4.3.2. Estructura de la frase	84
4.3.2.1. Estructura " <i>There be</i> "	84
4.3.2.2. Verbos perifrásticos	86
4.4. Posibles coloquialismos en la escritura académica de los alumnos chinos	90
4.4.1. Transferencia negativa de la lengua materna	90
4.4.2. Métodos tradicionales de enseñanza del inglés en China.....	94
5. CONCLUSIONES	96
5.1. Hallazgos principales.....	96
5.2. Limitaciones e investigaciones futuras.....	99
BIBLIOGRAFÍA.....	101

ÍNDICE DE TABLAS

<u>TÍTULO</u>	<u>PÁGINAS</u>
Tabla 1: Conocimiento del inglés en la población	27
Tabla 2: Número de hablantes nativos del inglés y proporción de usuarios en los principales países del mundo	28
Tabla 3: Información básica del corpus BASE	45
Tabla 4: Información básica del corpus BAWE.....	46
Tabla 5: Información básica del corpus LM	47
Tabla 6: Información básica del corpus TFM.....	48
Tabla 7: Visibilidad del escritor de los 4 corpus.....	59
Tabla 8: Resultados comparativos de la visibilidad del escritor de los 4 corpus	66
Tabla 9: Comparación de TTR de los 4 corpus	69
Tabla 10: Resultados comparativos de la diversidad léxica de los 4 corpus	69
Tabla 11: Comparación de la densidad léxica de los 4 corpus	70
Tabla 12: Resultados comparativos de la densidad léxica de los 4 corpus	71
Tabla 13: Comparación de la frecuencia del vocabulario coloquial de los 4 corpus	73
Tabla 14: Resultados comparativos de la frecuencia del vocabulario coloquial de los 4 corpus.....	77
Tabla 15: Comparación de la frecuencia de abreviatura y contracción de los 4 corpus.....	78
Tabla 16: Resultados comparativos de frecuencia de abreviaturas y contracción de los 4 corpus.....	81
Tabla 17: Comparación de la longitud promedio de la frase de los 3 corpus....	82
Tabla 18: Resultados comparativos de la longitud promedio de los 3 corpus ..	83
Tabla 19: Comparación del uso de la estructura "There be" en los 3 corpus....	84
Tabla 20: Resultados comparativos de la frecuencia de la estructura "There be" en los 3 corpus.....	84

Tabla 21: Waibel's 72 High-frequency Phrasal-Verb List en LOCNESS	86
Tabla 22: Comparación del uso de los 10 verbos perifrásticos en los 4 corpus	87
Tabla 23: Características coloquiales de las escrituras académicas de los alumnos chinos	91

ÍNDICE DE FIGURAS

TÍTULO	PÁGINAS
Figura 1: Teoría de la clasificación de los tipos de texto de Rice y Chippindale	26
Figura 2: El modelo de escritura académica de Paltridge et al	33
Figura 3: Modelo estructural de la escritura académica de Murray y Moore	34
Figura 4: Los cuatro corpus involucrados en esta investigación	44
Figura 5: Elementos del vocabulario más utilizados en la categoría coloquial de cada función retórica de Paquot.....	52
Figura 6: Características de la visibilidad del escritor de Han	58
Figura 7: Comparación de la visibilidad del escritor en los 4 corpus (por millón)	67

LISTA DE ABREVIATURAS

ILE: Inglés como lengua extranjera

BASE: British Academic Spoken English Corpus

BAWE: British Academic Written English Corpus

ESRC: Economic and Social Research Council

LM: Corpus de pregrado en lengua materna

LOCNESS: Louvain Corpus of Native English Essays

TFM: Corpus de trabajos fin de Máster

CNKI: China National Knowledge Infrastructure

L1: Lengua materna

L2: Lengua extranjera

AWL: Academic Writing List

1. INTRODUCCIÓN

En la presente investigación se explora el uso de lenguaje coloquial en la escritura académica del inglés mediante la comparación de la escritura académica de alumnos nativos ingleses y de alumnos chinos. En este primer apartado se presentan los antecedentes de la investigación, los objetivos y el significado del estudio, las preguntas de investigación y la estructura del trabajo.

1.1. Antecedentes

Con el continuo desarrollo de la globalización, se han generado enormes oportunidades para el desarrollo económico y para los intercambios culturales. Como idioma común internacional, el inglés es una herramienta importante para la globalización en muchos campos (generalmente en los campos académicos) y es cada vez más valorado por países de todo el mundo. Los expertos, científicos e incluso alumnos de diferentes campos académicos tienen más oportunidades de poder publicar sus investigaciones científicas y sus logros en las plataformas internacionales a través de artículos académicos en inglés.

Al mismo tiempo, con la importancia de la educación en inglés y con la mayor frecuencia de los intercambios internacionales, con el objetivo de facilitar la comunicación y la comprensión, el inglés coloquial se ha convertido en el centro de los alumnos de inglés, mientras que el inglés escrito no es tan considerado. Se está enfatizando cada vez más con el cultivo de las habilidades de comunicación coloquial de los alumnos y se están descuidando las expresiones escritas, lo que hace que los alumnos no presten la debida atención a la estandarización del inglés escrito. Por lo tanto, están apareciendo muchas características coloquiales en el inglés escrito de los alumnos de este idioma. Incluso en los alumnos avanzados que no presentan errores importantes, todavía mantienen el uso expresiones incorrectas y de coloquialismo en su

escritura académica.

En China, el inglés se ha convertido desde hace mucho tiempo en el segundo idioma más importante, ya que es una habilidad que los alumnos chinos deben dominar desde la infancia. Wang y Jenkins (2016:41) explican que "Desde principios del siglo XXI, el inglés se ha convertido en un idioma ampliamente utilizado por muchos usuarios chinos". Sin embargo, para los alumnos chinos, su conciencia para diferenciar entre el lenguaje coloquial y escrito es vaga, y no solo en los alumnos de inglés de bajo nivel, sino también en los alumnos avanzados de especializaciones en inglés, ya que en sus escritos aún se puede encontrar una tendencia coloquial, lo que afecta seriamente a la profesionalidad de la escritura académica.

Gracias a las referencias de estudios previos, me he encontrado con que este fenómeno no solo existe en la escritura académica de los alumnos extranjeros de inglés, sino que también en la de los alumnos nativos, los cuales tampoco pueden ignorar el coloquialismo de la escritura académica. No obstante, entre los alumnos nativos y los alumnos de inglés, el grado de coloquialismo en la escritura académica es diferente, así como sus características coloquiales. Pero para mí, el foco de la investigación es el hecho que estos alumnos usen inadvertidamente las características relevantes del coloquialismo en la escritura académica.

A nivel general, la situación de la escritura en inglés en China es preocupante. Ya sean alumnos nativos o alumnos chinos, la mayoría de las personas no están prestando suficiente atención a este tema. No obstante, esta situación no es inevitable, ya que encontrando la causa y tomando las sugerencias correspondientes puede mejorar esta situación. En este estudio, se combinará la investigación y los estudios previos y se compararán las características coloquiales de los alumnos nativos y de los alumnos chinos en la escritura

académica, estudiando el uso del coloquialismo en la escritura académica, analizando las razones de este fenómeno y proponiendo sugerencias.

1.2. Objetivos

A medida que el inglés se ha ido usando cada vez con más frecuencia en la vida diaria, gradualmente se han formado algunos hábitos coloquiales en los alumnos nativos y en los alumnos de inglés. No obstante, la mayoría de los autores no prestan demasiada atención a este tema. Con el paso del tiempo, estos hábitos coloquiales han ido difuminando gradualmente los límites entre el lenguaje escrito y el coloquial, haciendo que la escritura académica sea cada vez más coloquial. Por lo tanto, se debe poner énfasis en el coloquialismo de la escritura académica.

El objetivo general de este artículo es, en primer lugar, analizar el coloquialismo en la escritura académica por parte de los alumnos nativos y los alumnos chinos. Los objetivos específicos son explorar algunas de las características coloquiales que a menudo se pasan por alto en la escritura académica, analizar sus razones y dar sugerencias relevantes, con el fin de mejorar la escritura académica de los alumnos chinos.

A partir de estos objetivos, espero mejorar la capacidad de escritura académica en inglés de los alumnos chinos y proporcionar ideas para los métodos de enseñanza del inglés.

1.3. Preguntas de investigación

Este trabajo pretende responder a las siguientes preguntas de investigación:

- Pregunta 1: ¿Cuáles son las características coloquiales de la escritura

académica en inglés? La pregunta se analizará desde los siguientes 3 aspectos:

- a. Visibilidad del escritor
- b. Nivel léxico
- c. Nivel sintáctico

- Pregunta 2: ¿Cuál es el motivo de esta tendencia al coloquialismo?
¿Cambia con la mejora del nivel de inglés?

- Pregunta 3: En comparación con los alumnos nativos, ¿existe una tendencia coloquial en la escritura académica de los alumnos chinos con un alto nivel de inglés (es decir, la Maestría en inglés)? ¿Existen diferencias significativas en comparación con los alumnos nativos?

1.4. Estructura del trabajo

Desde la perspectiva teórica, se ha analizado el corpus académico relacionado con la escritura académica, basándonos en las investigaciones previas y analizando los tres aspectos de visibilidad del escritor, el nivel léxico y el nivel sintáctico, con el fin de reflejar las características distintivas del inglés coloquial en la escritura de los alumnos nativos y chinos. Por otro lado, esperamos utilizar esta investigación para proporcionar datos para el análisis de una tendencia coloquial en la escritura académica en inglés.

Desde una perspectiva práctica, esta investigación no solo se centra en los coloquialismos de la escritura académica, sino que también analiza sus principales razones y hace sugerencias al respecto. Al mismo tiempo, esta investigación adopta un método de análisis comparativo basado en el análisis de corpus, recopilando artículos académicos de alumnos chinos y alumnos nativos.

A través del análisis comparativo, se demostrará si existen rasgos coloquiales en sus escritos académicos, los cuales podrían reflejar el coloquialismo de la escritura académica en la actualidad.

Esta investigación consta de cinco capítulos.

En el capítulo 1 se presenta brevemente la situación general de la presente investigación, incluyendo los antecedentes de la investigación, los objetivos y el significado de la investigación, las preguntas de investigación y la estructura de la investigación.

El capítulo 2 es una revisión de la literatura, la cual proporcionará una base teórica para el trabajo. Este capítulo divide en: la escritura académica, el inglés como lengua extranjera, la escritura académica en una lengua extranjera, el lenguaje coloquial y, finalmente, explica la escritura académica del inglés en China.

El capítulo 3 es el método de investigación, en el que se presentan los métodos de recopilación y de análisis de datos. También se presentan los parámetros, las herramientas y los corpus utilizados en esta investigación. El corpus incluye los corpus autoconstruidos LM y TFM, y los corpus de referencia BASE y BAWE de *Sketch Engine*. Las herramientas de análisis incluyen *Sketch Engine*, *AntConc* y *WordSmith 4.0*.

El capítulo 4 incluye los resultados de la investigación y de la discusión. En este capítulo se discute y se analizan las características coloquiales de los cuatro corpus académicos desde tres aspectos: la visibilidad del escritor, el nivel léxico y el nivel sintáctico. Entre ellos, el nivel léxico se divide en diversidad léxica, densidad léxica, vocabulario coloquial, abreviaturas y contracciones. Por otro lado, el nivel sintáctico se divide en la longitud promedio de la frase, la estructura

"*There be*" y verbos perifrásticos. Finalmente, se analiza la posibilidad coloquial de la escritura académica de los alumnos chinos a partir de datos relevantes.

El capítulo 5 es la conclusión de toda la investigación. En este capítulo se presentan los principales hallazgos, como también las limitaciones e investigaciones futuras de esta investigación.

2. FUNDAMENTACIÓN TEÓRICA

En este capítulo, se exponen los conceptos relacionados con la escritura académica del inglés. Se ha dividido en las siguientes 5 partes: la escritura académica, el inglés como lengua extranjera, la escritura académica en una lengua extranjera, el lenguaje coloquial y la escritura académica de inglés en China. Las teorías incluidas se han desarrollado en torno a la escritura académica.

2.1. La escritura académica

Bailey (2003:8) explica que “la escritura académica está diseñada para cualquiera que esté estudiando en universidades y colegios universitarios de nivel medio en inglés y tenga que escribir ensayos y otras asignaciones para exámenes o cursos.”

2.1.1. Definición de la escritura académica

Copland y Creese (2015:161) argumentan que “la escritura académica sirve como una herramienta de comunicación para transmitir los conocimientos adquiridos en un campo de estudio específico”. Copland y Creese (2015) consideran que la escritura académica presenta un tono serio y, a menudo, demuestra teorías y hechos particulares relacionados con un argumento dado.

Según la investigación de Richards (2010:51), “escribir es una habilidad que se requiere en muchos contextos a lo largo de la vida, pero la escritura académica hace muchas de las cosas que la escritura personal no hace y tiene su propio conjunto de reglas y prácticas.”

Hinkel (2002:392) considera que la guía de escritura y sus estándares de

medición dependen de las normas y de los estándares de fonética y del texto, por lo que la escritura académica es el estándar más alto del inglés escrito. Sin embargo, la escritura académica tiene las características típicas del inglés escrito.

La escritura académica es la que tienes que hacer para tus cursos universitarios. Tus profesores pueden tener diferentes nombres para las asignaciones de una redacción académica (ensayo, artículo, artículo de investigación, trabajo final, etc), pero todas ellas tienen el mismo objetivo y principios. (Whitaker, 2009:2)

Irvin (2010:6) defiende que "la escritura académica es siempre una forma de evaluación que requiere demostrar el conocimiento y la habilidad para pensar, explicar y expresar ciertas habilidades de la materia". También cree que el éxito de la escritura académica depende del nivel de comprensión de la escritura y en la forma en la que se completan las tareas de escritura.

Biber (1991:299) considera que los artículos académicos son el estilo de escritura típico. También Šimėkaitė (2012) señala que la escritura académica es una habilidad más difícil para dominar, y que, en comparación con el lenguaje cotidiano, la estructura de la escritura académica es compleja y su estilo es más formal y objetivo.

En la investigación de Liao (2013), se encontró que, en la escritura académica, el autor intenta construir una identidad de la comunidad académica y cree que puede ser reconocida por los lectores a través de las estrategias discursivas en el vocabulario, la sintaxis y la estructura del lenguaje utilizado. Liao (2013:137) explica que "no importa si este esfuerzo tiene éxito o no, ya que el texto muestra ciertas características de la identidad del autor."

Según la investigación de Yu, Xu y Bai (2011), en la escritura académica, el autor

intenta construir una identidad dentro de la comunidad académica. Al mismo tiempo, a través de estrategias de discurso en el vocabulario, la sintaxis y la estructura del lenguaje utilizado, el texto mostrará ciertas características de la identidad del autor.

2.1.2. El artículo académico

Según la investigación de Yu, Xu y Bai (2011:42), “los artículos académicos son un estilo de escritura académica para el análisis teórico, el resumen y la presentación de informes de resultados de investigación académica innovadores.”

Reiss (1989) divide los textos en tres tipos:

- (1). Informativo: se enfoca en hechos puros, proporcionando información, conocimiento, opiniones, etc. El lenguaje es lógico y el enfoque de la comunicación es el contenido o el tema.
- (2). Expresivos: el autor utiliza un lenguaje estético. El estado del autor o el remitente original del mensaje y la forma del texto son el foco de dichos textos.
- (3). Vocativos u operativos: este tipo de texto está destinado a persuadir al lector o al destinatario del mensaje. Usa un lenguaje conversacional, ya que el objetivo de este tipo de texto es llegar a los lectores.

Según los diferentes tipos de texto, Rice y Chippindale (2001:685) también propuso los métodos de traducción correspondientes (véase figura 1).

Text types	Informative	Expressive	Operative
language function	present things and facts	express the attitude of the sender	infect the receiver of the text
language features	logical	aesthetic	dialogic
text focus	focus on content	focus on form	focus on infectious function
language expression	in Plain language, according to the requirements of Concise	adopt the method of imitation and be faithful to the original author	it is equivalent to the original text by means of compiling

Figura 1: Teoría de la clasificación de los tipos de texto de Rice y Chippindale

Fuente: Análisis de la escritura académica del inglés en China, Zhang R.F, (2020: 21)

Yu, Xu y Bai (2011) consideran que los artículos académicos son textos informativos y que se centran en expresar el contenido. Por lo tanto, en la redacción de artículos académicos, el contenido del artículo debe expresarse en un lenguaje plano, directo, conciso y claro de acuerdo con los hábitos de expresión.

Halliday (1994:55) explica que "los artículos académicos son los actos del habla social en los que el autor está muy involucrado". Al mismo tiempo, Ivanič (1998: 329) cree que "el artículo académico es en el que los autores utilizan diversas estrategias de discurso para negociar la identidad propia de la comunidad académica". Específicamente, Swales, (1990:84) indica que:

Los artículos académicos de las asignaturas diferentes se tienen características lingüísticas similares. Por ejemplo, la longitud promedio de una frase es cercana a las 25 palabras; la frecuencia de las frases atributivas es menor que las frases nominales, la cláusula adverbial de razón y tiempo; el tema de la cláusula principal se incluye muchos sustantivos abstractos y formas sustantivas.

Bhatia (1999) cree que las estructuras de nominalización complejas, la

nominalización y las frases en gerundio son características típicas de los trabajos académicos.

2.2. El inglés como lengua extranjera

2.2.1. El inglés como lengua franca

Hoy en día, se utilizan más de 6.000 idiomas diferentes en todo el mundo, pero el inglés se trata del idioma franco del mundo. Además, es considerado como idioma oficial en muchos países. Zhou (1994:28) explica que “el inglés es un puente de comunicación entre personas que hablan diferentes idiomas.” Se trata de uno de los idiomas oficiales de las Naciones Unidas, de la Unión Europea y de muchas otras organizaciones internacionales regionales y mundiales. Según algunas estadísticas, el inglés es el primer idioma de más de 375 millones de personas. Al mismo tiempo, el inglés es el idioma más estudiado como lengua extranjera, ya que tiene más de mil millones de personas que lo están aprendiendo actualmente (véase tabla 1).

Clasificación por población					
Lengua	Como hablante nativo	Rango	Como usuario de un segundo idioma	Rango	Total
Inglés	367 millones	3	898,4 millones	1	1.268 mil millones

Tabla 1: Conocimiento del inglés en la población (producción y traducción propias)

Fuente: Los datos provienen de <https://baike.baidu.com/item/英语>

Con la informatización de la vida social y la globalización económica, la importancia del inglés se ha vuelto cada vez más prominente. Zhou (1994) cree que, como uno de los portadores de información más importantes, el inglés se ha convertido en el idioma más utilizado en todas las áreas de la vida cotidiana y que el dominio del inglés se ha convertido en una habilidad esencial. En concreto explica:

Para finales de este siglo, la ortografía y la pronunciación del inglés se unificarán gradualmente; los cambios irregulares en el vocabulario y en la gramática se analizarán por analogía de reconstrucción; se usarán más siglas, omisiones y frases cortas; la conversión de voz será extremadamente flexible, con expresiones más concisas; la tasa de repetición de palabras comunes será bastante alta; los límites entre los sustantivos contables e incontables, los verbos transitivos y los verbos intransitivos, casi habrán desaparecido. Todo esto facilitará enormemente que las personas aprendan y utilicen el inglés como herramienta comunicativa (Zhou, 1994:28).

En países como China, el inglés ya es un curso de lengua extranjera obligatorio en las escuelas. La tabla 2 nos muestra los hablantes nativos del inglés:

Clasificación por hablantes nativos del inglés			
Rango	País	Hablantes nativos	Porcentaje
1	Estados Unidos	300.809.000	64,3%
2	Reino Unido	60.200.000	16,7%
3	Canadá	25.232.195	5,3%
4	Australia	20.581.334	4,7%
5	Irlanda	> 4.200.000	1,1%
6	Sudáfrica	3.673.203	1,3%
7	Nueva Zelanda	> 3.500.000	1,1%

Tabla 2: Número de hablantes nativos del inglés y proporción de usuarios en los principales países del mundo (producción y traducción propias)

Fuente: Los datos provienen de <https://baike.baidu.com/item/英语>

Según la definición de "Lengua Internacional" de Smith (1976), el idioma que se utiliza por personas de diferentes países para comunicarse se trata de una lengua internacional. Esta definición es obviamente muy vaga y no es suficiente para distinguir a una la lengua internacional de una lengua franca.

