



**UNIVERSIDAD NACIONAL DE  
CÓRDOBA**

**Facultad de Lenguas  
Maestría en Traductología**

**Tesis de Maestría**

**EL VOCABULARIO DE LIBROS DE TEXTO UNIVERSITARIOS  
DE INGENIERÍA EN INGLÉS: ESTUDIO DE CORPUS**

**Autora: Natalia Busso**

**Directora: Dra. Verónica L. Muñoz**

**Córdoba, septiembre de 2018**



*A mi esposo Mario,  
por acompañarme cada día con amor, humor e inteligencia.*

*A Sergio,  
amigo entrañable y autor serial de frases e imágenes memorables.*

## Agradecimientos

*Feeling gratitude and not expressing it is like wrapping a present and not giving it.*

William Arthur Ward

Son muchas las personas a quienes quiero agradecer por haberme acompañado, de una u otra manera, a lo largo del proceso de realización de esta investigación.

Para empezar, deseo expresar mi más profundo agradecimiento a la Dra. Verónica Muñoz por haber aceptado dirigirme y, especialmente, por su inmensa generosidad intelectual y nivel de compromiso. Agradezco sus lecturas tan atentas como enriquecedoras y, sobre todo, valoro el haberme transmitido la seriedad y rigurosidad que implica la labor académica.

En segundo lugar, quiero agradecer a aquellas personas que colaboraron de diferentes formas e hicieron posible este trabajo:

Al Decano de la Facultad de Ciencias Exactas, Físicas y Naturales (FCEFyN), Mgtr. Pablo Recabarren, por haberse interesado en la realización de esta tesis y haber puesto la institución a mi disposición.

Al Secretario de Graduados de la FCEFyN, Ing. Oscar Alberto Cáceres, por haberme contactado con profesores de esta casa de altos estudios, algunos de los cuales se convirtieron en asesores especialistas de la presente investigación.

Al Dr. Sergio Preidikman, profesor de la Facultad de Ciencias Exactas, Físicas y Naturales, por haber accedido a convertirse en informante clave del subdominio de la mecánica de los fluidos en este estudio, por hacerse tiempo siempre para reunirme y por explicarme con paciencia, y apelando a distintos recursos y estrategias, nociones complejas de la ingeniería.

Al Dr. Carlos Rodolfo Oldani, profesor de la Facultad de Ciencias Exactas, Físicas y Naturales, por haberse desempeñado como informante clave del subdominio de la ciencia de los materiales en esta investigación.

Al Director del Centro de Estudios y Tecnología del Agua de la FCEFyN, Dr. Carlos Marcelo García Rodríguez, y al Coordinador de la Unidad Técnica de Materiales del Instituto Nacional de Tecnología Industrial en Córdoba (INTI), Lic. Reinaldo Mancini, por haber respondido preguntas sobre la mecánica de los fluidos y la ciencia de los materiales, respectivamente, en una etapa inicial de elaboración de la tesis.

A la Dra. María Teresa Cabré, por haber tenido la enorme gentileza de contestar algunas consultas puntuales sobre terminología y semántica.

También quiero agradecer profundamente a aquellas personas que me acompañaron y brindaron su apoyo en las distintas etapas de desarrollo y escritura de esta tesis:

A mi esposo Mario, por ser mi compañero, asesor, lector y especialista en informática, estadísticas y cuestiones varias.

A mi madre Amanda y a mi tía Raquel, por haberme inculcado, desde muy pequeña, el amor por los libros y la curiosidad por aprender.

A mis amigas Alejandra, Daniela y Luciana, por estar pendientes, escucharme, aconsejarme y alentarme durante todo el proceso de elaboración de este trabajo.

A todos, ¡muchas gracias!

## Resumen

El libro de texto es uno de los géneros de mayor impacto en la formación universitaria de grado. La presente investigación tiene como objetivos principales caracterizar el vocabulario de libros de texto universitarios publicados en inglés en un dominio específico, la ingeniería, y producir listas de palabras técnicas y no técnicas, que ofrezcan a los estudiantes hispanohablantes de esta disciplina un repertorio de vocabulario útil para la lectura comprensiva en inglés. El estudio se sustenta en aportes provenientes de la teoría de análisis de los géneros, investigaciones sobre el vocabulario científico-académico en inglés y la lingüística de corpus. El corpus especializado y pedagógico compilado para esta investigación fue sometido a un análisis cuantitativo y cualitativo organizado en dos etapas: una descripción general del vocabulario y un análisis detallado de las palabras de alta frecuencia identificadas. La descripción general del vocabulario demostró que el corpus tiene una baja variación léxica. Este hallazgo es coherente con la función primordial del libro de texto universitario: introducir a los estudiantes de manera gradual en los conceptos, teorías y prácticas discursivas de una comunidad disciplinar. El análisis inicial de las palabras de alta frecuencia evidenció que la mayoría de ellas, incluidas las palabras generales y académicas, están relacionadas con la ingeniería como disciplina, y por consiguiente, forman parte del vocabulario técnico del corpus. La reclasificación de las palabras de alta frecuencia en técnicas y no técnicas permitió compilar listas de palabras con aplicaciones pedagógicas. Los resultados han aportado mayor información sobre la naturaleza y tipos de vocabulario de los libros de textos de ingeniería. Asimismo, se considera que los resultados contribuirán a la formulación de criterios más sólidos para la enseñanza del vocabulario, como así también para el diseño de material didáctico en cursos de inglés con fines específicos en el campo de la ingeniería.

## Índice

### Capítulo 1. Introducción

1.1 Contextualización del estudio	1
1.2 Estado de la cuestión	7
1.2.1 Estudios lingüísticos en el campo de la ingeniería	7
1.2.2 Estudios sobre el vocabulario en diferentes géneros y disciplinas	7
1.2.3 Estudios sobre el vocabulario en libros de texto universitarios de ingeniería	8
1.3 Objetivos del estudio	11
1.3.1 Objetivos generales	11
1.3.2 Objetivos específicos	11
1.4 Estructura y organización de la tesis	12

### Capítulo 2. Fundamentos teóricos

2.1 Teoría de los Géneros	13
2.1.1 Definición de género	13
2.1.2 Tradiciones académicas	14
2.2 El género libro de texto universitario	16
2.2.1 Contextualización del género	16
2.2.2 Definición	17
2.2.3 Caracterización	18
2.3 Vocabulario	21
2.3.1 Palabras técnicas o términos	23
2.3.1.1 Definición de palabras técnicas o términos	23
2.3.1.2 Identificación de las palabras técnicas o términos	26
2.3.1.2.1 Un continuum de especialización para la identificación de palabras técnicas o términos	26
2.3.1.2.2 Otros métodos empleados para la identificación de palabras técnicas o términos	28
2.3.1.2.3 Criterios semánticos y pragmáticos para la identificación de palabras técnicas o términos	31

2.4 Lingüística de corpus	33
2.4.1 La lingüística de corpus como metodología de investigación	33
2.4.2 El corpus como objeto de estudio	34
2.4.2.1 Definición	34
2.4.2.2 Criterios para el diseño y compilación del corpus	34
2.4.3 Tipos de corpus	37
2.4.4 Unidades de análisis en estudios de vocabulario basados en corpus	38
2.4.5 Técnicas y procedimientos de análisis	39
2.4.5.1 Técnicas cuantitativas: listas de palabras, rango y cobertura	39
2.4.5.2 Técnicas cualitativas: concordancias, uso y significado	42
2.5 Síntesis del capítulo	44
<b>Capítulo 3. Metodología</b>	
3.1 Materiales	45
3.1.1 El corpus	45
3.1.1.1 Preparación de los textos para su análisis	48
3.1.2 El programa informático de análisis de corpus	49
3.1.3 Listas de palabras	50
3.2 Recolección y análisis de los datos	51
3.2.1 Descripción léxica general del corpus	51
3.2.2 Análisis cuantitativo de las palabras de alta frecuencia en el corpus	52
3.2.2.1 Identificación de las palabras de alta frecuencia	52
3.2.2.2 Clasificación y cobertura de las palabras de alta frecuencia	52
3.2.2.3 Lemas de las palabras de alta frecuencia	53
3.2.3 Análisis cualitativo de las palabras de alta frecuencia en el corpus	53
3.3 Síntesis del capítulo	55
<b>Capítulo 4. Resultados</b>	
4.1 Descripción léxica general del corpus	56
4.1.1 Tamaño del corpus y tamaño del vocabulario del corpus	56
4.1.2 Relación entre tipos de palabras y número total de casos en el corpus	57
4.1.3 Rango de los tipos de palabras	57
4.1.4 Clasificación de los tipos de palabras del corpus	60

4.1.5 Cobertura de los distintos tipos de palabras	60
4.2 Identificación y descripción de las palabras de alta frecuencia	62
4.2.1 Selección de las palabras de alta frecuencia	62
4.2.2 Clasificación y cobertura de las palabras de alta frecuencia	63
4.2.3 Lemas de las palabras de alta frecuencia	64
4.2.4 Análisis cualitativo de las palabras de alta frecuencia	65
4.2.4.1 Palabras técnicas o términos	66
4.2.4.2 Palabras no técnicas	74
4.3 Síntesis del capítulo	83
<b>Capítulo 5. Discusión y conclusiones</b>	
5.1 Discusión	84
5.1.1 Descripción general del vocabulario del corpus	84
5.1.2 Análisis cuantitativo y cualitativo de las palabras de alta frecuencia en el corpus	87
5.2 Posibles aplicaciones del estudio	90
5.3 Limitaciones del estudio y recomendaciones para investigaciones futuras	92
5.4 Conclusiones finales	94
<b>Referencias bibliográficas</b>	97
<b>Apéndices</b>	
<b>Apéndice 1</b> Lista completa de lemas de alta frecuencia en el corpus	119
<b>Apéndice 2</b> Lista completa de palabras técnicas o términos de alta frecuencia en corpus	127
<b>Apéndice 3</b> Lista completa de palabras no técnicas de alta frecuencia en el corpus	130

## Índice de tablas y figuras

### Tablas

<b>Tabla 3.1</b> Descripción del corpus	48
<b>Tabla 4.1</b> Rango de los tipos de palabras en el corpus	58
<b>Tabla 4.2</b> Número de tipos de palabras, casos y cobertura de las cuatro listas de palabras	61
<b>Tabla 4.3</b> Los diez lemas más frecuentes de palabras generales, académicas y Otras palabras de alta frecuencia en el corpus	65
<b>Tabla 4.4</b> Palabras técnicas o términos de alta frecuencia en el corpus	73
<b>Tabla 4.5</b> Palabras no técnicas de alta frecuencia en el corpus	81

### Figuras

<b>Figura 4.1</b> Rango de las palabras gramaticales en el corpus	59
<b>Figura 4.2</b> Rango de las palabras restantes en el corpus	59
<b>Figura 4.3</b> Número de tipos de palabras en el corpus	60
<b>Figura 4.4</b> Relación entre tipos de palabras y su cobertura en el corpus	62
<b>Figura 4.5</b> Relación entre tipos de palabras de alta frecuencia y su cobertura en el corpus	64

## Lista de siglas

**AWL:** *Academic Word List*

**BEL:** *Basic Engineering List*

**BNC:** *British National Corpus*

**CIILTU:** Corpus de inglés de ingeniería de libros de texto universitarios

**COCA:** *Corpus of Contemporary American English*

**EAP:** *English for Academic Purposes*

**EC:** *Engineering Corpus*

**EEWL:** *Engineering English Word List*

**EIL:** *English as an International Language*

**ELF:** *English as a Lingua Franca*

**ESP:** *English for Specific Purposes*

**FCEFYN:** Facultad de Ciencias Exactas, Físicas y Naturales

**GrWL:** Lista de palabras gramaticales

**GSL:** *General Service List*

**KWIC:** *Key Word in Context*

**MICASE:** *Michigan Corpus of Academic Spoken English*

**SEEC:** *Student Engineering English Corpus*

**SEWL:** *Student Engineering Word List*

**STTR:** *Standardized Type-Token Ratio*

**TCT:** Teoría Comunicativa de la Terminología

**TGT:** Teoría General de la Terminología

**TTR:** *Type-Token Ratio*

**UFE:** unidades fraseológicas especializadas

**UNC:** Universidad Nacional de Córdoba

**UTP:** unidades terminológicas polilexemáticas o poliléxicas

## Capítulo 1

### Introducción

#### 1.1 Contextualización del estudio

Desde mediados del siglo pasado, y debido a factores predominantemente geopolíticos, económicos y sociales ampliamente documentados por la literatura, el inglés ha experimentado un crecimiento sin precedentes y ha alcanzado el estatus de lengua internacional para la comunicación en distintos dominios<sup>1</sup>, en especial de la ciencia y de la tecnología (Björkman, 2013; Crystal, 2003; Dudley-Evans y St. John, 1998; J. Flowerdew, 2000; Hyland, 2006a; Kaplan, 1993; Mauranen, Pérez-Llantada, y Swales, 2010; Seidlhofer, 2004; Swales, 2004; Tardy, 2004). El advenimiento del inglés como medio de comunicación global<sup>2</sup> ha causado una expansión acelerada de la enseñanza de dicha lengua a nivel mundial (Seidlhofer, 2001). En el nivel superior, el predominio del inglés en el rol de lengua vehicular para el avance y la diseminación del conocimiento y el intercambio académico (Björkman, 2013; Mauranen, 2018; Mauranen et al., 2010; A. Wood, 2001) ha contribuido al desarrollo sostenido del campo de Inglés con Fines Específicos (*English for Specific Purposes* - ESP) y de una de sus ramas, el Inglés con Fines Académicos (*English for Academic Purposes* - EAP) (Björman, 2011; Ferguson, 2007). El área de inglés con fines académicos se enfoca en la enseñanza del inglés con el objetivo de facilitar el estudio o la investigación a estudiantes con necesidades específicas en contextos académicos (J. Flowerdew y Peacock, 2001; Hamp-Lyons, 2011; Hyland y Hamp-Lyons, 2002).

Las universidades en América Latina no han resultado ajenas a la situación arriba descrita (Martínez, 2011; Salager-Meyer, Llopis de Segura, y Ramos, 2016), dado que son ámbitos donde circula una diversidad de textos académicos y profesionales de distinto grado de especialidad escritos en lengua inglesa. Esta realidad hace que los estudiantes deban adquirir y desarrollar, durante su trayecto formativo, las herramientas lingüísticas, discursivas y retóricas en inglés necesarias para acceder y aprehender los géneros y las convenciones retórico-discursivas de la comunidad científica en general y de la disciplina en particular (Hyland, 2004, 2015, 2017). Con el objetivo de facilitar el acceso de los estudiantes de nivel superior a materiales bibliográficos de diferente índole publicados en

---

<sup>1</sup> En su función de medio para la comunicación internacional, el inglés ha recibido diferentes denominaciones: Inglés como lengua internacional (*English as an International Language* - EIL), Inglés mundial (*World English*), Inglés como lengua global (*English as a global language*) e Inglés como lengua franca (*English as a Lingua Franca* - ELF).

<sup>2</sup> Crystal (2003) afirma que una lengua logra el estatus de *lengua global* cuando desarrolla un rol especial que es reconocido por los distintos países.

inglés, la mayoría de las universidades en Latinoamérica ofrecen cursos de lectura comprensiva en dicho idioma, en los cuales se abordan textos relevantes a cada campo disciplinar (Hyland, 2006a). Estos cursos se enmarcan en el área de inglés con fines específicos y están diseñados para preparar a distintos grupos de alumnos en el uso de géneros y habilidades específicas en diferentes disciplinas (Dudley-Evans y St. John, 1998; Hyland, 2006a; Hyland y Hamp-Lyons, 2002; Martínez, 2011).

Los cursos de inglés con fines específicos tienen como objetivo central la enseñanza de los recursos lingüísticos, prácticas discursivas y convenciones comunicativas según los propósitos y las necesidades particulares de comunidades académicas y profesionales determinadas (Hyland, 2002, 2015). La finalidad última es contribuir a que los estudiantes se conviertan en miembros competentes de una comunidad discursiva académica o profesional particular (Hüttner, Smit, y Mehlmauer-Larcher, 2009). La enseñanza de inglés con fines específicos posee las siguientes características distintivas: el análisis de las necesidades de los estudiantes como instrumento indispensable para el diseño curricular, el desarrollo ad hoc de materiales que respondan específicamente a las particularidades de cada situación, el enfoque en géneros y habilidades específicas, y el desarrollo de cursos, generalmente en un tiempo breve, donde se maximiza la instrucción de aspectos específicos (Benesch, 1996; Carkin, 2005; Dudley-Evans y St. John, 1998; L. Flowerdew, 2013; Hutchinson y Waters, 1987). Uno de los rasgos más singulares, que diferencia la enseñanza de inglés con fines específicos de otras situaciones y enfoques de enseñanza, es la relevancia que adquiere el análisis lingüístico, cuyo objetivo es ofrecer descripciones detalladas del uso de la lengua en distintas disciplinas, profesiones u ocupaciones (Basturkmen, 2003, 2010; Bhatia, 1993, 2004, 2008, 2012; Dudley-Evans y St. John, 1998; Johns y Dudley-Evans, 1991; Swales, 1990). Los estudios lingüísticos, principalmente los que se ocupan del análisis de los géneros y otros enfoques de investigación, sirven de base para identificar los aspectos relevantes en la enseñanza de la lengua en distintos contextos, como por ejemplo las disciplinas (Ha y Hyland, 2017; Hyland y Hamp-Lyons, 2002; Johns, 2013).

El concepto de *especificidad* adquiere una importancia medular en los cursos de inglés con fines específicos, aunque distintos autores refieran a diferentes concepciones y grados de especificidad<sup>3</sup>. Referentes en el área como Hyland (2002, 2015, 2017) advierten

---

<sup>3</sup> Algunos autores emplean el término *especificidad* en referencia a los propósitos y contextos de aprendizaje específicos que son propios de cualquier curso de inglés con fines específicos, mientras que otros académicos utilizan el término para señalar la especialización de los contenidos temáticos y el fenómeno de la

que cada disciplina o profesión emplea distintos recursos léxicos, gramaticales, discursivos y retóricos para crear y comunicar conocimiento especializado. Es decir, "el discurso académico no es uniforme ni monolítico" (Hyland, 2017, p. 21). Por lo tanto, los estudiantes tienen que desarrollar un conjunto de prácticas de alfabetización apropiadas, en función del contexto al que se insertan<sup>4</sup>. Especialistas como Hyland (2015, 2017) defienden la centralidad de ofrecer a los estudiantes un enfoque que sea más sensible a la especificidad de las disciplinas. De allí que se resalte la necesidad de estudiar las variaciones lingüísticas y retóricas que exhiben los géneros en los distintos campos disciplinares, dado que estas selecciones particulares reflejan prácticas sociales y epistemológicas muy diferentes (Berkenkotter y Huckin, 1995; Hyland, 2006b; Swales, 2004) y proyectan una identidad o cultura disciplinar (J. Flowerdew, 2011; J. Flowerdew y S. Wang, 2015; Hyland, 2015; Ha y Hyland, 2017). Es precisamente este enfoque el que motiva el presente trabajo.

En nuestro contexto, la Universidad Nacional de Córdoba, se ofrecen desde 1999 cursos de inglés con fines académicos en la mayoría de las carreras de grado en dos modalidades: obligatoria u optativa<sup>5</sup>. Tal como describe Martínez (2011), en relación a la enseñanza de inglés con fines académicos en el contexto latinoamericano, estos cursos se enfocan en general en el desarrollo de una única habilidad, la lectura, ya que tienen como objetivo principal preparar a los estudiantes para leer géneros especializados de circulación frecuente con fines académicos o profesionales. En el caso específico de la Facultad de Ciencias Exactas, Físicas y Naturales (FCEFYN), el Módulo de Idioma Inglés, que contextualiza nuestro estudio, se incorporó como materia curricular en el año 2000 y es dictado por docentes del Departamento de Idiomas con Fines Académicos de la Facultad de Lenguas. El Módulo se centra en el desarrollo de la lectocomprensión en inglés y es una materia obligatoria de régimen cuatrimestral, con cuatro horas semanales distribuidas en dos clases. Los estudiantes que cursan la asignatura, en su mayoría, se encuentran en el tercer año de la carrera de constructor o de alguna de las once ingenierías que forman parte

---

especificidad o variación disciplinar (ver Basturkmen, 2003; Huckin, 2003; Nhã, 2015 para explicaciones detalladas de las distintas concepciones y grados de especificidad).

<sup>4</sup> En la literatura en inglés, se designa con el término *academic literacies* al conjunto de habilidades y competencias (no solo aquellas relacionadas con la lectura y escritura) que deben desarrollar los estudiantes para tener éxito en la vida académica (Hamp-Lyons, 2011).

<sup>5</sup> En el año 1999, el Consejo Superior de la Universidad Nacional de Córdoba aprueba la creación del Módulo de Enseñanza del Idioma Inglés destinado a todos los estudiantes de las carreras de grado de esta casa de altos estudios, y recomienda a las distintas facultades y escuelas la incorporación en forma gradual del Módulo en sus planes de estudio (Busso y Moyetta, 2015).

de la propuesta académica de esta facultad<sup>6</sup> (Busso y Moyetta, 2015). Los grupos de clase son heterogéneos no solo respecto a su conformación, sino también por el dispar dominio de la lengua inglesa que poseen los estudiantes. A lo largo de un cuatrimestre, los alumnos son entrenados en la lectura en inglés principalmente de los siguientes géneros académicos y profesionales: libros de texto universitarios, manuales de instalación, manuales de mantenimiento, normas, y artículos de investigación. Dichos géneros han sido identificados y seleccionados por los docentes de la cátedra a partir de un relevamiento de la bibliografía en inglés incluida en los programas correspondientes a las asignaturas obligatorias de las diferentes carreras.

