El término en español:
hacia una ampliación del concepto a partir de su caracterización lingüística en cinco áreas de especialidad

Carlos Mario Pérez Pérez

Tesis para optar al título de Doctor en Lingüística

Asesor

Dr. Gabriel Quiroz Herrera

Profesor Titular

Programa de Traducción Inglés – Francés - Español

Escuela de Idiomas

Universidad de Antioquia

Doctorado en Lingüística
Facultad de Comunicaciones y Filología
Universidad de Antioquia
2023

UNIVERSIDAD
DE ANTIOQUIA
1803

Cita (Pérez-Pérez, 2023)

Referencia Estilo APA 7 (2020) Pérez-Pérez, C. M. (2023). *El término en español: hacia una ampliación del concepto a partir de su caracterización lingüística en cinco áreas de especialidad* [Tesis Doctoral, Universidad de Antioquia].

creative commons BY NC SA

Doctorado en Lingüística, Cohorte VIII.

Grupo de Investigación Traducción y Nuevas Tecnologías (TNT)

UNIVERSIDAD DE ANTIOQUIA
Facultad de Comunicaciones y Filología

Sistema de Bibliotecas

Biblioteca Carlos Gaviria Díaz

Repositorio Institucional: http://bibliotecadigital.udea.edu.co

Universidad de Antioquia - www.udea.edu.co

El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Antioquia ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por los derechos de autor y conexos.

A mi madre, Aleida María y mi padre Luis Emilio.

A mi abuela, María del Consuelo.

A mi tía Viviana.

A mis hermanos, Julián Andrés, María Isabel y Samuel.

Los frutos de este nuevo trabajo de investigación también serán cosechados en familia.

Agradecimientos

Ciertamente, *una tesis es aquello que hacemos mientras la vida pasa*. Un camino que día a día se recorre al vaivén de las cosas pequeñas, en la procura de construir nuevos horizontes. Un desafío que se vive en lo individual del entorno que nos rodea. Mis excusas para aquellos que, sin duda, quedarán por fuera de estas líneas, en su momento les he expresado mí gratitud.

Agradecer a Dios por regalarme vida para llevar a buen puerto este otro proyecto de investigación. Del Cielo los dones para afrontar los desafíos y la calma para avanzar en medio de los días de tormenta, la serenidad y prudencia en aquellos de sol intenso. La última piedra está a puertas de encontrar su lugar. A mi familia, estos años han estado llenos de novedades, unas muy felices otras no tanto. Hemos logrado sortear los azares del camino y crecer en el amor. Kyra, Nieve, Bruno y Polito... su cariño es inexplicable. A todos, gracias por entender las ausencias, los silencios... gracias por el tiempo que me han regalado. Esta tesis de doctorado les pertenece. Madre, el don de tu vida es el mejor regalo.

Profesor Gabriel, gracias por confiar en mí para este proyecto. Por las muchas discusiones y las largas jornadas de trabajo, sin importar el día ni la hora. Gracias por compartir el conocimiento y por el apoyo anímico en las altas y bajas. Gracias por creer en mi criterio en las decisiones tomadas. A mis colegas del TNT, seguiremos fortaleciendo nuestros saberes. Gracias por siempre estar atentos a las necesidades en el camino. A la Facultad de Comunicaciones y Filología, a todas las profesoras y profesores, en el disenso seguiremos construyendo academia. A la Escuela de Idiomas, por el apoyo en los momentos en que fue necesario. A la Universidad de Antioquia, en tus pasillos continuaré construyendo ciencia. A mis colegas del doctorado, un placer compartir con ustedes. A mis amigos, las alegrías compartidas tienen un matiz diferente, gracias.

En cada línea de este escrito está tu presencia, el tiempo que hemos compartido ha sido un privilegio. Tus reflexiones inevitablemente han enriquecido este trabajo. Al Universo por el don de tu presencia. No hay palabras que logren describir mis pensamientos a vos. Gracias por regalarme tus segundos. Todo lo que he querido decirte, ya te lo he expresado. En el tiempo veré tus alas abrirse a la grandeza. Mi cariño viaja a ti. ¡Gracias!

Al Mario de 12 años. Tu valentía nos trajo hasta aquí, ya puedes descansar.

Resumen

La terminología, como disciplina de la lingüística, se ocupa del análisis y la descripción de los lenguajes de especialidad, siendo el *término* el objeto de su estudio. Así, bajo la mirada paradigmática vigente, los estudios terminológicos se han centrado principalmente en construcciones nominales, dejando de lado a las demás unidades terminológicas tanto simples (adjetivo, adverbio y verbo) como complejas, además de enfocarse en áreas como la *medicina* y la *economía*, rasgos que se evidencian no solo en los manuales de terminología sino también en la mayoría de las investigaciones en el campo a nivel mundial, convirtiéndose en una mirada restrictiva del fenómeno que no refleja completamente las dinámicas de la comunicación humana en relación con el uso del lenguaje.

La presente tesis doctoral tiene como objetivo describir lingüísticamente los términos, simples y complejos, en español y su variación en función de los contextos de especialidad en los que se encuentran inmersos. Para ello, se compiló un corpus especializado para la variación terminológica horizontal conformado por 538 artículos de investigación recuperados de países donde el español es lengua oficial y en cinco áreas del conocimiento, a saber: *medicina*, *lingüística, veterinaria, computación* y *filosofía*. Un total de 37.055 candidatos a término, simples y complejos, fueron recuperados automáticamente utilizando *Termostat* como herramienta de extracción terminológica. Además, también se extrajeron candidatos correspondientes a sintagmas nominales especializados (84.367), metáforas (1.600) y siglas (4.263). Asimismo, se implementó el *Protocolo para la validación de términos simples y complejos* diseñado a partir de aspectos formales de tipo morfológico, sintáctico y semántico teniendo en cuenta aspectos micro contextuales prototípicos, en la expresión del conocimiento de especialidad, de cada uno de los campos del saber analizados.

La validación se realizó de forma canónica. Es decir, se revisó manualmente cada uno de los candidatos en los cinco subcorpus. En primer lugar, se seleccionaron aquellos candidatos temáticamente relacionados con el campo del saber. En segundo lugar, se determinaron los términos simples y complejos, teniendo en cuenta los criterios establecidos en dicho protocolo. Adicionalmente, se registró la información correspondiente a las definiciones lexicográficas a partir del Diccionario de la lengua española (DRAE) y de diccionarios de especialidad para validar el carácter terminológico del candidato analizado, a saber: el Diccionario Mosby, el Diccionario Académico de la Medicina de la Academia Nacional de Medicina de Colombia y el Diccionario de términos académicos de la Real Academia Nacional de Medicina de España y el Diccionario de cáncer del Instituto Nacional del Cáncer, para *medicina*; el Diccionario de lingüística moderna, para *lingüística*; el Agrovoc, para *veterinaria*; el Portal terminológico de Microsoft, para *computación* y el Diccionario de filosofía de Ferrater y el Diccionario AKAL de filosofía, para *filosofía*. En tercer lugar, se recuperó información contextual de cada candidato desde el Corpus de Referencia del Español Actual (CREA) y el Corpus Textual Especializado para la Variación Terminológica Horizontal (VaTeHo) compilado en el marco de la presente tesis doctoral.

Los resultados permiten validar las hipótesis de trabajo. En primera instancia, todas las unidades léxicas simples o complejas pueden ser catalogadas como términos en cuanto permiten recuperar el contenido de especialidad en función del campo del saber al cual pertenecen y, por ende, en segunda instancia, cada área del conocimiento emplea mecanismos propios en la materialización de sus saberes, elemento que hemos denominado *variación terminológica horizontal tipológica*. Asimismo, en línea con nuestros hallazgos, se presenta una propuesta de ampliación en la definición del concepto de *término*, aspecto que en la actualidad se sigue presentando como uno de los mayores retos en la terminología debido a la falta de consenso en las distintas tradiciones y escuelas de pensamiento terminológico. En última instancia, por un lado, otros elementos lingüísticos como las metáforas, las siglas y los sintagmas especializados también

pueden ser catalogados como términos, característica que es necesario corroborar en futuras investigaciones en el campo. Por otro lado, se hace necesario una revisión de las herramientas terminológicas que permitan la inclusión de esta mirada amplia del fenómeno al incorporar las demás unidades terminológicas, más allá de las nominales, lo que permitiría una mirada sino completa, más amplia de los distintos saberes humanos.

Palabras clave: *lingüística computacional, lingüística de corpus, término, terminología, variación terminológica horizontal.*

Abstract

Terminology, as a discipline of linguistics, deals with the analysis and description of specialized languages, having the *term* as the object of its study. Thus, under the paradigmatic view that is currently prevalent, terminological studies have mainly focused on nominal constructions, leaving aside both simple lexical units (adjective, adverb and verb) and other complex terminological units, as well as being focused on areas such as *Medicine* and *Economics*. This issue is evident not only in terminology manuals but also in the majority of worldwide research in the field. And therefore, a restrictive view of the phenomenon that does not fully reflect the dynamics of human communication in relation to the use of language.

This doctoral thesis aims at linguistically describing *Spanish terms* and their variation in accordance with the specialized contexts in which they are found. In order to do it so, a specialized corpus for horizontal terminological variation was compiled, consisting of 538 research articles from countries where Spanish is the official language, and form five areas of knowledge: *Medicine, Linguistics, Veterinary, Computing* and *Philosophy*. 37,055 simple and complex candidate terms were automatically recovered using *TermoStat* as the terminological extraction tool. Specialized syntagmas (84,367), metaphors (1.600) and acronyms (4,263) were also extracted. Furthermore, *El protocolo para la validación de términos simples y complejos,* based on morphological, syntactic, and semantic aspects in recovering the micro-contextual aspects of specialized expression in each of the analyzed fields of knowledge was implemented.

The validation process was performed in a canonical way. That is, each candidate was manually verified in the five subcorpora. Firstly, those candidates that were thematically related to the field of knowledge were selected. Secondly, the simple and complex terms

were determined based on the criteria established in the protocol. Additionally, the information corresponding to lexicographic definitions was recorded based on *Diccionario de la lengua española* (DRAE) and on specialized dictionaries to validate the terminological nature of the candidate analyzed: *Diccionario Mosby, Diccionario Académico de la Medicina de la Academia Nacional de Medicina de Colombia* and *Diccionario de términos académicos de la Real Academia Nacional de Medicina de España* and *Diccionario de cáncer del Instituto Nacional del Cáncer* for *Medicine, Diccionario de lingüística moderna* for *Linguistics, Agrovoc* for *Veterinary, Portal terminológico de Microsoft* for *Computing* and *Diccionario de filosofía de Ferrater* and *Diccionario AKAL de filosofía* for *Philosophy*. Thirdly, the contextual information of each candidate was recovered from the *Corpus de Referencia del Español Actual* (CREA) and the *Corpus Textual Especializado para la Variación Terminológica Horizontal* (VaTeHo). The latter was compiled in the framework of this doctoral work.

The results validate the research hypotheses. First of all, every simple or complex lexical units can be classified as term as it allows for the retrieval of specialized content in relation to the field of knowledge to which they belong. And therefore, secondly, each area of knowledge employs its mechanisms in the materialization of its knowledge, which we have referred to as *Variacion Terminológica Horizontal Tipologica*. Furthermore, in virtue of our findings, a proposal for an expansion in the definition of the concept of *term* is presented, an aspect that currently remains to be one of the greatest challenges in terminology due to the lack of consensus among different traditions and Schools of terminological thought. Finally, on the one hand, other linguistic elements such as metaphors, acronyms, and specialized syntagmas can also be classified as terms, a characteristic that needs to be further verified in future research in the field. On the other hand, a revision of the terminological tools is necessary to allow the inclusion of this broad view of the phenomenon by incorporating the other terminological units, beyond the nominal

ones, which would allow a broader, if not complete, view of the different human knowledge.

Keywords: *Computational Linguistics, Corpus Linguistics, Term, Terminology, Horizontal Terminological Variation.*

Contenido

Agradecimientos VII
Resumen IX
Abstract XII
Lista de figuras XIX
Lista de tablas XXI
Acrónimos, abreviaturas, convenciones y siglas XXIII

PARTE I: TEORÍA

Introducción: un estudio sobre terminología en español 31
0.1 Hipótesis 37
0.2 Objetivos de la tesis 38
0.2.1 Objetivo general 38
0.2.2 Objetivos específicos 38
0.3 Estructura de la tesis 39
1. Terminología: una mirada en conjunto 43
1.1 Enfoques de trabajo terminológico 45
1.2 Investigaciones en terminología 52
1.2.1 Terminología en Latinoamérica 55
1.3 Recapitulación 61
2. El término: caracterización de la unidad de estudio 65
2.1 Terminología: una disciplina de la lingüística 66
2.2 El término 70
2.3 Bases teóricas para un protocolo de validación de términos simples y complejos 76
2.3.1 Cuestiones preliminares en el diseño del protocolo de validación 77
2.3.1.1 Unireferencialidad del término *78*
2.3.1.2 Composicionalidad del término complejo *80*

2.3.1.3 Arbitrariedad del término ... *80*
2.3.1.4 Convencionalidad del término ... *82*
2.3.2 Unidades léxicas terminológicas ... 83
2.3.2.1 El adjetivo terminológico ... *84*
2.3.2.2 El adverbio terminológico ... *85*
2.3.2.3 El sustantivo terminológico ... *87*
2.3.2.4 El verbo terminológico ... *88*
2.3.2.5 Unidades complejas terminológicas ... *89*
2.4 Variación terminológica horizontal ... 90
2.5 Recapitulación ... 91

PARTE II: METODOLOGÍA

3. Metodología: un corpus del español para la variación terminológica horizontal ... 97
3.1 Conceptos generales en nuestra metodología de investigación ... 100
3.2 Metodología y corpus para la variación terminológica horizontal (VaTeHo) ... 102
3.2.1 Fase 01: Preliminares ... 103
3.2.1.1 Áreas del saber ... *104*
3.2.1.2 Tipo de texto ... *105*
3.2.1.3 Países y distribución del corpus ... *105*
3.2.1.4 Herramientas para la extracción y procesamiento de los datos ... *106*
3.2.1.5 Criterios de selección y limpieza de los datos ... *109*
3.2.2 Fase 02: Corpus especializado del español para el estudio de la variación terminológica horizontal (VaTeHo) ... 111
3.2.2.1 Conformación final del corpus VaTeHo ... *111*
3.2.2.2 Base de datos y gestión documental ... *115*
3.2.3 Fase 03: Extracción de los candidatos a término y muestra de trabajo 117
3.2.4 Fase 04: Aspectos metodológicos para la aplicación del protocolo de validación ... 120
3.2.4.1 Criterios para la determinación de unidades simples como términos ... *123*

3.2.4.2 Criterios para la determinación de unidades complejas como términos *123*
3.2.5 Fase 05: Resultados, análisis y escritura del informe final de investigación 126
3.3 Recapitulación 128

PARTE III: EL TÉRMINO EN ESPAÑOL

4. Resultados: una antesala al concepto de término en español 133
4.1 Generalidades de la extracción en *Termostat* *134*
4.1.1 Datos generales subcorpus de medicina 136
4.1.2 Datos generales subcorpus de lingüística 141
4.1.3 Datos generales subcorpus de veterinaria 145
4.1.4 Datos generales subcorpus de computación 149
4.1.5 Datos generales subcorpus de filosofía 152
4.1.6 VaTeHo: Procesamiento de los datos a partir de la extracción en bruto 156
4.1.7 VaTeHo: Muestra de candidatos a término para la validación 159
4.2 VaTeHo: Aplicación de un protocolo para la validación de candidatos a término 160
4.2.1 Hallazgos terminológicos en medicina 161
4.2.2 Hallazgos terminológicos en lingüística 171
4.2.3 Hallazgos terminológicos en veterinaria 180
4.2.4 Hallazgos terminológicos en computación 191
4.2.5 Hallazgos terminológicos en filosofía 198
4.3 Recapitulación 204
5. El término en español 207
5.1 Ampliación de nuestra unidad de estudio 208
5.1.1 Los términos simples en español 224
5.1.2 Los términos complejos en español 229
5.1.3 Otras unidades terminológicas 231
5.2 Recapitulación 235
6. Conclusiones: una mirada al futuro 237
6.1 Validación de las hipótesis y objetivos de investigación 239

6.2 Aportes de la investigación a la terminología 242
6.3 Limitaciones y perspectivas de trabajo futuro. 244

Referencias 247

Recursos lexicográficos 267

Programas y herramientas computacionales 269

Epílogo 271
Monarcas 271

Lista de figuras

Figura 1. Fases metodología. Diseño propio. 103

Figura 2. Matriz Base de Datos VaTeHo. 116

Figura 3. Campos base de datos VaTeHo. 117

Figura 4. Excel con los datos de la validación de los candidatos a término columnas A-J. 122

Figura 5. Interfaz corpus VaTeHoVe. 127

Figura 6. Interfaz de consulta VaTeHoVe. 128

Figura 7. Interfaz de consulta VaTeHoVe. 128

Figura 8. Total de candidatos a término por área del saber. 135

Figura 9. Nube de palabras medicina. 137

Figura 10. Nube de palabras lingüística. 141

Figura 11. Nube de palabras veterinaria. 145

Figura 12. Nube de palabras computación. 149

Figura 13. Nube de palabras filosofía. 153

Figura 14. Porcentajes limpieza temática. 157

Figura 15. Cálculos tamaño muestral. 160

Figura 16. Distribución de las estructuras sintácticas en la muestra de trabajo. 161

Figura 17. Cantidad de términos versus no términos. 204

Figura 18. Nube palabras medicina - Voyant Tools (Sinclair & Rockwell, 2022). 216

Figura 19. Nube palabras lingüística - Voyant Tools (Sinclair & Rockwell, 2022). 216

Figura 20. Nube palabras veterinaria - Voyant Tools (Sinclair & Rockwell, 2022). 217

Figura 21. Nube palabras computación - Voyant Tools (Sinclair & Rockwell, 2022). 217

Figura 22. Nube palabras filosofía - Voyant Tools (Sinclair & Rockwell, 2022). 218

Figura 23. Nube palabras medicina con lista de exclusión - Voyant Tools (Sinclair & Rockwell, 2022). 218

Figura 24. Nube palabras lingüística con lista de exclusión - Voyant Tools (Sinclair & Rockwell, 2022). 219

Figura 25. Nube palabras veterinaria con lista de exclusión - Voyant Tools (Sinclair & Rockwell, 2022). 219

Figura 26. Nube palabras computación con lista de exclusión - Voyant Tools (Sinclair & Rockwell, 2022). 220

Figura 27. Nube palabras filosofía con lista de exclusión - Voyant Tools (Sinclair & Rockwell, 2022). 220

Figura 28. Nube palabras medicina – adaptación propia en Word Cloud Generator (Davies, 2022). 221

Figura 29. Nube palabras lingüística – adaptación propia en Word Cloud Generator (Davies, 2022). 222

Figura 30. Nube palabras veterinaria – adaptación propia en Word Cloud Generator (Davies, 2022). 222

Figura 31. Nube palabras computación – adaptación propia en Word Cloud Generator (Davies, 2022). 223

Figura 32. Nube palabras filosofía – adaptación propia en Word Cloud Generator (Davies, 2022). 223

Figura 33. Comparativo criterios validación términos simples. 225

Figura 34. Comparativo criterios validación términos complejos. 230

Lista de tablas

Tabla 1. Google Académico: Investigadores en terminología. Elaboración propia. 53

Tabla 2. Relación de artículos por país y por área. 112

Tabla 3. Relación de tokens por país y por área. 114

Tabla 4. Estadística extracción TermoStat - medicina. 138

Tabla 5. Estadística extracción TermoStat - lingüística. 142

Tabla 6. Estadística extracción TermoStat - veterinaria. 146

Tabla 7. Estadística extracción TermoStat - computación. 150

Tabla 8. Estadística extracción TermoStat - filosofía. 154

Tabla 9. Conformación final de la muestra de trabajo por candidato y por área. 160

Tabla 10. Etiquetas términos en medicina según validación por protocolo. . . 162

Tabla 11. Etiquetas no términos en medicina según validación por protocolo. 168

Tabla 12. Etiquetas términos en lingüística según validación por protocolo. 171

Tabla 13. Etiquetas no términos en lingüística según validación por protocolo. 177

Tabla 14. Etiquetas términos en veterinaria según validación por protocolo. 181

Tabla 15. Etiquetas no términos en veterinaria según validación por protocolo. 187

Tabla 16. Etiquetas términos en computación según validación por protocolo. 191

Tabla 17. Etiquetas no términos en computación según validación por protocolo. 196

Tabla 18. Etiquetas términos en filosofía según validación por protocolo. . . .199

Tabla 19. Etiquetas no términos en filosofía según validación por protocolo. 203

Tabla 20. Cantidad de candidatos a término simples versus complejos. 214

Acrónimos, abreviaturas, convenciones y siglas

[...]	Omisión de una parte de contexto textual
[N+Adj]	Nombre + Adjetivo
[N+Adj+Adj]	Nombre + Adjetivo + Adjetivo
[N+Adj+Adj+Adj]	Nombre + Adjetivo + Adjetivo + Adjetivo
[N+Adj+Adj+Adj+Adj]	Nombre + Adjetivo + Adjetivo + Adjetivo + Adjetivo
[N+Adj+Conj+Adj]	Nombre + Adjetivo + Conjunción + Adjetivo
[N+N]	Nombre + Nombre
[N+Prep+N]	Nombre + Preposición + Nombre
[N+Prep+N+Adj]	Nombre + Preposición + Nombre + Adjetivo
Adj	Adjetivo
Adv	Adverbio
ARG	Argentina
BOL	Bolivia
CEMC	Corpus del Español Mexicano Contemporáneo
CHL	Chile
CINAN	Centro de Informática Aplicada a la Normalización
CITERM	Centro de Investigación Terminológica, de la Universidad Autónoma de Manizales, Colombia
COL	Colombia

COLTERM	Red Colombiana de Terminología, Colombia
Com	Computación
CRI	Costa Rica
CTTE	Centro de Traducciones y Terminología Especializada, Cuba
CUB	Cuba
DEM	Diccionario del Español de México
DIN	Deutsches Institut für Normung, Alemania
DRAE	Diccionario de la lengua española
DOM	República Dominicana
DUECh	Diccionario de uso del español de Chile
ECU	Ecuador
ESH	Sahara Occidental
ESP	España
Fil	Filosofía
GITT	Grupo de Investigación en Terminología y Traducción, Universidad de Antioquia, Colombia
GNQ	Guinea Ecuatorial
GTERM	Grupo en Terminología y Organización del Conocimiento, Universidad de la República de Uruguay, Uruguay
GTM	Guatemala
HND	Honduras
ID artículo	Identificación del artículo
INFOTERM	International Information Centre for Terminology

ISO	International Organization for Standarization
IULA	Instituto de Lingüística Aplicada, Universidad Pompeu Fabra, España
Lin	Lingüística
Med	Medicina
MEX	México
N	Nombre (Sustantivo)
NCI	Instituto Nacional del Cáncer, Estados Unidos
NIC	Nicaragua
NLTK	Natural Language Toolkit
No – Acus	No es término: el núcleo es deverbal y el modificador es objeto directo
No – AdjGeneral	No es término: adjetivo general
No – Deadj	No es término: el núcleo es de-adjetival
No - Ejemplo de	No es término: la secuencia es un ejemplo de un tipo de núcleo
No – Ncomún	No es término: es nombre común
No – Nomin	No es término: el núcleo es sujeto y el modificador es nombre deverbal o el núcleo es deverbal y el modificador es sujeto
No - Ocurre en	No es término: el núcleo ocurre en un lugar o momento indicado por el modificador
No – Partit	No es término: el núcleo es un partitivo del modificador
No – Prop	No es término: el núcleo es una propiedad del modificador
No – Unireferencial	No es término: el candidato no es unireferencial

No – Vsoporte	No es término: verbo soporte
PAN	Panamá
PDF	Portable Document Format
PDVSA Intevep	Centro de Investigaciones y Apoyo Tecnológico de Petróleos de Venezuela
PER	Perú
PLN	Procesamiento de Lenguaje Natural
POS	Part of Speech
Prep	Preposición
PRY	Paraguay
RITerm	Red Iberoamericana de Terminología
Sí – AdjDenominal	Término: adjetivo de-nominal
Sí – AdjGeneral	Término: adjetivo general
Sí – AdjRelacional	Término: adjetivo relacional
Sí -AdvAdjetival-Mente	Término: adverbio adjetival terminado en -mente
Sí – Arbitrario	Término: no cumple con los criterios de validación semántica o de propiedades compartidas, pero es un nombre que se le ha dado a un objeto o espécimen del área
Sí - Hecho de	Término: el núcleo está hecho del modificador
Sí - Metf hecho de	Término: el núcleo es una metáfora que está hecho del modificador
Sí – NcompCulta	Término: nombre por composición culta
Sí – Ncomún	Término: nombre común
Sí – NmorfSubj	Término: nombre por composición morfológica

Sí – NproductMorf	Término: nombre por productividad morfológica
Sí – Ntérmino	Término: nombre término
Sí - Tipo de	Término: el término es un tipo de *núcleo*
Sí - Tipo de/Hecho de	Término: ambos *Tipo de* o *Hecho de*
Sí – Vtérmino	Término: verbo término
SLV	El Salvador
SNEE	Sintagma Nominal Especializado Extenso
ST	Socioterminología
TBM	Terminología Basada en Marcos
TC	Terminología Cultural
TCT	Teoría Comunicativa de la Terminología
TE	Terminología Experimental
Termiazúcar	Grupo de Investigación Termiazúcar, Universidad del Valle, Colombia
TGT	Teoría General de la Terminología
TNT	Grupo de Investigación en Traducción y Nuevas Tecnologías, Universidad de Antioquia, Colombia
TS	Terminología Social
TSC	Terminología Sociocognitiva
TT	Terminología Textual
TV	Terminología Variacionista
TXT	Text File
UNESCO	Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura
UNICODE	Universal Character Encoding

Unifé	Universidad Femenina del Sagrado Corazón, Perú
URL	Universal Resource Locator
URY	Uruguay
UT	Unidad Terminológica
UTF-8	Universal Transformation Format-8
V	Verbo
VaTeHo	Variación Terminológica Horizontal
VaTeHoVe	Variación Terminológica Horizontal y Vertical
VEN	Venezuela
VenTerm	Asociación Venezolana de Terminología
Vet	Veterinaria
XLS	Microsoft Excel Spreadsheet File

PARTE I:
TEORÍA

DE ANTIOQUIA
ELLIN

Introducción: un estudio sobre terminología en español

"Allí estaban inmóviles al viento,
serenas a la espera del tiempo,
impasibles e imponentes,
silenciosas e indiferentes a una nueva realidad".

(Monarcas).

Hablar de terminología es hablar de la historia misma de la humanidad, tal cual lo expresaba Fedor de Diego al caracterizarla como un fenómeno lingüístico tan antiguo como el hombre mismo (1995, p. 13). De allí, una amalgama de aportes que han dotado al campo de un cuerpo teórico y metodológico para sustentar esta necesidad de nombrar las realidades existentes, en especial, aquellas propias de los saberes humanos y, en consecuencia, el desafío aún vigente de delimitar y definir su objeto de estudio: el *término*. Al respecto, Cabré (1995) introduce una mirada triple al observarlo como una manera de *saber* (lingüística), de *conocer* (filosofía) y de *transferir* y/o *comunicar* (disciplinas científico-técnicas) y, en consecuencia, el fundamento para la mirada poliédrica del *término* (Cabré, 2000) bajo la cual, actualmente, se observan muchos de los fenómenos terminológicos. Asimismo, para entender la terminología, su historia y su evolución es importante reconocer en ella un campo dinámico de investigación donde cada una de las escuelas de pensamiento terminológico han aportado sus propias concepciones sin llegar todavía a un consenso integrador, esto es, según afirma Cabré "muchos comienzos, que van circulando en vías paralelas y nunca llegan a encontrarse" (19 de noviembre de 2020).

A nivel teórico, luego de casi un centenario desde la concepción de la terminología clásica, fruto de los trabajos de la escuela de Viena, como una disciplina de la lingüística aplicada, su estudio aún se centra, principalmente, en el análisis de unidades nominales, relegando a un segundo plano otras unidades terminológicas como las simples (verbos, adjetivos, adverbios) y complejas (compuestos, sintagmas, frasemas, etc.), al igual que otras figuras lingüísticas, como las unidades metafóricas con carácter terminológico (*mouse* o *ratón* en la computación), como se evidencia en el contenido de los manuales disponibles en el área al día de hoy, a saber: Dubuc (1980)[1], Felber & Pitch (1984), Cabré (1992), Torruella & Sager (1993), Arntz & Picht (1995), Fedor de Diego (1995), Pavel & Nolet (2002), Cardero (2003), Cabré, Domènech & Estopà (2018) y Faber & L´Homme (2022). Asimismo, a pesar de que las principales escuelas de la terminología lograron a nivel teórico y metodológico una reformulación de los postulados de la Teoría General de la Terminología (TGT) (Wüster, 1998) y, en consecuencia, una ampliación del paradigma en el estudio terminológico vigente hasta el día de hoy, en muchas de sus investigaciones toman las unidades nominales como foco de estudio, siguiendo la tendencia clásica en la elección de la unidad de análisis.

A nivel metodológico, gracias a los avances tecnológicos, existe una gran cantidad de herramientas computacionales para el procesamiento automático de información que permiten abarcar grandes cantidades de datos en menor tiempo del empleado en los análisis manuales tradicionales. La implementación de estos programas, tales como *Termostat, BC_Term, Multiterm, UCREL, Wordnet, #LancsBox,* por mencionar algunos, ha permitido un estudio sistemático de distintas unidades con características terminológicas especializadas en diversas investigaciones a nivel mundial y, de este modo, se comenzó la discusión sobre la necesidad de ampliar el espectro de análisis en terminología a esas otras unidades lingüísticas relegadas o poco estudiadas. Además, desde el Procesamiento de Lenguaje Natural (PLN) se ha logrado

[1] El *Manuel Pratique de Terminologie* escrito por Robert Dubuc en el año 1992, fue traducido al español por la profesora Ileana Cabrera de la Pontificia Universidad Católica de Chile (UC) y publicado en el año de 1999.

avanzar en la interacción humano-máquina mediante el diseño y perfeccionamiento de tareas orientadas, entre otras finalidades, a la extracción de información a partir de corpus textuales (Sur, 2020; Tamayo, Burgos & Gelbukh, 2022a, 2022b; Tamayo, Angel & Gelbukh, 2022). No existe discusión sobre la importancia de la terminología para las diferentes áreas del conocimiento o como insumo esencial para el funcionamiento adecuado de muchos recursos tecnológicos. En efecto, gran parte del desarrollo de aplicativos en la *web* tienen un componente terminológico relevante para el almacenamiento y recuperación de la información. Igualmente, para la clasificación de textos y recuperación de la información, el estudio de las unidades terminológicas es de vital importancia para mejorar la precisión al reducir el *ruido* y el *silencio*[2] en sistemas de ese tipo, como lo son los buscadores *web* (Google, Bing, Yahoo! Search, etc.), los sistemas de gestión de bibliotecas, entre otros.

Desde la lexicografía, para la confección de diccionarios, el análisis de los tipos de unidades terminológicas es materia prima para saber qué se incluye o no en un diccionario además que permite caracterizar la naturaleza y el alcance de dicho recurso. Hasta ahora, la construcción de los diccionarios se ha realizado bajo el criterio del lexicógrafo o terminólogo sin tener en cuenta las realidades de cada área y las tendencias lingüísticas y estadísticas que configuran la forma como cada una de estas materializa su conocimiento específico. Para la traducción, la terminología es una de las disciplinas base que mejora la precisión y la calidad de las traducciones, al mismo tiempo que aumenta la eficiencia y el rendimiento del traductor. La relación entre ambas se ha construido, primordialmente, desde el aspecto práctico en cuanto para la traducción especializada se hace indispensable el aporte de la terminología mientras que la traducción se apoya de la terminología solo en contextos donde emergen necesidades neológicas desde un punto de partida lingüístico (Cabré, 2004, p. 90). De esta

[2] Los conceptos de *ruido* y *silencio* provienen del lenguaje computacional. Al respecto y con relación a la terminología, Porta (2006) define el *ruido* como "la proporción de falsos términos respecto al total de términos predichos" (p. 5) y el *silencio* como "la proporción de términos no reconocidos del total de términos predichos" (p. 6).

manera, la descripción de las unidades terminológicas es vital para la enseñanza de la terminología tanto para traductores como para los mismos terminólogos puesto que tener un criterio claro en la selección de dichas unidades para la toma de decisiones en traducción y terminología es un elemento determinante.

De este modo, los estudios, principalmente en tesis, se enfocan en el análisis de una sola unidad terminológica o parte de ella dentro de un área de conocimiento, preferentemente la medicina y la economía (Casademont, 2008); de un determinado tipo de texto o género textual (Tebé, 2005; da Cunha, 2008); de sintagmas terminológicos (Cortés, 2004; Quiroz, 2006, 2008; Sanz, 2011; Lopera, 2016; Vásquez, 2017; Quiroz & Arroyave, 2018); de colocaciones/frasemas (Bevilacqua, 2004; Rojas, 2014; Patiño, 2017), de adjetivos y participios (Folguerà, 2002; Salazar, 2011); de verbos (Lorente, 2002; Pérez-Pérez, 2018; Pérez-Pérez, Quiroz & Tamayo, 2020), de siglas (Giraldo, 2008; Rivas, 2017; Rivas, Quiroz, & Giraldo, 2021); por no mencionar otros estudios cuya unidad de análisis no es clara o es difusa donde se analizan casos de unidades poliléxicas (Burgos, 2014). Estos estudios con relación al análisis de una unidad terminológica o parte de ella en un área del conocimiento, normalmente inductivos, no podrían ser encuadrados como estudios de variación horizontal ya que ninguno analiza unidades terminológicas en varias áreas del conocimiento. Asimismo, algunos de estos estudios utilizan como corpus de contraste diccionarios especializados de diferentes áreas con el objetivo de validar las unidades encontradas, pero no hacen uso de corpus textuales para dicho fin.

Con relación a los estudios en variación terminológica estos se orientan, generalmente, a partir de dos criterios en pro de su tipologización. Por un lado, la variación determinada por la temática (horizontal) y, por otro lado, la variación según el grado de especialización (vertical). En línea con los objetivos de la presente investigación (§ 0.3), abordamos el eje horizontal de la variación. Al respecto, autores como Montero afirman que la variación horizontal está relacionada con dos aspectos esenciales: en primer lugar, lo temático y la perspectiva (2014, p. 5) que define cómo la arbitrariedad

de la primera permitiría la transferencia del conocimiento entre varias áreas de especialidad (Pavel, 1993, p. 21). En segundo lugar, la variación horizontal identifica las múltiples posibilidades de abordar un tema, otorgando el nivel de especialidad al discurso (Cabré, 1999, p. 18). Desde la perspectiva de Hoffmann, se comprende la manera como la *división horitzontal* de los lenguajes de especialidad hace referencia a la estructura de clasificación de los ámbitos comunicativos y sus correspondientes lenguajes de especialidad (1998, pp. 56-60). Lo anterior, corresponde a la misma naturaleza de la actividad humana en el ámbito terminológico.

Esta estructura es presentada como una escala donde se conjugan criterios de proximidad o lejanía y no simplemente como una lista de cada área, es decir, la variación terminológica horizontal permite realizar una clasificación de los lenguajes de especialidad más no su cuantificación (Rodríguez-Piñero & García, 2009, p. 917). Asimismo, Hoffmann, haciendo uso de medidas estadísticas, presenta una clasificación de los lenguajes de especialidad, donde la posición de cada uno se determina con relación al grado de identidad establecida con el sublenguaje de especialidad que le es menos distante. En este sentido, Domènech reafirma lo expresado por este autor agregando que este tipo de variación "es el fruto de la actividad de los interlocutores y sus saberes y, por lo tanto, puede variar en función de cada situación construida" (Domènech, 2006, p. 55). En ampliación a lo anterior, otros autores han abordado la variación horizontal desde diferentes puntos de vista, esto es: lo estético (Beneš, 1969), la ciencia, técnica, economía, política, entre otras (Schmidt, 1969), la milicia (Kolguixkin, 1970), los deportes (Riedel, 1975), entro otros.

En cuanto al concepto de término, su definición representa uno de los desafíos más importantes de la disciplina de cara a la integración teórica y operativa de este. Al respecto, normas como la ISO 1087 (2019), ISO 704 (2022) y autores como Felber & Pitch (1984), Cabré (1999), entre otros, dan cuenta de los acercamientos conceptuales y normativos con miras a una definición de la unidad de estudio. Al respecto, al momento de utilizar expresiones como *liofilizar*

o *carburador* no hay duda sobre el carácter terminológico de estos conceptos, incluso en contextos donde se desconoce su significado y/o la relación del concepto con el objeto al que designa, como ocurre en el segundo caso. Sin embargo, frente a casos como *atrás* o *adelante*, *cuchara, harina, gato*, entre otros, surge la duda sobre el carácter terminológico de estas unidades las cuales son observadas y definidas, incluso por expertos, bajo la óptica del *depende de*, la cual se define a la luz del contexto o el área temática. De este modo, se hace necesaria una propuesta de identificación de términos que contribuya con la aclaración de la ya descrita *frontera difusa* que suscita la identificación del valor terminológico de las unidades empleadas en los distintos contextos de interacción.

Teniendo como base el rastreo bibliográfico realizado hasta el momento de la entrega de la presente tesis doctoral, no se conoce un estudio que involucre un análisis lingüístico-descriptivo con un carácter horizontal que tomando un corpus de varias áreas del conocimiento analice las unidades terminológicas más allá del sustantivo. De esta manera, en la presente tesis doctoral[3] se analizan descriptivamente aspectos relacionados con la variación terminológica horizontal en español, entendida desde la óptica de las similitudes y diferencias en el plano lingüístico en las formas de expresión y materialización del conocimiento de especialidad. Para ello, se compiló un corpus textual de cinco (5) áreas del conocimiento (medicina, lingüística, veterinaria, computación y filosofía), atendiendo a criterios en pro de una mirada amplia e integradora del saber y, en consecuencia, de la terminología y su objeto de estudio, el término.

[3] Esta investigación hace parte del proyecto *Variación terminológica horizontal y vertical en español: hacia una descripción lingüística de las unidades terminológicas desde la lingüística de corpus* radicado por el Grupo de Investigación en Traducción y Nuevas Tecnologías (TNT), Escuela de Idiomas, Universidad de Antioquia y aceptado para financiación en la Convocatoria Programática Ciencias Sociales, Humanidades y Artes del Comité para el Desarrollo de la Investigación (CODI) de la Universidad de Antioquia para el año 2018 en el Acta nro. 2018-22970.

0.1 Hipótesis

Es importante destacar que la presente tesis se centra en explorar y describir el comportamiento de los términos a nivel horizontal, esto es, entre distintas áreas del conocimiento con miras a determinar si existen similitudes o diferencias a nivel lingüístico en los términos empleados en los diferentes campos del saber analizados. Así, por un lado, partimos de una concepción amplia del concepto de término, esto es, consideramos que cualquier unidad léxica ya sea simple o compleja tiene potencial terminológico (contrario a la concepción paradigmática de término que rige actualmente la mayoría de los estudios en terminología) y, por ende, puede ser catalogada como término según criterios morfosintácticos y semánticos, como se evidencia en los ejemplos en 1.

1. **Medicina**: *histológico* (Adj), *quirúrgicamente* (Adv), *cáncer* (N), *padecer* (V), *biopsia endoscópica* [N+Adj], *leucemia linfoblástica aguda* [N+Adj+Adj], *carcinoma de célula escamosa* [N+Prep+N+Adj].

   Lingüística: *lingüístico* (Adj), *semánticamente* (Adv), *término* (N), *sustantivar* (V), *aspecto verbal* [N+Adj], *diccionario escolar venezolano* [N+Adj+Adj], *corpus del español* [N+Prep+N].

   Veterinaria: *bovino* (Adj), *enzoóticamente* (Adv), *ganadería* (N), *ordeñar* (V), *carga ganadera* [N+Adj], *ganadería extensiva tradicional* [N+Adj+Adj], *cabeza de ganado* [N+Prep+N].

   Computación: *virtual* (Adj), *análogamente* (Adv), *aplicación* (N), *desarrollar* (V), *ambiente virtual* [N+Adj], *entorno de desarrollo* [N+Prep+N], *lenguaje de programación* [N+Prep+N].

   Filosofía: *filosófico* (Adj), *epistemológicamente* (Adv), *verdad* (N), *afirmar* (V), *ser humano* [N+Adj], *verdad ontológica* [N+Adj], *estado de cosa* [N+Prep+N], *género de conocimiento* [N+Prep+N].

Y, por otro lado, de la creencia que cada una de las áreas del conocimiento, gracias a la arbitrariedad propia de su fundamentación epistemológica, materializan de formas diversas o similares sus conceptos a través de los términos susceptibles de ser categorizados desde distintos puntos de vista (en nuestro caso particular desde el lingüístico) consideraciones que, en consecuencia, se traducen en las siguientes hipótesis:

a. Existen diferencias, a nivel lingüístico, en los términos simples o complejos utilizados por las diferentes áreas del conocimiento y, por ende, las características lingüísticas de estos términos varían de acuerdo con las particularidades de cada una de las áreas de conocimiento.

b. Todas las unidades lingüísticas, simples y complejas, más allá del nombre como categoría prototípica en el análisis terminológico, son susceptibles de contener el conocimiento de especialidad y, por lo tanto, ser catalogadas como términos.

0.2 Objetivos de la tesis

0.2.1 Objetivo general

Para comprobar las dos hipótesis arriba expuestas, se propone como objetivo general de esta tesis doctoral *describir el comportamiento lingüístico de los términos en español y su variación en función de los contextos de especialidad en los que se encuentran inmersos*.

0.2.2 Objetivos específicos

- Establecer los criterios específicos para la selección de las distintas áreas del conocimiento para la conformación del corpus de trabajo.

- Compilar un corpus especializado compuesto por cinco subcorpus de diferentes áreas del saber: medicina, lingüística, veterinaria, computación y filosofía.

- Extraer los candidatos a términos a partir del corpus especializado mediante la utilización de herramientas computacionales disponibles para tal propósito.

- Validar lingüísticamente los candidatos a términos a partir del protocolo para la validación de términos simples y complejos (Quiroz & Burgos, en prensa)

- Caracterizar los términos según criterios lingüísticos.

- Establecer las características prototípicas en las distintas áreas del conocimiento en la utilización de los términos.

- Ampliar el concepto de término a partir de la definición base utilizada en la presente tesis (Cabré, 1999b, p. 48) y los hallazgos en la aplicación del protocolo para la validación de candidatos a término simples y complejos (Quiroz & Burgos, en prensa).

0.3 Estructura de la tesis

La tesis está organizada en tres partes: en la primera parte, *Teoría*, se comienza con la introducción a este trabajo doctoral seguida del desarrollo del estado del arte y el marco teórico. En la segunda parte, *Metodología*, se detallan los aspectos metodológicos en relación con el corpus de trabajo y los criterios de compilación, selección de la muestra de trabajo y la validación de los datos. Por último, en la tercera parte, *Resultados*, se da cuenta de los resultados, la discusión y las conclusiones de trabajo. Además, al cierre de cada uno de los capítulos se presentará una síntesis con los detalles más relevantes de la información presentada o discutida a lo largo del segmento. Asimismo, en la estructura general de la tesis se incluye, al inicio, los *agradecimientos*, el *índice* general y los específicos (Figuras y Tablas), la *lista de siglas, acrónimos y abreviaturas* y, al final, se presentan los apartados correspondientes a las *referencias bibliográficas* y los *anexos*.

La primera parte está compuesta por tres capítulos. El primero, *Introducción*, contiene un esbozo general alrededor del campo de trabajo a partir de algunas reflexiones iniciales para delimitar el foco de acción y el alcance de la tesis. El segundo, *Estado del arte*, se presenta, de manera sucinta, una *mirada en conjunto* desde el inicio de los estudios disciplinares de la terminología hasta el presente. El objetivo principal de este capítulo es la integración de las distintas tradiciones teóricas y metodológicas en la construcción de una panorámica integradora en torno al *término* como unidad de estudio de la terminología. Para ello, se describe, brevemente, en primera instancia, cada una de las aportaciones de las principales escuelas de pensamiento terminológico, sus exponentes más representativos y sus bases de trabajo y, en segunda instancia, se presentan algunas investigaciones en el campo con miras a la construcción de esta *mirada común*. En el tercero, *Bases teóricas*, se aborda la definición de la terminología con el objetivo de contrastar y delimitar el campo de trabajo de la disciplina y su objeto de estudio. Luego, se presenta un recorrido a través de las distintas definiciones de término con el fin de caracterizar el tratamiento que desde las distintas tradiciones de pensamiento terminológico se le ha dado a este y, a partir de allí, presentar nuestra propuesta para la ampliación del concepto, vigente hoy en día, a partir de los hallazgos en esta investigación. Por último, se describen las bases conceptuales del protocolo para la validación de las unidades terminológicas simples y complejas (Quiroz & Burgos, en prensa).

La segunda parte está compuesta por el capítulo *Metodología*. La información desarrollada en este se dividió de la siguiente manera: a) el enfoque y el alcance metodológico de la investigación, b) los estadios de trabajo donde se indican los criterios empleados en la conformación del corpus especializado, los criterios definidos para la selección de las áreas de análisis, el tipo de textos a compilar y la organización de la información en las bases de datos, c) la descripción de las distintas etapas de validación de los candidatos a término y los programas computacionales empleados en las distintas etapas arriba descritas cuando fue necesario y, finalmente,

d) los aspectos estadísticos descriptivos para la caracterización de los datos analizados en la presente tesis.

La tercera parte está compuesta por los tres últimos capítulos de la tesis. El primero de ellos, *Resultados*, compila los hallazgos a partir de los datos recuperados del corpus de trabajo con relación a los términos simples, complejos y las demás unidades terminológicas al igual que se presentan los ejemplos que servirán de insumo para la discusión. En el segundo, *Discusión*, se retoman las hipótesis y objetivos de la tesis en torno a los resultados. Se realiza un análisis desde cada una de las áreas trabajadas y se plasman las similitudes y diferencias entre estas. Por último, en el capítulo *Conclusiones*, se sintetizan los aportes más significativos de la tesis en el área de trabajo al igual que se mencionan las principales dificultades y limitaciones en el desarrollo de la investigación para, en último lugar, proyectar las posibles líneas de trabajo futuro y las recomendaciones generales fruto de la presente tesis doctoral.

1. Terminología: una mirada en conjunto

"La historia es un incesante volver a empezar"
Tucídides

(Historiador ateniense, 460 a.C. – 396 a.C.).

No es tarea fácil compilar la historia de nuestra disciplina[4] en unas cuantas líneas. Sin embargo, es menester dar cuenta de la evolución de la terminología a lo largo del tiempo: su inicio en los principios fundacionales en Wüster (1931) y su desarrollo en los distintos postulados de cada una de las escuelas de pensamiento terminológico, cuyos aportes modelaron lo que hoy concebimos como nuestro objeto de estudio: el término[5]. La mirada, *una mirada en conjunto*, será integradora, es decir, un transitar por las distintas teorías y metodologías de trabajo, por investigaciones en el campo y sus diversas aportaciones en la construcción de esa puesta en común. Así, el principio rector del presente estado de la cuestión es, *grosso modo*, la reflexión de cada aportación no con el objetivo de discutir el impacto de cada una en comparación con las demás, sino de identificar cómo en la diversidad de acercamientos se logra ahondar en la observación de una realidad común: el conocimiento especializado. En este sentido, Cabré reafirma la necesidad de "encajar las piezas

4 Sobre la denominación de la terminología como disciplina, Cabré (2005) reconoce la importancia del aporte de Wüster al dotarla de este estatus y proporcionarle una teoría base de trabajo (Teoría General de la Terminología, TGT) (p. 3). No obstante, la autora prefiere referirse a la terminología como un "campo de conocimiento" para evitar la institucionalización inherente al uso del término "disciplina" desde una perspectiva social (p. 2). Para los objetivos del presente trabajo, nos referimos a la terminología como disciplina, en cuanto permite recuperar el alcance de su quehacer en el ámbito del conocimiento sin desconocer el trasfondo epistemológico que esto implica.

5 Para un acercamiento más amplio sobre el concepto de término, véase el Capítulo 02: El término: caracterización de la unidad de estudio.

del *puzzle*" para lograr entender que "todo lo que se ha dicho sobre la terminología es una contribución" (19 de noviembre 2020).

El punto de partida, sin duda, se concreta en los postulados de la Teoría General de la Terminología (TGT) propuesta por Eugene Wüster (1998) que marca, por un lado, la antesala del nacimiento de la escuela de Viena y el surgimiento paralelo de la escuela Rusa con Drezen y Lotte, de la escuela de Praga con Trubetzkoy y Drodz (con una fuerte influencia de los postulados saussureanos del lenguaje) y, por otro lado, el inicio de la concepción disciplinar de los estudios terminológicos, que ya en los siglos XVIII y XIX vislumbraba las primeras discusiones sobre la relación entre los conceptos y sus correspondientes formas y la notoria expansión de los términos en los ámbitos científicos (Sageder, 2010, p. 124).

En su trabajo, Wüster dotó a la naciente disciplina con las bases conceptuales y metodológicas necesarias para una sistematización del procesamiento y el trabajo con datos terminológicos de su objeto de estudio: el *término* y, en consecuencia, sembró las bases para la normalización de la terminología a través de distintos escenarios, entre ellos: a) la creación de asociaciones como el *International Information Centre for Terminology* (INFOTERM)[6] por la UNESCO en el año de 1971 y b) distintas normativas mediante la creación del comité técnico de terminología ISO/TC 37[7]. Además, con la llegada de los denominados *teóricos del giro terminológico* y en sintonía con los paradigmas en el estudio del lenguaje (desde el estructuralismo hasta el funcionalismo), se logró integrar una mirada desde la *Ciencia del Lenguaje* lo que permitió identificar y reflexionar en torno a los postulados de la TGT. Sin embargo, en línea con el propósito del presente capítulo, relacionamos los aportes de Wüster y el grupo de Viena en la consolidación de los orígenes[8] clásicos de la terminología

6 Para más información acerca de INFOTERM, véase: http://www.infoterm.info/.

7 Par más información acerca de la norma ISO/TC 37, véase: https://www.iso.org/committee/48104.html.

8 En sintonía con Cabré (1992) quien retoma la propuesta de Auger (1988), concordamos con los cuatro periodos en la evolución de la terminología moderna: los orígenes (1930 a 1960),

sin desconocer que las críticas a la TGT son, precisamente, el inicio de los posteriores enfoques teóricos y de las investigaciones derivadas en línea con estas nuevas propuestas de análisis.

La década de 1980 marca, por un lado, el inicio de la concepción de los manuales en terminología (Felber & Picht, 1984; Cabré, 1992; Torruella & Sager, 1993; Arntz & Picht, 1995; Fedor de Diego, 1995; Dubuc, 1999; Pavel & Nolet, 2002; Cardero, 2003; Cabré, Domènech & Estopà, 2018 y Faber & L'Homme, 2022) los cuales, por esencia, continúan su postura prescriptiva y normalizadora, en sintonía con los postulados wüsterianos y, por otro lado, la dinamización de las investigaciones en terminología desde perspectivas de corte lingüístico, cognitivo, social, cultural y computacional gracias a la proliferación de los campos de análisis en terminología junto a los avances de las distintas áreas convergentes. A continuación, se presenta un esbozo de los principales enfoques de trabajo terminológico posteriores a la TGT conforme al orden de aparición de estas nuevas propuestas de análisis, seguido de algunas consideraciones relevantes en torno a los investigadores en la disciplina y la diversidad de trabajos resultado de la aplicación y reflexión de los postulados en cada uno de los enfoques existentes. Por último, se describe el aporte latinoamericano en el desarrollo de la disciplina gracias a las particularidades del contexto del centro y sur de América.

1.1 Enfoques de trabajo terminológico

Como ya se mencionó, el punto de inflexión es, sin lugar a duda, la TGT, teoría bajo la cual se comenzó la consolidación de la terminología, a partir de la interacción entre sus postulados y las posteriores reflexiones teóricas y metodológicas en torno a estos. Sin dejar a un lado el aporte de las ya citadas escuelas Rusa y de Praga y otros pioneros

la estructuración (1960 a 1975), la eclosión (1975 a 1985) y la ampliación (desde 1985). Sin embargo, consideramos o un subperiodo de la *ampliación* o un quinto periodo *teorización y aplicación* marcado por la puesta en escena de los postulados teóricos y metodológicos en torno a la terminología en los trabajos de Cabré (1992, 1999), Fedor de Diego (1995), Temmerman (2000), Faber (2012), entre otros.

en el campo como Scholman o Holmstrom[9]. Así, para la TGT el centro de atención del trabajo terminológico se orienta hacia la asociación entre el *concepto* y su *denominación* de cara a un presupuesto final en la *normalización* y *estandarización* de los términos de especialidad y, por ende, aspectos relevantes atribuidos a estos como la *univocidad*, la *monosemia* y la *monorreferencialidad*. Así, el término bajo la mirada de la TGT se queda, necesariamente, en un plano bastante *restrictivo*, característica que sale a la luz gracias al surgimiento de nuevas perspectivas sociales y lingüísticas, al integrar en la discusión las regularidades de la comunicación profesional mediada por el lenguaje, característica que permite, asimismo, ahondar en la concepción natural del lenguaje de especialidad más allá del rótulo de lenguaje artificial y, en consecuencia, el inicio de los posteriores enfoques de trabajo terminológico en adelante detallados.

En primer lugar, la Socioterminología (ST) como fruto de la integración entre los principios de la sociolingüística y la terminología (Diallo, 2020, p. 25), le permitió a la última integrar el carácter social a su objeto de estudio, esto es, la posibilidad de identificar la dimensión social, temporal y cultural del término, además de reconocer su posición e integración al interior del lenguaje natural y, por ende, la desmitificación de la terminología como un lenguaje artificial y su categorización como un subconjunto de la lengua natural (Cabré, 1999b, p. 33). La ST logra afianzarse a partir de los trabajos de Guespin & Laroussi (1989), de la Universidad de Rouen, quienes logran materializar el estudio social de los términos como respuesta a la crítica de los postulados wüsterianos con relación a la no reflexión sobre el uso de la lengua en contextos reales de interacción. De esta manera, el interés de la ST, como lo expresa Gaudin, se centra en el análisis en torno al uso, aparición, circulación e implantación de los términos en escenarios de especialidad (2005, p. 81); aspectos que reafirman el cambio de paradigma en los estudios del lenguaje surgido en la

[9] Al respecto de los postulados de la TGT, Cabré presenta una interesante reseña en relación con las críticas más relevantes a esta teoría en su artículo denominado "Hacia una aproximación teórica de base comunicativa. Elementos para una teoría de la terminología: hacia un paradigma alternativo" (1999, pp. 69-92).

segunda mitad del siglo XX. Es importante mencionar, entre otros, los trabajos de Marcellesi & Gardin (1974), Gambier (1987), Auger (1994), Myking (2000) y Gentile (2007), como exponentes de la ST.

En segundo lugar, la Terminología Variacionista (TV), de origen canadiense y posterior a la Socioterminología de la escuela de Rouen, que adiciona a los principios socioterminológicos el componente pragmático, esto es, la identificación de los grupos sociales y los diferentes roles en la interacción. Para la TV, cuyo principal exponente es Auger (1974), los términos al igual que las unidades de la lengua natural se utilizan en función del contexto comunicativo, es decir, varían conforme parámetros temporales, sociales, culturales y según el grupo social o profesional (Cabré, 19 de noviembre 2020). En ese sentido, en la variación terminológica se contradice el principio clásico de la biunivocidad en función de la comunicación en los contextos de especialidad integrando el carácter de poliedricidad del objeto de estudio descrito por Cabré en su "teoría de las puertas", esto es, el análisis desde las perspectivas cognitivas, lingüísticas y socio-comunicativas (2000, p. 59). Otros autores de la TV que permitieron afianzar este enfoque de trabajo son: Faulstich (1998) y Freixa (2005).