Por lo tanto, Brutt-Griffler (1998:381) hizo una definición más específica y clara de una lengua franca internacional, ya que creyó que una lengua cosmopolita e internacional debería tener estas cuatro características:

- (1). Es un producto de la integración de la economía mundial, la tecnología y el desarrollo cultural.
- (2). Se suele convivir con los idiomas locales y es el idioma principal de grupos multilingües o bilingües.
- (3). A diferencia de una interlengua general, su uso no se limita a un pequeño número de élites involucradas en actividades sociales y económicas. Es una lengua objeto de aprendizaje y una herramienta utilizada por todas las clases de la sociedad.
- (4). Una lengua franca internacional no se transmite por personas que usan el idioma como lengua materna y que emigran a otras regiones, sino por muchas personas que aprenden el idioma en todo el mundo.

Se puede decir que el desarrollo del inglés hasta el día de hoy ha alcanzado plenamente los cuatro estándares de la lengua franca internacional. Hoy en día, el estatus internacional del inglés como lengua franca sigue siendo estable. Gracias al aprendizaje del inglés uno puede comunicarse fácilmente con personas de muchos países del mundo, ya que todos los países se esfuerzan por mantener el estatus del inglés. Zhou (1994:28) explica al respecto:

Debido a la creciente popularidad e importancia del aprendizaje del inglés en todo el mundo en los últimos años, el inglés seguirá siendo la lengua franca internacional y el "puente" para comunicarse con el mundo durante mucho tiempo.

2.2.2. El inglés en China

El entusiasmo de los chinos por aprender idiomas extranjeros siempre ha sido alto. Sun (1989:19) cree que hay una fuerte motivación detrás de este

entusiasmo, que son los llamados “ojos abiertos para ver el mundo”. Zhou (1994:28) considera que “como una lengua extranjera es una herramienta, una forma y una manera de entender el mundo, también es un puente hacia el mundo”. Wu (2014:26) explica que “es un hecho indiscutible que el inglés es una lengua internacional.” Sun (1989:19) añade a esta idea:

Con la tendencia de la globalización, China está cada vez más estrechamente conectada con el mundo. En China la mayoría de las personas utilizan el inglés de forma extensiva en la comunicación internacional, la educación, la cultura y el comercio internacional. Se ha convertido en una herramienta indispensable en el proceso de desarrollo y modernización de China.

Wu (2014:26) encontró que desde que el gobierno chino fundó el "Jing Shi Tong Wen Guan" (京师同文馆) para enseñar inglés en 1862, la historia de la educación en inglés ya tiene más de 150 años en China. Wu (2014:26) cree que el desarrollo del inglés en China ha pasado por cinco etapas, que son las siguientes:

- (1). En 1862, el inglés se entró oficialmente en China.
- (2). En 1903, el inglés ingresó en la educación universitaria, se convirtió oficialmente en una asignatura y tuvo alumnos formalmente.
- (3). En 1962, el inglés ingresó como asignatura del examen de ingreso a la universidad china.
- (4). En 1964, el inglés se fue catalogado oficialmente como primer idioma extranjero por el Ministerio de Educación de China, el cual pidió que aumentara el número de alumnos de inglés y que se ajustara la proporción de clases de idiomas extranjeros en las escuelas secundarias.

(5). En enero de 2001, el Ministerio de Educación de China decidió establecer cursos de inglés en las escuelas primarias como parte importante de la reforma de educación básica de principios del siglo XXI. Los cursos de inglés se ofrecen gradualmente en las escuelas primarias de varias ciudades y se estipula que el inglés debe aprenderse desde el tercer grado de la escuela primaria.

Zhou (1994) predijo que la '*English Hot*' se convertiría en una tendencia en China y, basándose en esta tendencia, el inglés se desarrollaría rápidamente en este país. Por otro lado, Wu (2014) ha concluido que China se ha convertido en el país con el mayor número de alumnos de inglés, ya que el número a menudo se ha mantenido por encima de los 6 millones en los últimos años, superando así a los Estados Unidos. Chen (2009:165) explica que:

No se puede comparar el chino con el inglés o con cualquier otro idioma extranjero. La capacidad del lenguaje y la sensibilidad son las mismas: la memoria, la capacidad de contraste, la capacidad de metáfora, la capacidad de transformación y la capacidad de escape están todas en la categoría del aprendizaje de un idioma a. Aprende un idioma extranjero es aprender otro tipo de pensamiento, ya que la esencia del lenguaje es la metáfora.

Con el impacto de la adhesión de China en la OMC (Organización Mundial del Comercio), el número de personas que aprenden inglés ha alcanzado un alto nivel. Chen (2009:168) considera que "la difusión rápida y generalizada del inglés es inseparable del desarrollo de la economía, la tecnología moderna y el desarrollo de la sociedad china." Sun (1989:23) cree que "el inglés puede ayudar a los chinos a comprender el mundo y a conectarse con los estándares internacionales."

En la investigación de Chen (2009:168), el autor se encontró que, con el

advenimiento de la era de la información, el inglés se ha desarrollado en muchos campos de la cultura china: Internet, el automóvil, la comida rápida y la telefonía móvil, los cuales se han ido aceptado por más y más personas y los cuales han jugado un papel en el desarrollo de la economía de China.

2.3. La escritura académica en una lengua extranjera

Hyland (2013:56) explica que "con la internacionalización de los intercambios académicos, el inglés se ha convertido en el idioma universal de comunicación en la academia". Liao (2013) cree que la capacidad de escritura académica en inglés es un reflejo directo de la calidad de la educación universitaria, la capacidad de investigación académica, la calidad profesional y el nivel de escritura académica en inglés de los alumnos, lo cual representa la manifestación de una calidad integral.

2.3.1. Discusión teórica sobre la escritura académica en las lenguas extranjeras

Las investigaciones tempranas sobre la escritura académica en una lengua extranjera se basan principalmente en la teoría de la escritura general de un idioma y de un segundo idioma. No fue hasta mediados y finales del siglo XX que se comenzaron a proponer ideas teóricas específicas para la escritura académica.

Paltridge et al (2009:187) propusieron un modelo de escritura académica en un segundo idioma basado en el modelo de la escritura cognitiva y en el modelo de producción del lenguaje escrito de los predecesores, con perspectivas comprensivas, sociales y textuales (véase figura 2).

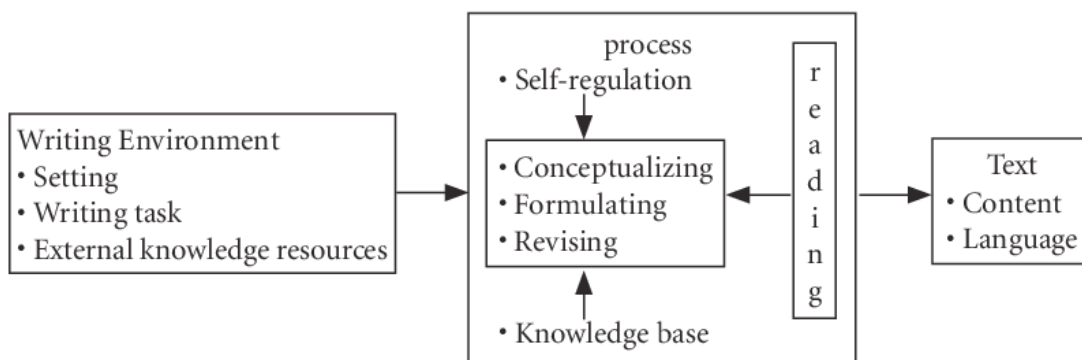


Figura 2: El modelo de escritura académica de Paltridge et al

Fuente: *Research Trends and Implications of Foreign Second Language Academic Writing*,

Bi J, Qin X.Q. y Song D.W, (2014:45)

毕劲, 秦晓晴, 宋德伟, p45, 国外二语学术写作研究趋势及其启示

La idea principal de este modelo es que los escritores académicos y los escritores generales deben pasar por diferentes subprocesos cognitivos internos para producir productos terminados. La escritura académica no es un comportamiento que ocurra en una burbuja, ya que también se verá afectado por los factores externos del escritor. Estos factores internos y externos trabajan juntos para producir el texto escrito. En la investigación de Bi, Qin y Song (2014:45), observaron que este modelo tiene tres características:

- (1). Se exploran todos los aspectos de la escritura académica en un segundo idioma desde múltiples perspectivas.
- (2). Se confirman los componentes de la escritura académica y el contenido de cada parte, y se proponen hipótesis sobre la relación entre ellos.
- (3). Se destaca el papel de los recursos de conocimiento externos.

Murray y Moore (2006:286) propusieron un modelo basado en la teoría de la motivación, la teoría de la participación y los datos empíricos (véase figura 3). En la investigación de Bi, Qin y Song (2014:45) se propone un modelo con dos características:

(1). Se enfatiza el papel de los factores motivacionales en la escritura académica.

(2). No es recursivo y la escritura se considera un proceso cíclico.

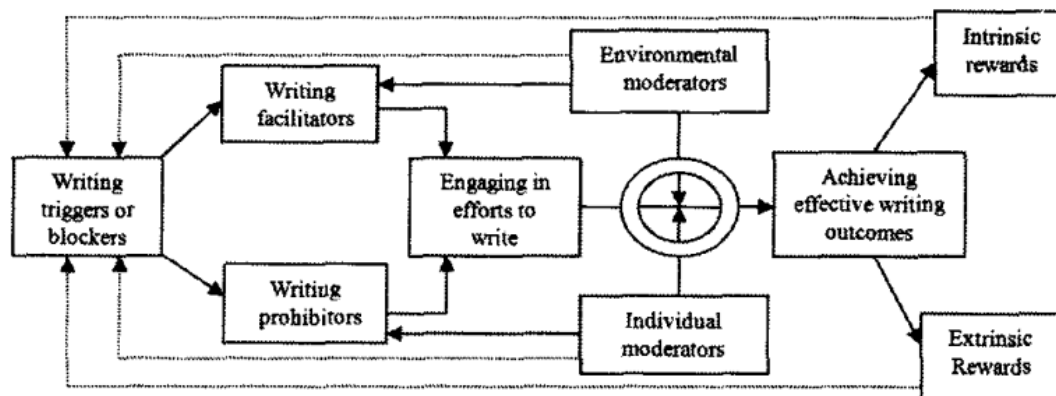


Figura 3: Modelo estructural de la escritura académica de Murray y Moore

Fuente: *Research Trends and Implications of Foreign Second Language Academic Writing*, Bi J, Qin X.Q. y Song D.W, (2014:45)

毕劲, 秦晓晴, 宋德伟, p45, 国外二语学术写作研究趋势及其启示

Por el contrario, Paltridge et al (2009:187) enfatizaron los factores internos de la escritura académica y la interacción entre los diversos procesos cognitivos de la escritura y su efecto sobre el texto. Sin embargo, Ellos prestaron menos atención a los factores externos, mientras que Murray y Moore (2006:286) sí que enfatizaron con los factores externos, así como con la interacción del proceso externo y las recompensas de la escritura académica. Por otro lado, Murray y Moore (2006:286) ignoraron el papel de los factores cognitivos, mientras que hacen una explicación breve sobre el texto del proceso interno y la función de la escritura académica. Por eso, Bi, Qin y Song (2014:45), consideran que estas dos teorías son complementarias y que la combinación de las dos ayudará a comprender completamente las características de la escritura académica.

2.3.2. La escritura académica del inglés

En general, la escritura académica del inglés trata de utilizar el inglés para la

escritura y la comunicación académicas. Cao y Lu (2016:490) creen que, para la escritura académica en inglés, el lenguaje utilizado debe ser complejo, formal, objetivo, y preciso, y también debe reflejar las características del lenguaje escrito y evitar el uso de lenguaje coloquial. La investigación de Cao y Lu (2016:490) demuestra que el lenguaje académico debe:

- (1). Evitar el uso de vocabulario coloquial, como ahora *"a lot of"*, *"thing"* y *"sort of"*.
- (2). Evitar usar las abreviaturas y las contracciones en inglés, como: *"can't"*, *"doesn't"* y *"shouldn't"*.
- (3). Evitar el uso de los verbos perifrásticos, como: *"carry on"* y *"find out"*.
- (4). Evitar usar subtítulos, números y formas de escritura en forma de lista y usar frases menos interrogativas al escribir.
- (5). Evitar usar pronombres muy subjetivos en primera persona, como: *"I"* y *"we"*, como también evitar usar la segunda persona *"you"* en la escritura académica.

Esta investigación se basará en estos estándares y seleccionará los parámetros relevantes.

2.4. El lenguaje coloquial

Yu (2008:107) explica que "En China, cada vez más personas están dirigiendo su atención al inglés coloquial y menos al inglés escrito". En su investigación, encontró que en los trabajos académicos de ILE (Inglés como lengua extranjera)

se empieza a aparecer con frecuencia el lenguaje coloquial. En esta sección se analiza principalmente la definición, las características del lenguaje y la tendencia coloquiales de la escritura académica de los alumnos chinos.

2.4.1. Definición del lenguaje coloquial

El lenguaje coloquial, a diferencia del lenguaje escrito, es el lenguaje utilizado en la comunicación oral, y es la primera forma de lenguaje comúnmente utilizada por los humanos. Pöchhacker (2016) cree que el lenguaje coloquial es "la forma del lenguaje original utilizada en la comunicación interpersonal directa (cara a cara)".

Como se define en el *Oxford Advanced Learner's English-Chinese Dictionary* (Wehmeier, 2004: 111), "El lenguaje coloquial es una palabra o una frase que se usa en un diálogo, pero no es una frase que se usa en la escritura formal".

Chafe (1982) señaló que el lenguaje coloquial está fragmentado, porque algunas unidades de pensamiento en el lenguaje coloquial a menudo están conectadas entre sí, pero otras veces no.

2.4.2. Las características del lenguaje coloquial

Las características del lenguaje coloquial son diversas y pueden clasificarse según diferentes estándares. A continuación, se analizan las características del lenguaje coloquial desde los niveles de vocabulario, sintaxis y texto.

(1). Características del vocabulario del lenguaje coloquial

El lenguaje coloquial tiene una baja diversidad léxica y una baja densidad léxica. La mayoría de las palabras que se utilizan en el lenguaje coloquial son cotidianas,

sencillas y específicas. También se suelen usar las abreviaturas y las contracciones en el lenguaje coloquial, como "I'm", "we're" y "can't" etc. Por otro lado, el lenguaje escrito necesita tener una diversidad léxica y una densidad léxica relativamente altas y se utilizan palabras más abstractas. Al mismo tiempo, Macaulay (2011:568) encontró que en el lenguaje coloquial aparecen más opiniones, mientras que en el lenguaje escrito se utilizan más estructuras pasivas y más formas nominales.

Según la investigación de Ure (1971:443), el nivel promedio del lenguaje coloquial es más bajo que en el caso del lenguaje escrito. Para ser precisos, en comparación con el lenguaje escrito, el lenguaje coloquial no tiene demasiado vocabulario y usa repetidamente un vocabulario simple. En general, este vocabulario solo se puede utilizar en conversaciones diarias para crear una atmósfera relajada entre el hablante y el oyente.

En la investigación de Ma y Wen (1999), demostraron la función del lenguaje coloquial mostrando que hay 50 contracciones por cada 1.000 palabras. Además, cuando la gente habla, suele dudar y hacer una pausa. En ese momento, los acrónimos informales o las abreviaturas a menudo se juegan un papel en hacer que el diálogo sea claro y lógico, relajar la atmósfera y proporcionar al hablante una buena estructura del contenido del diálogo.

(2). Las características sintácticas del lenguaje coloquial

En la clasificación de Krauthamer (1999:67), el uso de los pronombres personales se considera una de las características más destacadas del lenguaje coloquial a nivel morfológico. Además, Krauthamer (1999:67) señaló que el uso frecuente de los pronombres conduce a una interferencia coloquial ineficaz.

Biber (1991:1299) señala que el lenguaje coloquial es opuesto al lenguaje escrito.

Considera que el diálogo es una expresión coloquial típica, mientras que los artículos académicos son un estilo de escritura típico. Biber, Conrad y Reppen (1998:271) creen que la voz pasiva es menos frecuente en las conversaciones diarias, porque el lenguaje coloquial puede reflejar mejor el comportamiento, los pensamientos y las posiciones de las personas, y mostrar menos temas. Por el contrario, la forma coloquial pasiva suele verse en el lenguaje escrito, unas 8.500 veces por millón de palabras.

Hou (1988) propuso que entre las características del lenguaje coloquial se incluyeran el uso de muchos pronombres, el uso de muchas abreviaturas, el uso de una gran cantidad de verbos perifrásticos, el uso de *"and"* o *"but"* entre frases, y el uso de cláusulas menos complejas.

Al mismo tiempo, el lenguaje coloquial no tiene una estructura complicada. En el lenguaje coloquial, las personas tienden a usar más puntos suspensivos, las emisiones completas se expresan de forma nominal y usan patrones de oración más simples. Por otro lado, Qin (2002:30) creyó que el lenguaje coloquial tiene una estructura de oración más flexible. Liu (2007:42), consideró que las personas son más proclives a usar cláusulas simples en el lenguaje coloquial, con las que las personas tendrán un diálogo. En general, la gente suele elije frases cortas y simples para expresar sus opiniones.

(3). Las características del texto del lenguaje coloquial

El lenguaje coloquial es diferente del lenguaje escrito en el nivel del texto. Según la investigación de Halliday y Hasan (2002:317), en las conversaciones diarias las personas suelen expresar sus opiniones de forma más libre o repiten directamente las mismas frases en lugar de utilizar diferentes expresiones.

2.5. La escritura académica del inglés en China

Wen (2003:6) señaló que, en China, la enseñanza e investigación de la escritura académica en inglés ha recibido una atención cada vez mayor.

2.5.1. La situación actual de la escritura académica del inglés en China

En comparación con la investigación extranjera, la investigación sobre la escritura académica del inglés en China es muy escasa. Desde 1970 hasta 1990 solo se han llevado a cabo trabajos de introducción y algunas investigaciones especulativas. Como indica Rong (1991:5):

A mediados de 1990, con la profundización de las reformas en la enseñanza del inglés, se incluyó la escritura académica, la cual se convirtió en un tema candente en los círculos académicos chinos. Por otro lado, la investigación sobre la escritura académica de inglés se basó cada vez menos en datos, situación que continuó hasta principios de este siglo.

Adicionalmente, Qin (2002:26) apunta:

Desde 2000 hasta 2007, se publicaron un total de 570 artículos de investigación sobre la escritura académica en 8 revistas principales de idiomas extranjeros en China, pero todavía existe una grave escasez de literatura específica para la escritura académica.

Xu, Guo y Xu (2007:48) creen que la investigación sobre la relevancia del contenido a expresar en la escritura, sobre el pensamiento investigador sobre y la conexión profunda entre los elementos de la tesis no es suficientemente profunda. Ellos (2007) apuntan:

Solo un pequeño número de instituciones chinas de educación superior ofrecen cursos de escritura académica en inglés. Los

cursos se desarrollan en torno a formas lingüísticas y a cómo organizar idiomas, o en torno a cómo realizar una investigación académica.

En resumen, la investigación sobre la escritura académica del inglés en China tiene un largo camino por recorrer y aún quedan muchos problemas por resolver.

2.5.2. La tendencia coloquial de la escritura académica de los alumnos chinos

Muchos eruditos han estudiado el inglés escrito por los alumnos de inglés con diferentes niveles y antecedentes. Según Aquinaso (1982), en la escritura formal del inglés, a menudo ocurren un gran número de usos coloquiales.

Según los parámetros utilizados por Cobb (2003), los resultados muestran que, en la escritura académica de los alumnos chinos de inglés como lengua extranjera, el porcentaje de la función coloquial es muy elevado, siendo similar al de los alumnos de inglés como lengua extranjera en otros países. Esto demuestra que las expresiones coloquiales no se ven afectadas por el origen o la cultura nativa, ya que tienen sus propias características en la escritura académica de los alumnos avanzados.