De los distintos géneros académicos que son funcionales en la educación superior, el *libro de texto universitario*, también llamado *manual universitario*, es uno de los géneros de mayor impacto en la formación universitaria, particularmente a nivel de grado (Biber, 2006; Hyland, 1999; Myers, 1992; Parodi, 2010; Swales, 1995). Este rol especial puede atribuirse a que el libro de texto se encuentra, en palabras de Bondi, "en la intersección de dos órdenes del discurso, educativo y disciplinar" (2016, p. 327).<sup>7</sup> Es precisamente esta característica, la integración del discurso divulgativo-didáctico propio de los géneros pedagógicos y el discurso científico propio de cada disciplina, lo que diferencia a los libros de texto de otros géneros académicos. En la literatura, se lo describe como un texto iniciático, por ofrecer un panorama general de una disciplina (Cubo de Severino, 2007; Parodi, 2008a), que persigue una finalidad claramente didáctica (Johns, 1997; Love, 1991, 1993), al tener como propósito central facilitar el dominio de los conceptos y métodos básicos de un determinado campo disciplinar (Cubo de Severino, 2007; Hyland, 1999, 2000, 2005; Parodi, 2008a, 2010). Se lo caracteriza además como un texto canónico<sup>8</sup> que introduce a los estudiantes no solo en el conocimiento consolidado y consensuado de una disciplina, sino también en las normas, valores, creencias e ideologías de una cultura disciplinar (Hyland, 1999, 2005; Richardson, 2004; Swales, 1990). A pesar de que el libro de texto sea probablemente el género discursivo que los estudiantes leen con mayor frecuencia a lo largo de su etapa de formación, y pese a su papel central en la construcción

---

<sup>6</sup> Además de la carrera de constructor, las once ingenierías que conforman la propuesta académica de la FCEFyN son: ingeniería aeronáutica, ambiental, biomédica, civil, electrónica, en agrimensura, en computación, industrial, mecánica, mecánica electricista, y química.

<sup>7</sup> Esta visión sobre el libro de texto universitario es compartida también por autores como Hyland (2002), Orna-Montesinos (2010, 2012) y Swales (1995) que lo describen como un género híbrido o mixto, que integra el discurso académico y el profesional.

<sup>8</sup> Hyland (2005) afirma que los libros de texto universitarios son concebidos como ejemplares textuales conservadores que albergan paradigmas vigentes y hechos reconocidos y que, de este modo, transmiten una imagen de estabilidad y autoridad.

y comunicación de conocimientos disciplinares (Bondi, 2016), ha sido sometido a un número menor de investigaciones en comparación con otros géneros, tales como el artículo de investigación y el resumen del artículo de investigación (Hyland, 1999; Parodi, 2008a). Al respecto, Carkin (2005) afirma que el libro de texto universitario constituye el punto más débil de aquello que se conoce hasta el momento sobre los géneros académicos escritos. Dada la centralidad y relevancia de este género en la formación de grado, creemos necesaria la realización de estudios que se enfoquen en el análisis pormenorizado de distintos aspectos lingüísticos, para contribuir a obtener una descripción holística del libro de texto y, al mismo tiempo, reflejar sus variaciones o características específicas por campo o cultura disciplinar (Carkin, 2005).

En la lectura de géneros en inglés que versan sobre un área de especialidad, como los libros de texto universitarios publicados en dicha lengua, el vocabulario adquiere un rol fundamental por tratarse de un aspecto determinante para la comprensión de los estudiantes, particularmente en su condición de lectores en una lengua extranjera (Clark e Ishida, 2005; Eskey, 2005; Grabe, 2009; Grabe y Stoller, 2011; Laufer y Ravenhorst-Kalovski, 2010; Nation, 2001a; M. Wang y Koda, 2007). La relación entre vocabulario y lectura es tan estrecha que el primero es considerado uno de los indicadores más importantes al momento de evaluar el grado de dificultad de un texto y el nivel de comprensión alcanzado por los lectores (Hsueh-chao y Nation, 2000; Laufer, 1992; Ma y Lin, 2015; Nation, 2001a, 2006). Numerosos estudios han explorado la relación entre los diferentes componentes (combinados o por separado) que caracterizan al conocimiento del vocabulario y la comprensión lectora de estudiantes de inglés como lengua extranjera (Laufer, 1992; Ma y Lin, 2015; Meara, 1996; Nagy y Herman, 1987; Qian, 1999, 2002; Şen y Kuleli, 2015). Los autores se han concentrado especialmente en medir el conocimiento del vocabulario que los lectores poseen, es decir la estimación del número de palabras que conocen *-breadth of vocabulary knowledge* o *vocabulary size-* y en medir la profundidad o calidad del conocimiento sobre el vocabulario, es decir la estimación del conocimiento acerca de los diferentes aspectos que distinguen a cada palabra (por ej., significado, ortografía, morfología, categoría gramatical, pronunciación) *-depth of vocabulary knowledge*. Otros estudios se han propuesto determinar el umbral léxico (*lexical threshold*), esto es, el número mínimo de palabras (calculado como porcentaje), que los lectores necesitan conocer receptivamente para lograr una comprensión adecuada de determinados géneros (Hirsh y Nation, 1992; Hsu, 2011; Laufer y Ravenhorst-Kalovski, 2010; Nation, 2006).

En contextos de enseñanza de inglés con fines académicos, caracterizados por la especificidad de los géneros y las disciplinas, la relevancia del vocabulario se hace aún más visible para la enseñanza de la lectura. Se ha comprobado que el léxico constituye uno de los rasgos lingüísticos que refleja mayores diferencias entre las distintas áreas disciplinares (Bowker y Pearson, 2002; Cabré, 1999; Ciapuscio y Kuguel, 2002; Coxhead, 2016, 2018; Halliday, 1993, 1998, 2004; Ha y Hyland, 2017; Hyland, 2017; Kuguel, 2006, 2007; Pearson, 1998). En efecto, Carter y McCarthy (2013) y Woodward-Kron (2008) coinciden en señalar que el vocabulario y el conocimiento de una disciplina están íntimamente relacionados, ya que el aprendizaje de los significados de palabras especializadas es parte fundamental del proceso de adquisición de contenidos disciplinares. Es decir, los estudiantes universitarios demuestran una adecuada internalización de conceptos, fenómenos y relaciones entre fenómenos relativos a una esfera disciplinar mediante la comprensión e incorporación de vocabulario especializado (Woodward-Kron, 2008). Por ello, tal como destaca Orna-Montesinos (2010), es importante el análisis del vocabulario específico de una disciplina, dado que los campos disciplinares estructuran el mundo de manera diferente y que los miembros de una determinada comunidad disciplinar comparten los significados y usos del vocabulario característico de esa disciplina. Es por esa razón que el vocabulario despierta especial interés entre distintos profesionales de la lengua, como profesores de inglés con fines específicos, traductores, lexicólogos, lexicógrafos y terminólogos (Cabré, 1999; Sager, 1990, 1998). Consecuentemente, desde hace un tiempo se enfatiza la relevancia de emprender descripciones léxicas basadas en estudios de corpus en dominios y subdominios que den habida cuenta de la especificidad de las disciplinas (Chung y Nation, 2003; J. Flowerdew y Peacock, 2001; Ha y Hyland, 2017; Hyland, 2006a, 2017; Hyland y Tse, 2009; Martínez, Beck, y Panza, 2009; Muñoz, 2015). En lo que refiere a la enseñanza de inglés con fines específicos y académicos, Ha y Hyland (2017) y Hyland y Tse (2007, 2009) sostienen que el diseño de repertorios léxicos por campo disciplinar constituye la dirección hacia la cual deberían encaminarse las futuras investigaciones sobre vocabulario en el área. Es precisamente en esta línea investigativa que esperamos, modestamente, poder aportar con nuestro estudio.

En virtud de todo lo expuesto, y atendiendo a las necesidades particulares de los estudiantes de las distintas carreras de ingeniería en la Universidad Nacional de Córdoba, quienes han inspirado la presente investigación, consideramos apropiado contribuir a este espacio de indagación al realizar un estudio que brinde información lingüística detallada

del vocabulario de libros de texto universitarios publicados en inglés en el campo de la ingeniería.

## **1.2 Estado de la cuestión**

En los próximos apartados se presenta una revisión de estudios lingüísticos de distinta naturaleza en el área de la ingeniería, de investigaciones sobre el léxico en disciplinas y géneros diferentes y de estudios específicos sobre el vocabulario en el dominio de la ingeniería.

### **1.2.1 Estudios lingüísticos en el campo de la ingeniería**

El área de la ingeniería ha cobrado especial interés en los estudios lingüísticos, particularmente en la tradición anglosajona. Las investigaciones abocadas a este campo han explorado distintos aspectos y géneros académicos, tales como los rasgos lingüísticos más destacados en los informes de laboratorio en ingeniería civil (Fakhruddin y Attan, 2012), los mecanismos de cuantificación en presentaciones en conferencias y actas de conferencias en ingeniería hidráulica (Rowley-Jolivet, 2015), los recursos metadiscursivos en los capítulos de Resultados y Discusión de tesis de maestría en ingeniería química, civil, electricista, mecánica y en sistemas (J. Lee y Casal, 2014) y los adverbios de evaluación en tesis de doctorado en ingeniería civil, electricista, mecánica y química (Ahmad y Mehrjooseresht, 2012). Numerosos autores se han enfocado en estudiar el artículo de investigación, centrándose por ejemplo en la estructura retórica de las distintas secciones del artículo de investigación en las subdisciplinas de ingeniería de los materiales, ambiental, electricista, química y en computación (Maswana, Kanamaru, y Tajino, 2015), en los movimientos retóricos de la Introducción, Métodos, Resultados y Discusión en ingeniería en biomédica, civil y en sistemas (Kanoksilapatham, 2015), y en el uso y frecuencia de cláusulas de relativo en artículos de investigación de ciencia e ingeniería (Cho y K. Lee, 2016). Algunos investigadores han comparado diferentes aspectos lingüísticos en el artículo de investigación y otros géneros académicos, como las vallas retóricas en artículos de investigación y tesis doctorales de ingeniería electrónica y química (Koutsantoni, 2006).

### **1.2.2 Estudios sobre el vocabulario en diferentes géneros y disciplinas**

En el marco de la enseñanza de inglés con fines específicos, diversos autores han llevado a cabo estudios léxicos de distintos géneros académicos, con el objetivo de

clasificar los tipos de vocabulario y compilar listas de palabras específicas a cada disciplina para su aplicación pedagógica. En esta línea de investigación, la mayoría de los trabajos se han centrado en el vocabulario académico empleado en el artículo de investigación de distintos dominios (Khani y Tazik, 2013; J. Liu y Han, 2015; Martínez et al., 2009; Moini e Islamizadeh, 2016; Shabani y Tazik, 2014; Valipouri y Nassaji, 2013; Vongpumivitch, Huang, y Chang, 2009; P. Wang, 2017; J. Wang, Liang, y Ge, 2008; Yang, 2015). Por otra parte, existen estudios que han analizado el vocabulario académico utilizado en el artículo de investigación, el artículo de revisión y el libro de texto universitario, este último en calidad de corpus textual de referencia para el proceso de selección de palabras (Lei y D. Liu, 2016). A pesar del extenso cuerpo de estudios antes mencionado, son muy escasos los trabajos que se han enfocado en clasificar los tipos de vocabulario y compilar listas de palabras en otros géneros de la ciencia y géneros académicos escritos en inglés, como por ejemplo el artículo de semi-divulgación (Muñoz, 2015) y el libro de texto universitario (Hajiyeva, 2015; Hsu, 2011; D. Wood y Appel, 2014). Asimismo, se puede apreciar que prácticamente la totalidad de las investigaciones han seleccionado como objeto de análisis al vocabulario académico, no así al vocabulario técnico de una esfera disciplinar en particular.

### **1.2.3 Estudios sobre el vocabulario en libros de texto universitarios de ingeniería**

En la literatura existente se pueden encontrar diversos estudios publicados sobre el vocabulario empleado en el discurso de la ingeniería. Para el presente estudio, revisten particular importancia los trabajos realizados por Hsu (2014), Mudraya (2006) y Ward (2009) sobre el vocabulario de libros de texto universitarios en inglés en diversas subdisciplinas de la ingeniería. Estas investigaciones, basadas en análisis de corpus y metodología cuantitativa, comparten el objetivo central de proporcionar, con una motivación pedagógica, listas de palabras de alta frecuencia, especialmente académicas o subtécnicas<sup>9</sup>, en el campo de la ingeniería.

Hsu (2014) construyó un corpus de libros de texto universitarios de veinte áreas temáticas de la ingeniería con el fin de medir su carga léxica para estudiantes de inglés como lengua extranjera en el nivel superior. Sus resultados demuestran que, para lograr una comprensión aceptable de un libro de texto de ingeniería (95% de cobertura léxica), un

---

<sup>9</sup> Algunos investigadores (M. Baker, 1988; J. Flowerdew, 1993; Trimble, 1985) denominan palabras subtécnicas a aquellas que poseen tanto significados técnicos como no técnicos, como por ejemplo, *iron*, *force* y *strength*.

estudiante cuya primera lengua no es el inglés necesita conocer alrededor de 5.000 familias de palabras<sup>10</sup>. Utilizando la ocurrencia especializada<sup>11</sup>, dispersión uniforme, rango y frecuencia como criterios de selección léxica, Hsu diseñó una lista de 729 familias de palabras de ingeniería a la que denominó *Engineering English Word List* -EEWL. Esta lista proporciona una cobertura aproximada del 14% del total de palabras en el corpus y excluye vocabulario altamente técnico por no ser común a los distintos subcampos de la ingeniería.

Mudraya (2006) compiló un corpus de libros de texto universitarios de uso común en distintas ingenierías, llamado *Student Engineering English Corpus* -SEEC. Mudraya encontró que la mayoría de las palabras en el SEEC correspondían a las categorías de palabras subtécnicas y no técnicas (estas últimas pertenecientes al registro académico). Tomando la frecuencia de ocurrencia como criterio de selección de palabras, este investigador diseñó una lista, a la que rotuló *Student Engineering Word List* -SEWL, con las 1.260 familias de palabras más frecuentes.

Ward (2009) construyó un corpus de libros de texto universitarios de cinco áreas temáticas de la ingeniería y lo llamó *Engineering Corpus* -EC. Este especialista diseñó una lista, a la que denominó *Basic Engineering List* -BEL, que contiene 299 tipos de palabras (*word types*)<sup>12</sup>. En función del propósito principal que guió la investigación (asistir a estudiantes de ingeniería con escaso dominio del inglés), la lista sólo incluye palabras no técnicas, generales y fáciles de aprender en un año, seleccionadas atendiendo a los criterios de frecuencia y distribución. Los resultados muestran que BEL proporciona una cobertura promedio de alrededor del 16,5% del corpus de ingeniería (EC).

Si bien estos trabajos han aportado información descriptiva relevante, se advierte que, por un lado, han puntualizado el estudio de palabras académicas (por ej., *assume*, *example* y *hypothesis*) (Hsu, 2014; Mudraya, 2006; Ward, 2009) y solo en algunos casos de palabras técnicas (por ej., *formula*, *friction* y *spectrum*) (Hsu, 2014); por otro lado, han empleado criterios estrictamente cuantitativos sin considerar aspectos semánticos y

---

<sup>10</sup> La investigación de Hsu reveló que el umbral léxico (*lexical threshold*) para la lectura de libros de texto universitarios varía significativamente en las distintas ramas de la ingeniería. Por ejemplo, ingeniería civil e ingeniería mecánica requieren que los estudiantes conozcan aproximadamente 3.500 familias de palabras para alcanzar una comprensión lectora aceptable, mientras que ingeniería marítima e ingeniería en bioquímica suponen el conocimiento de alrededor de 8.500 familias de palabras.

<sup>11</sup> El criterio de ocurrencia especializada establecía que las familias de palabras incluidas en la EEWL no podían pertenecer a las primeras 2.000 palabras más frecuentes de la *General Service List* -GSL (West, 1953).

<sup>12</sup> Es decir, cada palabra se trata por separado, en lugar de considerarse parte de un lema o familia de palabras (Ward, 2009). Ward decidió adoptar el tipo de palabra como unidad de inclusión en BEL por entender que los estudiantes con un vocabulario muy limitado en inglés tienen muy escaso conocimiento de los mecanismos de formación de palabras en inglés.

pragmáticos relativos al empleo del vocabulario en géneros y disciplinas específicas. Es decir, en dichas investigaciones se han propuesto listas de vocabulario para uso pedagógico construidas a partir de la selección de palabras en base al criterio de frecuencia. Se han desestimado así los significados y usos que adquieren las palabras en géneros disciplinares, como el libro de texto.

Otro trabajo que adquiere especial relevancia en el marco de la presente investigación es el de Todd (2017). Este estudio se centró, al igual que los reseñados en el párrafo anterior, en el análisis de un corpus de libros de texto universitarios de ingeniería para la construcción de una lista de palabras con fines pedagógicos. No obstante, se diferencia de los estudios anteriores en que adopta la opacidad (polisemia) como criterio para la selección de las palabras. Todd define como palabras opacas aquellas que, en un contexto determinado, no adquieren su significado más usual<sup>13</sup> (el que presentan en la lengua general), y considera que son estas palabras las que generan mayores dificultades a los estudiantes ya que, habitualmente, estos no pueden acceder a sus significados de manera autónoma. Todd encontró que las palabras opacas tienden a ser palabras de alta frecuencia, generales o específicas de la disciplina. La lista final que produjo consta de 41 palabras opacas y responde a criterios cuantitativos (análisis de palabras clave) y criterios cualitativos (análisis del grado de opacidad o polisemia de las palabras). Aunque esta lista combina criterios cuantitativos y cualitativos, no contempla aspectos pragmáticos que refieren al uso de las palabras y no ofrece una descripción lingüística detallada de los tipos de vocabulario que caracterizan al género libro de texto universitario.

A pesar de esta diversidad de trabajos empíricos, no existen, hasta lo que conocemos, estudios publicados que hayan realizado una descripción pormenorizada de los distintos tipos de vocabulario que caracterizan a los libros de texto universitarios en el área de la ingeniería escritos en inglés. Asimismo, la revisión bibliográfica evidencia la necesidad de un mayor número de investigaciones basadas en corpus que den cuenta del vocabulario de libros de texto en ingeniería, a fin de compilar listas de palabras que asistan a los estudiantes universitarios hispanohablantes en el proceso de lectocomprensión en inglés. Con el objetivo de contribuir a este campo de investigación, en este estudio nos proponemos, en primer lugar, analizar y caracterizar el vocabulario de libros de texto de ingeniería escritos en inglés a partir del asesoramiento de especialistas en dicho campo y la integración de criterios cuantitativos (frecuencias) y cualitativos (semánticos y

---

<sup>13</sup> Por ejemplo, en el campo de la ingeniería la palabra *value* es más probable que designe un número específico en lugar de activar su significado más habitual de *valor* (Todd, 2017).

pragmáticos), que nos permitan lograr una descripción lingüística más acabada y, en segundo lugar, construir una lista de palabras que responda a las necesidades de los estudiantes en nuestro contexto y que trate sobre una disciplina específica, la ingeniería (Hyland y Tse, 2007, 2009).

### **1.3 Objetivos del estudio**

#### **1.3.1 Objetivos generales**

1.3.1.1 Avanzar en la descripción de géneros académicos de ingeniería escritos en inglés, particularmente de libros de texto universitarios.

1.3.1.2 Analizar y categorizar el vocabulario de un corpus de libros de texto universitarios en el área de la ingeniería.

#### **1.3.2 Objetivos específicos**

1.3.2.1 Diseñar y construir un corpus de libros de texto universitarios en el área de ingeniería escritos en inglés.

1.3.2.2 Determinar el tamaño del vocabulario del corpus.

1.3.2.3 Determinar la relación entre los tipos de palabras y el número total de palabras en el corpus.

1.3.2.4 Determinar el rango de las palabras en el corpus.

1.3.2.5 Identificar y clasificar los distintos tipos de vocabulario que componen el corpus en base a listas de palabras propuestas en la literatura.

1.3.2.6 Determinar la cobertura de los distintos tipos de palabras en el corpus.

1.3.2.7 Identificar las palabras de alta frecuencia.

1.3.2.8 Clasificar las palabras de alta frecuencia en base a listas de palabras propuestas en la literatura.

1.3.2.9 Determinar la cobertura de los distintos tipos de palabras de alta frecuencia.

1.3.2.10 Categorizar las palabras de alta frecuencia según criterios semánticos y pragmáticos.

1.3.2.11 Construir listas de palabras.

#### **1.4 Estructura y organización de la tesis**

Los demás capítulos que conforman la presente tesis responden a la organización que se especifica a continuación. El Capítulo 2 desarrolla los fundamentos teóricos que sustentan este estudio: análisis del género, literatura sobre el vocabulario en inglés y lingüística de corpus. El Capítulo 3 describe la metodología adoptada, detallando los criterios utilizados para la construcción del corpus, los materiales empleados para el análisis del corpus y los procedimientos seguidos para la recolección y el análisis de los datos. El Capítulo 4 informa sobre los resultados obtenidos. Por último, el Capítulo 5 discute los principales hallazgos en relación con la literatura reseñada y el marco conceptual en que se inscribe este trabajo, analiza aplicaciones pedagógicas del estudio, identifica limitaciones de la investigación, plantea futuras líneas de investigación, y esboza algunas conclusiones finales.