En tercer lugar, la Teoría Comunicativa de la Terminología (TCT), del Grupo de Investigación en Léxico y Tecnología (iula*Term*)[10] de la Universidad Pompeu Fabra, propuesta por Cabré, presenta una visión integradora a partir de las distintas teorías lingüísticas, del conocimiento y de la comunicación en torno a los estudios del lenguaje. En este sentido, en la denominada *teoría de las puertas* (Cabré, 2000), las unidades de análisis son poliédricas, esto es, son susceptibles de ser analizadas desde tres perspectivas distintas, a saber: semiótico-lingüística, cognitiva y comunicativa. En la TCT, para el análisis de dichas unidades, siguiendo la metáfora propuesta por la autora, se entra por la *puerta* lingüística, a partir de producciones tanto escritas como orales en contextos de comunicación especializada. En este sentido, la TCT parte del texto especializado

[10] Para más información sobre el Grupo de Investigación en Léxico y Tecnología (iula*term*), véase: https://www.upf.edu/web/iulaterm/presentacio.

para la identificación de las unidades que adquieren su valor terminológico precisamente en el contexto de especialidad en el cual están inmersas. De esta manera, una teoría terminológica, de base lingüística como lo es la TCT, debe tener la capacidad de detectar, describir, discriminar y analizar las unidades de análisis al igual que establecer tanto las relaciones como las estructuras conceptuales con miras a describir la dinámica de la comunicación de especialidad y predecir el comportamiento de tales unidades (Cabré, 2000, pp. 10-15).

En cuarto lugar, la Terminología Textual (TT) introduce la noción de texto al integrar los principios de la lingüística textual y la terminología. De esta manera, La TT hace uso de la metodología de la lingüística de corpus con miras a detectar, observar y analizar los términos en el texto favoreciendo el desarrollo de herramientas para el Procesamiento de Lenguaje Natural (PLN) e incrementar la eficiencia al procesar grandes cantidades de datos (Cabré, 19 de noviembre 2020). Ahora bien, con relación a los principales enfoques de la TT se tiene como referencia, por un lado, el trabajo de Collet (1997, 2004), precursora de esta y, por otro lado, el de Ciapuscio (1998). El primero, un enfoque de corte francés orientado al estudio de los fenómenos de reducción discursiva en el uso de los términos en el plano sintagmático. El segundo, un análisis de corte alemán orientado a la simplificación (reformulación) de los términos (entendidos estos como componentes del lenguaje natural en consonancia con los postulados de la TCT ya mencionados) en función del género textual donde aparecen.

En quinto lugar, la Terminología Sociocognitiva (TSC) de Temmerman (2000) retoma las variables sociales ya estudiadas desde la ST y las aplica al concepto, lo que permite la inclusión del aspecto cognitivo teniendo como base tanto la psicología como la filosofía cognitiva (Cabré, 19 de noviembre 2020). Para Temmerman, el concepto es entendido en el sentido de unidades de comprensión (*unit of understanding*) más allá de la visión tradicional donde estos (los conceptos) son definidos como estructuras claramente delimitables (2000, pp. 223-224). En la TST, Temmerman redefine

los principios tradicionales de la terminología con lo cual propende sentar las bases para un nuevo paradigma definiendo tres ejes de análisis, a saber: a) de las estructuras prototípicas, b) de los modelos cognitivos y c) del eje diacrónico (2000, p. 229). La autora establece una clara distinción entre el objeto de trabajo terminológico del terminográfico, sin profundizar en ello, que decanta en la reformulación del alcance del segundo. Para ella, si el objeto del primero está orientado a las metodologías y principios del trabajo en terminología y el segundo se centra en la descripción sistemática de la información terminológica entonces, desde la TSC, se hace imperativo reformular el trabajo del terminógrafo para que sus análisis partan no de la noción clásica de concepto sino de aquella propuesta por la TSC: *unit of understanding* (Temmerman, 2000, pp. 230-231). Finalmente, la TSC retoma los modelos cognitivos de Lakoff (1987) al igual que la teoría de los marcos en Fillmore (1982) en la definición de sus categorías de análisis.

En sexto lugar, la Terminología Cultural (TC) de Diki-Kidiri (2008), desarrollada en el contexto multilingüe africano, permitió identificar cómo la terminología, y por extensión la lengua, al servicio de la sociedad contribuye en su desarrollo. De este modo, en el marco de un programa de desarrollo lingüístico, las cuestiones terminológicas van de la mano con otros aspectos relacionados a los avances científicos y tecnológicos (p. 16). En principio, la TC hace uso de la caracterización sociolingüística clásica africana lo que permite identificar postulados de orden social explorados también en los trabajos desde la ST. Sin embargo, aquí se aborda el fenómeno desde lo que Cabré en su momento denominó "una instrumentalización de la terminología" (19 de noviembre 2020).

Diki-Kidiri presenta una pirámide compuesta por tres niveles en los cuales se representa la realidad lingüística de África, a saber: en el nivel inferior se encuentran las *lenguas vernáculas* en las cuales se concentra el conocimiento ancestral del saber y saber-hacer herencia del pasado; mientras que en un nivel intermedio se localizan las *lenguas vehiculares* resultado de los intercambios sociales que

trascienden a las etnias y a la expresión de realidades modernas importadas a las culturas locales y, por último, en el nivel superior se ubican las lenguas oficiales, en esencia las europeas heredadas debido a los procesos de colonización y que desempeñan un rol dominante en todos los sectores de la vida nacional. Por consiguiente, la TC identifica cómo las *lenguas vehiculares* absorben tanto elementos de las *oficiales* como de las *vernáculas* y de allí la importancia del eje cultural en la producción y la recepción de los términos en función del contexto de especialidad, evitando al máximo el desarraigo y conservando los valores del territorio (2008, pp. 17-18).

En séptimo lugar, la Terminología Basada en Marcos (TBM), del Grupo de Investigación LexiCon[11] de la Universidad de Granada, propuesta por Faber (2012), marca un punto de inflexión con relación al enfoque, de corte lingüístico, hasta ese momento predominante. Así, junto a la propuesta de Temmerman, la TBM aborda los términos desde una perspectiva cognitiva con miras a, por un lado, estructurar los dominios de especialidad a partir de la adaptación de la propuesta de Fillmore sobre los *marcos semánticos* (1982) y, por otro lado, la creación de las representaciones no lingüísticas teniendo como base el *significado*, entendido este a la luz de la lingüística cognitiva y su acercamiento no lingüístico o enciclopédico del conocimiento. No obstante, a pesar de marcar un cambio significativo en la forma de abordar el estudio de los términos, la TBM comparte concepciones con la TCT, por ejemplo, el traspasar la discusión de la distinción entre *palabra* y *término* con miras a un estudio de la terminología a partir del comportamiento de las unidades de especialidad en los textos (Faber, 2012, pp. 22-23). De esta manera, siguiendo la metáfora cabretiana de las puertas (Cabré, 2000, p. 14), la TBM aborda los términos desde la perspectiva cognitiva mientras las TCT lo hace desde la puerta lingüística (Faber, 2012, p. 14).

En último lugar, en el marco del Simposio de la Red Iberoamericana de Terminología (RITerm) realizado en la ciudad de Manizales, Colombia en 2018, surgen dos enfoques nuevos de investigación

[11] Para más información sobre el Grupo de Investigación LexiCon, véase: http://lexicon.ugr.es/.

terminológica, aún en observación, fruto de los trabajos del Grupo de Investigación en Traducción y Nuevas Tecnologías (TNT)[12] de Medellín y del Centro de Investigación Terminológica (CITERM)[13] de Manizales, los cuales fueron denominados, respectivamente, como *Terminología Social* (TS) y *Terminología Experimental* (TE) (Cabré, 19 de noviembre 2020). En primera instancia, la TS, de base lingüística, intenta describir cómo a través del discurso de especialidad se logra ideológicamente estandarizar la información en contextos principalmente políticos y sociales, es decir, cómo en el uso del lenguaje se construyen definiciones ideológicas sobre problemas de orden social, ya sea de forma explícita o implícita (Chiquito & Quiroz, 2017, p. 7). Para ello, el grupo TNT, en el marco del proyecto *Poverty, Language and Media* (POLAME)[14], analizó lingüísticamente las estrategias sobre cómo el discurso de la pobreza se naturaliza a través de los medios de comunicación escrita más importantes en Argentina, Brasil, Colombia y México (países de los cuales se compiló el corpus de trabajo[15]), los cuales definen las agendas de debate público sobre el tema en cuestión (Chiquito & Quiroz, 2017, p. 9). En segunda instancia, la TE, con foco en el paradigma cognitivo en cuanto busca identificar la manera como en el cerebro se procesa el discurso de especialidad, presentándose a sí misma como la posibilidad de un nuevo enfoque de análisis terminológico y como complemento de lo ya trabajado en la TCT (Gómez et al., 2019, p. 21).

12 Para más información sobre el Grupo de Investigación en Traducción y Nuevas Tecnologías (TNT), véase: https://www.udea.edu.co/wps/portal/udea/web/inicio/investigacion/grupos-investigacion/humanidades/tnt.

13 Para más información sobre el Centro de Investigación Terminológica (CITERM), véase: https://archivo.autonoma.edu.co/investigacion/grupos-de-investigacion/citerm.

14 El proyecto POLAME, bajo la coordinación de la Prof. Dra. Ana Beatriz Chiquito de la Universidad de Bergen, Noruega, analizó la forma cómo el discurso sobre la pobreza se difunde en los medios de comunicación escrita más importantes en Argentina, Brasil, Colombia y México. Para más información del proyecto POLAME, véase: https://www.uib.no/en/project/polame.

15 Para consultar el corpus POLAME, véase: https://www.uib.no/en/project/polame/88208/polame-korpus-login y/o http://grupotnt.udea.edu.co/polame-corpus/.

1.2 Investigaciones en terminología

En la actualidad, fruto de las bases teóricas y metodológicas establecidas gracias a los diferentes enfoques en el trabajo terminológico descritos en la sección anterior (§ 1.1), se cuenta con una gran variedad de trabajos de investigación, algunos de ellos de tesis doctorales o de maestría, enfocados en la descripción y análisis de diversos fenómenos de índole terminológico que, a su vez, permitieron la creación y/o consolidación de diversos grupos de investigación a nivel internacional. Además, en la presente tesis, se hará un énfasis especial en las principales escuelas terminológicas latinoamericanas (§ 1.2.1) con el espíritu de resaltar el desarrollo de la disciplina en el contexto americano. En adelante, algunas consideraciones alrededor de estos aspectos con miras a complementar nuestra *mirada en conjunto*.

Gracias a las nuevas tecnologías de la información, ha sido posible tener un acceso más rápido a grandes cantidades de material académico acompañado de una serie de datos estadísticos que permiten a la comunidad académica disponer de información relevante sobre el alcance y el impacto de los trabajos publicados y, por ende, de los autores representativos. Para nuestro caso en particular, se realizaron consultas en el portal de Google Académico (sin desconocer la existencia de otros recursos como BITRA[16]) teniendo *terminology* o *terminología* como palabras clave, según las particularidades de cada investigador, esto es, se tomó la información del perfil con la mayor cantidad de citaciones. Inicialmente, el portal de Google Académico permitió realizar una identificación, en orden descendente, de los investigadores más citados en el área y, por ende, los más representativos y con mayor visibilidad "en la construcción de las "áreas de investigación"" (Milard, 2010, p. 70). Asimismo, diversas investigaciones se centran en la importancia de la bibliometría en el impacto de las publicaciones, tanto de las revistas como de los autores (Torres-Salinas, Ruiz-Pérez & Delgado-López-Cázar, 2009; Silva Ayçaguer, 2012; Orduña-Malea, 2016).

[16] Para más información sobre estos portales de consulta bibliográfica consúltese, respectivamente: https://scholar.google.com/intl/es/scholar/about.html y https://dti.ua.es/es/bitra/introduccion.html.

En la Tabla 1, se compila la información sobre los principales investigadores en terminología y su índice de citaciones. Se realizó una distinción entre las citas totales y aquellas relacionadas directamente con trabajos e investigaciones en terminología con el propósito de identificar los factores arriba mencionados.

Tabla 1. Google Académico: Investigadores en terminología. Elaboración propia[17].

Autor	Filiación institucional	Citaciones totales	Citaciones en terminología
María Teresa Cabré	Universidad Pompeu Fabra	13160	13025
Pamela Faber Benítez	Universidad de Granada	5152	4979
Marie-Claude L'Homme	Universidad de Montreal	2927	2858
Kyo Kaguera	Universidad de Tokio	2761	2510
Gerhard Budin	Universidad de Viena	2378	2056
Rosa Estopà Bagot	Universidad Pompeu Fabra	1780	1737
Jorge Vivaldi Palatresi	Universidad Pompeu Fabra	1597	1498
Judit Freixa	Universidad Pompeu Fabra	1585	1000
Pilar León Araúz	Universidad de Granada	1440	1440
Maribel Tercedor Sánchez	Universidad de Granada	1052	819
Margarita Correia	Universidad de Lisboa	1011	196
Mercè Lorente Casafont	Universidad Pompeu Fabra	877	852
Clara Inés López Rodríguez	Universidad de Granada	835	682
Silvia Montero Martínez	Universidad de Granada	809	667

[17] La información sobre las métricas de citación de los principales autores fue recuperada en el perfil de *terminology, terminología* y *terminología* del Google Académico. Para más información, consúltese: https://scholar.google.com/citations?view_op=search_authors&hl=es&mauthors=label:terminology, https://scholar.google.com/citations?hl=es&view_op=search_authors&mauthors=label:terminología o https://scholar.google.com/citations?hl=es&view_op=search_authors&mauthors=label:terminologia.

Autor	Filiación institucional	Citaciones totales	Citaciones en terminología
Cleci Bevilacqua	Universidad Federal de Rio de Janeiro	715	177
Gladis de Barcellos Almeida	Universidad Federal de São Carlos	578	438
Rogelio Nazar	Universidad de Valparaíso	511	477
Chelo Vargas Sierra	Universidad de Alicante	452	320
Elsabe Taljard	Universidad de Pretoria	416	275
Miriam Buendía Castro	Universidad de Granada	400	332
Sabela Fernández Silva	Universidad de Valparaíso	376	372
Antonio San Martín	Universidad de Quebec en Trois-Rivières	240	240
Gabriel Quiroz Herrera	Universidad de Antioquia	233	138

Al observar los datos en la Tabla 1, y teniendo en cuenta las etiquetas de búsqueda utilizadas, es necesario mencionar que dentro de los datos compilados no se incluyeron, por un lado, los investigadores recuperados cuyo ámbito de investigación no es la terminología, tal es el caso de Hongfang Liu del departamento de inteligencia artificial e informática de la Mayo Clinic, Johan Kelly de la Universidad de Harvard o Graeme Miller de la Universidad de Sídney y, por otro lado, los investigadores enfocados principalmente en la lexicografía como Henning Bergenholts o Sven Tarp de la Universidad de Aarhus. Asimismo, podría afirmarse que la tradición española en terminología se consolida, en la actualidad, no solo como una de las más influyentes sino también como la que ha marcado ruta de trabajo en otras latitudes como se evidencia, por ejemplo, en las investigaciones de Judit Freixa, Gladis Maria de Barcellos Almeida, Rogelio Nazar, Sabela Fernández Silva o Gabriel Quiroz Herrera, cercanos a la TCT de Cabré, y Marie-Claude L'Homme o Antonio San Martín, cercanos a la TMB de Faber.

En este sentido, también podría establecerse estas dos posturas de trabajo terminológico como los *paradigmas* predominantes en la actualidad investigativa en el área. Otro dato que llama la atención tiene

que ver con que las citaciones pertenecientes a la Universidad Pompeu Fabra corresponden al 49 % de los datos compilados, mientras que el 24 % son relativos a la Universidad de Granada, quedando un 27 % repartido en las demás Universidades, a saber: de Montreal, de Tokio, de Viena, Lisboa, Federal de Rio de Janeiro, Federal de São Carlos, Pontifica Católica Universidad de Valparaíso, de Alicante, de Pretoria, de Quebec en Trois-Rivières y de Antioquia. Esto es interesante en cuanto ratifica, por un lado, la preponderancia, en la actualidad, de la investigación en España con relación al resto del mundo y, por otro lado, el posicionamiento del Instituto de Lingüística Aplicada (IULA) de la Universidad Pompeu Fabra, como uno de los centros de formación terminológica más importantes en las últimas décadas.

En cuanto a las investigaciones, las aportaciones han dado paso a debates interesantes con relación a los aspectos teóricos y metodológicos. Así, al hablar del primer elemento, surge la pregunta acerca de si realmente podemos hablar de postulados teóricos, en parte por la ausencia de una definición precisa del objeto de estudio en varios de los enfoques terminológicos. Este es, quizás, el mayor desafío que enfrenta nuestra disciplina de cara a una *mirada integradora del término*. Con relación al segundo elemento, se ha logrado un buen avance en este aspecto a partir, por ejemplo, del aprovechamiento de las herramientas provenientes de las nuevas tecnologías al servicio de la investigación lingüística y, por extensión, de la terminología. Sin embargo, esta integración también ha llevado a debates recientes sobre la necesidad de reenfocar los análisis a partir de puntualizaciones lingüísticas en balance con aspectos netamente estadísticos que en la actualidad han desvirtuado, en muchas ocasiones, nuestras investigaciones (Larsson, Egbert & Biber, 2022).

1.2.1 Terminología en Latinoamérica

Con aras de completar nuestra *mirada en conjunto*, es necesario hacer mención de la terminología en nuestro continente y presentar, hasta ahora, el desarrollo de la disciplina en el contexto americano que, en sus inicios, propendió por una integración con España, con la

creación de las primeras asociaciones, como la *Unión Internacional Iberoamericana de Bibliografía y Tecnología Científicas*, en la década de 1920 y, posteriormente, con el desarrollo investigativo fruto de la formación recibida desde y en varias de las escuelas europeas arriba descritas. Así, durante el siglo XX (en especial la segunda mitad) y hasta la actualidad, surgieron y se consolidaron importantes tradiciones terminológicas en distintos países del continente, entre ellos se destacan, en especial, los aportes de los grupos de México, Cuba, Colombia, Argentina y Venezuela; este último testigo de la creación oficial de RITerm, en la Universidad Simón Bolívar, con la firma del tratado de Caracas en 1988, marcando el inicio de la integración entre la terminología española y la iberoamericana (de Irazazábal, 2010, pp. 13-15). A continuación, se presentan algunas de las aportaciones latinoamericanas más relevantes a la construcción de la terminología, desde el centro hacia el sur del continente.

En primer lugar, la Escuela Mexicana que nace como respuesta a la necesidad del reconocimiento de su variedad dialectal respecto de la influencia norteamericana, de corte anglo y franco parlante (Lara, 2009, p. 37). En este sentido, la década de 1970 marcó el inicio de las investigaciones sobre el español mexicano gracias a los trabajos de Luis Fernando Lara. Su investigación sobre la relación entre la *palabra* y el *referente* exacto en vocablos pertenecientes a las ciencias o las técnicas, recuperados del Corpus del Español Mexicano Contemporáneo (CEMC), sembraron las bases de la metodología de trabajo terminológico en el contexto mexicano (Pozzi, 2009, p. 40), que decantaron en la publicación del Diccionario del Español de México (DEM) (Lara, 2022) que, en la actualidad, ya cuenta con su segunda edición[18]. Los trabajos en terminología, varios de ellos de investigaciones a nivel posgradual[19], son la evidencia en la evolución e interés en la disciplina. Al respecto, también se destacan los aportes

[18] Para más información sobre el Diccionario del Español de México (DEM), consúltese: https://dem.colmex.mx/Inicio.

[19] Son importantes las investigaciones de: Bogomilova, 2010 y Ritcher, 2010; a nivel de doctorado y maestría, respectivamente; por ser pioneras en trabajos con enfoque terminológico en el contexto mexicano. Para más información sobre otras investigaciones en la disciplina, consúltese: https://repositorio.colmex.mx/?locale=es o https://repositorio.unam.mx/.

de: Maria Pozzi (1997, 1999, 2003), Ana Maria Cardero (2004), Gerardo Sierra (2022), entre otros; en la consolidación de la disciplina, no solo desde el ámbito investigativo sino también en el avance en la formación terminológica en el contexto mexicano.

En segundo lugar, la Escuela Cubana de terminología cuyos orígenes se remontan a las investigaciones de Felipe Poey y Aloy a partir de un interés normalizador desde las ciencias naturales. Entre sus obras más representativas se encuentra las *Memorias sobre la historia natural de la isla Cuba* publicada, en dos volúmenes, entre los años 1851 y 1858. Además de Poey y Aloy, en la época clásica de la terminología cubana, sobresale la obra de Néstor Ponce de León (1883), un *Diccionario tecnológico inglés-español y español-inglés*. La segunda mitad del siglo XX trajo consigo un renovado interés por la normalización terminológica a partir de principios lingüísticos en un trabajo liderado desde organismos de orden estatal como el Ministerio de Agricultura y asociaciones como el Centro de Informática Aplicada a la Normalización (CINAN). La última década del siglo anterior estableció la "época dorada" de la terminología cubana con la creación del Centro de Traducciones y Terminología Especializada (CTTE) cuyo objetivo principal se enmarcó en la potenciación no solo de la comercialización de la terminología especializada sino también de la investigación en distintas esferas de interés nacional, es decir, la ciencia, la técnica y la economía. Sin embargo, la actualidad en la terminología cubana se ha reducido, posterior a la desaparición de la CTTE en 1998, a algunos trabajos de orden académico de limitado impacto en la sociedad cubana (Alpízar, 2009, p. 18).

En tercer lugar, la Escuela Colombiana de terminología fundada desde los finales de la década de 1980 gracias al significativo aporte de las investigadoras Cecilia Plested (1987, 2010), Plested et. al (2015) y Berta Nelly Cardona (1997), consideradas como las pioneras de la terminología en Colombia y cuyos aportes no solo se ven reflejados en la consolidación inicial de las posturas teóricas y metodológicas sino también en los procesos de formalización de la disciplina a través, por ejemplo, de la creación de redes de cooperación nacional

como la Red Colombiana de Terminología (Colterm) (Cardona, 1999). Con el surgimiento de los grupos de investigación Termiazúcar de la Universidad del Valle, el Grupo de Investigación en Terminología y Traducción (GITT) de la Universidad de Antioquia y el Centro de Investigación Terminológica (CIT) de la Universidad Autónoma de Manizales, se reafirmó el interés naciente por el trabajo terminológico en el contexto colombiano, evidenciado en la creación de importantes diccionarios y glosarios en ámbitos como la medicina, el derecho, la sociología, entre otros. Asimismo, varios de los investigadores, formados en los albores de la terminología colombiana, bajo la mirada de la Escuela clásica, completaron sus estudios posgraduales en universidades europeas como las de Barcelona, Pompeu Fabra, la Escuela de Negocios de Copenhague, The Norwegian School of Economics (NHH), entre otras.

En la actualidad, y como se mencionó en el apartado final de los enfoques de trabajo terminológico (§ 1.1), Colombia cuenta con dos enfoques nuevos en terminología, que se unen a los clásicos de corte europeo o norteamericano. En este sentido, el avance del trabajo terminológico en el contexto colombiano es reflejo de la dinámica actual en nuestro continente, el cual ha logrado transcender la etapa inicial de una replicación de los postulados clásicos y se ha permitido proponer y aportar a la disciplina elementos de orden teórico y metodológico, en lo que podríamos denominar la época del desarrollo terminológico no solo en Colombia sino en el contexto latinoamericano. En este sentido, a modo de ilustración, el grupo de investigación TNT de la Universidad de Antioquia, creado en el año 2008, cuenta en la actualidad con una serie de recursos[20] al servicio de la comunidad académica y general, a saber: corpus textuales en diferentes campos del saber, etiquetadores sintácticos (TNTagger), extractores terminológicos (BC-Term), validadores terminológicos (TNTvalidate), *software* de pronunciación en francés para hispanohablantes (Pronunciación francesa), entre otros. Además,

[20] Para más información sobre el banco de recursos propiedad del Grupo de Investigación TNT, consúltese: https://www.udea.edu.co/wps/portal/udea/web/inicio/investigacion/grupos-investigacion/humanidades/tnt.

unidos a los investigadores colombianos ya mencionados, pueden citarse los aportes de Emma Rodríguez (2011), Jhon Jairo Giraldo (2008), Gloria Montoya & Cecilia Zapata (2005), entre otros tantos investigadores que han contribuido con sus trabajos a la consolidación y fortalecimiento de la terminología colombiana. Al respecto, en el año 2002 fue publicado un artículo sobre algunos aspectos históricos de la terminología en Colombia en un número especial de la revista *Terminómetro* sobre la terminología en los países andinos (Quiroz, 2002, pp. 3-6).

En cuarto lugar, la Escuela Venezolana, fundada a principios de la década de 1980, considerada como la pionera en el trabajo terminológico hispanoamericano. En el desarrollo de su trabajo, esta escuela se caracterizó, en esencia, por una amplia formación en terminología y una importante producción terminológica evidenciada, por un lado, en la creación de glosarios y vocabularios en diversidad de áreas del saber y, por otro lado, en la materialización del primer y, hasta ahora, único manual de terminología en español, escrito en América, autoría de Alicia Fedor de Diego hacia 1995[21]. En su haber, la terminología venezolana se ha desarrollado en varias universidades, entre ellas: la Universidad Central de Venezuela, la Universidad Metropolitana, la Universidad de los Andes, la Universidad Simón Bolívar y la Universidad Católica Andrés Bello; al igual que cuenta con importantes organizaciones como la Asociación Venezolana de Terminología (VenTerm) o el Centro de Investigaciones y Apoyo Tecnológico de Petróleos de Venezuela (PDVSA Intevep). Entre otros investigadores venezolanos en la disciplina, podemos mencionar los trabajos de Tatiana Lugo (1982), Melva Márquez (2006), Luz Marina Rivas junto a Luz Mayorga y Diana Ravelo (2019).

En quinto lugar, la terminología en el contexto de las escuelas Peruana y Ecuatoriana. Por un lado, la tradición peruana tiene su aparición en la década de 1990 gracias a la formación en traducción

[21] En esta misma línea, el Grupo de Investigación TNT se encuentra en la etapa final de edición de un nuevo manual en terminología el cual será publicado por la editorial Routledge bajo el título *Terminología del español: el término.*

recibida por un grupo de profesoras, adscritas a la Universidad Femenina del Sagrado Corazón (Unifé), en la Universidad Católica de Chile. El impacto inicial de la formación terminológica en Perú se vio reflejado, principalmente, en la creación a nivel curricular de un curso específico en el área en los programas académicos en traducción. Es importante resaltar el trabajo de Ivana Suito (1998) y Rosa Luna (2012). Por otro lado, la terminología en Ecuador, de la mano de los trabajos de Pilar García-Donoso responde a una realidad, de por sí común en nuestro contexto, de reconocimiento de la importancia de la traducción y la terminología en el futuro del país (2002, p. 43). El avance de la disciplina en el Ecuador se ha visto vinculado, directamente, con los esfuerzos y disposición de la Pontificia Universidad Católica del Ecuador.

En sexto lugar, la terminología en el Cono Sur, esto es, las escuelas de Argentina, Chile y Uruguay. En primera instancia, la terminología en el contexto argentino trajo consigo aportaciones desde la lingüística del texto con Ciapuscio (2000), la sintagmación en los trabajos de Kuguel (2006) o Adelstein (2007) y la terminología en el ámbito de la traducción con Perassi & Ferreira (2016) o Baduy, Garda & Perassi (2016). De igual manera, el aporte de Daniel Prado a la consolidación de la terminología en el contexto iberoamericano es de vital reconocimiento, siendo uno de los pioneros en el establecimiento y consolidación de la terminología en nuestro continente y, en consecuencia, su reconocimiento como "el pilar fundamental de la terminología iberoamericana" (Cabré, 2012, p. 82). En segunda instancia, el aporte chileno a la terminología está marcado, inicialmente, por las investigaciones realizadas desde la Pontificia Universidad Católica de Chile por la Profesora Ileana Cabrera quien, con sus aportes desde lo teórico y metodológico, logró establecer el ámbito de la terminología en el contexto de Chile, desde la década de 1980 (Diéguez, 2014, p. 65). Además, los trabajos de Diéguez (2008) y la publicación del *Diccionario de uso del español de Chile DUECh* por la Academia Chilena de la Lengua (2001), entre otros, son reflejo del avance de la terminología en Chile. Por último, en tercera instancia, el aporte de Uruguay a la terminología gracias, por

un lado, a las investigaciones realizadas por el Grupo en Terminología y Organización del Conocimiento (GTERM) bajo la coordinación de Mario Barité, en aspectos teórico-metodológicos (2017) y, por otro lado, a la publicación del *Diccionario del español de Uruguay* por la Academia Nacional de Letras de Uruguay en el año 2017[22].

En último lugar, la Escuela Brasileña de terminología cuyo origen se remonta a principios del siglo XX como respuesta a un proceso de normalización técnica que originó, entre otros aspectos, la publicación de documentos como el *Manual de Resistência de Materiais* en 1905 o la conformación de organizaciones como la *Associação Brasileira de Normas Técnicas* en 1936. Además, la actividad normalizadora en los distintos comités en Brasil contó, principalmente, con profesionales en las áreas específicas careciendo de la presencia de lingüistas o terminólogos, situación que para la segunda mitad del siglo anterior se vio reversada por el aumento en las investigaciones en terminología en varias de las universidades que ya para el 2009 contaban con cerca de 19 grupos de investigación en la disciplina (de Barcellos, 2008). En la actualidad, Brasil cuenta con importantes investigadores que han permitido ampliar las discusiones en terminología desde elementos teóricos y metodológicos en ámbitos científico-técnicos unido a la integración con la lingüística computacional y de corpus en varias de sus investigaciones. Entre estos podemos citar los trabajos de: Krieger (2000), Bevilacqua (2004), Faulstich (2006), Barbosa (2006) e Isquerdo & Alves (2007).

1.3 Recapitulación

Hasta aquí, se ha presentado un recorrido por las aportaciones de los distintos enfoques de trabajo terminológico, sus principales exponentes y las bases con las cuales se configura cada una de las

[22] Es importante resaltar que, con relación a la publicación de los diccionarios de español en Hispanoamérica, Colombia cuenta con el Diccionario de Colombianismos publicado por el Instituto Caro y Cuervo en 2018. Para más información, consúltese: https://colombianismos.caroycuervo.gov.co/.

visiones de trabajo en la construcción de esta *mirada en conjunto*. En esencia, la investigación en terminología, posterior a Wüster y su grupo, se ha visto dividida en dos paradigmas intrínsecamente relacionados, producto de la dinámica en el estudio del lenguaje, lo que ratifica lo ya establecido por Cabré acerca de las particularidades de los lenguajes de especialidad heredadas de características de la lengua (1999, p. 56). Esto es, la dualidad entre lo lingüístico y lo cognitivo materializada en los postulados de la TCT de Cabré y la TBM de Faber (que a su vez retoma los expuesto por Temmerman en la TSC), respectivamente. Sobre esto, se podría afirmar que la TCT se configura como el paradigma rector en la mayor parte de las investigaciones en el campo desde su aparición a finales de la década del 90. Es de resaltar que la *teoría de las puertas* busca presentar un acercamiento integral del estudio de las unidades terminológicas que permitiese observar el fenómeno en su carácter poliédrico (Cabré, 2000, p. 10). Sin embargo, un estudio de tal magnitud implicaría un nuevo paradigma de investigación, en el cual se logre hacer uso de todas *las puertas* en función del análisis de la unidad de estudio, lo cual sobrepasaría los alcances de un trabajo claramente limitado, incluso, esta situación podría argumentarse desde la misma decisión en la TCT de enfocar los análisis solo desde la *puerta lingüística*.

La terminología en el contexto latinoamericano ha sido el resultado de un proceso no solo de normalización en las diferentes áreas del conocimiento sino también de una apuesta por la reivindicación de nuestros orígenes en cada una de las diversidades lingüísticas que conforman el mapa del lenguaje en nuestro continente. A nivel teórico y metodológico se ha tenido un desarrollo significativo al lograr integrar, al panorama europeo y norteamericano, importantes avances en ámbitos terminográficos y normativos, como se evidencia en una gran variedad de recursos como glosarios, corpus y *softwares*, pensados y ejecutados desde cada uno de nuestros contextos particulares y como complemento a las miradas paradigmáticas de mitades del siglo XX. Finalmente, esta *mirada en conjunto* propende por una observación a través de la historia reciente con el reconocimiento a la herencia terminológica que, hoy en día, continúa su recorrido con miras

a seguir caracterizando y descubriendo las particularidades de una puesta en común de los saberes humanos existentes y aquellos que surgirán a la par de la evolución de la humanidad.

En el Capítulo 02 se presentarán los aspectos teóricos alrededor del concepto de *término* y el tratamiento que se ha dado a este en las diferentes épocas de la terminología. En este sentido, la discusión se centrará en sus características lingüísticas a partir de los distintos enfoques de trabajo y las distintas normativas. Asimismo, se abordarán investigaciones en las cuales se ahonda sobre el estatus terminológico de las unidades lexicales más allá del sustantivo como categoría prototípica de análisis y, en consecuencia, se presentará una propuesta en la ampliación de la actual concepción de término con miras a una operativización del concepto que permita, entre otros, la automatización en la recuperación de la información terminológica en distintos campos del conocimiento.

2. El término: caracterización de la unidad de estudio

Respecto a mí, mi querido Sócrates,
después de muchas discusiones con nuestro amigo y con muchos otros,
no puedo creer que los nombres tengan otra propiedad,
que la que deben a la convención y consentimiento de los hombres.

Platón.
(El Crátilo o de la exactitud de los nombres).

A un poco más de 91 años desde la publicación de la tesis doctoral de Wüster (1931) la cual marcó, sin lugar a duda, un antes y un después en el entendimiento de la terminología como actividad humana, nuestra disciplina aún discurre entre los límites de las definiciones conceptuales en la delimitación y entendimiento mutuo: ¿qué se entiende por terminología? ¿Cuál es su unidad de estudio? ¿Cómo se concibe el término? ¿Cómo entender la relación entre palabra y término? Cuestionamientos que no solo reflejan el desafío vigente, sino que también permiten ahondar en las reflexiones a través de las cuales se ha consolidado la terminología como un ámbito inherente a la naturaleza humana. Lejos estamos de una mirada estática de sus fenómenos, en gran medida por la variedad de enfoques de análisis y la diversidad de pensamientos terminológicos, lo que ha permitido *una mirada en conjunto*[23] de una realidad común: la materialización del conocimiento de especialidad a través de las bondades del lenguaje.

El presente marco teórico tiene como objetivo un recorrido que permita la delimitación, desde lo lingüístico, del término como unidad

[23] Sobre esta *mirada en común* el Capítulo 01: *Terminología: una mirada en conjunto* presenta un recorrido integrador de los diferentes enfoques teóricos y metodológicos hasta ahora propuestos en torno a los debates terminológicos.

de estudio y análisis de la terminología y, en consecuencia, una actualización en la definición conceptual imperante en la actualidad. Para ello, se presentará, inicialmente, una definición de la terminología enmarcada en la directa relación de esta con la lingüística (ciencia en la cual asienta sus principios conceptuales y metodológicos) para, posteriormente, analizar las características, hasta ahora descritas, del término desde los postulados teóricos y normativos vigentes en el área. Al respecto, nuestra postura de análisis propende no solo por una mirada restrictiva del sustantivo como unidad prototípica, sino que integra las demás unidades léxicas como portadoras del conocimiento de especialidad. Para ello, se ha implementado en nuestro análisis un protocolo de validación tanto para las unidades simples como para las complejas (Quiroz & Burgos, en prensa) basado en criterios lingüísticos que, a su vez, permitirían una operativización en la identificación de los términos a partir de los hallazgos en el corpus compilado para los propósitos de la presente tesis. Por último, se integra a la discusión el concepto de variación terminológica entendido a la luz de las características prototípicas del término en función del área de especialidad y las regularidades en la expresión de los conceptos propios de las realidades epistémicas de cada campo del conocimiento.

2.1 Terminología: una disciplina de la lingüística

Al hablar de terminología es necesario ahondar en su relación con el ser humano, esto es, entender cómo esta se configura como un elemento intrínseco a la naturaleza humana, en cuanto su campo de acción va ligado a la necesidad de denominar las realidades existentes y, en nuestro caso particular, aquellas pertenecientes a los dominios de especialidad. En este sentido, el siglo XX se convierte en el punto de inflexión entre lo que podemos considerar, por un lado, como una primera concepción de la terminología basada en miradas filosóficas y/o científicas que buscaban, a partir de la asociación entre los conceptos y las denominaciones, un entendimiento y posterior normalización del lenguaje de especialidad y, por otro lado, como

una concepción disciplinar gracias a los aportes de las distintas escuelas terminológicas a nivel mundial. Ahora, en aras de justificar esta noción de la terminología como una disciplina independiente, partiremos de reflexiones de orden filosófico puesto que es allí, en la filosofía, donde se originan los entramados epistemológicos sobre la naturaleza de las ciencias, las disciplinas y las artes, sus alcances y las relaciones o diferencias que se establecen entre ellas.

El punto clave se centra en establecer cómo la terminología se constituye en una disciplina de la lingüística, en la medida que permite delimitar y regular su discurso el cual, a su vez, está ligado al objeto de su quehacer epistémico que no es otro que el lenguaje. Así, el concepto de ciencia, de origen latino, es el resultado de la evolución en la manera de concebir el conocimiento por parte del hombre y su afán de legitimar la verdad de los objetos por él observados. En un primer momento, resultado de la cristianización del mundo conocido, nace un paradigma teocéntrico delimitado por la introducción de la *inspiración divina* bajo la cual se validó el saber hasta que, en un segundo momento, con la llegada de la modernidad se introduce el *método científico* y, por ende, un nuevo paradigma antropocéntrico regido por la observación hecha por el hombre de la realidad circundante, de cara a la creación, comprobación y conclusión de hipótesis. Bajo esta mirada, la lingüística se consolida como la *ciencia del lenguaje*, en gran medida, por los postulados saussureanos (de Saussure et al., 1919:1967), en los cuales se le asigna un objeto de estudio que permite el nacimiento de diversas miradas, constituyendo las disciplinas del lenguaje. De esta manera, siguiendo la lógica foucaultiana (1970, 1993, 2013), a partir de la relación entre las prácticas sociales, entendidas como las actividades del día a día, que originan prácticas discursivas y, por ende, saberes, es factible afirmar que la terminología se encarga de legitimar el discurso de especialidad producto de las interacciones humanas en contextos donde se dialoga en torno al conocimiento humano.

En esta línea, nos encontramos frente a una nueva noción en el entendimiento de la clasificación de las ciencias que, tradicionalmente,

han sido agrupadas en dos categorías, a saber, *ciencias naturales* y *ciencias humanas*, para dar paso a una tercera concepción, *ciencias del espíritu* (las cuales a futuro serán denominadas como ciencias humanas y sociales), en la cual se intenta comprender la cultura y el accionar humano siendo Dilthey, en el siglo XIX, quien comienza el trasegar en busca de encontrar un método propio de análisis, en contraste con el establecido en las ciencias naturales, con el objetivo de "comprender la vida por sí misma" (1883:1944, p. 20). Al respecto, Foucault hace referencia a esta distinción al indicar que en "el eje conciencia-conocimiento-ciencia" el interés se centra en explicar desde el método científico, en cambio las ciencias sociales, en un primer momento llamada *ciencias del espíritu*, buscan comprender desde las prácticas sociales en el eje "práctica discursiva-saber-ciencia" (1970, p. 307). De esta manera, la lingüística se ve inmersa bajo esta nueva mirada adquiriendo un método fenomenológico y hermenéutico en el análisis del lenguaje. Así, reafirmamos que la terminología se configura como una disciplina autónoma de la lingüística, a pesar de que aún persiste cierto disenso en la lingüística al momento de considerarla como una disciplina independiente del lenguaje y, en ocasiones, ubicarla bajo la sombra de la lexicología.

La terminología, ya sea entendida como ciencia o disciplina, ha sido el fruto de las bases teórico-metodológicas de las diferentes escuelas y posturas en el trabajo terminológico posteriores a los postulados wüsterianos, que se consolidan como el punto de partida en el reconocimiento de esta como disciplina. Wüster concibe, entre otras características, la terminología como la "lexicología del lenguaje especializado" (1998, p. 162). En este sentido, en su trabajo, Wüster reconoce la terminología como una disciplina orientada a una finalidad práctica (metodológica) en función de la eliminación del carácter ambiguo en la comunicación y en el uso de los conceptos en entornos especializados. Posteriormente, orientó su interés hacía la reflexión teórica de su propuesta metodológica desarrollada, en mayor medida, a partir de los trabajos posteriores a su tesis doctoral (Wüster, 1931) lo que, en consecuencia, sentó las bases para esta concepción de la terminología. Asimismo, no desconocemos las posturas que entienden

la terminología como práctica, principalmente, en los trabajos de Sager al establecer su rechazo al estatus de la terminología como disciplina independiente y enmarcarla como "una serie de prácticas que han evolucionado en torno a la creación de términos, su recopilación y explicación y, finalmente, su presentación en forma impresa o por medios electrónicos" (1993, p. 19). En esta misma línea, autores como Dubuc (1980) concuerdan con la postura de Sager al entender la terminología como una práctica. Sin embargo, en sintonía con la presente investigación, en la cual se concibe la terminología como una disciplina de la lingüística en cuanto contribuye, a partir de una propuesta teórica y metodológica, a la construcción de esta como ciencia del lenguaje, haremos especial énfasis en esta mirada.

En sus trabajos, Felber denomina la terminología como "Terminology Science" y la define como un "campo de conocimiento inter y transdisciplinario que tiene que ver con los conceptos y sus representaciones (términos, símbolos, etc.)"[24] (1984, p. 1); concepción que también es compartida por Arntz & Picht en la publicación de su manual *Introducción a la Terminología* asociándola, incluso, con otras áreas del saber tales como la lingüística, la filosofía o la ingeniería del conocimiento (1995, p. 22). Para Cabré, la terminología, como disciplina, está marcada por su carácter interdisciplinario en el cual encuentra las bases para una delimitación de su campo de estudio al tomar elementos de las áreas con las cuales cohabita para definir el alcance de su quehacer investigativo (1999, p. 25). En Pavel & Nolet se introduce la noción de "language discipline" la cual se dedica al estudio científico de los conceptos y términos utilizados en los lenguajes de especialidad (2001, p. XVII). En Faber la terminología es entendida más allá de un espacio técnico o una instancia particular de la lengua de especialidad, es decir, una concepción compartida con los autores arriba mencionados en virtud de la definición de esta como una disciplina arraigada a la lingüística (2012, p. 9).

[24] Traducción propia. En el texto original: "Inter- and transdisclplinary field of knowledge dealing with concepts and their representations (terms, symbols, etc.) (Felber, 1984, p. 1).

En el ámbito latinoamericano, Fedor de Diego ahonda en la concepción de la terminología como ciencia al denominarla como "la ciencia que estudia la formación y el uso de los términos" (1995, p. 13). Para Cardero la terminología, como ciencia, además permite la integración entre el especialista y el terminólogo en función de la identificación del vocabulario de especialidad de forma sistemática (2004, p. 28). En Adelstein la terminología como disciplina discurre entre los límites de la semiótica y la lingüística con miras a la delimitación de su objeto de estudio (2007, p. 22). En línea con estas autoras, Faulstich reconoce el estatus disciplinar de la terminología e incorpora a esta concepción el uso de los corpus de especialidad en el trabajo terminológico (2006, p. 27). Por último, las demás escuelas latinoamericanas, a saber: mexicana, cubana, colombiana, venezolana, peruana, ecuatoriana, argentina, chilena, uruguaya y brasileña, aunque no explícitamente, concuerdan con la concepción disciplinar de la terminología.

2.2 El término

Uno de los mayores desafíos a los que aún se enfrenta nuestra disciplina se centra en la definición de su objeto de estudio, esto es, la construcción de una concepción del término que permita un entendimiento desde las distintas perspectivas de análisis y que, a su vez, logre integrar las diversas miradas en torno al concepto de especialidad. Además, en la identificación de qué es o no un término intervienen, entre otros aspectos, la percepción del terminólogo y la del especialista, esto es, el alcance desde cada eje de acción. Así, para el terminólogo es importante la identificación del término mientras que para el experto el concepto es el punto de encuentro y, en consecuencia, los enfoques tanto onomasiológicos como semasiológicos que conllevan a la diferenciación entre el alcance de trabajo de la lexicología/lexicografía y el de la terminología/terminografía (Cabré, 2007, cap. 1). Otro aspecto importante, ya mencionado por Adelstein en su tesis doctoral, está relacionado con el tratamiento del léxico de especialidad tanto desde la lingüística como desde la

terminología (2007, p. 22). En el presente apartado, se abordará el término como unidad de análisis de la terminología con el afán de observar sus características lingüísticas y, de este modo, orientar nuestra discusión hacia una ampliación del concepto hasta ahora propuesto y fundamentado bajo los postulados cabretianos desarrollados en la TCT (Cabré, 1998, p. 48).

En primera instancia, en relación con el rol del experto en la identificación del término, nuestra área ha destacado su papel en este proceso, sin embargo, debido a la normalización de los conceptos por parte de estos, parece ser que su mirada no es suficiente a la hora de identificar el carácter terminológico de las unidades empleadas en los contextos de especialidad y, en consecuencia, trabajos referentes en el área como el de Estopà (1999), en el cual se establece la mirada del experto como eje central en la identificación del término, es debatible tanto por elementos de orden metodológico, al tener en cuenta no solo los perfiles de experto definidos en su trabajo, sino también el instrumento de análisis empleado para el objetivo de su investigación y la representatividad de los datos obtenidos de cara a las conclusiones presentadas por la autora (pp. 361-369). En segunda instancia, al hablar de término es menester mencionar las ópticas desde las cuales se ha abordado, tradicionalmente, esta discusión. Por un lado, desde la academia dos perspectivas, la primera de orden semiótico y la segunda lingüístico y/o cognitivo. Por otro lado, desde lo normativo, consecuencia de las discusiones académicas, la creación de distintas normativas de orden internacional como aquellas promulgadas por la *International Organization for Standarization* (ISO), a saber: ISO 1087 (2019) o la ISO 704 (2022). En línea con Adelstein, concordamos con las ópticas en la comprensión de la definición y caracterización del término como unidad de estudio de la terminología, es decir, los postulados desde la terminología clásica entendida bajo la mirada en Wüster y la de los autores denominados como los terminólogos del "giro lingüístico de la terminología" (Adelstein, 2007, p. 27).

En Wüster el término es entendido conforme a la relación que se establece entre el concepto y su denominación, es decir, para este autor el término corresponde a una denominación de un concepto técnico (1998, p. 11). Sobre las características del término y su definición, varias investigaciones han retomado lo aportado por Wüster en su tesis y trabajos posteriores, entre estos, resaltamos el artículo de investigación sobre el *modelo de término wüsteriano* publicado por Adelstein en 1998. Por lo tanto, la visión fundacional de Wüster a través de las distintas épocas de su trabajo terminológico no solo permitieron definir el término desde lo formal en su comprensión como sinónimo de concepto, sino que estableció las bases sobre las cuales se trabajó la terminología clásica en busca de una unidad propia de análisis y, en consecuencia, el origen de la oposición entre *término* y *palabra*, que no es más que la distinción entre el lenguaje natural y de especialidad ahondada, entre otros, por Cabré (1998, p. 7).

Las características del término en la TGT están determinadas, en esencia, no por elementos de orden lingüístico de la unidad de análisis sino por la asociación del concepto que representan, al interior del área de especialidad, con el signo que lo materializa (enfoque onomasiológico) y, de esta manera, la posibilidad de lograr una comunicación más precisa y eficaz en contextos científico-técnicos. De allí, que el término en Wüster esté mediado por la arbitrariedad manifestada en la elección de las formas con las cuales el experto representa su conocimiento en un plano denominativo, dejando a un lado características propias de la lengua, es decir, al entender el término en el contexto de la lengua de especialidad, elementos como la univocidad, la monorreferencialidad o la homonimia se contrastan con la variación, la sinonimia y aspectos de orden social, cultural, afectivo, diacrónico, discursivo, entre otros. Además, desde esta perspectiva, los aspectos de orden morfosintáctico no son relevantes por lo que la fraseología fue excluida de los análisis al igual que otras unidades complejas sintácticamente y asuntos semánticos como las unidades metafóricas. Ahora bien, el aporte de la TGT en la definición del término permitió, a pesar de su carácter reduccionista, ampliamente debatido en los distintos enfoques de trabajo terminológico más allá

de las escuelas de Rusia, Viena y Praga, definir el carácter autónomo de la terminología y, en consecuencia, la modelación de una disciplina de trabajo con una base teórica y metodológica propia.

Con relación a las diferentes definiciones del término, en *La Terminología Avui Termes Textos i Aplicacions*, Cabré, Domènech & Estopà presentan una interesante compilación sobre este concepto (2018, pp. 67-74). Así, autores como Lotte (1961), Rondeau (1983), Rey (1979) concuerdan con que el término está determinado por su función denominativa en cuanto se considera como una representación sígnica del concepto, mientras que en Sager (1993), Cabré (1992, 1998), Faber (2012) el término se observa a partir de sus características lingüísticas, es decir, son entendidos como unidades lingüísticas. En este sentido, en la actualidad, el término es entendido bajo los postulados de la TCT y la teoría de las puertas (Cabré, 2000) en la cual subyace el principal replanteamiento a la propuesta wüsteriana del término que, a su vez, se ha configurado como la base para otras concepciones de base no lingüística en Temmerman (2000), Faber (2012) y Gómez et al. (2019). Así, para la TCT las denominadas unidades terminológicas (UT) deben cumplir con tres características esenciales: 1) estructura, 2) especificidad semántica y 3) necesidad en la estructura conceptual (Cabré & Estopà, 2005, p. 77). Además, en relación con la categoría gramatical de las UT, el estudio se ha decantado por aquellas nominales, entre otras razones, por la frecuencia de aparición y su capacidad denominativa, es decir, para Cabré & Estopà existen UT de tipo verbal, adjetival o adverbial, pero en las nominales recae la mayor frecuencia en la expresión del conocimiento en los textos especializados (2005, p. 82) y, en consecuencia, su predilección como unidad de análisis. Además, en Adelstein la predominancia de las UT nominales se debe, principalmente, a sus características prototípicas de denominación y referencialidad (2004, p. 78).

Desde el punto de vista cognitivo, en la TBM de Faber se concibe el término a partir de su concepción en la TCT al considerarlo a partir de las unidades nominales y en consonancia con la no distinción entre palabra y término ya superada por los postulados en Cabré (Faber, 2012,

p. 22). De este modo, en el desarrollo de su propuesta de investigación, la autora concibe el término como una UT por medio de la cual se vehicula el conocimiento de especialidad, a partir de un análisis lingüístico en el que se incluyen los niveles morfológico, sintáctico, semántico y pragmático y, en consecuencia, una mirada contextual de las UT (2012, p. 249). Lo interesante de este acercamiento teórico recae en la inclusión del aspecto cognitivo hasta el momento solo abordado por la TSC de Temmerman (2000). En Faber, la modelación del conocimiento especializado, a partir de la adaptación de la teoría de los Marcos de Fillmore (1982), busca de una mejor comprensión de los sistemas de conceptos dinámicos y complejos producto de la integración entre las descripciones lingüísticas y gráficas de las entidades especializadas (Faber, 2012, p. 31). Asimismo, el trabajo del grupo CITERM, de Manizales, fundamenta parte de su investigación en los aspectos cognitivos de la terminología[25].

En lo normativo, se evidencia uno de los logros más representativos del trabajo en Wüster. La creación del Comité Técnico 37 (TC 37)[26] por la *International Standarization Organization* (ISO) en el año de 1947 permitió la internacionalización de la terminología y, en consecuencia, la publicación de las normas ISO 704 (2022)[27] o la ISO 1087 (2019) en las cuales se establecen, por un lado, los principios y métodos básicos en la preparación y compilación de las terminologías (ISO 704, 2022) y, por otro lado, las definiciones y conceptos básicos en terminología y para el trabajo terminológico (ISO 1087, 2019)[28]. Además, en el contexto alemán también se cuenta con la norma DIN 2342[29] emitida por el *Deutsches Institut für Normung* (2022) en la cual se definen los conceptos básicos en terminología para el

[25] Para más información sobre los trabajos publicados por el grupo CITERM, consúltese: https://www.autonoma.edu.co/investigacion/grupos-de-investigacion/citerm.

[26] Para más información sobre la ISO/TC 37, consúltese: https://www.iso.org/committee/48104.html.

[27] Existe una adopción idéntica (IDT) por traducción al español de la norma ISO 704:2009 editada por el Instituto Colombiano de Normas Técnicas y Certificación (ICONTEC) en el 2013.

[28] Para más información sobre las Normas ISO 704 y 1087, consúltese: https://www.iso.org/standard/38109.html y https://www.iso.org/standard/62330.html, respectivamente.

[29] Para más información sobre la DIN 2342, consúltese: https://tienda.aenor.com/norma-din-2342-2011-08-140370128.

entendimiento del trabajo terminológico. El aspecto fundamental desde lo normativo recae en la búsqueda de la estandarización a nivel internacional de cara a la práctica además de dar cuenta de las bases teóricas en la concepción de las normas. Al respecto, en las normas ISO se identifica claramente la noción de término wüsteriana mientras que en las DIN se identifica el giro en esta concepción al definirlo bajo el entendimiento de UT (DIN 2342, 2022). Esta distinción es también mencionada por Adelstein en su trabajo de 2002 donde indaga sobre el estatuto lingüístico del término y su representación.

En general, la concepción actual de término es la consecuencia de las dos miradas paradigmáticas: la primera en Wüster con los planteamientos de la TGT y la segunda en Cabré desde la TCT, quien a su vez influyó en los postulados modernos en el trabajo terminológico. Así, hoy en día, el término se concibe, *grosso modo*, como una "unidad léxica cuya estructura corresponde a una unidad léxica de origen o producto de la lexicalización de un sintagma, que posee un significado específico en el ámbito al que se asocia y es necesaria en la estructura conceptual del dominio del que forma parte" (Cabré & Estopà, 2005, p. 77) a partir de enfoques tanto lingüísticos como cognitivos, donde la unidad prototípica de análisis se decanta por las unidades nominales. Frente a esto, a partir de los análisis en la presente tesis de trabajo concordamos, parcialmente, con la definición paradigmática de término toda vez que consideramos que no solo las unidades nominales deben ser entendidas como terminológicas, sino que cualquier unidad léxica[30] puede serlo. En este sentido, a pesar de que en la definición en Cabré se menciona el carácter de término como "unidad léxica" lo que hace referencia a cualquier unidad tanto simple como compleja, la realidad en las diferentes investigaciones ha demostrado que el análisis ha dejado a un lado o mencionado parcialmente las demás unidades más allá del nombre. Con relación a esto, a pesar de las razones para esta predilección, no hay una justificación,

[30] En relación con nuestra visión de *término* como una etiqueta en la cual se incluye tanto unidades simples como complejas en función de la cantidad de componentes, entendemos el concepto *unidad léxica* bajo la misma óptica, lo que incluye unidades simples como adjetivos, adverbios, sustantivos y verbos, al igual que unidades complejas (también podrían denominarse sintácticas) en las cuales se incluyen, por ejemplo, metáforas, sintagmas, entre otros.

ni desde la teoría ni la estadística, para dejar a un lado a esas otras unidades. Así, los análisis realizados en nuestra tesis nos permiten indicar que cualquier unidad léxica podría ser categorizada como término puesto que en estas se contiene el contenido temático que permite asociarlo con un nodo cognitivo dentro del lenguaje de especialidad al cual pertenece y, de allí, la tesis central del protocolo de validación de candidatos a término al igual que de la noción de *variación terminológica*, aspectos que se desarrollaran en los siguientes apartados (§ 2.3 y 2.4).

En suma, en esta tesis doctoral, se define el término como *todo tipo de unidad léxica, simple o compleja que, en función de cada discurso de especialidad, contiene y materializa los saberes propios de un campo del conocimiento a través de los diferentes mecanismos inherentes al lenguaje natural, a partir de elementos lingüísticos y cognitivos que, a su vez, responden a una función social mediada por aspectos de orden cultural.*

2.3 Bases teóricas para un protocolo de validación de términos simples y complejos

Resultado de las reflexiones teórico-metodológicas en torno a la concepción de la terminología y, por extensión, del término como unidad de análisis y con miras a una actualización de las concepciones paradigmáticas actuales en la comprensión de estos dos conceptos, en la presente tesis doctoral retomamos la propuesta de identificación y clasificación de términos, simples o complejos, desarrollada por Quiroz & Burgos (en prensa) en el marco de la investigación sobre *variación terminológica* en la cual se adscribe el presente trabajo. A continuación, se relacionan las bases teóricas en la construcción de este protocolo, las cuales permitieron, a su vez, una profundización en nuestra conceptualización, la cual será ampliada en el apartado de Resultados y Discusión (§§ 4-5). En primera instancia, se presentan los preliminares en la construcción del protocolo. En segunda instancia, las características en la delimitación de cada unidad léxica analizada y

su posterior denominación como *terminológico*, es decir: a) el sustantivo terminológico, b) el adjetivo terminológico, c) el verbo terminológico, d) el adverbio terminológico y d) las unidades complejas terminológicas. Por último, en tercera instancia, se relacionan algunas consideraciones generales para la comprensión del protocolo y su posterior aplicación metodológica en el corpus compilado en nuestra investigación para la Variación Terminológica Horizontal (VaTeHo).