Gilquin y Paquot (2008:48) señalaron que no importa qué idioma nativo usen los alumnos del inglés, ya que todos tienden a usar una gran cantidad de lenguaje coloquial en su escritura académica. Wen (2003:6) añade:

La escritura académica del inglés en el alumno es una mezcla de lenguaje coloquial y escrito. A veces, se muestran múltiples características coloquiales en el lenguaje escrito, y algunas veces, se muestran las características escritas en el lenguaje coloquial.

Con la ayuda de algunos corpus clave, Liu (2005:110) realizó una investigación

sobre las características coloquiales de la escritura académica del inglés de los alumnos universitarios chinos. Los resultados de la investigación mostraron que las características coloquiales son existentes en la escritura académica. Yu (2008:107) explica que:

El fenómeno del coloquialismo a menudo aparece en las escrituras académicas de los alumnos chinos avanzados. Por lo tanto, se puede ver que las características coloquiales todavía se pueden encontrar en las escrituras académicas de los alumnos con mayor dominio del inglés.

Wen, Ding y Wang (2003) realizaron una investigación sobre el lenguaje coloquial en la escritura académica de los alumnos chinos con inglés avanzado. Los resultados de la investigación demostraron que los alumnos chinos avanzados de inglés habían estudiado las diferencias entre el lenguaje coloquial y el escrito, pero el lenguaje coloquial todavía era usado con frecuencia en las escrituras académicas de los alumnos chinos. De igual forma, Wen, Ding y Wang (2003) consideran que los alumnos chinos del inglés no pueden usar el lenguaje coloquial y escrito de manera apropiada en la comunicación diaria y en la escritura académica.

El enfoque de esta investigación es analizar las razones de este fenómeno y hacer sugerencias relevantes a través del estudio del uso del lenguaje coloquial en la escritura académica del inglés.

3. METODOLOGÍA

Para medir el grado de inglés coloquial, se ha adoptado una combinación de métodos de investigación cuantitativos y cualitativos y métodos de análisis comparativos. Desde una perspectiva holística, el objeto de investigación son los alumnos nativos de inglés y los alumnos chinos de inglés avanzado. En el presente capítulo se presentan los métodos, el corpus y el diseño de investigación.

3.1. Métodos

Todos los datos de este estudio provienen del corpus y el propósito de la investigación se ha logrado mediante el análisis comparativo del BASE, BAWE y de dos corpus (LM y TFM) construidos por el autor. Debido a que los cuatro corpus son de diferentes tamaños, el análisis del Capítulo 4 se basa en un análisis estadístico por millón.

3.1.1. El método basado en el análisis de corpus

El método de investigación y de análisis basado en el análisis de corpus proporciona una nueva perspectiva para el aprendizaje de idiomas y para la enseñanza de lenguas extranjeras. Muchos investigadores han definido el concepto de corpus como una gran cantidad de texto. Leech (1991) definió un corpus como una base de datos; al mismo tiempo, Biber (1991) identificó cuatro características básicas del análisis basado en el corpus: una gran colección de principios, experiencia y texto natural, una combinación de métodos cuantitativo y cualitativo, y una tecnología automática e interactiva.

En los últimos años, ha habido algunos estudios relacionados con el uso de los métodos de corpus para analizar las características coloquiales utilizadas por los

alumnos del inglés en la escritura académica. Por ejemplo, Wen (2003) se basó en el corpus para estudiar la tendencia coloquial de los alumnos chinos del inglés en el lenguaje escrito, ya que una de las características de los académicos de inglés con alto nivel es que tienen una tendencia coloquial en su escritura académica. La manifestación específica es el uso excesivo de pronombres personales de primera y segunda persona, un uso en exceso de vocabulario de alta frecuencia y el uso de palabras que carecen de distinción y diversidad. La investigación de Feng y Wu (2011) demuestra que la mayoría de los alumnos chinos usan un inglés coloquial en su escritura académica.

En esta investigación se han utilizado cuatro corpus recopilados y establecidos por el autor, y se ha utilizado el método de análisis comparativo del corpus para obtener datos relevantes de manera rápida y efectiva. Este método apareció en la década de 1990 y se usa a menudo para comparar la producción literaria de alumnos de idiomas extranjeros con diferentes antecedentes lingüísticos. El uso de la comparación del corpus para estudiar a los académicos de inglés con un alto nivel comenzó en Europa. En 1998, el *"Learner English on Computer"*, escrito por Granger (1998:10) recopilaba los resultados de investigación pionera de académicos europeos en análisis comparativo de corpus.

Hoy en día, muchos académicos están dispuestos a adoptar el método del análisis comparativo del corpus en sus investigaciones. Su característica principal es que puede reflejar las características lingüísticas de los alumnos de una manera conveniente, objetiva y directa. Cobb (2003:395) señaló que el *"Computer Learning English"* de Granger es "el momento más importante en el desarrollo de este nuevo método de análisis de corpus comparativo".

3.1.2. El método cuantitativo y cualitativo

En esta investigación se combinan los métodos de análisis cuantitativo y

cualitativo para hacer una investigación más objetiva y extensa. El análisis cualitativo es la base del análisis cuantitativo, ya que los datos del análisis cualitativo sin ningún análisis cuantitativo son poco fiables e inexactos. La combinación de ambos puede ayudar a que esta investigación sea más precisa y científica, y que se pueda llegar a conclusiones extensas y profundas.

3.2. Corpus

Los corpus usados en esta investigación son los corpus profesionales BASE y BAWE, y dos corpus pequeños del propio autor (véase figura 4). Los datos son verdaderos y confiables y se pueden utilizar para el análisis y el resumen de esta investigación.



RECENTLY USED CORPORA			NEW CORPUS
British Academic Spoken English Corpus (BASE)	English	1,477,281	
British Academic Written English Corpus (BAWE)	English	6,968,089	
Corpus de pregrado en la lengua materna (LM)	English	500,519	
Corpus de trabajos fin de Máster (TFM)	English	325,439	

Figura 4: Los cuatro corpus involucrados en esta investigación

Fuente: <https://app.sketchengine.eu>

3.2.1. *British Academic Spoken English Corpus (BASE)*

El proyecto *British Academic English Spoken Corpus (BASE)* se llevó a cabo en la Universidad de Warwick y la Universidad de Reading de 2000 a 2005, el cual fue dirigido por Hilary Nessi (2001) (Universidad de Warwick), y presentado por

Paul Thompson (2001) (Universidad de Reading). Además, fue financiado por la Universidad de Warwick y Reading, BALEAP, EURALEX y la Academia Británica. Véase detalle en la tabla 3:

<i>Tokens</i>	1.756.545
Palabras	1.477.281
Frases	193

Tabla 3: Información básica del corpus BASE

Fuente: <https://app.sketchengine.eu>

BASE consta de 160 conferencias y 40 seminarios grabados en varios departamentos (grabación de videos de la Universidad de Warwick y grabación de audios de la Universidad de Reading). Contiene un total de 1.756.545 tokens (conferencias y seminarios) y el contenido involucra cuatro grupos (humanidades y artes, biología y medicina, ciencias físicas e investigación social y ciencia), los cuales tienen 40 conferencias y 10 seminarios. Este corpus incluye: grabaciones en video, audios de conferencias, seminarios grabados por diversos departamentos universitarios, grabaciones de vídeo de conferencias académicas, entrevistas con el personal académico sobre su trabajo y campos académicos (grabaciones de audio, transcripciones, notas de entrevistas), etc. El BASE también se puede encontrar a través de la interfaz de búsqueda del *Sketch Engine*. Esta versión es un archivo separado que describe las búsquedas relacionadas a través de las anotaciones de esta versión.

3.2.2. British Academic Written English Corpus (BAWE)

El *British Academic Written English Corpus* (BAWE) es una parte de la '*An Investigation of Genres of Assessed Writing in British Higher Education*'. Su número de recurso es el 2.539 y su objetivo es ayudar a los alumnos a aprender el género académico del inglés. Para lograr este propósito proporciona una biblioteca de recursos para aprender la escritura académica. El corpus es el

resultado de un proyecto de investigación sobre el género de la escritura académica en la educación superior británica. Además, el BAWE fue financiado por ESRC (*Economic and Social Research Council*) en 2004, y se creó juntamente con la Universidad de Warwick, la Universidad de Reading y la Universidad de Oxford Brookes entre 2004 y 2007.

El proceso de recopilación del corpus en BAWE proviene de 2.761 artículos académicos de alumnos de elevado nivel recopilados por las 3 universidades (Oxford Brooks, Universidad de Warwick, Universidad de Reading). Véase detalle en la tabla 4:

<i>Tokens</i>	8.336.262
Palabras	6.968.089
Frases	293.113

Tabla 4: Información básica del corpus BAWE

Fuente: <https://app.sketchengine.eu>

El contenido involucra cuatro áreas temáticas amplias (artes y humanidades, ciencias sociales, ciencias de la vida y ciencias físicas) y cuatro niveles de aprendizaje, y su contenido representa 30 materias principales. Por otro lado, el BAWE se puede encontrar a través de la interfaz de búsqueda de *Sketch Engine*. Esta versión es un archivo separado que describe las búsquedas relacionadas a través de las anotaciones de esta versión.

3.2.3. Corpus de pregrado en la lengua materna (LM)

Este corpus está compuesto por textos originales y auténticos recopilados para este trabajo. El propósito es medir el uso del inglés coloquial en la escritura académica de los alumnos nativos de pregrado y compararlo con el BAWE para analizar si este fenómeno cambiará con la mejora del dominio del inglés. El corpus completo tiene un total de 500.519 palabras (véase tabla 5). Se

denominará "LM" a la lengua materna a partir de ahora para facilitar la comprensión del texto.

<i>Tokens</i>	566.936
Palabras	500.519
Frases	21.977

Tabla 5: Información básica del corpus LM

Fuente: <https://app.sketchengine.eu>

El corpus es una recopilación de 10 ensayos de alumnos de la Universidad de Michigan. Estos textos son todos sobre escritura académica y provienen de la biblioteca de la Universidad de Michigan; y el corpus de *Louvain Corpus of Native English Essays* (LOCNESS).

El LOCNESS fue construido por el Centro de Lingüística de Corpus en Inglés de la Universidad Católica de Lovaina (Bélgica) y se puso a disposición del público en el año 1998. Los textos del corpus son ensayos producidos por hablantes nativos británicos y estadounidenses entre 1991 y 1995. El contenido incluye una tesis de nivel A para los alumnos británicos (60.209 palabras), las tesis de universitarios británicos (95.695 palabras) y las tesis de alumnos universitarios estadounidenses (168.400 palabras), con un total de 314.317 palabras. Para poder compararlo con el corpus de los alumnos internacionales de inglés, el corpus recopilado por los alumnos universitarios británicos y estadounidenses se ha limitado a los ensayos argumentativos.

3.2.4. Corpus de trabajos fin de Máster (TFM)

Este corpus también está compuesto por escritos recopilados para este trabajo. Su propósito es medir el uso del inglés coloquial en la escritura académica de los alumnos chinos de inglés. El corpus completo tiene un total de 325.439 palabras (véase tabla 6). Para facilitar la comprensión, los trabajos fin de Máster,

se denominarán "TFM" a partir de ahora.

<i>Tokens</i>	386.714
Palabras	325.439
Frases	13.998

Tabla 6: Información básica del corpus TFM

Fuente: <https://app.sketchengine.eu>

El corpus es una recopilación de 20 tesis de maestría de varias universidades de China recopiladas por el autor. Estos textos de maestría también tratan sobre la escritura académica y todas provienen del CNKI (*China National Knowledge Infrastructure*). Se debe enfatizar que, para reflejar los datos relevantes de una manera más intuitiva y efectiva, el autor solo ha adoptado el texto principal de cada artículo (excluyendo la portada, los agradecimientos, el resumen, el índice y la bibliografía) en el proceso de recopilación.

3.3. Diseño experimental

3.3.1. Parámetros utilizados en el estudio

En esta investigación se analizarán 3 aspectos y 8 parámetros, los cuales pueden probar el grado del coloquialismo en la escritura académica en diversos grados. Estos aspectos se explican a continuación.

3.3.1.1. La visibilidad del escritor

En el inglés coloquial, a menudo, se usan la primera persona, la segunda persona y otros recursos para resaltar una cierta conexión entre el contenido de las palabras, la audiencia y el hablante. En el campo de la investigación del análisis del discurso, Chafe (1982:36) intentó distinguir el lenguaje escrito y lenguaje coloquial a través de una determinada característica del lenguaje.

Como señaló este investigador, la mayor diferencia entre el lenguaje escrito y el lenguaje coloquial es que "el hablante interactúa directamente con la audiencia, pero el autor no". En comparación con el lenguaje escrito formal, el lenguaje coloquial contiene muchos pronombres en primera persona y citas directas.

La visibilidad del escritor refleja la mención de sí mismo del autor y la comunicación con los lectores por escrito, y a menudo, se utiliza para medir el grado de coloquialismo. En el marco del modelo principal de indicadores de medición, se acepta la investigación de Petch-Tyson (1998), en la cual la frecuencia de uso de los pronombres personales de primera y segunda persona, las palabras difusas, las partículas de énfasis y las palabras demostrativas de situación de escritura, son la base principal para analizar el grado de coloquialismo de la escritura académica.

3.3.1.2. El nivel léxico

Zhang (2018:28) cree que "el nivel léxico se puede dividir para reflejar la redacción y el profesionalismo de una escritura académica." Coxhead (2000: 221) considera que "el vocabulario es reconocido como la característica más notable que difiere todos los registros". Leki y Carson (1994: 92) explicaron que "el vocabulario es uno de los aspectos lingüísticos que los alumnos deben mejorar". Por otro lado, Laufer y Nation (1995: 307) propusieron que "una composición bien escrita hace un uso eficaz del vocabulario". En términos del nivel léxico, se analizarán 4 parámetros.

3.3.1.2.1. Diversidad léxica (TTR)

La diversidad léxica siempre ha sido un parámetro importante para medir la profesionalidad de una escritura académica. Read (2002) opina que la riqueza léxica es un criterio importante en la prueba de escritura del ILE. Cobb (2003:

399) propone que “la vaguedad surge de la falta de vocabulario y que el uso excesivo de vocabulario básico indica una infrautilización de un vocabulario más rico, preciso y variado”. En una escritura académica, la diversidad de vocabulario representa la capacidad académica del autor. Cuanto mayor sea la diversidad del vocabulario, menor es la repetitividad de las palabras utilizadas por el autor, lo cual hace que la escritura académica sea más formal y que sus características coloquiales sean menos obvias. Por el contrario, cuanto menor sea la diversidad de vocabulario, las características coloquiales de la escritura académica serán más evidentes.

En el análisis del corpus, el "TTR" es un indicador importante para medir la diversidad de vocabulario de un corpus. Biber (1991:182) explica que "el TTR en lenguaje coloquial es menor que el TTR en lenguaje escrito porque es espontáneo, y casi no hay tiempo de planificar y cambiar la elección de las palabras ". Por lo tanto, cuanto menor sea el valor de TTR, mayor será el grado coloquial del texto.

3.3.1.2.2. Densidad léxica

La densidad léxica se refiere a la proporción de palabras de contenido en un texto. Este concepto refleja la cantidad de información difundida por los materiales de texto. Por lo tanto, es un indicador importante para medir el contenido de información de un texto. Las palabras son directamente proporcionales a la cantidad de información, lo cual puede reflejar el género del texto, que es directamente proporcional al grado de la escritura académica. Cuanto mayor sea la densidad léxica, mayor es el grado de escritura, mayor es el grado de profesionalismo y menor cantidad de características coloquiales aparecen en la escritura académica.

Ure (1971) propuso por primera vez el concepto de la densidad léxica. Su método

de cálculo para la densidad léxica es el siguiente: densidad léxica = (palabras de contenido / palabras totales) * 100. Las palabras de contenido se suelen ser de contenido real, incluyendo sustantivos, verbos, adjetivos y adverbios. Ure (1971) utiliza diversas fórmulas para realizar estadísticas y análisis de los escritos recopilados y concluye que la densidad léxica de un texto coloquial está por debajo del 40% y que la densidad léxica del texto escrito está por encima del 40%. En el pasado, algunos investigadores también estudiaron el lenguaje coloquial y escrito desde la perspectiva de la densidad léxica. Por ejemplo, Song (2017) descubrió que la densidad léxica del lenguaje coloquial y el lenguaje escrito de su corpus eran del 61% y el 75%, respectivamente.

3.3.1.2.3. Vocabulario coloquial

Paquot (2010:150) define la característica principal del modo coloquial en los trabajos académicos como el "uso excesivo de al menos un elemento del vocabulario". Por lo tanto, podría resumirse como: el uso excesivo de un elemento del vocabulario coloquial en el lenguaje escrito o el uso excesivo del vocabulario coloquial en el lenguaje escrito. Este estudio se refiere a Paquot (2010) y en él se enumeran a 22 palabras/expresiones coloquiales en la escritura académica (véase figura 5).

Rhetorical function	Spoken-like overused lexical item
Exemplification	<i>like</i> <i>thanks to</i>
Cause and effect	<i>so</i> <i>because</i> <i>that/this is why</i>
Comparison and contrast	<i>look like</i> <i>like</i>
Concession	sentence-final adverb <i>though</i>
Adding information	sentence-initial <i>And</i> the adverb <i>besides</i> <i>I think</i>
Expressing personal opinion	<i>to my mind</i> <i>from my point of view</i> <i>it seems to me</i>
Expressing possibility and certainty	<i>really</i> <i>of course</i> <i>absolutely</i> <i>maybe</i> <i>I would like to/want/am going to talk about</i>
Introducing topics and ideas	<i>thing</i> <i>by the way</i>
Listing items	<i>first of all</i>
Quoting and reporting	<i>say</i>

Figura 5: Elementos del vocabulario más utilizados en la categoría coloquial de cada función retórica de Paquot

Fuente: *Spoken-Like Features in the Academic Writing of Advanced Chinese Learners of English*, Sun Y.Z, (2012:10)

孙亚真, p10, 中国高水平英语学习者学术写作中的口语化特征

3.3.1.2.4. Abreviaturas y contracciones

Biber (1999: 182) explica que "las abreviaturas y las contracciones se producen omitiendo letras y sonidos internos, los cuales suelen aparecer en las conversaciones en inglés y tienen fuertes características coloquiales".

En la escritura académica en inglés formal, los escritores a menudo deben cumplir con muchas reglas y formas estrictas, por lo que no es apropiado usar ni abreviaturas ni contracciones en este tipo de escritura más formal.

3.3.1.3. El nivel sintáctico

Chafe (1982) considera que el inglés escrito se caracteriza por la integración. En comparación con el inglés coloquial, el inglés escrito tiende a utilizar estructuras más complejas, como ahora patrones de frases fijas. Por otro lado, Macaulay

(2011:517) explica que "el inglés escrito tiende a utilizar frases más largas, con el uso de cláusulas." En términos del nivel sintáctico, se han elegido 3 parámetros para analizar, los cuales se explican a continuación.

3.3.1.3.1. Longitud promedio de la frase

Al analizar las características coloquiales de la escritura académica, la longitud promedio de las frases se puede utilizar como el estándar de medición. Carroll (1979:443) concluyó con que "la longitud promedio de una frase es el estándar más objetivo para medir la madurez del lenguaje". De forma similar, Huang y Jia (2013) observan que muchos estudios adoptan la longitud promedio de la frase para medir la complejidad del inglés coloquial y escrito.

En general, cuanto mayor sea el valor promedio de la longitud de una frase, más larga será la frase y más compleja será la estructura del artículo académico, lo cual también podría indicar que el artículo académico tiene un mayor grado de profesionalismo y que su lenguaje es más escrito y estandarizado. Por el contrario, las habilidades coloquiales son poco aplicables en la escritura académica, debido a que las frases largas con estructuras complejas raramente se utilizan para expresar opiniones subjetivas en el lenguaje coloquial.

3.3.1.3.2. Estructura "*There be*"

"*There be*" es una estructura utilizada para expresar una existencia o para presentar una nueva información del discurso en un texto. Las investigaciones sobre la estructura "*There be*" se han centrado en su abuso en la escritura y en su uso en el lenguaje coloquial. Song (2008:79) estudió la estructura "*There be*" en la escritura de los alumnos chinos de inglés y exploró la relación entre su uso excesivo y la transferencia de la lengua materna. Chen (2014:56) explica que:

Aunque los alumnos hayan aprendido a usar la estructura *'There be'* para revelar la existencia de los objetos, su abuso aumenta el coloquialismo de sus escrituras académicas.