## **Capítulo 2**

### **Fundamentos teóricos**

Este capítulo expone el marco teórico que sustenta la presente investigación. En la sección 2.1 se presenta la teoría de Análisis de los Géneros a partir de una revisión y comparación de tres escuelas o tradiciones académicas. En la sección 2.2 se define y caracteriza el género objeto de este estudio, el libro de texto universitario. La sección 2.3, por su parte, está centrada en explorar la literatura existente sobre la naturaleza y tipos de vocabulario, especialmente aquellos trabajos enfocados en el estudio del léxico en inglés. Por último, en la sección 2.4 se describe la Lingüística de Corpus y los lineamientos teórico-metodológicos que propone para la construcción y análisis de un corpus especializado.

#### **2.1 Teoría de los Géneros**

El presente estudio se vale de distintas nociones y aportes provenientes de la Teoría de Análisis de los Géneros, enfoque multidisciplinario desarrollado a partir de estudios en lingüística, retórica, teoría literaria, análisis crítico del discurso, sociología y psicología (Bawarshi y Reiff, 2010; Bhatia, 2002), que tiene como objetivo central examinar la relación que se establece entre la lengua y el contexto en textos escritos y orales (Hyland, 2002; Swales, 1990). Como su nombre lo indica, este enfoque se basa en el concepto de *género* que fuera acuñado por Bakhtin en el año 1920 y reformulado sucesivamente por numerosos académicos de diferentes disciplinas, tales como la literatura, la retórica y la lingüística (Swales, 1990).

##### **2.1.1 Definición de género**

Dentro del campo de la lingüística, se concibe en líneas generales a los géneros como textos con convenciones retórico-discursivas y léxico-gramaticales particulares que resultan de actividades sociales recurrentes y rutinarias llevadas a cabo en contextos culturales específicos. El género como construcción teórica posee el valor de ofrecer una teoría de la lengua con un fundamento social (Hyland, 2002) y propicia una mayor comprensión acerca de cómo y por qué los miembros de una comunidad dada producen determinadas clases de discurso (Freedman y Medway, 1994). Los estudios del género han contribuido así a la descripción de textos académicos y profesionales en diversas

disciplinas y a su correspondiente aplicación en contextos educativos<sup>14</sup> en primeras lenguas y en lenguas extranjeras (Bhatia, 2001, 2012; Hüttner et al., 2009; Hyon, 1996; Swales, 1990).

### 2.1.2 Tradiciones académicas

Desde sus inicios hace alrededor de cuatro décadas, la teoría del género se ha desarrollado de manera paralela en tres escuelas o tradiciones: la Nueva Retórica, la Escuela Australiana o Escuela de Sydney y el Inglés con Fines Específicos. Las tres escuelas han surgido en países diferentes, en respuesta a distintas necesidades, realidades y motivaciones; por lo tanto, difieren en varios aspectos. Estas divergencias han sido señaladas y reseñadas, entre otros autores, por Freedman y Medway (1994), Hyland (2002), Hyon (1996) y Johns (2002). A continuación, se ofrece una síntesis de las principales diferencias que se pueden advertir entre las tres tradiciones en relación con los aspectos que cada escuela privilegia, la metodología de análisis generalmente aplicada en cada orientación académica, los contextos donde se aplican los resultados obtenidos a partir de los estudios de género, y las conceptualizaciones del concepto *género* desarrollada por cada tradición académica.

La Nueva Retórica se origina en los Estados Unidos y cuenta entre sus principales representantes a investigadores tales como Bazerman (1988, 1994), Berkenkotter y Huckin (1995), Coe (1994, 2002), Freedman y Medway (1994) y Miller (1984, 1994). Esta tradición académica pone el énfasis en los contextos situacionales donde se emplean los géneros y en los propósitos sociales que estos persiguen. En consonancia con la importancia asignada a los aspectos contextuales y funcionales de los géneros, se privilegia el uso de técnicas etnográficas para el análisis de los textos. Los representantes de esta escuela consideran que el estudio de los géneros contribuye a mejorar la enseñanza de la escritura académica y profesional de estudiantes universitarios y profesionales novatos o principiantes cuya primera lengua es el inglés.

La Escuela Australiana, también denominada Escuela de Sydney, nace en Australia como una aplicación de la teoría del lenguaje elaborada por Halliday, y se desarrolla a partir de las contribuciones de referentes como Christie y Martin (1997), Eggins y Martin (2000), Feez (2002), Halliday y Hasan (1989), Martin y Rose (2008), entre otros. Esta

---

<sup>14</sup> Vale aclarar que la evidencia proporcionada por los estudios de género derivan en aplicaciones en ámbitos académicos y no académicos, es decir en comunidades discursivas que no se reducen exclusivamente a las disciplinares.

escuela se focaliza en los rasgos textuales de los géneros con el fin de explicar la relación existente entre la lengua y sus funciones en contextos sociales. El análisis de los géneros se aplica a la enseñanza en instituciones primarias y secundarias de Australia, como así también a la educación de inmigrantes adultos en dicho país. En este enfoque, el objetivo central de una pedagogía basada en el género es familiarizar a los estudiantes con los recursos lingüísticos necesarios para un desempeño exitoso en contextos académicos o laborales.

La tradición de Inglés con Fines Específicos, donde nos posicionamos en el presente estudio, surge en Inglaterra y tiene entre sus académicos más reconocidos a investigadores como Bhatia (1993, 2001, 2002, 2004, 2008, 2012), Dudley-Evans (1994), Dudley-Evans y St. John (1998), Hyland (2000, 2004, 2006a), Hyon (2002), Johns (2002) y Swales (1990, 2004). En este enfoque, que ha sido influenciado por los dos ya mencionados, el análisis de los géneros se centra en la descripción lingüística de los textos y pone especial énfasis en los elementos microtextuales así como en la macro organización de los textos. En esta escuela, se pretende familiarizar a estudiantes cuya primera lengua no es el inglés con las convenciones lingüísticas y funciones sociales de los géneros que necesitan leer y escribir en contextos tanto académicos como profesionales.

Por otra parte, cabe destacar que las tres tradiciones manifiestan diferencias en su conceptualización del constructo *género*. Estas radican esencialmente en el rol preponderante que una y otra escuela le asigna a la lengua y a la estructura del texto o al contexto y a la comunidad (Johns, 2002). La Nueva Retórica se concentra menos en los aspectos textuales (léxico-gramática, estructura retórica) y resalta, en cambio, el contexto situacional donde los géneros tienen lugar (Adam y Artemeva, 2002; Coe, 2002; Miller, 1984). Este enfoque concibe al género como una estrategia o herramienta social que adopta una forma típica de discurso y que se desarrolla en respuesta a un tipo recurrente de situación retórica (Coe y Freedman, 1998; Miller, 1984). Además, esta tradición defiende el carácter dinámico, en permanente evolución, del género como forma de acción condicionada por cambios en el contexto de situación (Miller, 1984).

Por su parte, la Escuela Australiana destaca la centralidad de los propósitos sociales de los géneros y la importancia de describir los rasgos léxico-gramaticales y las estructuras retóricas que han desarrollado los géneros para cumplir esos propósitos. En este enfoque, se define al género como un proceso social, orientado a alcanzar una meta y organizado en etapas (Martin, 1992, 1997); es decir, desde esta conceptualización, el género tiene un carácter interactivo (social), intencionado (movilizado por un propósito) y secuencial

(conformado por etapas). Es necesario aclarar que, a diferencia de la Nueva Retórica y la Escuela de Inglés con Fines Específicos, en las cuales los géneros son concebidos como eventos sociales culturalmente determinados y reconocidos por un nombre (por ej., libro de texto, artículo de investigación y reseña, en la comunidad académica), en la Escuela Australiana los géneros son entendidos como patrones retóricos, tales como narraciones y descripciones, que se podrían asociar someramente a las denominadas *secuencias* (Adam, 1992; Werlich, 1975, como se citan en Ciapuscio, 2003a) y *tipos de texto* (Paltridge, 1996, 2001), nociones desarrolladas por otros enfoques teóricos.

La tradición del Inglés con Fines Específicos pone el énfasis en las convenciones lingüísticas y en la estructura de los géneros. Los géneros son concebidos como eventos comunicativos reconocidos socialmente y utilizados por comunidades discursivas específicas, cuyos miembros comparten propósitos comunicativos culturalmente definidos (Bhatia, 1993, 2004; Swales, 1990, 2004). Dichos propósitos determinan tanto la estructura global del género como sus elecciones de contenido y estilo.

A pesar de las diferencias señaladas, varios autores han reconocido algunas coincidencias entre las tres tradiciones teóricas (Bhatia, 2004; Freedman y Medway, 1994; Hyland, 2004; Johns, 2002; Swales, 2009). En primer lugar, las tres tradiciones parten de una concepción del lenguaje como una actividad comunicativa socialmente situada, por lo tanto, acuerdan en la necesidad de entender los géneros como producto y medio para las interacciones entre miembros de una sociedad. En segundo término, los tres enfoques reconocen la existencia de una relación intrínseca entre lengua y contexto, donde la lengua responde a los contextos (sociales, culturales, situacionales e institucionales) y los contextos dan forma a la lengua, por ende, coinciden en considerar al contexto como un marco imprescindible para entender los géneros. Por último, las escuelas consideran que los géneros sirven propósitos comunicativos reconocidos por los miembros de la comunidad discursiva en que circulan. De este modo, la identificación de los propósitos comunicativos de los textos se convierte, en las tres orientaciones, en un concepto metodológico importante para el análisis de los géneros.

## **2.2 El género libro de texto universitario**

### **2.2.1 Contextualización del género**

A lo largo de la historia de la ciencia y la tecnología, han surgido una diversidad de géneros académicos y profesionales que vehiculizan conocimiento especializado y

cumplen propósitos comunicativos específicos. Dentro de los géneros académicos, el *libro de texto* desempeña un papel central en la transmisión de saberes disciplinares desde su popularización a comienzos del siglo XIX. El filósofo e historiador de la ciencia Thomas Kuhn (1971) afirma que es principalmente a través de la lectura de libros de texto, en su condición de instrumentos pedagógicos perpetuadores de la ciencia, que cada nueva generación de científicos aprende su profesión. Más aún, Kuhn sostiene que este género, al presentar un conjunto de conocimientos y teorías aceptados por la comunidad científica, refuerza paradigmas científicos predominantes y recrea una imagen de la ciencia en un momento histórico dado.

Específicamente en el ámbito de la educación superior, la interacción que el libro de texto universitario entabla con otros géneros del sistema de géneros académicos, en especial con el artículo de investigación, ha sido objeto de indagación por autores como Bazerman (1988), Hyland (2005) y Love (2001, 2013), entre otros. A diferencia de los artículos de investigación que comunican hallazgos realizados por investigadores y puestos a consideración de sus pares (es decir, conocimiento en proceso de construcción y legitimación o validación), los libros de texto universitarios transmiten conocimiento ya consolidado y consensuado entre los miembros expertos de una comunidad disciplinar. De allí que el género artículo de investigación persiga, además del propósito de difundir conocimientos nuevos, el de argumentar y persuadir a la comunidad de pares; mientras que el libro de texto universitario tiene una clara intencionalidad didáctica al servicio de diseminar conocimiento establecido. Hyland (1999), Myers (1992) y Swales (1995) señalan que estas diferencias podrían explicar por qué el artículo de investigación es considerado el género académico más prestigioso y representativo de la labor científica y, en cambio, el libro de texto universitario es asignado muchas veces un rol más periférico, incluso marginal. Sin embargo, el libro de texto universitario es el género por excelencia que proporciona a los autores el medio para difundir una visión general de un campo disciplinar (Hyland, 2000, 2005; Swales, 1995). Además, es el género que desempeña un papel clave en el entendimiento por parte de los estudiantes universitarios de un área de especialidad (Hyland, 2005), dado que les ayuda a formar una representación mental del conjunto de conocimientos propios de una esfera disciplinar (Bondi, 2016).

### **2.2.2 Definición**

Como se mencionara en la Introducción (ver sección 1.1), el *libro de texto universitario* o *manual universitario* es uno de los géneros académicos de mayor impacto

en la formación universitaria, particularmente a nivel de grado (Biber, 2006; Hyland, 1999; Myers, 1992; Parodi, 2010; Swales, 1995). Este género ha recibido diversas denominaciones en la literatura en inglés: *coursebook* (Hyland, 1999), *undergraduate coursebook* (Hyland, 2005), *undergraduate textbook* (Hyland, 2002), *introductory textbook* (Love, 1991, 1993; Richardson, 2004) y *university textbook* (Moore, 2002).

En consonancia con la noción de libro de texto científico propuesta por Kuhn (1971), los libros de texto universitarios son definidos de manera general como repositorios de los hechos, teorías y tradiciones de una disciplina (Hyland, 2005) y como compendios de información establecida y aceptada (Parkinson y Adendorff, 2004). Se los concibe como textos de carácter iniciático que presentan contenidos secuenciados y jerarquizados cuidadosamente, con el objetivo de ofrecer un panorama general de una disciplina (Cubo de Severino, 2007) en un mapa epistemológico ordenado de manera coherente (Hyland, 1999). Se aprecia en estos textos una manifiesta finalidad didáctica, ya que tienen como propósito central facilitar el dominio de los conceptos y métodos básicos de un determinado campo disciplinar (Cubo de Severino, 2007; Hyland, 1999, 2005; Parodi, 2008a, 2010).

### 2.2.3 Caracterización

Como es bien sabido, un aspecto fundamental en la descripción de cualquier género está dado por el rol social de los participantes y el tipo de interacción que entablan. En el libro de texto universitario, el destinador y el destinatario se encuentran posicionados prácticamente en los dos extremos de la escala de experticia, ya que existe una disparidad muy pronunciada en la competencia disciplinar por parte de uno y otro (Cubo de Severino, 2007; Kuguel, 2009). El autor de un libro de texto universitario es generalmente un profesor experto en una materia dada, que escribe con la finalidad de informar e instruir a estudiantes de la disciplina en cuestión. El escritor se presenta a sí mismo como una voz única de autoridad en un esfuerzo por poner el acento en la transmisión de conocimientos ya reconocidos y aceptados, en lugar de enfatizar la multiplicidad de sus fuentes, es decir, en un intento por "reducir la cacofonía de textos pasados" (Hyland, 2005, p. 101). En cuanto a los destinatarios de los libros de texto, son alumnos universitarios considerados semilegos o aprendices de especialistas por la etapa de formación académica en que se hallan. La noción de *semilego* es una categoría que designa a aquellos interlocutores que poseen ciertos conocimientos sistematizados sobre un área específica y que comprende distintos perfiles, entre ellos al estudiante del nivel superior (Ciapuscio, 2003a; Ciapuscio y

Kuguel, 2002; Kuguel, 2007, 2009; Pearson, 1998). Por lo expuesto, la relación que se establece en el libro de texto universitario entre escritor y lector es de tipo asimétrica, basada en el mayor conocimiento que posee el primero sobre una esfera disciplinar.

En relación al contexto de circulación de este género, determinado por su propósito pedagógico y audiencia, el libro de texto universitario se emplea en el ámbito de la educación superior. Se ha demostrado que este género es el más utilizado en la formación de estudiantes universitarios de diversas disciplinas (Parodi, 2008a, 2010, 2014). Los alumnos están expuestos de manera cotidiana a los libros de texto como material de lectura obligatoria, como material de consulta y como fuente de ejercicios y problemas para resolver (Granés y Caicedo, 1997, como se cita en Cabrera Castillo, 2012).

La estructura retórica general del libro de texto o manual universitario ha sido descrita con profundidad en la lengua española por Parodi (2008a, 2010, 2014). El modelo retórico elaborado por Parodi se basa en la *macromovida*, concebida como una unidad discursiva de mayor jerarquía que una *movida* (unidad del discurso con un propósito comunicativo particular que contribuye al propósito comunicativo global del género). El concepto de macromovida posibilita dar cuenta de la extensión de los textos que integran este género y el carácter recursivo de ciertas secciones obligatorias, como así también lograr un mayor nivel de abstracción en el análisis (Parodi, 2008a).

Parodi (2008a, 2010, 2014) identifica tres macromovidas en el género libro de texto universitario: preámbulo, conceptualización y ejercitación, y corolario; cada una de ellas se compone de un conjunto de movidas y pasos que pueden ser obligatorios u opcionales. El *preámbulo* se localiza en la primera parte del libro, donde se lo presenta y se ofrece información útil para guiar su lectura. Comprende secciones que revelan una preocupación especial por los destinatarios del texto, tales como un prólogo o prefacio, un índice temático o tabla de contenidos, un índice o tabla de símbolos y abreviaturas, y una introducción. La macromovida *conceptualización y ejercitación* constituye el corazón del libro de texto y su función central es brindar conceptos y definiciones, acompañados de una puesta en práctica y de recapitulaciones a modo de cierre. Esta macromovida se repite de manera cíclica a lo largo de los distintos capítulos que componen un libro de texto y conforma de este modo una estructura espiralada. El *corolario* se ubica en la parte final del libro y su función principal es complementar y profundizar los contenidos abordados mediante soluciones y respuestas a ejercicios y problemas, especificaciones en anexos, apéndices y glosarios, y guías en índices y referencias. El carácter espiralado y recursivo

de la macromovida central lleva a Parodi a denominar la estructura retórica del libro de texto universitario como una *colonia encadenada*.

En la actualidad, varios autores sostienen que los libros de texto universitarios constituyen un género de naturaleza mixta o híbrida por articular el discurso académico y el pedagógico o de divulgación (Bondi, 2016; Cubo de Severino, 2007; Hyland, 2000). El primero responde al propósito de este género de vehicular conocimiento especializado en el nivel superior y el segundo obedece a su finalidad didáctica de facilitar el acceso de estudiantes a un campo del saber y la consiguiente incorporación a la comunidad discursiva de una disciplina.

Dada su naturaleza mixta, el género libro de texto universitario exhibe en la superficie textual rasgos típicos del discurso académico y otros propios del discurso pedagógico. El uso extenso de nominalizaciones (Love, 1993; Myers, 1992), la utilización frecuente de sintagmas nominales complejos (Biber, 2006), la explotación de formas pasivas (Biber y Gray, 2016; Neilson, 2016), la escasa presencia de mitigadores o atenuadores (Hyland, 1994, 1999; Myers, 1992) y la abundancia de definiciones y clasificaciones (Cubo de Severino, 2007; Trimble, 1985) son elementos representativos del registro académico. Mientras que el empleo sostenido de preguntas retóricas como orientadoras del discurso (Bhatia, 1984, como se cita en Bhatia, 2013; Trimble, 1985), la explotación de las primeras y segundas personas para el establecimiento de un diálogo entre autor y lector (Kuguel, 2007), la incorporación de recursos metatextuales en pos de la simplificación del discurso (Hyland, 1999, 2005; Williams, 1985), la marcada presencia de elementos gráficos y visuales (multimodalidad o multisemiosis) en favor de la comprensión de cuestiones complejas (Boudon y Parodi, 2014; Kuguel, 2007; Parodi, 2008a, 2010) y la apelación a recursos de tratamiento como la explicación y la ejemplificación y a procedimientos de reformulación como las paráfrasis y las repeticiones (Bhatia, 1993; Ciapuscio, 1998, 2003b; Kuguel, 2007, 2009) son instanciaciones distintivas de un discurso con una clara perspectiva didáctica.

Por último, haciendo eco del creciente interés por estudiar la influencia de las culturas disciplinares sobre los géneros, se ha comenzado a explorar el impacto de la variación disciplinar en los libros de texto universitarios, especialmente en la naturaleza de sus contenidos, prácticas discursivas, realizaciones gramaticales, léxico especializado y enfoques de enseñanza adoptados (Bhatia, 2004; Bondi, 1999; Hyland, 1999). En el caso particular de los libros de texto universitarios pertenecientes al campo de la ingeniería, se puede advertir el carácter pragmático y aplicado de esta disciplina, y su propósito de

transferir el conocimiento abstracto en técnicas, métodos, dispositivos y productos que se ponen al servicio de la comunidad (Orna-Montesinos, 2012). Dado que en la ingeniería este género se centra en discusiones metodológicas y procesos que involucran principalmente entidades concretas, se observa un uso mucho más extendido que en otras esferas disciplinares de verbos en tiempos que indican presente y de estructuras pasivas que tienen como finalidad destacar el objeto que está siendo manipulado y opacar al agente que realiza la acción (Biber y Gray, 2016).

### 2.3 Vocabulario

El vocabulario es reconocido como uno de los aspectos más distintivos del discurso científico-técnico y, al mismo tiempo, como uno de los rasgos lingüísticos que exhibe mayor especificidad disciplinar (Bowker y Pearson, 2002; Cabré, 1999; Ciapuscio y Kuguel, 2002; Halliday, 1993, 1998, 2004). Mayormente, el vocabulario científico-técnico en inglés ha sido descrito, a los fines de su enseñanza y aprendizaje, a partir de la siguiente clasificación de las palabras: palabras generales o de alta frecuencia (West, 1953), palabras académicas (Coxhead, 2000), palabras de baja frecuencia, palabras no técnicas, y palabras técnicas o términos. El presente estudio se basó en este esquema de categorización de las palabras para analizar el vocabulario de libros de texto universitarios de ingeniería publicados en inglés. Por su centralidad en esta investigación, las palabras técnicas o términos son definidas y tratadas en una sección aparte.