2.3.1 Cuestiones preliminares en el diseño del protocolo de validación

En inicio, es importante recordar que en la presente tesis los análisis en la identificación de los términos se enfocan en el plano lingüístico, esto es, se tiene como ejes de trabajo aspectos que van desde lo morfológico y sintáctico hasta reflexiones de orden semántico y pragmático. En este sentido, los criterios definidos por los autores y, posteriormente, identificados en los candidatos a término extraídos a través del programa *Termostat* (para más detalles, véase la § 3. Metodología), recopilan las bases conceptuales tanto desde la lingüística como la terminología y, por ende, se incorporan referentes teóricos desde ambas áreas del conocimiento, entre ellos: Sager (1978), Nunberg et al. (1994), Alvar (1996), Wright & Budin (1997), Cabré (1998), Lorente (2002), Drouin (2003, 2010), Quiroz (2008), Quiroz & Arroyave (2014), Burgos (2014), Montero (2014), Pérez-Pérez (2018) y Szabó (2020). Asimismo, al inicio de la concepción del protocolo de validación (Quiroz & Burgos, en prensa) y, por extensión, de la fundamentación teórica de la presente tesis doctoral, se tomó como base la definición de término propuesta en la TCT (Cabré, 1998, p. 48) para la cual, fruto de los resultados y análisis realizados en nuestra investigación, se presentó una ampliación en la propuesta de definición de término (§ 2.2). Nuestra tesis central de trabajo, retomando lo expresado en el protocolo de validación, nos permite delimitar el término según cumpla o no con las características de *unireferencialidad*, *no composicionalidad*, *arbitrariedad* y *convencionalidad* y, en consecuencia, entenderlo como una etiqueta que ocupa un nodo cognitivo al interior de un campo de especialidad (Quiroz & Burgos, en prensa).

A razón de ello, uno de los principios rectores en el entendimiento del protocolo tiene que ver con la identificación de los conceptos, esto es, para lograr una aplicación rigurosa del protocolo es necesario tener un conocimiento contextual del área analizada y, por ende, la estratificación en los criterios de trabajo, es decir, a pesar de las claves sintácticas y de la representación del conocimiento, siempre debe prevalecer en la determinación de la unidad como término las nociones conceptuales (semánticas y pragmáticas) propias de cada campo del saber. En este sentido, el papel del experto es relevante en el proceso, pero no es determinante ni excluyente, entre otras, por las razones arriba presentadas a raíz de la crítica al trabajo de Estopà (1999). Para ello, en función de la compilación de los saberes disciplinares, las áreas cuentan con diversidad de recursos lexicográficos para este propósito los cuales, a raíz de las conclusiones en el presente trabajo, también deberían ser revisados en función de su entendimiento sobre qué es y qué no es un término y, por ende, de las entradas por estos registradas. A continuación, se relacionan las características base para la determinación de las unidades léxicas como términos y, por extensión, los criterios para su designación como no términos sino como unidades con carácter de especialidad.

2.3.1.1 Unireferencialidad del término

Aplicado a la terminología, la referencialidad remite a la relación entre el concepto y su equivalente en el plano lingüístico, aspecto que no debe confundirse con la monosemia o monoreferencialidad propuesta por la TCT e incluida en la norma ISO 704. En este orden de ideas, y en sintonía con lo expresado por Wright & Budin (1997), Montero (2014) y Burgos (2024), la unireferencialidad, como característica principal del término, indica la relación única entre la unidad lingüística y el nodo cognitivo al que hace referencia en el campo del conocimiento al que pertenece, sin importar si la unidad es simple (adjetivo, adverbio, sustantivo o verbo) o compleja. De allí, la cualidad del término como referencia a un único nodo cognitivo y su entendimiento como una etiqueta no composicional sin importar su extensión o cantidad de elementos constitutivos. Así, visto bajo esta óptica, formalmente,

no hay distinción entre un término simple o complejo, toda vez que este último es entendido como un solo elemento en el cual se incluyen los espacios en blanco que constituyen un carácter más de la etiqueta (Quiroz & Burgos, en prensa). Como resultado de estas consideraciones, las unidades en 2 y 3[31] son consideradas como término:

2. **Adjetivos**: *celular*, *gástrico* (Med); *sintáctico, gramatical* (Lin); *agropecuario*, *bovino* (Vet); *analógico*, *computacional* (Com); *empírico*, *filosófico* (Fil).

   Adverbios: *histológicamente*, *quirúrgicamente* (Med); *fonéticamente*, *semánticamente* (Lin); *enzoóticamente*, *fenotípicamente* (Vet); *análogamente*, *programáticamente* (Com); *cognoscitivamente*, *ontológicamente* (Fil).

   Sustantivos: *metástasis*, *quimioterapia* (Med); *diccionario*, *término* (Lin); *ganado*, *raza* (Vet); *aplicación*, *comando* (Com); *conocimiento*, *verdad* (Fil).

   Verbos: *diagnosticar*, *padecer* (Med); *definir*, *sustantivar* (Lin); *criar*, *ordeñar* (Vet); *automatizar*, *desarrollar* (Com); *afirmar*, *ser* (Fil).

3. **Términos complejos**: *cáncer de próstata* (Med, N+Prep+N), *célula tumoral* (Med, N+Adj); *análisis del discurso* (Lin, N+Prep+Adj), *descripción metalexicográfica* (Lin, N+Adj); *sistema de pastoreo* (Vet, N+Prep+Adj), *ganado bovino* (Vet, N+Adj); *diagrama en bloque* (Com, N+Prep+N), *capacitor electrolítico* (Com, N+Adj); *filosofía del lenguaje* (Fil, N+Prep+N), *hombre natural* (Fil, N+Adj).

[31] Los ejemplos aquí y en adelante presentados son tomados del corpus para la Variación Terminológica Horizontal (VaTeHo) compilado para los propósitos de esta tesis doctoral. Unido a los ejemplos, conforme a la necesidad de explicación, entre paréntesis se puede relacionar ya sea el área de conocimiento o la categoría gramatical o la estructura sintáctica. Para más información sobre este recurso y la descripción de las abreviaturas empleadas, véase el capítulo 3 Metodología.

2.3.1.2 Composicionalidad del término complejo

Toda vez que el análisis que nos proponemos en la presente tesis doctoral también abarca aspectos desde la semántica de los términos, la composicionalidad como característica de las unidades poliléxicas se concibe a sí misma como una de las tareas de mayor relevancia en el dominio terminológico. Así, al entender que el significado de las unidades complejas está determinado por su estructura y la suma de los significados de sus componentes (Szabó, 2020, párr. 1), los términos complejos están fuera de este análisis. En este sentido, la no composicionalidad del término tiene su arraigo no solo en la unireferencialidad sino también en el carácter arbitrario de estos. De este modo, siguiendo lo expresado por Quiroz & Burgos (en prensa) en relación con este criterio del término, al tener expresiones como *diagrama de ojo* nos referimos a un término del procesamiento de señales en las telecomunicaciones digitales, mientras que, en *extirpación de colón*, en la medicina, tenemos dos conceptos del área lo que constituiría una secuencia con contenido terminológico que podría funcionar, en algunos casos, como una explicación del término, en este particular para *colectomía*, característica que se complementa con el criterio de *arbitrariedad*. Asimismo, bajo la óptica del término como una etiqueta (Burgos, 2024), no existe distinción entre las unidades simples de las complejas. Sin embargo, por motivos de orden operativo y metodológico conservamos esta distinción en la denominación toda vez que es ampliamente aceptada por la terminología. A su vez, la composicionalidad de los términos también ha sido discutida por autores como McCray et. al (1994), Navigli et. al (2003), L'Homme (2004) y Burgos (2024).

2.3.1.3 Arbitrariedad del término

La arbitrariedad del término complejo está determinada por las elecciones tomadas por cada área del saber en las formas de expresar su conocimiento según las preferencias de la comunidad discursiva, permitiendo un sin número de posibilidades en la motivación y posterior designación de las unidades lingüísticas para la materialización de sus nodos cognitivos. Así, retomando las ejemplificaciones en Quiroz

& Burgos (en prensa), en la medicina *mano ajena* hace referencia a un problema neuropsicológico cuya denominación está motivada por el reflejo del trastorno, mientras que en *dispraxia diagonística* su origen es la base para la designación. Esta particularidad en la motivación para la denominación de las etiquetas permite establecer una gradación de la arbitrariedad en mayor y menor medida y, en consecuencia, la utilización de distintos recursos de la lengua en la construcción de la etiqueta. Así, por un lado, entre más arbitraria la denominación se emplean recursos de tipo metafórico o expresiones figurativas o por extensión (Quiroz & Burgos, en prensa), como en el caso de: *abordaje quirúrgico* (Med), *aparato genital* (Med), *cuerpo léxico* (Ling), *lengua materna* (Lin), *ambiente virtual* (Com), *código fuente* (Com), *cabeza de ganado* (Vet), *bomba de leche* (Vet), *escuela de Frankfurt* (Fil), *juego dialéctico* (Fil) y, por otro lado, entre menos arbitraria la etiqueta es más factible encontrar un acortamiento de las descripciones composicionales del término, como se ilustra en los siguientes casos tomados del protocolo de validación (Quiroz & Burgos, en prensa): *objeto indirecto* (Lin) que corresponde a *objeto tomado indirectamente por el verbo* o en *diccionario etimológico* (Lin) proveniente de *diccionario que describe el origen etimológico en sus entradas*.

En este sentido, otro elemento descrito en el protocolo de validación (Quiroz & Burgos, en prensa), en relación con la determinación del grado de arbitrariedad del término, es que se debe tener en cuenta la lectura composicional de este: si el grado de arbitrariedad de la secuencia es cero entonces estamos frente a una descripción, que en ocasiones puede solaparse con la definición del término, y no ante un término complejo. En esencia, a mayor arbitrariedad del término mayor será la distancia entre este y su descripción y, por ende, entre menor arbitrariedad menor distancia. Asimismo, con fines operativos, la distancia puede ser medida tanto de forma cualitativa al comparar tanto el término como su descripción y de manera cuantitativa utilizando medidas ya establecidas como la distancia de Levenshtein[32],

[32] Para más información sobre la distancia de Levenshtein, consúltese: https://sites.google.com/site/algoritmossimilaridaddistancia/distancia-de-levenshtein.

ampliamente utilizada en investigaciones de naturaleza lingüística[33] (Quiroz & Burgos, en prensa). Sin embargo, en la presente tesis no se ahondará en este aspecto puesto que consideramos es un elemento que debe ser tratado en una nueva investigación debido a la amplitud en la discusión lo que, metodológicamente, se traduciría como un problema en la delimitación de nuestro alcance de trabajo. A modo de ejemplo, retomando algunos de los casos presentes en el protocolo de validación, la distancia entre *diagrama de ojo* y su descripción es menor que aquella en *mano ajena*. A su vez, para *extirpación de colón* la distancia es cero y, por lo tanto, se confirma su caracterización como una descripción más no como un término complejo (Quiroz & Burgos, en prensa).

Otras particularidades de la arbitrariedad están relacionadas con la capacidad de predictibilidad de las características del término en distintos niveles como el sintáctico (las categorías gramaticales) o semánticas (características del núcleo de la etiqueta) y, por ende, el enfoque de análisis del término. Así, en *biopsia endoscópica* (Med), *cáncer colorrectal* (Med), *ganado bovino* (Vet), *recurso natural* (Vet) estamos hablando que el modificador es un tipo del núcleo, mientras que en mano ajena (Med) no se está hablando de un tipo de mano lo que requiere un análisis más allá de la relación sintáctica (Quiroz & Burgos, en prensa). En este sentido, una de las bases conceptuales que, posteriormente, permitirán la operativización de este protocolo de validación está ligada a la gramática de dependencias planteada por Tesnière (1959). Además, en aspectos de la variación terminológica, la homonimia en el discurso especializado, los nombres propios y grupos no nominales, también es factible observar la arbitrariedad del término.

2.3.1.4 Convencionalidad del término

La última de las características base en la delimitación de los términos, desarrolladas en el protocolo de validación (Quiroz & Burgos, en prensa), está orientada a la convencionalidad resultado de elementos tanto de la *composicionalidad* como de la *arbitrariedad*. En este

[33] Para investigaciones en la lingüística donde se ha empleado la distancia de Levenshtein, explórese: https://scholar.google.com/scholar?hl=es&as_sdt=0%2C5&q=distancia+levenshtein&btnG=.

sentido, el término simple o complejo se concibe como una etiqueta convencional que demanda por parte del experto no solo conocer el dominio de especialidad a la cual pertenece sino también el significado de la etiqueta, además de comprender los distintos mecanismos en la construcción de la etiqueta que puede incluir tanto unidades léxicas como funcionales. En este sentido, en relación con los términos complejos, para el experto no será suficiente la sumatoria de los componentes de la etiqueta toda vez que, como lo hemos establecido, su significado no está mediado por este aspecto más por la arbitrariedad en la constitución de la etiqueta (Quiroz & Burgos, en prensa). De esta manera, en la comprensión del término *cabeza de ganado* (Vet) no es suficiente la sumatoria de los significados denotativos de cada constituyente [cabeza+de+ganado] sino que se hace necesario identificar la asociación convencional entre la etiqueta y su definición: *Res, unidad de cualquier tipo de ganado* (Real Academia Española, 2022c), caso contrario en la comprensión de la descripción donde es suficiente el contenido denotativo de cada unidad monoléxica.

2.3.2 Unidades léxicas terminológicas

Discutidas y ejemplificadas las características base en la construcción del protocolo de validación, es pertinente ahondar en la concepción de cada unidad léxica, simple o compleja, como terminológica. El fin de este segmento no es el presentar una suerte de tratado donde se relacionen los pormenores de cada unidad sino una mirada prototípica que permita identificar los rasgos lingüísticos que las configuran como portadoras del conocimiento de especialidad, a partir de una revisión bibliográfica al respecto de cada unidad, para entender cómo desde la literatura disponible se ha hecho énfasis en el componente especializado atendiendo al postulado según el cual el lenguaje de especialidad se configura como un subconjunto de la lengua general (Cabré, 1998, p. 7). Sin embargo, es la regularidad que el foco de las investigaciones en terminología tienda a ser orientado hacia el estudio de las unidades nominales, lo que supone un vacío considerable en el tratamiento de esas otras unidades y, por ende,

la poca o incluso nula bibliografía en su tratamiento especializado como se evidencia, por ejemplo, con el adverbio terminológico; unido a la poca producción sobre y en español.

A continuación, se relacionan algunas consideraciones generales sobre cada unidad léxica y su potencial terminológico. Sin embargo, es importante mencionar que, en la posterior aplicación metodológica del protocolo de validación, se ahondará en nuestra caracterización y concepción de término a partir de las particularidades de cada unidad en función de las cinco (5) áreas del conocimiento analizadas (§§ 4-5). El orden en la presentación de los siguientes subsegmentos se estableció alfabéticamente según cada unidad, con la intensión de indicar que todas las unidades son igual de relevantes en la expresión del conocimiento especializado sin dejar a un lado el carácter prototípico de las unidades nominales con el cual concordamos.

2.3.2.1 El adjetivo terminológico

El adjetivo, como unidad autónoma y sintagmática, ha sido estudiado desde la lingüística en variedad de trabajos, ejemplo de ello: Alcina & Blecua (1975), Alarcos (1994), Hernanz & Brucart (1987), De la Peña (1985), Gili Gaya (1994), Hernández (1996), Demonte (1999), Dixon (2004, 2010), entre otros. Sin embargo, son pocos los estudios desde la terminología que indagan, desde el contexto de los lenguajes de especialidad, las características de esta unidad. Al respecto, Folguerà (2002) demuestra el funcionamiento terminológico del adjetivo tanto como unidad simple autónoma como inmersa en estructuras sintagmáticas en cuanto permite la lexicalización de los sintagmas (pp. 213-215). En Salazar (2011), el adjetivo se configura como la categoría que mejor establece la diferenciación entre el lenguaje especializado y el general en un área como la economía (p. 26). En Quiroz (2008), se ha descrito los procesos de sufijación en los adjetivos en contextos de especialidad determinando, por ejemplo, la relación cuantitativa con aquellos de la lengua general concluyendo que no hay mayores diferencias entre ambos contextos (p. 228). Asimismo, el autor establece, a partir de la caracterización morfológica, la predominancia de los sufijos *-ico* y *-al* en la formación

de los adjetivos en el corpus analizado (genoma humano) (p. 229). En relación con el adjetivo terminológico, Quiroz & Vásquez (2024) presentan una caracterización de esta unidad léxica y su carácter terminológico. De este modo, el adjetivo como unidad léxica es tanto portadora como transmisora del carácter de especialidad. En 4, se listan ejemplos de adjetivos término:

4. **Medicina**: *cervical*, *craneal*, *gástrico*, *histológico*, *médico*, *terminal*, *torácico*, *uretral*.

   Lingüística: *agramatical*, *deverbal*, *dialectal*, *funcional*, *metafórico*, *semántico*, *vocálico*.

   Veterinaria: *agronómico*, *biótico*, *endémico*, *genético*, *racial*, *sérico*, *trófico*, *zoonótico*.

   Computación: *analógico*, *artificial*, *computacional*, *métrico*, *secuencial*, *sigmoidal*.

   Filosofía: *aristotélico*, *cognoscitivo*, *epistémico*, *racional*, *simbólico*, *teleológico*.

2.3.2.2 El adverbio terminológico

El adverbio, contrario al sustantivo, es la unidad menos estudiada en el discurso de especialidad, incluso, hoy en día, no se encuentran referencias directas en la literatura con relación a su tratamiento y descripción en la terminología, lo que supone un importante desafío a la hora de analizar su comportamiento dentro del discurso disciplinar. Contrario a este panorama, desde la lingüística sí se encuentran referencias y avances importantes en la descripción de esta unidad y su carácter de complementariedad respecto de las demás unidades léxicas. Gramaticalmente, el adverbio se ha categorizado desde obras clásicas como en De Nebrija (1492:2010) a partir de su característica prototípica que subyace en la capacidad de modificar otras partes de la oración, elemento que se confirma siglos después por Alarcos (1969) al determinar la influencia oracional de este (p. 302), Además,

establece la falta de descripción del adverbio en las gramáticas actuales al considerar que solo se menciona su repercusión modificadora en las otras unidades (verbo, adjetivo y otros adverbios) quedando su análisis exclusivamente en un plano semántico (p. 301) y, de este modo, que la idea generalizada sobre esta unidad en la terminología este regida por esta óptica y, por ende, la falta de interés en su análisis al no considerarse como una unidad portadora de contenido de especialidad. Ahora bien, la dinámica observada en los análisis de la presente tesis doctoral, permitirían constatar que el adverbio cumple no solo su función sintáctico-semántica con el resto de las unidades, sino que por sí solo daría cuenta del conocimiento especializado en función del área disciplinar donde este inmerso como se discutirá más adelante (§§ 4-5). Asimismo, en Quiroz & Diaz (2024) se encuentra una revisión tanto bibliográfica como teórica de esta unidad y una propuesta en la determinación de su carácter terminológico. En 5, se listan algunos ejemplos de adverbios término, siendo los prototípicos aquellos terminados en *-mente*:

5. **Medicina**: *histológicamente*, *inmunohistoquimicamente*, *imagenológicamente*, *macroscópicamente*, *molecularmente*, *transoperatoriamente*.

   Lingüística: *dialectalmente*, *fonéticamente*, *metafóricamente*, *morfológicamente*, *postsintácticamente*, *semánticamente*, *sincrónicamente*, *sociolingüísticamente*.

   Veterinaria: *enzoóticamente*, *epizootiológicamente*, *estacionalmente*, *fenotípicamente*.

   Computación: *algorítmicamente*, *análogamente*, *cíclicamente*, *digitalmente*, *logarítmicamente*, *tecnológicamente*.

   Filosofía: *acríticamente*, *cognoscitivamente*, *epistemológicamente*, *ontológicamente*.

2.3.2.3 El sustantivo terminológico

El sustantivo, es por sí mismo, la unidad prototípica del análisis terminológico, lo que ha permitido que se convierta en el foco de atención de los terminólogos a la hora de presentar sus análisis (Cabré, 1998, Adelstein, 2007, entre otros autores). Sin embargo, a pesar de su prototipicidad, aún existe un vacío teórico y metodológico en la caracterización de esta unidad, en español, que incluya reflexiones desde la morfología, sintaxis, semántica y pragmática. Así, por analogía al concepto de *punto* en la matemática, el sustantivo como término se ha convertido en un axioma de la terminología: *el sustantivo es un término y punto.* Investigaciones en el campo indican, cuantitativamente, la prevalencia de los sustantivos en contraste con las demás unidades, por ejemplo, en un estudio sobre la terminología en la ganadería bovina, Galeano indica que, con relación a las unidades simples, el sustantivo representa el 59,6 % de un total de 16.553 extraídas de forma automática (2017, p. 134). El fundamento central en la determinación del sustantivo como término y, por ende, su prevalencia en los trabajos terminológicos radica en su potencial denominativo y referencial (Adelstein, 2004, p. 78) concordando con los fundamentos del área de cara a una estandarización y normalización del conocimiento especializado (Wüster, 1998; Cabré, 1999b). Al respecto de la determinación de los sustantivos terminológicos, Burgos & Vásquez (2024) presentan un análisis de esta unidad en relación con la expresión del conocimiento de especialidad y, por ende, la ratificación de su denominación como términos. A modo de ilustración de la capacidad nominal para contener el conocimiento especializado se listan los siguientes ejemplos en 6:

6. **Medicina**: *biopsia*, *carcinoma*, *metástasis*, *patología*, *próstata*, *quimioterapia*, *tumor.*

   Lingüística: *alfabeto*, *concordancia*, *diccionario*, *lengua*, *léxico*, *palabra*, *sílaba*, *sufijo.*

   Veterinaria: *camada*, *distomatosis*, *forraje*, *ganado*, *hato*, *leche*, *potrero*, *rebaño*, *vaca.*

Computación: *aplicación*, *capacitor*, *cliente*, *desarrollador*, *microcontrolador*, *usuario.*

Filosofía: *conocimiento*, *dialéctica*, *dogma*, *episteme*, *eticidad*, *fenomenología*, *verdad.*

2.3.2.4 El verbo terminológico

El verbo, considerado una de las palabras más complejas en cualquier lengua, es entendido como el núcleo del predicado y la palabra flexiva por excelencia (Alcoba, 1999, p. 4917). De allí, su relevancia en los análisis lingüísticos y, por extensión, la necesidad de ahondar en su rol al interior de los discursos de especialidad. Investigaciones en el área (L'Homme, 2002; Lorente, 2002; Costa & Silva, 2004; Galeano, 2017; Pérez-Pérez, 2018; Pérez-Pérez et al., 2019) establecen, desde diversos análisis y enfoques de trabajo, el carácter terminológico del verbo indicando su potencial en la construcción del conocimiento especializado, además de enfatizar en la ausencia de un criterio riguroso que justifique la exclusión de estos en recursos de orden terminográfico (Lorente, 2002, §. 5.4). Además, a partir del verbo, al igual que de los adjetivos, es posible identificar el comportamiento y significado de las unidades nominales (L'Homme, 2002, p. 32). Sin embargo, a pesar de la mención sobre el componente morfológico, sintáctico o semántico del verbo (L'Homme, 2002; Martínez & Castro, 2017; Perez-Pérez, 2018; Perez-Pérez et al., 2019), solo existe una propuesta que permite caracterizar esta unidad como término (Lorente, 2002). En esta misma dirección, Cardona, Pérez-Pérez & Zuluaga (2024) caracterizan el verbo terminológico presentando una descripción de esta unidad al introducir al ámbito de discusión, entre otras, las características aspectuales del verbo en función de las distintas áreas del saber. Al igual que se ha establecido en las demás unidades simples, el verbo tiene, por sí mismo, la capacidad de dar cuenta del discurso especializado en función del área de trabajo (Pérez-Pérez, 2018, Pérez-Pérez et al., 2019), como se evidencia en los ejemplos listados en 7:

7. **Medicina**: *anestesiar*, *curar*, *detectar*, *diagnosticar*, *fallecer*, *malignizar*, *metastatizar*.

   Lingüística: *adjetivar*, *derivar*, *elidir*, *glosar*, *inferir*, *lematizar*, *lexicalizar*, *prefijar*.

   Veterinaria: *abrevar*, *cercar*, *criar*, *cultivar*, *cruzar*, *fertilizar*, *ordeñar*, *parcelar*, *pastar*.

   Computación: *acoplar*, *compilar*, *ejecutar*, *modelar*, *notar*, *parametrizar*, *programar*.

   Filosofía: *argüir*, *aprehender*, *conjeturar*, *develar*, *fundamentar*, *negar*, *repensar*.

2.3.2.5 Unidades complejas terminológicas

En el estudio de las unidades complejas terminológicas los avances más representativos se orientan a su reconocimiento y extracción automática. Al respecto, Cabré, Estopà & Vivaldi presentan una interesante compilación de sistemas diseñados para tal propósito (2001, p. 53). Además, aún no existe un punto de encuentro entre los teóricos de la terminología que permita delimitar y caracterizar dichas unidades lo que ha llevado a no tener claros los límites entre los términos complejos y otras unidades propias de la fraseología. De este modo, lograr establecer criterios que permitan la caracterización de estas unidades es vital en cuanto garantizaría un mejor desempeño en las áreas donde la terminología es un insumo de vital importancia en el desarrollo del quehacer disciplinar, entre estas: la lexicografía, la traducción, el procesamiento del lenguaje natural. En este sentido, teniendo en cuenta las características arriba descritas (§ 2.3.1), Burgos (2024) afirma que los términos complejos son entendidos como etiquetas no composicionales con criterios de *convencionalidad* y *arbitrariedad*. Así, los términos complejos juegan un rol determinante en la configuración lingüística de cara a la materialización de los saberes propios de cada contexto disciplinar. Los ejemplos en 8 dan

cuenta de términos complejos en cada una de las áreas analizadas en la presente tesis doctoral:

8. **Medicina**: *cáncer de cuello uterino*, *carcinoma de célula escamosa*, *ganglio linfático regional*, *leucemia linfoblástica aguda*, *sarcoma de parte blando*, *terapia endocrina básica*, *vaciamiento central*.

   Lingüística: *acto de habla*, *análisis del discurso*, *concordancia de sujeto*, *descripción metalexicográfica*, *riqueza léxica*, *significado léxico*, *sistema lingüístico*, *variedad lingüística*.

   Veterinaria: *cabeza de ganado*, *carga animal*, *célula alimentadora inactivada*, *condición de stress calórico*, *densidad de tallo*, *época seca*, *estatus serológico*, *fiebre aftosa*, *ganadería bovina*.

   Computación: *algoritmo de inferencia*, *ambiente virtual*, *árbol regular*, *cadena de carácter*, *dispositivo móvil*, *entorno de desarrollo integrado*, *gramática de árbol regular*, *lenguaje de programación*.

   Filosofía: *ciencia intuitiva*, *concepción cartesiana*, *epistemología del testimonio*, *estado de cosa*, *filosofía del lenguaje*, *hombre natural*, *reflexión epistemológica*, *testimonio moral*.

2.4 Variación terminológica horizontal

Al hablar de variación horizontal nos ubicamos en uno de los dos ejes bajo los cuales se ha observado, tradicionalmente, el discurso de especialidad. Así, complemento del eje vertical de la variación, encargada de establecer las dinámicas terminológicas en la relación entre los distintos participantes de la comunicación desde la mirada del experto hasta el lego pasando por el aprendiz; la variación horizontal propende por una teorización a lo largo y ancho del saber humano. En

general, las investigaciones de Hoffmann (1998), Cabré (1999a), Pavel (2001), Domènech (2006), Rodríguez-Piñero & García (2009), Cabré et al. (2018) han permitido identificar que la variación horizontal permite la clasificación de los lenguajes de especialidad a partir de criterios de cercanía o lejanía de los términos otorgando el carácter de especialidad a las áreas del saber y las distintas maneras de abordar un tema conforme la naturaleza de cada una de ellas. Sin embargo, al momento de ser escrita la presente tesis doctoral, en terminología no se cuenta con un estudio que abarcando varias áreas del conocimiento más allá de las comúnmente estudiadas (medicina y economía) permita observar empíricamente el comportamiento terminológico y, a razón de ello, proporcionar una categorización horizontal de las dinámicas terminológicas por áreas del saber. Otras investigaciones importantes en aspectos horizontales se han centrado en la identificación de lo estético (Beneš, 1969), la ciencia, técnica, economía, política, entre otras (Schmidt, 1969), la milicia (Kolguixkin, 1970) y los deportes (Riedel, 1975).

2.5 Recapitulación

Nuestra *mirada en conjunto* (§ 1) nos ha permitido ahondar en el trasegar de nuestra disciplina a través de las distintas etapas en su desarrollo y consolidación. Las indagaciones teóricas en el presente capítulo nos permiten recorrer el pasado y el presente con el reto de proyectar el futuro de la terminología. De este modo, nos hemos propuesto dejar en claro el carácter autónomo de la terminología como disciplina de la lingüística, incluso, podríamos hablar de la *ciencia de los términos* a partir del estatus epistemológico que nos permite entender "el conjunto de ecuaciones" (como se citó en Wallerstein, 2006, p. 3) de cara a la construcción de ese sistema de control (Focault, 1970, p. 224), que para nuestro particular no es otro que el discurso de los términos y, como consecuencia, su configuración como un aporte significativo en la construcción de la lingüística y los lenguajes de especialidad en general y particular.

En el concepto de término, como principio y fin de la terminología, nuestra disciplina encuentra una ardua tarea en cuanto a una concepción que permita caracterizarlo y automatizarlo para, de este modo, responder y sacar beneficio de las dinámicas modernas, en las cuales la computación se ha convertido en un aliado más, sin olvidar la esencia de nuestro trabajo en la descripción lingüístico-terminológica de estas unidades ya sean simples o complejas. A su vez, los múltiples acercamientos, a partir de la mirada paradigmática aún vigente y propuesta por Cabré (2000) en su teoría de las puertas, se han decantado en dos frentes de trabajo desde lo cognitivo y lo lingüístico, pero en ambos subyace la misma concepción de término como "unidad léxica que ocupa un nodo cognitivo dentro de un dominio de especialidad" (Cabré, 1998, p. 48). Al respecto, una preponderancia teórica y cuantitativa de la clase nominal como unidad de análisis, con la cual concordamos, pero que, a su vez, creemos con certeza que cualquier unidad léxica tiene en sí misma la capacidad de portar y expresar el contenido de especialidad.

Nos permitimos presentar una propuesta de ampliación del concepto paradigmático de término en la cual explicitamos nuestra concepción de este, hasta ahora justificada teóricamente y la misma que, en la aplicación metodológica del protocolo para la validación de candidatos a término, discutiremos y ejemplificaremos a partir del corpus compilado para esta investigación. Una concepción que permitirá volver la mirada a esas otras unidades, comunes en el discurso cotidiano, pero olvidadas en el discurso especializado, contrario a la premisa ya establecida por Cabré en la cual ratifica los lenguajes de especialidad como un subconjunto de la lengua general (1998, p. 7) y, por ende, una razón más para no concordar con la falta de interés investigativo en su estudio. Al respecto, se presentaron las bases conceptuales en la construcción del protocolo para la caracterización de los términos (Quiroz & Burgos, en prensa), lo que constituiría una contribución en la definición operativa del término y el reconocimiento de sus características lingüísticas, *puerta* por la cual hemos ingresado en nuestros análisis, bases que igualmente permitirían caracterizar e indicar qué no es un término. Asimismo, en línea con nuestra mirada

integradora también presentamos un esbozo inicial sobre el carácter terminológico de los Sintagmas Nominales Especializados Extensos (SNEE), las metáforas y los fenómenos de reducción léxica (siglas) con carácter terminológico (§ 5.1.3).

En los capítulos siguientes (§§ 3-5), nos centramos en el español como nuestra lengua materna, recorriendo sus particularidades en la materialización del conocimiento universal. Una reflexión que nos permitió unir los dos lados del Océano Atlántico al lograr integrar en el corpus de trabajo una muestra importante de información a partir de los distintos países donde el español es lengua oficial y donde se contó con disponibilidad de datos para nuestro propósito de trabajo. Asimismo, hemos definido cinco áreas del conocimiento a partir de sus contrastes y, de esta manera, lograr una amplitud relevante en la horizontalidad terminológica hasta ahora no equiparable con ningún trabajo en la bibliografía de nuestra disciplina. Nos ubicamos en un plano descriptivo del fenómeno en cuanto los datos analizados nos permiten, precisamente, indagar de manera inicial en esta otra mirada del *término* como objeto de estudio de la terminología.

PARTE II:
METODOLOGÍA

DE ANTIOQUIA
ELLIN

3. Metodología: un corpus del español para la variación terminológica horizontal

> Although there certainly are questions of linguistic value
> that do not require inferential statistics to be answered,
> there is no denying that statistical techniques
> are valuable and often even necessary.
>
> (Larsson et al. 2022, p. 150).

Hasta ahora, los capítulos anteriores: 1 *Terminología: una mirada en conjunto* y 2 *El término: caracterización de la unidad de estudio*, nos han permitido establecer y delimitar nuestro campo de acción, esto es, una reflexión acerca del estado actual de la terminología en relación con la delimitación y caracterización del término como su objeto de estudio. Por un lado, los diferentes postulados teóricos y metodológicos a partir de las investigaciones a lo largo y ancho del espectro del español, como lengua que une los dos extremos del Atlántico, han permitido un diálogo común en torno a los lenguajes de especialidad y las particularidades de su materialización, a partir de los principios clásicos y modernos de la terminología. Por otro lado, los desafíos, de cara al futuro, en la búsqueda de precisiones que permitan ampliar la *mirada en conjunto* de nuestra disciplina y, por ende, el punto de partida para la presente tesis doctoral. Al respecto, hemos planteado dos hipótesis de trabajo en las cuales indagamos, en primer lugar, acerca del valor terminológico de todas las unidades léxicas con la intención de ampliar el foco de análisis hasta ahora orientado a las construcciones nominales y, en segundo lugar, observar, en las áreas del saber que componen el corpus, el comportamiento de las unidades terminológicas con la premisa que

estás varían dependiendo del campo de conocimiento y merecen un tratamiento diferencial contrario a la dinámica actual de la disciplina.

Al momento de escribir el presente trabajo, y teniendo en cuenta los manuales hasta ahora escritos, no se cuenta con una investigación que, haciendo uso de la lingüística computacional y de corpus, haya compilado un corpus multidisciplinar y que, geográficamente, contenga información de la mayoría de los países donde el español es lengua oficial, lo que constituiría, de entrada, un aporte a la terminología actual. Al respecto, concordamos, parcialmente, con Cabré en su postura sobre la jibarización de la terminología, debido al trabajo desde la lingüística computacional y el procesamiento del lenguaje, en cuanto a la observación de los términos en grandes volúmenes y su tratamiento como cadenas de caracteres (19 de noviembre 2020) y, por extensión, a otros análisis lingüísticos donde la observación de sus objetos de estudio puntuales parece tomar el mismo rumbo en cuanto a sus dinámicas de trabajo. Parcialmente, puesto que reconocemos que, de cierto modo, puede verse opacada la naturaleza de la terminología al entrar en contacto con estas otras disciplinas, pero también resaltamos las bondades de las tecnologías de la información al servicio del lenguaje, en cuanto han permitido tener acceso a mayor cantidad de datos que, sin duda, decantan en la posibilidad de ejemplificaciones más amplias de los fenómenos terminológicos. A su vez, también abogamos por un repensar esta simbiosis entre la terminología y lo computacional en cuanto a la necesidad de potenciar el análisis desde lo lingüístico más allá de solo la aplicación de modelos estadísticos que se quedan solo en la descripción de frecuencias y relaciones inferenciales, preocupación que Larsson et al. dejan entrever en su artículo, *On the status of statistical reporting versus linguistic description in corpus linguistics: A ten-year perspective*, al respecto del cambio significativo en el incremento del enfoque estadístico en las investigaciones de corte lingüístico (2022, p. 137).

Con relación al corpus de trabajo, un corpus especializado para la Variación Terminológica Horizontal (VaTeHo), luego de una serie de consideraciones de orden epistemológico, se definió trabajar con

un total de cinco áreas del conocimiento, a saber: medicina (Med), lingüística (Lin), veterinaria (Vet), computación (Com) y filosofía (Fil), con el propósito de abarcar, de manera horizontal, áreas a la vez distantes y afines en sus objetos de estudio. Además, se compiló un máximo de 10 artículos de investigación recuperados en revistas indexadas en los países donde el español es lengua oficial y se contó con disponibilidad de información. En este sentido, consideramos que nuestro corpus es relevante y constituye un aporte importante a los trabajos en terminología, toda vez que se logró compilar información de la mayoría de estos países. En este punto, importante cuestionar la forma como desde la Real Academia Española de la Lengua se ha propiciado un desbalance en el origen de los datos utilizados en recursos, bastante importantes para la investigación en español con fines tanto específicos como generales, como lo es el Corpus de Referencia del Español Actual (CREA) donde el 50 % de la información proviene de España mientras que el otro 50 % de la información se distribuye en las demás zonas lingüísticas tradicionales: andina, Antillas (caribeña), caribe continental, chilena, Estados Unidos, México y Centroamérica y Rio de la Plata (CREA, 2022b, parámetros habituales). No encontramos justificación ni desde lo teórico ni desde lo estadístico para esta decisión en la conformación del corpus CREA, más allá de una imposición de la variante peninsular sobre las demás.

En adelante, se presentará, en detalle, el camino metodológico que se ha recorrido de cara a la obtención de los resultados (§ 4) con los cuales ahondaremos en las premisas hasta ahora descritas con relación a las hipótesis y objetivos de trabajo en la posterior discusión de este trabajo de doctorado (§ 5). En primer lugar, se describe el enfoque de la investigación al igual que el alcance de nuestros análisis y la validez de nuestros datos. En segundo lugar, se detalla el corpus especializado VaTeHo a partir de: 1) los criterios para su construcción, 2) la adecuación de los datos, la estructuración de la base de datos y la codificación empleada en la marcación de la información y 3) los programas computacionales empleados en el procesamiento y extracción de la información. En tercer lugar, se presentará el componente metodológico del protocolo para la validación de candidatos a términos

(Quiroz & Burgos, en prensa) del cual ya se presentó su fundamentación teórica (§ 2.3) y, en cuarto lugar, una mirada general en la consolidación del informe escrito final de la presente investigación.

3.1 Conceptos generales en nuestra metodología de investigación

Nuestra investigación basada en corpus (*Corpus Based*) (Biber, 2015, pp. 196-201), con la cual aportamos en la construcción de este tipo de trabajos en nuestro contexto colombiano, donde son poco frecuentes en comparación con las dinámicas en otros escenarios investigativos a nivel internacional, se caracteriza a partir de dos elementos: en primer lugar, gracias al corpus especializado de trabajo, con el cual se logró hacer uso de las bondades de la lingüística computacional y de corpus, en la comprobación de las hipótesis de trabajo a través del cumplimiento de los objetivos planteados desde el inicio de la investigación lo que, en segundo lugar, nos permitió tener una mirada amplia del fenómeno aquí trabajado: el *término*. De este modo, esta investigación, dentro de los límites de una propuesta tipológica, es entendida como: *básica*, *empírica y descriptiva*.

En primer lugar, la concebimos como *básica* y *empírica* puesto que partimos de los fundamentos conceptuales en la comprensión y en el perfeccionamiento de las distintas técnicas y herramientas desde la lingüística computacional y de corpus para avanzar, sin mayores contratiempos, en la construcción y la compilación del corpus VaTeHo y su procesamiento a partir de la utilización de los recursos disponibles para estos propósitos, entre ellos: extractores terminológicos automáticos o semiautomáticos como *TermoStat, Multiterm, BC Term, One Click Terms, Terminology Extractor,* entre otros; procesadores o editores de texto como *EditPlus, Wordsmith, #Lancsbox, Simple Concordance Program* (SCP), *Nvivo*, entre otros; con los cuales se realizan diversidad de tareas entre las cuales se pueden listar: obtención de frecuencias, concordancias, correlaciones, listas de palabras clave, estructuras sintácticas, campos semánticos, y demás funciones de acuerdo a los

propósitos particulares de cada investigación lingüística y, por ende, la aplicación de métodos computacionales, estadísticos y lingüísticos a lo largo de los distintos estadios en el análisis del fenómeno estudiado.

En segundo lugar, es una investigación *descriptiva* toda vez que, haciendo uso de los datos compilados, nos permitimos analizar, caracterizar, describir, explicar e indagar sobre el concepto de término alrededor de la variación terminológica horizontal. En este sentido, al ser una investigación a partir de corpus también se logró ir más allá de las tendencias casuísticas de investigación, en gran medida, recurrentes en el contexto latinoamericano, al tener la posibilidad de ilustrar y ejemplificar con amplitud los puntos establecidos en cada una de las discusiones de la presente tesis doctoral. Además, gracias al acercamiento tanto deductivo como inductivo implementado según las necesidades de cada etapa de trabajo, nuestra investigación se aleja de las tendencias prescriptivistas comunes, por ejemplo, en los manuales de terminología hasta ahora publicados tanto en español como en otras lenguas (alemán, francés, catalán) y, por ende, una perspectiva descriptiva del fenómeno gracias a los datos de trabajo.

En cuanto a la concepción de corpus, la literatura disponible en el área es generosa en las conceptualizaciones. Autores como Leech (1992), Fillmore (1992), De Yzaguirre (1996), Biber et al. (1998), Martí & Llisterri (2000), Rojo (2002), Quiroz (2003), Baker et al. (2006), Parodi (2007, 2008), McEnery (2019), entre otros, incluyen en sus investigaciones los principios teórico-metodológicos de la lingüística computacional y de corpus afianzando sus análisis en la construcción de un campo de trabajo en conjunto con el lenguaje. Así, en primer lugar, para nuestra tesis de trabajo entendemos corpus en concordancia con la definición en Quiroz al concebirlo como una representación real del lenguaje a partir de textos tanto orales como escritos con finalidad de ser utilizados en estudios en las diferentes áreas del lenguaje (2003, p. 133). De esta manera, nuestro corpus de trabajo es una muestra real representativa de los saberes de cada una de las áreas analizadas, al menos en los datos compilados en nuestra tesis. En segundo lugar, al anclarnos en una metodología basada en

corpus, nos hemos permitido trabajar las hipótesis a partir de datos lingüísticos más allá de la introspección prototípica en nuestra área. Este enfoque nos permitió tener acceso a varias de las herramientas arriba mencionadas en una mirada amplia del fenómeno. Al respecto, en la fase inicial de la investigación se llevaron a cabo una serie de pruebas piloto con la finalidad de identificar, dentro de la amplia gama de herramientas para el procesamiento de lenguaje natural, cuáles respondían mejor de cara al objetivo de nuestro trabajo. Finalmente, el corpus VaTeHo[34] está disponible para su consulta ya sea con fines académico-investigativos o propósitos de documentación por parte de público especializado y no especializado.

3.2 Metodología y corpus para la variación terminológica horizontal (VaTeHo)

Esta tesis doctoral contó con cinco fases de trabajo: 1) concepción previa de la investigación, 2) construcción del corpus VaTeHo, 3) extracción de los candidatos a término, 4) consolidación de la muestra de trabajo, 5) análisis de datos y escritura de la tesis doctoral. Asimismo, cada una de estas etapas estuvo conformada por una serie subprocesos en atención a los objetivos de investigación propuestos. Así, podemos establecer, a partir de estas fases, dos macro etapas metodológicas: una primera orientada a la construcción del corpus (Fase 1, 2) y una segunda de cara a su utilización según los propósitos en el análisis del tema en cuestión y su posterior socialización en la escritura del documento final de esta tesis doctoral (Fase 3, 4 y 5). En la Figura 1, se ilustra el flujo de trabajo a partir de cada una de las fases y subprocesos de investigación. Posteriormente, se hará una descripción en detalle de cada una de las fases con la información más relevante que permita una mirada global y la consolidación general del ámbito metodológico de esta tesis doctoral.

[34] El corpus para la Variación Terminológica Horizontal (VaTeHo) construido en el desarrollo de esta tesis doctoral se encuentra alojado para su consulta en el servidor del Grupo de Investigación en Traducción y Nuevas Tecnologías (TNT), Escuela de idiomas, Universidad de Antioquia. Para más información, consúltese: http://grupotnt.udea.edu.co/VaTeHoVe/.

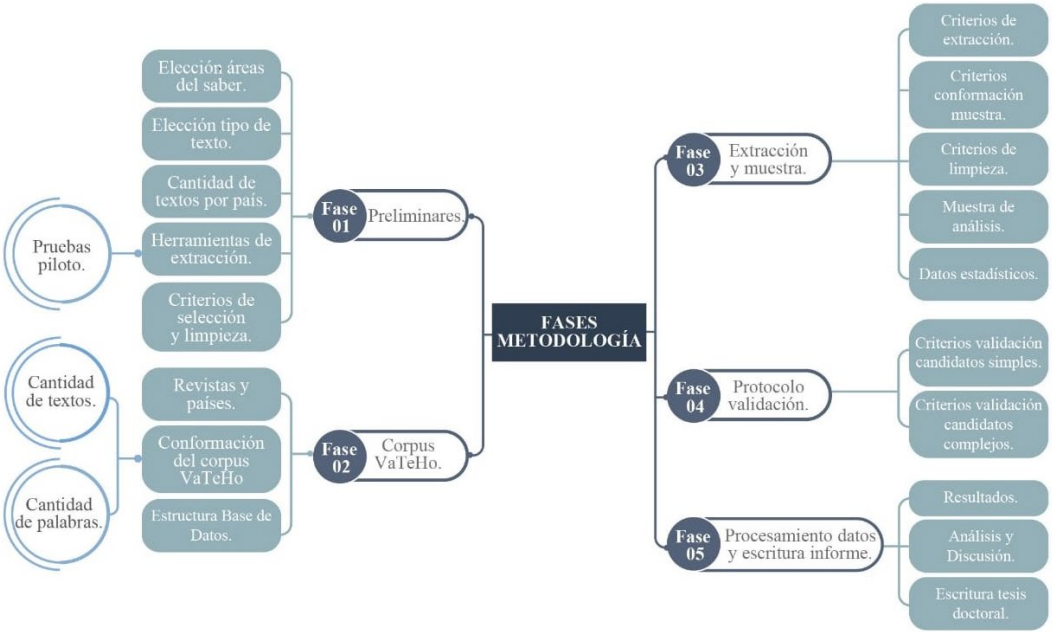


Figura 1. Fases metodología. Diseño propio.

3.2.1 Fase 01: Preliminares

Un reto importante en la construcción del corpus de trabajo consistió en la definición de los criterios para la compilación de la información, en gran medida por la necesidad de construir un recurso que nos permitiese observar el fenómeno de la variación terminológica, por un lado, a nivel horizontal en un amplio contraste de saberes y, por otro lado, en una cantidad justa, balanceada y representativa de la diversidad lingüística del español a nivel mundial. En consecuencia, las discusiones preliminares estuvieron orientadas a: 1) la elección de las áreas del saber que, finalmente, conforman el corpus VaTeHo, 2), el tipo de texto a trabajar, 3) la determinación de los países de los cuales se obtendrían los datos, 4) la cantidad de información a compilar y 5) las herramientas computacionales para la extracción y para el procesamiento de los candidatos a término.

3.2.1.1 Áreas del saber

Los trabajos en terminología han privilegiado el estudio de sus fenómenos, principalmente, en datos pertenecientes a la medicina y/o la economía, prueba de ello las investigaciones realizadas por Adelstein (2007), Quiroz (2008), Giraldo (2008), Casademont (2008), entre otros. De esta manera, en la compilación del corpus VaTeHo se tomó la decisión de incluir cinco áreas del saber en las cuales se pudiese observar de manera amplia el fenómeno de la variación terminológica en el eje horizontal y explorar otros campos del saber no incluidos, hasta el momento, en otras investigaciones, además de observarlos en un mismo trabajo, lo que representa un aporte a la construcción de la terminología desde elementos teórico-metodológicos y un recurso representativo para la consulta. Para la elección de estas áreas de trabajo se tuvo en cuenta tres criterios generales. En primer lugar, se debía incluir en la muestra final áreas de alto impacto para la sociedad y, por ende, áreas muy analizadas no solo desde la terminología sino también desde otras perspectivas de investigación. En este sentido, se definió una lista preliminar donde se incluyó *medicina*, *derecho*, *economía*, *biología* y *veterinaria*. En segundo lugar, áreas de bajo impacto y, a su vez, menos analizadas. Aquí, se listaron, inicialmente, *filosofía*, *lingüística*, *música* y *sociología*. En tercer lugar, áreas de interés social, esto es, áreas de importancia social, en la cuales se incluyeron, preliminarmente, *computación* y *medio ambiente*.

De esta manera, se conformó el listado con: *medicina, lingüística, veterinaria, computación* y *filosofía*. La decisión de incluir solo cinco áreas obedece a criterios de orden metodológico relacionados con el tiempo disponible para la investigación y la conclusión de esta. Asimismo, una vez definidas las áreas de trabajo, se procedió con la elección de las subáreas de especialidad a trabajar. Para lo cual se establecieron las palabras clave para la búsqueda de información en los distintos recursos utilizados para la compilación de los textos a trabajar. Así, en medicina se utilizó *oncología* y/o *cáncer*, en lingüística *léxico, lexicografía* y/o *lexicología*, en veterinaria *ganadería*, en computación *software* y en filosofía *epistemología* y/o *conocimiento*.

3.2.1.2 Tipo de texto

La decisión del tipo de texto a utilizar para la compilación de la información corresponde al objeto de estudio de nuestro trabajo de investigación, esto es, el *término*. Así, al elegir el artículo de investigación con la estructura prototípica de Introducción, Métodos, Resultados y Discusión (IMRyD) garantizamos que los candidatos a término correspondían a contextos de comunicación con un alto nivel de especialidad. Al respecto, la clasificación según el nivel de especialidad en Hoffman (1998, pp. 72-82) permite reafirmar esta postura acerca de la presencia de unidades terminológicas en contextos de comunicación entre expertos, situación prototípica de este tipo de textos y el canal de publicación de estos: las revistas indexadas. Además, el artículo de investigación, al concebirse bajo la clasificación de género científico, permite la interacción con lectores especializados y, por ende, la constitución de las comunidades científicas y académicas conforme los distintos campos de conocimiento (Frías, 1996, cap. 4). Por otro lado, la elección de este tipo de textos también nos permitió contar con una validación inicial a partir de las dinámicas propias de las revistas académicas, esto es, la evaluación por pares de los artículos garantiza, por sí misma, una primera validación por parte de expertos en cada una de las áreas observadas al igual que la validación por la comunidad discursiva, aspecto que se complementará con la aplicación de los criterios establecidos en el protocolo validación en nuestro trabajo (Quiroz & Burgos, en prensa).

3.2.1.3 Países y distribución del corpus

Otro de los aportes del corpus VaTeHo a los estudios en terminología está relacionado con la elección de los países de los cuales se obtuvo la información, siendo un corpus representativo *del español* y *no en español*, es decir, la conformación final del corpus permitió incluir datos provenientes de la mayor parte de los países donde el español es lengua oficial. Es importante señalar que en algunos de estos países no se logró recuperar información en las áreas de conocimiento definidas, específicamente: Guinea Ecuatorial y Sahara Occidental.

Asimismo, en el caso de Estados Unidos y Puerto Rico, se decidió no incluirlos en la base de datos, en principio, por el contacto del español con una variante dominante, en este caso, el inglés, además de no contar con el estatus de lengua oficial. En este punto, también se tuvo en cuenta las particularidades de otros países en los cuales se habla español, pero su estatus aún no está oficialmente definido y, por lo tanto, también se dejaron por fuera, a saber: Filipinas, Belice, Trinidad y Tobago.

Con relación a la distribución y cantidad de los datos a compilar, se discutieron varias posibilidades con miras a garantizar un balance de la información con la cual cada país aportará en igual medida a la compilación final. En este sentido, se consideró asignar la cantidad de artículos por país teniendo en cuenta: 1) la población, 2) la cantidad de universidades, 3) la cantidad de grupos de investigación en las áreas de conocimiento del corpus, 4) la cantidad de revistas indexadas y 5) la cantidad de doctores en las áreas de conocimiento del corpus. Sin embargo, revisadas y analizadas cada una de estas opciones, el resultado en la aplicación generaba un desbalance toda vez que siempre habría países con mayor porcentaje lo que contrastaba con nuestro propósito y concordaba con la postura de una o varias variantes dominantes como se mencionó para el CREA. Finalmente, se decidió construir el corpus a partir de la misma cantidad de textos por cada país y de esta manera garantizar, en principio, un aporte de datos igual por cada país.

3.2.1.4 Herramientas para la extracción y procesamiento de los datos

Dentro de las ventajas de la lingüística computacional y de corpus sobresale la posibilidad de tener una gran variedad de herramientas disponibles para los distintos propósitos de investigación[35]. Para nuestro particular, fue necesario la utilización de un extractor de terminología de los cuales, en la actualidad, hay diversidad de opciones tanto para trabajo en la nube como instalables en computador.

[35] Para un panorama más amplio de herramientas para el procesamiento de datos lingüísticos, consúltese https://linguistlist.org/sp/software-index.cfm.

La pregunta central en este punto de la investigación se orientó a determinar cuál de las opciones disponibles era la más indicada. Al respecto, se llevaron a cabo una serie de siete ensayos piloto con el fin de determinar la eficacia de varias herramientas según los resultados de las extracciones realizadas. Para este proceso, se compiló una muestra de trabajo conformada por un total de 162 textos en medicina (oncología) y 82 en lingüística (léxico) según los mismos criterios definidos para la compilación final (§ 3.2.1.5). Igualmente, se realizó el proceso de adecuación (§ 3.2.1.5) de los textos antes de proceder con la extracción en los programas a evaluar. No se balanceó la muestra puesto que el procesamiento se dio por separado toda vez que el propósito consistió en probar los recursos. Durante las pruebas piloto se utilizaron las siguientes herramientas:

En primera instancia, *BC_Term*[36] (2014), programa autoría del Profesor Diego Burgos de la Universidad *Wake Forest*. Este sistema permite la extracción automática de términos tanto en español como en inglés, a partir de una serie de reglas en torno al núcleo nominal. El resultado de la extracción arrojo una cantidad considerable de datos que no corresponden directamente con los propósitos de nuestro trabajo por lo que la limpieza de los datos obtenidos representaría un esfuerzo no menor de cara al avance del cronograma de trabajo y, por ende, se descartó la implementación de esta herramienta en nuestro trabajo. En segunda instancia, se utilizó *Multiterm*[37] (2022), propiedad de la compañía para servicios lingüísticos *RWS Group*. Esta herramienta permite el procesamiento de terminología a partir de corpus de referencia. Por ende, más que un extractor, en el estricto sentido de la palabra, esta herramienta cumple una función como gestor de terminología por lo cual no cumple con nuestro objetivo. En tercera instancia, *TBXTools*[38] (2015), creado por los profesores Antoni Oliver y Mercé Vàsquez de la Universidad Abierta de Cataluña. A diferencia del *BC_Term*, este extractor utiliza medidas estadísticas para la recuperación de los términos a partir de *Scripts* escritos en *Python*

[36] Para más información, consúltese: http://grupotnt.udea.edu.co/BC-Term/.

[37] Para más información, consúltese: https://www.trados.com/products/multiterm-desktop/.

[38] Para más información, consúltese: https://sourceforge.net/projects/tbxtools/.

unido al uso del *Natural Language Toolkit* (NLTK). Igualmente, la extracción arrojo un ruido[39] considerable lo que generaría el mismo problema en relación con la limpieza.