3.3.1.3.3. Verbos perifrásticos

En inglés, los verbos perifrásticos son frases que expresan acciones, como ahora *"take on"* o *"make up"*. Biber (1991:403) explica que "el término se aplica a dos o tres estructuras diferentes pero relacionadas: verbos y partículas y / o preposiciones, los cuales forman una unidad semántica".

Los verbos perifrásticos son más comunes en el lenguaje coloquial y tienen un fuerte componente coloquial. Comparados con sus correspondientes verbos individuales, son menos formales. En la escritura formal, es preferible usar verbos únicos, como *"examine"* y *"devise"*, en lugar de los verbos perifrásticos, como ahora *"look at"* y *"come up with"*. Por lo tanto, los verbos perifrásticos deben evitarse en la escritura académica.

3.3.2. Herramientas utilizadas en el estudio

Cuando es necesario analizar el corpus, con tal de poder extraer información útil, se hace necesario utilizar herramientas de cálculo para su correcto procesamiento. En esta investigación se han utilizado el *Sketch Engine* (2003), *AntConc* (2004) y *WordSmith 4.0* (1996) para obtener y discutir los datos científicos.

3.3.2.1. Sketch Engine

Sketch Engine es un software de análisis de texto y un administrador de corpus desarrollado por el *Lexical Computing Limited* desde 2003. Su propósito es permitir que las personas que estudian los comportamientos lingüísticos (autores

de diccionarios, investigadores de lingüística de corpus, traductores, alumnos de idiomas, etc.) puedan aprender idiomas mediante búsquedas de texto y su lema es "corpus para todos".

Sketch Engine proporciona corpus en más de 90 idiomas, incluyendo los dos corpus de inglés académico utilizados en esta investigación, BAWE y BASE. Tanto lingüistas, traductores, así como alumnos, profesores y amantes del idioma pueden utilizarlo. Los usuarios pueden incluso crear su propia base de datos en un idioma en especial (corpus de texto) y utilizar la herramienta de análisis de texto incorporada para analizarla, al igual que los corpus LM y TFM del autor. El propósito principal del *Sketch Engine* siempre ha sido su tecnología de diccionario, en la cual se pueden generar listas de palabras para varios idiomas (y muchas otras listas: n-gramas, palabras clave y términos) y también puede usarse para la búsqueda de términos. Por lo tanto, el *Sketch Engine* ha proporcionado una gran ayuda para esta investigación.

En esta investigación, como herramienta principal para la recolección y el análisis de datos, se ha usado el *Sketch Engine* para recopilar toda la información básica y los datos de los cuatro corpus: BASE, BAWE, LM y TFM. Se han incluido "Tokens", "Palabras", "Frases", así como sustantivos, verbos, adjetivos, adverbios y otros datos necesarios en esta investigación. A través de la función "Concordancia", se puede encontrar rápidamente la frecuencia y la frecuencia por millón de un vocabulario en específico, como también del vocabulario coloquial de alta frecuencia, como ahora "so", "because" y "I think" etc; abreviaturas, contracciones como "-s", "-re" y "-n't" y verbos perifrásticos como "go on" y "find out" etc, que deben estudiarse en la investigación. Además, en la presente investigación, todos los ejemplos provienen de *Sketch Engine*, lo cual ha sido de gran ayuda para calcular la visibilidad del escritor, la densidad léxica y la frecuencia del vocabulario de los cuatro corpus.

3.3.2.2. AntConc

AntConc es un software gratuito desarrollado por Laurence Anthony (2014), de la Universidad de Waseda. Este software se utiliza para implementar la lingüística del corpus y el aprendizaje basado en los datos creado por Laurence Anthony. *AntConc* es fácil de usar, ya que, a través de la lista de herramientas, todas las palabras del corpus se pueden contar y mostrar en una lista ordenada. Esto puede ayudar a los usuarios a encontrar rápidamente las palabras más utilizadas en el corpus y la información del corpus que necesitan, y así permitir una investigación más aplicable y eficiente.

En esta investigación, el *AntConc*, como herramienta para ayudar al *Sketch Engine*, se ha utilizado principalmente para analizar los datos relevantes de los corpus LM y TFM que recopilé y construí. A través de la lista de herramientas, todas las palabras del corpus se pueden contar y mostrar en una lista ordenada. Además, se pueden encontrar rápidamente las palabras más utilizadas en los dos corpus y la información del corpus que necesitan, a fin de lograr una investigación más aplicable y eficaz. A través de la función "Lista de palabras", se pueden encontrar rápidamente tanto las palabras de alta frecuencia en ambos corpus y su uso de colocación, como el parámetro "verbo perifrástico" necesario en esta investigación.

3.3.2.3. WordSmith 4.0

WordSmith es un software de recuperación de corpus profesional, desarrollado por Mike Scott (1996), una lingüista de la Universidad de Liverpool, Reino Unido. Este software incluye *Concord* (una herramienta de recuperación contextual); *WordList* (una herramienta de recuperación de lista de frecuencia de palabras); *KeyWords* (una herramienta de recuperación de palabras clave); *Splitter* (una herramienta de segmentación de texto); *Text Converter* (una herramienta de

reemplazo de texto) y *Viewer* (una herramienta de búsqueda de texto). Los primeros tres programas son las principales herramientas de recuperación de texto y los últimos tres son herramientas auxiliares. Cobb (2003) explica que "*WordSmith* es la herramienta más utilizada para el análisis del corpus general". La versión de *WordSmith 4.0* utilizada por el autor es completamente gratuita, fácil de operar y en la cual es fácil de consultar datos relacionados.

En esta investigación se ha utilizado principalmente la función vocabulario para dividir el vocabulario de los dos corpus LM y TFM, y se han recopilado datos estadísticos de cada corpus, como ahora tokens, tipos y frases, con el fin de calcular el TTR y la longitud promedio de la frase de LM y TFM.

4. RESULTADOS Y DISCUSIÓN

En este capítulo se han utilizado los métodos y las herramientas de investigación propuestos en el Capítulo 3 para presentar los resultados de la investigación y las razones detrás de esta discusión. En este capítulo se ha utilizado un método de análisis comparativo para comparar las similitudes y las diferencias de los cuatro corpus. Al analizar la visibilidad del escritor, el nivel léxico y el nivel sintáctico en el corpus, se han analizado los coloquialismos en la escritura académica en inglés de los alumnos chinos.

4.1. Resultados de los coloquialismos de visibilidad del escritor en el inglés académico

En esta sección, se han seleccionado algunas características de la visibilidad del escritor como estándares y se ha referido al estándar de medición de Han (2013:31) (véase figura 6).

Table 5.1 The keywords used in the data retrieval

category	words
First person reference (singular/plural)	I, I*, me, my, mine, we, we*, us, our, ours
Second person reference (specific, general)	you, you*, your, yours
fuzziness words	kind of, sort of, and so on
references to situation	here, now, this essay, this article
speaker's mental processes	think, believe
emphatic particles	just, really
monitoring of information flow	you know, I mean

Figura 6: Características de la visibilidad del escritor de Han

Fuente: *Writer/reader Visibility in Chinese Advanced ILE Learner Argumentative Writings- A Corpus-based Study*, Han, Y.N, (2013:31)

韩娅楠, p31, 中国高级英语学习者议论文中的作者/读者显现度--基于语料库的研究

Al comparar las diferencias cuantitativas y cualitativas, entre los niveles de

visibilidad del escritor en los cuatro corpus, se ha podido analizar el grado de coloquialismo en la escritura académica. Véanse los resultados en la tabla 7:

Característica	BASE (palabras: 1.477.281)	BAWE (palabras: 6.968.089)	LM (palabras: 500.519)	TFM (palabras: 325.439)
Pronombres en primera persona (singular) <i>I, Me, My, Mine</i>	24.404	19.047	3.114	548
Pronombres en primera persona (plural) <i>We, Us, Our, Ours</i>	19.812	21.164	2.219	1.279
Pronombres en segunda persona (singular y plural) <i>You, Your, Yours</i>	36.333	4.885	1.146	175
Total	80.549(545)	45.096(65)	6.479(129)	2.002(62)
Palabras vagas <i>Kind of, Sort of, And so on</i>	5.922	1.331	231	293
Referencias a la situación <i>Here, Now, This essay, This article</i>	7.721	6.159	560	112
Proceso mental del hablante <i>Think, Believe</i>	6.457	5.291	1.011	327
Partículas enfáticas <i>Just, Really</i>	8.313	3.677	746	145
Seguimiento del flujo de información <i>You know, I mean</i>	5.457	87	39	8
Total	33.870(229)	16.545(24)	2.587(52)	885(27)
Total	113.969(771)	61.641(88)	9.066(181)	2.887(89)

Nota: El número entre paréntesis es la frecuencia por millón.

Tabla 7: Visibilidad del escritor de los 4 corpus

Fuente: Todos los datos provienen de *Sketch Engine*

La Tabla 7 enumera las características de visibilidad del escritor de los cuatro

corpus: BASE, BAWE, LM y TFM. Desde la perspectiva de las siete características específicas, los cuatro corpus muestran distintos grados de tendencia coloquial y de características específicas, utilizadas de manera desigual. Además, se puede ver que la razón principal de la diferencia en las características de visibilidad del escritor es el uso de los pronombres personales, mientras que la diferencia en otros factores no es significativa. Por lo tanto, el autor ha observado las diferencias en la visibilidad del escritor en los cuatro corpus del uso específico de los pronombres personales.

4.1.1. Análisis de BASE

En la Tabla 7, se puede ver que BASE tiene características de visibilidad del escritor. Como un corpus de lenguaje coloquial, BASE tiene 771 (por millón) de este tipo, lo que representa más que la suma de los otros tres corpus escritos, ya que es 8,7 veces más representativo que BAWE, 4,3 veces que LM y 8,6 veces que TFM. En la tabla de características de visibilidad del escritor de BASE, los pronombres personales de primera / segunda persona son las características más importantes que se presentan en la tabla 7. Entre las 771 características, el uso del primer / segundo pronombre personal es de 545 (por millón). En el uso de los pronombres en primera persona, "I" apareció un total de 21.643 veces (147 por millón), ya que es el pronombre en primera persona más utilizado, lo cual indica que los alumnos nativos tienden a utilizar la forma singular para expresarse. Este punto de vista refleja una identidad más clara del autor y un mayor grado de responsabilidad. El uso específico en BASE es el siguiente:

P. ej.

(1). I wanted to say to you this morning is that I think today's session is going to be the last of the lecture sessions for the course or if it isn't there'll only be about ten minutes next week.

(2). I know I've said that about ten times, but people do sometimes forget it even though I say it ten times.

("I"= El escritor expresa sus ideas a otros lectores, destacando la identidad del

escritor.)

En el uso de pronombres de segunda persona, "you" fue el pronombre que apareció con mayor frecuencia, con un total de 33.127 (224 por millón) veces. Según las investigaciones de Petch-Tyson (1998:108), el pronombre en segunda persona "you" refleja la interacción entre el autor y el lector, y su uso excesivo puede hacer que el artículo parezca informal. Por lo tanto, se puede reflejar el alto grado de interacción entre el autor y el lector en este corpus de lenguaje coloquial, destacando la identidad del lector. El uso específico en BASE es el siguiente:

P. ej.

(1). *As I say you can convert to something if you want to after the so-called impact in other words once the clutch has been engaged then you've got the two systems rotating together.*

(2). *You can easily work out from the dynamics where that position happens to be in the case of a cricket bat.*

("You"= La interacción entre el escritor y el lector, destacando la identidad del lector.)

Además, en los otros parámetros restantes, "think", "just", "now", "sort of" y "here" son las palabras y las frases más utilizadas, lo que también demuestra, que, en BASE, la mención del autor y la interacción con los lectores son más frecuentes. Además, también se muestra que hay un fenómeno muy obvio de pronombres personales de primera y segunda persona en el corpus del lenguaje coloquial. A continuación, el autor comparará y analizará los tres corpus escritos basándose en los datos de las características de BASE. El uso específico en BASE es el siguiente:

P. ej.

(1). *I think I mentioned to you last week the purposes of this, so you know what's happening.*

(2). *Now I'm going to work in some slightly unorthodox units here just for convenience.*

- (3). *When it's sort of it will be easier to throw under across.*
(4). *I think we'll carry on as normal and ignore what's going on over here.*

4.1.2. Análisis de BAWE

En la Tabla 7, se puede observar, como en un corpus de inglés académico, las características de visibilidad del escritor de BAWE son significativamente diferentes que en BASE. Entre ellos, solo hay 88 (por millón), lo cual indica que la visibilidad del escritor de BAWE es la más baja de los cuatro corpus. Por lo tanto, se demuestra que, entre los cuatro corpus, BAWE contiene el lenguaje coloquial más bajo y el texto escrito más formal. El uso específico en BAWE es el siguiente:

P. ej.

(1). *I think that there is definitely some heritability of intelligence in the white population.*

(2). *I believe there may be an initial flaw in Blackmore's argument in that she appears to take the existence of memes for granted and expounds her detailed argument from this premise when this fact is not necessarily clear.*

("I"= El escritor expresa su idea a otros lectores, destacando la identidad del escritor.)

Sin embargo, el uso de pronombres personales de primera y segunda persona en la tabla de características de visibilidad del escritor de BAWE sigue siendo su característica más obvia. Entre los 88 (por millón), el uso de pronombres personales de primera y segunda persona representa 65 (por millón), lo cual es un 73,9% del total. Entre ellos, "I" aparece un total de 14.204 (20 por millón) veces. Aunque es mucho más bajo que en BASE, sigue siendo el pronombre de primera persona es más utilizado en BAWE, lo cual indica nuevamente que los trabajos académicos de los alumnos nativos tienden a expresar sus propias opiniones, a reflejar su propia autoría. Además, en el uso de pronombres de segunda persona, "you" aparece con mayor frecuencia, con un total de 3.297 (5 por millón) veces, en comparación con BASE (224 por millón), lo cual indica que,

para los alumnos nativos, en los trabajos académicos no se destaca la identidad del lector. Entre los otros parámetros, *"think"* y *"believe"* son las palabras más utilizadas. En comparación con BASE, también se muestra que los trabajos académicos de los alumnos nativos están más inclinados a la expresión de los deseos subjetivos del autor. En general, BAWE tiene el grado más débil de coloquialismo, lo que sugiere que el inglés coloquial rara vez se usa en los trabajos académicos de los alumnos nativos de inglés.

4.1.3. Análisis de LM

LM es el corpus de pregrado de lengua materna recopilado para este estudio. Como colección de dos corpus, incluye la tesis de alumnos nativos recopilada por el autor y la tesis de pregrado en lengua nativa del corpus LOCNESS. En la Tabla 7, se puede ver que, aunque las características de visibilidad del escritor de LM sean más bajas que en BASE, se muestra que como corpus de artículos académicos, en comparación con el corpus del lenguaje coloquial, sus características coloquiales no son demasiado obvias. LM tiene 181 veces (por millón) de visibilidad del escritor más altas que los otros dos corpus de escritura académica.

En la tabla de características de visibilidad del escritor de LM, el uso de los pronombres de primera y segunda persona es de 129 (por millón) veces, lo cual representa el 71,3%, por lo que también es su característica más obvia. En el uso de los pronombres en primera persona, *"I"* apareció 2.181 (44 por millón) veces, y en el uso de pronombres en segunda persona, *"you"* apareció 898 (18 por millón) veces, ambos casos más bajos que BASE, pero superiores a los otros dos corpus académicos. Entre los parámetros restantes, *"think"* y *"just"* son las palabras más utilizadas. El uso específico en LM es el siguiente:

P. ej.

(1). *I think that the standard of writing, the way it has to be, I think it changed my actually talking, constantly working on a certain topic and the way how the teacher wants it formatted, I think it stuck in my head somehow.*

(“I”= *Presentar opiniones subjetivas a otros lectores, destacando la identidad del escritor.*)

(2). *See what you can accomplish if you only try, how far you can get if you see the light.*

(“You”= *La interacción entre el escritor y el lector, destacando la identidad del lector.*)

(3). *It has a population of well under 1.000 students, and just over sixty percent of students receive Free and Reduced Lunch, which is the most common measure of socio-economic status (SES) in American public schools.*

Debido a que en LOCNESS se recopilan los trabajos académicos de los alumnos de pregrado de lengua materna, se puede ver a partir de los datos que, cuando se usan los pronombres en primera persona, los alumnos de pregrado muestran características de lenguaje más informales y coloquiales en su escritura académica. Ello que implica que tanto si estos autores son alumnos de idiomas como si son alumnos nativos, la capacidad del idioma afecta a la calidad de la escritura. En comparación con los autores con un mayor dominio del idioma, los alumnos de pregrado prefieren utilizar más estrategias de visibilidad del escritor. Además, la presente investigación ha demostrado que una mayor educación y formación profesional puede ayudar a que la escritura académica sea más estándar y formal.

4.1.4. Análisis de TFM

Se puede observar en la Tabla 7, que como corpus académico de inglés, el TFM es también el único corpus de escritura académica de alumnos chinos. Sus características de visibilidad del escritor son menores que en BASE y LM, ya que son básicamente las mismas que en BAWE (88 por millón), lo que indica que TFM tiene el grado más bajo de lenguaje coloquial y un mayor grado de lenguaje escrito formal.

El uso de pronombres personales de primera y segunda persona en la tabla de características de visibilidad del escritor de TFM es también su característica más obvia. El uso de los pronombres personales de primera y segunda persona representa 62 (por millón), lo que representa el 70%, proporción más baja de los cuatro corpus. La principal diferencia con los tres corpus anteriores, es el uso de los pronombres de primera y segunda persona, ya que los tres primeros corpus tienen la mayor cantidad de apariciones de "I", mientras que en el TFM, la ocurrencia de "we" es la mayor, ya que es de 882 (27 por millón) veces. Además, se demuestra que, cuando se usa el pronombre personal "we", existen grandes diferencias entre los alumnos nativos y los alumnos chinos. Los alumnos nativos son más propensos a usar la forma del pronombre singular en primera persona para expresar sus opiniones, lo que refleja una clara conciencia subjetiva; mientras que los alumnos chinos prefieren usar la forma plural de los pronombres en primera persona para expresar opiniones grupales y usar grupos para expresar sus propias opiniones. El uso específico en TFM es el siguiente:

P. ej.

(1). These selected 39 potential shell nouns have a high frequency of occurrences in both of the two corpora, but we cannot draw a conclusion that the potential shell nouns are used in high frequency as shell nouns.

(2). By introducing the background information collected from news reports we can find the commonness of plagiarism phenomena in modern society, which challenges the reputation and the authority of higher education institutions.

("We"= Expresar las opiniones subjetivas del grupo a otros lectores, destacando la identidad del escritor.)

(3). Some studies have been made to examine the use of shell nouns in academic discourse.

(4). Therefore, the writer makes the table about the proportion of shell noun usage of 13 selected nouns in the two corpora in order to find out which nouns are used as shell nouns in the two corpora.

(5). Researchers have diverse opinions on the roles of language proficiency and experience in academic writing as well as their relationships with success in academic publishing.

Al mismo tiempo, en TFM, la visibilidad del escritor en algunas colocaciones fijas

es mayor que en los tres primeros corpus. Por ejemplo en el uso de : *"the use of"* (9 por millón), *"in order to"* (5 por millón), y *"as well as"* (4 por millón), lo que comporta que, en comparación con los alumnos nativos, los alumnos chinos se muestran más inclinados a utilizar un patrón fijo en la escritura escrita, lo cual está relacionado con los métodos educativos chinos. Para los alumnos chinos, la escritura académica se lleva a cabo estrictamente de acuerdo con las reglas de la escritura. Por lo tanto, desde la visibilidad del escritor, las características coloquiales de TFM son las mismas que en BAWE, que es la más baja entre los cuatro corpus.