Las *palabras generales*, también llamadas de alta frecuencia, se definen como palabras no especializadas que ocurren de manera frecuente en una variedad de usos de la lengua (Coxhead y Nation, 2001; Nation, 2001b, 2008). West (1953) se refirió a este grupo de palabras como vocabulario de servicio general<sup>15</sup> ya que son de servicio, en el sentido de utilidad, independientemente de para qué se emplea la lengua en un contexto y momento dado (Chung y Nation, 2003). Estas palabras representan aproximadamente 2.000 familias de palabras que incluyen en su mayoría palabras de contenido (nombres, adjetivos, verbos y adverbios) y casi la totalidad de las palabras gramaticales en inglés (alrededor de 176 familias de palabras) (Nation, 2001a), tales como *body*, *good*, *catch* y *about*. Las palabras generales son un grupo relativamente reducido de palabras que cubren, generalmente, el 80% de las palabras en un texto. Las 100 palabras más frecuentes de las palabras generales proporcionan una cobertura de alrededor del 50% de las palabras de un texto (Nation,

---

<sup>15</sup> De allí la designación de la lista de palabras creada por West (1953) como *General Service List* (GSL) (cfr. Capítulo 3, sección 3.1.3).

2005). La importancia de estas palabras reside así en su alta probabilidad de aparición en un rango muy diverso de usos de la lengua (Nation, 2001b).

Las *palabras académicas* (también denominadas desde otros enfoques palabras subtécnicas, palabras semitécnicas, vocabulario científico generalmente útil o léxico especializado no técnico<sup>16</sup>) son aquellas palabras que ocurren de manera relativamente frecuente en un amplio repertorio de textos académicos en distintas disciplinas, pero que no están comprendidas dentro de las palabras generales ni son específicas de una disciplina en particular (Nation, 2001a, 2001b). Son ejemplos de palabras académicas *assess*, *data*, *empirical* y *hence*. Tal como aclaran Coxhead y Nation (2001), estas palabras también aparecen en otros tipos de textos, no académicos, pero con una frecuencia visiblemente menor. En su mayoría, las palabras académicas provienen etimológicamente del latín, del griego o del francés, y se caracterizan por ser multisilábicas, admitir mecanismos de afijación (prefijos y sufijos) y poseer algún grado de formalidad (Nation, 2005). Generalmente, las palabras académicas se utilizan para desarrollar un argumento, organizar el discurso o la información, referirse a obras de otros autores, describir actividades científicas y trabajar con datos (P. Meyer, 1990, como se cita en Nation, 2001a). El vocabulario académico proporciona usualmente una cobertura de entre el 8,5-10% de las palabras en textos académicos. Si se computan las palabras académicas y las palabras generales, se obtiene una cobertura total de aproximadamente el 90% de las palabras en un texto académico (Coxhead y Nation, 2001; Nation, 2001b).

Las *palabras de baja frecuencia* son aquellas que tienen una frecuencia de aparición muy baja y un rango<sup>17</sup> muy limitado (Coxhead y Nation, 2001). Suelen aparecer una o dos veces en un texto o tener una frecuencia y rango moderados, debido a esto quedan excluidas del grupo de palabras de alta frecuencia. No obstante, estas palabras conforman el grupo más numeroso de palabras en inglés, que comprende palabras técnicas de diferentes áreas disciplinares, nombres propios y todas las palabras que no son de alta frecuencia o académicas (Nation, 2001a). Nation (2008) señala que el vocabulario de baja frecuencia en inglés está integrado por más de 100.000 familias de palabras y que las palabras de baja frecuencia, típicamente, cubren alrededor del 10% de las palabras de un texto académico.

Las *palabras no técnicas* en textos especializados han sido definidas por P. Meyer (1997) como las palabras utilizadas para construir la estructura semántico-pragmática de

---

<sup>16</sup> (Ver Chung y Nation, 2003; Hyland y Tse, 2007; Nation, 2001a).

<sup>17</sup> El concepto de rango se define en este capítulo en el apartado 2.4.5.1.

un texto. El autor observa que las palabras no técnicas representan una porción importante del vocabulario de textos en contextos especializados y que, por lo general, se emplean para hacer referencia a entidades, propiedades y procesos que no son específicos de un campo disciplinar. Tal como advierte P. Meyer, si bien las palabras no técnicas denotan conceptos que están de algún modo relacionados con una disciplina determinada, las referencias que establecen son de tipo más general y pueden ser asociadas con una diversidad de contextos especializados. P. Meyer propone clasificar las palabras no técnicas en cinco categorías: tiempo, aspecto y modalidad (por ej., *currently*, *prolonged* y *can*), elementos de la temática descripta (por ej., *increase*, *large* y *problem*), relaciones entre entidades (por ej., *although*, *involve* e *interchange*), elementos de la práctica académica (por ej., *applicable*, *measure* y *experimentation*), y el dominio del texto (por ej., *above*, *foregoing* y *show*). Esta propuesta de categorización sustentará la clasificación de las palabras no técnicas de alta frecuencia relevadas en el corpus construido para la presente investigación.

### **2.3.1 Palabras técnicas o términos**

#### **2.3.1.1 Definición de palabras técnicas o términos**

Las *palabras técnicas* o *términos*, también denominadas unidades terminológicas, términos técnicos, léxico especializado y vocabulario especializado<sup>18</sup>, han sido abordadas y conceptualizadas en la literatura desde diferentes perspectivas, dos de las cuales adquieren una importancia central en función de los objetivos de esta investigación: la perspectiva de la lingüística aplicada (Chung y Nation, 2003, 2004; Coxhead y Nation, 2001; Nation, 2001a, 2001b, 2008, entre otros) y la perspectiva de la terminología (Bowker y Pearson, 2002; Cabré, 1999, 2003; Ciapuscio, 2003a; Pearson, 1998; Sager, 1990, 1998; Temmerman, 2000, entre otros). Estos dos enfoques proporcionan definiciones de las palabras técnicas o términos y proponen distintos procedimientos y técnicas para su reconocimiento en los textos.

Desde la perspectiva de la lingüística aplicada, se han propuesto distintas definiciones de las palabras técnicas. Chung y Nation (2004) las definen como palabras que son muy frecuentes en un dominio especializado pero que son menos frecuentes o inexistentes en otros campos. Nation (2008) postula que las palabras técnicas tienen cuatro

---

<sup>18</sup> (Ver Chung y Nation, 2004).

características distintivas que determinan su naturaleza y proporcionan las bases para su identificación, a saber: 1) están relacionadas con un área temática; 2) ocurren en un dominio especializado; 3) son partes del sistema de conocimientos de una disciplina; y 4) pueden aparecer en dominios técnicos que están relacionados con otros dominios o que guardan alguna relación con actividades de la vida cotidiana. Las palabras técnicas comprenden: palabras asociadas con una disciplina en particular, siglas, acrónimos, abreviaturas, fórmulas químicas, símbolos y palabras de baja frecuencia tales como nombres propios cuyos significados están estrechamente vinculados con un tema específico (Nation, 2001a). Se estima que el vocabulario técnico de una disciplina puede estar integrado por 1.000 a 5.000 palabras (Nation, 2008). En un estudio sobre el vocabulario técnico de dos disciplinas, Chung y Nation (2003) afirman que las palabras técnicas constituyen una proporción importante de las palabras de un texto especializado, que supera ampliamente el 5% sugerido anteriormente por Nation (2001a). Estos autores sostienen que al menos el 20% de las palabras en la mayoría de los textos especializados son palabras técnicas.

Desde la perspectiva de la terminología, tradicionalmente se han definido y caracterizado los términos en oposición a las palabras<sup>19</sup>. Wüster, fundador de la llamada Teoría General de la Terminología (TGT) en los años 1960, concebía al *término* (objeto de estudio de la terminología) como la forma de una unidad de designación, es decir como la denominación de un concepto perteneciente a un dominio de especialidad<sup>20</sup>, mientras que entendía a la *palabra* (objeto de estudio de la lingüística) como la forma de un signo lingüístico de la lengua general (Adelstein, 2007). Desde una postura crítica al enfoque prescriptivista de la TGT y su concepción idealizada del término, Cabré, creadora de la Teoría Comunicativa de la Terminología (TCT), concibe a las unidades terminológicas o términos<sup>21</sup> como unidades léxicas que forman parte del lenguaje natural y de la gramática de una lengua. Las unidades léxicas no son inicialmente ni palabras ni términos sino

---

<sup>19</sup> El estatuto semiótico (signo primario, forma de un signo, y signo secundario) que distintos enfoques teóricos en el seno de la terminología y de la lingüística le asignan al término y a la palabra determina la conceptualización y consiguiente caracterización que se hacen de una y otra entidad (ver Capítulo 1 de Adelstein (2007) para una revisión pormenorizada de la oposición término/palabra según diversas aproximaciones teóricas).

<sup>20</sup> Dentro de esta línea tradicional de la terminología, un término, para ser considerado como tal, debía reunir los siguientes atributos: univocidad (monosemia, mononimia), carácter esencialmente nominal, pertenencia a un único dominio o disciplina, definición dentro del sistema conceptual de un determinado dominio o disciplina, ausencia de valor estilístico y de connotaciones (Adelstein, 2007; Adelstein y Cabré, 2002).

<sup>21</sup> Si bien es posible realizar una distinción entre *unidad terminológica* y *término* (donde término alude solo al aspecto formal de la unidad terminológica), en la presente investigación *término* designa a una unidad de forma y contenido, y funciona entonces como sinónimo de *unidad terminológica* (Adelstein, 2007).

potencialmente términos o no términos, dado que su valor especializado se activa<sup>22</sup> en función de su uso en un contexto y una situación determinada (Cabré, 1999, 2000). En otras palabras, la TCT no considera que existan diferencias ontológicas entre términos y palabras, sino que defiende la existencia de unidades léxicas que potencialmente pueden actualizarse como términos o palabras, de acuerdo con distintos factores sintácticos, textuales, pragmáticos, situacionales y sociales (Adelstein, 2007, 2009; Adelstein y Cabré, 2002, 2003; Cabré, Feliu y Tebé, 2001; Cabré y Estopà, 2005). Este cambio sustancial de perspectiva en la terminología postula analizar los términos en su condición de unidades multidimensionales<sup>23</sup> y dinámicas de representación y comunicación del conocimiento especializado (Cabré, 2002a, 2002c) y estudiar su empleo en contextos reales de aparición, es decir, en los textos especializados (análisis *in vivo* en lugar de análisis *in vitro*) (Cabré, 2002a, Cabré y Estopà, 2002; Ciapuscio, 2003a). De todo lo expuesto se desprende la necesidad de llevar a cabo investigaciones basadas en corpus que permitan obtener evidencia empírica acerca del uso del vocabulario en géneros y disciplinas específicas.

La multiplicidad de sentidos que pueden activar ciertas palabras se explica en función de dos propiedades semánticas: la polisemia y la homonimia. Estos dos conceptos, que se utilizan para describir esencialmente palabras que tienen más de un significado en contextos distintos (Cruse, 2000; McCarthy, 1990), guiaron la clasificación de las palabras técnicas y no técnicas en el presente estudio. Como señala Cabré (1999), la polisemia constituye un aspecto fundamental del significado de las palabras y es uno de los mecanismos más productivos para extender el vocabulario de una lengua. La mayoría de las palabras son polisémicas, a excepción de los términos altamente especializados que generalmente son monosémicos con un solo significado, referente y función (Carter, 1987). Una palabra polisémica es una palabra que posee varios sentidos distintos pero que están relacionados semánticamente y asociados al potencial de significado de la palabra (Cruse, 2006; McCarthy, 1990). Todos estos sentidos comparten un núcleo de significado o sentido básico (Cruse, 2000, 2006; Geeraerts, 1990). Adelstein y Cabré (2002) postulan la necesidad de reconocer, en las descripciones del léxico de los dominios de especialidad,

---

<sup>22</sup> Como señalan Adelstein y Cabré (2002), el concepto de *activación* de valores especializados o no especializados evidencia el enfoque cognitivo de la Teoría Comunicativa de la Terminología. Se presupone que la información semántica (organizada en rasgos, módulos de rasgos o paquetes de información) correspondiente a cada sentido de una unidad léxica es seleccionada a través de un proceso cognitivo.

<sup>23</sup> En la TCT los términos son definidos como objetos *poliédricos* ya que son a la vez unidades de conocimiento, unidades de comunicación y unidades lingüísticas (Cabré, 2002b, 2003, 2008, 2009).

distintos tipos o procesos de polisemia tales como *polisemia interdominio*<sup>24</sup>, *polisemia intradominio*<sup>25</sup>, *terminologización* y *desterminologización*<sup>26</sup>, entre otros. Por su parte, el fenómeno de la homonimia se produce cuando una misma forma lingüística representa conceptos totalmente diferentes y significados lexicalizados que no están relacionados semánticamente de modo alguno, y por lo tanto se tratan como palabras diferentes (Cabré, 1999; Cruse, 2000, 2006; Kostina, 2011). Es necesario aclarar que la homonimia de la terminología tradicional (TGT) se trata en la Teoría Comunicativa de la Terminología como polisemia (Adelstein y Cabré, 2002, 2003). Al respecto, Adelstein y Cabré observan que la TCT propone un modelo polisémico de descripción y tratamiento de las unidades léxicas que posibilita diferenciar los sentidos especializados de los no especializados y que da debida cuenta de la flexibilidad de las unidades terminológicas.

### **2.3.1.2 Identificación de las palabras técnicas o términos**

Por las razones antes explicitadas, diferenciar las palabras técnicas o términos del resto de las palabras dista de ser una labor sencilla (Chung y Nation, 2004; Coxhead y Nation, 2001; Pearson, 1998). Desde la lingüística aplicada y la terminología se han propuesto métodos muy diversos para la identificación de las palabras técnicas en contextos especializados. A continuación se presentan los métodos más utilizados y se señalan sus principales dificultades.

#### **2.3.1.2.1 Un continuum de especialización para la identificación de palabras técnicas o términos**

Tanto la lingüística aplicada como la terminología proponen métodos para la identificación de las palabras técnicas o términos que se basan en el postulado de un *continuum* o *gradación de especialización* (Ciapuscio, 2003a; Coxhead y Nation, 2001; Guantiva, Cabré, y Castellà, 2008; Hoffmann, 1985, como se cita en Pearson, 1998), esto

---

<sup>24</sup> La *polisemia interdominio* se refiere a los distintos sentidos que un término adquiere en diferentes dominios especializados. Este tipo de polisemia implica que el significado adquirido por un término en un dominio se ha generado a partir del que tiene en otro dominio (Meyer y Mackintosh 2000, como se cita en Adelstein y Cabré, 2002).

<sup>25</sup> La expresión *polisemia intradominio* se utiliza cuando una palabra desterminologizada reingresa o se filtra nuevamente en un dominio de especialidad (Meyer y Mackintosh 2000, como se cita en Adelstein y Cabré, 2002).

<sup>26</sup> La *terminologización* se refiere al proceso por el cual palabras de la lengua general pasan a funcionar como términos en un dominio dado, y la *desterminologización* señala el proceso inverso por el que términos de un dominio de especialidad se incorporan al lenguaje común como expresiones de uso corriente (Russo, 2008). En síntesis, los procesos de terminologización y desterminologización establecen relaciones polisémicas entre usos especializados y no especializados de las unidades léxicas (Adelstein y Cabré, 2002).

es, las palabras técnicas pueden ubicarse en una escala que va desde el nivel más alto de especialización en un extremo hasta el nivel más bajo en el otro extremo. En la presente investigación, la noción de grados de especialización de las palabras es especialmente relevante para el análisis del vocabulario en inglés de libros de texto universitarios, en su carácter de ejemplares de un género utilizado con fines pedagógicos. Esta gradación permite dar cuenta no solo de las palabras altamente especializadas específicas de una disciplina, en este caso de la ingeniería, sino también de aquellas palabras originalmente clasificadas como generales o académicas en las listas de palabras propuestas en la literatura (cfr. Capítulo 3, sección 3.1.3) pero que adquieren significados especializados según su uso en géneros y disciplinas específicas.

En el marco de la lingüística aplicada, Coxhead y Nation (2001) proponen una escala de clasificación de las palabras técnicas conformada por cuatro categorías de palabras que varían desde las más técnicas a las menos técnicas. El uso de esta escala implica, esencialmente, decidir cuán estrechamente el significado de una palabra está asociado a un área temática, campo o disciplina en particular. En esta escala, la categoría 1 comprende palabras que no aparecen, o rara vez lo hacen, fuera de una determinada esfera de especialidad, tales como *galvanometer* en electrónica y *morpheme* en lingüística aplicada. Estas palabras son claramente técnicas ya que se utilizan en un único campo disciplinar. La categoría 2 incluye palabras que se emplean tanto dentro como fuera de un campo específico pero con significados diferentes, como por ejemplo en el caso de *caution*, cuyo significado en el área del derecho (caución) difiere del adquirido en un contexto de seguridad laboral (precaución). Las palabras de esta categoría se consideran técnicas debido a que el significado especializado que activan en un campo dado no puede ser accedido a través del significado más general que adquieren cuando son utilizadas fuera de dicha esfera de especialidad. Por su parte, la categoría 3 abarca palabras que se emplean tanto dentro como fuera de un campo determinado pero la mayoría de los usos con un significado particular se producen en cada campo específico. A diferencia de la categoría anterior, el significado especializado de estas palabras en una disciplina específica puede ser accedido fácilmente a partir de su significado fuera de ella. Son ejemplos de palabras correspondientes a la categoría 3 *dispersion* y *frequency* en la lingüística de corpus y *energy* y *resistance* en el campo de la electrónica. Por último, la categoría 4 está formada por palabras que son más comunes en un campo determinado que en cualquier otro contexto. Estas palabras tienen un menor grado de especialización que las de las categorías 1, 2 y 3, aunque cualquier persona familiarizada con la disciplina en cuestión tendrá una

idea más acabada y precisa de sus significados. Algunos ejemplos de palabras pertenecientes a la categoría 4 son *word* en el ámbito de la lingüística aplicada y *load* en el área de electrónica. Las palabras correspondientes a las categorías 3 y 4 son menos especializadas en virtud de que su forma y su significado no están restringidos a un único dominio. Además, es posible acceder con facilidad al significado de las palabras comprendidas en estas dos categorías por medio de su significado fuera de un campo determinado. La aplicación de esta escala de clasificación de palabras técnicas, tal como lo reconocen sus propios autores (Coxhead y Nation, 2001), lleva a los investigadores a depender de su intuición y conocimiento de una disciplina dada.

En el ámbito de la terminología, Hoffmann (1985, como se cita en Pearson, 1998) propone un esquema de categorización de las palabras técnicas que, en líneas generales, es similar al elaborado por Coxhead y Nation (2001). Según Hoffmann, en los textos de especialidad se pueden distinguir tres categorías de palabras: vocabulario específico de una disciplina, vocabulario no específico de una disciplina y vocabulario general. El vocabulario específico de una disciplina está integrado por términos que se utilizan de manera exclusiva en un solo dominio especializado y que se caracterizan por ser monosémicos. El vocabulario no específico de una disciplina comprende términos que, si bien establecen una referencia especializada, se emplean en más de un campo disciplinar. El vocabulario general incluye palabras que, aunque no aportan una referencia especializada en ningún dominio, se perciben de algún modo como especializadas por el hecho de aparecer en textos de una disciplina en particular.

#### **2.3.1.2.2 Otros métodos empleados para la identificación de palabras técnicas o términos**

Además del criterio del continuum de especialización, en la literatura se han propuesto otros métodos para la identificación de las palabras técnicas, entre ellos se destacan: la comparación de la frecuencia y el rango de palabras en un corpus especializado y en un corpus general, la identificación de combinaciones de palabras, las consultas con especialistas, el uso de diccionarios técnicos, y la extracción automática de términos.

Un método ampliamente utilizado para el reconocimiento de las palabras técnicas consiste en comparar la frecuencia de ocurrencia y el rango o consistencia de una palabra en un corpus especializado y en uno general o en un corpus de otra área disciplinar (Chung, 2003; Chung y Nation, 2004; Coxhead y Nation, 2001; Nation, 2001a; Pearson,

1998). La comparación de corpus permite detectar palabras clave<sup>27</sup> que aparecen con una frecuencia inusualmente elevada en un corpus especializado. Las palabras técnicas que son importantes en un campo específico exhibirán una mayor frecuencia y rango en un corpus especializado (Bowker y Pearson, 2002; Chung, 2003; Chung y Nation, 2004). Sin embargo, este método tiene una limitación: aquellas palabras que tienen la misma forma pero distinto significado se consideran una única palabra al momento del cálculo de frecuencias. Es decir, no se da debida cuenta de la multiplicidad de significados que las palabras pueden activar en contextos diferentes.

Por otra parte, la identificación de la combinación y coaparición frecuente de palabras y de los patrones típicos resultantes en campos especializados proporciona información valiosa sobre el vocabulario de una disciplina, más si se tiene en cuenta que los términos pueden ser unidades léxicas simples o complejas (Adelstein y Kuguel, 2005; Cabré, 1999), siendo estas últimas también llamadas *unidades terminológicas polilexemáticas* o *unidades terminológicas poliléxicas* (UTP)<sup>28</sup> (Adelstein, 2007; Cabré, Estopà, y Lorente, 1996). Estas combinaciones recurrentes de palabras adquieren valor sintagmático y terminológico debido a que denotan conceptos específicos de un ámbito de especialidad (Cabré, 1999; McCarthy, 1990). Además de las unidades terminológicas complejas o polilexemáticas, existen otras combinaciones de palabras que son frecuentes en el discurso especializado, aunque no se correspondan con conceptos establecidos de un campo determinado, y se denominan *unidades fraseológicas especializadas* (UFE)<sup>29</sup>. Cabré et al. (1996) sugieren que el estudio de las unidades fraseológicas especializadas permite una mejor identificación de los términos y de sus límites conceptuales (precisión semántica), y contribuye a acercar la terminología a los actos comunicativos de los hablantes, es decir, al contexto en que los términos se emplean.