Por último, en cuarto lugar, se llevó a cabo el piloto con la herramienta *TermoStat* (2022), creada por el Profesor Patrick Drouin de la Universidad de Montreal (2003, 2010). Esta opción fue la elegida para nuestro trabajo, entre otras razones, porque la extracción, a pesar de tener ruido considerable, permitió un alto porcentaje de acierto en los términos recuperados. Este recurso emplea para su extracción, además de reglas sintácticas y medidas estadísticas, corpus especializados y no especializados de contraste. Además, permite realizar extracciones en cinco (5) idiomas, entre ellos, el español. Otro dato importante de *TermoStat*, tiene que ver con la cantidad de citaciones disponibles en Google Académico. Así, al momento de la consulta, esta herramienta contaba con alrededor de 945 citaciones según las métricas de este portal. Por último, es importante aclarar que, aunque la documentación del programa indica que es un extractor de términos, bajo la óptica de nuestra investigación, estos serán considerados como candidatos a término, puesto que según los resultados en la aplicación del protocolo para la validación de nuestra tesis (Quiroz & Burgos, en prensa) se confirmarán o no como tal.

En esta misma línea, se utilizaron los programas *Antconc*[40] (2022), autoría del profesor Laurence Anthony de la Universidad de Waseda; *#Lancsbox*[41] (2020), diseñado por los profesores Vaclac Brezina, Pierre Weill-Tessier y Tony McEnery de la Universidad de Lancanster; *WordSmith*[42] (2022), creado por Mike Scott de la Universidad de

[39] El concepto de ruido y silencio corresponden a índices en la extracción de datos. El primero, corresponde a la cantidad de falsos términos respecto al total de términos predichos, mientras que el segundo corresponde a los términos no reconocidos del total de términos predichos (Porta, 2006, pp. 5-6).

[40] Para más información, consúltese: https://www.laurenceanthony.net/resume.html.

[41] Para más información, consúltese: http://corpora.lancs.ac.uk/lancsbox/.

[42] Para más información, consúltese: https://www.lexically.net/wordsmith/.

Oxford; *EditPlus*[43], de ES-Computing; Excel[44] (2022) y Word[45] (2022), de Microsoft; *Simple Concordace Program* (SCP)[46] (1997), diseñado por Alan Reed; *Voyant Tools*[47] (2003), creado por Stéfan Sinclair y Geoffrey Rockwell; *Word Cloud Generator*[48] (2022), autoría de Jason Davies; para distintos procesamientos de los datos de trabajo, entre los cuales se destacan: almacenamiento y organización de la base de datos, búsqueda de concordancias y contextos de aparición de los términos, identificación de nubes conceptuales, entre otros, procesos que se verán reflejados, implícita o explícitamente, en las secciones de Resultados (§ 4) y Discusión (§ 5).

3.2.1.5 Criterios de selección y limpieza de los datos

Luego de establecidas las áreas del conocimiento y el tipo de texto para la conformación de nuestro corpus, se procedió a establecer los parámetros para la compilación de los datos. En este sentido, la recolección de los artículos se realizó de forma canónica, es decir, el corpus fue constituido de forma manual a razón de las particularidades en la selección que impedían un buen rendimiento de las búsquedas automáticas o semiautomáticas. En este punto, se consultó con el ingeniero Antonio Tamayo[49] la posibilidad de generar un *script* en lenguaje de programación que permitiera este proceso. Las conclusiones de las indagaciones al respecto decantaron por el proceso manual el cual permitiría un mejor aprovechamiento del tiempo y el recurso humano disponible para tal labor. Los criterios para la compilación del corpus se definieron teniendo en cuenta la

43 Para más información, consúltese: https://www.editplus.com/.

44 Para más información, consúltese: https://www.microsoft.com/es-es/microsoft-365/excel.

45 Para más información, consúltese: https://www.microsoft.com/es-es/microsoft-365/word.

46 Para más información, consúltese: http://www.textworld.com/scp/#Description.

47 Para más información, consúltese: https://voyant-tools.org/docs/.

48 Para más información, consúltese: https://www.jasondavies.com/wordcloud/about/.

49 El ingeniero Antonio Tamayo es Magíster en Ingeniería por la Universidad de Antioquia y Doctor en Inteligencia Artificial y Procesamiento de Lenguaje Natural por el Instituto Politécnico Nacional de México. Asimismo, es miembro del Grupo de Investigación en Traducción y Nuevas Tecnologías (TNT) de la Escuela de Idiomas, Universidad de Antioquia, Colombia. Para más información sobre el perfil académico y profesional del profesor Tamayo, consúltese: https://scholar.google.com/citations?user=643heS8AAAAJ&hl=es.

intención de lograr construir un corpus del español y, por lo tanto, estos se orientaron, en esencia, a garantizar que la información recuperada correspondiera con esta lengua. Es importante precisar que cada artículo compilado está publicado en revistas de cada uno de los países, es decir, la información compilada en Colombia (COL) corresponde a revistas científicas colombianas y así para cada variedad dialectal.

De este modo, la recolección de los datos correspondería, en primera instancia, a la recopilación de artículos de revistas científicas en cada una de las cinco áreas de conocimiento y en cada uno de los países donde el español es lengua oficial, a partir de las palabras clave definidas para cada subcorpus (§ 3.2.1.1). En segunda instancia, estar escrito por hablantes nativos del español. Para ello se identificó que al menos uno de los autores del artículo fuese hablante nativo de la variedad dialectal y/o en su defecto el manual de estilo de la editorial de publicación o la intervención en el proceso de redacción del artículo de alguna universidad del país de compilación. En este punto, nos encontramos, por ejemplo, con artículos publicados por hablantes de alguna variedad dialectal en específico, pero cuya publicación se realizó en un país distinto al nativo y, por lo tanto, no fueron tenidos en cuenta en la conformación final corpus. En tercera instancia, los documentos deberían estar en formato digital y ser de libre acceso tanto para la consulta como para la descarga. En este punto, cada archivo está almacenado en formato *.pdf* y *.txt* en función de las configuraciones básicas necesarias para el posterior procesamiento de la información en *TermoStat* y demás procesadores de información utilizados (§ 3.2.1.4). Por último, en cuarta instancia, no se estableció ningún límite de palabras para la elección y posterior compilación de los artículos.

El último elemento en las decisiones preliminares consistió en los criterios para la adecuación inicial de los artículos compilados. En este sentido, se procuró omitir la mayor cantidad de información contenida en estos que no estuviese directamente relacionada con los objetivos de investigación. De este modo, se eliminaron de cada uno de los artículos la información correspondiente a la revista y

los autores, los resúmenes y las palabras clave en lenguas diferentes al español, los agradecimientos y las declaraciones de conflicto de interés, la lista de referencias bibliográficas y demás elementos según las particularidades de cada artículo. La compilación final del corpus, a partir de las consideraciones teórico-metodológicas arriba descritas, permite definirlo como un corpus especializado del español.

3.2.2 Fase 02: Corpus especializado del español para el estudio de la variación terminológica horizontal (VaTeHo)

De acuerdo con las diferentes aristas hasta ahora descritas en las secciones anteriores, el corpus especializado del español para la Variación Terminológica Horizontal es el resultado de la aplicación de consideraciones de orden teórico y metodológico que permitieron por un lado, conservando los límites de la prudencia, la construcción de un recurso novedoso y representativo para la terminología toda vez que abarca cinco áreas de conocimiento más allá de las tradicionalmente estudiadas en diversas investigaciones y, por otro lado, un corpus del español balanceado conforme la realidad lingüística de este. En adelante, se detallan los elementos mínimos para la comprensión de la estructura del corpus al igual que la gestión documental de la base de datos y la información relacionada con este.

3.2.2.1 Conformación final del corpus VaTeHo

El corpus VaTeHo está constituido por un total de 538 artículos científicos distribuidos en cinco áreas del saber y compilados en 19 países de habla española. Se estableció que por país se compilaría un total de diez artículos por cada área. Sin embargo, al momento de realizar la búsqueda de información, la realidad en investigación de algunas regiones no permitió recuperar la totalidad de la información e incluso en algunas áreas no fue posible obtener datos. De este modo, nuestro corpus también es un recurso en el cual la disponibilidad de la información se constituye como un criterio de construcción. Igualmente, a pesar de la poca o ausencia de datos, sí fue posible obtener datos de las regiones lingüísticas tradicionales del español,

lo que permitió observar el fenómeno en cuestión a lo largo y ancho del espectro de esta lengua. Asimismo, con relación a las revistas científicas utilizadas para la obtención de los artículos, estas fueron ubicadas a partir de una búsqueda de información en bases de datos y en portales de universidades donde se oferta formación en las áreas de conocimiento del corpus. En total se emplearon 197 revistas distribuidas por áreas y, por ende, por países de la siguiente manera: 94 en medicina, 24 en lingüística, 36 en veterinaria, 18 en computación y 25 en filosofía. En la Tabla 2, se presenta la información sobre la conformación final del corpus. De este modo, se obtuvo la cantidad de artículos por área en cada país y la totalidad de artículos por país al igual que la cantidad por área.

Tabla 2. Relación de artículos por país y por área.

País	Medicina	Lingüística	Veterinaria	Computación	Filosofía	Total art. país
Argentina	10	10	10	1	9	40
Bolivia	10	0	10	0	1	21
Chile	10	10	5	10	7	42
Colombia	10	10	10	10	9	49
Costa Rica	10	10	10	10	10	48
Cuba	10	0	0	10	8	28
Ecuador	10	0	2	10	8	29
El Salvador	0	0	0	0	2	2
España	10	10	3	10	8	41
Guatemala	2	0	0	0	9	11
Guinea Ecuatorial	0	0	0	0	0	0
Honduras	10	0	0	0	1	11
México	10	10	10	10	9	49
Nicaragua	2	0	0	2	0	3
Panamá	10	0	0	10	3	23
Paraguay	10	1	0	0	0	11

País	Medicina	Lingüística	Veterinaria	Computación	Filosofía	Total art. país
Perú	10	10	7	10	10	47
República Dominicana	7	0	0	0	5	12
Sahara Occidental	0	0	0	0	0	0
Uruguay	10	4	1	0	10	25
Venezuela	10	7	5	10	10	42
Total art. área	161	82	73	103	119	538

Asimismo, en la Tabla 3, se relaciona la cantidad de *tokens* (entendidos como cualquier unidad lexical sin incluir la puntuación) en la misma lógica de la Tabla 2. Un total de 3.057.973 tokens constituyen nuestro corpus. En cuanto a la representatividad del corpus, aunque es pequeño en comparación con otros recursos, de corte general, como: el *Longman/Lancaster English Language Corpus*[50] con 30 millones de palabras, el *British National Corpus* (BNC)[51] con más de 100 millones de palabras, el *Corpus of Contemporary American English*[52] (COCA) con más de 1 billón de palabras y, en español, el *Corpus de Referencia del Español*[53] (CREA) con más de 170 millones de formas, el *Corpus del Español*[54] con más de 2 billones de palabras y el *Sketch Engine*[55] con un corpus de más de 9 billones de palabras y una versión actualizada, aún en modo prueba, con más de 17 billones de palabras; nuestro corpus al ser especializado es considerado representativo. Además, al hablar de representatividad no solo se debe tener en cuenta el tamaño del recurso sino también la capacidad del corpus de dar cuenta del

50 Para más información, consúltese: http://www.pearsonlongman.com/dictionaries/corpus/lancaster.html.

51 Para más información sobre este recurso, consúltese: https://www.english-corpora.org/bnc/.

52 Para más información sobre este recurso, consúltese: https://www.english-corpora.org/coca/.

53 Para más información sobre este recurso, consúltese: https://www.rae.es/banco-de-datos/crea.

54 Para más información sobre este recurso, consúltese: https://www.corpusdelespanol.org/compare_larger.asp.

55 Para más información sobre este recurso, consúltese: https://www.sketchengine.eu/corpora-and-languages/corpus-list/.

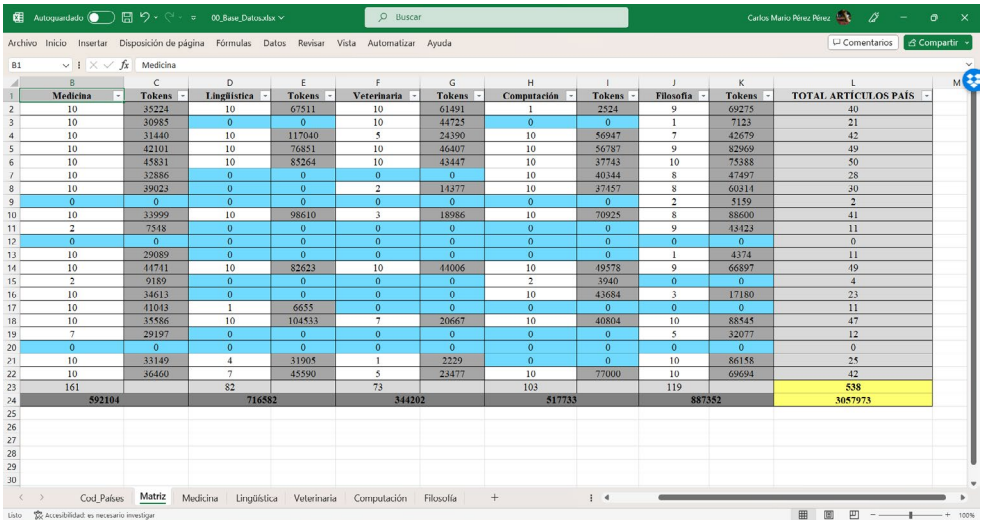
mayor rango de variabilidad en la población analizada, entendida esta desde la óptica de la estadística y, por ende, la posibilidad de observar este criterio tanto desde lo situacional como desde lo lingüístico, como factores determinantes de la representatividad en el diseño del corpus (Biber, 1993, p. 243).

Tabla 3. Relación de tokens por país y por área

País	Medicina	Lingüística	Veterinaria	Computación	Filosofía	Total Tokens país
Argentina	35224	67511	61491	2524	69275	236025
Bolivia	30985	0	44725	0	7123	82833
Chile	31440	117040	24390	56947	42679	272496
Colombia	42101	76851	46407	56787	82969	305115
Costa Rica	45831	85264	43447	37743	75388	287673
Cuba	32886	0	0	40344	47497	120727
Ecuador	39023	0	14377	37457	60314	151171
El Salvador	0	0	0	0	5159	5159
España	33999	98610	18986	70925	88600	311120
Guatemala	7548	0	0	0	43423	50971
Guinea Ecuatorial	0	0	0	0	0	0
Honduras	29089	0	0	0	4374	33463
México	44741	82623	44006	49578	66897	287845
Nicaragua	9189	0	0	3940	0	13129
Panamá	34613	0	0	43684	17180	95477
Paraguay	41043	6655	0	0	0	47698
Perú	35586	104533	20667	40804	88545	290135
República Dominicana	29197	0	0	0	32077	61274
Sahara Occidental	0	0	0	0	0	0
Uruguay	33149	31905	2229	0	86158	153441
Venezuela	36460	45590	23477	77000	69694	252221
Total Tokens área	592104	716582	344202	517733	887352	3.057.973

3.2.2.2 Base de datos y gestión documental

La parte final de la Fase 02 consistió en la conformación de la base de datos y la gestión documental, entendida esta como las decisiones relacionadas con la codificación de cada uno de los archivos compilados y su almacenamiento en las distintas carpetas y subcarpetas de trabajo. Además, todos los archivos concernientes a la base de datos fueron alojados tanto en *Google Drive* como en *Dropbox* de manera que se tuviese un soporte de la información en caso de alguna inconsistencia de estos servidores. En el proceso de estratificación y organización de la información se utilizó, principalmente, el programa *Excel* puesto que este facilitó la ordenación no solo de los metadatos relacionados con cada artículo, sino que también permitió una integración directa con los archivos arrojados por *TermoStat* con los resultados de las extracciones. A su vez, durante los distintos momentos de trabajo con los datos, se fueron asignando una serie de colores según era necesario resaltar algún tipo de información puntual, sin embargo, no se hará una descripción en detalle de este elemento puesto que no refleja ningún tipo de análisis o resultado relevante, solo se utilizó como ayuda visual para el manejo puntual de la información, contrario a lo propio en la sección de Resultados (§ 4) donde los colores utilizados sí pueden marcar un punto enfoque necesario para el posterior análisis y discusión de los resultados (§ 5). Por un lado, la Figura 2, ilustra la matriz general del corpus, en la cual se compila la información sobre el total de este y los datos sobre los artículos y la cantidad de *tokens*.

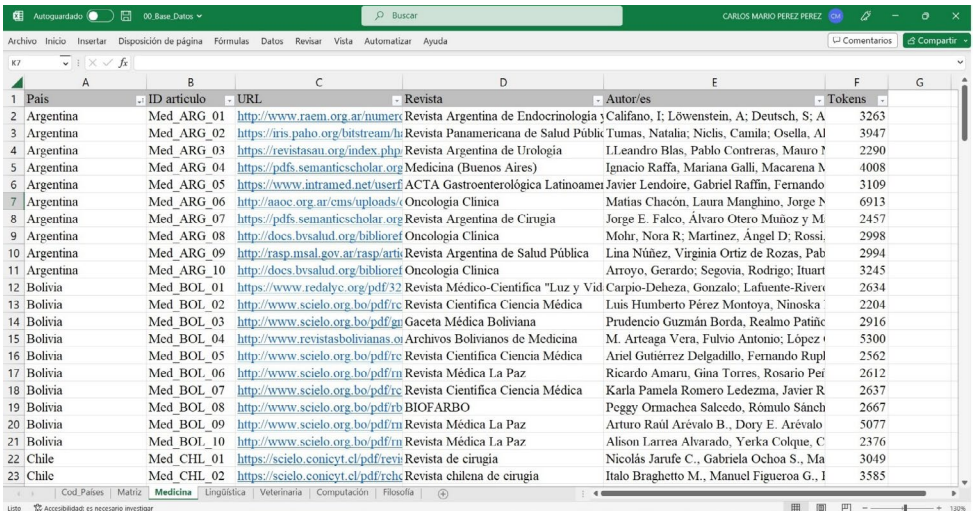


Medicina	Tokens	Lingüística	Tokens	Veterinaria	Tokens	Computación	Tokens	Filosofía	Tokens	TOTAL ARTÍCULOS PAÍS
10	35224	10	67511	10	61491	1	2524	9	69275	40
10	30985	0	0	10	44725	0	0	1	7123	21
10	31440	10	117040	5	24390	10	56947	7	42679	42
10	42101	10	76851	10	46407	10	56787	9	82969	49
10	45831	10	85264	10	43447	10	37743	10	75388	50
10	32886	0	0	0	0	10	40344	8	47497	28
10	39023	0	0	2	14377	10	37457	8	60314	30
0	0	0	0	0	0	0	0	2	5159	2
10	33999	10	98610	3	18986	10	70925	8	88600	41
2	7548	0	0	0	0	0	0	9	43423	11
0	0	0	0	0	0	0	0	0	0	0
10	29089	0	0	0	0	0	0	1	4374	11
10	41741	10	82623	10	44006	10	49578	9	66897	49
2	9189	0	0	0	0	2	3940	0	0	4
10	34613	0	0	0	0	10	43684	3	17180	23
10	41043	1	6655	0	0	0	0	0	0	11
10	35586	10	104533	7	20667	10	40804	10	88545	47
7	29197	0	0	0	0	0	0	5	32077	12
0	0	0	0	0	0	0	0	0	0	0
10	33149	4	31905	1	2229	0	0	10	86158	25
10	36460	7	45590	5	23477	10	77000	10	69694	42
161		82		73		103		119		538
592104		716582		344202		517733		887352		3057973

Figura 2. Matriz Base de Datos VaTeHo.

Por otro lado, la Figura 3, da cuenta de la organización de los artículos con la información correspondiente a los metadatos en la identificación de cada uno de estos, a saber: país, identificación (ID) del artículo, URL de portal *Web* donde se recuperó el archivo, revista, autores y cantidad de tokens. La misma información se presenta por cada una de las áreas trabajadas y se almacenó en pestañas diferenciadas. En este sentido, para nombrar cada archivo se siguió la siguiente estructura: *área* (tres primeras letras, mayúscula inicial), *código país* (según la Norma ISO 3166-1, codificación de tres caracteres, mayúscula sostenida), el *número secuencial* de cada artículo (número arábigo de dos dígitos) y la *extensión del archivo*, como se evidencia en los siguientes ejemplos: Med_ARG_01.txt, Lin_COL_02.txt, Vet_CRI_03.txt. Por último, la jerarquización del contenido en cada carpeta y subcarpeta corresponde a la lógica en el avance de las Fases de la investigación. De este modo, la base de datos se divide en dos carpetas principales: 1) la del corpus y 2) la de los datos de procesamiento. Todos los archivos se nombraron utilizando guion bajo entre las distintas etiquetas. Además, la diferencia entre los archivos almacenados en la carpeta UTF-8 y UNICODE corresponde exclusivamente a la codificación del archivo dependiendo de la herramienta utilizada, así, por ejemplo,

para el procesamiento de los datos en *TermoStat* es necesario utilizar la codificación UTF-8. En 9, se listan algunos ejemplos de la denominación de estos elementos:



	A	B	C	D	E	F	G
1	País	ID artículo	URL	Revista	Autor/es	Tokens	
2	Argentina	Med_ARG_01	http://www.raem.org.ar/numero	Revista Argentina de Endocrinologia y	Califano, I; Löwenstein, A; Deutsch, S; A	3263	
3	Argentina	Med_ARG_02	https://iris.paho.org/bitstream/hi	Revista Panamericana de Salud Públic	Tumas, Natalia; Niclis, Camila; Osella, Al	3947	
4	Argentina	Med_ARG_03	https://revistasau.org/index.php/	Revista Argentina de Urologia	LLeandro Blas, Pablo Contreras, Mauro I	2290	
5	Argentina	Med_ARG_04	https://pdfs.semanticscholar.org	Medicina (Buenos Aires)	Ignacio Raffa, Mariana Galli, Macarena M	4008	
6	Argentina	Med_ARG_05	https://www.intramed.net/userfi	ACTA Gastroenterológica Latinoamer	Javier Lendoire, Gabriel Raffin, Fernando	3109	
7	Argentina	Med_ARG_06	http://aaoc.org.ar/cms/uploads/d	Oncologia Clinica	Matias Chacón, Laura Manghino, Jorge M	6913	
8	Argentina	Med_ARG_07	https://pdfs.semanticscholar.org	Revista Argentina de Cirugia	Jorge E. Falco, Álvaro Otero Muñoz y M	2457	
9	Argentina	Med_ARG_08	http://docs.bvsalud.org/biblioref	Oncologia Clinica	Mohr, Nora R; Martínez, Ángel D; Rossi,	2998	
10	Argentina	Med_ARG_09	http://rasp.msal.gov.ar/rasp/artic	Revista Argentina de Salud Pública	Lina Núñez, Virginia Ortiz de Rozas, Pab	2994	
11	Argentina	Med_ARG_10	http://docs.bvsalud.org/biblioref	Oncologia Clinica	Arroyo, Gerardo; Segovia, Rodrigo; Ituart	3245	
12	Bolivia	Med_BOL_01	https://www.redalyc.org/pdf/32	Revista Médico-Cientifica "Luz y Vid	Carpio-Deheza, Gonzalo; Lafuente-Rivero	2634	
13	Bolivia	Med_BOL_02	http://www.scielo.org.bo/pdf/rc	Revista Cientifica Ciencia Médica	Luis Humberto Pérez Montoya, Ninoska	2204	
14	Bolivia	Med_BOL_03	http://www.scielo.org.bo/pdf/gm	Gaceta Médica Boliviana	Prudencio Guzmán Borda, Realmo Patiño	2916	
15	Bolivia	Med_BOL_04	http://www.revistasbolivianas.or	Archivos Bolivianos de Medicina	M. Arteaga Vera, Fulvio Antonio; López	5300	
16	Bolivia	Med_BOL_05	http://www.scielo.org.bo/pdf/rc	Revista Cientifica Ciencia Médica	Ariel Gutiérrez Delgadillo, Fernando Rupl	2562	
17	Bolivia	Med_BOL_06	http://www.scielo.org.bo/pdf/rm	Revista Médica La Paz	Ricardo Amaru, Gina Torres, Rosario Per	2612	
18	Bolivia	Med_BOL_07	http://www.scielo.org.bo/pdf/rc	Revista Cientifica Ciencia Médica	Karla Pamela Romero Ledezma, Javier R	2637	
19	Bolivia	Med_BOL_08	http://www.scielo.org.bo/pdf/rb	BIOFARBO	Peggy Ormachea Salcedo, Rómulo Sánch	2667	
20	Bolivia	Med_BOL_09	http://www.scielo.org.bo/pdf/rm	Revista Médica La Paz	Arturo Raúl Arévalo B., Dory E. Arévalo	5077	
21	Bolivia	Med_BOL_10	http://www.scielo.org.bo/pdf/rm	Revista Médica La Paz	Alison Larrea Alvarado, Yerka Colque, C	2376	
22	Chile	Med_CHL_01	https://scielo.conicyt.cl/pdf/revi	Revista de cirugía	Nicolás Jarufe C., Gabriela Ochoa S., Ma	3049	
23	Chile	Med_CHL_02	https://scielo.conicyt.cl/pdf/rchc	Revista chilena de cirugía	Italo Braghetto M., Manuel Figueroa G., I	3585	

Figura 3. Campos base de datos VaTeHo.

9. **Carpeta raíz corpus:** Corpus_Horizontal.

   Carpetas áreas: 01_Med, 02_Lin, 03_Vet, 04_Com, 05_Fil.

   Carpeta raíz datos: Datos_Procesamiento.

   Archivos datos procesamiento: Med_All_ TermoStat_Estadistica.txt, Lin_COL_All_BC_Term.txt, Com_ESP_All_MultiTerm.txt.

3.2.3 Fase 03: Extracción de los candidatos a término y muestra de trabajo

Una vez consolidado el corpus de trabajo, se procedió con la definición de los parámetros de la extracción de los candidatos a término, siendo este proceso clave en la obtención de la información para nuestros análisis. De este modo, como ya se estableció, se utilizó el programa *TermoStat* para este fin (Drouin, 2003, 2010). Asimismo,

se recurrió a literatura disponible en el área para identificar otras posibles estructuras sintácticas con posibilidad de ser consideradas como candidatas a término. De esta manera, los trabajos de Folguerà (2002), Lorente (2002), Bevilacqua (2004), Boyan (2005), Quiroz (2006, 2008), Giraldo (2008), Mira (2009, 2010), Sanz (2011, 2012), Tercedor et al. (2012), Burgos (2014), Rojas (2014), Quiroz & Arroyave (2014), Lopera (2016), Patiño (2017), Vásquez (2017), Rivas (2017), Galeano (2017), Zuluaga et al. (2017, 2021), Tamayo (2018), Pérez-Pérez (2018), Pérez-Pérez et al. (2020), Rivas et al. (2021); sirvieron como referencia según las particularidades de cada una de las unidades léxicas exploradas en estos trabajos.

Para la extracción en *TermoStat* fue necesario configurar cada subcorpus de manera que cumpliesen con los requisitos obligatorios del sistema para el procesamiento de la información, en específico, el programa solo acepta un único documento en formato de texto (.txt) por proceso, lo que no nos permitió medir la dispersión. En este sentido, se procedió a conformar un archivo de texto donde se agruparon todos los artículos en limpio (véase § 3.2.1.5) por cada una de las áreas de trabajo. De este modo, se crearon cinco archivos que conforman cada uno de los subcorpus de trabajo. En general, la interfaz de trabajo en *TermoStat* es bastante amigable con el usuario, lo que facilita el montaje, procesamiento, obtención y posterior descarga de la información. En este sentido, de los cinco bloques de información que arroja el sistema, a saber: 1) lista de términos, 2) nube semántica, 3) estadística, 4) estructuras sintácticas y 5) bigramas, que a su vez ofrecen otros análisis más en detalle, como son la recuperación de contextos y concordancias en 4, aunque se recuperó toda esta información por cada subcorpus, nos centramos en 1 y 3, principalmente. Las demás posibilidades se utilizaron, parcialmente, para completar algunos elementos de nuestros análisis. En total se extrajeron un total aproximado de 37.055 candidatos a término que conforman nuestra población de datos para los posteriores análisis. Asimismo, se estableció como parámetro para la selección de la muestra un nivel de confianza del 95 %, con un intervalo de confianza de +/- 4 % y homogeneidad del 0,5 %.

En línea con lo anterior, en aras de un mejor entendimiento de cómo funciona *TermoStat*, debe tenerse presente que la herramienta hace uso de una serie de medidas estadísticas orientadas al reconocimiento de la frecuencia de los candidatos a término a partir de los textos sometidos al proceso de extracción en contraste con los corpus de referencia, tanto especializados como generales, utilizados por el sistema. El proceso de extracción se divide en tres etapas: 1) *etiquetado del texto con TreeTager*[56] para desambiguar sintácticamente las palabras del texto, 2) la *extracción de las estructuras sintácticas* a partir de un conjunto de reglas predeterminadas por el sistema y, 3) la ponderación y la selección de los *candidatos a términos* teniendo en cuenta su frecuencia a partir de una puntuación asignada por el sistema en relación con el umbral de aceptabilidad establecido (Drouin, 2010). Asimismo, en relación con las medidas estadísticas empleadas por *TermoStat*, el sistema emplea una serie de índices y pruebas (índice de frecuencia, especificidad, chi-cuadrado, *log-likehood* y *log-odds ratio*) que le permiten la recuperación de candidatos a término tanto simples como complejos. En general, el sistema asigna valores a las unidades léxicas marcando las de mayor frecuencia como los candidatos a término, teniendo como base la comparación en el comportamiento de dichas unidades entre los textos a analizar y los corpus de referencia utilizados por *TermoStat*.

Una vez realizada la extracción con *TermoStat* se procedió con la limpieza de los datos recuperados. Es menester recordar que la precisión general de la extracción realizada por *TermoStat* se establece alrededor del 81 % (Drouin, 2003, p. 113), lo que implica que necesariamente debe darse un proceso de revisión de los candidatos extraídos. En este sentido, se establecieron, inicialmente, los siguientes criterios: 1) revisar el etiquetaje, es decir, que cada candidato contara con la correcta etiquetación de su categoría gramatical entendida como el tipo de palabra; 2) revisar que el candidato correspondiera temáticamente con el área de origen, en este sentido, todos los candidatos relativos a la *estadística* fueron omitidos para el análisis final; 3) se omitieron todos

[56] Para más información sobre este recurso, consúltese: https://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/.

los candidatos a término relativos a maquetación de información, esto es, por ejemplo: *figura, tabla, gráfico*, entre otros; 4) los verbos de soporte fueron omitidos; 5) se omitieron los candidatos en idiomas distintos al español salvo si pertenecen al área analizada, como por ejemplo *software* en el área de la computación; 6) se omitieron las unidades de medida, de tiempo y espacio; y 7) los nombres propios no temáticos. Al respecto, como se verá más adelante en la sección de Discusión (§ 5) algunos de estos criterios fueron repensados en función de los resultados obtenidos en nuestro análisis.

3.2.4 Fase 04: Aspectos metodológicos para la aplicación del protocolo de validación

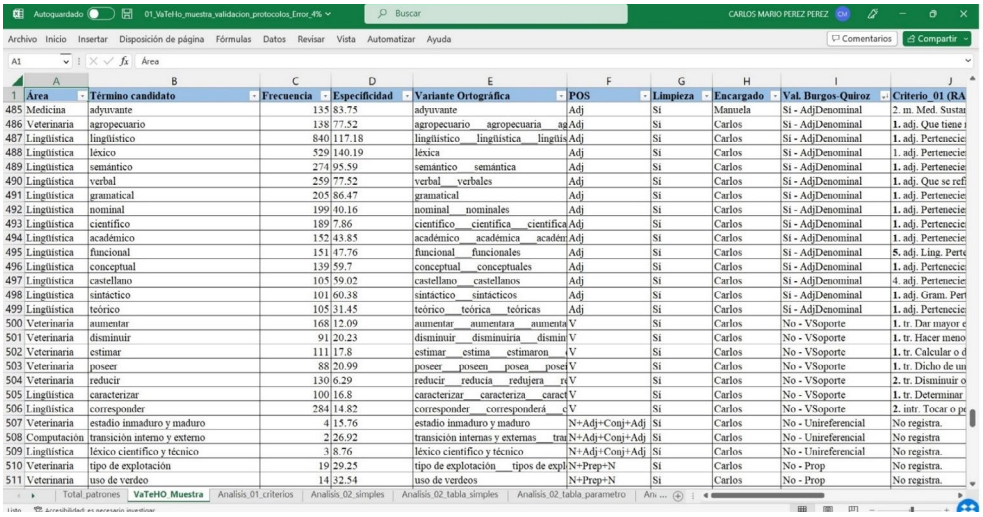
Teniendo como base los aspectos teóricos en la construcción del protocolo para la validación de los candidatos a término (§ 2.3), en nuestro caso los recuperados a partir del corpus VaTeHo, la aplicación metodológica del protocolo (Quiroz & Burgos, en prensa) es el resultado de los desafíos teórico-metodológicos en la delimitación y entendimiento acerca de qué y qué no se considera como término. Para ello, se implementaron una serie de procesos que permitieron la etiquetación y clasificación de los candidatos en la muestra de trabajo procesada como se evidencia tanto en los Resultados (§ 4) como en la Discusión (§ 5). En esencia, el instrumento está diseñado para responder a las particularidades de las unidades simples y complejas, por lo tanto, en su aplicación se podría hablar de dos subprocesos. En primera instancia, se definieron los criterios que determinan qué se considera término en los candidatos simples y, por contraste, qué candidatos no se consideran como tal y, en segunda instancia, el mismo procedimiento para las unidades complejas. Un aspecto clave en la aplicación del protocolo fue que no se eliminó ningún candidato a término de los obtenidos en la extracción, sino que se identificaron en la base de datos con *Sí* o *No* para indicar si cumplían o no con los parámetros arriba descritos. Este elemento metodológico contribuyó en gran medida en la discusión final de la presente tesis doctoral en dos elementos, a saber: 1) la caracterización

de aquellas unidades no terminológicas y 2) el refinamiento del protocolo de validación.

Para el manejo de la información sobre la organización y validación de los candidatos a término, se creó una hoja de Excel conformada por las siguientes columnas que a su vez corresponden con cada uno de los procesos en la aplicación del protocolo. En este punto no hay distinción entre unidades simples o complejas toda vez que en ambas se aplicaron los mismos pasos de validación. Así, la distribución de columnas corresponde a: a) *área del saber*, b) *término candidato*, c) frecuencia de aparición en el subcorpus, d) *especificidad*[57], (al respecto de este criterio, a partir de la propuesta de Lafon (1980), se asigna un puntaje a cada candidato de acuerdo a su comportamiento, en función del contraste entre el corpus de consulta y los corpus de referencia de la herramienta, que a su vez permite caracterizar el vocabulario de especialidad respecto del general), e) *variante ortográfica*, f) *categoría gramatical* (*Part of Speech*, POS), g) *limpieza* (§ 3.2.3), h) *encargado*[58], i) *validación* según el protocolo, j-k) validación por diccionarios, l-n) *contextos de aparición*, tanto en nuestros datos como en corpus de referencia como el CREA, o) *aparición* en otras áreas distintas a la inicial y, p) *comentarios generales*. La Figura 4, a modo de ejemplo, ilustra la organización del archivo de Excel con los datos validados, en las columnas A-J.

[57] Con relación a estas medidas estadísticas utilizadas por *TermoStat*, consúltese el artículo publicado por Drouin (2003) al respecto de este aspecto, titulado "Term extraction using non-technical corpora as a point of leverage".

[58] En la aplicación del protocolo de validación se contó con el apoyo de los miembros del Grupo de Investigación en Traducción y Nuevas Tecnologías (TNT), Escuela de Idiomas, Universidad de Antioquia, Colombia. La distribución de este trabajo se dividió de la siguiente manera: aproximadamente el 90 % de los datos fueron validados por mí, como investigador a cargo de la Variación Horizontal y el 10 % restante se distribuyó entre los demás miembros del grupo que apoyaron esta labor. Puntualmente, Dr. Gabriel Quiroz, Mg. Felipe Zuluaga y la traductora Manuela Hernández. Un reconocimiento y agradecimiento por su aporte en esta labor.



	A	B	C	D	E	F	G	H	I	J
1	Área	Término candidato	Frecuencia	Especificidad	Variante Ortográfica	POS	Limpieza	Encargado	Val. Burgos-Quiroz	Criterio_01 (RA
485	Medicina	adyuvante	135	83.75	adyuvante	Adj	Si	Manuela	Si - AdjDenominal	2. m. Med. Sustar
486	Veterinaria	agropecuario	138	77.52	agropecuario__agropecuaria__ag	Adj	Si	Carlos	Si - AdjDenominal	1. adj. Que tiene
487	Lingüística	lingüístico	840	117.18	lingüístico__lingüística__lingüís	Adj	Si	Carlos	Si - AdjDenominal	1. adj. Pertenecie
488	Lingüística	léxico	529	140.19	léxica	Adj	Si	Carlos	Si - AdjDenominal	1. adj. Pertenecie
489	Lingüística	semántico	274	95.59	semántico__semántica	Adj	Si	Carlos	Si - AdjDenominal	1. adj. Pertenecie
490	Lingüística	verbal	259	77.52	verbal__verbales	Adj	Si	Carlos	Si - AdjDenominal	1. adj. Que se refi
491	Lingüística	gramatical	205	86.47	gramatical	Adj	Si	Carlos	Si - AdjDenominal	1. adj. Pertenecie
492	Lingüística	nominal	199	40.16	nominal__nominales	Adj	Si	Carlos	Si - AdjDenominal	1. adj. Pertenecie
493	Lingüística	científico	189	7.86	científico__científica__científica	Adj	Si	Carlos	Si - AdjDenominal	1. adj. Pertenecie
494	Lingüística	académico	152	43.85	académico__académica__académ	Adj	Si	Carlos	Si - AdjDenominal	1. adj. Pertenecie
495	Lingüística	funcional	151	47.76	funcional__funcionales	Adj	Si	Carlos	Si - AdjDenominal	5. adj. Ling. Perte
496	Lingüística	conceptual	139	59.7	conceptual__conceptuales	Adj	Si	Carlos	Si - AdjDenominal	1. adj. Pertenecie
497	Lingüística	castellano	105	59.02	castellano__castellanos	Adj	Si	Carlos	Si - AdjDenominal	4. adj. Pertenecie
498	Lingüística	sintáctico	101	60.38	sintáctico__sintácticos	Adj	Si	Carlos	Si - AdjDenominal	1. adj. Gram. Per
499	Lingüística	teórico	105	31.45	teórico__teórica__teóricas	Adj	Si	Carlos	Si - AdjDenominal	1. adj. Pertenecie
500	Veterinaria	aumentar	168	12.09	aumentar__aumentara__aumenta	V	Si	Carlos	No - VSoporte	1. tr. Dar mayor e
501	Veterinaria	disminuir	91	20.23	disminuir__disminuiría__dismin	V	Si	Carlos	No - VSoporte	1. tr. Hacer meno
502	Veterinaria	estimar	111	17.8	estimar__estima__estimaron__	V	Si	Carlos	No - VSoporte	1. tr. Calcular o d
503	Veterinaria	poseer	88	20.99	poseer__poseen__posea__posei	V	Si	Carlos	No - VSoporte	1. tr. Dicho de un
504	Veterinaria	reducir	130	6.29	reducir__reducía__redujera__re	V	Si	Carlos	No - VSoporte	2. tr. Disminuir o
505	Lingüística	caracterizar	100	16.8	caracterizar__caracteriza__caract	V	Si	Carlos	No - VSoporte	1. tr. Determinar
506	Lingüística	corresponder	284	14.82	corresponder__correspondería__c	V	Si	Carlos	No - VSoporte	2. intr. Tocar o pe
507	Veterinaria	estadio inmaduro y maduro	4	15.76	estadio inmaduro y maduro	N+Adj+Conj+Adj	Si	Carlos	No - Unireferencial	No registra.
508	Computación	transición interno y externo	2	26.92	transición internas y externas__tra	N+Adj+Conj+Adj	Si	Carlos	No - Unireferencial	No registra
509	Lingüística	léxico científico y técnico	3	8.76	léxico científico y técnico	N+Adj+Conj+Adj	Si	Carlos	No - Unireferencial	No registra.
510	Veterinaria	tipo de explotación	19	29.25	tipo de explotación__tipos de expl	N+Prep+N	Si	Carlos	No - Prop	No registra.
511	Veterinaria	uso de verdeo	14	32.54	uso de verdeos	N+Prep+N	Si	Carlos	No - Prop	No registra.

Figura 4. Excel con los datos de la validación de los candidatos a término columnas A-J.

Otro aspecto relevante en la validación de los candidatos fue el contraste con diccionarios. En este punto, se utilizó el *Diccionario de la lengua española* (DRAE) (Real Academia Española, 2022a) para contrastar cada uno de los candidatos en todas las áreas. Al respecto, a pesar de que el DRAE es un diccionario de lengua general, este también presenta marcas de especialidad en diversidad de entradas, además es un recurso normativo y referencia para el español, por lo tanto, era necesaria su utilización en nuestra investigación. En cuanto a los diccionarios de especialidad utilizados, estos fueron elegidos conforme su representatividad al interior de cada campo del conocimiento analizado, de este modo, para medicina se utilizó el *Mosby* (Anderson, 2003) y el *Diccionario de cáncer* del Instituto Nacional del Cáncer (2022); para lingüística el *Diccionario de lingüística moderna* (Alcaraz & Martínez, 1997); para veterinaria el *Agrovoc* (Food and Agriculture Organization of the United Nations, 2022); para computación el *Portal lingüístico de Microsoft* (Microsoft, 2022) y para filosofía el *Diccionario de filosofía* (Ferrater, 1978) y el *Diccionario AKAL de filosofía* (Audi, 2004).

3.2.4.1 Criterios para la determinación de unidades simples como términos

En primer lugar, son *adjetivos terminológicos* aquellos denominales y deverbales como en *gástrico*, *histológico* (Med), *nominal*, *semántico* (Lin), *bovino*, *ruminal* (Vet), *artificial*, *virtual* (Com), *aristotélico*, *epistémico* (Fil), además de los colores como en *azul*, *blanco*, *rojo*, *verde*, (Med, Vet). Asimismo, no son *adjetivos terminológicos* aquellos calificativos como *grande*, *joven* (Med, Vet). En segundo lugar, son *adverbios terminológicos* aquellos derivativos terminados en el sufijo *-mente* como en *imagenológicamente*, *quirúrgicamente* (Med), *alveopalatalmente*, *dialectalmente* (Lin), *enzoóticamente*, *epizootiológicamente* (Vet), *análogamente*, *paramétricamente* (Com), *acríticamente*, *ontológicamente* (Fil). A su vez, no son *adverbios terminológicos* los provenientes de adjetivos calificativos o los generales como en *adicionalmente*, *comúnmente*, *generalmente* (Med, Lin, Vet).

En tercer lugar, son *sustantivos terminológicos* aquellos que por composición o derivación pertenecen a una de las áreas analizadas como en *cardiopatía, histología*, (Med), *coloquialismo, declarativo* (Lin), *agropecuario, botánico* (Vet), *algoritmo, generador* (Com), *aristotelismo, epistemología* (Fil). Asimismo, no son *sustantivos terminológicos* aquellos comunes como en *población, riesgo* (Med), *conocimiento, papel* (Lin), *alimento, agua* (Vet), *herramienta, información* (Com). En cuarto lugar, son *verbos terminológicos* aquellos derivados por prefijación o sufijación, compuestos por formantes patrimoniales o de origen grecolatino como en *detectar, diagnosticar* (Med), *adjetivar, definir* (Lin), *criar, ordeñar* (Vet), *desarrollar, representar* (Com), *afirmar, colegir* (Fil). No son *verbos terminológicos* los de soporte, discursivos, conectores, fraseológicos (Lorente, 2002) como en *utilizar, visualizar* (Med), *agrupar, relatar* (Lin), *comparar, evidenciar* (Vet), *crear, utilizar* (Com), *agregar, postular* (Fil).

3.2.4.2 Criterios para la determinación de unidades complejas como términos

Para la determinación de las unidades complejas como términos el protocolo de validación (Quiroz & Burgos, en prensa) establece dos

momentos. En primera instancia, es necesaria una verificación de los rasgos semánticos de los constituyentes de la secuencia, condición dada por las relaciones de los componentes en función de las unidades de menor arbitrariedad. De esta manera, para determinar este elemento se procede con cuatro verificaciones iniciales a cada uno de los candidatos a términos complejos, las cuales se relacionan entre sí jerárquicamente, esto es, para avanzar a la siguiente debe cumplirse con la anterior regla de lo contrario el candidato no se tiene en cuenta para la segunda parte del análisis en la cual se determina si es un término o, por el contrario, otro tipo de unidad con carácter de especialidad. En este sentido: a) se verifica que el candidato es *unireferencial* (2.3.1.1) y *no composicional* (2.3.1.2), b) se observa el grado de *arbitrariedad* (2.3.1.3) del candidato, esto es, a mayor grado de arbitrariedad mayor posibilidad de ser un término, c) se lleva cada elemento a su categoría primitiva sin que cambie su significado como en *extirpación del colon*, donde las categorías primitivas son *extirpar* y *colon* y, por ende, se obtiene *extirpar el colon* y, d) si el candidato tiene premodificación, articulo definido entre sus componentes o si no es un sintagma nominal, no se tiene en cuenta para la posterior etapa de validación.

En segunda instancia, una vez completados los pasos anteriores, se procede con la verificación semántica de los candidatos, es decir, la relación entre sus componentes debe ser de estricto orden semántico. Para ello, en primer lugar, se verifica la condición de *independencia del núcleo*, es decir, si el núcleo significa lo mismo en cualquier contexto entonces se descarta el candidato, como sucede en *creatividad lingüística*, donde el núcleo *creatividad* significa lo mismo en cualquier contexto. En segundo lugar, se verifica la condición de *propiedades compartidas*, esto es, si el candidato es un término entonces al menos una de las características del núcleo es modificada por alguna del modificador. Si esta condición no se cumple, el candidato es descartado para la siguiente etapa de la verificación, como ocurre en *centro de interés* donde el modificador interés no aporta ninguna de sus características al núcleo. Ahora bien, realizada la verificación semántica, se procede con asignación de las etiquetas

a los candidatos, es decir, se clasifican como términos o no según las siguientes consideraciones. Así, son *términos complejos* aquellos candidatos que la relación entre sus componentes es[59]:

a. *Arbitraria*: el candidato no cumple con los criterios de *validación semántica* o *propiedades compartidas*, pero es una unidad dentro del área del saber a la cual pertenece, como sucede en: *exceso de peso* (Med), *disponibilidad léxica* (Lin), *capacidad de carga* (Vet), *entorno de desarrollo* (Com), *estado de cosa* (Fil).

b. *Hecho de*: el núcleo está hecho del modificador. Como sucede en: *médula ósea* (Med), *corpus del español* (Lin), *emisión de metano* (Vet), *base de datos* (Com), *escuela de Frankfurt* (Fil).

c. *Metf hecho de*: el núcleo es una metáfora hecha por el modificador. Como sucede en: *familia lingüística* (Lin).

d. *Tipo de*: el término es un tipo de núcleo. Como sucede en: *metástasis hepática* (Med), *argumento léxico* (Lin), *vaca lechera* (Vet), *lenguaje de programación* (Com), *testimonio moral* (Fil).

e. *Tipo de / Hecho de*: puede interpretarse tanto como *b* o *d*. Como sucede en los ejemplos anteriores.

Asimismo, no son términos complejos aquellos candidatos que la relación entre sus componentes está determinada por y, por lo tanto, son descartados:

a. *No-acus*: el núcleo es deverbal y el modificador es un objeto directo. Como sucede en: *consumo de tabaco* (Med), *conocimiento léxico* (Lin).

[59] Los ejemplos presentados corresponden a los candidatos validados según la muestra de trabajo (§ 4.2.7) y, por lo tanto, pueden variar en la cantidad y áreas ejemplificadas. Para el caso de la lingüística, los ejemplos se toman de las ejemplificaciones dadas en el protocolo de validación (Quiroz & Burgos, en prensa).

b. *No-Deadj*: el núcleo es deadjetival. Como sucede en: *habilidad lectora* (Lin), *diversidad vegetal* (Vet).

c. *No-ejemplo de*: el candidato es un ejemplo de un tipo de núcleo. Como sucede en: *cátedra de dermatología* (Med), *diccionario del español* (Lin).

d. *No-meron*: el núcleo es parte del modificador. Como sucede en: *morfemas de la lengua* (Lin).

e. *No-nomin*: el núcleo es sujeto y el modificador es nombre deverbal o viceversa. Como sucede en: *control del cáncer* (Med), *lengua receptora* (Lin), *consumo de materia seca* (Vet).

f. *No-ocurre en*: el núcleo ocurre en un lugar o momento indicado por el modificador. Como sucede en: *tarea escolar*, *sílaba final* (Lin).

g. *No-partit*: el núcleo es un partitivo del modificador. Como sucede en: *número de ganglio* (Med), *cantidad de palabra* (Lin), *número de vaca* (Vet).

h. *No-prop*: el núcleo es una propiedad del modificador. Como sucede en: *tamaño del tumor* (Med), *variedad lingüística* (Lin), *uso de verdeo* (Vet).

i. *No-ambig*: el candidato no es un término. Esta etiqueta, aunque funcional, puede llegar a convertirse en un parámetro muy arbitrario.

3.2.5 Fase 05: Resultados, análisis y escritura del informe final de investigación

La Fase 05 constituye el final del camino metodológico en la presente investigación doctoral. En esta última etapa de trabajo se integran los dos bloques de información centrales del trabajo, es decir, los aspectos teóricos (§§ 1 y 2) y la fundamentación metodológica (§ 3), integración que permitió la obtención de los

resultados ((§ 4) y el análisis y discusión de estos (§ 5). Así, el presente escrito se constituye como la evidencia sistemática de nuestro estudio en torno al fenómeno lingüístico de la variación y, en consecuencia, una propuesta de ampliación del concepto de término hasta ahora prototípico en la terminología. Por último, a partir del corpus VaTeHoVe[60] se construyó la interfaz de consulta disponible para investigadores en el área y demás público interesado en el tema de trabajo. Las Figuras 5 a 7 ilustran una versión preliminar de la interfaz final de consulta de este corpus.

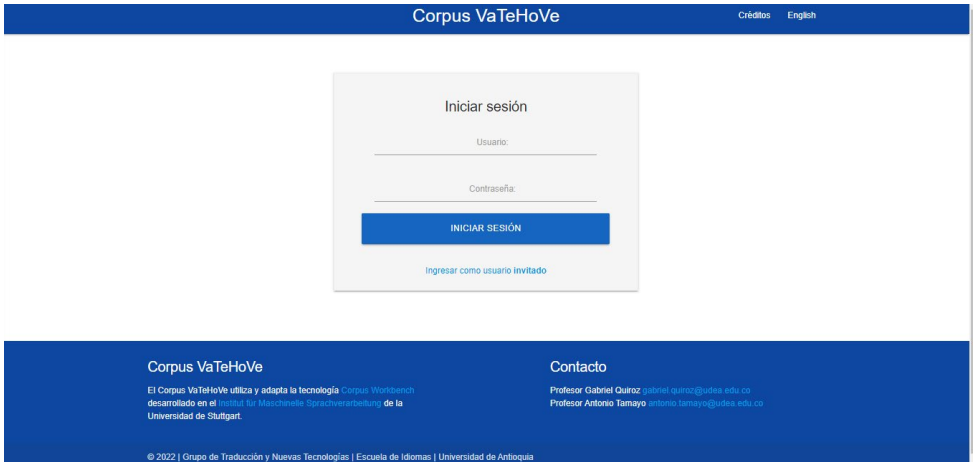


Figura 5. Interfaz corpus VaTeHoVe.

60 Este recurso comprende tanto la variación horizontal como vertical. Este corpus es el resultado del proyecto macro en el cual se enmarca la presente tesis doctoral: Variación Terminológica Horizontal y Vertical en español: hacia una descripción lingüística de las unidades terminológicas desde la lingüística de corpus", del Grupo de Investigación en Traducción y Nuevas Tecnologías (TNT), Escuela de Idiomas, Universidad de Antioquia, Colombia; financiado por el Comité para el Desarrollo de la Investigación (CODI) de la misma universidad, en Convocatoria Programática 2018 Área Ciencias Sociales, Humanidades y Artes, Acta nro. 2018-22970.



Figura 6. Interfaz de consulta VaTeHoVe.



Figura 7. Interfaz de consulta VaTeHoVe.

3.3 Recapitulación

En relación con el segmento metodológico se han detallado las diferentes *fases* de trabajo desde la concepción inicial de la presente investigación pasando por la compilación del corpus textual especializado VaTeHo y culminando con las reflexiones en torno a los resultados gracias a la implementación del protocolo para la validación

de términos simples y complejos (Quiroz & Burgos, en prensa). En este sentido, la presente investigación de corte descriptiva hace uso de herramientas computacionales que permitieron optimizar el tiempo empleado en distintos procesos sin dejar a un lado la validación manual necesaria durante todas las etapas de trabajo. Por un lado, el corpus VaTeHo quedó conformado por un total de 538 artículos de investigación en cinco áreas del saber: *medicina*, *lingüística*, *veterinaria*, *computación* y *filosofía*, a partir de 197 revistas científicas. Además, el corpus VaTeHo se configura como un recurso representativo del fenómeno en estudio con cerca de 3.057.973 *tokens*. Un corpus del español en cuanto logró abarcar la realidad lingüística de esta lengua al contener información de todas las zonas lingüísticas tradicionales a lo largo y ancho del planeta toda vez que la información compilada fue recuperada de la mayoría de los países donde el español es lengua oficial.

Por otro lado, la implementación del protocolo para la validación de términos permitió tener un panorama más amplio en la concepción del objeto de estudio a partir de la caracterización de las unidades tanto simples como complejas. En este sentido, junto con las bases teóricas en la construcción de este instrumento (§ 2.3) el componente metodológico aquí descrito permite operativizar a partir de los criterios sintácticos, semánticos, pragmáticos y contextuales la identificación de dichas unidades. Asimismo, la creación de la base de datos a partir de los distintos bloques de información y, en consecuencia, la puesta en funcionamiento de la interfaz de consulta del corpus VaTeHo permitirá la visibilización de nuestro trabajo en distintos contextos tanto nacionales como internacionales y, a su vez, la interacción en los distintos escenarios de discusión académica en torno al fenómeno analizado en nuestro trabajo de investigación. Finalmente, en los siguientes apartados (§§ 4-5), se presentarán los hallazgos y las discusiones al respecto de nuestras hipótesis y objetivos de investigación gracias a la integración teórico-metodológica de esta tesis doctoral.

PARTE III:
EL TÉRMINO EN ESPAÑOL

DE ANTIOQUIA
ELLIN

4. Resultados: una antesala al concepto de término en español

"Insanity is doing the same thing
over and over again
and expecting different results."

(*Sudden Death*, Rita Mae Brown).

Las secciones anteriores a este capítulo (§§ 0-3) nos permitieron hacer un recorrido por los principales umbrales (Bachelard, 1957, p. 45) en el estudio de la terminología. Por un lado, los principios y postulados fundacionales que configuran la mirada clásica de nuestra disciplina. Por otro lado, el camino transitado, hasta hoy, en la concepción moderna y en la construcción del área alrededor del término. El objetivo central de la presente tesis doctoral nos ha permitido repensar la disciplina a partir de la ampliación en la forma cómo se ha observado el término. Así, nuestro planteamiento metodológico (§ 3) se presenta como una respuesta al desafío de seguir construyendo sobre las bases aportadas por todas y cada una de las escuelas de pensamiento en las diferentes latitudes alrededor del planeta. Además, la consolidación del corpus de trabajo permitió tener una mirada amplía del fenómeno de la variación terminológica al incluir cinco áreas del saber y compilar los datos a partir de la mayor cantidad de países donde el español es lengua oficial. De este modo, una vez superada la construcción teórico-metodológica de esta tesis doctoral, los resultados aquí presentados, ejemplificados y en adelante discutidos, se constituyen en un aporte más en el entendimiento común del término como objeto de estudio de la terminología.