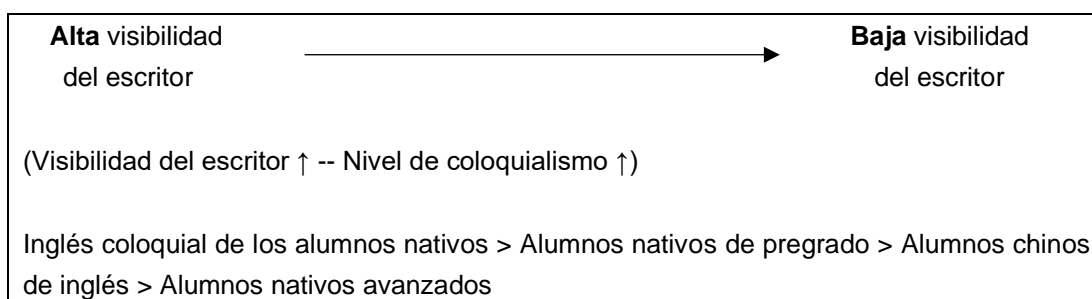


Tabla 8: Resultados comparativos de la visibilidad del escritor de los 4 corpus

Fuente: Extraído de los resultados de la Tabla 7

Los resultados del análisis muestran que en la visibilidad del escritor, BASE> LM> TFM> BAWE, lo que indica que el ranking del grado de coloquialismo en estos corpus es: corpus coloquial en lengua nativa > corpus de alumnos de pregrado en lengua nativa > corpus de alumnos chinos de inglés > corpus de alumnos avanzados en lengua nativa (véase tabla 8).

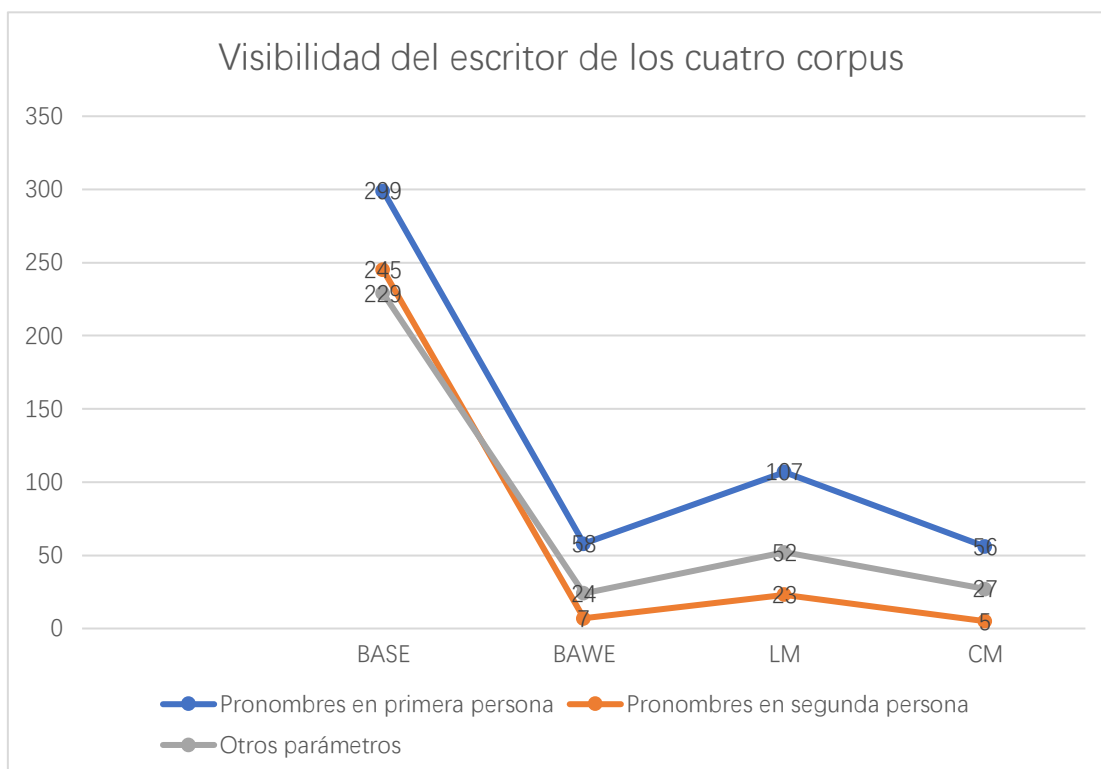


Figura 7: Comparación de la visibilidad del escritor en los 4 corpus (por millón)

Fuente: Microsoft Excel (producción propia)

Como característica principal de la visibilidad del escritor, el uso de los pronombres personales de primera y segunda persona es la característica más obvia presentada en los cuatro corpus (véase figura 7). En BASE, BAWE, LM y TFM aparecen respectivamente, 545/65/129/62 (por millón) veces en el texto, ya que son las características más comunes de visibilidad del escritor en cada corpus. Comparando el recuento promedio de esta característica, los alumnos nativos usan principalmente la forma singular de los pronombres en primera persona en su escritura académica, mientras que los alumnos chinos usan principalmente la forma plural en primera persona en su escritura académica. Esto demuestra que los alumnos nativos tienen una gran conciencia de sí mismos, mientras que los alumnos chinos tienden a expresar opiniones de grupo.

Además, a través de la comparación entre BAWE y LM, el autor entiende que es necesario comparar el corpus de los alumnos con diferente dominio del inglés en el proceso de investigación. Además, se puede reflejar directamente que, con

la mejora de la habilidad del inglés del autor, la visibilidad del escritor también mantiene una tendencia a la baja, lo que indica que la visibilidad del escritor está muy relacionada con el dominio del inglés del autor, mostrando una obvia correlación negativa. De forma similar, Wen, Ding y Wang (2003:269) encontraron que, con la mejora del dominio del inglés, las características coloquiales de la escritura en inglés de los alumnos avanzados de inglés son cada vez menores.

En resumen, a través del análisis comparativo de la visibilidad del escritor en los cuatro corpus y, basado en las características coloquiales de BASE, el estudio encontró que tanto los alumnos nativos como los alumnos chinos tienen características coloquiales relacionadas en la escrituras académica.

4.2. Resultados de los coloquialismos léxicos en el inglés académico

En esta sección se utilizarán los datos relevantes de BASE para comparar y analizar las características coloquiales de los corpus de BAWE, LM y TFM en cuanto a diversidad de vocabulario, densidad de vocabulario y frecuencia de vocabulario.

4.2.1. Diversidad léxica (TTR)

En esta sección, se ha usado el Type Token Ratio (TTR) para analizar la diversidad léxica. El TTR se calcula según el "*Tokens*" y el "*Types*" del corpus. Su fórmula de cálculo es: $TTR = (Type / Token) * 100$. Para hacer una comparación más precisa, a través del *Sketch Engine* y el *WordSmith 4.0*, se calculó el TTR de los cuatro corpus (véase tabla 9). Basándonos en el valor TTR de BASE, se compara y analiza el grado coloquial de los otros tres corpus escritos.

	BASE	BAWE	LM	TFM
<i>Tokens</i>	1.756.545	8.336.262	566.936	386.714
<i>Types</i>	185.932	1.396.718	72.568	55.678
TTR	10,6	16,8	12,8	14,4

Tabla 9: Comparación de TTR de los 4 corpus (por millón)

Fuente: Todos los datos provienen de *Sketch Engine* y *WordSmith 4.0*

La diversidad léxica de los cuatro corpus se puede determinar fácilmente a partir de la Tabla 9. El TTR en orden descendente es BAWE (16,8)> TFM (14,4)> LM (12,8)> BASE (10,6), por lo que la diversidad léxica de los cuatro corpus en orden descendente es BAWE> TFM> LM> BASE, lo que indica que BAWE, una base de datos de alumnos avanzados en lengua nativa, es la más profesional en cuanto a la escritura académica. Según Biber (1999:184), “cuanto menor es el valor del TTR, mayor es el grado de coloquialismo del texto”. Ya que se trata de un corpus de lenguaje coloquial, BASE tiene el TTR más bajo y el grado de coloquialismo más alto. En base al TTR de BASE, se ha encontrado que el TTR de los otros tres corpus es menor que el TTR de BASE, pero el grado de coloquialismo sigue siendo diferente, ya que el orden es TFM> LM> BAWE, lo que significa que, entre los tres corpus escritos, BAWE tiene el menor grado de coloquialismo.

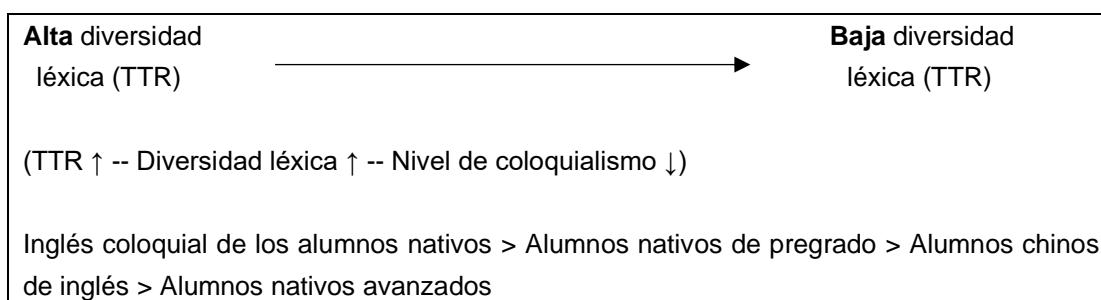


Tabla 10: Resultados comparativos de la diversidad léxica de los 4 corpus

Fuente: Extraído de los resultados de la Tabla 9

Al mismo tiempo, se observa en la tabla 10 que los resultados reflejados por la diversidad léxica de los cuatro corpus son los mismos que los resultados del análisis previo de la visibilidad del escritor. El análisis comparativo también se

basa en BASE, y los resultados muestran que los otros tres corpus escritos tienen características coloquiales con diferentes grados. Pero en contraste, el grado de profesionalismo es mucho mayor que el corpus coloquial de BASE. Por otro lado, en LM, al ser un corpus de trabajo académico de alumnos de pregrado, tiene una baja diversidad de vocabulario y tiene el grado más alto de lenguaje coloquial entre los tres corpus escritos. Las características coloquiales de la escritura académica de los alumnos de pregrado nativos son significativamente más altas que en los autores con un mayor dominio del inglés. Además, en TFM, como un corpus de escritura académica para los alumnos chinos avanzados de inglés, presenta una diversidad de vocabulario (TTR) más alta que en LM y más baja que en BAWE, refleja que los alumnos chinos de inglés son altos que los alumnos nativos de pregrado en la escritura académica, pero son bajos que los alumnos avanzados nativos.

4.2.2. Densidad léxica

Basándonos en las conclusiones de Ure (1971:452) y con la ayuda del *Sketch Engine* y el *AntConc*, hemos obtenido la densidad de vocabulario de los cuatro corpus (véase tabla 11). Los resultados son los siguientes:

	BASE	BAWE	LM	TFM
Sustantivo	332.942	1.902.560	115.249	106.916
Verbo	62.669	1.139.136	90.430	28.830
Adjetivo	102.472	639.843	41.548	36.635
Adverbio	23.360	338.328	28.078	13.760
Total	551.443	4.019.863	275.305	186.141
Todas las palabras	1.447.281	6.968.089	500.519	325.439
Densidad léxica	38,1%	57,7%	55%	57,1%

Tabla 11: Comparación de la densidad léxica de los 4 corpus (por millón)

Fuente: Todos los datos provienen de *Sketch Engine* y *AntConc*

Los datos de la tabla 11 muestran que la densidad léxica de los tres corpus de la escritura académica de BAWE, LM y TFM son de 57,7%, 55% y 57,1%, por

los cuales todos son superiores al valor crítico del 40% sugerido por Ure (1971), lo que indica que estos tres corpus no tienen características coloquiales. Para identificar la tendencia coloquial en los tres corpus, es necesario utilizar un corpus de referencia. Por lo tanto, se ha hecho un análisis comparativo basado en la densidad léxica de BASE (38,1%), un corpus de lenguaje coloquial, que muestra que los tres corpus tienen una tendencia coloquial, pero que el grado de coloquialismo es aceptable.

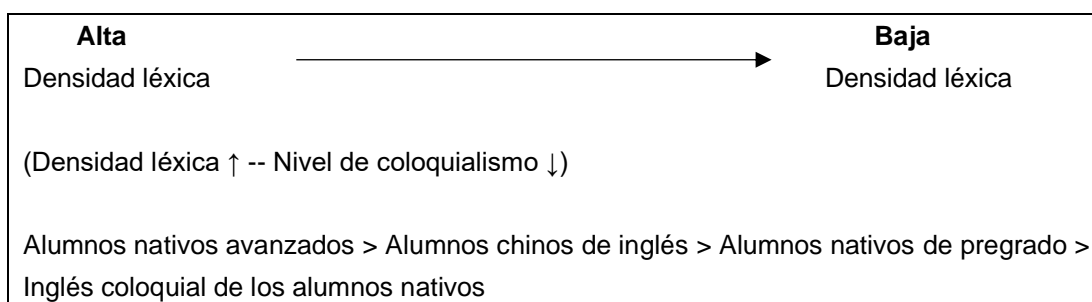


Tabla 12: Resultados comparativos de la densidad léxica de los 4 corpus

Fuente: Extraído de los resultados de la Tabla 11

En la tabla 12, se puede ver que BAWE tiene la mayor densidad léxica, seguido por TFM y LM. La densidad léxica de BASE del 40%, en concordancia con la conclusión de Ure (1971:452), "la densidad léxica del texto coloquial está por debajo del 40%". Este resultado es lo mismo que el anterior (visibilidad del escritor y diversidad léxica). Hasta cierto punto, esto puede explicar que los trabajos académicos de los alumnos nativos avanzados sean relativamente profesionales y que la escritura académica de los alumnos chinos de inglés sea más formal que la de los alumnos nativos de pregrado. En combinación con la investigación de Jiang (2018:29), se observa que el modelo de educación tradicional del inglés de China debería conceder una gran importancia al aprendizaje de vocabulario.

4.2.3. Frecuencia de vocabulario

En esta sección, se han utilizado las dos características principales del lenguaje

coloquial (el vocabulario coloquial y las abreviaturas de palabras) como indicadores para medir el grado de coloquialismo de la escritura académica. Con base en la frecuencia de uso de estas palabras en BASE, se compararán y analizarán las características coloquiales de estos indicadores de los tres corpus de la escritura académica.

4.2.3.1. Frecuencia de vocabulario coloquial

En base a la figura 5, se analizarán los 22 vocabularios coloquiales de la escritura académica de Paquot (2010:150). Además de estas 22 palabras coloquiales, esta investigación también se ha referido a Sun (2012: 10), añadiendo otras dos palabras "say" y "all in all", palabras utilizadas para citar e informar. Finalmente, ninguno de los 24 elementos del vocabulario coloquial eran palabras del AWL (*Academic Writing List*). Combinando estas palabras coloquiales, se ha utilizado el *Sketch Engine* para calcular la frecuencia de uso de estas palabras (por millón) en los cuatro corpus y para calcular la proporción de estas palabras coloquiales. Con base a los datos relevantes del corpus académico coloquial, en BASE, se han comparado los datos relevantes de los otros tres corpus escritos para estudiar sus características coloquiales. Los datos se muestran en la tabla 13:

Elemento léxico	BASE	BAWE	LM	TFM
<i>like</i>	2.753	450	1.937	584
<i>thanks to</i>	3	12	12	0
<i>so</i>	8.903	1.243	1.644	1.621
<i>because</i>	2.499	828	2.660	732
<i>that is why</i>	9	6	12	5
<i>this is why</i>	11	9	20	0
<i>look like</i>	112	14	28	5
<i>though</i>	252	246	850	114
<i>And</i>	23.734	25.035	21.856	28.530
<i>besides</i>	5	33	55	357
<i>I think</i>	1.674	47	320	202
<i>to my mind</i>	3	1	0	5
<i>from my point of view</i>	5	0	0	0
<i>it seems to me</i>	22	2	12	0
<i>really</i>	1.404	94	403	52
<i>of course</i>	692	46	115	26
<i>absolutely</i>	124	16	20	8
<i>maybe</i>	372	36	162	59
<i>I would like to talk about</i>	2	0	0	0
<i>I am going to talk about</i>	0	0	0	0
<i>thing</i>	2.472	204	798	197
<i>by the way</i>	42	2	0	21
<i>first of all</i>	137	18	40	72
<i>say</i>	3.572	404	1.375	393
<i>all in all</i>	1	4	4	23
Total	48.804	28.749	32.321	33.006
Porcentaje	56,9%	12%	56,57%	38,67%

Tabla 13: Comparación de la frecuencia del vocabulario coloquial de los 4 corpus (por millón)
Fuente: Todos los datos provienen de *Sketch Engine*

Se puede ver claramente a partir de los datos de la tabla 13 que las proporciones de estas palabras coloquiales en los cuatro corpus son muy diferentes. Entre ellos, en BASE, la frecuencia de uso de estas palabras coloquiales fue de 48.804 (por millón), lo cual representa un 56,9%. Este es un dato bastante alto, lo que indica que las características coloquiales en el corpus BASE son evidentes. Esto también está relacionado con su naturaleza, ya que, como un corpus académico de lengua coloquial, los textos (conferencias, seminarios) y los contenidos en BASE son más académicos que en los otros corpus de lengua coloquial general,

por lo que estas palabras coloquiales todavía se utilizan con frecuencia en sus textos. Entre ellas, las más utilizadas son "and" (23.734 por millón) y "so" (8.903 por millón), lo que demuestra que los alumnos nativos están acostumbrados a usar algunos conectivos y a expresar causalidad en su lenguaje académico coloquial. Por otro lado, también se indica que, en los hábitos coloquiales de los alumnos nativos, prefieren usar "so" y "because" (2.499 por millón). Al expresar opiniones subjetivas, los alumnos nativos prefieren utilizar la estructura de "I think" (1.674 por millón) en el lenguaje coloquial, en lugar de la estructura de "I would like/ am going to talk about". El uso específico en BASE es el siguiente:

P. ej.

(1). *The lecture is being recorded and I think I mentioned to you last week the purposes of this, so you know what's happening. (La expresión de opinión subjetiva del escritor)*

(2). *Is very useful in this problem because you can use it at every stage in what happens.*

En BAWE, la frecuencia de uso de estas palabras coloquiales es de 28.749 (por millón), lo que representa un 12% del total. Esta proporción es la más baja entre los cuatro corpus, lo que indica que todavía hay características coloquiales en BAWE, pero en comparación con BASE, sus características coloquiales no son tan evidentes. En comparación con LM y TFM, los artículos académicos de los alumnos nativos avanzados son más formales y estandarizados. De los datos, se puede observar que, entre estas características coloquiales, al igual que en BASE, "and" (25.035 por millón) es también la palabra más coloquial, con un uso incluso superior que en BASE (23.734 por millón), lo que muestra que, los alumnos nativos avanzados a menudo usan conjunciones como "and" en sus escritos académicos. Igualmente, los alumnos nativos avanzados no usan estructuras como "I would like to talk about" en su escritura académica. Pero a diferencia de BASE, la frecuencia de uso de vocabulario coloquial que expresa opiniones subjetivas, como "I think", en BAWE (47 por millón) es más baja que en BASE (1.674 por millón), lo que indica que, para los alumnos nativos

avanzados, la escritura académica del autor no expresa las opiniones subjetivas del autor, lo cual también es una de las características de la escritura académica. El uso específico en BAWE es el siguiente:

P. ej.

(1). Racism is still a problem within our society today, and many ethnic minorities face inequalities in many areas, including education, housing, and employment.

(2). There is also the idea that White Europeans wished to continue their superiority over colonized societies, so sustained the ideas associated with the inferior 'Other' in order to maintain it.