La consulta con especialistas en un dominio dado es otro método utilizado para la identificación y selección de palabras técnicas (Bowker y Pearson, 2002; Cabré, 1999;

---

<sup>27</sup> Nuestra traducción para *keywords*, que deriva de la noción de *keyness* desarrollada en la lingüística de corpus.

<sup>28</sup> Cabré et al. definen a una *unidad terminológica polilexemática* como "una unidad léxica propia de un dominio de especialidad formada por más de un lexema" (1996, p. 5).

<sup>29</sup> Uno de los problemas centrales en torno a las unidades fraseológicas especializadas ha sido su distinción respecto de las unidades terminológicas polilexemáticas (UTP). Cabré et al. (1996) proponen los siguientes criterios diferenciadores: a) desde un punto de vista semiótico, las unidades terminológicas polilexemáticas (UTP) se asocian de forma estable a elementos referenciales; b) desde un punto de vista morfológico, están formadas al menos por dos unidades léxicas autónomas; c) desde la perspectiva de su uso, aparecen en un ámbito temáticamente especializado; y d) desde una perspectiva sintáctica, ocupan en un frase la posición de constituyente autónomo mínimo. Por el contrario, las unidades fraseológicas especializadas (UFE) son unidades sintácticas, no léxicas, constituidas por más de un lexema que son muy frecuentes en textos de dominios especializados.

Chung y Nation, 2004; Kuguel, 2007; Schmitt, 2010). Los expertos, por su amplia experiencia y profundo conocimiento de un ámbito de especialidad, son fuentes autorizadas para el reconocimiento de vocabulario propio del campo en cuestión<sup>30</sup>. Este método implica un análisis léxico detallado que, si bien puede insumir mucho tiempo y supone la participación de especialistas que usualmente tienen escasa disponibilidad horaria, puede resultar muy fructífero, particularmente en estudios del vocabulario de un dominio de especialidad basados en corpus.

El uso de diccionarios técnicos o de glosarios especializados representa otra forma de establecer si una palabra puede ser considerada o no un término (Chung y Nation, 2004). En este método se clasifica como término toda palabra que aparezca en un diccionario técnico o glosario de una determinada área de especialidad. Tal como señalan Chung y Nation (2004), una de las mayores desventajas de emplear un diccionario como parámetro para el reconocimiento de vocabulario especializado es desconocer los criterios específicos adoptados por los compiladores para la selección de las palabras a ser incluidas. Otra limitación, también identificada por estos autores, se relaciona con la dificultad de encontrar diccionarios técnicos que registren una lista lo suficientemente completa de términos correspondientes a distintas subdisciplinas con altos niveles de especialización.

Otro método usado en la actualidad para distinguir vocabulario especializado es el de la extracción automática de términos o reconocimiento automático de términos (Bowker y Pearson, 2002; Cabré, Estopà, y Vivaldi, 2001; Chung y Nation, 2004; Estopà, 2009). Este método consiste en la identificación automática de candidatos a términos en un corpus mediante el uso de un programa informático. Luego sigue una instancia de revisión manual donde el investigador selecciona de la lista resultante de candidatos solo aquellos que tienen valor terminológico (Bowker y Pearson, 2002). La extracción automática reviste al menos dos dificultades: por un lado, algunos de los candidatos incluidos en las listas no son términos (este problema se conoce como *ruido*); por otro lado, algunos términos no son reconocidos inicialmente por los programas informáticos disponibles y quedan por lo tanto excluidos de las listas (este problema se denomina *silencio*).

---

<sup>30</sup> Cabré et al. señalan en este sentido que "el reconocimiento de los términos de un ámbito de especialidad es una tarea que sólo pueden resolver de manera directa los especialistas en la materia" (1996, p. 2).

### 2.3.1.2.3 Criterios semánticos y pragmáticos para la identificación de palabras técnicas o términos

Además de los procedimientos y técnicas antes mencionados, distintos autores defienden el uso de criterios semánticos y pragmáticos para la identificación de palabras técnicas (Cabré, 1999; P. Meyer, 1997; Pearson, 1998), ya que estos dos campos, la semántica y la pragmática, representan la división más importante en el tratamiento del significado lingüístico (Cruse, 2006). Aunque no existen definiciones de ambos campos totalmente consensuadas, Cruse (2006) señala que la *semántica* se ocupa de los recursos estables de significado de una lengua en cuanto sistema, mientras que la *pragmática* trata sobre el uso de ese sistema para la comunicación en ocasiones y contextos particulares. Nuestra investigación adopta así un enfoque que conjuga la consulta a especialistas con el uso de criterios semánticos y pragmáticos para distinguir las palabras técnicas de las no técnicas en un corpus de libros de texto universitarios del campo de la ingeniería.

Desde una perspectiva semántica, los términos se caracterizan por tener una forma lingüística o designación que hace referencia a un concepto del mundo real; es decir, los términos son concebidos como unidades de referencia porque tienen un propósito que es esencialmente referencial (Cabré, 1999). Los conceptos pueden referirse al mundo externo o interno, a una realidad concreta o abstracta, colectiva o individual, y pueden clasificarse en objetos o entidades, procesos, operaciones y acciones, propiedades, estados y cualidades, y relaciones (Cabré, 1999; Sager, 1990). Uno de los rasgos más distintivos de los términos reside en que se utilizan para designar conceptos pertenecientes a dominios especializados (Cabré, 1999; Chung y Nation, 2004). Dentro de un área de especialidad, los conceptos se organizan en conjuntos jerarquizados llamados sistemas conceptuales y estos sistemas reflejan la visión del mundo real que tienen los expertos de esa disciplina o esfera de actividad. Por lo tanto, la terminología de una disciplina, en nuestro estudio la ingeniería, es la representación lingüística de su organización conceptual (Cabré, 1999; Cabré y Estopà, 2005; Pearson, 1998).

Además de la aplicación de criterios semánticos para la identificación de palabras técnicas, es necesario recurrir a criterios pragmáticos debido a que la carga semántica de una palabra varía en función del contexto de uso (Carter, 1987; Cruse, 2000; P. Meyer, 1997). El vocabulario técnico se emplea generalmente en contextos específicos, es decir en situaciones comunicativas especializadas (Cabré, 1999). Como señalan Atkins y Rundell (2008), el nivel de especialización del contexto determina cambios en el significado de una

palabra y este se vuelve cada vez más acotado y restringido a un dominio en particular<sup>31</sup>. En otras palabras, mientras más específico sea el campo disciplinar, más especializado será el vocabulario.

Para poder identificar vocabulario técnico a partir de criterios pragmáticos, Bowker y Pearson (2002), Cabré (1999) y Pearson (1998) proponen considerar los niveles de especialización de las situaciones comunicativas. Estos autores identifican al menos tres posibles contextos donde se puede llevar a cabo una comunicación especializada, definidos principalmente por el grado de competencia de los participantes sobre una determinada área del conocimiento. En el primer tipo de situación, la comunicación se produce entre expertos; es decir, los interlocutores son personas que tienen formación y experiencia en un campo especializado (Bowker y Pearson, 2002). Este contexto comunicativo particular favorece el uso de vocabulario altamente especializado con el que los participantes, en su condición de expertos en un dominio dado, están debidamente familiarizados. La terminología empleada solo se explica en el caso de que algún concepto esté siendo redefinido, o bien en el caso de que se esté acuñando un nuevo término. En el segundo tipo de situación, la comunicación se da entre expertos y semilegos (estos últimos también llamados aprendices de especialistas) (ver aquí sección 2.2.3). En este tipo de contexto, los expertos también emplean palabras técnicas, aunque la densidad terminológica es menor y algunos términos, que se presumen desconocidos o no comprendidos acabadamente por parte de los lectores, son objeto de tratamiento mediante el uso de palabras de la lengua general. En el tercer tipo de situación, la comunicación tiene lugar entre expertos relativos y no expertos<sup>32</sup>. Los no expertos (también denominados no iniciados) son definidos por Pearson (1998) como adultos que tienen una educación o formación general y que no están involucrados, ya sea de manera profesional o por medio de sus intereses personales, con un campo especializado determinado. En este contexto, la densidad terminológica es mucho menor que en el segundo tipo de situación. Como no se asume que los interlocutores compartan conocimientos específicos sobre un área disciplinar, los autores pueden elegir emplear palabras de la lengua general para describir conceptos y, cuando utilizan términos, no pretenden que el lector alcance el mismo nivel de comprensión de dicho vocabulario especializado, sino más bien una idea general.

---

<sup>31</sup> Este proceso es denominado *especialización* o *restricción semántica* (Atkins y Rundell, 2008; Geeraerts, 2010). Geeraerts (2010) define la *especialización semántica* como un tipo de cambio léxico-semántico por medio del cual un ítem léxico desarrolla un nuevo significado que adopta una relación de subordinación con respecto al significado original.

<sup>32</sup> Nuestra traducción para *relative experts* y *non-experts* (Pearson, 1998).

Como sugieren distintos académicos del campo de la lingüística aplicada y de la terminología, en la presente investigación se llevó a cabo la identificación y clasificación de palabras técnicas y no técnicas en libros de texto universitarios del dominio de la ingeniería a partir de la noción del continuum de especialización. Con el propósito de reconocer los significados especializados que activaban las palabras a lo largo del continuum, este estudio se valió de la participación de especialistas (cfr. Capítulo 3, sección 3.2.3) y la integración de criterios semánticos y pragmáticos.

## **2.4 Lingüística de corpus**

### **2.4.1 La lingüística de corpus como metodología de investigación**

Para el análisis del vocabulario de libros de texto universitarios de ingeniería publicados en inglés se utilizó metodología de corpus, la cual proporcionó los lineamientos, criterios, herramientas y procedimientos para el diseño y relevamiento del corpus compilado en el presente estudio (cfr. Capítulo 3, sección 3.1.1). La lingüística de corpus ha cobrado particular relevancia a lo largo de las últimas décadas en su condición de enfoque metodológico para el estudio de las lenguas que ofrece "oportunidades revolucionarias para la descripción, análisis y, enseñanza de discursos de todo tipo" (Parodi, 2008b, p. 95). Los estudios de corpus han adquirido una importancia creciente en áreas tales como el análisis del discurso, la estilística, la lexicología, la lexicografía, la lingüística forense, la psicolingüística, la sociolingüística, la terminología, la traducción y, particularmente en la lingüística aplicada, por la diversidad de aplicaciones pedagógicas en el contexto de la enseñanza de lenguas (Hunston, 2002; Kennedy, 1998; Stubbs, 2004). La lingüística de corpus aporta principios, procedimientos, herramientas y técnicas innovadoras para el estudio de grandes cantidades de material lingüístico auténtico, y los hallazgos obtenidos redundan en descripciones más confiables de la naturaleza, estructura, usos y variación de la lengua (Biber, Conrad, y Reppen, 1998; Biber y Reppen, 2015; Kennedy, 1998; Sinclair, 1991, 2005; Stubbs, 2004).

Los estudios basados en corpus emplean una metodología que combina técnicas analíticas cuantitativas y cualitativas, esto es, distintas mediciones estadísticas y la interpretación de los investigadores. El análisis cuantitativo requiere generalmente el empleo de computadoras y paquetes informáticos para el procesamiento automático de datos, mientras que el análisis cualitativo se realiza usualmente de manera manual, dado que consiste en la interpretación y organización sistemática de los datos por parte de los

investigadores, sin mediación de programas informáticos. Una combinación de ambos enfoques proporciona información más rica y detallada del fenómeno lingüístico objeto de estudio, como así también resultados más confiables y generalizables (Evison, 2010; McEnery y Wilson, 2001).

## **2.4.2 El corpus como objeto de estudio**

### **2.4.2.1 Definición**

El objeto de análisis de esta metodología de investigación lingüística es precisamente el corpus<sup>33</sup>, concebido como una compilación sistemática de textos auténticos, usualmente aunque no exclusivamente en formato digital, que son seleccionados atendiendo a ciertos criterios externos para el estudio de determinados elementos o rasgos lingüísticos, de manera que se puedan realizar conclusiones o generalizaciones sobre los usos de una lengua en base a evidencia empírica (Hunston, 2002; Sinclair, 2005). De la definición se desprenden cuatro características distintivas de un corpus. Primero, el diseño y compilación de un corpus obedece a una planificación y estructuración en base a criterios externos; es decir, no se trata de una colección de textos realizada sin un orden o una lógica determinada (Kennedy, 1998; Sinclair, 2005). En segundo término, un corpus es una compilación de textos auténticos elegidos para caracterizar una lengua; esto significa que no son textos especialmente creados o procesados pedagógicamente para el estudio o enseñanza de la lengua (Cheng, 2012). Tercero, un corpus es una muestra que ha sido diseñada para representar una totalidad, ya sea una lengua o variedad de lengua determinada (Biber, 2008). En cuarto lugar, un corpus es construido con un propósito lingüístico particular que determina su diseño, tamaño y naturaleza (Cheng, 2012; Kennedy, 1998).

### **2.4.2.2 Criterios para el diseño y compilación del corpus**

El diseño y la compilación del corpus tienen una importancia medular. Como observa Sinclair (1991), las decisiones acerca de la estructura interna del corpus y su preparación sistemática son aspectos muy importantes ya que van a tener un rol decisivo en la calidad y validez de los resultados de la investigación. En relación a la estructura del corpus, los investigadores deben decidir qué variedad de la lengua y qué géneros serán

---

<sup>33</sup> La palabra *corpus* proviene del latín y significa cuerpo (Baker, Hardie, y McEnery, 2006).

analizados, y deben además establecer qué procedimientos se utilizarán en la selección de los textos (Biber, 2008; Sinclair, 1991). En esta etapa de planificación, la lengua o variedad de lengua a ser estudiada debe ser descripta de manera sistemática, mediante la identificación y definición de los géneros y registros a ser explorados a partir de variables no lingüísticas<sup>34</sup>, las cuales incluyen aspectos relacionados con la producción, recepción y circulación de los textos en contextos particulares (Biber, 2008; Kennedy, 1998; Sinclair, 2005). Dicho de otro modo, un corpus debe reflejar una correspondencia entre la lengua que está siendo explorada y el tipo de material compilado (Biber, 1993).

Una vez identificados y definidos los géneros, se inicia la etapa de selección de los ejemplares textuales, denominada *procedimiento de muestreo* por Kennedy (1998). Aunque la selección de los textos está siempre condicionada por cuestiones de factibilidad y disponibilidad (Hunston, 2002; Stubbs, 2004), esta debe realizarse siguiendo un número de parámetros, claramente diferenciados entre sí, para el diseño de un corpus que sea representativo de la lengua o variedad a ser analizada. A continuación se presenta una síntesis de los criterios que deben considerarse para la compilación de un corpus, sugeridos por Biber (2008), Hunston (2002), C. Meyer (2004), Pearson (1998) y Sinclair (1991, 2005).

Dos principios fundamentales que guían el diseño y la construcción de un corpus son la *representatividad* y el *balance*. Como se mencionara más arriba, todo corpus es una muestra que pretende ser representativa de una totalidad y, como tal, debería exhibir características similares a la lengua que representa<sup>35</sup>. Biber (2008) define la representatividad como el grado en que una muestra incluye el rango completo de variables de una población, en el sentido de una lengua o variedad de lengua. Por otro lado, se considera que un corpus es balanceado cuando está estructurado en diferentes secciones que contienen proporciones iguales de determinadas variables, como por ejemplo la inclusión de textos con el mismo grado de especialización (Hunston, 2002; Kennedy, 1998; Sinclair, 2005; Stubbs, 2004). La representatividad y el balance de un corpus son objetivos difíciles de precisar y, más aún, de alcanzar (Sinclair, 2004). Kennedy (1998) reconoce que estas dos nociones dependen del criterio de los investigadores y solo pueden ser

---

<sup>34</sup> Algunas de las variables extralingüísticas, también llamadas criterios externos, a ser consideradas al momento del diseño del contenido y estructura de un corpus son: el modo de los textos (escrito, oral); el tipo de texto (o género); el dominio de los textos (por ej., académico o periodístico); el tema de los textos; el origen de los textos; los participantes; la fecha de publicación de los textos o de producción del evento de habla; la lengua, lenguas o variedad/es de lengua del corpus, entre otras (Cheng, 2012; Kennedy, 1998; Sinclair, 2004).

<sup>35</sup> Sinclair (1991) destaca que ningún corpus, sin importar cuán grande sea o cuán meticulosamente haya sido diseñado, puede tener exactamente las mismas características que la lengua a la cual representa.

aproximadas. No obstante, se han establecido ciertos parámetros que deben tenerse en cuenta al momento del diseño y compilación de un corpus: tamaño de un corpus, disponibilidad de textos en formato electrónico, extensión de los textos, uso de ejemplares textuales completos, período de publicación y filiación de los autores. Estos parámetros son explicados de manera sucinta debajo.

El tamaño de un corpus<sup>36</sup> depende de los objetivos de la investigación para la que se construye dicho corpus, la metodología empleada (Sinclair, 1991, 2005), la disponibilidad de recursos electrónicos, y el tiempo con el que se cuenta para la compilación de los textos (C. Meyer, 2004). Generalmente, los corpus que contienen entre medio millón a un millón de palabras son llamados pequeños (Hunston, 2002), como por ejemplo el corpus construido para el presente estudio (cfr. Capítulo 3, sección 3.1.1). Sin embargo, la distinción entre corpus pequeños y grandes se ha tornado bastante difusa (Gavioli, 2005); un corpus que es considerado grande hoy puede no serlo mañana ya que, con los avances computacionales, los corpus son cada día más grandes (Ragan, 2001). Sinclair (1991) afirmó inicialmente que un corpus debe ser tan grande como sea posible. Pero luego Sinclair (2001) expresó que el tamaño de un corpus no determina la calidad de una investigación, dado que los corpus pequeños son apropiados para estudios que persiguen objetivos de investigación muy específicos. Por otra parte, Hunston (2002), Kennedy (1998) y Koester (2010) coinciden en señalar que un corpus pequeño es adecuado para el análisis de vocabulario de alta frecuencia<sup>37</sup>. Otra ventaja de usar corpus pequeños es que permiten establecer relaciones más estrechas entre el corpus y el contexto específico donde circula y cobra sentido la lengua representada por el corpus (Koester, 2010). Además, los corpus pequeños son más adecuados para ser analizados manualmente, mientras que los corpus grandes son más apropiados para ser procesados de manera automática con la asistencia de programas informáticos (C. Meyer, 2004).

Un criterio, eminentemente pragmático, a tener en cuenta al momento del diseño y la construcción de un corpus es el de la disponibilidad de textos en formato electrónico. Como observa Reppen (2010), crear un corpus de textos escritos es una tarea más sencilla que la de construir un corpus de textos orales porque, en la actualidad, un número importante de textos escritos ya se encuentran disponibles en formato electrónico, como es el caso de los textos seleccionados para conformar el corpus del presente estudio.

---

<sup>36</sup> El tamaño de un corpus se establece en función del número total de palabras (*tokens*).

<sup>37</sup> Koester (2010) señala que en un corpus pequeño es posible analizar todas las ocurrencias de las palabras de alta frecuencia, mientras que en un corpus grande el volumen de datos para estas palabras se volvería difícil de manejar y se estudiaría sólo una muestra.

Otros parámetros que se deben atender son la extensión de los textos seleccionados y el uso de textos completos. Con respecto a la longitud de los textos, Sinclair (2005) destaca que, desde un punto de vista lingüístico, no agrega valor a un corpus que los textos seleccionados sean de la misma extensión. Como observa Sinclair (1991), y para referirnos al uso de textos completos, tiene mayor relevancia que los ejemplares textuales seleccionados sean documentos completos que reflejen la extensión típica de los textos que el corpus intenta representar.

El período de publicación de los textos recolectados y la filiación de los autores son dos variables que también contribuyen a la representación de un corpus. En relación al período de publicación, Sinclair (1991) sostiene que es importante compilar textos que hayan sido escritos dentro de un tiempo determinado. Kennedy (1998) y C. Meyer (2004) señalan que esto es particularmente importante en el caso de corpus sincrónicos, los cuales se caracterizan por representar los usos de la lengua en un período dado. En lo concerniente a la filiación de los escritores, en los corpus especializados (ver aquí 2.4.3) es importante seleccionar textos cuyos autores se desempeñan en un mismo contexto o ámbito para asegurar la generalización de los datos relevados (C. Meyer, 2004).

### 2.4.3 Tipos de corpus

En función de los propósitos que orientan una investigación, algunos investigadores hacen uso de corpus textuales ya disponibles, como por ejemplo el *British National Corpus* (BNC), el *Corpus of Contemporary American English* (COCA), o el *Michigan Corpus of Academic Spoken English* (MICASE). Por el contrario, otros investigadores emplean corpus especialmente diseñados y construidos para responder preguntas de investigación específicas, como es el caso de nuestro estudio. Según Hunston (2002), los corpus pueden clasificarse en varios tipos<sup>38</sup>: generales, especializados, comparables, paralelos, pedagógicos, de monitoreo, diacrónicos y sincrónicos. En el presente trabajo, se construyó un corpus especializado para el análisis del vocabulario de un género, el libro de texto universitario, en un campo disciplinar, la ingeniería, y en dos subdisciplinas, la ciencia de los materiales y la mecánica de los fluidos.