Haciendo uso de las ventajas de la lingüística computacional y de corpus, al servicio del procesamiento del lenguaje natural, nuestra investigación discurre entre lo manual y lo automático: lo primero evidenciado en la forma canónica en la cual se compilaron los 538 artículos que conforman el corpus especializado VaTeHo y, a su vez, en la manera como fueron revisados, uno a uno, los candidatos a término en las distintas Fases de trabajo (§ 3.2) y, lo segundo, en cuanto estos se recuperaron, automáticamente, gracias al procesamiento realizado en *TermoStat*, siendo este el único proceso automático que se llevó a cabo a lo largo de la investigación. Asimismo, como ya se estableció, el corpus VaTeHo se conformó gracias a la compilación de los artículos científicos en 197 revistas indexadas en cada una de las áreas de trabajo, a saber: medicina, lingüística, veterinaria, computación y filosofía. Ahora bien, esta distribución de los artículos obedece, principalmente, a la disponibilidad de información en cada uno de los países donde se realizó la búsqueda de información, bajo los criterios de selección establecidos para tal propósito (§ 3.2.1.5). El corpus VaTeHo quedó conformado por 3.057.973 *tokens*. A continuación, se presentarán los pormenores en los resultados obtenidos desde la extracción inicial, pasando por la consolidación de la muestra de trabajo hasta la aplicación del protocolo para la validación de los candidatos a término (Quiroz & Burgos, en prensa).

4.1 Generalidades de la extracción en *Termostat*

Para el procesamiento de cada subcorpus con *TermoStat* fue necesario la creación de cinco archivos independientes, esto es, se unificó en un solo documento de texto todos los artículos por área luego de llevarse a cabo la adecuación inicial de estos, según los parámetros establecidos, con el objetivo de eliminar toda la información no relevante para el propósito de la investigación contenida en los artículos (§ 3.2.1.5). Así, se tiene un documento por área los cuales se denominaron: *Med_All*, *Lin_All*, *Vet_All*, *Com_All* y *Fil_All*. De este modo, fruto del procesamiento, el programa arrojó un total aproximado de 37.055 candidatos a término distribuidos

en cada uno de los cinco subcorpus procesados, como se ilustra en la Figura 8. Es importante mencionar que este valor corresponde a lo que se denominó la *extracción en bruto*, es decir, la información arrojada por la herramienta sin ningún tipo de filtro, salvo los mínimos requeridos por la interfaz para el procesamiento, a saber: definición del idioma del texto a procesar (español) y los términos a extraer (simples y complejos).

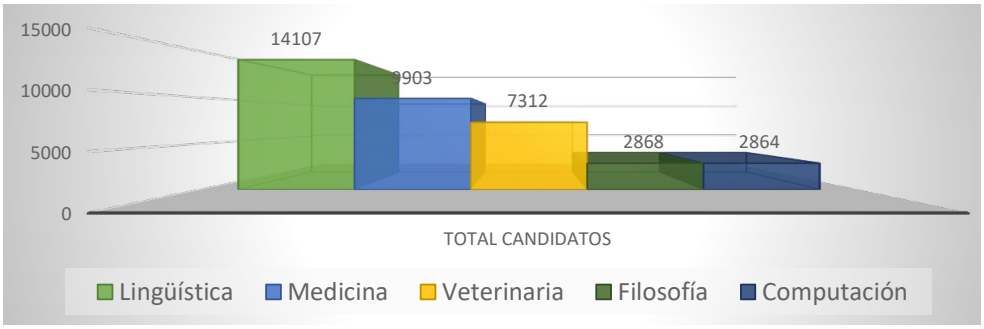


Figura 8. Total de candidatos a término por área del saber.

En línea con estos resultados, a partir de este procesamiento también se logró recuperar otra información importante para la descripción de los datos obtenidos. Por un lado, se recuperó el *campo semántico* de cada área lo que permitió corroborar que la selección de los artículos concuerda con las subáreas de trabajo establecidas, esto es: medicina (*oncología* y/o *cáncer*), lingüística (*léxico, lexicología* y/o *lexicografía*), veterinaria (*ganadería*), computación (*software*) y filosofía (*epistemología* y/o *conocimiento*). Por otro lado, *TermoStat* permite recuperar la información estadística sobre la extracción realizada además de los *contextos de aparición* y las *colocaciones* de los candidatos a término. Esta información es relevante puesto que permite fijar un punto de partida y comparación con los resultados finales de nuestra tesis doctoral en función de la aplicación de nuestro protocolo de validación y, en consecuencia, la posibilidad de aportar desde nuestro entendimiento qué es y qué no es un término en el mejoramiento de este tipo de herramientas de procesamiento

automático al servicio de la terminología. Así, en este segmento se presentará la información conforme es recuperada por la herramienta de manera automática.

Al respecto, como se ampliará en la discusión (§ 5), hay candidatos recuperados automáticamente por el programa que no corresponden, bajo la mirada del protocolo de validación, a términos y, en consecuencia, se consolidan como un insumo importante con miras a una posible actualización y retroalimentación de la herramienta. En adelante, se detallan las generalidades en la extracción por cada una de las áreas del saber. Para la ejemplificación, además de la categoría gramatical y/o la estructura sintáctica, junto al término se anotará, entre paréntesis, la información sobre su etiqueta de validación, conforme al protocolo. Es importante mencionar en este punto que muchos de los candidatos a término recuperados bajo la estructura [N+N] están mal etiquetados. En este sentido, la mayoría de estas unidades corresponden a la estructura [N+Adj] o, en menor medida, a otras estructuras como [Pron+N]. Esta particularidad se ahondará igualmente en el segmento de discusión. Por último, en las ejemplificaciones donde se presentan los contextos de aparición de la unidad analizada se indicará la codificación del artículo del cual fue recuperado el ejemplo, para su identificación al interior de la base de datos.

4.1.1 Datos generales subcorpus de medicina

El subcorpus de medicina está conformado por 592.104 *tokens* a partir de 161 artículos compilados. En primera instancia, la Figura 9 ilustra la nube de palabras de esta área. Los conceptos que sobresalen en la construcción temática en medicina son, entre otros, *cáncer*, *cáncer gástrico*, *metástasis*, *neoplasia* y *patología*. A manera de ilustración, se presentan los ejemplos 10 a 14 en VaTeHo.



Figura 9. Nube de palabras medicina.

10. El fenómeno de envejecimiento poblacional ha sido también vinculado al incremento de la morbimortalidad por *cáncer* (NTérmino). (Med_ARG_01).

11. El *cáncer gástrico* (Tipo de) se diagnosticó en estadio IV en 65 % de los casos. La mejor supervivencia se presentó en casos de cáncer de colon sometidos a cirugía. (Med_COL_02).

12. Se demostró *metástasis* (NTérmino) a ganglios axilares en 59 casos (45%). (Med_MEX_10).

13. En la mayoría de los casos la *neoplasia* (NTérmino) es de tipo papilar y de bajo grado de malignidad. (Med_CUB_02).

14. Cualquier *patología* (NTérmino) que, de manera secundaria, pudiera interferir en el correcto metabolismo de la vitamina D [...]. (Med_ESP_07).

En segunda instancia, en la Tabla 4 se compila la información estadística descriptiva a partir de los candidatos recuperados. Inicialmente, es importante señalar que las *unidades complejas* son las de mayor frecuencia al tomar el dato por *types*[61], característica que comparte con los datos recuperados en computación. Unida a la información estadística de cada una de las estructuras sintácticas recuperadas, *TermoStat* proporciona una lista de ejemplos que serán

61 Entiéndase *types* como el número de palabras diferentes según el total de palabras (*tokens*) (Brainerd, 1972, 507). En este caso, los types son equivalentes a los términos ya sea simples o complejos recuperados según las reglas de *TermoStat*.

etiquetados conforme se explicó previamente. En 15 a 25, se presentan algunos ejemplos en cada categoría.

Tabla 4. Estadística extracción TermoStat - medicina.

	ÁREA	MEDICINA			
	POS	*Types*		*Tokens*	
		Frecuencia	%	Frecuencia	%
Estructuras sintácticas	N	2579	26,0 %	98637	55,8 %
	N+Adj	2543	26,0 %	15771	8,9 %
	N+Prep +N	1781	18,0 %	10748	6,1 %
	Adj	1146	12,0 %	29500	16,7 %
	N+N	665	7,0 %	3555	2,0 %
	V	448	5,0 %	13826	7,8 %
	N+Prep +N+Adj	401	4,0 %	1384	0,8 %
	N+Adj+Adj	201	2,0 %	726	0,4 %
	Adv	83	1,0 %	2278	1,3 %
	N+Adj+Conj+Adj	52	1,0 %	209	0,1 %
	N+Adj+Adj+Adj	3	0,0 %	6	0,0 %
	N+Adj+Adj+Adj+Adj	1	0,0 %	2	0,0 %
	Total de artículos	**161**			
	Total de palabras	**592104**			
Types	Total de candidatos a término	**9903**			
	Unidades simples	**44 %**			
	Unidades complejas	**58 %**			
Tokens	Total de candidatos a término	**176642**			
	Unidades simples	**81,7 %**			
	Unidades complejas	**18,3 %**			

15. **N términos**: *cáncer* (NTérmino), *diagnóstico* (NTérmino), *enfermedad* (NTérmino), *paciente* (NTérmino), *tratamiento* (NTérmino), *tumor* (NTérmino).

N no términos: *año* (NComún), *caso* (NComún), *edad* (NComún), *estudio* (NComún).

16. **[N+Adj] términos**: *cáncer gástrico* (Tipo de), *detección temprana* (Tipo de), *estadio clínico* (Hecho de), *ganglio linfático* (Tipo de), *tumor primario* (Tipo de), *tratamiento sistémico* (Tipo de).

    [N+Adj] no términos: *tipo histológico* (No-Prop), *característica patológica* (No-Prop), *clínica alemana* (No-Prop).

17. **[N+Prep+N] términos**: *calidad de vida* (Arbitraria), *cáncer de próstata* (Tipo de), *diagnóstico de cáncer* (Tipo de), *factor de riesgo* (Hecho de), *tasa de mortalidad* (Tipo de).

    [N+Prep+N] no términos: *años de edad* (No-Meron), *momento del diagnóstico* (No-Ocurre en), *servicio de fisiatría* (No-Prop), *tipo de cáncer* (No-Prop).

18. **Adj términos**: *alto* (AdjGeneral), *bajo* (AdjGeneral), *clínico* (AdjDenominal), *positivo* (AdjGeneral), *gástrico* (AdjRelacional), *negativo* (AdjGeneral).

    Adj no términos: *grande* (AdjGeneral), *pequeño* (AdjGeneral).

19. **[N+N] términos**: *grupo control* (Tipo de), *sector salud* (Hecho de).

    [N+N] que son [N+Adj] términos: *carcinoma epidermoide* (Tipo de), *célula tumoral* (Tipo de), *médico especialista* (Tipo de), *respuesta patológica* (Tipo de), *tamaño tumoral* (Hecho de).

    [N+N] que son [Pron+N] no términos: *nuestro país* (No-Ambig), *nuestro estudio* (No-Ambig).

20. **V términos**: *anestesiar* (VTérmino), *desinflamar* (VTérmino), *detectar* (VTérmino), *diagnosticar* (VTérmino), *malignizar* (VTérmino), *padecer* (VTérmino), *sospechar* (VTérmino).

    V no términos: *incluir* (VSoporte), *mostrar* (VSoporte), *observar* (VSoporte), *presentar* (VSoporte), *realizar* (VSoporte), *utilizar* (VSoporte).

21. **[N+Prep+N+Adj] término**: *cáncer de cuello uterino* (Tipo de), *carcinoma de célula escamosa* (Tipo de), *factor de crecimiento epidérmico* (Tipo de), *índice de masa corporal* (Tipo de/ Hecho de), *mortalidad por cáncer cervical* (Tipo de/Hecho de), *problema de salud pública* (Tipo de/Hecho de), *sarcoma de parte blanda* (Tipo de), *virus del papiloma humano* (Tipo de).

    [N+Prep+N+Adj] no términos: *número de ganglio linfático* (No-Partit).

22. **[N+Adj+Adj] términos**: *ganglio linfático regional* (Tipo de), *leucemia linfoblástica aguda* (Tipo de), *sistema nervioso central* (Tipo de), *terapia endocrina básica* (Tipo de), *tratamiento loco regional* (Tipo de), *vaciamiento central profiláctico* (Tipo de).

23. **Adv términos**: *histológicamente* (Sí-AdvAdjetival-Mente), *imagenológicamente* (Sí-AdvAdjetival-Mente), *inmunohistoquimicamente* (Sí-AdvAdjetival-Mente), *quirúrgicamente* (Sí-AdvAdjetival-Mente).

    Adv no términos: *aproximadamente* (No-AdvGral-Mente), *donde* (No-AdvRelativo), *localmente* (AdvAdjetival-Mente), *luego* (No-AdvGral), *previamente* (No-AdvGral-Mente), *principalmente* (No-AdvGral-Mente), *respectivamente* (No-AdvGral-Mente).

24. **[N+Adj+Conj+Adj] no términos**: *característica clínica y patológica* (no unireferencial), *institución pública y privada* (no unireferencial), *ovario voluminoso e irresecable* (no unireferencial), *recurso físico y humano* (no unireferencial), *salud sexual y reproductiva* (no unireferencial), *tumor benigno y maligno* (no unireferencial).

25. **[N+Adj+Adj+Adj] términos**: *cáncer gástrico difuso hereditario* (Tipo de), *proyección medio latero oblicua* (Tipo de).

4.1.2 Datos generales subcorpus de lingüística

El subcorpus de lingüística está conformado por 716.582 *tokens* a partir de 82 artículos compilados. Al igual que con medicina, la Figura 10 ilustra la nube de palabras de esta área, siendo los conceptos más representativos en la construcción temática en este campo del saber, entre otros, *diccionario*, *disponibilidad léxica, lengua*, *palabra* y *sufijo*. En 26 a 30 se presentan ejemplos de aparición.



Figura 10. Nube de palabras lingüística.

26. La idea tras esto es que el *diccionario* (NTérmino) es testimonio y depósito del saber, de la experiencia, de las distintas formas de concebir el mundo asentadas [...]. (Lin_CHL_09).

27. En España, Ortolano (2005) buscó identificar la *disponibilidad léxica* (Arbitrario) de los estudiantes de séptimo y noveno grado de bachillerato. (Lin_COL_02).

28. [...] poder de ascendencia yucateka -por darles un nombre- en el noroeste de la península de Yucatán como hablantes de esta *lengua* (NTérmino). (Lin_MEX_06).

29. Así en su enseñanza comienzan por la frase o *palabra* (NTérmino) y terminan en las sílabas y letras. (Lin_URY_03).

30. Ello porque, por el tiempo en que escribía el cronista, el *sufijo* (NTérmino) referido estaba en vías de completar un proceso de

simplificación en favor de -na (cf. González Holguín [1607] 1975 [...]. (Lin_PER_10).

La Tabla 5 compila la información estadística descriptiva en el área de lingüística. Al respecto, a diferencia de la medicina y computación, para esta área las *unidades simples* son las de mayor frecuencia según el número de *types*, característica que comparte con los datos recuperados en filosofía. En 31 a 41 se presentan los ejemplos suministrados, automáticamente, por *Termostat* para cada categoría.

Tabla 5. Estadística extracción TermoStat - lingüística.

	ÁREA	**LINGÜÍSTICA**			
	POS	***Types***		***Tokens***	
		Frecuencia	**%**	**Frecuencia**	**%**
Estructuras sintácticas	N	5576	40,0 %	107625	58,1 %
	N+Adj	2987	21,0 %	13368	7,2 %
	N+Prep +N	1542	11,0 %	5794	3,1 %
	Adj	1592	11,0 %	30612	16,5 %
	N+N	598	4,0 %	2488	1,3 %
	V	1283	9,0 %	20790	11,2 %
	N+Prep +N+Adj	240	2,0 %	736	0,4 %
	N+Adj+Adj	106	1,0 %	106	0,1 %
	Adv	148	1,0 %	3495	1,9 %
	N+Adj+Conj+Adj	31	0,2 %	91	0,0 %
	N+Adj+Adj+Adj	4	0,0 %	8	0,0 %
	N+Adj+Adj+Adj+Adj	0	0,0 %	0	0,0 %
Total de artículos		**82**			
Total de palabras		**716582**			
Types	Total de candidatos a término	**14107**			
	Unidades simples	**61 %**			
	Unidades complejas	**39 %**			
Tokens	Total de candidatos a término	**185113**			
	Unidades simples	**87,8 %**			
	Unidades complejas	**12,2 %**			

31. **N términos**: *diccionario* (NTérmino), *español* (NTérmino), *lengua* (NTérmino), *léxico* (NTérmino), *palabra* (NTérmino), *significado* (NTéWrmino), *voz* (NTérmino).

    N no términos: *ejemplo* (NComún).

32. **[N+Adj] términos**: *categoría gramatical* (Tipo de/Hecho de), *diccionario bilingüe* (Tipo de), *lengua extranjera* (Tipo de), *lengua indígena* (Tipo de), *lengua materna* (Tipo de), *sistema lingüístico* (Tipo de), *unidad léxica* (Tipo de).

    [N+Adj] no términos: *tarea escolar* (No-Ocurre en).

33. **[N+Prep+N] términos**: *acto de habla* (Tipo de), *adverbio de foco* (Tipo de), *corpus de anécdotas* (Hecho de), *clase de palabra* (Arbitrario), *formación de palabras* (Arbitrario).

    [N+Prep+N] no términos: *centro de interés* (No-Independencia de núcleo), *número de palabras* (No-Partit), *objeto de estudio* (No-Independencia de núcleo).

34. **Adj términos**: *gramatical* (AdjDenominal), *lingüístico* (AdjDenominal), *léxico* (AdjDenominal), *nominal* (AdjDenominal), *semántico* (AdjDenominal), *sintáctico* (AdjDenominal), *verbal* (AdjDenominal).

    Adj no términos: *diferente* (AdjGeneral), *mismo* (AdjGeneral), *siguiente* (AdjGeneral), *solo* (AdjGeneral).

35. **[N+N] que son [N+Adj] términos**: *campo léxico* (Metf hecho de), *lengua española* (Arbitraria), *lengua maya* (Arbitraria), *vocabulario lunfardo* (Hecho de).

36. **V términos**: *adjetivar* (VTérmino), *definir* (VTérmino), *sustantivar* (VTérmino).

V no términos: *aparecer* (VSoporte), *encontrar* (VSoporte), *incluir* (VSoporte), *mostrar* (VDiscursivo), *presentar* (VDiscursivo).

37. **[N+Prep+N+Adj] término**: *concordancia de objeto directo* (Arbitraria), *discurso de divulgación científico* (Tipo de).

    [N+Prep+N+Adj] no términos: *alumno del método sintético* (No-Ejemplo de), *hablantes del maya yucateko* (No-Ejemplo de), *léxico del cuerpo humano* (No-Ejemplo de), *léxico del habla culta* (No-Ejemplo de).

38. **[N+Adj+Adj] términos**: *altura tonal baja* (Tipo de), *discurso científico semiespecializado* (Tipo de), *maya yucateco colonial* (Tipo de), *sílaba final libre* (Tipo de), *sufijo apreciativo diminutivo* (Tipo de).

39. **Adv términos**: *dialectalmente* (Sí-AdvAdjetival-Mente), *fonéticamente* (Sí-AdvAdjetival-Mente), *léxicamente* (Sí-AdvAdjetival-Mente), *metafóricamente* (Sí-AdvAdjetival-Mente), *semánticamente* (Sí-AdvAdjetival-Mente).

    Adv no términos: *donde* (No-AdvRelativo), *generalmente* (No-AdvGral-Mente), *posteriormente* (No-AdvGral-Mente), *respectivamente* (No-AdvGral-Mente), *solo* (AdvGral).

40. **[N+Adj+Conj+Adj] no términos**: *léxico científico y técnico* (no unireferencial), *morfología nominal y verbal* (no unireferencial), *semántica estructural y cognitiva* (no unireferencial).

41. **[N+Adj+Adj+Adj] términos**: *atlas lingüístico etnográfico gráfico*, *realización bilabial fricativa sonora* (Tipo de).

4.1.3 Datos generales subcorpus de veterinaria

El subcorpus de veterinaria está conformado por 344.202 *tokens* a partir de 73 artículos compilados. Al igual que con las dos áreas arriba descritas (medicina y lingüística), la Figura 11 ilustra la nube de palabras de esta área. Los conceptos que sobresalen en la construcción temática en veterinaria son, entre otros, *finca*, *ganadería*, *leche*, *pastoreo* y *vaca*. En 42 a 46 se presentan ejemplos de aparición de estos términos en VaTeHo.



Figura 11. Nube de palabras veterinaria.

42. La evaluación se realizó en una *finca* (NTérmino) de Frigocolanta ubicada en santa Rosa de Osos, Antioquia (latitud 6,613015, longitud -74.438933), a 2587 msnm [...]. (Vet_COL_05).

43. La GE convierte a la *ganadería* (NTérmino) y a la agricultura en una pieza clave del ecosistema y devuelve al ganadero su protagonismo como conservador y artífice del paisaje. (Vet_ESP_03).

44. La producción promedio de *leche* (NTérmino) a 305 días de lactancia (DL-305) obtenida en municipio de Achacachi, provincia Omasuyos, fue de 2743.0 kg [...]. (Vet_BOL_10).

45. El *pastoreo* (NTérmino) produce cambios en la fisonomía y en la composición florística relacionados con las estrategias de las plantas frente a la herbivoría [...]. (Vet_ARG_06).

46. De igual forma, este grupo presenta la más baja productividad medida por el índice de número de terneros por *vaca* (NTérmino). (Vet_CHL_03.txt).

La Tabla 6 compila la información estadística en el área de veterinaria. Al respecto, a diferencia de las demás áreas, para esta área la distribución entre las *unidades simples* y *complejas* es igual según *types*, es decir, 50 % en ambos casos. En 47 a 56 se presentan los ejemplos suministrados, automáticamente, por *Termostat* para cada categoría.

Tabla 6. Estadística extracción TermoStat - veterinaria.

	ÁREA	**VETERINARIA**			
	POS	***Types***		***Tokens***	
		Frecuencia	**%**	**Frecuencia**	**%**
Estructuras sintácticas	N	2392	33,0 %	58960	59,1 %
	N+Adj	1564	21,0 %	7660	7,7 %
	N+Prep +N	1335	18,0 %	5508	5,5 %
	Adj	752	10,0 %	15511	15,5 %
	N+N	440	6,0 %	1745	1,7 %
	V	457	6,0 %	8318	8,3 %
	N+Prep +N+Adj	207	3,0 %	655	0,7 %
	N+Adj+Adj	78	1,0 %	237	0,2 %
	Adv	50	1,0 %	1114	1,1 %
	N+Adj+Conj+Adj	36	0,5 %	85	0,1 %
	N+Adj+Adj+Adj	1	0,0 %	2	0,0 %
	N+Adj+Adj+Adj+Adj	0	0,0 %	0	0,0 %
Total de artículos		73			
Total de palabras		344202			
Types	Total de candidatos a término	7312			
	Unidades simples	50 %			
	Unidades complejas	50 %			
Tokens	Total de candidatos a término	99795			
	Unidades de simples	84,1 %			
	Unidades de complejas	15,9 %			

47. **N términos**: *animal* (NTérmino), *leche* (NTérmino), *peso* (NTérmino), *producción* (NTérmino).

    N no términos: *año* (NComún), *corregimiento* (NComún).

48. **[N+Adj] términos**: *época seca* (Tipo de), *estrés calórico* (Tipo de), *ganadería bovina* (Tipo de), *ganadería extensiva* (Tipo de), *materia seca* (Tipo de), *peso vivo* (Tipo de), *sistema productivo* (Tipo de), *producción ganadera* (Tipo de).

49. **[N+Prep+N] términos**: *cabeza de ganado* (Arbitraria), *emisión de metano* (Tipo de), *ganancia de peso* (Tipo de), *producción de leche* (Tipo de), *producción de metano* (Tipo de).

    [N+Prep+N] no términos: *número de animal* (No-Partit), *mes de edad* (No-Partit).

50. **Adj términos**: *ganadero* (AdjDenominal), *genético* (AdjDenominal), *natural* (AdjDenominal), *productivo* (AdjDenominal).

    Adj no términos: *diferente* (AdjGeneral), *potencial* (AdjGeneral), *total* (AdjGeneral).

51. **[N+N] términos**: *bienestar animal* (Hecho de), *carga animal* (Arbitraria).

    [N+N] que son [N+Adj] términos: *ganado lechero* (Tipo de), *mastitis subclínica* (Tipo de), *pasto ovillo* (Tipo de), *vaca lechera* (Tipo de).

52. **V términos**: *abrevar* (VTérmino), *criar* (VTérmino), *incubar* (VTérmino), *ordeñar* (VTérmino), *parasitar* (VTérmino), *pastar* (VTérmino), *producir* (VTérmino).

V no términos: *determinar* (VSoporte), *encontrar* (VSoporte), *indicar* (VDiscursivo), *mostrar* (VDiscursivo), *obtener* (VSoporte), *observar* (VDiscursivo), *presentar* (VDiscursivo), *realizar* (VSoporte), *utilizar* (VSoporte).

53. **[N+Prep+N+Adj] término**: *condición de estrés calórico* (Tipo de), *ganancia de peso diario* (Arbitraria), *sistema de producción ganadero* (Tipo de), *población de bovino criollo* (Tipo de/Hecho de), *tallo de ballico perenne* (Hecho de).

    [N+Prep+N+Adj] no términos: *aplicación de tecnología inadecuada* (No-Nomin), *cantidad de materia seca* (No-Nomin), *característica de composición corporal* (No-Nomin), *exportación de carne argentina* (No-Nomin).

54. **[N+Adj+Adj] términos**: *bosque húmedo premontano* (Tipo de), *carga óptima económica* (Tipo de), *célula alimentadora inactivada* (Tipo de), *ganado blanco orejinegro* (Tipo de), *población rural dispersa* (Tipo de), *sistema mixto familiar* (Tipo de).

55. **Adv términos**: *enzoóticamente* (Sí-AdvAdjetival-Mente), *epizootiológicamente* (Sí-AdvAdjetival-Mente), *estacionalmente* (Sí-AdvAdjetival-Mente), *fenotípicamente* (Sí-AdvAdjetival-Mente).

    Adv no términos: *aproximadamente* (AdjGeneral-Mente), *debajo* (AdvLugar), *donde* (AdvRelativo), *generalmente* (AdjGeneral-Mente), *principalmente* (AdjGeneral-Mente), *posteriormente* (AdjGeneral-Mente), *respectivamente* (AdjGeneral-Mente).

56. **[N+Adj+Conj+Adj] no términos**: *actividad económica y productiva* (No-Unireferencial), *estadio inmaduro y maduro* (No-Unireferencial), *mecha hemilustre y monocapa* (No-Unireferencial), *peso vivo y ancho* (No-Unireferencial), *región ocular y periorbital* (No-Unireferencial), *rendimiento productivo y económico* (No-Unireferencial).

4.1.4 Datos generales subcorpus de computación

El subcorpus de computación está conformado por 517.733 *tokens* a partir de 103 artículos compilados. La Figura 12 ilustra la nube de palabras de esta área. Al igual que en las demás áreas se presentan los conceptos sobresalientes en la construcción temática del campo. Entre otros, *algoritmo*, *ambiente virtual*, *conjunto finito*, *gramática de árbol* y *transductor*. En 57 a 61 se presentan ejemplos de aparición de estos términos en VaTeHo.



Figura 12. Nube de palabras computación.

57. […] Tamaño de los vectores de parámetros número entero y el *algoritmo* (NTérmino) con el que se han extraído la información de la forma de onda. El *algoritmo* (NTérmino) se pone entre llaves, como el resto de palabras clave. (Com_CHL_10.txt),

58. El *ambiente virtual* (Tipo de) de aprendizaje facilitó un estímulo para observar con mayor atención diseños y significados en los tejidos, en general, y aplicar conocimientos frente a los originales […]. (Com_CHL_07).

59. […] que consiste de una gramática de árbol regular y un *conjunto finito* (Arbitrario) de transductores. (Com_CHL_09).

60. Analizando cada una de las gramáticas, podemos aún eliminar e identificar no terminales, con el fin de encontrar las producciones

generativas que dan origen a la *gramática de árbol* (Tipo de) regular. (Com_CHL_09).

61. El *transductor* (NTérmino) es un array lineal de 128 canales con una frecuencia central de 5MHz y un ancho de banda del 65 % (PA 8617-B101, Imasonic). (Com_ESP_04).

La Tabla 7 compila la información estadística en el área de computación. En 62 a 71, se presentan los ejemplos suministrados, automáticamente, por Termostat para cada categoría.

Tabla 7. Estadística extracción TermoStat - computación.

	ÁREA	**COMPUTACIÓN**			
	POS	***Types***		***Tokens***	
		Frecuencia	**%**	**Frecuencia**	**%**
Estructuras sintácticas	N	686	24,0 %	7859	59,1 %
	N+Adj	471	16,0 %	875	6,6 %
	N+Prep +N	861	30,0 %	1342	10,1 %
	Adj	166	6,0 %	1086	8,2 %
	N+N	247	9,0 %	307	2,3 %
	V	176	6,0 %	1384	10,4 %
	N+Prep +N+Adj	160	6,0 %	194	1,5 %
	N+Adj+Adj	48	2,0 %	63	0,5 %
	Adv	25	1,0 %	153	1,2 %
	N+Adj+Conj+Adj	23	1,0 %	26	0,2 %
	N+Adj+Adj+Adj	2	0,1 %	2	0,0 %
	N+Adj+Adj+Adj+Adj	0	0,0 %	0	0,0 %
Total de artículos		103			
Total de palabras		517733			
Types	Total de candidatos a término	2865			
	Unidades simples	37 %			
	Unidades complejas	64 %			
Tokens	Total de candidatos a término	13291			
	Unidades simples	78,9 %			
	Unidades complejas	21,1 %			

62. **N términos**: *árbol* (NTérmino), *conjunto* (NTérmino), *método* (NTérmino), *modelo* (NTérmino), *sistema* (NTérmino), *proceso* (NTérmino).

    N no términos: *calidad* (NComún), *juego* (NComún), *proyecto* (NComún).

63. **[N+Adj] términos**: *ambiente virtual* (Metf hecho de), *capa lógica* (Tipo de), *conjunto finito* (Arbitraria), *dispositivo móvil* (Tipo de), *método ágil* (Tipo de), *programa computacional* (Tipo de), *transición interna* (Tipo de).

64. **[N+Prep+N] términos**: *gramática de árbol* (Metf hecho de), *base de datos* (Arbitrario), *algoritmo de inferencia* (Tipo de), *transductor de árbol* (Metf hecho de), *entorno de desarrollo* (Tipo de).

65. **Adj términos**: *adimensional* (AdjDenominal), *binario* (AdjDenominal), *interno* (AdjDenominal), *móvil* (AdjDenominal), *principal* (AdjDenominal), *real* (AdjDenominal), *virtual* (AdjRelacional).

    Adj no términos: *completo* (AdjDenominal), *diferente* (AdjGeneral), *siguiente* (AdjGeneral).

66. **[N+N] términos**: *alumno evaluador* (Hecho de), *árbol top-down* (Arbitraria), *árbol generador* (Metf hecho de), *archivo fuente* (Tipo de), *sistema generador* (Tipo de), *código fuente* (Arbitraria).

    [N+N] que son [N+Adj] términos: *estado ocupado* (Tipo de), *robot manipulador* (Tipo de), *sistema motriz* (Tipo de).

67. **V términos**: *automatizar* (VTérmino), *desarrollar* (VTérmino), *ejecutar* (VTérmino), *modelar* (VTérmino), *parametrizar* (VTérmino).

V no términos: *corresponder* (VSoporte), *indicar* (VSoporte), *mostrar* (VDiscursivo), *observar* (VDiscursivo), *presentar* (VDiscursivo), *realizar* (VSoporte), *utilizar* (VSoporte).

68. **[N+Prep+N+Adj] término**: *algoritmo de inferencia gramatical* (Tipo de), *entorno de desarrollo integrado* (Tipo de), *gramática de árbol regular* (Tipo de), *modo de oscilación relevante* (Tipo de), *prototipo de robot móvil* (Tipo de), *técnica de inteligencia artificial* (Tipo de).

69. **[N+Adj+Adj] términos**: *disparo secuencial alternado* (Tipo de), *esquina superior derecha* (Arbitrario), *esquina superior izquierda* (Arbitraria), *interfaz gráfica amigable* (Tipo de), *red neuronal especializada* (Tipo de).

70. **Adv términos**: *análogamente* (AdvDeadjetival-Mente), *intuitivamente* (AdvDeadjetival-Mente), *paramétricamente* (AdvDeadjetival-Mente).

    Adv no términos: *anteriormente* (AdvTiempo-Mente), *donde* (AdvLugar), *eventualmente* (AdvModo-Mente), *luego* (AdvTiempo), *posteriormente* (AdvTiempo-Mente), *respectivamente* (AdvGral-Mente), *usualmente* (AdvModo-Mente).

71. **[N+Adj+Conj+Adj] no términos**: *capacitor electrolítico y cerámico* (No-Unireferencial), *entrada lógico y natural* (No-Unireferencial), *generador eólico y fotovoltaico* (No-Unireferencial), *lado izquierdo y derecho* (No-Unireferencial), *parte real e imaginaria* (No-Unireferencial), *transición interna y externa* (No-Unireferencial).

4.1.5 Datos generales subcorpus de filosofía

El subcorpus de filosofía está conformado por 887.352 *tokens* a partir de 119 artículos compilados. La Figura 13 ilustra la nube de palabras de esta área. Los conceptos que sobresalen en su construcción

temática son, entre otros, *conocimiento*, *epistemología*, *filosofía*, *hombre natural* y *testimonio moral*. En 72 a 76, se presentan ejemplos de aparición de estos términos en VaTeHo.



Figura 13. Nube de palabras filosofía.

72. En Martí, la pregunta por la subjetividad nuestro-americana nos dirige al interrogante por el *conocimiento* (NProductMorf). (Fil_ARG_01).

73. La sociología complementaría a la *epistemología* (NProductMorf) manteniendo ésta su independencia y capacidad explicativa, Con Kuhn, por ejemplo, las reconstrucciones racionales de la ciencia. (Fil_ESP_06).

74. La historia de la *filosofía* (NCompCult), del derecho y de la política registra casos excepcionales. (Fil_DOM_03).

75. El *hombre natural* (Tipo de) es el sujeto de la resistencia y de la liberación en Nuestra América. (Fil_ARG_01).

76. Es altamente plausible que el *testimonio moral* (Tipo de) pueda proporcionar razones explicativas para una creencia u acción. (Fil_ARG_09).

La Tabla 8 compila la información estadística en el área de computación. En 77 a 86 se presentan los ejemplos suministrados, automáticamente, por *Termostat* para cada categoría.

Tabla 8. Estadística extracción TermoStat - filosofía.

	ÁREA	FILOSOFÍA			
	POS	*Types*		*Tokens*	
		Frecuencia	%	Frecuencia	%
Estructuras sintácticas	N	904	32,0 %	21360	50,7 %
	N+Adj	783	27,0 %	3109	7,4 %
	N+Prep +N	259	9,0 %	857	2,0 %
	Adj	428	15,0 %	7916	18,8 %
	N+N	79	3,0 %	231	0,5 %
	V	262	9,0 %	6839	16,2 %
	N+Prep +N+Adj	48	2,0 %	118	0,3 %
	N+Adj+Adj	20	1,0 %	47	0,1 %
	Adv	60	2,0 %	1577	3,7 %
	N+Adj+Conj+Adj	25	1,0 %	68	0,2 %
	N+Adj+Adj+Adj	0	0,0 %	47	0,1 %
	N+Adj+Adj+Adj+Adj	0	0,0 %	0	0,0 %
Total de artículos		119			
Total de palabras		887352			
Types	Total de candidatos a término	2868			
	Unidades simples	58 %			
	Unidades complejas	43 %			
Tokens	Total de candidatos a término	42169			
	Unidades de simples	89,4 %			
	Unidades de complejas	10,6 %			

77. **N términos**: *ciencia* (NComún), *concepto* (NComún), *conocimiento* (NProductMorf), *idea* (NComún), *sentido* (NComún), *teoría* (NComún), *verdad* (NComún).

    N no términos: *caso* (NComún), *forma* (NComún), *relación* (NComún).

78. **[N+Adj] términos**: *autonomía epistémica* (Tipo de), *ciencia social* (Tipo de), *conocimiento científico* (Tipo de), *hombre natural* (Tipo de), *testimonio moral* (Tipo de), *teoría social* (Tipo de), *ser humano* (Hecho de), *reflexión epistemológica* (Tipo de).

79. **[N+Prep+N] términos**: *epistemología del testimonio* (Tipo de), *escuela de Frankfurt* (Arbitraria), *filosofía del lenguaje* (Tipo de), *campo de estudio* (Metf hecho de).

    [N+Prep+N] no términos: *áreas de estudio* (No-Nomin), *definición de idea* (No-Nomin), *tipo de conocimiento* (No-Prop).

80. **Adj términos**: *científico* (AdjDenominal), *empírico* (AdjDenominal), *histórico* (AdjDenominal), *natural* (AdjDenominal), *social* (AdjDenominal), *teórico* (AdjDenominal).

    Adj no términos: *mismo* (AdjGeneral), *propio* (AdjGeneral).

81. **[N+N] términos**: *conocimiento mapuche* (Tipo de), *experimentación animal* (Tipo de).

    [N+N] que son [N+Adj] términos: *deseo martiano* (Tipo de), *época moderna* (Tipo de), *teoría verificacionista* (Tipo de/ Hecha de).

    [N+N] que son [Pron+N] no términos: *nuestra creencia* (No-Ambig), *nuestro día* (No-Ambig).

82. **V términos**: *afirmar* (VTérmino), *entender* (VSoporte), *evidenciar* (VSoporte), *ser* (VComún).

83. **[N+Prep+N+Adj] término**: *criterio de compromiso ontológico* (Tipo de), *dialéctica de fuerza productiva* (Tipo de), *doctrina del flujo perpetuo* (Tipo de), *proceso de inflexión epistémica* (Tipo de), *proceso de refutación empírica* (Tipo de).

    [N+Prep+N+Adj] no términos: *cuestión del testimonio moral* (No-Ambig), *éxito de argumento concreto* (No-Ambig).

84. **[N+Adj+Adj] términos**: *campo filosófico argentino* (Tipo de), *filosofía analítica anglosajona* (Tipo de), *modelo epidemiológico clásico* (Tipo de), *proceso educativo intercultural* (Tipo de/ Hecho de).

[N+Adj+Adj] no términos: *campo filosófico argentino* (No-Ejemplo de).

85. **Adv términos**: *acríticamente* (Adv-Deadjetival-Mente), *cognoscitivamente* (Adv-Deadjetival-Mente), *epistemológicamente* (Adv-Deadjetival-Mente), *ontológicamente* (Adv-Deadjetival-Mente).

    Adv no términos: *así* (AdvDem), *allá* (AdvDem), *allí* (AdvDem), *como* (AdvRel), *donde* (AdvLug), *entonces* (AdvDem), *siempre* (AdvFre).

86. **[N+Adj+Conj+Adj] no términos**: *actividad social y cultural* (No-Unireferencial), *autonomía práctica y epistémica* (No-Unireferencial), *epistemología contextual e intercultural* (No-Unireferencial), *facultad sensitiva e intelectiva* (No-Unireferencial), *idea clara y distinta* (No-Unireferencial), *institución científica y política* (No-Unireferencial), *medio natural y espiritual* (No-Unireferencial), *refutación progresiva y parcializada* (No-Unireferencial).

4.1.6 VaTeHo: Procesamiento de los datos a partir de la extracción en bruto

Una vez compilados los datos más relevantes en la extracción automática realizada con *TermoStat* se procedió a establecer los criterios bajo los cuales se constituyó la muestra de trabajo con miras a la aplicación del protocolo de validación (§ 4.2). Como se mencionó, la extracción automática realizada con esta herramienta arrojó un total aproximado de 37.055 candidatos a término. Es de conocimiento general, en el ámbito de la lingüística computacional y de corpus, que la mayoría, por no mencionar todas, las herramientas automáticas utilizadas en procesos de extracción o recuperación de información presentan un margen de error en sus resultados. Así, teniendo esto en cuenta, se aplicó una serie de condiciones para lograr una limpieza general de la *extracción en bruto* a partir de VaTeHo que permitió, por un lado, revisar y ajustar los errores en la extracción y, por otro lado, reducir los datos iniciales en un porcentaje cercano al 34,20 %,

esto es, el porcentaje de ruido en los datos obtenidos (§ 3.2.3). Así, el objetivo con esta limpieza temática se orientó a obtener una lista de candidatos depurada que permitió generar la muestra de trabajo y, en consecuencia, una aplicación más efectiva y eficaz del protocolo de validación a partir de la correcta implementación de estas condiciones de limpieza. La Figura 14 representa los resultados de esta limpieza inicial. Además, en 87 a 92 se listan ejemplos de los candidatos omitidos según las condiciones ya descritas. Estos se presentan con las etiquetas recuperadas a partir de la extracción.

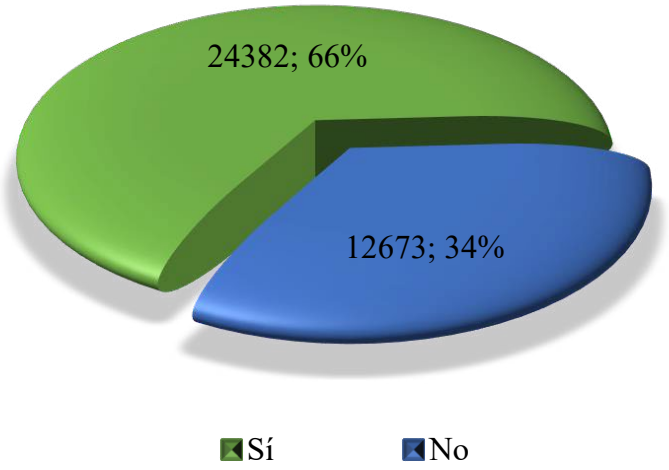


Figura 14. Porcentajes limpieza temática.

87. **Revisión etiquetaje**: *aquello* (N), *cognoscitivo* (V), *criollo* (V), *fique* (V), *Haverkate* (V), *leve* (Adv), *Llorente* (V), *nuestra* (N), *nuestra época* ([N+N]), *sinónimo* (Adj).

88. Omisión de candidatos que no corresponden al área temática:

    a) **Medicina**: *algoritmo* (N), *análisis de correlación* ([N+Adj]), *combustión del diésel* ([N+Prep+N]), *paquete estadístico* ([N+Adj]).

b) **Lingüística**: *astronómico* (Adj), *botánico* (Adj), *celular* (Adj), *coeficiente de confiabilidad* ([N+Prep+N]), *desviación standard* ([N+Adj]), *mamífero carnívoro* ([N+Adj]), *ombligo* (N).

c) **Veterinaria**: *análisis de correlación* ([N+Prep+N]), *análisis descriptivo* ([N+Adj]), *brecha tecnológica* ([N+Adj]), *coeficiente de heterocigocidad* ([N+Prep+N]), *vector de coeficiente* ([N+Prep+N]).

d) **Computación**: *artesanía regional* ([N+Adj]), *asignatura de historia prehispánica* ([N+Prep+N+Adj]), *filosofía* (N), *precolombino* (N), *proteína* (N).

e) **Filosofía**: *adverbio modal* ([N+Adj]), *biología molecular* ([N+Adj]), *agroquímico* (N), *cuantitativo* (Adj), *geométrico* (Adj), *matemático* (Adj), *neurobiología* (N).

89. **Omisión de candidatos relativos a la diagramación**: *figura* (N), *gráfico* (N), *gráfico anterior* ([N+Adj]), *ilustración* (N), *tabla de contingencia* ([N+Prep+N]), *tabla siguiente* ([N+Adj]).

90. **Omisión de candidatos en otras lenguas distintas el español**: *correctly* (N), *fishman* (V), *introduction* (V), *language dictionary* ([N+N]), *organized* (V), *phone* (V), *recently created* ([Adv+V]), *verba dicendi* ([N+V]).

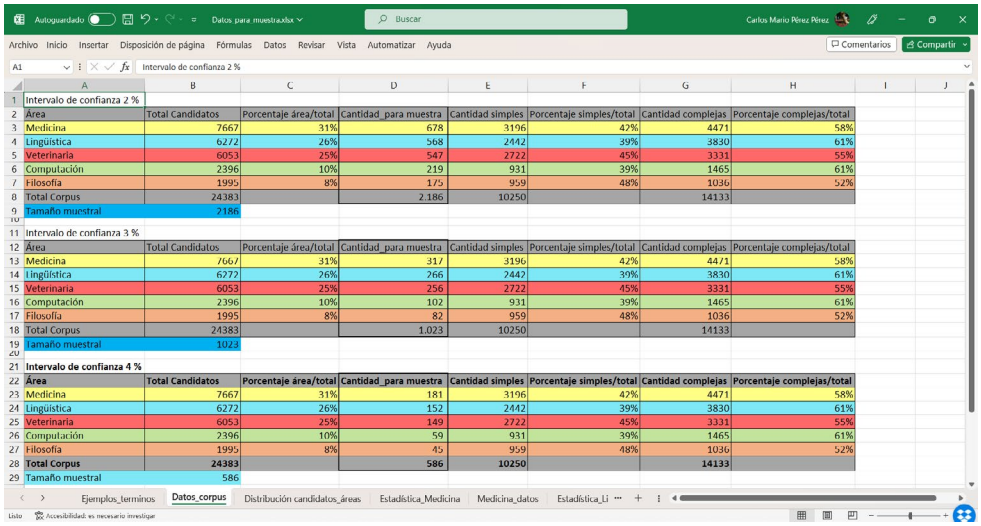
91. **Omisión de candidatos relativos a unidades de medida, tiempo y espacio**: *cm de altura* ([N+Prep+N), *cm de diámetro* ([N+Prep+N]), *mg* (N), *ml* (N).

92. **Omisión de errores de conversión**: *b[a* (Unknown), *invierno primav* ([N+N]), *manej* (N), *nse rvació* ([N+N]), *nte nido* ([N+N]), *qu'el rey* ([Prep+Art+N]), *rganizacio ne* ([N+N]).

4.1.7 VaTeHo: Muestra de candidatos a término para la validación

Luego de la revisión manual de cada uno de los candidatos extraídos y finalizada la limpieza temática, se procedió con la conformación de la muestra de datos para la aplicación del protocolo de validación. Inicialmente, se realizó un piloto de validación con el objetivo de refinar el procedimiento paso a paso y observar, en la realidad, la dinámica de trabajo en este estadio de la investigación. El tiempo promedio de validación de cada uno de los candidatos osciló alrededor de los cinco minutos. Así, se realizó el cálculo del tamaño muestral a partir los siguientes parámetros: nivel de confianza del 95 %, con un intervalo de confianza del 4 % y homogeneidad del 0,5 %. Asimismo, se calculó la muestra con intervalos de confianza de 2 y 3 % para, de esta manera, poder identificar la variación en la cantidad de datos según estos parámetros y tomar la decisión final en la conformación de la muestra. Al respecto, esta decisión responde a criterios metodológicos con relación al tiempo necesario para la validación *versus* el tiempo disponible y el recurso humano para tal propósito.

En este sentido, la validación del intervalo de confianza del 2 % tomaría 192 horas, la del 3 % 85 horas y la del 4 % 49 horas, aproximadamente. En la Figura 15 se muestran los cálculos muestrales según cada intervalo de confianza. Asimismo, de acuerdo con los cálculos muestrales y con miras a la distribución de los 586 candidatos a términos a validar, se procedió con la asignación de la información a compilar por cada una de las áreas trabajadas y por cada una de las unidades extraídas. De este modo, la muestra quedo conformada por 181 candidatos en medicina correspondientes al 31,05 % del total de la muestra, 152 en lingüística el 25,93 %, 149 en veterinaria el 25,08 %, 59 en computación el 10,06 % y 45 en filosofía el 8,02 %. Se realizó un muestro estratificado, aleatorio e intencionado tomando los candidatos por orden descendente de mayor a menor frecuencia. La Tabla 9 compila la distribución final de candidatos muestrales.



Intervalo de confianza 2 %							
Área	Total Candidatos	Porcentaje área/total	Cantidad_para muestra	Cantidad simples	Porcentaje simples/total	Cantidad complejas	Porcentaje complejas/total
Medicina	7667	31%	678	3196	42%	4471	58%
Lingüística	6272	26%	568	2442	39%	3830	61%
Veterinaria	6053	25%	547	2722	45%	3331	55%
Computación	2396	10%	219	931	39%	1465	61%
Filosofía	1995	8%	175	959	48%	1036	52%
Total Corpus	24383		2.186	10250		14133	
Tamaño muestral	2186						

Intervalo de confianza 3 %							
Área	Total Candidatos	Porcentaje área/total	Cantidad_para muestra	Cantidad simples	Porcentaje simples/total	Cantidad complejas	Porcentaje complejas/total
Medicina	7667	31%	317	3196	42%	4471	58%
Lingüística	6272	26%	266	2442	39%	3830	61%
Veterinaria	6053	25%	256	2722	45%	3331	55%
Computación	2396	10%	102	931	39%	1465	61%
Filosofía	1995	8%	82	959	48%	1036	52%
Total Corpus	24383		1.023	10250		14133	
Tamaño muestral	1023						

Intervalo de confianza 4 %							
Área	Total Candidatos	Porcentaje área/total	Cantidad_para muestra	Cantidad simples	Porcentaje simples/total	Cantidad complejas	Porcentaje complejas/total
Medicina	7667	31%	181	3196	42%	4471	58%
Lingüística	6272	26%	152	2442	39%	3830	61%
Veterinaria	6053	25%	149	2722	45%	3331	55%
Computación	2396	10%	59	931	39%	1465	61%
Filosofía	1995	8%	45	959	48%	1036	52%
Total Corpus	**24383**		**586**	**10250**		**14133**	
Tamaño muestral	586						

Figura 15. Cálculos tamaño muestral.

Tabla 9. Conformación final de la muestra de trabajo por candidato y por área.

Patrón/cantidad_área	Medicina	Lingüística	Veterinaria	Computación	Filosofía	Total
Adj	23	20	15	3	6	67
Adv	1	1	1	0	0	3
N	48	33	43	17	14	155
N+Adj	53	56	35	8	16	168
N+Adj+Adj	4	2	2	1	0	9
N+Adj+Conj+Adj	1	1	1	1	0	4
N+N	7	6	9	3	1	26
N+Prep +N	32	23	30	20	5	110
N+Prep +N+Adj	8	4	5	3	1	21
V	4	6	8	3	2	23
Total	**181**	**152**	**149**	**59**	**45**	**586**

4.2 VaTeHo: Aplicación de un protocolo para la validación de candidatos a término

La muestra de trabajo quedó conformada por 248 candidatos a término simples y 338 complejos que dan cuenta de 10 estructuras sintácticas con predominancia de las estructuras nominales atendiendo a la mirada paradigmática de la terminología. De este modo, el 80,54 % de los candidatos analizados corresponde a este tipo de construcciones.

La Figura 16 da cuenta de la distribución total de los candidatos. Asimismo, en los siguientes segmentos se detallan los resultados por cada una de las áreas del saber junto con ejemplos recuperados del corpus para ilustrar cada uno de los hallazgos. Con relación a la muestra es importante mencionar que esta se definió, principalmente, por criterios de frecuencia. Al respecto, consideramos que el factor frecuencia es fundamental en el establecimiento del discurso de los distintos saberes. Así, reconocemos en nuestra tesis que son tan importantes las unidades con mayor frecuencia al igual que las de menor en cuanto en ambos extremos se consolida y materializa el conocimiento de cada área del conocimiento. En consecuencia, en algunos de los casos a desarrollar en adelante, se tomarán ejemplos más allá de la muestra analizada en función de ilustrar con suficiencia cada condición descrita, según las particularidades de cada una de las unidades analizadas y para evitar la tendencia del paradigma. Por último, no se realizó ningún tipo de corrección ortotipográfica de los contextos recuperados con el objetivo de conservar la originalidad de la fuente.

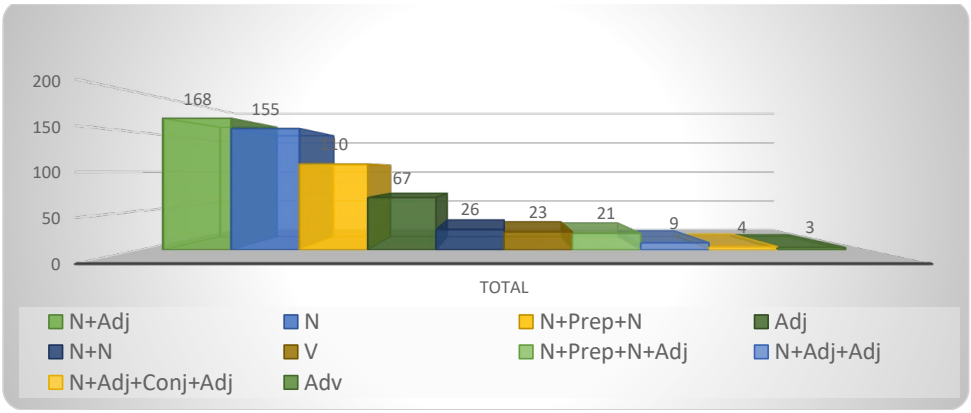


Figura 16. Distribución de las estructuras sintácticas en la muestra de trabajo.

4.2.1 Hallazgos terminológicos en medicina

De los 181 candidatos analizados en medicina, se determinó que 142 corresponden a *términos* que dan cuenta directamente de un

concepto al interior del área, mientras que los 39 candidatos restantes corresponden a unidades que contribuyen en la construcción del discurso temático más no representan un nodo cognitivo directamente. Por un lado, la Tabla 10 da cuenta los *términos* en medicina y la categoría según el protocolo. Además, en 93 a 105 se desarrollan ejemplos de cada categoría con sus correspondientes contextos de aparición. Por otro lado, en la Tabla 11 se relaciona la información sobre las unidades *no término* en el área seguida de las correspondientes ejemplificaciones en 106 a 110.

Tabla 10. Etiquetas términos en medicina según validación por protocolo.

	Criterio	Tipo de candidato	Medicina
	AdjDenominal	Simple	4
	AdjGeneral	Simple	9
	AdjRelacional	Simple	7
	AdvAdjetival-Mente	Simple	1
	NCompCulta	Simple	3
	NComún	Simple	0
	NMorfSubj	Simple	5
Validar	NProductMorf	Simple	7
	NTérmino	Simple	28
	VTérmino	Simple	4
	Arbitrario	Complejo	8
	Hecho de	Complejo	9
	Metf hecho de	Complejo	0
	Tipo de	Complejo	60
	Tipo de/Hecho de	Complejo	4

93. **AdjDenominal**:

— Los resultados oncológicos ofrecidos por las diferentes vías de abordaje (*vaginal*, *abdominal*, laparoscópica) se

consideran equiparables si son practicadas por personal con suficiente experiencia (7). (Med_URY_10.txt).

— […] el remanente hepático debía cumplir con: tener al menos 2 segmentos hepáticos en continuidad, irrigación portal, *arterial* y drenaje biliar (o al intestino) y corresponder al menos al 20-25% del volumen hepático […] (Med_CHL_01. txt).

— El diagnóstico histopatológico principal fue de tipo *ductal* en 91% de los casos, de los cuales, 7% fueron in situ. (Med_ COL_08.txt).

94. **AdjGeneral**:

— Este marcador tumoral se hizo *positivo* (mas de 4 ng.) en el 45% y en el 20% en el Hospital del seguro y en el I.G.B.J respectivamente. (Med_BOL_03.txt).

— Los 9 pacientes recibieron 4 o más ciclos de quimioterapia previo a la cirugía del *primario* con un rango de 4 a 24 ciclos. (Med_CHL_09.txt).

— La escala de Glogaw se utiliza para estimar el grado de foto envejecimiento clasificándolo en 4 tipos. El tipo 1 (*temprano*) se caracteriza por ausencia de arrugas, cambios pigmentarios leves, ausencia de queratosis y es típico de las personas de entre 20 y 30 años. (Med_PAN_03.txt).

95. **AdjRelacional**:

— El comportamiento *epidemiológico* de los tumores malignos del país ha presentado un cambio en los últimos años; el cáncer de mama ha aumentado su incidencia hasta colocarse como […] (Med_CRI_04.txt).

— Cuando se incorpora el resultado del estudio *histológico* se logra la localización del tumor primario en más del 70 % de las ocasiones. (Med_CUB_06).

— Los estándares de quimioterapia y tratamiento *quirúrgico* se cumplieron en el grupo de pacientes reportadas en este estudio. (Med_MEX_02).