En LM, la frecuencia de uso de estas palabras coloquiales es de 32.321 (por millón), lo que representa un 56,57%. Esta proporción es más alta que en BAWE y TFM, lo que indica que los alumnos de pregrado nativos, todavía tienen características del coloquialismo en su escritura académica, lo cual es más obvio en comparación con los alumnos nativos avanzados y los alumnos chinos de inglés. Entre ellos, la palabra "*and*" es la más utilizada, con una frecuencia de uso de 21.856 (por millón). Lo interesante es que esta frecuencia de uso es la más baja entre los cuatro corpus, incluso más baja que en el corpus de lenguaje coloquial de BASE. Al mismo tiempo, la frecuencia de uso de "*so*" (1.644 por millón) y "*because*" (2.660 por millón) también es mayor en este vocabulario coloquial, lo que indica que las escrituras académicas de los alumnos de pregrado nativos expresan causalidad y usan el vocabulario coloquial con más frecuencia. La frecuencia de uso de la estructura "*I would like/ am going to talk about*" para expresar opiniones subjetivas es de 0, y la frecuencia de uso de "*I think*" (320 por millón) también es menor que BASE, pero es mayor que en BAWE y TFM. La diferencia es que la frecuencia de uso de "*though*" en LM es de 850 (por millón), la más alta entre los cuatro corpus, lo que indica que los alumnos pregrados nativos están acostumbrados a una "concesión" en la escritura académica. El uso específico en LM es el siguiente:

P. ej.

(1). *These actions are evil and because someone wanted money, or wanted more money than they had, they resorted to evil to attain it.*

(2). *Later in this interview, he noted: I think I write normally.*

(3). *Though the students in this study described a lot of conversation from their families and high school teachers about being prepared for college.*

En TFM, la frecuencia total de estas palabras coloquiales es de 33.006 (por millón), lo que representa un 38,67%. Esta proporción es menor que en BASE y LM, pero mayor que en el corpus de alumnos nativos de BAWE, lo que muestra que, en la escritura académica, en comparación con los alumnos nativos, los alumnos chinos están más acostumbrados a usar este vocabulario coloquial, lo que es inseparable de la forma y del método en que se enseña el inglés en China (Jiang, 2018:33). Como en los tres primeros corpus, la palabra "and" es la más utilizada, con una frecuencia de 28.530 (por millón) incluso mucho más alta que en BASE (23.734 por millón). En combinación con la investigación anterior, el análisis muestra que, para los alumnos chinos, sus trabajos académicos expresan una característica de "fórmula", y una frecuencia de uso del mismo tipo de palabras, frases, etc. A diferencia de los otros tres corpus, la frecuencia de uso de "thanks to" en TFM es de 0, lo que indica que los alumnos chinos de inglés no usan esta estructura en sus escrituras académicas. Al mismo tiempo, la frecuencia de uso de "besides" (357 por millón) es más alta que en los otros tres corpus, lo que indica que, en comparación con los alumnos nativos, los alumnos chinos avanzados usan más estructuras para complementar el contenido en sus escrituras académicas, como "besides" para complementar el contenido del artículo. El uso específico en TFM es el siguiente:

P. ej.

(1). *As people inevitably express their emotions, opinions and attitudes, and even their judgments in the language that they use.*

(2). *Besides , there are some studies proving that shell nouns are not only common in general English discourse, but also prevalent in academic English discourse.*

A través del análisis de la frecuencia del uso del vocabulario coloquial en los cuatro corpus, a diferencia de BASE, se ha observado que los otros tres corpus académicos tienen características de coloquialismo (usando palabras coloquiales en diversos grados). En la tabla 14, se puede ver claramente que la frecuencia del vocabulario coloquial en la escritura académica de los alumnos nativos es menor que la de los alumnos chinos de inglés.

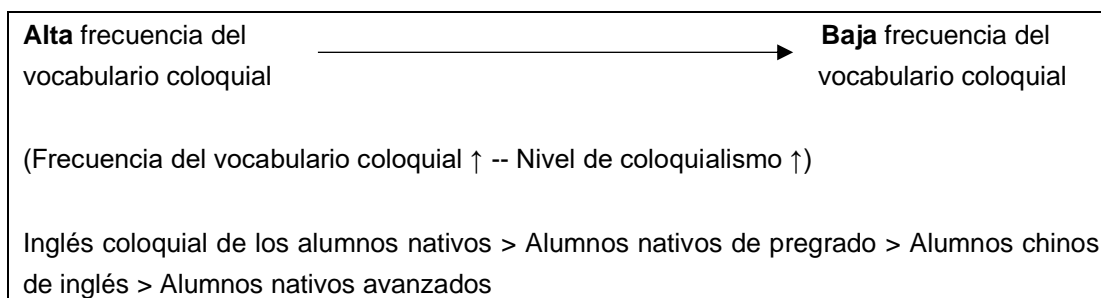


Tabla 14: Resultados comparativos de la frecuencia del vocabulario coloquial de los 4 corpus
Fuente: Extraído de los resultados de la Tabla 13

Al mismo tiempo, con la mejora de la calidad profesional de los alumnos nativos, la frecuencia del vocabulario coloquial en la escritura académica de los alumnos avanzados nativos es menor que en los alumnos de pregrado nativos. En el uso de vocabulario coloquial de alta frecuencia, no hay diferencias obvias entre los cuatro corpus, ya que todos usan la conjunción "and" con una frecuencia alta, lo que indica que cuando los autores de los escritos académicos necesitan usar una conjunción, a menudo se elige la palabra coloquial "and", lo que también es la característica coloquial más obvia entre los tres corpus. A la vez, en las opciones de uso de palabras coloquiales, los tres corpus tienen diferencias obvias, ya que los alumnos de pregrado de lengua materna (LM) utilizan "though" en la escritura académica para expresar sus concesiones y los alumnos chinos (TFM) utilizan "besides" en la escritura académica para completar el contenido del artículo. Por el contrario, en BAWE, como el nivel es más alto en la escritura académica en inglés, es evidente que los alumnos nativos avanzados (BAWE) utilizan este vocabulario coloquial con menos frecuencia en sus escritos académicos.

4.2.3.2. Frecuencia de abreviaturas y contracciones

Para analizar mejor las características del lenguaje coloquial de los cuatro corpus, se ha utilizado la frecuencia de "abreviatura y contracción" como indicador a analizar. La abreviatura y contracción utilizadas en este estudio adoptan principalmente la gramática Longman del inglés coloquial y los escritos de Biber et al (1999: 185). Los resultados obtenidos con la ayuda del *Sketch Engine* se muestra en la tabla 15:

Abreviatura y contracción	BASE	BAWE	LM	TFM
<i>-n't</i> (P. ej.: <i>isn't</i>)	5.416	259	1.292	26
<i>let's</i>	80	0	12	0
<i>I'm</i>	0	0	0	0
<i>-re</i> (P. ej.: <i>we're</i>)	3.835	11	71	0
<i>-s</i> (P. ej.: <i>what's</i>)	12.383	3.497	743	10
<i>-ve</i> (P. ej.: <i>you've</i>)	3.280	13	95	0
<i>-d</i> (P. ej.: <i>I'd</i>)	729	12	32	0
<i>-ll</i> (P. ej.: <i>you'll</i>)	1.791	10	12	0
Total	27.515	3.770	2.257	36
Porcentaje	57,8%	11,99%	37,93%	38,9%

Tabla 15: Comparación de la frecuencia de abreviatura y contracción de los 4 corpus (por millón)
Fuente: Todos los datos se provienen de *Sketch Engine*

A partir de los datos de la Tabla 15, es obvio que BAWE, LM y TFM tienen una brecha significativa con BASE en cuanto a la frecuencia de abreviatura y contracción. Entre ellos, BAWE tiene la menor frecuencia de uso.

En BASE, la frecuencia de uso de estas palabras de contracción es de 5.418 (por millón), lo que representa el 57,8%. Excepto por la forma de contracción "*I'm*", la frecuencia de uso es de 0. Como un corpus de lenguaje coloquial académico nativo, BASE usa la forma contraída "*-s*" con mayor frecuencia, ya que su frecuencia es de 12.383 (por millón). El uso de "*-s*" en BASE es mayor que del "*is*", ya que las contracciones se muestran principalmente como "*what's*", "*that's*" y "*it's*". Al mismo tiempo, la frecuencia de uso de la forma negativa "*-n't*"

(5.416 por millón) en BASE también es más frecuente en el uso de contracción en todo el corpus, especialmente en el caso de "isn't". El propósito de esta abreviatura es hacer una descripción negativa del contenido, lo cual es una característica coloquial muy importante. El uso específico en BASE es el siguiente:

P. ej.

(1). *Today's session is going to be the last of the lecture sessions for the course or if it isn't there'll only be about ten minutes next week.*

(2). *I think I mentioned to you last week the purposes of this, so you know what's happening.*

(3). *We're going to do is take a look at a problem which involves rotational momentum and that's the next problem.*

(4). *It's a concept that perhaps is a little bit difficult to perceive what we are going to be doing here is looking at a problem.*

En BAWE, la frecuencia de uso de estas contracciones es de 3.770 (por millón), lo que representa el 11,99%, el uso más bajo entre los cuatro corpus. También se utiliza la forma de contracción de alta frecuencia "-s" (3.497 por millón), que es la misma que las características de BASE; pero la mayor diferencia entre los dos corpus es que el uso de "-s" es completamente diferente. Aquí "-s" se usa como posesivo para expresar propiedad, como "today's" y "people's". Esta es una característica coloquial muy importante de BAWE. Se explica que, en la escritura académica de los alumnos nativos avanzados, al expresar propiedad no existe un concepto claro de lenguaje escrito y coloquial, y se acostumbra a utilizar lenguaje coloquial en lugar del lenguaje escrito. El uso específico en BAWE es el siguiente:

P. ej.

(1). *Overall colonialism has impacted on the racist ideas prevalent in today's society.*

(2). *However he takes it one step further by saying that these decisions are influenced by people's background experiences.*

En LM, la frecuencia de uso de estas contracciones es de 2.257 (por millón), lo

que representa el 37,93%. Esto es más bajo que en BASE y en TFM, pero más alto que en BAWE, un corpus de escritura académica para alumnos nativos avanzados. Esto es interesante porque en el análisis anterior, LM se utiliza como base de datos en la lengua materna de los alumnos de pregrado, y el nivel de coloquialismo en sus diversos análisis es superior al de TFM, un corpus de alumnos chinos de inglés. Creo que la razón de este fenómeno puede ser la vaguedad de conciencia de estilo de estos alumnos chinos de inglés, ya que se forma una conciencia de lenguaje solidificada, que no es fácil de cambiar, por lo que los alumnos de pregrado nativos prestan atención y evitan el uso de estas abreviaturas. Al mismo tiempo, a diferencia de los otros tres corpus, se puede ver que las abreviaturas negativas como "-n't" en LM se utilizan con mayor frecuencia (1.292 por millón). Creo que esto se debe a los hábitos lingüísticos y al descuido de los alumnos de pregrado. El uso específico en LM es el siguiente:

P. ej.

(1). *I personally don't agree with that.*

(2). *By arguing that affirmative action hasn't worked, opponents of the argument are really defending the need for such programs.*

(3). *They ignore the fact that blacks can't take part in the system until they are allowed in.*

En TFM, la frecuencia de uso de estas abreviaturas y contracciones es del 38,9%, la cual es más alta que en los tres corpus escritos. Entre ellos, los alumnos chinos del inglés solo usan "-n't" (26 por millón) y "-s" (10 por millón), que es lo mismo que LM, lo que significa que en comparación con los alumnos nativos, los alumnos chinos suelen usar estas abreviaturas y tienen más coloquialismo en su escritura académica. El uso específico en TFM es el siguiente:

P. ej.

(1). *When I was writing an English essay, I knew I needed to use certain words, but I couldn't remember them for a while.*

(2). *And eventually one's success in completing the task, and planning how to complete the task.*

A través de analizar la frecuencia de uso de las abreviaturas y las contracciones en los cuatro corpus, con base en BASE, se ha observado que los textos académicos de los otros tres corpus tienen las características de coloquialismo (el uso de abreviaturas). En combinación con la tabla 16, se puede ver que, en el uso de abreviaturas y contracciones, los alumnos nativos avanzados (BAWE) prestan más atención a evitar esta característica coloquial en su escritura académica.

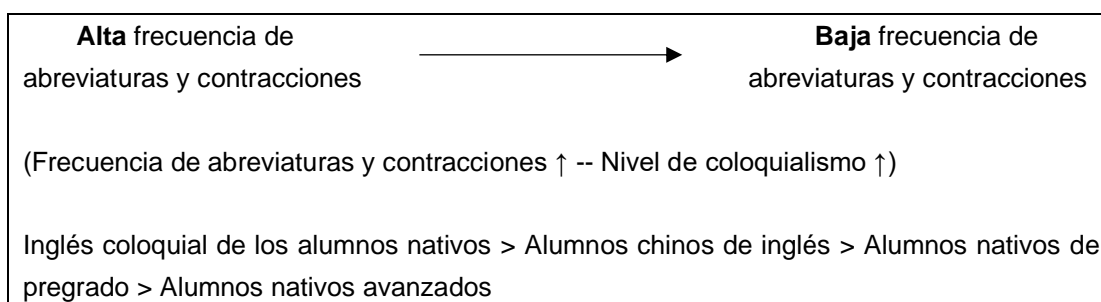


Tabla 16: Resultados comparativos de frecuencia de abreviaturas y contracción de los 4 corpus
Fuente: Extraído de los resultados de la Tabla 15

Mientras, los alumnos nativos de pregrado (LM) son más cuidadosos que los alumnos chinos de inglés (TFM) en el uso de las abreviaturas y las contracciones, lo que demuestra que la tendencia coloquial en el lenguaje escrito no es inversamente proporcional al nivel de inglés de los alumnos.

4.3. Resultados de los coloquialismos sintácticos en el inglés académico

A nivel sintáctico, existe una gran diferencia entre el inglés coloquial y el escrito. En general, la escritura académica profesional debe basarse en las reglas del lenguaje escrito y tratar de evitar el uso de expresiones coloquiales. En esta sección, se analizan las características coloquiales de los cuatro corpus de académica desde dos aspectos: longitud promedio de la frase y estructura de la frase.

4.3.1. Longitud promedio de la frase

Tomarse ha tomado como estándar la longitud promedio de la frase y se ha realizado un análisis comparativo de los tres corpus académicos de escritura. Se ha utilizado el *Sketch Engine* y el *WordSmith 4.0* en el proceso de conteo. De acuerdo con el número de palabras y frases en los tres corpus, se puede calcular la longitud promedio de la frase. Los resultados se muestran en la tabla 17:

	BAWE	LM	TFM
Palabra	8.336.262	566.936	325.439
Frase	293.113	21.977	13.998
Longitud promedio de la frase	28,5	25,8	23,24

Tabla 17: Comparación de la longitud promedio de la frase de los 3 corpus
Fuente: Todos los datos se provienen de *Sketch Engine* y *WordSmith 4.0*

Cabe señalar que, en el análisis de la longitud promedio de las frases, BASE se ha utilizado como un corpus de lenguaje coloquial. El número de frases completas es de sólo 193. El número calculado por estos datos no es exacto. Al mismo tiempo, en este análisis, el autor ha tenido que centrarse en tres corpus de escritura académica, por lo que no se han adoptado los datos relevantes de BASE.

Según se puede ver en la Tabla 17, en la longitud promedio de frase, la brecha entre BAWE, LM y TFM es clara. Entre ellos, el número total de palabras en el corpus de inglés escrito de alumnos nativos avanzados (BAWE) es de 8.336.262, mientras que el número total de palabras de alumnos de pregrado en lengua materna (LM) es de 566.936 y el número total de palabras de los alumnos chinos de inglés (TFM) es de 325.439. Por lo tanto, el número de palabras en el corpus de los alumnos chinos de inglés es menor que el número de palabras en el corpus de los alumnos de pregrado nativos es menor que el número de palabras en el corpus de alumnos avanzados nativos. El número de frases en el corpus de los alumnos nativos avanzados (BAWE) es de 293.113, y la longitud promedio

de las frases es de aproximadamente 28,5. El número de frases en el corpus de los alumnos nativos de pregrado (LM) es de 21.977, y la longitud promedio de las frases es de aproximadamente 25,8. Además, el número de frases de los alumnos chinos de inglés (TFM) es 13.998, y la longitud promedio de la frase es de aproximadamente 23,24, por lo que se puede decir que la longitud promedio de la frase del corpus de los alumnos nativos avanzados es mucho más larga que la longitud promedio de los alumnos chinos. Al mismo tiempo, la longitud promedio de las frases del corpus de los alumnos nativos avanzados es más larga que la longitud promedio de la frase del corpus de los alumnos nativos de pregrado.

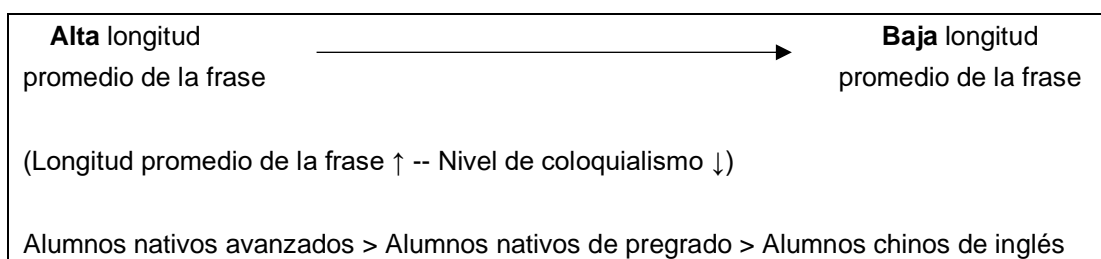


Tabla 18: Resultados comparativos de la longitud promedio de los 3 corpus

Fuente: Extraído de los resultados de la Tabla 17

En la tabla 18, se muestra que, en comparación con los alumnos chinos de inglés, a los alumnos nativos les gusta expresar sus opiniones en frases complejas, mientras que los alumnos chinos suelen expresar sus opiniones en frases cortas y sencillas. Además, se necesita un lenguaje más complejo en los artículos académicos, por lo que a menudo aparecen frases largas y complejas. Por el contrario, las personas suelen utilizar frases cortas en el lenguaje coloquial para facilitar la comunicación. Por lo tanto, este resultado puede indicar que, en comparación con los alumnos nativos del idioma inglés, los alumnos chinos tienen una orientación coloquial más alta. Por lo tanto, en comparación con los hablantes nativos avanzados, los hablantes nativos de pregrado tienen una tendencia coloquial más alta.

4.3.2. Estructura de la frase

Esta investigación parte de las estructuras "There be" y del verbo perifrástico, comparando y analizando las características coloquiales de los cuatro corpus.