En el caso de los estudios que tienen como finalidad indagar los géneros (académicos o profesionales) que un grupo de estudiantes necesita leer y/o escribir en contextos específicos, adquiere especial importancia la distinción entre corpus generales

---

<sup>38</sup> Dado los objetivos de nuestro estudio, nos detendremos únicamente en la distinción entre corpus generales y corpus especializados.

construidos con objetivos múltiples versus corpus especializados compilados con objetivos específicos<sup>39</sup>. Gavioli (2005), Hunston (2002) y C. Meyer (2004) definen a un *corpus general* como un corpus diseñado y recolectado para representar a toda una lengua (por ej., inglés) o a una variedad geográfica (por ej., inglés británico). Generalmente, este tipo de corpus está conformado por una variedad de géneros y usos de la lengua oral y escrita, es construido para atender propósitos múltiples, y tiene un tamaño mucho mayor que un corpus especializado. En contraposición, un *corpus especializado* es definido por Sinclair (1991) como un corpus que representa una muestra de una variedad determinada de la lengua, incluye un número limitado de géneros, y es compilado con el propósito de responder preguntas de investigación específicas. A diferencia de los corpus generales, los corpus especializados aportan evidencia empírica sobre patrones lingüísticos en contextos de uso específicos (Koester, 2010; O’Keeffe, McCarthy, y Carter, 2007).

#### 2.4.4 Unidades de análisis en estudios de vocabulario basados en corpus

La noción de *palabra* como unidad de análisis adquiere especial relevancia en los estudios de vocabulario basados en corpus, especialmente en el análisis cuantitativo. Esto se debe a que el concepto de palabra<sup>40</sup> admite interpretaciones muy diversas que se manifiestan en definiciones diferentes: palabra, lema o lexema, familia de palabras, *token* y *type*<sup>41</sup>, entre las categorías más utilizadas. En esta investigación, las unidades de análisis empleadas fueron casos (*tokens*), tipos (*types*) y lemas o lexemas.

En el campo de la lingüística de corpus, Sinclair (1991) define a las *palabras* como formas gráficas distintivas que están constituidas por una secuencia ininterrumpida de letras<sup>42</sup>. Como señala Carter (1987), esta definición ortográfica de palabra privilegia la forma de la palabra y no es sensible a distinciones de significado y de función gramatical. Una noción teórica que contribuye a solucionar parcialmente los problemas recién mencionados es la de lema o lexema. Carter (1987) y Singleton (1999) conceptualizan a

---

<sup>39</sup> Tribble (2002, como se cita en Koester, 2010) sostiene que los corpus generales no satisfacen las necesidades de estudiantes y profesores de inglés con fines específicos o académicos, debido a que proporcionan una cantidad demasiado grande de datos en un espectro extremadamente amplio. Dicho de otro modo, los datos están muy poco focalizados para ser de ayuda a estudiantes con propósitos de aprendizaje específicos.

<sup>40</sup> Según Singleton (1999), el concepto de *palabra* dependerá en gran medida del nivel de abstracción con el que esté operando un hablante o escritor dado, el nivel lingüístico que esté siendo analizado, y el grado en el cual el contenido semántico sea tratado como un criterio determinante.

<sup>41</sup> De aquí en adelante los términos en inglés *token* y *type* serán reemplazados por equivalentes en español ya utilizados en la literatura (ver Dúcculi, Muñoz, y Beck, 2016; Garofolo, 2012; Garofolo y Picchio, 2016): *caso* y *tipo de palabra*, respectivamente.

<sup>42</sup> Esta es la definición que adoptan los programas informáticos para el conteo de palabras y cálculo de frecuencias.

los *lemas* o *lexemas* como unidades abstractas conformadas por grupos de palabras relacionadas, tales como *eat, eats, eating, ate, eaten* o *quick, quicker, quickest* (Hunston, 2002). Los lemas o lexemas incluyen los siguientes elementos: una forma base, sus flexiones (que indican información gramatical tales como persona, tiempo verbal y número en el caso de verbos, y forma comparativa y superlativa en el caso de adjetivos), sus formas reducidas o contraídas (*n't*), y variantes ortográficas (por ej., *favor, favour*). Todas estas formas, que usualmente comparten el mismo significado básico y generalmente pertenecen a la misma categoría gramatical, constituyen en cierto modo una entidad léxica única (Biber et al., 1998; Cruse, 2000; Hunston, 2002; Nation, 2001a, 2016; Singleton, 1999). La palabra base de un lema puede ser la forma base sin flexiones o una palabra miembro del lema con mayor frecuencia de ocurrencia en un corpus (Sinclair, 1991). El concepto de lema guarda estrecha relación con la noción de *familia de palabras*, definida por Nation (2001a) como un conjunto de palabras que contiene una palabra base y sus formas flexionadas, como así también sus formas derivadas, las cuales corresponden a diferentes categorías gramaticales con significados diferentes.

La noción de palabra ortográfica está muy relacionada con los conceptos de caso y tipo de palabra empleados en la lingüística de corpus. Un *caso* se define como una cadena lineal de caracteres delimitada en cada lado por un espacio en blanco o un signo de puntuación (Hunston, 2002; Kennedy, 1998; Nation, 2016; Sinclair, 1991, entre muchos otros). Cuando la unidad de conteo en los estudios de corpus es el caso, el programa informático cuenta cada palabra presente en el corpus, incluso las repeticiones. Otra forma de contar palabras supone considerar cada palabra diferente en el corpus, sin repeticiones, como un tipo de palabra. Un *tipo de palabra* se define como cada palabra con identidad diferente que aparece en un texto o corpus, sin contar repeticiones (Hunston, 2002; Kennedy, 1998; Nation, 2016; Singleton, 1999, entre muchos otros). Adoptar los tipos de palabras como unidad de conteo en un corpus trae aparejada la dificultad de omitir casos de homonimia y polisemia.

## **2.4.5 Técnicas y procedimientos de análisis**

### **2.4.5.1 Técnicas cuantitativas: listas de palabras, rango y cobertura**

Los estudios de vocabulario basados en corpus utilizan un conjunto de técnicas para el procesamiento automático de los datos. Las listas de palabras son uno de los métodos más empleados para visualizar y analizar vocabulario en un corpus, y pueden estructurarse

de distintas maneras, las dos más comunes son por orden alfabético y por frecuencia de ocurrencia ascendente o descendente (Nation, 2001b; Sinclair, 1991). Las listas de palabras son consideradas la herramienta más básica en los estudios de corpus, dado que constituyen la primera aproximación a un corpus y proporcionan un panorama general del vocabulario relevado (Scott, 2010). Asimismo, permiten al investigador identificar y seleccionar elementos lingüísticos específicos para su posterior análisis, más detallado y conforme a los objetivos de la investigación, como por ejemplo diferenciar palabras de alta frecuencia y de baja frecuencia.

Las listas de palabras proporcionan información cuantitativa en conteos de frecuencia que indican el número total de casos y de tipos de palabras en un corpus. El número total de casos provee información sobre el tamaño de un corpus, en tanto que el número de tipos de palabras determina el tamaño del vocabulario de un corpus.

Además del tamaño del vocabulario y del corpus, las listas de palabras generadas en los programas informáticos muestran distintas medidas estadísticas tales como: relación entre tipos y casos (comúnmente referenciada por su sigla en inglés, *TTR -type-token ratio*), relación estandarizada entre tipos y casos (usualmente referenciada por su sigla en inglés, *STTR -standardized type-token ratio*<sup>43</sup>), rango o consistencia de las palabras, y cobertura de las palabras. La *relación entre tipos y casos* (TTR) es una medida que determina la *variación léxica* (Laufer y Nation, 1995) o *diversidad léxica* (Malvern, Richards, Chipere, y Durán, 2004; Schmitt, 2000) de un corpus, entendida como la variedad de vocabulario de un corpus. La relación entre tipos (*types*) y casos (*tokens*) se calcula dividiendo el número total de tipos de palabras por el número total de casos y como resultado se obtiene un valor comprendido entre 0 y 1, que también puede ser expresado como un porcentaje (Laufer y Nation, 1995). Los valores más altos (más cercanos a 1) indican una mayor variación o diversidad léxica en un corpus, mientras que los valores más bajos (más próximos al 0) señalan una mayor repetición de palabras y por ende una menor variación en el vocabulario (Cheng, 2012; Malvern et al., 2004; Schmitt, 2000).

Como señalan diferentes autores (Chipere, Malvern, y Richards, 2004; Cobb y Horst, 2015; Malvern et al., 2004; Schmitt, 2000), la relación entre tipos y casos es una medida que depende claramente de la longitud de los textos analizados. Por lo tanto, el tamaño de un corpus tendrá una influencia directa en la proporción entre tipos y casos. Mientras más pequeño sea el corpus, más alto será el valor de TTR (mayor variabilidad

---

<sup>43</sup> Esta relación estandarizada entre tipos y casos también es referida en inglés como *mean segmental type-token ratio -MSTTR* (ver Malvern et al., 2004).

léxica) y, por el contrario, mientras más grande sea, más bajo será el valor obtenido (menor variabilidad léxica). P. Baker, Hardie, y McEnery (2006), Cheng (2012) y Malvern et al. (2004) explican que este fenómeno se produce porque las palabras tienen mayores probabilidades de repetirse en textos extensos y en corpus grandes. Dicho de otra manera, en general, el número de tipos de palabras no aumenta tan rápidamente como el número de casos en textos extensos (Schmitt, 2000).

A fin de neutralizar la influencia de la extensión de los textos, se puede calcular la *relación estandarizada entre tipos y casos* (STTR), medida estadística que resulta particularmente apropiada en corpus conformados por textos de diferente longitud (McEnery y Hardie, 2012; Scott, 2004). Esta relación se determina calculando la proporción tipo-caso en segmentos textuales consecutivos de igual extensión que no coinciden necesariamente con la longitud de los textos individuales que componen un corpus. Luego, el programa calcula la relación tipo-caso promedio a través de todos los segmentos del corpus y, de ese modo, se obtiene un porcentaje que se constituye en un indicador más fiable del grado de variación léxica de un corpus (P. Baker et al., 2006). Para obtener este índice estandarizado es necesario configurar el programa informático de análisis lingüístico para que segmente el corpus en un número determinado de intervalos del mismo número de casos<sup>44</sup>, tomando generalmente como parámetro la longitud del texto más pequeño (Malvern et al., 2004; Scott, 2004).

Una lista de palabras también proporciona información sobre el *rango* de las palabras en un corpus, a veces llamado *consistencia* (Scott, 2001; Scott y Tribble, 2006). Esta medida muestra la distribución de las palabras en los distintos textos que integran un corpus, es decir, indica la ocurrencia de una palabra a lo largo de las diferentes secciones de dicho corpus (Nation, 2001a). Las palabras consistentes tienen un rango amplio ya que ocurren de manera frecuente en una amplia variedad de textos del corpus, mientras que las palabras inconsistentes tienen un rango más acotado dado que están restringidas a algunos textos (Nation, 2001b). En un corpus general, el rango aporta información sobre cuán consistentemente una palabra dada se utiliza en diferentes tipos de textos. Por el contrario, en un corpus especializado, el rango permite identificar aquellas palabras que caracterizan un género (Scott y Tribble, 2006).

---

<sup>44</sup> En el caso del programa de análisis lingüístico utilizado en esta investigación, *WordSmith Tools 7.0*, la opción configurada por defecto computa la relación estandarizada entre tipos y casos cada 1.000 palabras del corpus.

Además del rango o consistencia de las palabras, otra medida estadística que se calcula en los estudios de vocabulario basados en corpus es la *cobertura* de una lista de palabras en un corpus. Esta medida consiste en calcular el porcentaje de casos en un corpus que aparecen en una lista de palabras específica (Nation, 2001b; Nation y Kyongho, 1995). La cobertura se obtiene sumando todas las ocurrencias de los tipos de palabras de una lista determinada que aparecen en un corpus, y luego convirtiendo los números en porcentajes que reflejan una proporción respecto del número total de casos.

En síntesis, las listas de palabras generadas por paquetes de análisis lingüístico proporcionan información estadística valiosa sobre el vocabulario de un corpus pero presentan algunas limitaciones. Una de las más importantes es que las listas de palabras ofrecen solo datos cuantitativos y muestran los tipos de palabras fuera de su contexto natural de aparición, es decir, los textos. Es así que las listas no distinguen palabras que tienen la misma forma pero pertenecen a distintas categorías gramaticales<sup>45</sup> (Bowker y Pearson, 2002; Stubbs, 2004). Otra de las deficiencias más relevantes es que los programas informáticos no distinguen los diferentes significados de una misma palabra y computan todas las ocurrencias de una palabra ortográfica con distintos significados como la misma palabra (Bowker y Pearson, 2002; Stubbs, 2004). Para dar debida cuenta de estas dificultades, es necesario combinar el enfoque cuantitativo con técnicas cualitativas, como por ejemplo, el análisis de concordancias.

#### **2.4.5.2 Técnicas cualitativas: concordancias, uso y significado**

La producción e interpretación de concordancias es una de las técnicas cualitativas más importantes para el estudio de vocabulario en corpus textuales<sup>46</sup>. Una vez que se genera de manera automática una lista de palabras de un corpus y se seleccionan los elementos lingüísticos que serán sometidos a un análisis más pormenorizado, se procede en una próxima etapa a explorar dichos ítems léxicos en su contexto real de uso, para poder así determinar sus significados y observar su comportamiento. Sinclair (1991) define a una *concordancia* como una colección de todas las ocurrencias de una palabra, cada una de ellas en su propio contexto textual, es decir el cotexto o entorno léxico-gramatical inmediato de una palabra, que suele comprender cuatro palabras hacia la izquierda y la derecha (Scott y Tribble, 2006; Sinclair, 1991).

---

<sup>45</sup> Este fenómeno se suele denominar *ambigüedad gramatical* (Bowker y Pearson, 2002).

<sup>46</sup> La concordancia se considera una técnica analítica útil tanto para la formulación como la verificación de hipótesis lingüísticas.

La mayoría de los programas informáticos disponibles permiten generar concordancias mediante el ingreso de un *ítem de búsqueda* que puede ser una *palabra clave* o *nodo* o incluso una frase (P. Baker et al., 2006; Evison, 2010). Las concordancias obtenidas se exhiben en la pantalla en *formato KWIC* (*Key Word in Context*). En el formato KWIC, el ítem de búsqueda se visualiza en el centro de cada línea (concordancia) y el número de palabras que constituyen el cotexto hacia la izquierda y hacia la derecha es preestablecido por el investigador, según los objetivos de la investigación (Bowker y Pearson, 2002; Kennedy, 1998; Sinclair, 1991). Este formato posibilita además ordenar alfabéticamente las líneas de una concordancia hacia la izquierda o derecha del ítem de búsqueda.

Las concordancias suponen un reordenamiento novedoso de las palabras de un corpus ya que se leen como una secuencia de fragmentos dispuestos de manera vertical (Hoey, Mahlberg, Stubbs, y Teubert, 2007). En el eje horizontal, una línea o concordancia representa una instancia de uso de la lengua (*parole* en términos saussurianos); mientras que en el eje vertical, una concordancia muestra co-ocurrencias repetidas de palabras y estos patrones son evidencia del sistema de una lengua (*langue* en términos saussurianos) (Hoey et al., 2007; Stubbs, 2004; Tognini-Bonelli, 2001). Al revelar patrones formales o tendencias en el sistema de una lengua, las concordancias proporcionan evidencia de prácticas sociales recurrentes en una comunidad discursiva (Hoey et al., 2007).

Como señala Hunston (2002), las concordancias aportan datos lingüísticos<sup>47</sup> que permiten realizar cuatro tipos de observaciones: aquello que es central y típico, distinciones de significado, relaciones entre significados y patrones, y distintos tipos de detalles. Un corpus proporciona información sobre el comportamiento recurrente de las palabras y revela sus usos típicos, tales como patrones de co-selección y los significados más frecuentes. Por otra parte, el análisis de concordancias posibilita establecer distinciones entre palabras con significados similares (por ej., cuasi-sinónimos), como así también determinar diferentes significados de una misma palabra. Mautner (2009) señala que las concordancias permiten elevar a la superficie textual un rango completo de elementos de significado que serían imposibles de relevar y catalogar sin la evidencia proporcionada por estudios de corpus.

---

<sup>47</sup> Sinclair propone en su libro *Reading Concordances* (2003) seguir un procedimiento que consta de siete etapas al momento de trabajar con concordancias (iniciar, interpretar, consolidar, informar, reciclar, registrar resultados y repetir). El autor sostiene que la realización de todas estas etapas permite dar cuenta de los datos hallados de manera sistemática y exhaustiva.

Observar los distintos significados y los diferentes usos de palabras en su contexto original adquiere una importancia central en investigaciones de naturaleza cualitativa que intentan atender fenómenos como la polisemia y la homonimia. En el presente estudio, la generación e interpretación de concordancias permitió identificar los distintos sentidos que las palabras más frecuentes en el corpus pueden activar, como así también contribuyó a diferenciar significados técnicos de no técnicos.

## **2.5 Síntesis del capítulo**

En este capítulo hemos expuesto los fundamentos teóricos que sustentan y guían la presente investigación. La teoría de análisis de los géneros nos ha proporcionado los conceptos y lineamientos necesarios para una definición y caracterización del género libro de texto universitario. La literatura existente sobre vocabulario, especialmente los estudios que exploran el léxico en inglés, ha aportado las nociones y herramientas para analizar y describir los distintos tipos de vocabulario relevados en la investigación. La lingüística de corpus ha provisto los principios teórico-metodológicos para la construcción y análisis de un corpus especializado. En el próximo capítulo se detallarán la metodología adoptada, los materiales utilizados y las técnicas y procedimientos seguidos en las distintas etapas del análisis cuantitativo y cualitativo que propone este estudio.

## Capítulo 3

### Metodología

Este capítulo detalla la metodología adoptada y los materiales utilizados para la realización del presente estudio. La sección 3.1 se refiere a los criterios que se siguieron para el diseño y construcción del corpus, al paquete de herramientas informáticas empleado para explorar el léxico del área disciplinar seleccionada, y a las listas de palabras usadas para el análisis de los tipos de vocabulario. La sección 3.2 desarrolla las etapas de recolección y análisis de los datos, y explica las diferentes técnicas y procedimientos que se utilizaron en el análisis cuantitativo y cualitativo del corpus.

### 3.1 Materiales

#### 3.1.1 El corpus

En función de los objetivos planteados en este estudio, se construyó un corpus especializado y pedagógico (Hunston, 2002; Sinclair, 1991). El corpus, de ahora en adelante denominado CIILTU (*Corpus de inglés de ingeniería de libros de texto universitarios*), se considera especializado porque es representativo de un género, el libro de texto universitario, un registro, el discurso académico-científico, un dominio, la ingeniería, dos subdominios, la ciencia de los materiales y la mecánica de los fluidos, y una lengua, el inglés. Además, se trata de un corpus pedagógico porque está conformado por libros de texto que se utilizan en el nivel superior, puntualmente en la Universidad Nacional de Córdoba. Los libros compilados son usados como bibliografía de estudio y de consulta en asignaturas troncales de distintas carreras de ingeniería y como material de lectura en cursos de lectura comprensiva en inglés destinados a estudiantes de dicha universidad en este campo disciplinar. Más específicamente, el libro *A Brief Introduction to Fluid Mechanics* (Young, Munson, Okiishi, y Huebsch, 2011) forma parte de la bibliografía recomendada por la cátedra *Mecánica de los fluidos* para las carreras de Ingeniería Aeronáutica, Ingeniería Industrial, Ingeniería Mecánica e Ingeniería Mecánica Electricista en la Facultad de Ciencias Exactas, Físicas y Naturales (FCEFyN) de la Universidad Nacional de Córdoba. El segundo libro, *Materials Science and Engineering: an Introduction* (Callister y Rethwisch, 2010), se emplea como bibliografía de estudio en las cátedras *Materiales I y II* dictadas en las carreras de Ingeniería Aeronáutica, Ingeniería Mecánica e Ingeniería Mecánica Electricista, en la asignatura *Materiales Ferrosos y No*

*Ferrosos* en la Tecnicatura de Mecánica Electricista y en la cátedra *Metalurgia* correspondiente a la carrera de Ingeniería Química; todas carreras que ofrece la FCEfyN de la Universidad Nacional de Córdoba.<sup>48</sup>

Dos criterios fundamentales guiaron el diseño y la construcción del presente corpus: el balance y la representatividad (Biber, 2008; Hunston, 2002; Kennedy, 1998; Sinclair, 1991, 2005). Debido a que el CIILTU fue planeado como un corpus especializado en términos de género, registro, dominio, subdominios y lengua, el balance del corpus estuvo dado por su estructuración interna homogénea. La decisión metodológica de indagar sólo dos subdominios de la ingeniería, ciencia de los materiales y mecánica de los fluidos, obedeció principalmente a dos motivos. En primer lugar, las áreas temáticas seleccionadas proporcionan nociones esenciales en diversas ramas de la ingeniería. En segundo lugar, el análisis cualitativo de las palabras de alta frecuencia, como se explica más adelante en el capítulo, se basó en criterios semánticos y pragmáticos que implicaron observar de manera detallada los significados que activaban estas palabras en los textos, labor que se realizó de manera manual. Por ello, delimitar el corpus a dos subdominios nos permitió sistematizar con mayor grado de certeza la clasificación semántico-pragmática de las palabras de alta frecuencia. A fin de compilar un corpus que fuera representativo, se consideraron parámetros centrales propuestos por distintos autores (Biber, 2008; C. Meyer, 2004; Pearson, 1998; Sinclair, 1991, 2005): disponibilidad de textos en formato electrónico, período de publicación, extensión de los textos, uso de ejemplares textuales completos y filiación de los autores.