96. **AdvAdjetival-Mente**:

— El ovario puede originar tumores *histológicamente* muy diferentes, con un comportamiento biológico agresivo, que en su mayoría al presentarse en mujeres jóvenes pre y posmenopáusicas [...]. (Med_VEN_04).

— Se estudiaron *inmunohistoquimicamente* cortes seriados (4-pm-grosor) de cada muestra para determinar la expresión del VEGF y sus receptores mediante anticuerpos de conejo policlonales anti-human VEGF. (Med_HND_09).

— Existen estudios que evidencian que la mortalidad en los adultos mayores intervenidos *quirúrgicamente* depende del retraso en la cirugía y las condiciones previas, y que la edad avanzada (mayor de 70 u 80 años) [...]. (Med_PAN_08).

97. **NCompCulta**:

— Un paciente falleció después del segundo curso de *quimioterapia*, debido a obstrucción intestinal. (Med_PER_05).

— El tratamiento con *radioterapia* consistió en la administración de 70-74 Gy sobre la localización primaria del tumor en régimen de fraccionamiento normal o hiperfraccionamiento. (Med_ESP_08).

— [...] algunos pacientes, puesto que supone una alternativa válida y con ciertas localizaciones como la orofaringe parece ser superior y aporta claras ventajas no precisar ni *traqueotomía* ni mandibulotomía por ejemplo. (Med_ESP_08).

98. **NMorfSubj**:

— Categoría 6: lesiones con malignidad demostrada mediante *biopsia*, previa a terapias definitivas (cirugía, radioterapia o quimioterapia), y por lo tanto no se debe confirmar su malignidad. (Med_BOL_10).

— El cuadro clínico de esta *patología* esta en dependencia del segmento intestinal afectado así en los de localización derecha suele presentarse como un cuadro de anemia y sangrado crónico oculto. (Med_PRY_03).

— Los nodulos linfáticos también se consideran factor *pronóstico* de sobrevida, en este estudio se asociaron con dificultad para lograr una resección sin márgenes infiltrados (p=0.023) con riesgo aumentado de R1 de dos veces [...]. (Med_HND_08).

99. **NProductMorf**:

— El pronóstico de estos pacientes sin *tratamiento* es malo, con sobrevida menor a 30% al año y menor al 5% a los 5 años[3]. Sin embargo, con *tratamiento* adecuado el pronóstico mejora [...]. (Med_CHL_09).

— La mediana de tamaño de la *lesión* fue de 20 mm y sólo 19.4 % presentaba metástasis extrapulmonares previo al procedimiento. (Med_PAN_01).

— El enfoque terapéutico realizado se basó en *cirugía* (Cx), radioterapia (RT), quimioterapia (QT), concomitante (QT/RT) o secuencial (QT+RT o RT+QT) y sus diferentes combinaciones [...]. (Med_COL_01).

100. **NTérmino**:

— En Europa, teniendo en cuenta ambos sexos, fue el segundo tipo de *cáncer* con mayor incidencia, según las estimaciones del año 2014. (Med_ESP_03).

— Para cada uno de ellos se documentó el sitio del tumor primario, presencia de *metástasis*, edad, sexo e intensidad del dolor así como efectos adversos relacionados con la morfina. (Med_URY_09).

— Al definir la técnica quirúrgica fue imprescindible la localización del *tumor* dependiendo de los cuadrantes mamarios. (Med_CUB_04).

101. **VTérmino**:

— La importancia de la palpación radica en que puede detectar cualquier alteración que no haya sido identificada por la inspección. (Med_ECU_02).

— Si bien, en la actualidad es prácticamente imposible diagnosticar con precisión el número y la extensión anatómica de metástasis de ganglios linfáticos [...]. (Med_PER_09).

— Existen diversos factores de riesgo para padecer este tipo de cáncer, entre los que se encuentran la edad avanzada, el sexo masculino, factores genéticos [...]. (Med_GTM_02).

102. **Arbitrario**:

— Las técnicas paliativas han mejorado la *calidad de vida* de los pacientes, sin prolongar la sobrevida a mediano ni largo plazo. (Med_URY_03).

— Debido a que el comportamiento y al *estilo de vida* que contribuyen al riesgo y al pronóstico del cáncer, se reconoce la necesidad de incluir a las familias [...]. (Med_ECU_06).

— En la región sureste del país, la relación entre el *exceso de peso* y los tumores mamarios, donde la prevalencia de sobrepeso y obesidad es muy alta, el riesgo podría estar incrementado en comparación con otras poblaciones. (Med_MEX_04).

103. **Hecho de**:

— En estos casos no queda claro si el *factor de riesgo* es genético o adquirido por exposición a otras sustancias. (Med_PAN_02.txt).

— La implementación del tamizaje está destinada a disminuir las *tasas de mortalidad*, mas no las de *incidencia* (12). (Med_PER_02.txt).

— Para saber si el *tratamiento hormonal* se relacionó con un mayor control estudiamos la supervivencia libre de RB a los 5 años en casos con HT [...]. (Med_ESP_10).

104. **Tipo de**:

— El informe de anatomía patológica de la *biopsia endoscópica* fue realizado en forma independiente y no influyó en el informe de la biopsia de la pieza operatoria. (Med_CHL_04).

— El diagnóstico de *cáncer cervicouterino* se cree que está relacionado a una conducta sexual promiscua y a no cumplir con estándares religiosos. (Med_NIC_02).

— La *cirugía reconstructiva* ha evolucionado muchísimo con el uso de colgajos que permiten movilizar tejido con características similares al territorio tisular perdido, buscando así el mejor resultado funcional [...] (Med_HND_07.txt).

105. **Tipo de/Hecho de**:

— Las nuevas técnicas de *citometría de flujo* e inmunohistoquímica permiten actualmente a la citología no solo realizar diagnóstico, sino también analizar el comportamiento biológico del tumor [...]. (Med_PER_08).

— Los pacientes fueron pesados y tallados y se les calculó el *Índice de Masa Corporal* para clasificarlos de acuerdo al tipo de neoplasias malignas. (Med_DOM_05).

— El cáncer de mama constituye un *problema de salud pública* a nivel mundial y representa el tipo de cáncer más frecuente en el sexo femenino. (Med_VEN_08).

Tabla 11. Etiquetas no términos en medicina según validación por protocolo.

	Criterio	Tipo de candidato	Medicina
	No - AdjGeneral	Simple	3
	No - NComún	Simple	5
	No - VSoporte	Simple	0
	No - Acus	Complejo	3
	No - Deadj	Complejo	0
Descartar	No - Ejemplo de	Complejo	0
	No - Nomin	Complejo	0
	No - Ocurre en	Complejo	0
	No - Partit	Complejo	2
	No - Prop	Complejo	19
	No - Unireferencial	Complejo	0

106. **No – AdjGeneral**:

— Una medición más fina de los riesgos debido a la exposición a plaguicidas requeriría un estudio más *grande* realizado en las personas, probablemente un estudio de casos y controles […]. (Med_CRI_10).

— El paciente más *joven* de nuestro grupo era un adolecente de 17 años con una adenocarcinoma de la unión esófago gástrica que duró 12 meses con síntomas. (Med_COL_03).

— Por esta razón, es especialmente útil para este tipo de estudio, que utilizan un número *pequeño* de pacientes. (Med_VEN_03.txt).

107. **No – Ncomún**:

— En el modo secuencial, una *serie* de imágenes se adquieren de forma secuencial en un período de tiempo, que se puede ajustar previamente en la cámara. (Med_ECU_09.txt).

— Es necesario realizar un adecuado *registro* de las evaluaciones y del seguimiento de los pacientes en la historia clínica. (Med_VEN_01.txt).

— De esto la importancia de la *educación* para la salud de modo de que estos temores no sean una barrera para la realización de chequeos de rutina en las mujeres. (Med_NIC_02.txt).

108. **No – Acus**:

— La mamografía con contraste se realizó mediante la *adquisición de imágenes* de baja y alta energía en forma sucesiva durante una sola compresión […]. (Med_CHL_08).

— Nuestros resultados podrían deberse a la *ubicación* de la población estudiada en una zona con antecedente de ser endémica de bocio, con alta incidencia de bocio multinodular. (Med_VEN_01.txt).

— En cuanto a la *distribución de pacientes* según las variables clínicas iniciales no se describieron diferencias entre los grupos aunque hubo un predominio del estado funcional general entre 0 y 1 en los dos casos. (Med_CUB_05).

109. **No – Partit**:

— En el grupo 2 se observa una amplia mayoría de pacientes con tratamiento radiante, una menor *cantidad de prostatectomías radicales* que en el grupo 1, y ya se comienza a observar un incremento de las terapias de privación androgénica. (Med_ARG_03).

— El *número de ganglios linfáticos* metastásicos oscilaron entre 0-29 (media, 3,9±5,8) y el *número de ganglios linfáticos resecados* comprendió los rangos de 1366 (media, 35,34±12,60). (Med_PER_09).

— Es el tumor que mayor *número de muertes* produce en la mujer en Cuba,[3] donde se diagnostican actualmente más de 20 mil nuevos casos al año; de ellos, más de 2 200 se localizan en las mamas femeninas. (Med_CUB_01).

110. **No – Prop**:

— Este *tipo de tumor* tiene características que se pueden estudiar con las técnicas de inmunohistoquímica, como son la expresión de receptores en las células neoplásicas [...] (Med_CRI_04.txt).

— El cáncer pulmonar es la *causa de muerte* por tumores malignos más frecuente. En Panamá, ocupó la tercera *causa de muerte* por cáncer en 2009. (Med_PAN_02).

— En Venezuela, el cáncer de mama se ubicó en el segundo lugar de *mortalidad por cáncer* en mujeres en 2005, año en que ocurrieron 1 441 defunciones por esta patología y se diagnosticaron 3 408 casos nuevos [...]. (Med_VEN_10).

4.2.2 Hallazgos terminológicos en lingüística

Con relación a los hallazgos en lingüística, de los 152 candidatos analizados, se determinó que 102 corresponden a *términos* que representan directamente un concepto al interior del área, mientras que los 50 candidatos restantes corresponden a unidades que contribuyen en la construcción del discurso temático más no representan un nodo cognitivo directamente. Por un lado, la Tabla 12 da cuenta los *términos* en lingüística y la categoría según el protocolo. Además, en 111 a 120 se desarrollan ejemplos de cada categoría con sus correspondientes contextos de aparición. Por otro lado, en la Tabla 13 se relaciona la información sobre las unidades *no término* en el área seguida de las correspondientes ejemplificaciones en 121 a 126.

Tabla 12. Etiquetas términos en lingüística según validación por protocolo.

	Criterio	Tipo de candidato	Lingüística
	AdjDenominal	Simple	13
	AdjGeneral	Simple	3
	AdjRelacional	Simple	4
	AdvAdjetival-Mente	Simple	1
	NCompCulta	Simple	0
	NComún	Simple	0
	NMorfSubj	Simple	0
Validar	NProductMorf	Simple	3
	NTérmino	Simple	29
	VTérmino	Simple	4
	Arbitrario	Complejo	22
	Hecho de	Complejo	3
	Metf hecho de	Complejo	0
	Tipo de	Complejo	41
	Tipo de/Hecho de	Complejo	0

111. **AdjDenominal**:

— Por ello, aunque el uso de esta perífrasis es perfectamente *gramatical* no es la forma adecuada para la interpretación deseada puesto que el mensaje transmitido no es, ni mucho menos, el mismo. (Lin_ESP_06).

— También se empleaba la reduplicación (pudiendo ser un enfático o un distributivo; no es claro) y un sufijo {-k^{w}aw}, cuya marca de palabra era *nominal* (es decir, terminaba con absolutivo). (Lin_MEX_10.txt).

— El diccionario— presentan la marca mf, o sea, masculino y femenino, la lengua fuente, el guaraní, no posee flexión; por tanto, no hace falta indicarla, así si define con un equivalente *adjetival*, el lema habrá de suponer que sea de la misma categoría. (Lin_PRY_01.txt).

112. **AdjGeneral**:

— También en La Nación, se elide el agente, junto con el *auxiliar* de las construcciones pasivas. (Lin_CRI_10.txt).

— En el corpus estudiando, este moralizador aparece en primera persona *plural*, incluyente, lo que refuerza la idea de que esta forma tiene una función de presentación de la información como "negociada" con el enunciatario. (Lin_VEN_07).

— Por ejemplo, en aquellos casos en los cuales el rol temático del argumento eliminado no se encuentra *implícito* (i.e., ha sido eliminado junto con el argumento) la operación debe tener lugar en el léxico. (Lin_ARG_06.txt).

113. **AdjRelacional**:

— En lo que se refiere a los tipos de definición, plantea Cabakian, predomina la de tipo enciclopédico frente a la de tipo *lingüístico*, "dado que se describen los entes de la realidad a los que se asocian las palabras seleccionadas como lemas" (p. 9). (Lin_URY_01.txt).

— Por todo esto, es válido decir que el parlache es un dialecto social de carácter *argótico*, que han creado, como ya se dijo, los jóvenes de los sectores marginales y populares de Medellín, para expresar la nueva realidad a la que se enfrentan cada día. (Lin_COL_05.txt).

— Se dice que un patrón es *productivo* cuando puede aplicarse a múltiples bases para producir nuevas palabras. (Lin_MEX_03).

114. **AdvAdjetival-Mente**:

— Se seleccionaron todas las palabras que han sido modificadas *morfológicamente* por un sufijo apreciativo, sea este diminutivo, aumentativo o despectivo. (Lin_VEN_06).

— [...] que posee entradas plenamente especificadas con información gramatical; en el segundo es una lista de piezas que dan contenido fonológico (y conceptual) a una estructura sintáctica y que se insertan *postsintácticamente*. (Lin_ESP_05).

— Descartadas ambas etimologías, formal y *semánticamente* gratuitas, es hora de sugerir una alternativa de interpretación que supere las deficiencias mencionadas. (Lin_PER_10).

115. **NProductMorf**:

— Por eso se prefiere entonces la *concordancia* plural, que es la que permite interpretar que embarazada se predica de un elemento 'mujer', parte de una multiplicidad de elementos iguales, y no de la mitad de cada entidad [...]. (Lin_ESP_10).

— Se consideran como tales ciertas unidades lingüísticas cuya *definición* está compuesta por rasgos semánticos muy genéricos, y por ello están dotadas de una intensión mínima y una extensión máxima (Koch y Oesterreicher 2007). (Lin_ARG_02).

— Se preocupa por un aspecto en particular: la *etimología* o el origen de las voces, cuando éstas se vinculan con las lenguas indígenas de la región. (Lin_URY_01).

116. **NTérmino**:

— Se entiende como macroestructura del *diccionario* como el conjunto de las entradas que conforman su nomenclatura, o sea, el conjunto de palabras que recoge, incluye tanto la selección y la ordenación de las entradas que [...]. (Lin_PRY_01).

— Lo más importante de esta teoría es que la *palabra* queda unida en su contorno a otra *palabra*, englobadas en una unidad semántica mayor denominada red, misma que permite, de cierta forma, engarzar las unidades menores con las mayores. (Lin_MEX_01).

— Dicho esto, resulta hasta cierto punto predecible dar cuenta del proceso de expansión del *término*, ya que se puede ir directamente al momento de aparición del objeto en cuestión para calcular el inicio del uso de la palabra. (Lin_MEX_02).

117. **VTérmino**:

— Por otra parte, *adjetivar* una realidad es llenarla de contenido descriptivo; si bien la nominalización pretende nombrarla, la adjetivación dota a la realidad de características propias de acuerdo con los valores axiológicos de la [...]. (Lin_COL_08).

— El léxico que son capaces de reunir y *definir* será el más representativo para mostrar cómo las nacientes repúblicas se seguían entendiendo lingüísticamente, hijas del hispanismo colonial y cómo, además, comenzaban a consolidar un nuevo léxico coloquial, referencial [...]. (Lin_URY_04).

— En las oraciones se observa una tendencia consistente a *elidir* la última vocal de las palabras componentes, exceptuándose la última palabra, la que siempre conserva su vocal: [...]. (Lin_CRI_08).

118. **Arbitrario**:

— Esta investigación tuvo en cuenta el axioma hallidyano de que el *análisis del discurso* sin gramática no pasa de ser un comentario de texto y en tal sentido se examinaron dos manuales escolares de historia a partir del funcionamiento de los adverbios en -mente [...]. (Lin_CHL_06).

— Como se ha tratado de explicar en este trabajo, la *disponibilidad léxica* presenta un orden cognitivo que se vislumbra en distintos niveles de organización, que va desde lo más general hasta lo más específico o viceversa. (Lin_MEX_01).

— Por último, se trata de un diccionario reducido o restringido, de palabras nuevas, pues estas no se han registrado con anterioridad en una *obra lexicográfica* y la mayoría de ellas han tenido que ser acuñadas para este propósito, puesto que

no existían con anterioridad en español las palabras que designaran los conceptos en cuestión. (Lin_CRI_05).

119. **Hecho de**:

— Es decir, en el *corpus de anécdotas* los niños emplearon cláusulas transitivas con 2 argumentos léxicos solamente en los casos en que debían codificar referentes discursivos nuevos. (Lin_MEX_04).

— Destacamos, en todo caso, que las marcas autonímicas combinadas están ausentes en el *discurso de divulgación científica* (N2), a excepción de un 1,2% de compuestos híbridos que presentan texto normal para el determinado en español y cursiva para el determinante en inglés. (Lin_CHL_02).

— Recordemos que la teoría de redes relacionales toma como objeto al *sistema lingüístico* de cada individuo real, con lo cual se admite que haya hablantes reales en cuyos sistemas sí se conectan los nodos para toy y para JUGUETE y al mismo tiempo [...]. (Lin_CHL_10).

120. **Tipo de**:

— No existe, en verdad, ningún tipo de *acto de habla* que deje de situarse en estas coordenadas de manifestación si bien, como ya dije, puede ocurrir que las modalidades elegidas correspondan a más de una coordenada temporal, espacial o social [...]. (Lin_PER_02).

— En definitiva, el *diccionario escolar* se convertirá en protagonista para el descubrimiento de la lengua, en vehículo para el hallazgo de sus tesoros y en promotor para el logro de su uso adecuado, expresivo [...]. (Lin_VEN_04).

— El criterio primordial para aceptar una palabra como *neologismo humorístico* es que la palabra resultante

constituya una nueva unidad semántica; esto quiere decir que, si ha sido creada por analogía a partir de algún modelo, su significado debe ser distinto [...]. (Lin_ARG_08).

Tabla 13. Etiquetas no términos en lingüística según validación por protocolo.

	Criterio	Tipo de candidato	Lingüística
	No - AdjGeneral	Simple	0
	No - NComún	Simple	1
	No - VSoporte	Simple	2
	No - Acus	Complejo	0
	No - Deadj	Complejo	1
Descartar	No - Ejemplo de	Complejo	6
	No - Nomin	Complejo	0
	No - Ocurre en	Complejo	2
	No - Partit	Complejo	2
	No - Prop	Complejo	14
	No - Unireferencial	Complejo	1

121. **No – Ncomún**:

— La *enseñanza* en el aspecto léxico debe convertir el input natural, desorganizado o espontáneo de la etapa adquisitiva pre-escolarizada en un input organizado, graduado y sistematizado para extender el conocimiento léxico. (Lin_PER_02.txt).

— [...] los actos cotidianos de significación, la gente representa la estructura social, afirmando sus propias posiciones y sus propios papeles, lo mismo que estableciendo y trasmitiendo sistemas comunes de valor y de *conocimiento*". (Lin_VEN_05.txt).

— Unos mecanismos de procesamiento lector eficientes proveen al *estudiante* de un acceso a los contenidos de los textos sin que ello importe un costo cognitivo que suponga una merma importante para la acción de los mecanismos de procesamiento [...]. (Lin_URY_02).

122. **No – Vsoporte**:

— [...] un ejemplo de cada país y de cada tipo de fuente que refleje el uso de los términos que se pretenden ilustrar; esto es posible gracias a una función que permite "*agregar* nuevo campo". (Lin_COL_09.txt).

— La búsqueda en la enciclopedia (en-searching, Borer 2013) permite *asignar* contenido léxico a las representaciones lingüísticas. (Lin_ESP_05.txt).

— Alexiadou 2010 también señala la necesidad de los índices (numéricos) para *diferenciar* entre formas fonológicamente idénticas, pero correspondientes a raíces distintas, como en Vbank ('banco': institución financiera) y Abank ('orilla': costa de un río). (Lin_ESP_05.txt).

123. **No - Ejemplo de**:

— Desde nuestro punto de vista, las diferencias encontradas en el procesamiento lector entre universitarios españoles y uruguayos están directamente relacionadas con las diferencias en consistencia de la *ortografía del español hablado* en Uruguay y el hablado en España. (Lin_ URY_02.txt).

— Como una última observación, todos los nombres para el *léxico del cuerpo humano* prefijan morfemas de posesión de manera opcional, de acuerdo con el contexto y con el poseedor referido. (Lin_COL_01).

— Muchos de los tangos que se escuchan aquí tienen letras escritas en lunfardo y a través de éstas, llegan a los hablantes muchas *palabras del lunfardo*, algunas creadas a través del vesre. (Lin_COL_05.txt).

124. **No - Ocurre en**:

— No se encontraron realizaciones consonánticas de /y/ en *posición codal*, las que sólo aparecen marginalmente en *posición frontal*. (Lin_CRI_08.txt).

— En otras palabras, el *"español americano"* existe en un nivel léxico, pero acompañado de glosas que buscan atenuar las diferencias regionales. (Lin_URY_01.txt).

— Una característica importante es que las consonantes que le siguen a una *vocal final* no contribuyen al peso de la sílaba, por lo que son consideradas extramétricas, tal es el caso de las reglas que no tienen en cuenta esas consonantes para la asignación [...]. (Lin_ARG_04).

125. **No – Partit**:

— Zanata (2010), citando a Bejoint (2002), expresa que el diccionario bilingüe en la sección para la comprensión debe ser abundante e incluir la mayor *cantidad de palabras* raras, nombres propios, los alomorfos, formas conjugadas, entre otros. (Lin_PRY_01).

— [...] sistematiza y pone a disposición de los miembros de la comunidad hablante el léxico de la lengua, o sea, el *conjunto de palabras*, entendidas como unidades de referencia a los hechos de experiencia que hemos aprendido a reconocer como unidades de sentido. (Lin_CHL_09).

— Utiliza el coeficiente que resulta de dividir el *número de palabras* diferentes de un corpus entre un segmento

determinado del texto, que debe tener 100 palabras gráficas. (Lin_CRI_09).

126. **No – Prop**:

— Se puede observar que, aunque las diferencias en el *promedio de palabra* por informante no son altas, sí aparecen diferencias asociadas a determinados centros de interés, las cuales están en relación con el área de procedencia. (Lin_COL_02.txt).

— La consideración de *datos gramaticales* sugiere una mayor diversificación al interior de ese complejo dialectal. (Lin_PER_05.txt).

— En el trabajo del año 1995 presenta una revisión de la evidencia empírica generada en las últimas 4 décadas, donde se apoya la idea de que el *tipo de lectura eficaz*, que caracteriza a los lectores avanzados, utiliza intensivamente el acceso a la información ortográfica almacenada en el léxico ortográfico. (Lin_URY_02.txt).

4.2.3 Hallazgos terminológicos en veterinaria

De los 149 candidatos analizados en veterinaria, se determinó que 121 corresponden a *términos* en la expresión temática del área, mientras que los 28 candidatos restantes corresponden a las unidades que aportan en la construcción del discurso temático más no representan un nodo cognitivo directamente. Por un lado, la Tabla 14 da cuenta los *términos* en veterinaria y la categoría según el protocolo. Además, en 127 a 137 se desarrollan ejemplos de cada categoría con sus correspondientes contextos de aparición. Por otro lado, en la Tabla 15 se relaciona la información sobre las unidades *no término* en el área seguida de las correspondientes ejemplificaciones en 138 a 144.

Tabla 14. Etiquetas términos en veterinaria según validación por protocolo.

	Criterio	Tipo de candidato	Veterinaria
	AdjDenominal	Simple	1
	AdjGeneral	Simple	6
	AdjRelacional	Simple	6
	AdvAdjetival-Mente	Simple	1
	NCompCulta	Simple	0
	NComún	Simple	0
	NMorfSubj	Simple	0
Validar	NProductMorf	Simple	7
	NTérmino	Simple	34
	VTérmino	Simple	3
	Arbitrario	Complejo	14
	Hecho de	Complejo	5
	Metf hecho de	Complejo	2
	Tipo de	Complejo	48
	Tipo de/Hecho de	Complejo	0

127. **AdjDenominal**:

— La contaminación de los recursos hídricos superficiales y subterráneos con residuos de origen doméstico, industrial y *agropecuario*, es uno de los principales problemas ambientales del sur de Chile, debido a que en esta área las aguas superficiales y subterráneas son una importante fuente de agua de bebida [...] (Vet_CHL_01.txt).

— [...] indican que han sido los ganaderos y agricultores tradicionales los que con su saber *ecológico* han mantenido este importante legado en la dehesa y en los espacios serranos de montaña. (Vet_ESP_03.txt).

— Una estrategia reproductiva que determine partos concentrados, permite ordenar el flujo *zootécnico* de la finca, mejorar la tasa de crecimiento de los reemplazos y concentrar todos los esfuerzos y recursos posibles en una época más favorable del año; ello posibilita elevar la eficiencia [...] (Vet_ECU_02.txt).

128. **AdjGeneral**:

— Este tiempo debe ser optimizado rigurosamente, debido a que el confinamiento en un camión sin movimiento puede ser incluso más *estresante* que el viaje, sobre todo en viajes de corta y mediana duración. (Vet_MEX_06).

— La mejora (reducción de insumos y/o incremento de producto) que debería experimentar cada productor *ineficiente* para convertirse en *eficiente* puede calcularse al comparar los valores observados y objetivos de éstos. (Vet_CHL_03.txt).

— En este proceso se calienta la leche a 30- 40° C. mayor sea el calentamiento más *seco* es el queso llamándose queso de pasta cocida [...] (Vet_BOL_08.txt).

129. **AdjRelacional**:

— Esto permite que la agricultura gane terreno sin que se deje de producir ganado *bovino*. (Vet_ARG_03.txt).

— En definitiva, entre los factores clave para lograr mayor productividad con el potencial *genético* de las vacas, se encuentran una mayor eficiencia de uso de las pasturas (más producción y cosecha de MS/ha) [...] (Vet_ECU_02.txt).

— Para IEP la tendencia fue a bajar con el tiempo lo cual implicó una mejora en el comportamiento *reproductivo* de las vacas durante el periodo evaluado. (Vet_VEN_05.txt).

130. **AdvAdjetival-Mente**:

— Así, a partir de los nueve meses de edad son necesarias las constantes reinfecciones para que el bovino desarrolle una inmunidad adquirida y permanezca una zona como estable *enzoóticamente* para la babesiosis (Jonsson et ál., 2008). (Vet_COL_10).

— Así una zona es considerada *epizootiológicamente* estable para la babesiosis bovina cuando el 75% o más de los bovinos entre 3 y 9 meses de edad son serorreactivos para Babesia bovis y/o Babesia bigemina [...]. (Vet_COL_10).

— En este sentido, Wallmo et al. (1977) señalan que la CCG estacional sería equivalente al número de animales necesario para consumir la mitad de la biomasa forrajera disponible y producida *estacionalmente*. (Vet_ESP_03).

131. **NProductMorf**:

— La intensa actividad extractiva de especies forestales de mayor talla y valor condujo a la degradación del bosque natural con la *arbustización* de los estratos bajos. (Vet_ARG_08).

— [...] producción agropecuaria al depender del clima de la zona donde se produce, encuentra una amenaza en el cambio climático y efecto invernadero, esperándose impactos negativos respecto a los cultivos, la *ganadería* y también en la salud humana. (Vet_BOL_04).

— El primero es la expansión de la *agriculturización*, tomando 14 millones de hectáreas ganaderas con destino de siembra, produciendo el incipiente corrimiento del stock ganadero [...] (Vet_ARG_01.txt).

132. **NTérmino**:

— Reducir la caída de productividad durante la época seca la tolerancia del *forraje* a la sequía y la suplementación animal durante la época seca parecen ser dos de las estrategias más convenientes para reducir la caída de productividad durante la época seca. (Vet_PER_04).

— Esto puede explicarse por el hecho de que quizás los animales, que han llegado a esta edad en el *hato*, se han mantenido en él precisamente por su habilidad productiva; de hecho se podría tratar de individuos superiores en los *hatos*. (Vet_CRI_10).

— La ovulación en la *vaca* se clasifica como espontánea y se presenta entre 24 a 32 horas posterior al inicio del *estro* (Senger, 2003). (Vet_CRI_07.txt).

133. **VTérmino**:

— *Cebar* novillos en trópico de altura puede ser una actividad productiva exitosa, pero es necesario identificar, tanto el nivel de inclusión, como la fuente energética adecuada para mejorar la relación energía:proteína [...]. (Vet_COL_05).

— Para los productores que conforman el conglomerado c2 el costo de alimentación representa un 44% de los costos de producción, lo sigue el costo de *fertilizar* pradera que tiene una importancia de un 17% en la composición de los costos de producción. (Vet_CHL_02).

— Las vacas en producción, son ordeñadas bajo el sistema con ternero hasta el final de lactancia, luego que el productor ha culminado de *ordeñar*, deja un cuarto de ubre o menos para la alimentación del ternero que representa 2 a 3 kg de leche/día practica que se realiza durante los primeros 3 a 5 [...]. (Vet_BOL_10).

134. **Arbitrario**:

— Por otra parte, probablemente el conocido antagonismo entre producción y reproducción podría estar afectando, en vista de la reducción en la producción de leche y posible mejora de la *condición corporal*, y a mejoras en el manejo del proceso de inseminación durante los años de estudio. (Vet_VEN_05).

— [...] (Quintana y Díaz, 2005), donde se encontró una mejor respuesta en *ganancia de peso diario* para novillos en pastoreo suplementados, cuando se tomó como base para establecer el suplemento un 1% del peso vivo, en contraste a emplear un 2%; atribuyendo estos resultados al efecto [...]. (Vet_CRI_06).

— [...] el agente afecta animales bien condicionados, con lo cual se favorece su crecimiento, y una vez empieza la multiplicación de la bacteria en el organismo del animal, se presenta *pérdida de peso* y la consecuente aparición de la diarrea (estadios avanzados de la enfermedad) [...]. (Vet_COL_08).

135. **Hecho de**:

— [...] los procesos más variables son aquellos englobados bajo el epígrafe otros, puesto que incluye actividades muy heterogéneas cuya presencia o ausencia determina grandes cambios en el consumo (cintas de alimentación, *bombeo de agua*, mezcladores o molinos...). (Vet_ESP_02).

— [...] incremento de valor de la carne vacuna permitió mejorar la competitividad de las otras carnes, tan importantes estratégicamente para la economía del país, como lo es el caso de la *carne porcina*, aviar y en menor medida ovina. (Vet_ARG_04).

— Pérdida de la capacidad de regeneración del monte en las pocas lagunas existentes, porque gran parte del ganado se concentra constantemente alrededor del mismo sitio, afectando la biodiversidad y *cobertura vegetal*. (Vet_BOL_01.txt).

136. **Metf hecho de**:

— El número de unidades muestreadas representó 5% del total de productores de cada zona; se seleccionaron aquellos que más *cabezas de ganado* tenían. (Vet_MEX_01.txt).

— El análisis de correlación indica que los establecimientos que incrementaron su porcentaje de destete fueron aquellos con mayor participación de ovinos en la *carga animal* en el inicio del período de análisis (Tabla 3). (Vet_ARG_09).

— Sin embargo, solo la mitad de ellos estarían dispuestos a reducir la *carga ganadera* si se vieran recompensados por ingresos provenientes de esas o de otras actividades complementarias. (Vet_ARG_10).

137. **Tipo de**:

— La *cepa de garrapatas*, utilizada en este ensayo fue proporcionada por el Laboratorio de Garrapatas del Servicio Nacional de Salud Animal (SENASA); originalmente proviene de México. (Vet_CRI_09).

— La industria de la *carne bovina* en Chile es un sector pequeño en producción, presenta una reducida *masa ganadera* y ha sido uno de los sectores más afectados por las importaciones en este último tiempo (Catrileo, 2005). (Vet_CHL_03.txt).

— Las sabanas de suelos ácidos son reconocidas como un ecosistema poco adecuado para el sostenimiento de una *ganadería extensiva*, requiriendo, como es el caso de las zonas afectadas por el SPB, más de 10 Ha por unidad animal. (Vet_VEN_03).

Tabla 15. Etiquetas no términos en veterinaria según validación por protocolo.

	Criterio	Tipo de candidato	Veterinaria
Descartar	No - AdjGeneral	Simple	2
	No - NComún	Simple	2
	No - VSoporte	Simple	5
	No - Acus	Complejo	0
	No - Deadj	Complejo	1
	No - Ejemplo de	Complejo	0
	No - Nomin	Complejo	4
	No - Ocurre en	Complejo	0
	No - Partit	Complejo	5
	No - Prop	Complejo	2
	No - Unireferencial	Complejo	1

138. **No – AdjGeneral**:

— En muchos tipos de vegetación, el nivel de calcio se mantiene relativamente *alto* durante todo el año, pero el nivel de fósforo decrece fácilmente y para la mitad del período de invierno, alrededor del 85% del contenido de fósforo, puede perderse. (Vet_VEN_03).

— Este número, que representa el 22% del total de estancias turísticas de toda la región Patagónica (Plan General Estratégico de Turismo Sustentable, 2009), es *bajo* teniendo en cuenta la extensión de la provincia, su potencial turístico (Scovenna, 2000) y el número de establecimientos ganaderos (947 establecimientos, Williams, 2009). (Vet_ARG_10.txt).

— En este mismo sentido, Mora (1997) determinó menor producción de leche en vacas seropositivas que en seronegativas al BLV en Costa Rica; sin embargo, la muestra analizada no fue lo suficientemente *grande* para sustentar estadísticamente el resultado, solamente reflejó una tendencia. (Vet_CRI_10.txt).

139. **No – Ncomún**:

— La duración del *día* oscila entre las 11h 50m, durante el "invierno ecológico" y las 12h 30m durante el "verano ecológico". (Vet_COL_03.txt).

— La producción ganadera representaba el único ingreso económico en su *hogar* para el 16% de los encuesta-dos; solamente el 10% de los encuestados afirmaron que la explotación ganadera representa más del 75% de sus ingresos [...] (Vet_ARG_10.txt).

— El recubrimiento de pisos con paja o aserrín es una práctica común, aunque controvertida, debido a que puede favorecer las caídas al hacer el piso menos adherente, dificulta la *limpieza* del remolque y por su posible implicación en la diseminación de agentes patógenos. (Vet_MEX_06.txt).

140. **No – Vsoporte**:

— La reconstrucción y la fusión se realizaron por el método «sandwich», que consiste en adherir una célula somática a una de las mitades enucleadas y posteriormente *adicionar* un segundo citoplasto enucleado. (Vet_PER_01.txt).

— Estas escalas pueden ser críticas para el bienestar animal porque incluyen múltiples operaciones de carga y descarga, manejo para *clasificar* por peso, sexo, raza o tipo comercial, mezcla social y dobles transportes en un mismo día, en una semana o un mes. (Vet_MEX_06.txt).

— De acuerdo con estos autores, si no hay impedimentos físicos o de comportamiento, los animales tienden a *optimizar* el uso de los recursos forrajeros sin intervención humana. (Vet_ARG_06.txt).

141. **No – Deadj**:

— *Cambios mensuales* en la densidad de tallos de ballico perenne (Ba) asociado con ovillo (Ov) y trébol blanco (Tr) en diferentes proporciones: a) 70 % de ballico perenne (20-70-10), b) 20 % de [...] (Vet_MEX_08.txt).

— La aplicación de una política de tipo de cambio alto, con la consecuente caída del *salario real* y la recomposición de la tasa de ganancia del sector privado, permitieron que la Argentina recuperase el superávit de la balanza comercial de la mano de las exportaciones agropecuarias. (Vet_ARG_03.txt).

— [...] ha venido implementando en programas de mejoramiento genético de bovinos de carne alrededor del mundo es el uso de mediciones de composición corporal por la técnica de ultrasonido de *tiempo real*, la cual ha mostrado ser un predictor correlacionado con las mediciones efectuadas directamente en la canal, con la producción de carne y con el grado de acabado. (Vet_COL_04.txt).

142. **No – Nomin**:

— Una vez conocido el *consumo eléctric*o de cada equipo, puede estimarse el consumo de cada uno de los tres principales procesos de producción en una explotación de vacuno de leche: ordeño (12,6 kWh), enfriamiento de la [...]. (Vet_ESP_02).

— Esto indica que el manejo, limpieza y desinfección de las jaulas de alumbramiento son muy importantes para reducir la coccidiosis en lechones, lo cual incluye lavado con *agua*

caliente a alta presión, flameado, aplicación de cal, uso de desinfectantes adecuados [...] (Vet_CRI_02.txt).

— El *conteo total* de los nematodos adultos en el tracto gastrointestinal de los animales parasitados fue, en promedio, 2175 ± 445. (Vet_MEX_03.txt).

143. **No – Partit**:

— Finalmente, de acuerdo con la condición corporal, la mayor *cantidad de animales* se encontró en el rango de score 3-4, y la seroprevalencia fue más alta en los animales clasificados entre 2-3 (tabla 3). (Vet_COL_08).

— En la primera, se comparó la kg 305-d según el estatus serológico por raza, así como por *número de lactancia* mediante la prueba de Kruskal-Wallis, utilizando el programa estadístico Infostat®. (Vet_CRI_10).

— Como se ve en el Cuadro 2, existe una gran dispersión de tamaños en la muestra, tanto en lo referido a los kilos de terneros producidos al año, como al *número de vacas* y también a la superficie y costos directos. (Vet_CHL_03).

144. **No – Prop**:

— No se registran cambios de magnitud en los valores promedio de la superficie propia ni de la superficie ganadera, pero se observa un aumento en el *uso de verdeos*, una disminución en la duración del servicio de monta y un incremento en el uso de la suplementación energético-proteica. (Vet_ARG_09.txt).

— En la tabla 3 se muestran los resultados serológicos hallados, de acuerdo con el *tipo de explotación* al cual pertenecen los hatos evaluados dentro del estudio. (Vet_COL_07.txt).

— La segunda etapa fue el diseño de muestreo y la aplicación de una entrevista estructurada a 225 consumidores de carne de bovino en seis *ciudades del estado de Puebla*. (Vet_MEX_04.txt).

4.2.4 Hallazgos terminológicos en computación

De los 53 candidatos analizados en esta área, se determinó que 47 corresponden a *términos* que dan cuenta directamente de un concepto al interior del área, mientras que los 6 candidatos restantes corresponden a unidades que contribuyen en la construcción del discurso temático más no representan un nodo cognitivo directamente. Por un lado, la Tabla 16 da cuenta los *términos* en computación y la categoría según el protocolo. Además, en 145 a 153 se desarrollan ejemplos de cada categoría con sus correspondientes contextos de aparición. Por otro lado, en la Tabla 17 se relaciona la información sobre las unidades *no término* en el área seguida de las correspondientes ejemplificaciones en 154 a 156.

Tabla 16. Etiquetas términos en computación según validación por protocolo.

	Criterio	Tipo de candidato	Computación
	AdjDenominal	Simple	1
	AdjGeneral	Simple	0
	AdjRelacional	Simple	1
	AdvAdjetival-Mente	Simple	0
	NCompCulta	Simple	0
	NComún	Simple	0
	NMorfSubj	Simple	0
Validar	NProductMorf	Simple	1
	NTérmino	Simple	16
	VTérmino	Simple	2
	Arbitrario	Complejo	5
	Hecho de	Complejo	4
	Metf hecho de	Complejo	6
	Tipo de	Complejo	18
	Tipo de/Hecho de	Complejo	0

145. **AdjDenominal**:

— Esto se debe a que el acceso a las memorias se hace *secuencial*, y mientras que por sus entradas de reloj no llegue un flanco de reloj, no habrá nada que mostrar. (Com_CUB_09.txt).

— Los objetos existen en un sistema *computacional* y están sujetos a restricciones y acciones del sistema, poseen un ciclo de vida: se crean y se destruyen, pueden ser asignados a otros objetos y tienen un comportamiento. (Com_VEN_02).

— La función utilizada es del tipo *sigmoidal* por simplicidad [19]. (Com_CHL_03.txt).

146. **AdjRelacional**:

— Una vez obtenido el resultado de la conversión lo envía por el puerto USB a una velocidad de 115200 baudios y luego cambia el multiplexor al siguiente canal *analógico* para convertir el canal vertical a la misma frecuencia. (Com_CUB_07).

— [...] se investigó e implementó cada una de las etapas involucradas en este detector, incluyendo la preparación de la base de datos de imágenes, el entrenamiento del sistema, utilizando algoritmos de aprendizaje *automático* y la implementación del algoritmo de detección. (Com_PAN_07).

— Simulación: Simulan hechos y/o procesos en un entorno *interactivo*, permitiendo al usuario modificar parámetros y ver cómo reacciona el sistema ante el cambio producido. (Com_MEX_09.txt).

147. **NProductMorf**:

— Sin embargo, los sistemas de excitación rápidos tienden a reducir el *amortiguamiento* de las oscilaciones (principalmente entre generadores cercanos), por lo que se introdujeron los llamados estabilizadores de potencia [...] (Com_CHL_02.txt).

— Con este fin se utilizó la tarjeta de desarrollo STM32FI-DISCO, aunque este *microcontrolador* no está diseñado para procesamiento digital de imágenes (PDI) como es el caso de los *microcontroladores* de Texas Instruments [...] (Com_ECU_10.txt).

— La *bisimulación* entre dos sistemas A y B, establece la relación entre cada estado de A y su correspondiente en el sistema B, y relaciona las transiciones observables de A y su correspondencia (Com_CHL_01).

148. **NTérmino**:

— Con este *dispositivo* se aprovecha la versatilidad de la realidad virtual no inmersiva en cuanto a la capacidad de aumentar la retroalimentacion de accion-observacion y así promover la activación [...] (Com_VEN_09.txt).

— En esta fase se evalúa la usabilidad del *software* mediante un cuestionario que se ha elaborado para el método y que se aplica a una muestra de usuarios del sistema que realizan su voto electrónico presencial. (Com_PER_01).

— La versión final del modelo computacional permite al *usuario* realizar una simulación y obtener resultados del estado de cargabilidad del transformador en análisis, esto permite obtener reportes y realizar comparaciones

de estados con el uso de otro tipo de transformador [...]. (Com_ECU_04).

149. **VTérmino**:

— Este principalmente consistía en *automatizar* más funciones de recolección de datos (tiempo y defectos) y generación de informes que las herramientas existentes, proveer un enfoque centrado en PSP que sea portable y entregue privacidad, [...]. (Com_CHL_06).

— El rendimiento total del proceso es el porcentaje de defectos encontrados y corregidos antes que los desarrolladores comiencen a *compilar* y probar sus programas. (Com_CHL_06.)

— Algunos programas permiten *modelar* desde cero y a manera de "dibujo" objetos que podemos encontrar en el mundo real y generarlos de manera digital según la creatividad del usuario. (Com_MEX_05).

150. **Arbitrario**:

— Se utiliza el sistema TreeBaG para interpretar un tipo particular de *generador de árbol* top-down, llamado td-generador, el cual está compuesto de una gramática de árbol regular y un transductor de árbol. (Com_CHL_09).

— En este trabajo se presenta la situación actual de este tipo de juegos, el modelamiento del presente juego, la *interfaz del juego* en base a su modelamiento previo y los resultados obtenidos mediante la evaluación del software en una muestra de alumnos de la misma ciudad. (Com_CHL_07).

— Por otra parte, el conjunto de árboles de derivación de una gramática de contexto libre pertenece a un *lenguaje de árbol regular*. (Com_CHL_09).

151. **Hecho de**:

— Después se creó una *base de datos* para registrar las características de la red (coordenadas, etiquetas para nodos y tramos, longitud, diámetro, rugosidad absoluta, coeficiente de pérdida menor global y la distribución interna de caudales). (Com_NIC_02).

— Como se puede observar, el evento únicamente trasmite el nombre del lugar mediante un parámetro de tipo *cadena de caracteres* denominado placeName, mientras que el servicio requiere de dos argumentos, la velocidad -un decimal- y el lugar que es denominado como place y que es también una cadena de caracteres. (Com_ESP_03).

— En sintonía con la metáfora del árbol, el *conjunto de datos* original es el nodo raíz, los subconjuntos son nodos, y los subconjuntos no particionados son las hojas, y estas hojas son definidas por la secuencia de las reglas de partición. (Com_CHL_05).

152. **Metf hecho de**:

— *Capa de Almacenamiento*: En la última capa se encuentra el servidor de base de datos en donde es almacenada la información generada por el sistema. (Com_CHL_06).

— La infraestructura de computación paralela híbrida entre procesadores y tarjetas gráficas es un *entorno de desarrollo* que permite el diseño, implementación y ejecución de soluciones paralelas. (Com_CUB_10).

— [...] libre (Chomsky) basado en cadenas de caracteres y las gramáticas de árbol regulares, sin embargo esto no debería sorprendernos, ya que todo lenguaje de árbol regular es una proyección de un *lenguaje de árbol* de derivación de una gramática de contexto libre basado en cadenas de caracteres. (Com_CHL_09).

153. **Tipo de**:

— El *código fuente* de estas funciones está disponible en tres lenguajes de programación: Java, C++ y MatLab, en el caso específico de este trabajo se utilizó el código en C++ (Li, Tang, Omidvar, Yang, & Qin, 2013). (Com_CUB_08.txt).

— El controlador de movimiento propuesto para el AEGIR implementa una arquitectura neuronal que está compuesta por una *red neuronal* de mapeo de dirección auto organizativa y una *red neuronal* para el comportamiento de evasión de obstáculos, ambas de inspiración neurobiológicas. (Com_ESP_06).

— Este trabajo propone el desarrollo de un prototipo de *robot móvil* para la Investigación de técnicas de Inteligencia Artificial que permita la navegación autónoma en entornos de prueba tipo indoor. (Com_ARG_01).

Tabla 17. Etiquetas no términos en computación según validación por protocolo.

	Criterio	Tipo de candidato	Computación
Descartar	No - AdjGeneral	Simple	1
	No - NComún	Simple	1
	No - VSoporte	Simple	0
	No - Acus	Complejo	0
	No - Deadj	Complejo	0
	No - Ejemplo de	Complejo	0
	No - Nomin	Complejo	0
	No - Ocurre en	Complejo	0
	No - Partit	Complejo	0
	No - Prop	Complejo	2
	No - Unireferencial	Complejo	1

154. **No – AdjGeneral**:

— Además la herramienta se encuentra *disponible* para utilizarlo en diversas plataformas o sistemas operativos, tales como: Unix, Linux, Windows XP y DOS. (Com_CHL_10).

— *Comprobable*, si, y solo si, existe algún proceso rentable finito con que una persona o la máquina pueda verificar que el producto reúne el requisito. (Com_CHL_08.txt).

— Esta propiedad es *relevante*, por que conviene que se ejecuten en la mayor variedad de arquitecturas, tienen una vida media mayor de diez años y la plataforma de ejecucion puede variar. (Com_ESP_08.txt).

155. **No – Ncomún**:

— Para realizar las mediciones, se dispuso de un sistema de coordenadas polares sobre un *afiche* alineado con la línea base de los sonares, como muestra la Figura 7. (Com_ARG_01.txt).

— La idea general surge dentro de un proyecto de investigación y desarrollo de software con la finalidad de determinar si es posible desarrollar una herramienta de software que permita al *alumno* o *egresado* realizar una autoevaluación antes de presentar el examen de Egreso [...] (Com_MEX_10.txt).

— Como el paciente debe tener la capacidad de coordinar sus movimientos con las acciones que ocurren en la escena y requiere de un cierto grado de concentracion, el *juego* sirve a su vez como un instrumento para tratar a pacientes con desordenes de atencion y autismo. (Com_VEN_09).

156. **No – Prop**:

— Para lograr una ventaja competitiva el *análisis de datos* se convierte en una estrategia clave para la mayoría de organizaciones. (Com_ECU_08).

— El usuario tendrá acceso a un menú principal que le proporcionará el *comienzo del juego*. (Com_CHL_07).

— En el *área de procesamiento* de señales, al referirnos a lo anterior, hacemos referencia a la extracción de parámetros, los cuales nos servirán para entrenar los HMM y lógicamente efectuar el reconocimiento. (Com_CHL_10).

4.2.5 Hallazgos terminológicos en filosofía

De los 45 candidatos analizados en filosofía, se determinó que las 45 unidades analizadas corresponden a *términos* que dan cuenta directamente de un concepto al interior del área. Por un lado, la Tabla 18 da cuenta los *términos* en filosofía y la categoría según el protocolo. Además, en 157 a 164 se desarrollan ejemplos de cada categoría con sus correspondientes contextos de aparición. Por otro lado, en la Tabla 19 se relaciona la información sobre las unidades *no término* en el área como insumo para la posterior discusión sobre la caracterización de las unidades clasificadas como término y no término (§ 5).

Tabla 18. Etiquetas términos en filosofía según validación por protocolo.

	Criterio	Tipo de candidato	Filosofía
	AdjDenominal	Simple	6
	AdjGeneral	Simple	0
	AdjRelacional	Simple	0
	AdvAdjetival-Mente	Simple	0
	NCompCulta	Simple	0
	NComún	Simple	10
	NMorfSubj	Simple	0
Validar	NProductMorf	Simple	1
	NTérmino	Simple	3
	VTérmino	Simple	2
	Arbitrario	Complejo	3
	Hecho de	Complejo	2
	Metf hecho de	Complejo	0
	Tipo de	Complejo	18
	Tipo de/Hecho de	Complejo	0

157. **AdjDenominal**:

— Una particular y común concepción de las vinculaciones entre lo *empírico* (observaciones y experimentación) y lo *teórico* (esquemas conceptuales o paradigmas) aparece en las propuestas de ambos. (Fil_ARG_04.txt).

— Sin embargo, Quine va a argumentar que la idea de que se puede trazar una línea de demarcación precisa entre estos dos tipos de enunciados es “un artículo metafísico de fe”, un dogma no *empírico* de los empiristas. (Fil_COL_01.txt).

— Este enfoque, de carácter empírico y a la vez *filosófico*, es considerado como la corriente europea en este campo (Escuela de Bruselas). (Fil_VEN_01.txt).

158. **NComún**:

— Por cierto, los textos cartesianos admitirían tres posibles respuestas a la pregunta por la naturaleza de la *idea*: la *idea* es un acto o evento mental, una entidad lógica, o bien, una disposición. (Fil_ARG_06).

— El primero de estos sentidos equivale a lo que Descartes llamaba *realidad* objetiva y es el que entra en juego en el concepto de adecuación. (Fil_ARG_06).

— La propuesta de la conducencia a *verdad* no es un infalibilismo. La conexión requerida con la *verdad* puede darse entendida como aquello que el sujeto obtiene cuando hace un máximo esfuerzo por alcanzar la *verdad* (justificación deontológica). (Fil_URY_01).

159. **NProductMorf**:

— [...] abordado la temática de la intuición en la filosofía de Spinoza desde la perspectiva de la problemática de lo que en la Ética aparece como ciencia intuitiva o tercer género de *conocimiento*. (Fil_ARG_06).

— Como es notorio, por el compromiso con el realismo ético cognitivista fuerte, Sócrates, al menos en Metafísica y *Epistemología* de la Ética, se encuentra en las antípodas del relativismo subjetivista y convencionalista del sofista Protágoras de Abdera tal como se describe en Teeteto [...]. (Fil_URY_10).

— El problema preliminar, que no es otro que el de la inducción, y que lo llevó a oponer el criterio de *falsabilidad* al de verificabilidad, lo conduciría necesariamente a considerar como no válido el concepto de ciencia manejado por Bloomfield y sus seguidores: la ciencia como pura y simple descripción que lleva al [...]. (Fil_VEN_05).

160. **NTérmino**:

— En este sentido podemos diferenciar entre la corriente *ecocentrista* (predominante en la biología de la conservación) y la perspectiva denominada *biocentrismo*, desde la cual se destaca el valor intrínseco de cada organismo (Lindenmayer & Burgman 2005). El *biocentrismo* no constituye una perspectiva mayormente considerada en los libros de texto. (Fil_ARG_03.txt).

— No se debe desconocer la diferencia cualitativa existente entre el enfoque gnoseológico (filosófico) del conocimiento y el enfoque epistemológico del *positivismo*, del empirismo lógico y de las corrientes y escuelas agnósticas, pesimistas e idealistas que desde un pensamiento "puramente científico" han combatido las mejores tradiciones [...]. (Fil_CUB_08).

— [...] sus actos y con ello independencia de la figura de Dios, ya que su finalidad no es un ideal en Dios, sino su visión del mundo se transforma lentamente en un *relativismo*. (Fil_GTM_05).

161. **VTérmino**:

— [...] pensar el mundo histórico real, el mundo en curso en la historia que los seres humanos padecen y hacen, para intentar así ser pensamiento de ese mundo, esto es, esfuerzo por *aprehender* su tiempo reflexivamente". (Fil_GTM_08).

— Una primera observación: podemos *conjeturar* que, para Sócrates, el Dios es sabio porque es omnisciente; sobre Sócrates y la omnisciencia del Dios ver (Jenofonte, Memorabilia, libro I: 46-51). (Fil_URY_10).

— Así propone reparar en las experiencias de sufrimiento moral cotidiano, para *develar* una moralidad escondida en los sentimientos e intuiciones de injusticia de aquellos

individuos o grupos que sufren formas de menosprecio y exclusión, aunque sean expresadas de forma difusa y prediscursiva. (Fil_URY_04).

162. **Arbitrario**:

— Por lo demás, esta interpretación podría ayudarnos a arrojar más luz sobre la cuestión de cuál es el *género de conocimiento* que opera en la Etica, aunque, claro está, es una cuestión que deberemos dejar para otra ocasión. (Fil_ARG_06.txt).

— [...] medidas de probabilidad en el intervalo cerrado 0< q< 1, donde 0 es la falsedad, 1 la verdad y q un número racional intermedio que nos da la probabilidad de un *estado de cosas* (no es preciso recurrir a números irracionales). (Fil_PER_01).

— En la sección anterior mostramos la relevancia de las categorías presentes en el debate sobre *epistemología del testimonio* en relación con la cuestión del testimonio moral, a partir de un conjunto de consideraciones generales. (Fil_ARG_09.txt).

163. **Hecho de**:

— Para Nietzsche, el *ser humano* a finales del siglo decimonono era un ser moral, un ser de valores y un ser en una sociedad constituida principalmente por el Estado. (Fil_GTM_05).

— Roig nos señala que en la *noción martiana* de hombre natural encontramos la punta de lanza para una crítica política a las subjetividades. (Fil_ARG_01).

— Siguiendo esta línea de reflexión, creo que es posible, haciendo un uso un poco salvaje de la semántica de los tiempos históricos, tomar a "*normalidad filosófica*" como un concepto social, en el sentido en que lo entiende Reinhart Koselleck. (Fil_ARG_02).

164. **Tipo de**:

— En cambio, el testimonio reviste peligro para las tesis de *autonomía epistémica* y práctica si asumimos que proporciona razones de segunda persona, basadas en relaciones de autoridad y responsabilidad entre un hablante y su audiencia. (Fil_ARG_09).

— Este giro se relaciona con los aportes contemporáneos de la *filosofía del lenguaje* que ha cuestionado la razón como fundamento, el sujeto como fuente de conocimiento válido y la historia como linealidad ascendente. (Fil_ECU_06).

— Por otro lado, gran parte del interés que suscita la cuestión del *testimonio moral* Comprender, raudamente, significa para el agente adquirir la habilidad práctica de aplicar ciertos preceptos morales más o menos generales a casos nuevos, semejantes pero no idénticos [...]. (Fil_ARG_09).

Tabla 19. Etiquetas no términos en filosofía según validación por protocolo.