4.3.2.1. Estructura "There be"

En este punto se analiza el uso de la estructura "There be" en los tres corpus escritos para estudiar sus propias características coloquiales. Cabe señalar que BASE como corpus de lengua coloquial no participa en este análisis. Los resultados se muestran en la siguiente tabla:

	BAWE	LM	TFM
"There be" (frecuencia)	13.167	950	667
Frase	293.113	21.977	13.998
Porcentaje	4,49%	4,32%	4,76%

Tabla 19: Comparación del uso de la estructura "There be" en los 3 corpus

Fuente: Todos los datos se provienen de *Sketch Engine*

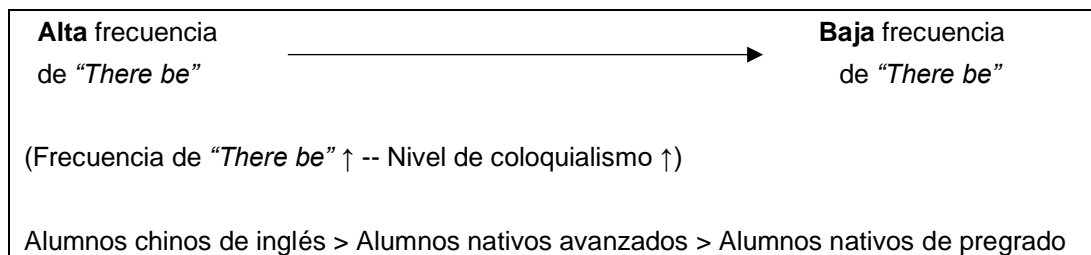


Tabla 20: Resultados comparativos de la frecuencia de la estructura "There be" en los 3 corpus

Fuente: Extraído de los resultados de la Tabla 19

En las tablas 19 y 20 se puede ver que en el uso de la estructura "There be", no existe una brecha obvia entre los tres corpus de escritura académica. En BAWE, la estructura "There be" aparece 13.167 veces, lo que representa el 4,49%; la estructura "There be" en LM aparece 950 veces, lo que representa el 4,32%; y la estructura "There be" en TFM apareció 667 veces, lo que representa el 4,76%. Esto muestra que en el uso de la estructura "There be", los artículos académicos

de los alumnos chinos avanzados de inglés (TFM) se muestran más coloquiales que los alumnos nativos. No obstante lo que se debe enfatizar es que para la mayoría de los alumnos chino de inglés, la estructura "There be" es una estructura básica con la que entran en contacto cuando aprenden la expresión en inglés por primera vez. Debido a la transferencia de idiomas negativa, los alumnos chinos están acostumbrados a utilizar esta estructura fija para expresar el significado de existencia y posesión, por lo que este dato tiene limitaciones. Además, no se puede juzgar completamente que la tendencia coloquial de los alumnos chinos de inglés avanzado en la escritura académica sea mayor que en los alumnos nativos. Al mismo tiempo, para los alumnos nativos, la naturaleza del lenguaje escrito de las escrituras académicas de alumnos nativos avanzados (BAWE) es más sólida que los alumnos nativos de pregrado (LM). El uso específico es el siguiente:

P. ej. (BASE corpus)

- (1). *There is a chart at the back of your notes showing the comparison between translational motion and rotational motion.*
- (2). *There are too many of them, so I hope that you agree that I have got the correct number of forces.*
- (3). *There's a useful formula given here which you can work out for yourself.*

P. ej. (BAWE corpus)

- (1). *There is a connection between one of the many stories unearthed by Sinclair.*
- (2). *There are three reasons for this.*
- (3). *There's also the influence of economic factors to consider.*

P. ej. (LM corpus)

- (1). *In cities, where all traffic jams start, there is no room for more roads.*
- (2). *There are three main ways of doing this: sharing cars, using public transport and walking or cycling.*

P. ej. (TFM corpus)

- (1). *Among the various lexical features, there is a special type of nouns which has attracted considerable attention in the last two decades, which is named as shell nouns in this study.*
- (2). *However, there are some researchers mentioning the evaluation function or interpersonal function of shell nouns.*

Al buscar ejemplos, el autor descubrió que en con comparar LM y TFM, en BAWE se usa la forma abreviada "there's" con más frecuencia, mientras que la estructura "there's" rara vez se usa en LM y TFM. Combinando los hallazgos previos en las abreviaturas y contracciones (véase tabla 15), se corroboró una vez más la conclusión del autor. Los alumnos nativos avanzados (BAWE) no tienen un fuerte sentido de la escritura en el uso de las abreviaturas y las contracciones, por lo que muestran una tendencia coloquial.

4.3.2.2. Verbos perifrásticos

Esta parte de la investigación se ha basado en la investigación relevante de Jiang (2018: 53) y se refiere a los 72 verbos perifrásticos de alta frecuencia en el LOCNESS de Waibel (2007:87) (véase tabla 21).

<i>bring about</i>	<i>bring down</i>	<i>bring in</i>	<i>bring on</i>	<i>bring out</i>	<i>bring up</i>
<i>carry out</i>	<i>come about</i>	<i>come across</i>	<i>come in</i>	<i>come back</i>	<i>come down</i>
<i>come around</i>	<i>come off</i>	<i>come on</i>	<i>come out</i>	<i>come up</i>	<i>find out</i>
<i>get along</i>	<i>get away</i>	<i>get back at</i>	<i>get down</i>	<i>get off</i>	<i>get on</i>
<i>get out</i>	<i>get through</i>	<i>get up</i>	<i>give up</i>	<i>go about</i>	<i>go ahead</i>
<i>go along</i>	<i>go down</i>	<i>go in</i>	<i>go off</i>	<i>go on</i>	<i>go out</i>
<i>go over</i>	<i>go up</i>	<i>look up</i>	<i>look up to</i>	<i>make up</i>	<i>pick up</i>
<i>point out</i>	<i>put away</i>	<i>put down</i>	<i>put forward</i>	<i>put in</i>	<i>put off</i>
<i>put on</i>	<i>put up</i>	<i>run out</i>	<i>set down</i>	<i>set in</i>	<i>set off</i>
<i>set up</i>	<i>stand up</i>	<i>take apart</i>	<i>take back</i>	<i>take down</i>	<i>take in</i>
<i>take off</i>	<i>take on</i>	<i>take out</i>	<i>take over</i>	<i>take up</i>	<i>turn down</i>
<i>turn in</i>	<i>turn on</i>	<i>turn out</i>	<i>turn up</i>	<i>look at</i>	<i>come up with</i>

Tabla 21: Waibel's 72 High-frequency Phrasal-Verb List en LOCNESS

Fuente: *A Corpus-based Study on Colloquialisms in Academic Writing by MA English Postgraduates*, Jiang, Y.S, (2018:70)

蒋艳双, p70, 基于语料库的英语硕士研究生学术写作口语化现象研究

Para la conveniencia de la investigación, se han seleccionado 10 de ellos, basándose en los datos de BASE. Gracias a ellos, se han analizado las características coloquiales de los cuatro corpus en el uso del verbo perifrástico. Los resultados se muestran en la tabla 22:

Verbo perifrástico	BASE	BAWE	LM	TFM
<i>go on</i>	410	47	194	18
<i>carry out</i>	41	187	47	184
<i>point out</i>	53	96	111	197
<i>take on</i>	39	28	47	10
<i>bring up</i>	15	9	83	5
<i>give up</i>	47	12	47	10
<i>bring about</i>	9	35	59	26
<i>find out</i>	86	34	67	261
<i>make up</i>	39	51	28	62
<i>set up</i>	97	70	43	39

Tabla 22: Comparación del uso de los 10 verbos perifrásticos en los 4 corpus (por millón)

Fuente: Todos los datos se provienen de *Sketch Engine*

A partir de los datos de la tabla 22, se puede ver que los cuatro corpus usan verbos perifrásticos en diversos grados y que hay una gran diferencia en los hábitos de su uso.

En BASE, se puede ver en la tabla 22 que el uso de estos 10 verbos perifrásticos es relativamente normal. A excepción del caso de *"go on"*, ya que su frecuencia total de uso " " es de 410 (por millón). Esta frecuencia es la más alta entre los cuatro corpus. *"go on"* significa continuar, y se usa principalmente en el lenguaje coloquial. A veces, incluso puede tener diferentes significados para diferentes ocasiones, por lo que es de uso flexible y es un término muy coloquial. Los materiales de BASE provienen principalmente de conferencias y seminarios académicos. Esto puede indicar que los alumnos nativos prefieren usar *"go on"* en lugar de *"continue"* en el inglés académico coloquial. Esta característica se debe a los hábitos lingüísticos de los alumnos nativos. El uso específico es el siguiente:

P. ej.

(1). *I think we'll carry on as normal and ignore what's going on over here.*

(2). *What actually goes on in these situations can be obtained.*

En BAWE, el uso de estos 10 verbos perifrásticos es relativamente desigual,

excepto por la frecuencia total de "carry out" que es de 187 (por millón) veces, y el resto de los verbos perifrásticos pueden ignorarse. El verbo perifrástico más utilizada en BAWE, "carry out", tiene casi la misma frecuencia que TFM, y es la más alta entre los cuatro corpus. "carry out" significa completar y ejecutar. Es más coloquial que "complete". El punto interesante es que la frecuencia de "complete" es de 187 veces (por millón), lo misma que "carry out". Esto muestra que para los alumnos nativos avanzados, al elegir utilizar "complete" o "carry out", la naturaleza escrita de los dos no se puede distinguir claramente. La elección de cuál utilizar depende enteramente de los hábitos lingüísticos del autor, ya que la comprensión coloquial de "carry out" es vaga. El uso específico es el siguiente:

P. ej.

(1). *He claimed he 'would never be a cause of trouble and dissent in the State'. It seems, from this, to have been rather a passionate, naturalistic sense of justice that drove him to carry out the corpus than a positivistic, legal course of action.*

(2). *Up to 24 hours were given to complete the sheet, if needed.*

En LM, la frecuencia de uso de estos 10 verbos perifrásticos es bastante uniforme. Esta situación es similar a la de BASE. También se muestra que en el uso del verbo perifrástico, el coloquialismo de los alumnos nativos de pregrado en las escrituras académicas es más cercano al del corpus nativo coloquial. Entre ellos, el verbo perifrástico más utilizado es también "go on", y su total de veces es de 194 (por millón). En comparación con "go on", "continue" es más formal en LM con una frecuencia total de 419 veces (por millón), que es mayor que "go on", lo que indica que los alumnos nativos de pregrado usan estos verbos perifrásticos coloquiales para expresar sus opiniones en trabajos académicos. Sin embargo, en comparación con sus hábitos coloquiales, los alumnos de pregrado nativos tienen la conciencia de distinguir entre el verbo perifrástico y las palabras escritas en la escritura académica. El uso específico es el siguiente:

P. ej.

(1). *Sartre was particularly against letting oneself be ruled by the past: what's done is done, life must go on.*

(2). *Until the rail service improves people will continue to use the ever increasingly clogged up roads.*

En TFM, el uso de estos 10 verbos perifrásticos es muy desigual. Los alumnos chinos de alto nivel usan "*find out*", "*point out*" y "*carry out*" de manera muy intensiva, con una frecuencia total de 261, 197 y 184 veces (por millón). El resto de verbos perifrásticos aparecen menos en la escritura académica de los alumnos chinos de inglés. Esto puede deberse a que los verbos perifrásticos son difíciles de comprender. En el proceso de aprendizaje, debido a la falta de una estructura china similar a los verbos perifrásticos, éstos pueden evitarse intencionalmente. Además, Spears (2007: 211) explica que "averiguar en un diccionario suele entenderse como la expresión coloquial de aprender algo o algo de alguien". En chino, "*find out*", "*point out*" y "*carry out*" a menudo aparecen en la escritura, especialmente en los trabajos académicos. Esta expresión da a los lectores una sensación muy profesional; pero en el inglés escrito, es una expresión muy coloquial, por lo que debe evitarse tanto como sea posible. Esto también corrobora el descubrimiento de Jiang (2018:56), por lo que, los alumnos chinos de maestría en inglés transformarán los hábitos de escritura en chino a los hábitos de escritura académica en inglés. El uso específico es el siguiente:

P. ej.

(1). *To analyze clearly, the writer will make another table to find out the higher proportions of shell noun usage.*

(2). *In addition, Li (2001) studies the use of epistemic modality in academic discourse and points out its interpersonal meaning.*

(3). *The whole process which lasted for 15 weeks was only carried out in the Experimental Class.*

Combinando el análisis del autor sobre el uso del verbo perifrástico en las cuatro bases de datos, se puede ver que, tanto los alumnos nativos como los alumnos chinos de inglés, independientemente de la frecuencia, están acostumbrados a

usar este verbo perifrástico coloquial en sus escritos académicos. Para los alumnos nativos, esto puede deberse a sus hábitos lingüísticos y su conocimiento vago de los verbos perifrásticos; mientras que para los alumnos chinos de inglés, llevar los hábitos de escritura china a la escritura académica en inglés es la razón principal del coloquialismo.

4.4. Posibles coloquialismos en la escritura académica de los alumnos chinos

A través del análisis anterior, se ha deducido que existen diferentes grados de lenguaje coloquial en la escritura académica de los alumnos chinos de inglés. Esta tendencia coloquial es más obvia en la escritura académica de los alumnos nativos avanzados, pero como alumnos avanzados de inglés, esta tendencia coloquial es más débil que la escritura académica de alumnos los de pregrado nativos. Para mejorar la capacidad de escritura de los alumnos chinos en la escritura académica en inglés, es necesario explorar y discutir las posibles razones del inglés coloquial. Combinando las características coloquiales previamente analizadas de la escritura académica de los alumnos chinos, se ha visto que esto se ve afectado en gran medida por la transferencia negativa de la lengua materna. Al mismo tiempo, también se ve muy afectado por los métodos tradicionales de educación en inglés de China.

4.4.1. Transferencia negativa de la lengua materna

En el proceso del aprendizaje de una L2, no se puede ignorar el fenómeno de la transferencia de idiomas. Se trata de un fenómeno muy complejo. Odlin (1989: 115) define la transferencia de idiomas como "el efecto de las similitudes y diferencias entre el idioma de destino y cualquier otro idioma previamente adquirido". Ellis (1985: 120) cree que la hipótesis es que cuando hay una diferencia entre la L1 y L2, el conocimiento de la L1 del alumno interferirá con su

aprendizaje de la L2, y las similitudes entre la L1 y L2 pueden ayudar a los alumnos en el aprendizaje de la L2, proceso llamado transferencia de idioma. Este tipo de transferencia se puede dividir en dos tipos: transferencia positiva y transferencia negativa. Se considera transferencia positiva los hábitos del lenguaje, las reglas del lenguaje y los patrones del lenguaje de la lengua materna, que promueven positivamente el aprendizaje de L2; mientras que la transferencia negativa se considera el hábito lingüístico, las reglas y los patrones de lenguaje de la lengua materna, por los que se puede interferir y dificultar el aprendizaje de L2 por parte de los alumnos.

Para los alumnos chinos, su aprendizaje del inglés se ve muy afectado por la transferencia negativa de su lengua materna. Debido a las diferencias obvias entre el inglés y el chino, si uno está acostumbrado a escribir inglés en la forma, estructura gramatical y reglas del chino, surgirán los problemas de escritura correspondientes. En combinación con la encuesta de Song (2017:79), se encuentra que la mayoría de los alumnos chinos creen que es imposible completar su escritura escrita en inglés sin la forma de pensar china. Esto refleja la mentalidad de los alumnos chinos en la escritura en inglés. La dependencia excesiva de la lengua materna se considera la escritura como un proceso de traducción más que como un proceso creativo. autora continuación se enumeran las características del inglés coloquial en las escrituras académicas de alumnos chinos afectados por la transferencia cultural negativa (véase tabla 23):

Coloquialismos	Ejemplos
Visibilidad del escritor	<i>"I", "we", "you"</i>
Vocabulario coloquial	<i>"mean", "think"</i>
Longitud de la frase	(corto y sencillo)
Verbo perifrástico	<i>"find out", "point out", "carry out"</i>

Tabla 23: Características coloquiales de las escrituras académicas de los alumnos chinos

Fuente: Extraído de los resultados de la Tabla 7, la Tabla 13, la Tabla 17 y la Tabla 22

En términos de visibilidad del escritor, Wen y Wang (2003:269) creyeron que el

modelo de pensamiento L1 tiene un impacto negativo en la escritura del L2 de los alumnos, que se manifiesta en el uso excesivo de pronombres en primera persona y de pronombres en segunda persona. Esta investigación también se ha identificado este problema. Se puede ver en la tabla 7 que hay un uso excesivo de los pronombres de primera y segunda persona en la escritura académica de los alumnos chinos. Es común en chino expresar opiniones en primera persona, y "I" es la perspectiva del autor en su realidad. "we" y "you" se refieren a la perspectiva del lector en la realidad, y no es la perspectiva del interlocutor. Esto es diferente a los hábitos del idioma inglés, lo que provoca que las escrituras académicas de los alumnos chinos muestren una fuerte tendencia coloquial. El uso específico es el siguiente:

P. ej.

(1). I decided to use similar methods to evaluate the intertextual expressions appeared in this thesis. ("I"= el propio autor)

(2). But we cannot draw a conclusion that the potential shell nouns are used in high frequency as shell nouns. ("we"= grupo de lectores)

(3). The third sentence means the person that you spoke to has the permission to leave. ("you"= lector específico)

En términos del vocabulario coloquial, en combinación con los datos de la tabla 13, para los vocabularios coloquiales de alta frecuencia utilizados por los alumnos chinos en sus escrituras académicas, se puede encontrar que su elección se debe en gran medida a la transferencia negativa de su lengua materna, como "mean", "think" y "want". En chino, estas palabras son descripciones más escritas cuando se expresan opiniones en la escritura académica. Muchos alumnos utilizan estas palabras en sus escrituras académicas; pero en el inglés no lo son, ya que estas palabras se clasifican como vocabulario coloquial y su uso excesivo afectará a la redacción de la escritura académica. El uso específico es el siguiente:

P. ej.

(1). This chapter means to draw a sketch of the whole thesis. (12 por millón)

(2). He thinks that on the one hand citation is more than the display of authors. (8,6 por millón)

(3). The author wants to clarify that the shell nouns are dealt with their shell contents in this thesis. (2,5 por millón)

En términos de la longitud de las frases, en combinación con los datos de la tabla 18, para la escritura académica en inglés, los autores suelen utilizar frases largas y complejas para expresar sus opiniones, que pueden reflejar la naturaleza escrita de la escritura académica. Las frases cortas a menudo aparecen en los textos coloquiales de los nativos para facilitar la comunicación. Sin embargo, en chino ocurre todo lo contrario. En la escritura china, las expresiones cortas pueden hacer que las opiniones sean más intuitivas y concisas, lo que refleja un mayor profesionalismo y lo que a menudo lleva a los alumnos chinos a elegir frases cortas para expresar sus opiniones. Esto también provoca que las escrituras académicas de los alumnos chinos muestren un cierto grado de tendencia coloquial. El uso específico es el siguiente:

P. ej.

Firstly, you can listen to music by phone, because music can relax yourself and make you happy.

(El vocabulario, la gramática y la expresión son simples)

En términos del verbo perifrástico, en combinación con los datos de la tabla 22, se puede encontrar que a los alumnos chinos les gusta usar "*find out*", "*point out*" y "*carry out*" en la escritura académica. En inglés, "*找出*" significa dos palabras, "*find*" y "*out*". El significado de "*find*" en inglés es "*discover*". Por el contrario, "*discover*" expresa el significado de "*找出*" en chino. Sin embargo, los alumnos chinos se ven afectados por la transferencia de idiomas negativa y creen que ambos no tienen el mismo significado, por lo que eligen usar el Verbo perifrástico "*find out*", ya que creen que es más profesional. En comparación con la

frecuencia de "*find out*" (261 por millón), "*discover*" (111 por millón) se utiliza menos en las escrituras académicas de los alumnos chinos, lo que muestra que, la mayoría de los alumnos chinos no eligen utilizar "*discover*" en la escritura académica. Pero, de hecho, el uso excesivo de "*find out*" es solo una expresión coloquial. Su uso específico es el siguiente:

P. ej.

(1). *The aim of this thesis is to discover the transformation of intertextuality from citation to plagiarism in the data.*

(2). *He aims to find out a new way.*

(Se puede ver claramente que la característica coloquial de "find out")

Debido a las características anteriores, el proceso de escritura académica de los alumnos chinos en inglés se ha convertido en un proceso de escritura en chino y luego de traducción al inglés. Por lo tanto, muchos fenómenos coloquiales aparecen en la escritura académica.

4.4.2. Métodos tradicionales de enseñanza del inglés en China

Para los alumnos chinos, aunque comiencen a aprender inglés cuando son pequeños, en la enseñanza regular en el aula, el aprendizaje se centra en aprender diferentes palabras y su uso, estructuras de frases complejas y algunas reglas gramaticales específicas. Por el contrario, pocos profesores enseñan a los alumnos conocimientos relacionados con la lingüística del inglés y las diferencias del uso entre el inglés y el chino. En opinión de los profesores de inglés chino, creen que los alumnos no necesitan dominar las diferencias entre el inglés coloquial y escrito. En el proceso de aprendizaje del inglés, el estilo del inglés no es una parte del contenido de la enseñanza. Es precisamente debido a este método tradicional de educación que es difícil para los alumnos chinos tener una comprensión clara de la escritura formal del inglés. Muchos alumnos chinos ignoran el inglés coloquial, lo que indirectamente conduce a la tendencia

al coloquialismo en la escritura académica.