En relación a la disponibilidad de los ejemplares textuales en formato electrónico, los dos libros de texto universitarios que conforman el CIILTU están disponibles en formato "pdf". En lo relativo al criterio de período, los libros seleccionados fueron publicados en el mismo lustro, años 2010 y 2011. Esto garantizó que los resultados hallados reflejaran los usos del léxico de la ingeniería en el género objeto de análisis en un determinado período. Con respecto a la extensión de los textos, se siguió lo propuesto por Sinclair (2005), quien observa que, desde un punto de vista lingüístico, no se justifica compilar textos de idéntico tamaño, sino garantizar que los textos seleccionados sean auténticos y reflejen la extensión real. La extensión de los capítulos de ambos libros varía entre 19.094 y 4.937 palabras. En lo concerniente a la filiación de los escritores, se

---

<sup>48</sup> Esta información se recolectó mediante entrevistas informales con docentes de las cátedras mencionadas y a través de la consulta de la bibliografía consignada en los programas de dichas asignaturas.

eligieron libros de texto cuyos autores se desempeñan en instituciones académicas de un país de habla inglesa, tal lo especificado en el prefacio de los libros.

Con el objetivo de realizar una selección textual fundada, se consideraron además las características prototípicas del libro de texto universitario, en especial aquellas referidas al propósito comunicativo del género y a su contexto de producción, recepción y circulación. Los dos libros de texto que se eligieron para compilar el CIILTU comparten, en su condición de representantes de un género de naturaleza mixta (ver sección 2.2.3), un propósito comunicativo intrínsecamente dual: por un lado, transmitir conocimiento especializado sobre una esfera disciplinar, en este caso, la ingeniería, y por otro lado, facilitar el acceso de estudiantes universitarios a un conjunto de saberes, métodos, técnicas, procedimientos y valores propios de esta disciplina. En cuanto a los productores textuales, estos dos libros de texto fueron escritos en total por seis autores que se desempeñan como profesores (dos de ellos eméritos y uno decano asociado) e investigadores en diferentes universidades de los Estados Unidos, todos con una extensa y reconocida trayectoria tanto profesional como académica en el campo de la ingeniería. Ambos textos están destinados (según lo explicita cada uno en su prefacio) principalmente a estudiantes universitarios que se aproximan a estos subdominios de la ingeniería por primera vez. También contemplan entre su audiencia a profesores de esta esfera disciplinar, a quienes les ofrecen diversos recursos, materiales y sugerencias para sus clases. Es así que el contexto natural de circulación de los libros de texto seleccionados es el ámbito universitario, donde son empleados por estudiantes y docentes de ingeniería dentro y fuera de los Estados Unidos. A continuación, la Tabla 3.1 sintetiza la información proporcionada sobre el corpus diseñado y compilado para la presente investigación (adaptada de Biber, 2008).

Tabla 3.1.

*Descripción del corpus*

Caracterización	CIILTU
Género	libro de texto universitario
Registro	académico-científico
Dominio	ingeniería
Subdominio	ciencia de los materiales y mecánica de los fluidos
Lengua	inglés
Variedad dialectal	inglés americano
Período (fecha de publicación)	2010-2011
Canal	escrito
Tamaño	402.505 palabras (33 capítulos o textos)
Formato	publicado
Autores	6 profesores/investigadores especialistas en las respectivas áreas
Lectores	estudiantes y docentes universitarios
Contexto de circulación	circulación restringida al ámbito universitario
Naturaleza de los contenidos	información fáctica
Propósitos	informar e instruir

**3.1.1.1 Preparación de los textos para su análisis**

Como punto de partida para la construcción del corpus, se dispuso de los libros de texto seleccionados en formato "pdf". Cada capítulo de ambos libros se copió y guardó en archivos separados de *Microsoft Word* (extensión "docx"). Los textos correspondientes a los treinta y tres capítulos se prepararon de modo que fueran legibles y pudieran ser procesados por el programa *WordSmith Tools 7.0* (Scott, 2016). Este proceso implicó la eliminación de página de legales, dedicatorias, agradecimientos, información personal de los autores, prefacio, tablas de contenidos, listas de abreviaturas y de símbolos, títulos de tablas y figuras, epígrafes, palabras clave y frases destacadas provistas en los márgenes, superíndices, notas a pie de página, resúmenes de ecuaciones y referencias bibliográficas de los capítulos, como así también la remoción de ilustraciones, fotos, tablas, gráficos y figuras. Por último, las versiones "docx" se guardaron como archivos de texto sin formato

(extensión "txt") para ser procesadas por el programa. En las versiones "txt" se etiquetaron direcciones de páginas de internet para que no fueran incluidas en el conteo de palabras. Finalmente, cada archivo se rotuló con un código alfanumérico para facilitar su identificación en el corpus (por ej., *I-FM*).

### 3.1.2 El programa informático de análisis de corpus

Como se adelantara en la sección anterior, para el análisis lingüístico del corpus se utilizó el paquete de herramientas *WordSmith Tools 7.0* (Scott, 2016). Este programa cuenta con distintas aplicaciones que permiten, entre otras acciones, localizar e identificar palabras en un corpus, realizar conteos, generar listas de palabras y proporcionar una multiplicidad de datos estadísticos (Kennedy, 1998). En este estudio se usaron cuatro de las herramientas que ofrece el programa, a saber, *WordList*, *Match List*, *Auto-join Tool* y *Concord*.

Mediante la herramienta *WordList* se generó automáticamente una lista de palabras que hizo posible establecer el tamaño del corpus, representado por el número total de casos (*tokens*), y el tamaño del vocabulario del corpus, expresado por el número total de tipos de palabras (*types*), como así también determinar la frecuencia de ocurrencia de cada tipo de palabra. Esta herramienta se empleó además para obtener otros datos estadísticos: el rango de las palabras y la relación estandarizada entre tipos y casos (conocida por su sigla en inglés, *STTR -standardized type-token ratio*). El rango de las palabras (también llamado consistencia) muestra la distribución de las palabras en los distintos textos del corpus, mientras que la relación estandarizada entre tipos y casos mide la frecuencia promedio de cada tipo de palabra respecto al número total de casos y se constituye en un indicador del grado de variación léxica del corpus (cfr. Capítulo 2, sección 2.4.5.1). La herramienta *Match List* se utilizó para agrupar los diferentes tipos de palabras del corpus según listas existentes en la literatura (ver aquí 3.1.3). Esta herramienta correlaciona las palabras de una lista con otra lista determinada, y se obtiene como resultado una lista de correspondencias marcadas que contiene aquellas palabras comunes a ambas listas, o una lista de correspondencias no marcadas que incluye aquellas palabras que aparecen en una lista pero no en la otra. Con la herramienta *Auto-Join* se lematizaron todas las palabras de alta frecuencia del corpus. El programa agrupa de manera automática las diferentes formas flexionadas de un mismo lema (cfr. Capítulo 2, sección 2.4.4). Finalmente, a través de la herramienta *Concord* se examinaron las palabras de alta frecuencia en su contexto de uso para establecer sus significados más recurrentes en el CIILTU.

### 3.1.3 Listas de palabras

Para la clasificación de las palabras del corpus, se utilizaron tres listas de palabras disponibles en la literatura: la General Service List -*GSL* (West, 1953), la Academic Word List -*AWL* (Coxhead, 2000), y una lista de palabras gramaticales, de aquí en adelante denominada -*GrWL*, que fuera extraída en el año 2004 de la página institucional de Paul Nation en la Universidad de Victoria<sup>49</sup>.

La *General Service List* y la *Academic Word List* son listas de palabras construidas sobre la base de datos estadísticos de frecuencia y rango. La lista *GSL*, creada por West (1953), comprende las 2.000 familias de palabras más frecuentes en inglés e incluye palabras léxicas y palabras gramaticales que tienen una alta probabilidad de aparecer en un amplio rango de usos de la lengua y en diferentes tipos de textos (Coxhead y Nation, 2001; Nation, 2001b). Las 2.000 familias de palabras más frecuentes en inglés cubren usualmente el 80% de las palabras de un texto. Cabe aclarar que, en función de los objetivos del presente estudio, se usó una versión de la *GSL* sin palabras gramaticales y se recurrió a una lista separada que está conformada exclusivamente por palabras gramaticales, como se mencionó anteriormente. La lista *AWL*, creada por Coxhead (2000), consiste en 570 familias de palabras y cubre usualmente entre el 8,5 y el 10% de las palabras de un texto. Esta lista incluye aquellas palabras que son recurrentes en un amplio rango de textos académicos de diferentes disciplinas, excluyendo las 2.000 familias de palabras más frecuentes (Nation, 2001a). Por último, la lista de palabras gramaticales *GrWL* empleada en nuestra investigación contiene 308 palabras funcionales: pronombres, determinantes, verbos auxiliares, preposiciones, conjunciones coordinantes y subordinantes, y 63 números en "forma alfabética" (Biber, Johansson, Leech, Conrad, y Finegan, 2000, p. 279). La inclusión en esta lista de un conjunto restringido de 63 números (ordinales y cardinales), desde el cero al diecinueve y del veinte al millón en decenas, está fundada teóricamente ya que los mismos constituyen un grupo reducido de palabras gramaticales, categorizadas como determinantes, que pueden ser utilizadas para formular un conjunto ilimitado de formas más complejas que representan distintos tipos de expresiones numéricas (Biber et al., 2000).

---

<sup>49</sup> <http://www.victoria.ac.nz/lals/about/staff/paul-nation>

### 3.2 Recolección y análisis de los datos

Una primera exploración del corpus proporcionó una descripción general del vocabulario de los libros de texto universitarios de ingeniería escritos en inglés a partir de los siguientes datos estadísticos: número total de casos y de tipos de palabras, relación estandarizada entre tipos y casos, rango o consistencia y cobertura de las listas de palabras. Luego se llevó a cabo un análisis cuantitativo y cualitativo más pormenorizado del corpus, con el propósito de identificar y clasificar las palabras de alta frecuencia del CIILTU. El análisis cuantitativo arrojó información sobre la frecuencia y cobertura de los distintos tipos de palabras de alta frecuencia, mientras que el análisis cualitativo implicó la interpretación de los datos cuantitativos recolectados y la formulación de una propuesta de clasificación y subclasificación de las palabras de alta frecuencia basada en la combinación de criterios semánticos (significado) y pragmáticos (uso) y el asesoramiento de especialistas en el campo disciplinar.

#### 3.2.1 Descripción léxica general del corpus

La lista de palabras generada automáticamente, a través de la herramienta *WordList* del programa descrito en la sección 3.1.2, reveló información sobre el número total de tipos y casos, el rango de las palabras y la relación estandarizada entre tipos y casos en el CIILTU. Por las razones explicadas en el Capítulo 2 (cfr. sección 2.4.5.1), se calculó la relación estandarizada entre tipos y casos (STTR) dado que la relación entre tipos y casos (TTR) es sensible al tamaño del corpus, particularmente si este contiene textos de diferente extensión, como es el caso de los capítulos de los libros de texto representados en el CIILTU. Se optó por tomar como base para el cálculo la opción por defecto del programa, esto es computar la relación entre tipos y casos cada 1.000 palabras en el corpus.

Posteriormente, utilizando las tres listas de palabras presentadas en la sección 3.1.3 y la herramienta *Match List* del programa, se clasificaron las palabras del corpus en palabras generales, palabras académicas y palabras gramaticales. Las palabras restantes, es decir aquellas que no se correspondían con ninguna de las listas empleadas, se agruparon en una cuarta lista, a la que se denominó "Otras palabras". Se procedió luego a determinar, por separado, la cobertura de las cuatro categorías de palabras calculando el porcentaje de casos que cada lista cubre en el corpus (Nation, 2001a, 2001b).

### **3.2.2 Análisis cuantitativo de las palabras de alta frecuencia en el corpus**

#### **3.2.2.1 Identificación de las palabras de alta frecuencia**

Como paso previo a la identificación de las palabras de alta frecuencia en el corpus, se eliminaron de la lista general los números escritos en forma digital y las palabras gramaticales. La exclusión de ambos elementos tiene un fundamento teórico-metodológico: evitar que la alta frecuencia tanto de los números como de las palabras gramaticales en el corpus distorsione los resultados. Específicamente, las palabras gramaticales conforman un grupo relativamente estable, en oposición a las palabras léxicas o de contenido que integran un conjunto que cambia de manera constante. Esto puede explicar, entre otras razones, por qué las palabras gramaticales exhiben una alta frecuencia de aparición en una diversidad de textos, mientras que las palabras léxicas tienen una ocurrencia estrechamente relacionada con la temática que abordan los textos (Biber et al., 2000).

Respecto al criterio adoptado para identificar las palabras más frecuentes en el corpus, Nation (2005) sostiene que la línea divisoria entre las palabras de alta frecuencia y las de baja frecuencia es trazada de manera arbitraria por el investigador. Así pues, en nuestro estudio, el punto de corte fue establecido mediante el cálculo de la mediana, al igual que en estudios previos sobre la temática (Muñoz, 2015). Se tomó la decisión de estimar la mediana, en lugar de la media, dado que esta medida estadística muestra tendencias centrales que resumen datos con valores muy dispares, en contraste con la media, que los distorsiona (Baroni, 2009; Levine y Stephan, 2010). Se calculó la mediana sumando las frecuencias de todas las palabras de la lista y luego dividiendo el total por dos. El valor medio obtenido permitió identificar en la lista una palabra que representaba el centro de distribución que separaba las palabras en dos grupos: palabras de alta frecuencia y de baja frecuencia. Luego, se seleccionaron las palabras de alta frecuencia, es decir, aquellas palabras cuya frecuencia era igual o mayor a la frecuencia de la palabra que sirvió como punto de corte.

#### **3.2.2.2 Clasificación y cobertura de las palabras de alta frecuencia**

Para determinar la cobertura de las palabras de alta frecuencia, se llevó a cabo el mismo procedimiento que se utilizó para la cobertura de las listas de palabras en la primera etapa de la investigación. Se hizo uso nuevamente de las listas de palabras disponibles en la literatura, la General Service List -*GSL* (West, 1953), la Academic Word List -*AWL*

(Coxhead, 2000), y la lista de palabras gramaticales *-GrWL*. También se utilizó la herramienta *Match List* para clasificar las palabras de alta frecuencia en tres listas: palabras generales, palabras académicas y Otras palabras. Una vez obtenidas las listas, se calculó la cobertura de cada una de ellas.

### **3.2.2.3 Lemas de las palabras de alta frecuencia**

La última etapa del análisis cuantitativo consistió en agrupar las palabras de alta frecuencia en lemas. Para ello, se adoptó la noción de lema propuesta por Nation (2001a), quien sostiene que un lema se compone de una palabra base o núcleo, sus formas flexionadas, que usualmente pertenecen a la misma categoría gramatical, y sus formas reducidas o contraídas (cfr. Capítulo 2, sección 2.4.4). Para la lematización automática de las palabras de alta frecuencia, se recurrió a la aplicación *Auto-Join* del programa *WordSmith Tools*. Al detectarse algunas imprecisiones, se llevaron a cabo revisiones y se realizaron ajustes de manera manual. Una vez finalizado el proceso, se seleccionó como palabra núcleo de cada lema a aquella con mayor frecuencia en el corpus (Sinclair, 1991).

### **3.2.3 Análisis cualitativo de las palabras de alta frecuencia en el corpus**

La primera etapa del análisis cualitativo de las palabras de alta frecuencia consistió en realizar una reclasificación de los lemas en dos categorías, palabras técnicas y palabras no técnicas, a partir de sus significados y usos específicos en el corpus. Para clasificar las palabras técnicas, se siguieron los criterios semánticos y pragmáticos postulados por Cabré (1999), Muñoz (2015) y Sager (1990). Para categorizar las palabras no técnicas, se adoptaron los criterios semánticos y pragmáticos propuestos por P. Meyer (1997) y Muñoz (2015). Sobre la base de dichos criterios y mediante el uso de la herramienta *Concord* del programa *WordSmith Tools*, se observaron y analizaron las palabras en su contexto real de aparición con el propósito de identificar sus posibles usos y significados.

La segunda etapa del análisis cualitativo implicó llevar a cabo una subclasificación de las palabras técnicas y no técnicas de alta frecuencia en el corpus. Se emplearon las categorías formuladas por Cabré (1999), Muñoz (2015) y Sager (1990) para subclasificar las palabras técnicas, mientras que para las palabras no técnicas se utilizaron las categorías diseñadas por P. Meyer (1997) y Muñoz (2015). En ambos casos, fue necesario crear categorías ad hoc para las palabras que no estaban comprendidas en ninguna de las categorías existentes.

Para la realización de ambas etapas del análisis cualitativo fue necesario recurrir al asesoramiento de informantes clave<sup>50</sup>, uno para cada subdominio de la ingeniería abordado en este estudio. Por *informante clave*<sup>51</sup> se entiende a aquella persona cuya posición en un determinado ámbito le confiere conocimiento especializado sobre una temática (Payne y Payne, 2004) y un profundo entendimiento de su contexto, por lo que se convierte en una fuente primaria de información en una investigación (Taylor, Bogdan, y DeVault, 2016). Los informantes clave, también denominados *testigos expertos*, se diferencian de los informantes generales en al menos dos aspectos: disponen de mayor información y generalmente gozan de un rol más visible en su contexto por ocupar posiciones formales de autoridad (Payne y Payne, 2004). Dada la naturaleza del presente trabajo, se consultaron dos informantes clave. En su carácter de especialista en el área de mecánica de los fluidos, se recurrió a un ingeniero mecánico aeronáutico<sup>52</sup> (Informante Clave 1) que actualmente se desempeña como profesor titular en la Facultad de Ciencias Exactas, Físicas y Naturales de la Universidad Nacional de Córdoba. Como especialista del área de ciencia de los materiales, se consultó a un ingeniero metalúrgico<sup>53</sup> (Informante Clave 2) que es en la actualidad profesor titular plenario de la Facultad de Ciencias Exactas, Físicas y Naturales en la Universidad Nacional de Córdoba. Con el Informante Clave 1 se llevaron a cabo

---

<sup>50</sup> Nuestra traducción del término *key informant* (Bloor y F. Wood, 2006; Payne y Payne, 2004; Taylor et al., 2016).

<sup>51</sup> El uso de *informantes clave* es un procedimiento típico en los métodos de investigación cualitativa de las ciencias sociales. Los informantes clave le ofrecen al investigador su visión y perspectiva de los procesos sociales objeto de estudio (Bloor y F. Wood, 2006; Payne y Payne, 2004; Taylor et al., 2016). Como nuestra investigación se centra en un proceso social determinado, la comunicación de conocimiento científico ya consolidado en libros de texto universitarios, en una disciplina específica, la ingeniería, en dos subdominios, la ciencia de los materiales y la mecánica de los fluidos, decidimos solicitar el asesoramiento de informantes clave para el análisis de los datos recolectados. En concordancia con otros autores (Atkinson, 2004; Hyland, 2000, 2005; Muñoz, 2017), concebimos a las disciplinas como *culturas*, definidas como grupos sociales cohesivos integrados por miembros que comparten objetivos comunes, prácticas y acciones rutinarias y competencias desarrolladas en contextos específicos, como por ejemplo entornos académicos e institucionales. Estas culturas están estrechamente relacionadas con la noción de *comunidad discursiva*, entendida como una red socio-retórica creada por un grupo de personas para interactuar y alcanzar metas comunes mediante el uso de géneros específicos (Swales, 1990).

<sup>52</sup> El Ing. Sergio Preidikman es Doctor en Ingeniería Mecánica. Se desempeña como profesor titular de las cátedras Cálculo Estructural I y III de la Facultad de Ciencias Exactas, Físicas y Naturales (UNC), y como profesor investigador invitado de *Virginia Polytechnic Institute and State University*. Es investigador independiente del Instituto de Estudios Avanzados en Ingeniería y Tecnología (IDIT-CONICET). Además, es miembro de la *American Academy of Mechanics*, el *American Institute of Aeronautics and Astronautics*, la *American Association for Wind Engineering* y la Asociación Argentina de Mecánica Computacional (información presentada en esta tesis con la autorización del Dr. Preidikman).

<sup>53</sup> El Ing. Carlos Rodolfo Oldani es Doctor en Ciencias de la Ingeniería. Se desempeña como profesor titular plenario de las cátedras Materiales I y Biomateriales de la Facultad de Ciencias Exactas, Físicas y Naturales (UNC) y es director del Laboratorio de Investigación y Desarrollo en Materiales del Departamento de Materiales de la mencionada facultad (información presentada en esta tesis con la autorización del Dr. Oldani).

cuatro reuniones en las que se analizaron y categorizaron las palabras técnicas de alta frecuencia en el CIILTU referidas a mecánica de los fluidos, se realizaron modificaciones en las categorías de análisis previstas, y se seleccionaron los elementos más representativos para las distintas categorías. Con el Informante Clave 2 se mantuvo un encuentro en el que fue posible analizar y clasificar las palabras técnicas de alta frecuencia relativas a ciencia de los materiales y se eligieron los elementos prototípicos para cada una de las categorías empleadas.

Como producto final del análisis cualitativo se construyeron dos listas que contienen el vocabulario característico de los libros de texto universitarios en inglés en el campo de la ingeniería, particularmente en las áreas de ciencia de los materiales y mecánica de los fluidos. Una lista está integrada por el vocabulario técnico del corpus y comprende palabras que están relacionadas a nivel semántico dado que representan conceptos de la ingeniería en cuanto dominio especializado. La otra lista consiste en las palabras no técnicas del corpus e incluye aquellas palabras que son usadas para describir y establecer relaciones entre palabras especializadas en esta esfera disciplinar, y palabras que hacen referencia a la investigación y al quehacer científico-académico en general y a la práctica de la ingeniería en tanto profesión.