	Criterio	Tipo de candidato	Filosofía
Descartar	No - AdjGeneral	Simple	0
	No - NComún	Simple	0
	No - VSoporte	Simple	0
	No - Acus	Complejo	0
	No - Deadj	Complejo	0
	No - Ejemplo de	Complejo	0
	No - Nomin	Complejo	0
	No - Ocurre en	Complejo	0
	No - Partit	Complejo	0
	No - Prop	Complejo	0
	No - Unireferencial	Complejo	0

4.3 Recapitulación

La consolidación de un corpus textual especializado con propósitos de análisis de la variación terminológica horizontal se constituye, por sí mismo, en un aporte significativo de cara a los estudios en terminología. De este modo, partiendo de una compilación de 538 artículos de investigación, un total aproximado de 3.057.973 tokens y una población de 37.055 candidatos a término, los resultados aquí descritos se conciben, por un lado, como el puente entre las bases teóricas consolidadas a lo largo de la historia de la terminología (§§ 1-2) y la propuesta metodológica de esta tesis doctoral (§ 3) y, por otro lado, como la antesala para la discusión en torno al término como objeto de estudio de nuestra disciplina. Una vez implementado el protocolo de validación de candidatos a término (Quiroz & Burgos, en prensa) en la muestra de datos (586 candidatos) se confirmaron como términos el 79,18 % de las unidades analizadas como se ilustra en la Figura 17.

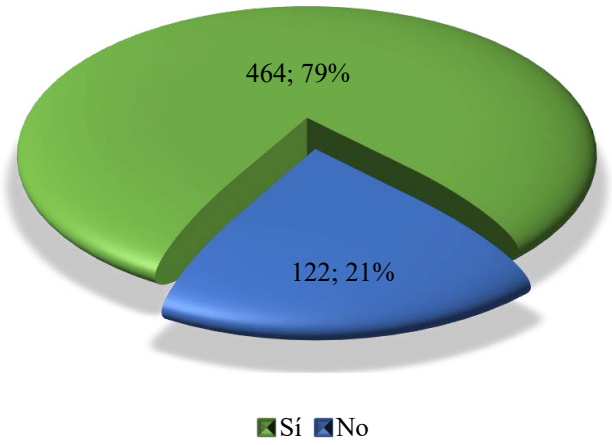


Figura 17. Cantidad de términos versus no términos.

En un primer momento, a partir de la extracción realizada con *Termostat*, se logró recuperar no solo los candidatos a término a validar sino también la información temática en la consolidación

del corpus VaTeHo en torno a las áreas y conceptos clave empleados en su compilación. Así, en el apartado de *Nuage* se puede constatar este aspecto que a su vez será discutido más adelante (§ 5.1) toda vez que permite plantear la necesidad de reevaluar este tipo de ayudas visuales de carácter terminológico. A su vez, con base en la información estadística de cada subcorpus se presentó un primer acercamiento a los términos utilizados en cada una de las áreas analizadas en contraste con los candidatos no término, pero que soportan la construcción temática debido a su carácter de especialidad. Por último, se logró identificar las inconsistencias en la extracción lo que sin duda servirá como insumo en una posterior reseña de la herramienta. En un segundo momento, se describió el recorrido realizado desde la población de datos hasta la muestra utilizada para la aplicación del protocolo de validación. En este sentido, las decisiones estuvieron marcadas, en esencia, por un criterio de tiempo y recurso humano para llevar a buen puerto la validación y posterior discusión de los resultados obtenidos. En tercer momento, se listan los términos y los criterios con los cuales fueron caracterizados como tal junto con los candidatos no término según las particularidades conforme la aplicación del protocolo de validación.

En este sentido, los resultados aquí presentados nos permitieron, a la luz de las hipótesis y objetivos propuestos para esta tesis doctoral (§ 0), reafirmar nuestra concepción del término como una ampliación de la definición actual de este. En el siguiente segmento, se presentarán las diferentes discusiones y reflexiones fruto de la integración del componente teórico y metodológico con miras a la consolidación de las conclusiones de esta tesis doctoral.

5. El término en español

"Si he logrado ver más lejos ha sido porque
me he subido sobre hombros de gigantes"

(Sir Isaac Newton).

El trasegar por los demás capítulos en esta tesis doctoral nos trae hasta aquí. Un recorrido por la historia de la terminología basado en las diferentes posturas y aportaciones de las escuelas de pensamiento en torno a un objeto común de estudio: el *término*. Una *mirada en conjunto* que nos permite observar el panorama de nuestra disciplina con el reto de una continuación, una actualización de nuestros saberes compartidos de cara al futuro de los estudios terminológicos. Un marco teórico que da cuenta de la amplitud de concepciones y reflexiones sobre el tema, un marco metodológico pensado e implementado para dar respuesta a los interrogantes aún latentes en nuestro ámbito, unos resultados que permiten ensanchar el horizonte de trabajo gracias a la ampliación del concepto de *término* fruto de las discusiones que configuran el cuerpo del presente trabajo de investigación doctoral. Una línea entre el pasado y el presente que permite observar con detenimiento los pormenores de la terminología. No partimos de la nada, construimos con aquellas voces de multitud de investigadoras e investigadores alrededor del mundo que dedicaron o aún siguen dedicando su tiempo alrededor de los desafíos en el entendimiento de las formas de materialización de los saberes humanos a través del lenguaje. Un reconocimiento, un agradecimiento, una responsabilidad de continuar con el legado por ustedes construido.

En el presente capítulo se compilan las reflexiones y discusiones a partir de nuestras hipótesis de trabajo, por un lado, las diferencias, a nivel lingüístico, de las características de los términos utilizados por las diferentes áreas del saber según las formas de expresar su conocimiento y, por otro lado, la concepción del término más allá y no solo a partir de unidades nominales sino de cualquier unidad léxica presente en cada discurso de especialidad. Asimismo, se dialoga con el objetivo principal de nuestro trabajo en la caracterización de qué entendemos por término y su variación según las diferentes áreas del saber, compiladas en el corpus textual especializado para la Variación Terminológica Horizontal (VaTeHo) para, posteriormente, indagar sobre las miradas paradigmáticas en los estudios terminológicos modernos basados en los puntos comunes y distantes de acuerdo con los hallazgos de nuestra investigación (§ 4 Resultados). Por último, los logros alcanzados y los retos pendientes que nacen gracias a las conclusiones en esta tesis doctoral.

5.1 Ampliación de nuestra unidad de estudio

En la actualidad, los estudios en terminología se han decantado por dos posturas paradigmáticas que han determinado las líneas de trabajo desde finales de la década de 1990. En este sentido, la mayoría de las investigaciones, por no decir todas, se fundamentan en las bases conceptuales a partir de la TCT de Cabré (2000) y en menor medida en la TBM de Faber (2012). Al respecto, la primera, de orientación lingüística, propende por una concepción del término desde la denominada teoría de las puertas (Cabré, 2000) según la cual estos pueden ser observados a partir de elementos lingüísticos, cognitivos o comunicativos (p. 10). Además, en su definición de término, Cabré los cataloga como "una unidad léxica que ocupa un nodo cognitivo dentro de un dominio especializado" (1999, p. 48). La segunda, de orientación cognitiva, intenta representar cómo funciona el conocimiento de especialidad en la mente del experto, a partir de la teoría de los marcos semánticos de Fillmore (1982). Ahora, con relación a su definición de término, Faber indica que estos son unidades lingüísticas que

expresan su significado conceptual al interior del marco de los textos de especialidad (Faber, 2012, p. 11). De este modo, inicialmente, no se presenta una distinción relevante en la concepción del término en ambos postulados salvo la *puerta* de acceso a su estudio, situación que se aclarara teniendo en cuenta el trabajo más reciente sobre la TBM, donde se presenta una definición más explícita que, en principio, logra subsanar el faltante en cuanto a una concepción del término que reflejase los principios y alcances epistémicos de este enfoque terminológico. En consecuencia, Faber concibe el término como:

> the natural language designation of an object, process, action, or attribute within a specialized knowledge domain and its conceptualization. Used for purposes of specialist communication and knowledge acquisition, a term is a relational unit that triggers a rich knowledge representation in which visual, auditory, and linguistic data coalesce and converge in a frame, which takes the form of a semantic network of related concepts (2022, p. 355).

En relación con la teoría de las puertas (Cabré, 2000), esta es concebida, por la autora, como una teoría de los términos que logra integrar el carácter poliédrico de estos, teniendo como punto de partida una base de análisis claramente de orientación lingüística. Así, aunque la teoría plantea, por lo menos, tres puertas de entrada para el análisis de las unidades terminológicas, la TCT se decanta por una construcción a partir de elementos lingüísticos, es decir, a pesar de plantear el término como una conjunción de aspectos semióticos y lingüísticos, cognitivos y comunicativos, que a su vez ratifican la terminología como una disciplina de conocimientos necesariamente interdisciplinarios, la elección de análisis corresponde al trayecto lingüístico. Además, otros elementos clave en la comprensión de los postulados en este acercamiento teórico se remiten, por un lado, a la caracterización de los términos debido a su uso en contextos de especialidad, en contraste con las palabras del léxico general, siendo esto un importante aporte de esta teoría al campo de trabajo y, por otro lado, la concepción de las unidades terminológicas como vehículos del conocimiento de especialidad y, por ende, portadores de

un significado especializado. En este sentido, teniendo en cuenta que la clave de este acercamiento está en comprender que, desde la puerta lingüística, el análisis de las unidades terminológicas debe permitir, por lo menos, los otros dos acercamientos (cognitivo y comunicativo), se hace necesario ahondar desde la TCT en esos otros aspectos para lograr un desarrollo completo de lo presentado en esta teoría.

En esta misma línea, en las diferentes investigaciones en terminología, salvo unas cuantas (§ 2), el foco de atención han sido las unidades nominales en campos de conocimiento prototípicos como la medicina y la economía y, en consecuencia, una mirada parcializada del fenómeno. En este sentido, reconocemos el aporte de las diferentes escuelas, a nivel teórico y metodológico, evidenciado en los hallazgos en sus trabajos y la amplia gama de publicaciones en las diferentes revistas y portales académicos. Sin embargo, no concordamos con esta mirada restrictiva del fenómeno. Si partimos de la premisa cabretiana sobre el lenguaje de especialidad como un subconjunto de la lengua general, no encontramos justificación para dejar a un lado a las demás unidades léxicas más allá del sustantivo y las estructuras nominales. Así, se hace necesario una ampliación en la concepción de nuestro objeto de estudio común que permita observar las dinámicas del discurso de especialidad desde la óptica de los adjetivos, adverbios, verbos y demás unidades complejas que han sido tratadas tangencialmente o excluidas en los análisis. Además, reconocemos el carácter prototípico y estadístico del sustantivo, lo que no va en contravía del carácter terminológico de las demás unidades como se ha identificado en la presente tesis doctoral.

En cuanto a la noción de variación terminológica horizontal, la literatura en el área ofrece una mirada mínima del fenómeno, es decir, se ha observado a partir de la noción temática en las diferentes áreas del saber y la relación entre estas introduciendo un carácter, podemos llamarlo, limítrofe entre cada una de ellas. En este sentido, el trabajo de Hoffmann (1998) se presenta como el punto de referencia en esta concepción que, inclusive, hoy en día sigue siendo el referente para esta caracterización. Sin embargo, con los enfoques epistemológicos

actuales, en los cuales las miradas positivistas se vienen cuestionando y replanteando, es necesario introducir en esta discusión la hibridación cada vez mayor de las áreas lo que, por sí mismo, implica un reto mayúsculo en la identificación de los términos propios de cada campo del saber. Ahora, fruto de los hallazgos en nuestra investigación concebimos la variación terminológica horizontal en el mismo sentido temático, pero además la observamos desde las características prototípicas de cada área analizada en el corpus. Es decir, reconocemos en esta variación no solo la delimitación en función de la epistemología de cada área, sino que también identificamos que cada una utiliza mecanismos lingüísticos propios en la materialización de sus saberes a través de los términos y sus estructuras. En este sentido, no es posible seguir analizando terminológicamente las áreas como un todo igual, más bien entender las dinámicas de cada campo de conocimiento y ahondar en sus particularidades. Allí, sin duda, habrá puntos comunes, pero también características propias que nos permiten hablar de una *variación terminológica horizontal tipológica*.

Nuestra visión de término está orientada a una ampliación del espectro de análisis de la terminología. Nos ocupamos de lograr integrar, a razón de nuestros hallazgos, otros puntos a observar en el panorama ya establecido por nuestra disciplina. Partimos de la compilación de un corpus en cinco áreas del conocimiento, cada una con un tema específico (§ 3.2.1.1), siendo esto un aporte importante en la comprobación de nuestras hipótesis y, en general, a la terminología. Un corpus especializado de más de tres millones de palabras y 37 mil candidatos a término extraídos automáticamente con *TermoStat*. Un análisis entre lo automático y lo manual que ratifica la necesaria simbiosis entre el humano y la máquina, nada está al alcance de las herramientas y programas computacionales que no haya sido primero observado, pensado, diseñado e implementado por la mente humana, quizás en menor proporción, pero con una mirada transdisciplinar en el reconocimiento de los saberes humanos. Un protocolo de validación (Quiroz & Burgos, en prensa) que consolidó nuestra concepción del fenómeno y, quizás, se presenta como la antesala para nuevas investigaciones en terminología. Criterios lingüísticos que parten

de reflexiones sintácticas, semánticas y pragmáticas que a su vez permiten caracterizar aquellas unidades que en medio del discurso especializado no son término, más cumplen otros roles significativos en la construcción temática de cada campo de conocimiento. En suma, como ya lo hemos mencionado, concebimos el término como *todo tipo de unidad léxica, simple o compleja que, en función de cada discurso de especialidad, contiene y materializa los saberes propios de un campo del conocimiento a través de los diferentes mecanismos inherentes al lenguaje natural, a partir de elementos lingüísticos y cognitivos que, a su vez, responden a una función social mediada por aspectos de orden cultural.*

En primer lugar, hablamos de *todo tipo de unidad léxica simple o compleja* toda vez que reconocemos el término, en primera instancia, como una etiqueta lo que, necesariamente, implica una delimitación más allá de la cantidad de componentes. Sin embargo, teniendo en cuenta las distinciones en el área conservamos la diferenciación operativa entre *simples* y *complejas* para su delimitación por número de componentes. A su vez, trascendemos la frontera de lo nominal en la ratificación del contenido terminológico presente en todas y cada una de las unidades léxicas analizadas al interior de nuestro corpus y las diferentes áreas del saber que lo componen. En segundo lugar, concebimos que el término *contiene y materializa los saberes propios de un campo del conocimiento* distanciándonos de la concepción vehiculante bajo la cual este se ha observado hasta el momento, toda vez que en esta óptica adquiere una labor operativa más su esencia se ve difuminada por la noción de tránsito, es decir, al hablar de este como vehículo del conocimiento de especialidad implícitamente se le está asignado un principio y un final en el trayecto de la expresión del saber. Así, afirmamos que el término contiene el saber de cada área lo que le faculta la materialización de dichos conocimientos, es decir, el término es el concepto al que hace referencia al interior de cada área del saber.

En tercer lugar, concordamos con la noción de los lenguajes de especialidad como subconjuntos de la lengua general (Cabré, 1998, p. 7). De este modo, el término cumple su función principal *a través de*

los diferentes mecanismos inherentes al lenguaje natural. En esencia, cada área del saber elige, a partir de sus fundamentos epistemológicos, las formas en las cuales expresar su conocimiento a través de las elecciones basadas en las posibilidades y alternativas propias de la lengua general y, por ende, contrario a las críticas en relación con la concepción del lenguaje de especialidad como artificial, los hallazgos de la presente tesis doctoral permiten contribuir en la desmitificación de esta postura al corroborar, gracias a los ejemplos presentados, que estos hacen uso de la lengua general de cara a la materialización de sus saberes. Asimismo, reconocemos que el abordaje en la caracterización y entendimiento del término puede darse tanto desde el plano *lingüístico* como desde el *cognitivo*, lo que en esencia se corresponde con la manera en que se ha asumido el estudio del lenguaje. Finalmente, en cuarto lugar, reafirmamos la *función social* y el *carácter cultural* de los términos, las dinámicas que esto representa y el impacto en los distintos contextos en los cuales están inmersos.

En consonancia con lo hasta ahora descrito en este trabajo doctoral y, por decisiones de orden metodológico, en relación con el alcance de la investigación y el tiempo disponible en el cierre de esta, hemos caracterizado, desde el plano lingüístico, los términos a la luz de su entendimiento teniendo como base nuestra concepción de la unidad de análisis plasmada en la propuesta de definición arriba detallada. Es menester aclarar que aunque el alcance de las discusiones de nuestro trabajo se basan en un ámbito descriptivo, toda vez que el fenómeno ha sido analizado haciendo uso de elementos desde la estadística descriptiva más no inferencial, esto no invalida ni contradice los hallazgos ni las conclusiones de nuestro trabajo, puesto que a partir de la observación de la información obtenida se logró realizar un análisis cualitativo y, por ende, una explicación a la luz de las hipótesis y objetivos de trabajo. La validez es por lo tanto endógena y corresponde a los datos compilados en función de esta tesis doctoral, elemento que a su vez se configura como el primer desafío a futuro en cuanto a la necesidad de ampliar los datos de análisis de manera que se continue con la implementación del protocolo de validación final aquí utilizado. En este sentido, un total de 586 unidades, entre simples y complejas,

sirvieron de insumo para la aplicación de los criterios establecidos en el protocolo de validación para la determinación de los términos.

Un primer elemento a destacar está relacionado con las frecuencias en ambos tipos de unidades, lo que permitiría establecer que las dinámicas en el uso ya sea de unidades simples o complejas se corresponde entre los núcleos del saber, esto es, en las áreas de la salud y las ciencias de la computación sería más frecuente el uso de unidades complejas sobre las simples en la expresión de su saber, mientras que las humanidades (en las cuales enmarcamos la lingüística y la filosofía) la frecuencia sería inversa. Esta información es relevante en cuanto permitiría, de cara a una automatización en el proceso del reconocimiento y extracción de términos, repensar y replantear las instrucciones dadas a este tipo de herramientas, introduciendo la noción de agrupación por núcleos del saber y, en consecuencia, seguir caracterizando las particularidades de cada campo del conocimiento en sintonía con nuestra propuesta de *variación terminológica horizontal tipológica*. En la Tabla 20 se ilustra esta situación según los datos en relación con las unidades recuperadas automáticamente con *TermoStat* y las muestrales validadas según el protocolo.

Tabla 20. Cantidad de candidatos a término simples versus complejos.

Tipo de candidato		**Medicina**	**Lingüística**	**Veterinaria**	**Computación**	**Filosofía**
Candidatos recuperados por TermoStat	Simple	4256	8599	3651	1053	1654
	Compleja	5647	5508	3661	1812	1214
Candidatos muestra de análisis	Simple	66	52	51	21	22
	Compleja	76	50	70	33	23

Un segundo elemento está relacionado con el paradigma actual en la concepción del término y su directa influencia en el diseño de diversas formas de presentar los datos terminológicos gracias a las ventajas de las nuevas tecnologías de la información. En este sentido, en una actualidad cada vez más multimodal, los diferentes programas para el procesamiento de terminología están integrando herramientas para la creación de *nubes de palabras* (Reyes, 2020, p. 56) que permiten identificar los conceptos de mayor frecuencia en los distintos temas objeto de análisis, a modo de ejemplo *TermoStat* con su apartado de *Nuage* o *Voyant Tools*[62] con *Cirrus*. En este punto, como ya hemos mencionado, concordamos con la predominancia de las unidades nominales como aquellas de mayor frecuencia y, por ende, las prototípicas en la terminología. Sin embargo, consideramos que la inclusión de las otras unidades léxicas terminológicas, tanto simples como complejas, en la creación de estos recursos permitiría una mirada, sino completa, más amplia del contenido de cada campo del conocimiento analizado. Para ilustrar este aspecto, se presentan a continuación tres *nubes* por cada uno de nuestros subcorpus que permiten reflexionar acerca de la necesidad de revisar y refinar estas otras formas de presentación de la información, a partir de las configuraciones e instrucciones dadas a las herramientas automáticas empleadas para tal propósito como las arriba mencionadas.

Inicialmente, se presentan las *nubes* obtenidas a partir de la herramienta *Voyant Tools* donde se emplearon los datos con el total de los artículos compilados en cada uno de los cinco subcorpus luego de la limpieza inicial (§ 3.2.1.5). Así, en las Figuras 18 a 22 se recupera la información según las configuraciones actuales mediadas por el paradigma, esto es, unidades simples de tipo nominal. En estas *nubes* se logra observar el *ruido* generado por la herramienta aportando a la versión final palabras de uso general no terminológico debido a la frecuencia de aparición en los textos analizados.

[62] Para más información sobre *Voyant Tools*, consúltese: https://voyant-tools.org/docs/#!/guide/about.



Figura 18. Nube palabras medicina - Voyant Tools (Sinclair & Rockwell, 2022).



Figura 19.Nube palabras lingüística - Voyant Tools (Sinclair & Rockwell, 2022).



Figura 20. Nube palabras veterinaria - Voyant Tools (Sinclair & Rockwell, 2022).



Figura 21. Nube palabras computación - Voyant Tools (Sinclair & Rockwell, 2022).



Figura 22. Nube palabras filosofía - Voyant Tools (Sinclair & Rockwell, 2022).

En este sentido, se hace necesario la utilización de una *lista de exclusión* que permita dejar por fuera de la presentación de los datos aquellas unidades que cumplen otras funciones al interior de los textos como los conectores o las conjunciones y otras de uso general. Las Figuras 23 a 27, elaboradas a partir de los mismos datos, dan cuenta de la incorporación de *listas de exclusión* para cada subcorpus lo que permite comparar ambas extracciones y obtener una *nube* más precisa, aunque no completa.



Figura 23. Nube palabras medicina con lista de exclusión - Voyant Tools (Sinclair & Rockwell, 2022).



Figura 24. Nube palabras lingüística con lista de exclusión - Voyant Tools (Sinclair & Rockwell, 2022).



Figura 25. Nube palabras veterinaria con lista de exclusión - Voyant Tools (Sinclair & Rockwell, 2022).



Figura 26. Nube palabras computación con lista de exclusión - Voyant Tools (Sinclair & Rockwell, 2022).



Figura 27. Nube palabras filosofía con lista de exclusión - Voyant Tools (Sinclair & Rockwell, 2022).

Las Figuras 28 a 32 dan cuenta de la propuesta de *nube* basada en nuestra mirada amplia del término. Para ello, se tomaron los datos a partir de los términos validados en la muestra de trabajo, seleccionando por cada una de las unidades los cinco con mayor frecuencia. El objetivo al presentar la comparación entre las tres *nubes* por cada subcorpus, subyace en la posibilidad de ilustrar

cómo en la ampliación del entendimiento del objeto de estudio, se lograría una caracterización más precisa del saber analizado, toda vez que en este tipo de esquemas los conceptos clave son la base para la definición conceptual de los temas trabajados y, por extensión, la validación del aporte de todas y cada una de las distintas unidades terminológicas en las cuales se contiene el discurso de especialidad. Por último, se empleó la herramienta *Word Cloud Generator*[63] para la elaboración de nuestra propuesta. Es importante mencionar que la configuración de ambas herramientas se basa en el algoritmo diseñado por Jason Davies.



Figura 28. Nube palabras medicina – adaptación propia en Word Cloud Generator (Davies, 2022).

[63] Para más información sobre *Word Cloud Generator* autoría de Jason Davies, consúltese: https://www.jasondavies.com/wordcloud/about/.



Figura 29. Nube palabras lingüística – adaptación propia en Word Cloud Generator (Davies, 2022).



Figura 30. Nube palabras veterinaria – adaptación propia en Word Cloud Generator (Davies, 2022).

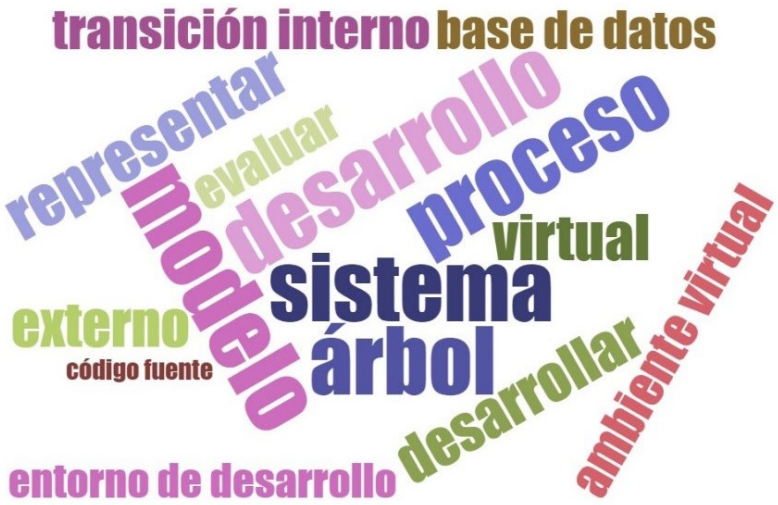


Figura 31. Nube palabras computación – adaptación propia en Word Cloud Generator (Davies, 2022).

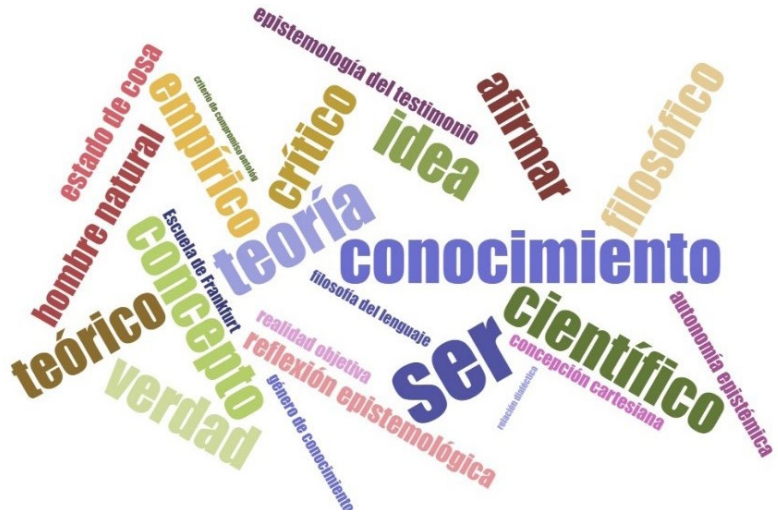


Figura 32. Nube palabras filosofía – adaptación propia en Word Cloud Generator (Davies, 2022).

Un tercer elemento tiene que ver con lo que hemos denominado *opacidad terminológica*, es decir, el reconocimiento del valor terminológico o no de una unidad al interior de los distintos contextos de especialidad. De esta manera, a la luz de los diferentes términos

analizados, surge el interrogante sobre qué características permiten que un usuario, ya sea experto o lego, identifique una unidad como término o la descarte como tal. Al respecto, en unidades como *biopsia* o *carcinoma* (Med), *adjetivar* o *sustantivar* (Lin) el carácter terminológico es identificado incluso sin ser necesario recuperar el concepto al cual hace referencia, mientras que en unidades como *cosa* u *objeto* (Fil) su valor terminológico requiere un análisis más en detalle, además de demandar por parte del usuario un conocimiento contextual y/o referencial del área analizada. En este sentido, es necesario ahondar en este rasgo de los términos de manera que sea posible la identificación, desde los diversos enfoques de trabajo, de los diferentes aspectos con el objetivo de su caracterización.

Un cuarto elemento está relacionado con la caracterización, desde lo lingüístico, de los términos simples y complejos. En los siguientes segmentos se presentan algunas particularidades de las distintas unidades analizadas en la muestra de trabajo y la consolidación de nuestra propuesta de análisis a partir de la ampliación del concepto de término y la introducción de lo que hemos denominado *variación terminológica horizontal tipológica*, además de otros hallazgos en función de la aplicación del protocolo de validación (Quiroz & Burgos, en prensa). Como se indicó, para el análisis se tomaron 538 candidatos a término de los cuales 258 son simples y 290 complejos distribuidos según la frecuencia de aparición de cada unidad al interior del corpus (véase § 4.1.7).

5.1.1 Los términos simples en español

La determinación de los términos simples, teniendo como base los quince criterios recuperados a partir de la implementación del protocolo de validación, permitió consolidar el carácter terminológico autónomo no solo del *sustantivo* sino también del *adjetivo*, el *adverbio* y el *verbo* lo que, a su vez, complementa y amplifica el panorama y espectro de análisis para este tipo de unidades. En esencia, el contenido de especialidad es expresado en y por cada una de estas unidades en función de los diferentes campos del saber analizados en los cinco

subcorpus. Hasta ahora, se ha ejemplificado cada unidad por cada área, pero al momento de integrar los resultados obtenidos se identifica la tendencia en la expresión del saber, es decir, el cumplimiento o no de los criterios establecidos en el protocolo de validación los cuales, como ya se estableció, dan respuesta a las particularidades hasta ahora aportadas desde las diferentes tradiciones en la terminología y complementadas con nuestra propuesta de término. De esta manera, en la Figura 33 se integran los resultados de la validación de los términos simples.

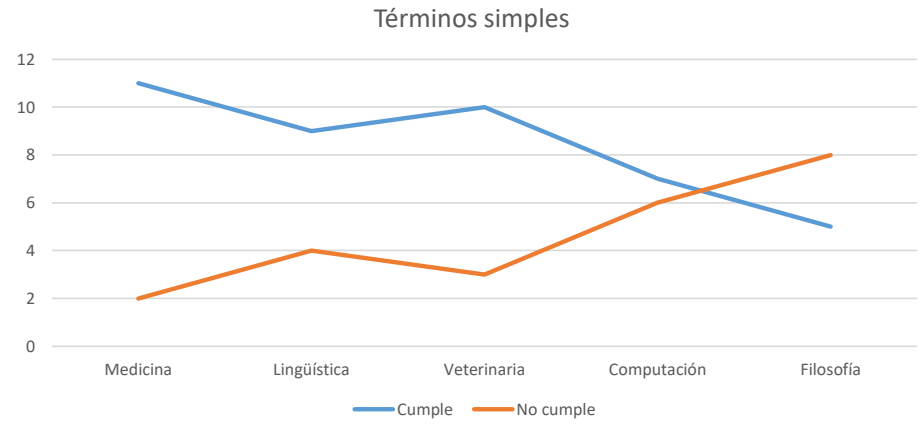


Figura 33. Comparativo criterios validación términos simples.

Como puede observarse en la Figura 33, para el caso de *medicina* esta se concibe como un área del saber prototípica en la cual los criterios establecidos en la determinación de qué es y no un término se cumplen con la regularidad esperada, mientras que en el caso de *filosofía* la tendencia es inversa, es decir, en esta área los términos responden a características no regulares bajo la mirada del protocolo. De este modo, para ilustrar nuestro punto, para esta última los nombres generales sí son términos, contrario a lo indicado tradicionalmente por la terminología hasta el momento de la publicación de la presente tesis doctoral. A modo de ejemplo, tenemos *cosa* como un sustantivo general etiquetado como término para la *filosofía*. Además, también podría afirmarse que existen características propias en cada área del saber, lo que nos permite introducir a la discusión sobre la variación

en terminología la noción de *tipologicidad terminológica*, es decir, hay características propias de cada área que le permiten distinguirse y a la vez integrarse en función de los demás campos del saber.

En este sentido, la *variación terminológica horizontal tipológica* está determinada por características inherentes a las dinámicas adoptadas por cada saber en la manera de configurar, expresar y representar su conocimiento. Asimismo, creemos que, en función de los núcleos del saber (las ciencias de la salud, las ciencias humanas, las ciencias exactas, entre otras), las características deben ser compartidas gracias a la naturaleza epistemológica común. A la luz de nuestros datos, la *medicina* y la *veterinaria* comparten características similares como se evidencia en la Figura 33 mientras que la *computación* y la *filosofía* se apartan de las características prototípicas ratificando la necesidad de integrar en los análisis una mirada tanto individual/particular como en conjunto en virtud de los núcleos del saber. Para el caso de la *lingüística* esta área se encuentra más cerca de ser un área prototípica como la *medicina* en contraposición con lo ya indicado para la *filosofía*. Al observar cada término simple en relación con cada subcorpus podría concluirse que toda unidad es factible de ser término, más allá de sus características intrínsecas, dependiendo del campo del conocimiento, validando implícitamente uno de los criterios determinantes en la construcción del protocolo de validación, a saber: *la arbitrariedad* (2.3.1.3). En adelante, se presentan algunos hallazgos en relación con las unidades simples que permiten ilustrar lo hasta aquí planteado. Además, los ejemplos que soportan las subsiguientes reflexiones están listados en el capítulo de Resultados (§ 4).

En primera instancia, en relación con los adjetivos término (§2.3.2.1) en el corpus VaTeHo se logró no solo identificar su función terminológica sino además su rol en la determinación del área temática, característica que comparte no solo con los sustantivos sino también con todas las unidades léxicas aquí estudiadas, toda vez que hacen referencia a un concepto al interior de cada área del saber. En ese sentido, los adjetivos término son aquellos derivativos o relacionales ya sean denominativos o deverbales. Además, los adjetivos calificativos,

considerados como generales, también pueden ser terminológicos gracias a su función sintáctica al completar el contenido terminológico de la estructura compleja en la cual están inmersos. En consecuencia, su diferenciación subyace en que mientras los primeros funcionan autónomamente en la recuperación del sentido de especialidad, los segundos necesitan de un núcleo para complementar su significación. Otra característica importante con algunos adjetivos tiene que ver con la posibilidad de ser tanto adjetivos como sustantivos, siendo clave la posición al interior de la oración para asignar su rol sintáctico, como el caso de *léxico* (Lin) ejemplificado en 165 y 166.

165. Por ejemplo, la evolución del *léxico* (N_Término) indígena puede influir en la variación social, como es el caso de «tote» 'revólver'. (Lin_COL_04).

166. El conocimiento *léxico* (Adj_Denominal) presentó correlaciones significativas y positivas de tamaño considerable con las inferencias en general [...] (Lin_ARG_01).

Otro elemento destacado con los adjetivos está relacionado con los colores, los cuales también pueden funcionar sintácticamente como sustantivos. En nuestros datos se corrobora el carácter terminológico de estas unidades, por un lado, al hacer referencia a conceptos temáticos en áreas como la *medicina*, la *veterinaria* o *computación* y, por otro lado, en áreas como la *lingüística* y la *filosofía*, donde sin hacer referencia a conceptos propios del campo siguen conservando su rol como términos. De este modo, en sintonía con lo hasta ahora planteado en relación con las *tipologicidad* en función de cada área del saber, esta información va en soporte del análisis terminológico por núcleos del saber y no de manera global. En 167 se presentan ejemplos de colores término, mientras que en 168 se ejemplifica su función terminológica sin estar directamente relacionados con el área analizada.

167. Esta finca posee vacas mestizas de doble propósito resultado del cruce alterno de Brahman *rojo* o Gir (Bos indicus) y Holstein *rojo* (Bos taurus). (Vet_VEN_04.txt).

168. Este último destaca con recuadros de color *amarillo* los venezolanismos descritos, haciendo que el usuario pueda visualizarlos con extrema facilidad y completar exitosamente la búsqueda. (Lin_VEN_04.txt).

En segunda instancia, los adverbios terminológicos (§2.3.2.2) son aquellos deadjetivales terminados en el sufijo *-mente*, lo que supone un primer acercamiento a la caracterización de esta unidad en el discurso de especialidad. Asimismo, en los datos observados, contrario a la característica de los adjetivos generales terminológicos, los adverbios que provienen de este tipo de adjetivos no cumplen rol como términos más su función está orientada a la construcción temática debido a su capacidad sintáctica de modificar, entre otras, construcciones de orden nominal. En tercera instancia, los sustantivos terminológicos (§2.3.2.3) ampliamente estudiados y caracterizados en la literatura, se consolidan como la unidad prototípica de la terminología de allí que el 84,12 % de los candidatos muestrales analizados correspondan con estructuras nominales y de estos el 31,44 % son sustantivos. Así, el sustantivo término hace referencia a nombres que bien sea por composición o derivación pertenecen a un campo del saber en específico. Sin embargo, para el caso de *filosofía*, teniendo en cuenta los resultados de la validación, los sustantivos generales son términos en esta área. En 169 y 170 se listan algunos ejemplos de sustantivos generales término para este campo del saber.

169. Bacon afirma: La mano desnuda o el entendimiento abandonado a él mismo no tiene un gran poder, la *cosa* se hace gracias a los instrumentos y a los apoyos que son tan necesarios a la inteligencia como a la mano [...] (Fil_CHL_05.txt).

170. Aquí también hay gradación, lo que explica que mientras la noción de *objeto* es estática, la de objetividad es claramente dinámica. (Fil_CRI_02.txt).

Por último, en cuarta instancia, los verbos terminológicos (§2.3.2.4) se configuran como una unidad con características

sintácticas y semánticas relevantes, lo que los sitúa en un lugar de privilegio junto con los sustantivos. Así, a través del verbo no solo se logra identificar los rasgos distintivos de un área del saber en contraste con las demás, sino que también se constituye como el núcleo base para las interacciones con las demás unidades, gracias a su función sintáctica en relación con el predicado además de su capacidad flexiva por excelencia ya descrita, por ejemplo, en los trabajos de Alcoba (1999, p. 4917). Unido a esto, con el verbo término se puede recuperar los procesos realizados por las distintas áreas al igual que auxilian en la construcción temática de cada campo del saber y, en consecuencia, concordamos con la tipología presentada por Lorente, en especial, en aquellos que la autora denomina verbos terminológicos (Lorente, 2002, § 8.4).

5.1.2 Los términos complejos en español

Para el caso de los términos complejos (§2.3.2.5), teniendo como base los trece criterios recuperados a partir del protocolo de validación y en la misma línea descrita con las unidades simples, la aplicación del instrumento confirma el carácter prototípico de la *medicina* en contraste con la *filosofía* como áreas en los dos extremos de los saberes analizados gracias a las características ya mencionadas (§ 3.2.1.1). De esta manera, la tendencia se mantiene confirmando que el análisis terminológico sí debe ser orientado a una concepción mediada por los núcleos del saber y no bajo el concepto global de conocimiento de especialidad en el cual todas las áreas se han venido trabajando de la misma manera, en especial, las predominantes (medicina, economía) en los análisis en varias de las investigaciones de mayor impacto en nuestra área lo que, en consecuencia, conlleva a la consolidación del criterio de *arbitrariedad* como un factor clave en la caracterización de las unidades complejas como términos, condición que también se evidencia al constatar que 14,50 % de los términos complejos validados fueron etiquetados bajo el criterio de *arbitrariedad*. La Figura 34 integra los resultados de la validación de los términos complejos en cada uno de los subcorpus analizados.

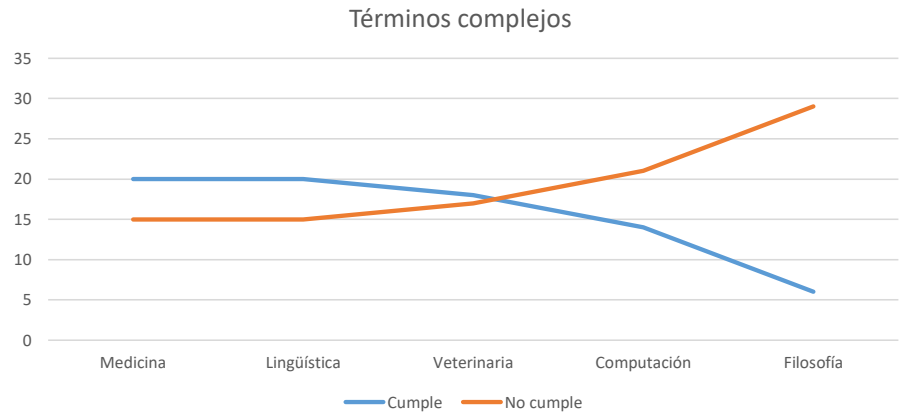


Figura 34. Comparativo criterios validación términos complejos.

Al respecto, al concatenar ambas tendencias en las Figuras 33 y 34, es importante indicar que la regularidad entre las unidades simples y complejas se mantiene, siendo esto un elemento esperado en la expresión del saber en cada área, donde ambos tipos de unidades conservan la misma dinámica de cara a una estructuración tipológica del saber, es decir, epistemológicamente, los términos concebidos en cada campo del conocimiento se rigen por las mismas premisas indistinto de su construcción según la cantidad de componentes. Este elemento, hasta ahora no mencionado explícitamente en el área, reafirma la idea del término como etiqueta y de allí que su construcción sea constante tanto en aquellos simples como complejos, distinción que conservamos en nuestro trabajo por motivos de orden operativo en pro de la distinción de ambos tipos, como ya se mencionó en nuestro marco teórico (§ 2.3.1.2). En general, el mayor porcentaje de los términos complejos se corresponde con unidades donde el núcleo se constituye como un *tipo de* término en relación con la noción de propiedades compartidas. Asimismo, los datos analizados nos permitieron observar particularidades, hasta ahora tangencialmente descritas en la literatura en terminología, pero ampliamente abordados desde otros campos de la lingüística, en relación con los términos metafóricos, los sintagmas terminológicos y las reducciones léxicas, en especial las siglas. No obstante, haremos una mención preliminar de estas unidades toda vez que por la amplitud en la discusión y en

relación con el alcance de nuestro escrito, se hace necesario dedicar, incluso, otras investigaciones en terminología a su estudio individual.

5.1.3 Otras unidades terminológicas

En primera instancia, los términos metafóricos cuyo potencial de análisis puede orientarse desde lo lingüístico y no solamente desde elementos de orden cognitivo como ha sido lo tradicional en los trabajos fraseológicos sobre este tipo de elementos (Lakoff & Johnson, 1980; Ruiz de Mendoza & Hernández, 2011). Por un lado, los términos metafóricos (Burgos, 2024) están presentes en los cinco subcorpus de trabajo, lo que también aporta un elemento más en la relación intrínseca de la lengua general con la de especialidad. Por otro lado, algunos hallazgos preliminares indican que en áreas como la computación hay una alta frecuencia de términos metafóricos en contraste con una menor frecuencia en áreas como la filosofía. De nuevo, es necesario un análisis en mayor profundidad que permita confirmar estas pesquisas iniciales, lo que constituirá un importante avance en la caracterización de unidades de origen fraseológico con carácter terminológico. Por último, consideramos que las metáforas al igual que los términos son considerados como etiquetas lo que permite su caracterización a la luz del protocolo de validación y, por ende, el acceso a su análisis a partir de aspectos sintácticos y semánticos. En este sentido, un total de 1.600 metáforas candidatas a término fueron recuperadas a partir del corpus. En 171 se listan algunos términos metafóricos en nuestro corpus.

171. **Medicina**: *abordaje quirúrgico* (Metf hecho de), *arco de movimiento* (Metf hecho de), *blanco terapéutico* (Metf hecho de), *agente infeccioso* (Metf hecho de), *aparato digestivo* (Metf hecho de).

Lingüística: *acceso léxico* (Metf hecho de), *alargamiento vócalico* (Metf hechode), *almacenamiento ortográfico* (Metf hecho de), *campo léxico* (Metf hecho de), *familia lingüística* (Metf hecho de).

Veterinaria: *árbol filogenético* (Metf hecho de), *bomba de leche* (Metf hecho de), *cabeza de ganado* (Metf hecho de), *carga animal* (Metf hecho de), *marcador genético* (Metf hecho de).

Computación: *ambiente virtual* (Metf hecho de), *banco de filtro* (Metf hecho de), *barra de herramientas* (Metf hecho de), *capa de almacenamiento* (Metf hecho de), *puerto USB* (Metf hecho de).

Filosofía: *agente moral* (Metf hecho de), *circulo hermenéutico* (Metf hecho de), *clima filosófico* (Metf hecho de), *juego dialéctico* (Metf hecho de), *plano ontológico* (Metf hecho de).

En segunda instancia, los Sintagmas Nominales Especializados Extensos (SNEE) son considerados como un elemento clave en el discurso especializado en función de la interacción entre expertos (Newmark, 1988, pp. 39-41), a pesar de la poca proliferación de investigaciones orientadas a su estudio a partir de la gramática o en español como lengua de análisis, condición que se contrasta con el creciente interés, desde la terminología, en la descripción lingüística y funcional de estos (Montero, 1995; Cardero, 2003, 2004; Cortés, 2004; Quiroz, 2008, 2014; Sanz, 2012). En este sentido, los SNEE se definen como un tipo de estructura nominal a partir de un sustantivo núcleo modificado por más de dos unidades (Quiroz, 2008, p. 21). Ahora bien, el valor terminológico de los SNEE aún se presenta como un desafío toda vez que estos suelen ser observados como unidades composicionales y de allí que se asigna su valor especializado más no su categorización como términos (Newmark, 1988; Alvar, 1996; Quiroz, 2008; Cabezas, 2019; entre otros) salvo en los trabajos de Martí (2009) quien los considera como términos sintagmáticos lo que, en consecuencia, bajo el criterio de *composicionalidad* (§ 2.3.1.2), confirmaría la tendencia prototípica al definirlos como no términos. Al respecto, tomando los datos en VaTeHo, (Quiroz, 2024) logró recuperar un total de 84.367 SNEE candidatos a término a partir de un aproximado de sesenta estructuras de los cuales obtuvo una primera muestra de trabajo compuesta por 67.562 candidatos, en 10 estructuras sintácticas, distribuidos de la siguiente manera: 17.093 en *medicina*, 13.267 en *lingüística*, 9.126 en *veterinaria*, 11.323 en *computación* y 16.753 en *filosofía*.

Al respecto, en consonancia con los hallazgos de este autor, se ha logrado indagar en esta característica de los SNEE donde, inicialmente, sí existen SNEE terminológicos a partir de una reevaluación del criterio de composicionalidad (§ 2.3.1.2). De esta manera, es posible plantear una ampliación de este criterio toda vez que es también factible hablar de una *composicionalidad terminológica* en la cual el resultado de los componentes del sintagma sería también considerados como una etiqueta sin importar el número de pre o posmodificadores, condición que estaría determinada por la prevalencia del criterio de arbitrariedad en cuanto la unidad es catalogada como término si esta representa un concepto al interior del campo del conocimiento dejando a un lado la aplicación de los demás criterios en el protocolo de validación (§ 2.3). Así, se establecería una diferencia entre *extirpación de colon* y *Virus de (la) Inmunodeficiencia Humana* (VIH), donde el primero constituye una descripción especializada y el segundo es considerado como un término (para nuestro caso un sintagma terminológico). De nuevo, es necesario un estudio más en detalle que corrobore no solo esta característica de los sintagmas sino también la noción introducida en este trabajo doctoral sobre la *composicionalidad terminológica*.

En tercera instancia, las siglas, como uno de los muchos procesos de reducción léxica. Al igual que las metáforas y los sintagmas nominales, estas han sido documentadas y categorizadas ampliamente desde otros ámbitos y, en menor medida, desde el terminológico (Alvar, 1996; Gutiérrez, 1998; Zuluaga et al. 2017; Rivas, 2017; Rivas et al., 2021). En este sentido, los fenómenos de reducción léxica se conciben como un elemento lingüístico utilizado desde hace varios siglos como una estrategia que permite, por un lado, atender a la economía del lenguaje y del espacio y, por otro lado, generar procesos nemotécnicos pertinentes en diferentes entornos de interacción especializada (Gutiérrez, 1998; Figueroa & Silva, 2000). Asimismo, los procesos de reducción léxica, tanto simples como complejos, también pueden generar *oscuridad* al momento de la trasmisión del contenido toda vez que el no conocimiento del significado de estas por parte del interlocutor puede desencadenar en la no comprensión de la unidad en cuestión (Fijo, 2003, p. 91). En el ámbito terminológico,

la reducción léxica manifestada, principalmente, en las *siglas*, es el reflejo de la subespecialización del conocimiento prototípica de áreas como la medicina y la necesidad de una compresión de dichos saberes en función de la transmisión del contenido de especialidad.

Al respecto, en la caracterización de las siglas se han observado distintos rasgos gramaticales, a saber: a) a nivel morfológico (Grange & Bloom, 2000; Alcaraz, 2002; Giraldo, 2006, entre otros) en relación con su extensión, las formas plurales y el género; b) a nivel sintáctico en su categorización como sustantivos (Rodríguez, 1987; Cannon, 1989; Alcaraz, 2002; entre otros) o a partir del cambio de categoría gramatical (a adjetivos) por cuestiones de orden estilístico o expresivo (Rodríguez, 1987) y c) a nivel semántico en la representación del elemento al cual hace referencia, esto es, siglas nominales o conceptuales (Figueroa & Silva, 2000). De este modo, las siglas se constituyen como uno de los elementos léxicos distintivos a la hora de observar la variación terminológica en el eje vertical tal como ha sido expresado, por ejemplo, en los trabajos de Freixa (2005), Giraldo (2010), Fernández-Silva & Becerra (2015), Guantiva, Cabré & Casrellà (2008), entre otros. Ahora, si se parte del presupuesto que las siglas son consideradas como sustantivos y, por ende, es factible catalogarlas como términos, entonces por extensión su forma extendida, que en su mayoría supera los tres *tokens*, también debería ser considerada como tal. Sin embargo, a este punto, no hay un estudio que permita corroborar este aspecto e incluso proporcione más elementos en pro de una caracterización de las siglas como términos. Asimismo, al igual que para los SNEE, Quiroz, Zuluaga & Rivas (2024) lograron recuperar un total de 4.236 siglas candidatas a término desde el corpus VaTeHo. A su vez, obtuvieron una muestra inicial para posteriores análisis de 1.699 siglas distribuidas en cada subcorpus de la siguiente manera: 677 en *medicina*, 230 en *lingüística*, 565 en *veterinaria*, 178 en *computación* y 49 en *filosofía*. En 172 se presentan algunas siglas recuperadas del corpus VaTeHo con sus respectivas formas extendidas.

172 **Medicina**: IAM (*Infarto Agudo de Miocardio*), HTA (Hipertensión Arterial), AIB (*Análisis de Impedancia Bioeléctrica*), CCGaS (*Carcinoma Gástrico de Célula en Anillo de Sello*).

Lingüística: AFI (*Alfabeto Fonético Internacional*), CNRTL (*Centre National de Ressourses Textuelles et Lexicales*), CORDE (*Corpus Diacrónico del Español*), DCECH (*Diccionario crítico Etimológico Castellano e Hispánico*).

Veterinaria: AACREA (*Asociación Argentina de Consorcios Regionales de Experimentación Agrícola*), ASOCEBU (*Asociación Colombiana de Criadores de Ganado Cebú*), CISA (*Control Integral de Salud Animal*), CONAMA (*Comisión Nacional del Medio Ambiente*).

Computación: CEM (*Campos Electromagnéticos*), CONELEC (*Consejo Nacional de Electricidad*), DD (*Disco Duro*), IBM (*International Business Machines*), LEDs (*Light Emitting Diodes*).

Filosofía: CRP (*Crítica de la Razón Pura*), ICPIC (*Consejo Internacional para la Investigación Filosófica con Niños*), LD (*Lógica Difusa*), TTP (*Tratado Teológico-Político*).

5.2 Recapitulación

En la presente tesis doctoral la concepción del término en español está orientada a una mirada amplia y no restrictiva del fenómeno. Así, en primer lugar, se ha indicado que todas las unidades léxicas, ya sean simples o complejas, son susceptibles de ser términos, más allá de las restricciones de las que han sido objeto, no solo en su estudio en las diferentes investigaciones en terminología, sino debido a criterios globalizantes que parecen dejar en el olvido los fundamentos epistemológicos de cada campo del saber, característica que impacta de forma directa en las dinámicas de comunicación del conocimiento de especialidad mediado por las elecciones en las distintas formas para su

materialización. En este sentido, en segundo lugar, el estudio de cada una de las cinco (5) áreas analizadas permitió observar sus características particulares en contraste con las demás, lo que abre la discusión hacia la existencia de una regularidad temática que, posiblemente, es compartida por los distintos núcleos del saber y, en consecuencia, la necesidad de definir criterios no generales sino tipológicos en los distintos análisis en terminología. Asimismo, en tercer lugar, se hace pertinente una reflexión en relación con las diferentes herramientas para el procesamiento de datos terminológicos, las instrucciones y configuraciones dadas a estas de cara a incluir esta mirada amplia del término gracias a los resultados en nuestro trabajo.

En cuarto lugar, una perspectiva de futuro en la continuación de nuestros análisis orientados a seguir caracterizando aquellas unidades poco estudiadas desde la terminología, a saber: los términos metafóricos, los sintagmas terminológicos y las siglas. Así, la mirada amplificadora de nuestro objeto de estudio se expande a otros desafíos que permitirían completar el panorama de la comunicación de especialidad. Incluso, bajo esta perspectiva, también se hace necesario ahondar en la caracterización de las unidades fraseológicas e identificar en ellas su potencial terminológico, reto aun mayor teniendo en cuenta que en la fraseología su caracterización es un tema pendiente de solucionar. Por último, en quinto lugar, en línea con el elemento anterior, fruto de nuestro trabajo emergen nuevos conceptos que demandan un acercamiento en mayor detalle para lograr su corroboración y, por ende, nuevas perspectivas de entendimiento de nuestro objeto común. De este modo, la *variación terminológica horizontal tipológica*, la *composicionalidad terminológica* y *la opacidad terminológica* se introducen como nuevos conceptos pendientes por explorar en nuevos estudios en terminología.

6. Conclusiones: una mirada al futuro

"We can only see a short distance ahead,
but we can see plenty there that needs to be done."

(Alan Turing).

El presente capítulo marca el punto final del recorrido trazado en la observación y la caracterización de los *términos* a partir de los campos del saber analizados en esta tesis doctoral. Un recorrido en la comprobación de nuestras hipótesis de trabajo guiados por los objetivos que a finales del año 2018 comenzaron a orientar nuestro trasegar. Un camino a la luz de una *mirada en conjunto* en la continuación de nuestra labor terminológica. Un paso a paso en la comprensión de lo dicho hasta ahora en terminología, reflexiones a lo largo del globo terráqueo y nuestro propósito de unir ambos lados del Atlántico a través del estudio del término en español y sus dinámicas en la materialización de los saberes humanos. Una mirada teórica desde las diferentes escuelas de pensamiento terminológico que decantan en nuestra propuesta de ampliación del concepto de *término* más allá de las construcciones nominales prototípicas y los campos del saber privilegiados en las investigaciones (*medicina* y *economía*). Una apuesta metodológica a partir de la implementación del protocolo para la validación de términos simples y complejos (Quiroz & Burgos, en prensa) en la descripción lingüística del término en función de sus características sintácticas, semánticas, pragmáticas y contextuales compartidas con la lengua general, desde un ámbito natural al ser humano en su apuesta por materializar las formas como se concibe el mundo a través de las diferentes realidades epistemológicas. Unos resultados con los cuales buscamos aportar en la construcción de nuestra disciplina. Un punto que no marca el final, sino que antecede

el camino a seguir gracias a los retos aún vigentes y los desafíos que nacen en este trabajo de investigación doctoral.

En este sentido, nuestro trabajo de investigación se presenta como un insumo no solo a la terminología sino también a las demás disciplinas que confluyen en torno a los lenguajes de especialidad. En primer lugar, desde lo teórico, esta tesis doctoral proporciona una reflexión conceptual que permite observar transversalmente el desarrollo de la terminología desde sus albores hasta el momento de la publicación del presente manuscrito. Un *paradigma integrador* con el cual resaltamos la importancia de las distintas perspectivas de análisis terminológico en la construcción y caracterización del *término* como nuestro objeto de estudio común, más allá de las concepciones alternativas que han marcado la dinámica de investigación en terminología posterior a la TGT. En segundo lugar, desde lo metodológico, la implementación del protocolo de validación permitió observar el panorama terminológico de manera amplia al incluir en el análisis criterios para la caracterización de todo tipo de unidad léxica con carácter terminológico. Asimismo, las decisiones en la compilación del corpus textual especializado VaTeHo son una guía que responde al reto aún vigente en la construcción de recursos terminológicos en español que den respuesta a la realidad lingüística de esta lengua a lo largo y ancho del planeta. Por último, en tercer lugar, resultado de la integración teórico-metodológica de nuestro trabajo, sin olvidar que la naturaleza de este tipo de investigación terminológica (lingüística) no propende por aspectos didácticos, nuestra tesis doctoral es un insumo que permite enriquecer la formación en terminología en áreas afines como la traducción, al propiciar contenido susceptible de ser debatido y analizado en espacios de formación que, sin lugar a duda, impactan directamente en el desempeño profesional de las actuales y futuras generaciones.

En adelante se presentan las ideas finales de nuestra tesis doctoral. En primer lugar, se abordan las hipótesis y objetivos de investigación a la luz de los resultados y reflexiones presentadas a lo largo de cada capítulo trabajado. Una evaluación del impacto de nuestro trabajo y el

aporte en la consolidación de nuestra mirada en torno a los desafíos de la terminología. En segundo lugar, los aspectos a futuro producto de las limitaciones de nuestra investigación, las cuales permiten proyectar nuevas líneas de trabajos en el campo y, en consecuencia, abrir el horizonte para otros retos en la construcción de nuestro objeto de estudio común: el *término*.