Al mismo tiempo, los profesores de inglés chino prestan gran atención a la llamada "estructura fija", la colocación fija de palabras y frases y a la expresión fija de frases en el proceso de enseñanza del inglés. Estos se convierten en una especie de "fórmula estándar", que hace que los alumnos tengan un modo de pensamiento fijo, con una dependencia y un uso excesivos. Por lo tanto, para los alumnos, el proceso de escritura es similar al proceso de "llenar los espacios en blanco". La "aplicación de fórmulas" es un fenómeno común en la escritura académica, que hace que muchas expresiones irregulares se utilicen en exceso y que se forme gradualmente un hábito de lenguaje. Al igual que los hallazgos previos del autor en la investigación, el uso excesivo de la estructura "*There be*" por parte de los alumnos chinos se deben a este modo de pensar. Hay otro ejemplo: los alumnos nativos rara vez usan la estructura "*as far as I am concerned*", pero a menudo aparece en los escritos académicos de los alumnos chinos. La frecuencia total de esta estructura en TFM es 75 veces (por millón). Por el contrario, en el corpus de los alumnos nativos casi no se usa esta estructura, y en muchos casos se utiliza por separado. El uso específico en TFM es el siguiente:

P. ej.

(1). *As far as academic writing is concerned, focus can only be placed on the linguistic type lexical device and grammatical device.*

(2). *As far as I am concerned, if students want to get a better achievement, they'd better leave their cell phones at home.*

5. CONCLUSIONES

En el presente estudio se ha comparado el corpus del inglés escrito, el corpus del lenguaje coloquial y el corpus de alumnos chinos en inglés, y se han comparado estos corpus a través de la visibilidad del escritor, el nivel léxico y el nivel sintáctico, y finalmente explorando el uso del inglés coloquial en la escritura académica. En el capítulo 4, se analizarán y discutirán 6 aspectos de la visibilidad del escritor, la diversidad del vocabulario, la densidad del vocabulario, la frecuencia del vocabulario, la longitud promedio de la frase y la estructura de la frase. Este capítulo es la conclusión de toda la investigación, dividida en 3 partes: los hallazgos principales, las limitaciones e investigaciones futuras.

5.1. Hallazgos principales

El propósito de esta investigación es analizar el uso coloquial de la escritura académica en inglés, mediante el análisis comparativo de las características de los cuatro corpus, por las cuales se ha respondido a las cuatro preguntas de investigación planteadas en el Capítulo 3. Se han obtenido las siguientes conclusiones:

En primer lugar, en términos de la visibilidad del escritor, los alumnos nativos y los alumnos chinos de inglés tienen diferentes grados de coloquialismo. La característica más obvia es el uso excesivo de pronombres de primera y segunda persona. La escritura académica de los alumnos nativos es principalmente en primera persona del singular, mientras que la escritura académica de los alumnos chinos de inglés está principalmente en la forma plural de la primera persona. Esto muestra que los alumnos nativos tienen una gran conciencia de sí mismos, mientras que los alumnos chinos tienden a expresar opiniones grupales. Además, la visibilidad del escritor disminuye con la mejora del dominio del inglés, lo que significa que con el dominio del inglés, el nivel de inglés coloquial en la

escritura académica también cambia.

En segundo lugar, en términos de nivel léxico, se ha analizado comparativamente la diversidad de vocabulario, la densidad de vocabulario, la frecuencia de vocabulario coloquial y la frecuencia de las abreviaturas y contracciones, y se encontró que, en la escritura académica de los alumnos nativos, aunque la diversidad de vocabulario sea mayor que para los alumnos chinos, la tasa de repetición es baja y es más profesional. Además, la densidad de vocabulario de los alumnos nativos es casi un 0,6% menor que la de los alumnos chinos, lo que significa que la cantidad de información textual es más coloquial que en los alumnos chinos. Al mismo tiempo, los alumnos nativos utilizan más vocabularios abreviados en la escritura académica, lo que puede deberse a sus hábitos lingüísticos. Lo que los alumnos chinos y los alumnos nativos tienen en común es que utilizan palabras coloquiales de alta frecuencia en diferentes grados en sus escrituras académicas. Se cree que, en los alumnos chinos, debido a la transferencia negativa de su lengua materna, los alumnos chinos usan las palabras coloquiales con más frecuencia en las escrituras académicas.

En tercer lugar, en términos del nivel sintáctico, se ha analizado comparativamente la longitud promedio de la frase y la estructura de las frases, y se ha encontrado que, la longitud promedio de las frases de las escrituras académicas de los alumnos nativos es mucho mayor que la de los alumnos chinos, lo que puede indicar que las frases de los alumnos nativos son más complicadas que las de los alumnos chinos, lo cual puede deberse a los diferentes hábitos lingüísticos del chino y el inglés. A los chinos les gusta usar frases más cortas y sencillas para expresar sus opiniones, mientras que a los alumnos nativos les gusta usar frases más largas. En la estructura de la frase, los alumnos chinos están acostumbrados a algunas estructuras fijas en la escritura académica. Creen que el uso de estas estructuras puede mejorar la

escritura académica, pero realmente están afectados por la transferencia cultural negativa y los hábitos lingüísticos.

En resumen, ya sean alumnos nativos o alumnos chinos, existe un cierto grado de uso coloquial en su escritura académica. En comparación con los alumnos nativos, el lenguaje coloquial de la escritura académica de los alumnos chinos de inglés de alto nivel (profesionales de inglés) no es muy diferente, pero aparece en diferentes aspectos. Esto puede deberse a la influencia de la transferencia negativa de la lengua materna entre los alumnos chinos. Además, en comparación con los alumnos nativos avanzados y con los alumnos chinos de inglés avanzados, la escritura académica de los alumnos nativos de pregrado tiene más características coloquiales y tiene una tendencia coloquial más obvia, lo que demuestra que, con la mejora del dominio del inglés, la tendencia coloquial en la escritura académica va disminuyendo, lo que puede deberse a los diferentes hábitos lingüísticos de los alumnos nativos y de los alumnos chinos. Por lo tanto, en combinación con las posibles razones dadas en el Capítulo 4, se puede concluir que:

Para los alumnos nativos, la mayoría de las características coloquiales en su escritura académica son causadas por sus propios hábitos lingüísticos. El hábito lingüístico inherente provoca la aparición de fenómenos coloquiales en la escritura académica. Por lo tanto, los alumnos nativos deben distinguir correctamente la parte coloquial del hábito lingüístico en la escritura académica y tratar de evitar este hábito lingüístico inherente a la escritura académica.

Para los alumnos chinos, las características coloquiales de su escritura académica se derivan de la transferencia negativa de su lengua materna y de los métodos tradicionales de enseñanza del inglés. Los profesores deben crear un entorno de aprendizaje de inglés puro, usar el inglés tanto como sea posible en el proceso de enseñanza, guiar a los alumnos a que piensen activamente en

inglés y superar la transferencia negativa de la lengua materna en el aprendizaje del inglés en la mayor medida. Al mismo tiempo, los profesores deben tomar la iniciativa para cambiar. El modelo tradicional de enseñanza del inglés utiliza el pensamiento fijo para restringir a los alumnos. Desde el principio, los alumnos chinos deben comprender la diferencia entre el inglés coloquial y el escrito, reconocer la diferencia entre el chino y el inglés, y dar importancia al cultivo de las habilidades de la escritura.

5.2. Limitaciones e investigaciones futuras

En este estudio se han comparado cuatro corpus de escritura académica por parte de los alumnos nativos y alumnos chinos para estudiar las características coloquiales de la escritura académica. Todos los datos y procesos se han basado en la literatura y las teorías relevantes, y se refieren a resultados de investigaciones anteriores de otros autores. Por eso esta investigación es completamente científica. Sin embargo, no se pueden ignorar algunas limitaciones de este estudio.

En primer lugar, de los 4 corpus utilizados en este estudio, BASE (1.477.281), BAWE (6.968.089), LM (500.519) y TFM (325.439), BASE y BAWE, usados directamente desde el *Sketch Engine* tienen muchas más palabras y son mucho más grandes que los dos corpus construidos por el autor. Aunque los resultados de los cálculos de la investigación se calculan por millón, estos aún afectan a la precisión de los datos. Si se pudiera construir un corpus más grande, los resultados serían más convincentes. Al mismo tiempo, para estudiar la relación entre el grado de coloquialismo en la escritura académica y el nivel de inglés, el autor estableció la LM como una base de datos terminológica para los alumnos de pregrado nativos. Sin embargo, debido a limitaciones de tiempo, el autor encontró menos información, por lo que se fusionó con el corpus LOCNESS. No todos los artículos están escritos por alumnos de pregrado, lo que ha afectado a

la precisión de los resultados.

En segundo lugar, los parámetros seleccionados en esta investigación son la visibilidad del escritor, la densidad del vocabulario, la diversidad del vocabulario, la frecuencia del vocabulario, la longitud y la estructura de la frase. Esta investigación ha tenido como objetivo explorar estas características coloquiales a través de estos parámetros. Sin embargo, debido a la limitación de tiempo y de las herramientas de análisis, este estudio no ha analizado parámetros más representativos como la "estructura pasiva".

En tercer lugar, a pesar de utilizar herramientas de lenguaje como *WordSmith 4.0* y *AntConc* para la investigación, han podido haber pequeñas omisiones y errores en el cálculo manual de los datos, lo que también afecta a la precisión de los resultados.

Por lo tanto, se recomienda disponer del tiempo suficiente para futuras investigaciones, recopilar una gran cantidad de corpus y ampliar el número de palabras en el corpus lo máximo posible. También se debería optar por utilizar herramientas que puedan ocuparse del análisis de la estructura de la frase y utilizar parámetros más descriptivos para el análisis comparativo. Al mismo tiempo, en el proceso de recopilación y de análisis de datos, se deberían utilizar más programas de computación y se deberían reducir los cálculos manuales para mejorar aún más la validez y la confiabilidad de los resultados.

BIBLIOGRAFÍA

Akinnaso, F. N. (1982). On the differences between spoken and written language. *Language and speech*, 25(2), 97-125.

Bailey, S. (2003). *Academic writing: A practical guide for students*. Psychology Press.

Bhatia, V. K. (1999). Integrating products, processes, purposes and participants in professional writing. *Writing: Texts, processes and practices*, 21-39.

Bi, J., Qin, X.Q., y Song, D.W. (2014). *Las tendencias de investigación en la escritura académica de una segunda lengua extranjera y su explicación*.

Biber, D., y Conrad, S. (1999). Lexical bundles in conversation and academic prose. *Language and Computers*, 26, 181-190.

Biber, D. S. (1998). Conrad y R. Reppen. *Corpus linguistics: investigating language structure and use*. Cambridge: Cambridge University Press.

Biber, D. (1991). *Variation across speech and writing*. Cambridge University Press.

Brutt-Griffler, J. (1998). Conceptual questions in English as a world language: Taking up an issue. *World Englishes*, 17(3), 381.

Carroll, J. M. (1979). 'Mere length' and sentence comprehension units: An argument from ambiguity bias. *Research on Language & Social Interaction*, 12(3-4), 443-450.

Cao, R.Y., y Lu, J. (2016). Análisis de la correlación entre las diferencias de estilo en el inglés y el chino y la tendencia de la escritura coloquial en inglés de los alumnos universitarios chinos. *Shandong Social Sciences*, (S1), 490-491.

- Chafe, W. (1982). Integration and involvement in speaking, writing, and oral literature. *Spoken and written language: Exploring orality and literacy*, 35-54.
- Chen, Y.J. (2014). *Un análisis basado en el corpus del análisis de errores de los verbos predicados dobles en los estudiantes universitarios chinos y comparación del uso del idioma en los hablantes nativos*. Southwest University for Nationalities.
- Chen, X.Y. (2009). Investigación e importancia de la localización en el inglés y el inglés chino. *Revista de la Universidad de Guangxi para Nacionalidades: Edición de Filosofía y Ciencias Sociales*, (5), 163-168.
- Cobb, T. (2003). Analyzing late interlanguage with learner corpora: Quebec replications of three European studies. *Canadian Modern Language Review*, 59(3), 393-424.
- Copland, F., y Creese, A. (2015). ETHICAL ISSUES IN LINGUISTIC ETHNOGRAPHY. *Ethics in Applied Linguistics Research: Language Researcher Narratives*, 161.
- Coxhead, A. (2000). A new academic word list. *TESOL quarterly*, 34(2), 213-238.
- Ellis, R. (1985). Sources of variability in interlanguage. *Applied linguistics*, 6(2), 118-131.
- Feng, R.L., y Wu, X.Y. (2011). Un estudio comparativo basado en el corpus sobre la tendencia coloquial de los alumnos universitarios en la escritura del inglés. *English Teachers*, (9), 47-52.
- Gilquin, G., y Paquot, M. (2008). Too chatty: Learner academic writing and register variation. *English Text Construction*, 1(1), 41-61.

- Granger, S. (1998). *The computer learner corpus: a versatile new source of data for SLA research* (3-18). na.
- Halliday, M. A. K., y Hasan, R. (2002). *Cohesion in English Revision VI*. Beijing: Foreign Language Teaching and Research Press.
- Halliday, M. A. K. (1994). Spoken and written modes of meaning. *Media texts: Authors and readers*, 7, 51-73.
- Han, Y.N. (2013). *Visibilidad del escritor / lector en los artículos de argumentación para estudiantes chinos avanzados en inglés: un estudio basado en el corpus*. Universidad de Ciencia y Tecnología de Tianjin.
- Hinkel, E. (2002). *Second language writers' text: Linguistic and rhetorical features*. Routledge.
- Hou, W.R. (1988). Investigación estilística y traducción. *Enseñanza e investigación de lenguas extranjeras: Lenguas extranjeras Bimestral*, 18-27.
- Huang, Z.R., y Jia, C.N. (2013). La aplicación y los problemas de la longitud promedio de las oraciones en la investigación de la adquisición del lenguaje. *Revista de la Universidad de Yangtze: Edición de Ciencias Sociales*, (1), 95-97.
- Hyland, K. (2013). 5 ESP and Writing. *The handbook of English for specific purposes*, 95.
- Irvin, L. L. (2010). What Is "Academic" Writing?. *Writing spaces: Readings on writing*, 1, 3-17.
- Ivanič, R. (1998). *Writing and identity: The discursal construction of identity in academic writing*. John Benjamins.

- Jiang, Y.S. (2018). *Una investigación basada en el corpus sobre la oralización de la escritura académica en los posgrados en inglés*. Universidad Marítima de Dalian.
- Krauthamer, H. S. (1999). *Spoken language interference patterns in written English* (Vol. 42). P. Lang.
- Laufer, B., y Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied linguistics*, 16(3), 307-322.
- Leech, G., y Garside, R. (1991). Running a grammar factory: the production of syntactically analysed corpora or treebanks. *English Computer Corpora: selected papers and research guide*, 15-32.
- Leki, I., y Carson, J. G. (1994). Students' perceptions of EAP writing instruction and writing needs across the disciplines. *TESOL quarterly*, 28(1), 81-101.
- Liao, X. (2013). La tendencia coloquial de las especialidades en la escritura en inglés: un estudio basado en el corpus. *Inglés de ultramar*, (21), 137-138.
- Liu, X.J. (2007). *Características del vocabulario del inglés coloquial*. Universidad Marítima de Dalian.
- Liu, X.B. (2005). La tendencia coloquial de los estudiantes universitarios chinos en el inglés escrito. *Investigación en tecnología y educación superior*, 24 (1), 109-112.
- Macaulay, M. (2011). *Processing varieties in English: An examination of oral and written speech across genres*. UBC Press.
- Ma, G.H., y Wen, Q.F. (1999). Un estudio sobre los factores que influyen en la capacidad de escritura en inglés en los alumnos universitarios. *Enseñanza e*

investigación de lenguas extranjeras: lenguas extranjeras bimestral, (4), 34-39.

Murray, R., y Moore, S. (2006). *The handbook of academic writing: A fresh approach*. McGraw-Hill Education (UK).

Odlin, T. (1989). *Language transfer* (Vol. 27). Cambridge: Cambridge University Press.

Paltridge, B., Harbon, L., Hirsh, D., Shen, H., Stevenson, M., Phakiti, A., y Woodrow, L. (2009). *Teaching academic writing: An introduction for teachers of second language writers*. Ann Arbor, MI: University of Michigan Press.

Paquot, M. (2010). *Academic vocabulary in learner writing: From extraction to analysis*. Bloomsbury Publishing.

Petch-Tyson, S. (1998). Reader/writer visibility in EFL persuasive writing.

Pöchhacker, F. (2016). *Introducing interpreting studies*. Shanghai: Shanghai Foreign Language Education Press.

Qin, X.B. (2002). *Un bosquejo del estilo y estilo en inglés*. Shanghai Foreign Language Education Press.

Reiss, K. (1989). Text Types, Translation Types and Translation Assessment In chesterman, A (Ed.) Readings in Translation Theory. *Helsinki: Finn Lectura*. S, 105-115.

Read, J. (2002). *Assessing vocabulary*. Cambridge: Cambridge University Press.

Richards, J. C. (2010). Academic writing for publication: Some points of departure for success. *Literacy Practice and Research*, 35(3), 50-53.

- Rice, W. R., y Chippindale, A. K. (2001). Intersexual ontogenetic conflict. *Journal of Evolutionary Biology*, 14(5), 685-693.
- Rong, P. (1991). El inglés de China existe objetivamente. *Revista de PLA University of Foreign Languages*, (1), 1-8.
- Smith, L. E. (1976). English as an international auxiliary language. *RELC journal*, 7(2), 38-42.
- Song, J. (2017). *Investigación sobre la tendencia coloquial de la escritura en inglés en los alumnos de secundaria*. Universidad Normal del Noreste.
- Song, Y. (2008). Un estudio basado en el corpus de la adquisición de la estructura There-be en los alumnos chinos de inglés. *Journal of Anhui Radio and TV University*, 2008 (3), 77-81.
- Spears, R. A. (2007). *McGraw-Hill's essential American idioms dictionary: [the up-to-date guide to the idioms of modern American English]*. McGraw-Hill.
- Šimčikaitė, A. (2012). Spoken discourse markers in learner academic writing. *Kalby studijos*, (20), 27-33.
- Sun, Y.Z. (2012). *Características coloquiales en la escritura académica de los estudiantes chinos de inglés con un elevado nivel*. Universidad de Ciencia y Tecnología de Huazhong.
- Sun, L. (1989). El estudio de las variantes de países en inglés y el inglés en China. *Lenguas extranjeras*, (2), 19-25.
- Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.
- Ure, J. (1971). Lexical density and register differentiation. *Applications of linguistics*, 443, 452.

- Waibel, B. (2007). *Phrasal verbs in learner English: A corpus-based study of German and Italian students* (Doctoral dissertation, Verlag nicht ermittelbar).
- Wang, Y., y Jenkins, J. (2016). "Nativeness" and intelligibility: Impacts of intercultural experience through English as a lingua franca on Chinese speakers' language attitudes. *Chinese Journal of Applied Linguistics*, 39(1), 38-58.
- Wen, Q.F., Ding Y.R., y Wang, W.Y. (2003). La tendencia coloquial en el inglés escrito por los alumnos universitarios chinos: un análisis comparativo del corpus de los estudiantes de inglés con un alto nivel. *Enseñanza e Investigación de Idiomas Extranjeros: Bimestralmente Idiomas Extranjeros*, 35(4), 268-274.
- Wehmeier, S. (2004). Oxford Advanced Learner's English-Chinese Dictionary. *Beijing: The Commercial Press and Oxford University Press*, 1, 111-1.
- Whitaker, A. (2009). Academic Writing Guide 2010: A Step-by-Step Guide to Writing Academic Papers. *City University of Seattle*.
- Wu, J.H. (2014). Retrospect and Prospect: Reflections on the Status Quo of Chinese English Studies. *Contemporary Foreign Language Studies*, 5, 24-25.
- Xu, Y.Z., Guo, L.H., y Xu, T. (2007). Indispensable para la enseñanza del género en las tesis académicas-Una encuesta de la introducción de la tesis de maestría en las especialidades del inglés. *Idiomas extranjeros: chino e inglés*, (4), 47-51.
- Yu, Q.M. (2008). La tendencia coloquial de los alumnos universitarios no ingleses en la escritura del inglés: un estudio basado en el corpus. *Boletín de*

educación vocacional: Journal of Jiangsu Normal University of Technology,
23(8), 106-109.

Yu J.P., Xu G.F., y Bai F.X. (2011). Problemas y contramedidas de la escritura académica en inglés de los alumnos chinos. *Teaching Research*, (6), 41-44.

Zhang, J.H. (2018). *Oralization of Written English by Chinese Scholars*. Xi'an International Studies University.

Zhang, R. (2020). Análisis de la escritura académica del inglés en China.

Zhou, H.Z. (1994). Características del inglés en el siglo XXI. *Conocimiento de enciclopedia*, (3), 13.