6.1 Validación de las hipótesis y objetivos de investigación

Nuestra investigación en variación terminológica horizontal contó con dos hipótesis de trabajo las cuales se corroboraron en el marco de los datos aquí compilados y, por ende, el alcance de nuestra investigación fue de corte descriptivo y su validación endógena. En este sentido, se hace necesario ampliar los datos de trabajo de manera que a partir de la inferencia estadística se pueda corroborar significativamente lo hasta ahora descrito. Al respecto, la no aplicación de estadística inferencial no invalida ni resta valor a los hallazgos de nuestro trabajo toda vez que en consonancia con Larsson et al. la naturaleza de los datos lingüísticos no siempre demanda un análisis inferencial para responder a sus cuestionamientos (2022, p. 150). Asimismo, no desconocemos el aporte y la importancia de la estadística en una actualidad donde las masas de información requieren su mirada, pero también concordamos con la premisa de volver al origen y a la esencia de la investigación en lingüística y no olvidar que más allá de los valores significativos o no, en la delicadeza de la descripción cualitativa seguimos construyendo nuestra ciencia.

En la primera hipótesis de trabajo, enmarcada en el ámbito de la variación terminológica, se indagó sobre la existencia de diferencias, en el plano lingüístico, en las características del término en función de los distintos campos del conocimiento. Al respecto, *sí existen diferencias* lo que permite ampliar el panorama actual de trabajo al identificar la necesidad de repensar los análisis hasta ahora tradicionales donde las prácticas terminológicas han encerrado el conocimiento en una suerte de única mirada desconociendo las bases

epistémicas de cada saber y sus elecciones en la materialización de sus conocimientos. En este sentido, se logró identificar que áreas como la *medicina* cumplen con los rasgos prototípicos hasta ahora descritos en la literatura en terminología, mientras que la *filosofía* se aparta e introduce nuevas características hasta ahora miradas tangencialmente como el carácter terminológico de los nombres comunes o los adjetivos generales. Además, los datos analizados permitieron observar que en saberes afines como la *medicina* y la *veterinaria* estas características se comparten, lo que permite incluir en la discusión una perspectiva de análisis por *núcleos del saber*, elemento que aportaría en un mejoramiento de las herramientas terminológicas en función de sus configuraciones en la ejecución de los distintos procesos para los cuales fueron diseñadas. A su vez, la *lingüística* se configura como un saber más cercano a lo prototípico y la *computación* como un área entre lo y no prototípico. En consecuencia, estas distinciones nos permiten hablar de una *variación terminológica horizontal tipológica*.

En la segunda hipótesis de trabajo, donde se cuestionó la mirada del paradigma actual en la terminología, se discutió acerca de la categorización de todas las unidades léxicas, fuesen simples o complejas, como términos más allá de las construcciones nominales. Los resultados nos permitieron identificar que, en efecto, todas estas unidades hasta ahora observadas tangencialmente *sí son términos*. Así, es pertinente hablar del *adjetivo término*, el *adverbio término* y el *verbo término*, además de las *estructuras complejas término* entre las cuales se observan otros fenómenos lingüísticos como el *metafórico, de reducción léxica* (en nuestro caso especial las siglas) y la *sintagmación*. Esta última nos permitió ahondar en uno de los criterios base del protocolo con relación a la composicionalidad e introducir el concepto de *composicionalidad terminológica* a partir, por ejemplo, de la estructura [N+Adj+Conj+Adj] donde casos como *léxico científico y técnico* (Lin) no es un término sino dos elementos en contraste con *ovario voluminoso e irresecable* (Med) donde se estaría haciendo referencia a un tipo de ovario. Este aspecto se configura

como otro de los retos a futuro toda vez que consideramos puede ser un tema de investigación individual.

En relación con el objetivo general de investigación, hemos *logrado describir lingüísticamente* el comportamiento de los términos en español a partir de cinco contextos de especialidad diferentes (*medicina, lingüística, veterinaria, computación* y *filosofía*) y, por ende, su análisis en el ámbito de la *variación terminológica horizontal*. En este sentido, la consolidación del corpus VaTeHo es el reflejo del cumplimiento de los objetivos específicos a través de cada una de las etapas de la investigación. De este modo, en primer lugar, en las discusiones llevadas a cabo para la *definición de los criterios de compilación del corpus* subyace uno de los pilares fundamentales en su categorización como un corpus especializado toda vez que, por un lado, se abarcan saberes de mayor y menor impacto y de interés social, lo que nos permitió observar la *variación terminológica* en un amplio espectro del conocimiento humano y, por otro lado, la inclusión de información en la mayoría de los países donde el español es lengua oficial. En segundo lugar, la *extracción de los candidatos a término*, a partir del uso de *TermoStat*, nos proporcionó un espacio adicional de análisis toda vez que en el procesamiento de los datos identificamos una serie de rasgos, como la ausencia de estructuras sintácticas existentes en el español, con los cuales se podría optimizar la herramienta para los procesos de extracción en esta lengua.

Unido a esto, la utilización de herramientas computacionales al servicio del Procesamiento de Lenguaje Natural nos permitió reflexionar sobre la importancia de estos recursos en unión con la labor del lingüista en la procura de un análisis no automático sino apoyado en las nuevas tecnologías. El correcto balance entre estos dos componentes permite que nuestras investigaciones tengan un mayor rango de acceso a la información en menor tiempo del utilizado en los análisis manuales y, por ende, una optimización de los recursos y la posibilidad de observar los fenómenos trabajados desde perspectivas que se escapan a los análisis casuísticos, aún recurrentes en muchas tradiciones de investigación en nuestro contexto latinoamericano.

En tercer lugar, la *validación lingüística* de los candidatos extraídos automáticamente, al igual que aquellos recuperados de forma manual e inductiva, para los propósitos de ejemplificación de nuestras discusiones, teniendo como punto de partida el protocolo para la validación de términos simples y complejos (Quiroz & Burgos, en prensa), es otro de los pilares con los cuales consolidamos nuestra mirada amplia del término y, en consecuencia, la propuesta de *ampliación del concepto*.

6.2 Aportes de la investigación a la terminología

Guardando los límites de la prudencia, la presente tesis doctoral se configura como un aporte a la discusión en el campo de la terminología. En primer lugar, la compilación del corpus textual especializado permite al contexto de investigación terminológica hacerse de un nuevo recurso disponible para su consulta y posterior exploración. Por un lado, el corpus VaTeHo representa un recurso *del español*, balanceado en cuanto a su composición y que apuesta por una representación conforme a la realidad lingüística de esta lengua a lo largo y ancho del planeta. Por otro lado, las cinco áreas del saber que conforman el corpus permitieron una caracterización lingüística del saber teniendo en cuenta las distintas tendencias en relación con los núcleos del conocimiento y, a razón de ello, otra perspectiva de análisis en la variación terminológica horizontal en procura de ampliar los postulados clásicos que emergieron bajo las premisas positivistas del siglo pasado, que para la actualidad ya no son tan vigentes debido a la hibridación cada vez mayor de los campos del saber. En segundo lugar, con nuestra propuesta para la ampliación de la definición del *término* se abren las puertas para una mirada complementaria de la realidad terminológica, al introducir en los análisis otros puntos de investigación, dibujando nuevos elementos en el actual panorama de trabajo terminológico y, de esta manera, contribuir en una terminología mucho más cercana a las realidades y dinámicas de la comunicación humana. Creemos, firmemente, que esta mirada amplia del término

enriquecerá nuestras discusiones y permitirá nuevas perspectivas de análisis en complemento de aquellas actuales y vigentes.

En tercer lugar, nuestra investigación en terminología, desde la lingüística computacional y de corpus, sigue contribuyendo en la consolidación de este tipo de investigaciones, a nivel doctoral, en nuestro contexto colombiano y latinoamericano. Asimismo, en nuestra *mirada en conjunto* subyace la reivindicación de la terminología en Latinoamérica, esto es, una puesta en escena de los avances y las aportaciones desde cada una de las escuelas de pensamiento que permitieron la consolidación de nuestra disciplina desde una fundamentación teórica y metodológica, que a su vez es reflejo de una cruzada identitaria en defensa de nuestras variedades dialectales en sintonía con el español ibérico. En cuarto lugar, cada uno de los aspectos desarrollados en esta tesis doctoral se entrelazan en función de una *ampliación del paradigma* actual de trabajo no solo desde la *puerta lingüística* por la cual hemos ingresado, sino a partir del entendimiento del término en lo cognitivo, social y cultural. Una vez más, la terminología más allá de su falsa concepción como un lenguaje artificial es un ámbito más en el que el lenguaje enriquece las dimensiones del ser humano. Hay tantas miradas como posibilidades de entender las dinámicas de los discursos de especialidad, en el disenso seguiremos consolidando el final del rompecabezas terminológico.

Por último, en quinto lugar, teniendo presente que esta tesis doctoral se enmarca en el proyecto de investigación para la Variación Horizontal y Vertical del Grupo de Investigación en Traducción y Nuevas Tecnologías (TNT), es pertinente mencionar que unido a este trabajo doctoral, se contará con la publicación por la editorial *Routledge* de un manual en terminología (2024) bajo la coordinación de los doctores en lingüística Gabriel Quiroz, Diego Burgos y el magíster en traducción Felipe Zuluaga. En dicho manual se abordará el término en español desde aspectos teóricos, metodológicos y pedagógicos en consonancia con nuestra mirada amplia de término y con miras a la consolidación de un nuevo aporte a la terminología a partir de los distintos capítulos contenidos en este y que propenden por un análisis

y caracterización de los fenómenos terminológicos a través de una fundamentación desde una perspectiva de corte lingüística.

6.3 Limitaciones y perspectivas de trabajo futuro

Una vez alcanzados los propósitos de nuestra investigación, es menester indagar sobre los desafíos a futuro que nacen gracias a las diferentes reflexiones en torno a nuestra mirada amplía del ámbito terminológico, un acercamiento que nos permitió observar, desde el pasado, el presente de la terminología y proyectar el devenir de nuestro campo de trabajo e investigación. De este modo, las distintas limitaciones son el punto inicial para el trabajo a futuro y la responsabilidad implícita de continuar ahondando en nuestra propuesta de término y en la ratificación de las hipótesis de investigación y, de esta manera, lograr una consolidación de nuestra propuesta teórica y metodológica alrededor del *término* y la *terminología*.

En primer lugar, se hace necesario continuar trabajando en el corpus VaTeHo, por un lado, a partir de la inclusión no solo de otras áreas del saber, desde las ciencias, las técnicas y las artes, sino también de la ampliación de la información hasta ahora compilada y, por otro lado, la ausencia de información en algunos de los países seleccionados para recolectar los datos es una limitante importante, pero que se escapa del margen de acción, toda vez que es un factor externo no controlable teniendo como base los criterios establecidos para la conformación del corpus, en los cuales se trató de garantizar un corpus balanceado del español. En línea con lo anterior, queda planteada la discusión sobre el *término metafórico*, los *sintagmas terminológicos* y el término en los distintos casos de *reducción léxica* (como el aquí abordado en relación con las *siglas*) y, en consecuencia, la apuesta a una caracterización teórica y un acercamiento metodológico, que permitan un desarrollo riguroso de los conceptos de *variación terminológica horizontal tipológica, composicionalidad terminológica* y *opacidad terminológica* los cuales, en el desarrollo

de este trabajo de investigación, fueron mencionados tangencialmente, en esencia, por la necesidad de delimitar el alcance de este debido al tiempo disponible para su conclusión.

En segundo lugar, es necesario trascender el alcance descriptivo de la investigación al llevar a cabo pruebas estadísticas inferenciales que permitan obtener conclusiones en términos de representatividad del fenómeno trabajado. De esta manera, es necesaria una validación exógena de los datos desde la cual el *protocolo de validación de términos simples y complejos* pueda ser replicado en otras investigaciones y a la luz de diversas perspectivas de análisis terminológico. Asimismo, otro aspecto sobre el cual se debe profundizar está relacionado con el rol del experto en la validación de la información en contraste con el uso de diccionarios especializados y teniendo en cuenta los procesos de desterminologización y el procesamiento del contenido de especialidad del experto en comparación con el realizado por un lego. Al respecto, los resultados de la presente investigación, eventualmente, cuestionaran la forma como hasta el día de hoy se viene compilando diversidad de recursos lexicográficos especializados, cuyas entradas son catalogadas, *per se*, como términos bajo los principios restrictivos del paradigma terminológico y, de allí, sus características prototípicas.

Por último, en tercer lugar, esta tesis doctoral propende por una mirada integradora de la terminología, lo que en sí demanda la necesidad de continuar con esta perspectiva de trabajo a partir de nuevas investigaciones y de su divulgación y socialización en distintos escenarios de discusión académica, tanto a nivel nacional como internacional. Una perspectiva complementaria de la terminología que también nos invita a una reflexión desde el plano didáctico (Gómez, Arroyave & Muñoz, 2024), una puesta en escena en las aulas, en torno a la transmisión y discusión académica de los distintos elementos desarrollados en esta tesis doctoral. Finalmente, una mirada al futuro de la *ciencia de los términos*, las puertas siguen abiertas a la discusión en terminología.

Referencias

Adelstein, A. (1998). Representación de unidades terminológicas: el modelo de término wüsteriano. *Organon*, *12*(26).

Adelstein, A. (2002) "Estatuto lingüístico del término y modelos semióticos de representación". En: M. Fuentes & J. García (Eds.) *Texto, terminología y traducción* (pp. 99-144). Ediciones Almar.

Adelstein, A. (2004). *Unidad léxica y valor especializado: estado de la cuestión y observaciones sobre su representación* [Tesis de maestría, Universidad Pompeu Fabra].

Adelstein, A. (2007). *Unidad léxica y significado especializado: modelo de representación a partir del nombre relacional madre* [Tesis doctoral, Universidad Pompeu Fabra].

Alarcos, E. (1969). Aditamento, adverbio y cuestiones conexas. *Archivum, 19,* 301-329.

Alarcos, E. (1994). *Gramática de la lengua española*. Espasa-Calpe.

Alcaraz, M. (2002). Las siglas en el discurso biomédico escrito en inglés: análisis y aplicaciones didácticas. *The ESPecialist, Pesquisa Em Línguas Para Fins Específicos. Descrição, Ensino E Aprendizagem, 23*(1), 37–5.

Alcina, F., & Blecua, J. (1975). *Gramática Española*. Editorial Ariel.

Alcoba, S. (1999). La flexión verbal. En I. Bosque. & V. Demonte (Eds.). *Gramática Descriptiva de la Lengua Española* (pp. 4915-4992).

Alpízar, R. (2009). La terminología en Cuba unos lustros después. En A. Lau (Rev.), *Panorama sobre la terminología en América Central, México y el Caribe*, (16-21). Unión Latina.

Alvar, M. (1996). *La formación de palabras en español*. Arco Libros.

Arntz, R., & Picht, H. (1995). *Einführung in die Terminologiearbeit*. Traducción del alemán: De Irazazábal, Amelia et al., *Introducción a la terminología*. Fundación Germán Sánchez Ruipérez, Pirámide.

Auger, P. (1974). Observation de la synonymie dans la terminologie minière. En M. Héroux (Ed.). *La normalisation linguistique: Actes du Colloque International de Terminologie,* 25-33.

Auger, P. (1988). La terminologie au Québec et dans le monde, de la naissance à la maturité. En *Actes du sixième colloque OLF-STQ de terminologie. L'ère nouvelle de la terminologie*, 27-59.

Auger, P. (1994). Implantabilité et acceptabilité terminologiques: les aspects linguistiques d'une socioterminologie de la langue du travail. *Terminologies nouvelles*, *12*, 47-57.

Bachelard, G. (1957). *La poética del espacio*. Fondo de Cultura Económica.

Baduy, M., Garda, M., & Perassi. M. (2016). Terminología e investigación para la formación en traducción. En M. del Valle & E. Grimaldi (Comp.), *Políticas lingüísticas y lenguas extranjeras en el Nivel Superior*, (pp. 481-491). Universidad Nacional del Litoral.

Baker, P., Hardie, A., & McEnery, T. (2006). *A Glossary of Corpus Linguistics*. Edinburgh University Press.

Barbosa, M. (2006). Para uma etno-terminologia: recortes epistemológicos. *Ciência e Cultura*, *58*(2), 48-51.

Barité, M. (2017). La definición en terminología. En S. Álvarez & M. Barité (Coords.), *Teoría y praxis en terminología* (pp. 45-67). Ediciones Universitarias.

Beneš, E. (1969). Zur Typologie der Stilgattungen der wissenschaftlichen Prosa. *Deutsch als Fremdsprache*, *6*(3), 225-233.

Bevilacqua, C. (2004). *Unidades fraseológicas especializadas eventivas: descripción y reglas de formación en el ámbito de la energía solar* [Tesis doctoral, Universidad Pompeu Fabra].

Biber, D. (1993). Representativeness in Corpus Design. *Literary and Linguistic Computing, 8*(4), 243-257.

Biber, D. (2015). Corpus-Based and Corpus-Driven Analyses of Language Variation and Use. En B. Heine & H. Narrog (Eds.), *The Oxford Handbook of Linguistic Analysis 2nd. Edition* (pp. 193-224). Oxford University Press.

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus Linguistics. Investigating Language Structure and Use*. Cambridge University Press.

Bogomilova, E. (2010). *La terminología del derecho ambiental mexicano. Identificacion y procesos de formación* [Tesis doctoral, El Colegio de México].

Boyan, A. (2005). *Contrastive Aspects of Terminological Metaphor* [Tesis doctoral, Universidad de Sofia].

Brainerd, B. (1972). On the Relation between Types and Tokens in Literary Text. *Journal of Applied Probability*, *9*(3), 507-518.

Burgos, D. (2014). *Towards an Image-term Co-ocurrence Model for Multilingual Terminology Alignment and Cross-language Image Indexing* [Tesis doctoral, Universidad Pompeu Fabra].

Burgos, D. (2024). El término metáforico. En G. Quiroz, D. Burgos, & F. Zuluaga (Eds.), *Terminología del español: el término*. Routledge.

Burgos, D. (2024). Términos complejos. En G. Quiroz, D. Burgos, & F. Zuluaga (Eds.), *Terminología del español: el término*. Routledge.

Burgos, D., & Vásquez, D. (2024). El nombre terminológico. En G. Quiroz, D. Burgos, & F. Zuluaga (Eds.), *Terminología del español: el término*. Routledge.

Cabezas, M. (2019). *Los compuestos nominales en terminología: formación, traducción y representación* [Tesis doctoral, Universidad de Granada].

Cabré, T. (1992). *La terminología: teoría, els mètodes, les aplicacions*. Editorial Empuries.

Cabré, T. (1995). La terminología hoy: concepciones, tendencias y aplicaciones. *Ciência da Informação, 24*(3).

Cabré, T. (1998). *Terminology: Theory, Methods, and Applications* (Vol. 1). John Benjamins Publishing.

Cabré, T. (1999a). El discurs especialitzat o la variació funcional determinada per la temática. En M. Cabré (Ed.), *La terminología: representación y*

comunicación. Elementos para una teoría de base comunicativa y otros artículos (pp. 151-173). Instituto Universitario de Lingüística Aplicada.

Cabré, T. (1999b). Hacia una aproximación teórica de base comunicativa. Elementos para una teoría de la terminología: hacia un paradigma alternativo. En M. Cabré (Ed.), *La terminología: representación y comunicación. Elementos para una teoría de base comunicativa y otros artículos* (pp. 69-92). Instituto Universitario de Lingüística Aplicada.

Cabré, T. (2000). Terminologie et linguistique: la théorie des portes. *Terminologies nouvelles*, (21), 10-15.

Cabré, T. (2004). La terminología en la traducción especializada. En: C. Gonzalo & V. García (Eds.) *Manual de documentación y terminología para la traducción especializada* (pp. 89-122). Arco/Libros.

Cabré, T. (2005). A Terminologia, uma disciplina em evolução: passado, presente e alguns elementos de futuro. *Debate Terminológico*, (1).

Cabre, T. (2007). La Terminología, una disciplina en evolución: pasado, presente y algunos elementos de futuro. *Debate terminológico,* (1).

Cabré, T. (2012). Daniel Prado, el pilar fundamental de la terminología iberoamericana. *Debate Terminológico*, (08), 82-83.

Cabré, T. (19 de noviembre 2020). Tendencias en terminología. En M. Monteagudo (Moderadora), *Ciclos de eventos virtuales RITerm 2020.*

Cabré, T., & Estopà, R. (2005). Unidades de conocimiento especializado: caracterización y tipología. En Documenta Universitaria (Eds.), *Coneixement, llenguatge i discurs especialitzat* (pp. 69 – 94).

Cabré, T., Estopà, R., & Vivaldi, J. (2001). Automatic Term Detection: A Review of Current Systems. En D. Bourigault, C. Jacquemin & M. L'Homme (Eds.), *Recent Advances in Computational Terminology* (pp. 53-87). John Benjamins Publishing Company.

Cabré, T., Domènech, O., & Estopà, R. (2018). *La terminologia avui: termes, textos i aplicacions*. Editorial UOC.

Cannon, G. (1989). Abbreviations and Acronyms in English Word-Formation. *American Speech 64*(2), 99-127.

Cardero, A. (2003). *Terminología y procesamiento.* Ediciones ACATLAN.

Cardero, A. (2004). *Lingüística y terminología*. UNAM.

Cardona, B. (1997). *Manual metodológico para la elaboración del tesauro agrícola Colombiano*. EIB.

Cardona, B. (1999). Red colombiana de terminología -Colterm-. *Revista Interamericana de Bibliotecología*, *22*(1), 129-138.

Cardona, M., Pérez-Pérez, C. M., & Zuluaga, F. (2024). El verbo terminológico. En G. Quiroz, D. Burgos, & F. Zuluaga (Eds.), *Terminología del español: el término*. Routledge.

Casademont, A. (2008). *Sintaxi i semàntica verbals en el discurs d'especialitat: elements per a l'activació del valor terminològic* [Tesis doctoral, Universidad Pompeu Fabra].

Chiquito, A., & Quiroz, G. (2017). *Pobreza, lenguaje y medios en América Latina*. Peter Lang.

Ciapuscio, G. (1998). La terminología desde el punto de vista textual: selección, tratamiento y variación. *Organon*, *12*(26).

Ciapuscio, G. (2000). Hacia una tipología del discurso especializado. *Revista iberoamericana de discurso y sociedad*, *2*(2), 39-71.

Collet, T. (1997). La réduction des unités terminologiques complexes de type syntagmatique. *Meta: journal des traducteurs/Meta: Translators' Journal*, *42*(1), 193-206.

Collet, T. (2004). What's a Term? An Attempt to Define the Term within the Theoretical Framework of Text Linguistics. *Linguistica Antverpiensia, New Series–Themes in Translation Studies*, *3*.

Cortés, J. (2004). *Análisis lingüístico de términos comparados en inglés técnico agrícola* [Tesis de doctorado, Universidad Complutense de Madrid].

Costa, R., & Silva, R. (2004). The Verb in the Terminological Collocations. Contribution to the Development of a Morphological Analyser: MorphoCom. En M. Lino, M. Xavier, F. Ferreira, R. Costa & R. Silva (Eds). *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC'04)* (pp. 1531-1534). European Language Resources Association (ELRA).

Da Cunha, I. (2008). *Hacia un modelo lingüístico de resumen automático de artículos médicos en español*. [Tesis doctoral, Universidad Pompeu Fabra]. Tesis Doctorals en Xarxa. https://www.tesisenred.net/handle/10803/7508.

De Barcellos, G. (2008). *Terminología en Brasil* [Diapositivas de Power Point]. Instituto de Lingüística Aplicada (IULA), Universidad Pompeu Fabra. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwiGw4Sc4JH6AhXKfjABHUjiDR4QFnoECAUQAQ&url=http%3A%2F%2Fwww.iula.upf.edu%2Fmaterials%2F081203almeida.pdf&usg=AOvVaw0t5YvaSk9bpaIZCU5An2c1.

De Irazazábal, A. (2010). Situación de la terminología española. En T. Cabré (Ed.), *Terminología y modelos culturales*, 13-30. Documenta Universitaria.

De la Peña R. (1985). *Gramática teórica y práctica de la lengua castellana*. UNAM.

De Nebrija, A. (1492:2010). *Gramática de la lengua castellana*. Linkgua Digital.

De Saussure, F., Bailly, C., & Séchehaye, A. (1919:1967). *Cours de linguistique générale. Collection Payothèque*. Payot & Rivages.

De Yzaguirre, Ll. (1996). Ingeniería lingüística y terminología. *Terminómetro. Monográfico: La terminología en España,* 69-71.

Demonte, V. (1999). El adjetivo: clases y usos. La posición del adjetivo en el sintagma nominal. En I. Bosque & V. Demonte (Eds.), *Gramática descriptiva de la lengua española. Volumen 1: sintaxis básica de las clases de palabras* (pp. 129-215). Espasa.

Deutsches Institut für Normung. (2022). *Vocabulary of Terminology* (DIN 2342). https://tienda.aenor.com/norma-din-2342-2011-08-140370128.

Diallo, S. (2020). Socioterminología: orígenes, objeto y aplicaciones. *Informatio. Revista del Instituto de Información de la Facultad de Información y Comunicación*, *25*(2), 22-40.

Diéguez, M. (2008). Análisis terminológico de sitios web localizados del inglés al español: uso de técnicas de amplificación y elisión. *Tradumática*, (6).

Diéguez, M. (2014). Ileana Cabrera: traductora y terminóloga. *Debate Terminológico*, (12), 65-67.

Diki-Kidiri, M. (2008). *Le vocabulaire scientifique dans les langues africaines: Pour une approche culturelle de la terminologie*. Karthala Editions.

Dilthey, W. (1883:1944). *Introducción a las ciencias del espíritu*. Fondo de cultura económica.

Dixon, R. (2004). Adjective Classes in Typological Perspective. En R. Dixon & A. Aikhenvald (Eds.) *Adjective classes: A cross-linguistic typology*, (pp. 1-49). Oxford University Press.

Dixon, R. (2010). *Basic Linguistic Theory, Vol. II*. Oxford University Press.

Domènech, O. (2006). *Textos especialitzats i variació vertical: la diversitat terminològica com a factor discriminant del nivell d'especialització d'un text* [Tesis doctoral, Universitat Pompeu Fabra].

Drouin, P. (2003). Term Extraction Using Non-technical Corpora as a Point of Leverage. *Terminology*, *9*(1), 99-117.

Drouin, P. (2010). *Présentation du logiciel TermoStat 3.0*. Observatoire de linguistique Sens-Texte. http://termostat.ling.umontreal.ca/doc_termostat/doc_termostat.html.

Dubuc, R. (1980). *Manuel Pratique de Terminologie*. Linguatech.

Dubuc, R. (1999). *Manuel pratique de terminologie*. Traducción del francés: Cabrera, Ileana., *Manual práctico de terminología*. RIL editores.

Estopà, R. (1999). *Extracció de terminologia: elements per a la construcció d'un SEACUSE (Sistema d'Extracció Automàtica de Candisats a Unitats de Significació Especialitzada)*. [Tesis doctoral, Universidad Pompeu Fabra].

Faber, P. (2012). *A Cognitive Linguistics View of Terminology and Specialized Language*. Mouton De Gruyter.

Faber, P., & L'Homme, M. (2022). *Theoretical Perspectives on Terminology: Explaining Terms, Concepts, and Specialized Knowledge*. John Benjamins Publishing Company.

Faulstich, E. (1998). Principes formels et fonctionnels de la variation en terminologie. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, *5*(1), 93-106.

Faulstich, E. (2006). A socioterminologia na comunicação científica e técnica. *Ciência e Cultura*, *58*(2), 27-31.

Fedor de Diego, A. (1995). *Terminología. Teoría y práctica*. Equinoccio-Unión Latina.

Felber, H., & Picht, H. (1984). *Métodos de terminografía y principios de investigación terminológica*. Instituto Miguel de Cervantes.

Felber, H. (1984). *Terminology Manual*. Unesco and Infoterm.

Fernández-Silva, S., & Becerra, N. (2015). La variación terminológica en la comprensión y producción de textos académicos: Propuesta de representación en un diccionario especializado de aprendizaje de Psicología. *Ibérica,* (30), 183-202.

Figueroa, B., & Silva, T. (2000). *Un diccionario de especialidad: las siglas*. En M. Casal, G. Conde, J. Lago, L. Pino & N. Rodríguez. (Eds.), *La lingüística francesa en España camino del siglo XII*, (pp. 455-467). Arrecife.

Fijo, M. (2003). *Las siglas en el lenguaje de la enfermería: análisis contrastivo inglés-español por medio de fichas terminológicas* [Tesis doctoral, Universidad Pablo de Olavide].

Fillmore, C. (1982). Frame Semantics. *Linguistics in the Morning Calm*, 110-137.

Fillmore, C. (1992). Corpus Linguistics or Computer-aided Armchair Linguistics. En J. Svartvik (Ed.) *Directions in Corpus Linguistics* (pp. 35-60). De Gruyter.

Folguerà, R. (2002). *Adjectius en el discurs espacialitzat: Una primera descripció dels adjectius en els textos del Genoma Humà* [Tesis de maestría, Universitat Pompeu Fabra].

Food and Agriculture Organization of the United Nations (2022). *Agrovoc*. Recuperado noviembre 07, 2022, de https://agrovoc.fao.org/browse/agrovoc/en/.

Foucault, M. (1970). *La arqueología del saber*. Siglo XXI Editores.

Foucault, M. (1993). *Esto no es una pipa. Ensayo sobre Magritte*. Anagrama.

Foucault, M. (2013). *¿Qué es usted, profesor Foucault? Sobre la arqueología y su método*. Siglo XXI Editores.

Freixa, J. (2005). La variació denominativa a través de la variació vertical. *Coneixement, llenguatge i discurs especialitzat. Institut Universitari de Lingüística Aplicada*, 111-129.

Frías, M. (1996). *Procesos creativos para la producción de textos*. Editorial Magisterio.

Galeano, M. (2017). *La terminología de la ganadería* [Tesis de maestría, Universidad de Antioquia].

Gambier, Y. (1987). Problèmes terminologiques des pluies acides: pour une socio-terminologie. *Meta 32*(3), 314-320.

García-Donoso, P. (2002). Traducción y terminología en la Pontificia Universidad Católica del Ecuador. En B. Osio (Dir.), *La terminología en los países andinos,* (43-45). Unión Latina.

Gaudin, F. (2005). La socioterminologie. *Langages* (157), 80-92.

Gentile, A. (2007). *La circulation sociale du discours de la psychanalyse en langue espagnole: un fonctionnement polynomique?* [Tesis doctoral, Universidad de Ruan].

Gili Gaya, S. (1994). *Curso superior de sintaxis española*. VOX.

Giraldo, J. (2006). *Sistemas de detección y extracción semiautomática de siglas Estado de la cuestión* [Reporte de investigación, Université de Toulouse - Le Mirail].

Giraldo, J. (2008). *Análisis y descripción de las siglas en el discurso especializado de genoma humano y medio ambiente* [Tesis doctoral, Universidad Pompeu Fabra].

Giraldo, J. (2010). Hacia una revisión del concepto de siglación. *Panace@, 11*(31), 70-76.

Gómez, N., Muñoz, C., & Arroyave, A. (2024). La enseñanza de la terminología: el uso de talleres como alternativa metodólogica en el aula. En G. Quiroz, D. Burgos, & F. Zuluaga (Eds.), *Terminología del español: el término*. Routledge.

Gómez, Y., Naranjo, M., Giraldo, D., Pineda, J., & Suárez, M. (2019). Terminología experimental y traducción especializada: un estudio sobre la distribución de la atención visual y la presión de tiempo. *Terminàlia*, 20-29.

Grange, B., & Bloom, D. (2000). Acronyms, Abbreviations and Initialisms. *BJU international 86*(1), 1-6.

Guantiva, R., Cabré, T., & Casrellà, J. (2008). *Classification of Specialized Texts Based on Their Terminologies*. Ikala, *13*(1), 15-39.

Guespin, L. & Laroussi F. (1989). Glottopolitique et standardisation terminologique. *La Banque des mots*, número especial, 5-21.

Gutiérrez, B. M. (1998). *La ciencia empieza en la palabra. Análisis e historia del lenguaje científico*. Península.

Hernández, C. (1996). *Gramática funcional del español actual*. Gredos.

Hernanz, M., & Brucart, J. (1987). *La sintaxis*. Crítica.

Hoffmann, L. (1998). *kommunikationsmittel fachsprache: eine einführung*. Traducción del alemán: Herrerias, Mercè., *Llenguatge d'especialitat. Selecció de textos*. Institut Universitari de Lingüística Aplicada.

International Organization for Standarization (ISO). (2019). *Terminology Work and Terminology Science – Vocabulary* (ISO 1087). https://www.iso.org/standard/62330.html.

International Organization for Standarization (ISO). (2022). *Terminology Work - Principles and Methods* (ISO 704). https://www.iso.org/standard/79077.html.

Isquerdo, A., & Alves, I. (2007). *As ciências do léxico: lexicologia, lexicografia, terminologia* (Vol. 3). Editora Humanitas.

Kolguixkin, A. (1970). *Lingvística v voiénnom dele*. Moscou.

Krieger, M. (2000). Terminologia revisitada. *Delta: documentação de estudos em lingüística teórica e aplicada*, *16*, 209-228.

Kuguel, I. (2006). *La semántica del léxico especializado: los términos en textos de ecología* [Tesis doctoral, Universidad de Buenos Aires].

Lafon, P. (1980). "Sur la variabilité de la fréquence des formes dans un corpus". *MOTS* (1), 128-165.

L'Homme, M. (2002). What can Verbs and Adjectives Tell us about Terms? *Terminology and Knowledge Engineering*, 28-33.

L'homme, M. C. (2004). *La terminologie: principes et techniques*. Pum.

Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. University of Chicago Press.

Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. University of Chicago.

Lara, L. (2009). Para la historia de los estudios terminológicos en México. En A. Lau (Rev.), *Panorama sobre la terminología en América Central, México y el Caribe*, (37-38). Unión Latina.

Larsson, T., Egbert, J., & Biber, D. (2022). On the Status of Statistical Reporting Versus Linguistic Description in Corpus Linguistics: A Ten-year Perspective. *Corpora*, *17*(1), 137-157.

Leech, G. (1992). Corpora and Theories of Linguistic Performance. En J. Svartvik (ed.) *Directions in Corpus Linguistics* (pp. 105-122). De Gruyter.

Lopera, G. (2016). *Sintagmas nominales con premodificación compleja con guiones en un corpus lexicográfico especializado bilingüe* [Tesis de maestría, Universidad de Antioquia].

Lorente, M. (2002). Verbos y discurso especializado. *Estudios de lingüística del español*, *16*. https://84.88.27.11/index.php/Elies/article/view/195493.

Lotte, D. (1961). *La formación del sistema de términos científico-técnicos. Textos de terminólogos de la Escuela Rusa*. Institut Universitari de Lingüística Aplicada.

Lugo, T. (1982). *Semántica de las palabras compuestas con un elemento activo. Comparación ruso-español* [Tesis doctoral, Universidad Patricio Lumumba].

Luna, R. (2012). La terminología en el campo de los trastornos del espectro autista (TEA). *Lengua, societat i comunicación*, 31-38.

Marcellesi, J., & Gardin, B. (1974). *Introduction à la sociolinguistique: La linguistique sociale*. Larousse.

Márquez, M. (2006). Los anglicismos terminológicos integrales en los textos especializados del español. *Estudios de lingüística aplicada*, (43), 11-29.

Martí, M., & Llisterri, J. (2000). La ingeniería lingüística en la Sociedad de la Información. *Digithum*. https://joaquimllisterri.cat/publicacions/FDS2000.html.

Martí, M. (2009). La discriminación de los sintagmas terminológicos en los glosarios especializados. *Lingüística Aplicada Actual, 1*(31), 61-88.

Martínez, S. M., & Castro, M. B. (2017). Clasificación semántica de colocaciones verbales para la adquisición y codificación de conocimiento experto: El caso de los riesgos naturales. *Revista Española de Lingüística Aplicada*, *30*(1), 240-272.

McCray, A. T, Srinivasan, S., & Browne, A. (1994). Lexical methods for managing variation in biomedical terminologies. In *Proceedings of the Annual Symposium on Computer Application in Medical Care* (pp. 235-239). American Medical Informatics Association.

McEnery, T. (2019). *Corpus Linguistics*. Edinburgh University Press.

Milard, B. (2010). Las citaciones científicas: redes de referencias en universos de referencias. El ejemplo de los artículos de química. *Redes. Revista hispana para el análisis de redes sociales*, *19*, 69-93.

Mira, G. (2009). *La traducción de las unidades fraseológicas en Rosario Tijeras: análisis pragmalingüístico y contrastivo* [Tesis de maestría, Universidad de Antioquia].

Mira, G. (2010). La traducción de las unidades fraseológicas desautomatizadas en rosario tijeras. *Núcleo, 22*(27), 217-237.

Montero, B. (1995). *La estructura del grupo nominal complejo en el inglés científico escrito. Sus componentes premodificadores y sus correspondencias en español* [Tesis doctoral, Universidad de Valencia].

Montero, S. (2014). Los nombres escuetos y el artículo Ø. *Verba hispánica*, *22*(1), 37-58.

Montoya, G., & Zapata, C. (2005). El taller: una estrategia para la normalización de términos y conceptos en un trabajo terminológico. *Análisis*, *6*(1), 25-35.

Myking, J. (2000). Sosioterminologi-ein modell for Norden? En A. Nuopponen, B. Toft & J. Myking (Eds.): *I terminologins tjänst. Festskrift för Heribert Picht på 60-årsdagen*, 92-109.

Navigli, R., Velardi, P., & Gangemi, A. (2003). Ontology learning and its application to automated terminology translation. *IEEE Intelligent systems*, *18*(1), 22-31.

Newmark, P. (1988). *A Textbook of Translation*. Oxford Pergamon Press.

Nunberg, G., Sag, I., & Wasow, T. (1994). Idioms. *Language*, *70*(3), 491-538.

Orduña-Malea, E. (2016). La revolución Google Scholar: Destapando la caja de Pandora académica. *La revolución google scholar*, 1-268.

Parodi, G. (2007). *Lingüística de Corpus y Discursos Especializados: Puntos de Mira*. Ediciones Universitarias de Valparaíso.

Parodi, G. (2008). Lingüística de Corpus: Una introducción al ámbito. *Revista de Lingüística Teórica y Aplicada, 46*(1), 93-119.

Patiño, P. (2017). *Description and Representation in Language Resources of Spanish and English Specialized Collocations from Free Trade Agreements* [Tesis doctoral, NHH University].

Pavel, S. (1993). Neology and Phraseology as Terminology-in-the-Making. En H. Sonneveld & K. Loening (Eds.), *Terminology Applications in Interdisciplinary Communication* (pp. 21-34). John Benjamins Publishing Company.

Pavel, S., & Nolet, D. (2001). *Handbook of Terminology*. Public Works and Government Services.

Pavel, S., & Nolet, D. (2002). *Manual de terminología*. Bureau de la Traduction.

Perassi, M., & Ferreira, A. (2016). Hacia una formación en traducción especializada: realidades y desafíos. *Íkala, revista de lenguaje y cultura*, *21*(1), 49-61.

Pérez-Pérez, C. M. (2018*). Metodología para la caracterización de patrones [verbo + objeto directo] sobre pobreza en la prensa colombiana para propósitos de Análisis de Sentimientos* [Tesis de maestría, Universidad de Antioquia].

Pérez-Pérez, C. M., Quiroz, G., & Tamayo, A. (2020). ¿Tiene carácter predictivo la estructura predicativa [verbo + objeto directo]? hacia una caracterización sintáctico-semántica para propósitos de análisis de sentimientos. *Lingüística y Literatura*, *41*(78), 11-35.

Plested, M. (1987). *Untersuchungen zu Problemen der Ertmittlung und Bestimmung von DDR-spezifischen Lexemen und ihren Äquivalenzen im Spanischen. 2Bde* [Tesis doctoral, Universidad Humboldt].

Plested, M. (2010). Editorial: Una mirada de conjunto. *Ikala, 15*(25), 9-17.

Plested, M., Mejia. M., Nicod, T., Gutiérrez, G., & Betancur-Jaramillo, A. (2015). Armonización conceptual en las ciencias forenses: una mirada terminológica desde la clínica forense. *Colombia Forense, 2*(1), 11-19.

Poey y Aloy, F. (1851). *Memorias sobre la historia natural de la isla de Cuba: accompañadas de sumarios latinos y extractos en frances* (Vol. 1). Barcina.

Poey y Aloy, F. (1858). *Memorias sobre la historia natural de la isla de Cuba: accompañadas de sumarios latinos y extractos en frances* (Vol. 2). Barcina.

Porta, J. (2006). Técnicas cuantitativas para la extracción de términos en un corpus. En E. Rambla, M. Alcina & E. Valero (Eds.), *Terminología y Sociedad del Conocimiento* (pp. 147-178) *Escuela Politécnica Superior–Universidad Autónoma de Madrid*.

Pozzi, M. (1997). El valor actual de la terminologia. *Íkala, Revista de Lenguaje y Cultura*, *2*(1-2), 109-121.

Pozzi, M. (1999). The Concept of "Concept" in Terminology: A Need for a New Approach. En P. Sandrini (Ed.), *The '99 Terminology and Knowledge Engineering* (pp. 28-42). TermNet.

Pozzi, M. (2003). Hacia la armonización de la terminología usada en las normas del comité técnico iso/tc37: identificación semi-automática de términos problemáticos en un corpus. *Estudios de Lingüística Aplicada,* (38), 109-127.

Pozzi, M. (2009). La investigación terminológica en el Colegio de México. En A. Lau (Rev.), *Panorama sobre la terminología en América Central, México y el Caribe*, (39-42). Unión Latina.

Quiroz, G. (2002). La terminología en Colombia: panorama general y algunas pautas históricas. *Terminómetro. Número especial. La terminología en los países andinos*, *6*, 3-6.

Quiroz, G. (2003). Preparación y procesamiento de un corpus para la creación de materiales en clase de español para propósitos específicos. *Actas del X Congreso Brasileño de Profesores de Español*, 131-150.

Quiroz, G. (2006). Using an English-Spanish Parallel Corpus to Solve Complex Premodification in Noun Phrases. En M. Gotti & S. Sarcevic (Eds.). *Translation of Specialized Text. Linguistic Insights series* (pp. 367-390). Peter Lang.

Quiroz, G. (2008). *Los sintagmas nominales extensos especializados en inglés y en español: descripción y clasificación en un corpus de genoma* [Tesis doctoral, Universitat Pompeu Fabra].

Quiroz, G. (2024). El sintagma nominal extenso especializado. En: G. Quiroz, D. Burgos, & F. Zuluaga (Eds.), *Terminología del español: el término*. Routledge.

Quiroz, G., & Arroyave, A. (2014). On Premodified Terms in Five Specialized Dictionaries. En G. Quiroz & P. Patiño (Eds.), *LSP in Colombia: advances and challenges* (pp. 137-154). Peter Lang.

Quiroz, G., & Arroyave, A. (2018). Sintagmas nominales con premodificación compleja: algunos aspectos de su traducción del inglés al español. En Colín, Marisela (Coord.). *Traducción de textos especializados. Nuevos enfoques, nuevas metodologías* (pp. 196-255). Universidad Nacional Autónoma de México.

Quiroz, G., & Burgos, D. (en prensa). *Protocolo de validación de términos simples y complejos. (Artículo inédito)*. Universidad de Antioquia, Wake Forest University.

Quiroz, G., & Diaz, G. (2024). El adverbio terminológico. En G. Quiroz, D. Burgos, & F. Zuluaga (Eds.), *Terminología del español: el término*. Routledge.

Quiroz, G., & Vásquez D. (2024). El adjetivo terminológico. En: G. Quiroz, D. Burgos, & F. Zuluaga (Eds.), *Terminología del español: el término*. Routledge

Quiroz, G., Burgos, D., & Zuluaga, F. (2024). *Terminología del español: el término*. Routledge.

Quiroz, G., Zuluaga, F., & Rivas, N. (2024). Procesos de abrevaciones en terminología. En G. Quiroz, D. Burgos, & F. Zuluaga (Eds.), *El término en español*. Routledge.

Rey, A. (1979). *La terminologie noms et notions*. Presses Universitaires de France.

Reyes, I. (2020). Nube de palabras. *Eutopía*, *12*(33), 55-56.

Riedel, G. (1975). *Syntaktische Besonderheiten der sprachlichen Darstellung von sportlichen Bewegungsabläufen im Russischen* [Tesis doctoral, Alemania].

Ritcher, M. (2010). La presencia del francés en la terminología gastronómica del español de México. [Tesis maestría, El Colegio de México].

Rivas, N. (2017). *Siglas dentro de siglas: el fenómeno de las siglas anidadas en el discurso científico y su comportamiento en las traducciones* [Tesis de maestría, Universidad de Antioquia].

Rivas, N., Quiroz, G., & Giraldo, J. (2021). Identification and Characterization of Nested-abbreviated Terms in Scientific Discourse. *Terminology*, *27*(2), 219-253.

Rodríguez, E. (2011). La terminología y la adquisición de conocimiento especializado. *Lenguaje*, (31), 93-117.

Rodríguez, F. (1987). Naturaleza sintáctica de las formas siglares. El cambio funcional. *Estudios de lingüística, 4*, 139-148.

Rodríguez-Piñero, A., & García, M. (2009). Lenguas de especialidad y lenguas para fines específicos, precisiones terminológicas y conceptuales e implicaciones didácticas. En A. Vera & I. Martínez (Eds), *El español en contextos específicos: enseñanza e investgación Vol. II* (pp. 907-932). Fundación Comillas.

Rojas, J. (2014). *Las unidades fraseológicas en textos especializados* [Tesis de maestría, Universidad de Antioquia].

Rojo, G. (2002). Sobre la lingüística basada en el análisis de corpus. *Hizkunza-corpusak. Oraria eta geroa*, 1-17.

Rondeau, G. (1983). *Introduction à la terminologie*. Gaëtan Morin.

Ruiz de Mendoza, F., & Hernández, L. (2011). The Contemporary Theory of Metaphor: Myths, Developments and Challenges. *Metaphor and Symbol, 26*(3), 161–185.

Sageder, D. (2010). Terminology Today: A Science, an Art or a Practice? Some Aspects on Terminology and its Development. *Brno Studies in English, 36*(1), 123-134.

Sager, J. (1978). Commentary by Prof. Juan Carlos Sager. En G. Rondeau (Ed.), *Actes Table Ronde sur les Problemes du Decoupage du Terms, Montreal* (pp. 39-74). AILA-Comterm.

Sager, J. (1993). *Curso práctico sobre el procesamiento de la terminología*. Ediciones Pirámide.

Salazar, H. (2011). *Descripción y representación de los adjetivos deverbales de participio en el discurso especializado* [Tesis doctoral, Universidad Pompeu Fabra].

Sanz, L. (2011). *Análisis contrastivo de la terminología de la teledetección. La traducción de compuestos sintagmáticos nominales del inglés al español* [Tesis doctoral, Universidad de Salamanca].

Sanz, L. (2012). Approaching Secondary Term Formation through the Analysis of Multiword Units: An English–Spanish Contrastive Study. *Terminology 18*(1), 105-127.

Schmidt, W. (1969). *Charakter und gesellschaftliche Bedeutung der Fachsprachen*. Sprachpflege.

Sierra, G. (2022). Detección automática de léxico especializado en corpus del español. En G. Parodi, P. Cantos-Gómez, & C. Howe (Eds), *Lingüística de corpus en español / The Routledge Handbook of Spanish Corpus Linguistics* (pp. 450-462). Routledge.

Silva Ayçaguer, L. (2012). El índice-H y Google Académico: una simbiosis cienciométrica inclusiva. *Acimed*, *23*(3), 308-322.

Suito, I. (1998). Terminología y traducción. En *Actas del II Congreso Nacional de Investigaciones Lingüístico-Filológicas* (Vol. 1), 251. Universidad Ricardo Palma.

Sur, C. (2020). RBN: Enhancement in Language Attribute Prediction Using Global Representation of Natural Language Transfer Learning Technology like Google BERT. *SN Applied Sciences*, *2*(22), 1-15.

Szabó, Z. (2020, Agosto 17). *Compositionality*. Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/compositionality/.

Tamayo, A. (2018). *Metodología para el análisis automático de sentimientos en documentos a partir de técnicas de aprendizaje automático y lingüística computacional* [Tesis de maestría, Universidad de Antioquia].

Tamayo, A., Angel, D., & Gelbukh, A. (2022). Detección del alcance de las negaciones en español usando Conditional Random Fields. *Research in Computing Science, 5*(151), 199-212.

Tamayo, A., Burgos, D., & Gelbukh, A. (2022a). ParTNER: Paragraph Tuning for Named Entity Recognition on Clinical Cases in Spanish using mBERT+ Rules. En *CEUR Workshop Proceedings*, *3202*.

Tamayo, A., Burgos, D., & Gelbukh, A. (2022b). Nlp-cic-wfu at socialdisner: Disease Mention Extraction in Spanish Tweets using Transfer Learning and Search by Propagation. En *Proceedings of The Seventh Workshop on Social Media Mining for Health Applications, Workshop & Shared Task*, 19-22.

Tebé, C. (2005). *La representació conceptual en terminologia: l'atribució temàtica en els bancs de dades terminològiques* [Tesis doctoral, Universidad Pompeu Fabra].

Temmerman, R. (2000). Towards New Ways of Terminology Description. *Towards New Ways of Terminology Description*, 1-274.

Tercedor, M., Ureña, J. & Prieto, J. (2012). Grasping Metaphoric and Metonymic Processes in Terminology. *Jostrans - The Journal of Specialised Translation* 18, 187-205.

Tesnière, L. (1959). Éléments de syntaxe structurale. Librairie C. Klincksieck.

Torres-Salinas, D., Ruiz-Pérez, R., & Delgado-López-Cózar, E. (2009). Google Scholar como herramienta para la evaluación científica. *El profesional de la información*, *18*(5), 501-510.

Torruella, J., & Sager, J. (1993). *Curso práctico sobre el procesamiento de la terminología*. Pirámide.

Vásquez, D. (2017). *Sintagmas nominales con premodificación compleja con guiones en un corpus de medicina* [Tesis de maestría, Universidad de Antioquia].

Wallerstein, I. (2006). *Abrir las ciencias sociales*. Siglo XXI Editores.

Wright, S., & Budin, G. (1997). *Handbook of Terminology Management. Vol. 1, Basic Aspects of Terminology Management.* John Benjamins Publishing Company.

Wüster, E. (1931). *Internationale Sprachnormung in der Technik: Besonders in der Elektrotechnik. (Die Nationale Sprachnormung und ihre Verallgemeinerung).* [Tesis Doctoral, Universidad de Stuttgart].

Wüster, E. (1998). *Einführung in die allgemeine Terminologielehre und terminologische Lexikographie*. Traducción del alemán: Nokerman, Anne-Cécile., *Introducción a la teoría general de la terminología y a la lexicografía terminológica*. Institut Universitari de Lingüística Aplicada.

Zuluaga, J., Tamayo, A., & Quiroz, G. (2017). La especialización del discurso en los medios en Colombia: el caso de los fenómenos de reducción léxica en la prensa de la pobreza. En A. Chiquito & G. Quiroz (Eds.), *Poverty, Language and Media* (pp. 175-193). Peter Lang.

Zuluaga, J., Quiroz, G., & Tamayo, A. (2021). La especialización del discurso mediático: el caso de las unidades de reducción léxica en la prensa latinoamericana sobre pobreza. En A. Chiquito, L. Pinardi & G. Offenbeck (Eds.), *El lenguaje de la pobreza: Comunicación multifacética* (pp. 101-128). Clacso.

Recursos lexicográficos

Academia Chilena de la Lengua. (2001). *Diccionario de uso del español de Chile (DUECh)*. Mn Editorial.

Academia Nacional de Letras de Uruguay. (2017). *Diccionario del español de Uruguay*. Ediciones de la Banda Oriental.

Alcaraz, E., & Martínez, M. (1997). *Diccionario de lingüística moderna*. Ariel.

Anderson, D. (2003). *Diccionario Mosby: medicina, enfermería y ciencias de la salud*. Elsevier.

Audi, R. (2004). *Diccionario AKAL de Filosofía*. Cambridge University Press.

Ferrater, J. (1978). *Diccionario de Filosofía*. Editorial Sudamericana.

Henríquez, M., & Rozo, N. (2018). *Diccionario de Colombianismos*. Instituto Caro y Cuervo.

Instituto Nacional del Cáncer. (2022). *Diccionario de cáncer del NCI*. Recuperado noviembre 07, 2022, de https://www.cancer.gov/espanol/publicaciones/diccionarios/diccionario-cancer.

Lara, L. (2022). *Diccionario del español de México*. El Colegio de México.

Microsoft (2022). *Portal lingüístico de Microsoft*. Recuperado noviembre 07, 2022, de https://www.microsoft.com/es-es/language.

Ponce de León, N. (1883). *Diccionario Tecnológico inglés-español y español-inglés*. Editorial Néstor Ponce de León.

Real Academia Española. (2022a). *Diccionario de la lengua española* (DRAE). Recuperado noviembre 07, 2022, de https://www.rae.es/.

Real Academia Española. (2022b). *Corpus de Referencia del Español Actual (CREA)*. Recuperado noviembre 07, 2022, de https://www.rae.es/banco-de-datos/crea.

Real Academia Española. (2022c). Cabeza de ganado. En *Diccionario Panhispánico del Español Jurídico*. Recuperado octubre 20, 2022, de https://dpej.rae.es/lema/cabeza-de-ganado#:~:text=Gral.,de%20cualquier%20tipo%20de%20ganado.

Rivas, L., Mayorga, L., & Ravelo, D. (2019). *Diccionario de colombianismos y venezolanismos*. Editorial El Tiempo.

Programas y herramientas computacionales

Anthony, L. (2022). *AntConc* (Versión 4.1.4) [Programa para computador]. Waseda University. https://www.laurenceanthony.net/software.

Brezina, V., Weill-Tessier. P., & McEnery, T. (2020). *#LancsBox* [Programa para computador]. Universidad de Lancaster. http://corpora.lancs.ac.uk/lancsbox/.

Burgos, D. (2014). *BC-Term* [Programa en línea]. Wake Forest University. http://grupotnt.udea.edu.co/BC-Term/.

Davies, J. (2022). *Word Cloud Generator* [Programa en línea]. https://www.jasondavies.com/wordcloud/.

Drouin, P. (2022). *TermoStat* [Software en línea]. http://termostat.ling.umontreal.ca/.

Microsoft. (2022). *Excel* [Programa para computador]. https://www.microsoft.com/es-es/microsoft-365/excel.

Microsoft. (2022). *Word* [Programa para computador]. https://www.microsoft.com/es-es/microsoft-365/word.

Oliver, A., & Vàsquez, M. (2015). *TBXTools* [Programa para computador]. Universidad Abierta de Cataluña. https://sourceforge.net/projects/tbxtools/.

Reed, A. (1997). *Simple Concordance Program (SCP)* [Programa para computador]. http://www.textworld.com/scp/#Description.

RSW Group. (2022). *Multiterm* [Programa para computador]. https://www.trados.com/products/multiterm-desktop/.

Scott, M. (2022). WordSmith Tools (Versión 8) [Programa para computador]. https://www.lexically.net/wordsmith/.

Sinclair, S., & Rockwell, G. (2003). *Voyant Tools* (Versión 2.6.1) [Software en línea]. https://voyant-tools.org/?lang=es.

Epílogo

Monarcas

Al despertar un grito entre lo mágico y el temor. Allí estaban inmóviles al viento, serenas a la espera del tiempo, impasibles e imponentes, silenciosas e indiferentes a una nueva realidad. Han pasado algunos días —metamorfosis, suelen llamarlo— en tu rostro la inmensidad de lo nuevo, lo inimaginable, lo bello, lo maravilloso. En tus ojos el reflejo de lo inexplicable, la alegría de la vida. Con el sol nacen los primeros destellos de luz. Un par de tazas de café, mientras con tu mirada grabas en la memoria cada momento. Sus alas se abren y la magia toma forma entre blancos, negros y naranja. Un par de suspiros escapan de tus labios, néctar y ambrosía. Vuelan los sueños, los deseos y las añoranzas, entre sus alas se marchan los segundos, los minutos, las horas de un pasado incierto, pero de nuevo, lleno de magia, instantes que no se olvidarán. ¡Mientras observas su vuelo, mis sentidos perciben tu realidad... tú eres la Magia en mis días!

¡Para vos!