### **3.3 Síntesis del capítulo**

En este capítulo hemos detallado los materiales (corpus, listas de palabras empleadas, programa informático y herramientas utilizadas) y los procedimientos usados para la recolección de los datos y el análisis del vocabulario mediante la combinación de técnicas cuantitativas y cualitativas. En el capítulo que sigue nos centraremos en los resultados obtenidos y presentaremos un esquema de clasificación y subclasificación de las palabras técnicas y no técnicas de alta frecuencia relevadas en el corpus, que se basa en el asesoramiento proporcionado por los informantes clave y la integración de criterios semánticos y pragmáticos.

## Capítulo 4

### Resultados

Este capítulo tiene como objetivo central presentar los resultados del análisis llevado a cabo sobre el vocabulario de libros de texto universitarios escritos en inglés correspondientes al dominio de la ingeniería. El capítulo se estructura en dos secciones principales: una descripción general del vocabulario del corpus, y un análisis detallado, cuantitativo y cualitativo, de las palabras de alta frecuencia identificadas en el corpus. En la sección 4.1 se proporciona información sobre el tamaño del corpus, el tamaño del vocabulario del corpus, la relación estandarizada entre los tipos de palabras y el número total de casos, el rango de las palabras y la cobertura de las diferentes listas de palabras utilizadas en el estudio. Por su parte, en la sección 4.2 se ofrece la frecuencia, cobertura y lemas de las palabras de alta frecuencia categorizadas según las listas propuestas en la literatura. Seguidamente, se propone una reclasificación de las palabras de alta frecuencia relevadas a partir de la combinación de criterios semánticos y pragmáticos y el asesoramiento proporcionado por especialistas en la disciplina.

#### 4.1 Descripción léxica general del corpus

##### 4.1.1 Tamaño del corpus y tamaño del vocabulario del corpus

La lista generada con el programa informático *WordSmith Tools 7.0* nos permitió obtener información acerca del tamaño del corpus, medido a base del número total de casos (*tokens*), y del tamaño del vocabulario del corpus, calculado a base del número total de los distintos tipos de palabras (*types*) en el corpus. La lista mostró que el CIILTU contiene 402.505 casos y 11.221 distintos tipos de palabras.<sup>54</sup> Dado el tamaño del corpus construido para el presente estudio, cabe señalar que es un corpus pequeño (Hunston, 2002; Sinclair, 1991), muy apropiado para el análisis cualitativo del vocabulario de alta frecuencia (Hunston, 2002; Kennedy, 1998; Koester, 2010), uno de los objetivos principales de este estudio.

---

<sup>54</sup> Tal como se señalara en el Capítulo 2, sección 2.4.4, empleamos las expresiones "casos" y "tipos de palabras" para referirnos a los términos en inglés "tokens" y "types", respectivamente.

#### **4.1.2 Relación entre tipos de palabras y número total de casos en el corpus**

Debido a que el CIILTU está conformado por 33 capítulos de distinta extensión, se optó por calcular la relación estandarizada entre los tipos de palabras y el número total de casos (cfr. Capítulo 3, sección 3.2.1). Tal lo calculado por el programa, la relación tipos de palabras-número de casos en nuestro corpus es de 33,75%, lo que indica que hay un promedio aproximado de 34 tipos de palabras nuevas cada 100 palabras (casos) presentes en el corpus. Esta frecuencia promedio revela una baja variabilidad de tipos de palabras o baja variación del vocabulario en el CIILTU, es decir, una alta repetición de palabras en el corpus.

#### **4.1.3 Rango de los tipos de palabras**

El rango o consistencia de los distintos tipos de palabras se determinó en función de su distribución a lo largo de los 33 textos (capítulos de libros) que integran el corpus (Tabla 4.1). Se observó que, del total de tipos de palabras en el corpus (11.221), solo 117 aparecieron en todos los textos. De estos tipos, la mayoría, 71, son palabras gramaticales, lo que demuestra la aparición frecuente de las palabras gramaticales en el corpus. Además, se identificaron 860 tipos de palabras que tienen un rango igual o mayor a 17, es decir que se encuentran en más de la mitad de los textos del corpus, predominando en este rango las palabras léxicas, dado que se identifican solamente 127 palabras gramaticales. Tal como se muestra en la Tabla 4.1, es posible advertir que en el rango de 32 a 29 textos, la distribución de las palabras gramaticales y de las palabras restantes no es balanceada, con un 18% y un 82% respectivamente. Esta diferencia se acentúa en el rango más bajo, de 4 textos a 1 texto, pues la distribución se vuelve extremadamente desigual, ya que las palabras gramaticales representan solo un 0,4%, mientras que el resto de las palabras constituyen un 99,6% sobre el total de palabras en dicho rango.

Tabla 4.1.

*Rango de los tipos de palabras en el corpus*

Rangos	Palabras gramaticales	Palabras restantes	Número total de tipos de palabras en cada rango
<b>33</b>	71	47	118
	60,2%	39,8%	100%
<b>32 a 29</b>	26	119	145
	17,9%	82,1%	100%
<b>28 a 25</b>	12	140	152
	7,9%	92,1%	100%
<b>24 a 21</b>	10	194	204
	4,9%	95,1%	100%
<b>20 a 17</b>	8	233	241
	3,3%	96,7%	100%
<b>16 a 13</b>	14	352	366
	3,8%	96,2%	100%
<b>12 a 9</b>	9	597	606
	1,5%	98,5%	100%
<b>8 a 5</b>	13	1.181	1.194
	1,1%	98,9%	100%
<b>4 a 1</b>	32	8.163	8.195
	0,4%	99,6%	100%
<b>Número total de tipos de palabras en el corpus</b>	<b>195</b>	<b>11.026</b>	<b>11.221</b>

Por otra parte, es interesante observar individualmente, por un lado, la distribución del total de palabras gramaticales y, por otro lado, la distribución del total del resto de las palabras en el corpus. La Figura 4.1 se focaliza en las palabras gramaticales e ilustra cómo el porcentaje del total de este tipo de palabras se concentra en el primer rango, de 33 textos, donde aparecen alrededor de un 36,5% del total de palabras gramaticales del corpus (195). Luego, la presencia de palabras gramaticales disminuye de manera muy marcada y finalmente aumenta levemente en el último rango (4-1 textos), donde aparecen aproximadamente el 16,5% del total de palabras gramaticales del corpus. Por su parte, la Figura 4.2 se centra en las palabras restantes del corpus, que exhiben un comportamiento inverso al de las palabras gramaticales. En el rango superior (33 textos), el resto de las palabras representa aproximadamente un 0,5% en relación al porcentaje total de este tipo de palabras en el corpus (11.026). En los rangos siguientes el porcentaje incrementa de

manera muy paulatina hasta alcanzar alrededor de un 11% en el rango 8-5 textos. En el rango inferior, de 4 a 1 textos, esta categoría de palabras evidencia un aumento muy significativo ya que representa un 74% sobre el total de este tipo de palabras.

A modo de síntesis, es posible observar que las palabras de alto rango en el corpus son en su gran mayoría palabras gramaticales, mientras que las palabras con rango bajo están representadas en general por el resto de las palabras del corpus.

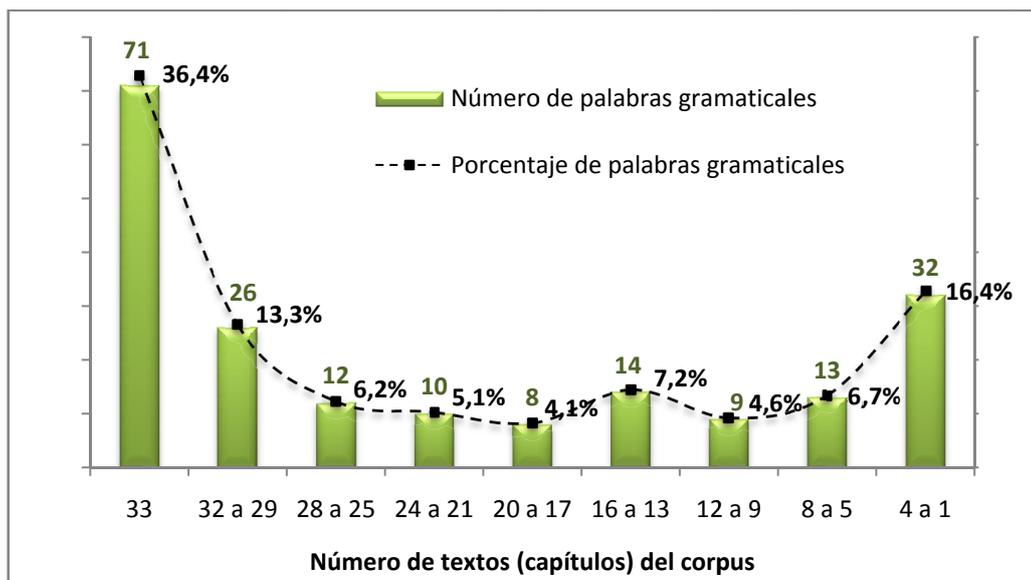


Figura 4.1. Rango de las palabras gramaticales en el corpus

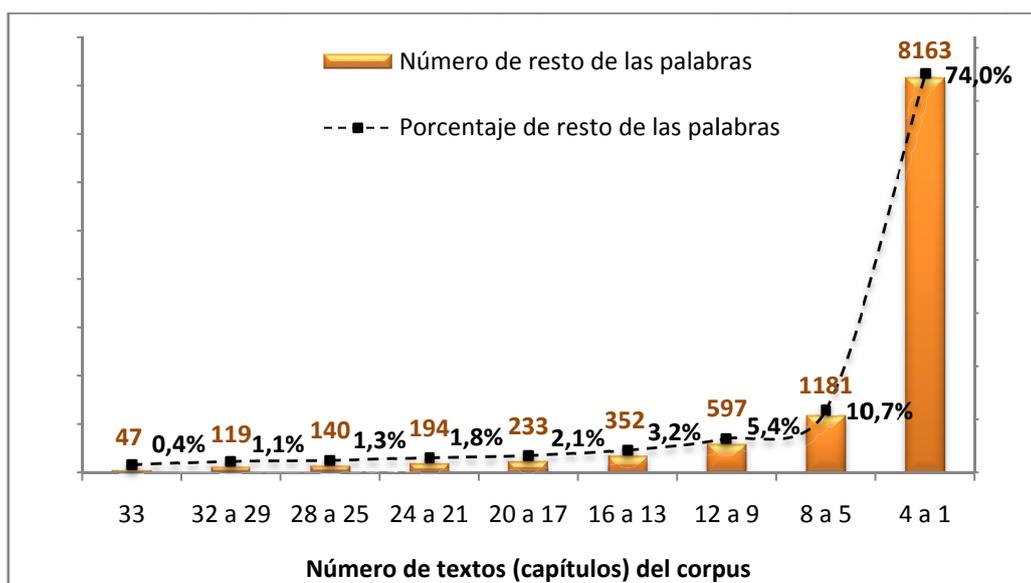


Figura 4.2. Rango de las palabras restantes en el corpus

#### 4.1.4 Clasificación de los tipos de palabras del corpus

Para realizar la clasificación de los tipos de palabras presentes en el corpus se emplearon tres listas de palabras disponibles en la literatura: la *General Service List -GSL* (West, 1953), la *Academic Word List -AWL* (Coxhead, 2000) y una lista de palabras gramaticales -*GrWL*, extraída en el año 2004 de la página institucional de Paul Nation en la Universidad de Victoria (cfr. Capítulo 3, sección 3.1.3). Mediante el uso de la herramienta *Match List* del programa informático *WordSmith Tools*, se identificaron en el corpus 3.183 tipos de palabras generales, es decir palabras que pertenecen a la lista de palabras generales *GSL* (por ej., *example*, *rate* y *use*); 1.406 tipos de palabras académicas, esto es, que aparecen en la lista de palabras académicas *AWL* (por ej., *data*, *function* y *required*); y 195 tipos de palabras gramaticales que integran la lista de palabras gramaticales *GrWL* (por ej., *the*, *of* y *and*). Los 6.437 tipos de palabras restantes se agruparon en una cuarta lista denominada Otras palabras. Esta última lista está formada principalmente por palabras específicas de la disciplina, abreviaturas, siglas, acrónimos y símbolos (por ej., *ft*, *laminar*, *polymers* y *MPa*). Como se observa en la Figura 4.3, más de la mitad de los tipos de palabras del corpus quedan comprendidos en la lista de Otras palabras, superando en número a las palabras generales, académicas y gramaticales en conjunto.

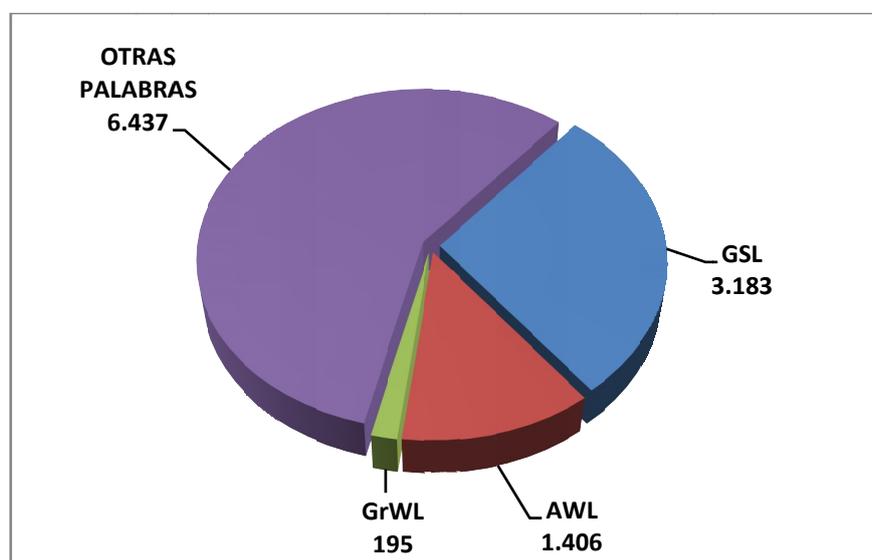


Figura 4.3. Número de tipos de palabras en el corpus

#### 4.1.5 Cobertura de los distintos tipos de palabras

Una vez clasificadas las palabras en las cuatro listas (*GSL*, *AWL*, *GrWL* y *Otras palabras*), se procedió a calcular la cobertura proporcionada por cada una de estas categorías en el CIILTU. Se encontró que las 3.183 palabras generales tienen una cobertura

de alrededor del 27% de la totalidad de palabras del corpus, las 1.406 palabras académicas de aproximadamente el 9%, las 195 palabras gramaticales del 45% y las restantes 6.437 que pertenecen a la lista de Otras palabras de alrededor del 19%. Se observó que las palabras generales y las palabras gramaticales combinadas cubren el 72% del corpus, porcentaje que se encuentra muy cercano al 75-80% de cobertura sugerido usualmente en la literatura (Coxhead y Nation, 2001; Nation, 2001a, 2001b). Por otra parte, al observar el porcentaje acumulado de cobertura de las palabras generales, gramaticales y académicas se obtiene una cobertura del 81%, es decir de la gran mayoría del corpus. La Tabla 4.2 sintetiza los tipos, casos y cobertura de las cuatro listas.

Tabla 4.2.

*Número de tipos de palabras, casos y cobertura de las cuatro listas de palabras*

LISTAS DE PALABRAS	TIPOS	CASOS	COBERTURA
GSL	3.183	109.189	27,13%
GrWL	195	181.661	45,13%
AWL	1.406	37.136	9,23%
OTRAS PALABRAS	6.437	74.519	18,51%
<b>TOTAL</b>	<b>11.221</b>	<b>402.505</b>	<b>100%</b>

La relación entre los tipos de palabras y su cobertura en el corpus se ilustra de manera más evidente en la Figura 4.4. El eje vertical a la izquierda muestra el número de tipos de palabras en tanto que el eje vertical a la derecha representa la cobertura que estas proporcionan. En este gráfico se puede advertir una marcada diferencia entre la lista de palabras gramaticales y la lista de Otras palabras. Mientras un número muy reducido de palabras gramaticales, 195, proporciona una cobertura del 45% del CIILTU, un número visiblemente mayor de palabras correspondientes a la categoría de Otras palabras, más de 6.400, dan una cobertura significativamente más baja, cercana al 19% del corpus. Esto se debe a que las palabras gramaticales conforman un grupo relativamente estable que tiene una alta probabilidad de aparecer en una amplia variedad de textos (Biber et al., 2000). Con respecto a la cobertura de las palabras académicas, cabe destacar que el porcentaje que proporcionan en el corpus coincide con el 9% relevado en estudios previos (Nation, 2001a, 2001b; Martínez et al, 2009; Valipouri y Nassaji, 2013) que indagaron en géneros y

disciplinas diferentes y se encuentra próximo al 10% informado en otras investigaciones (Chen y Ge, 2007).

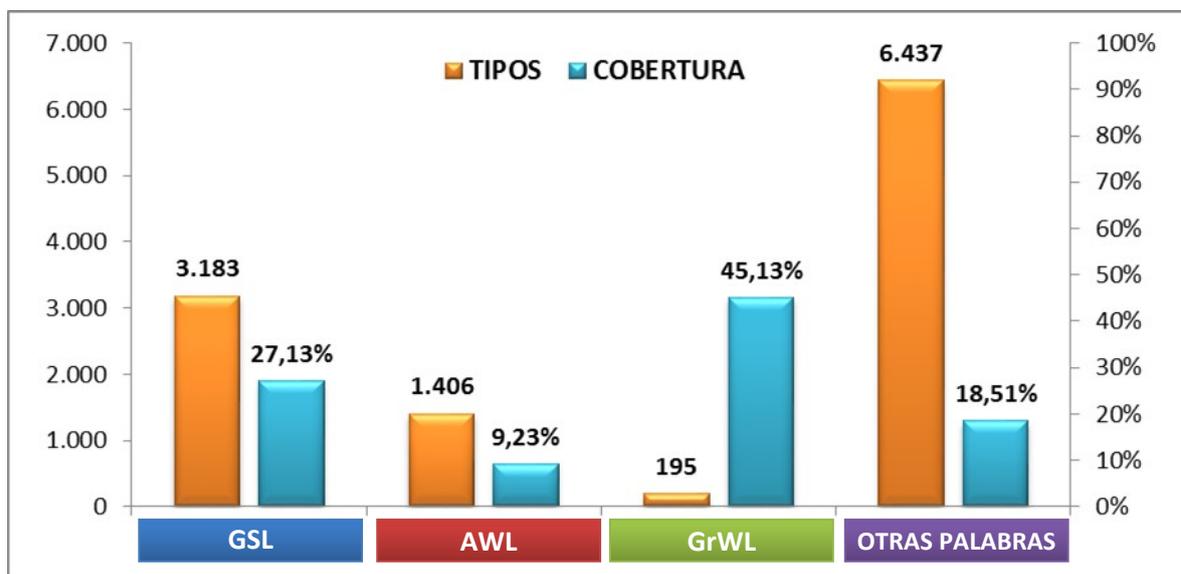


Figura 4.4. Relación entre tipos de palabras y su cobertura en el corpus

## 4.2 Identificación y descripción de las palabras de alta frecuencia

### 4.2.1 Selección de las palabras de alta frecuencia

Luego de describir de manera general el vocabulario del corpus, se emprendió la segunda etapa de análisis que consistió en la identificación y clasificación de las palabras de alta frecuencia en el CIILTU de acuerdo con listas existentes en la literatura, en una primera instancia, y una nueva clasificación basada en criterios semánticos y pragmáticos, en una segunda instancia. Para lograr estos objetivos, fue necesario previamente eliminar todas las palabras gramaticales y los números en formato digital de la lista de palabras obtenida en la primera etapa de análisis, tal como se especifica en el Capítulo 3. Así, la lista de palabras del corpus se redujo a 11.026 tipos que representan 220.844 casos.

A fin de distinguir las palabras de alta frecuencia de aquellas de baja frecuencia en la nueva lista, se optó por determinar el punto de corte mediante el cálculo de la mediana, en lugar de la media que fuera usada en otros estudios de esta naturaleza (por ej., Hyland y Tse, 2007). Siguiendo a Muñoz (2015) y como se señala en el Capítulo 3, tomamos la mediana dado que es un indicador confiable de tendencia central por su capacidad para sintetizar datos con valores muy dispares, como en el caso de las frecuencias extremas que se advierten al comienzo y al final de las listas de palabras de un corpus. Para calcular esta

medida estadística, utilizando el programa *Excel*, se dividieron por la mitad los 220.844 casos de la lista de palabras de manera de obtener un valor medio, 110.422, que posibilitara segmentar en partes iguales el número de casos en la lista de palabras del corpus. Se determinó que dicho valor medio se alcanzaba por las frecuencias acumuladas de los primeros 372 tipos de la lista de palabras. Se identificó, así, que el punto de corte que separaba las palabras de alta frecuencia de las de baja frecuencia estaba dado por el tipo en la posición número 372 de la lista, con una frecuencia de 129 casos. Por lo tanto, en este estudio se consideraron palabras de alta frecuencia aquellas con una frecuencia igual o superior a 129 casos, es decir, palabras que aparecieron al menos 129 veces en el corpus.

El cálculo de la mediana, en lugar de la media, permitió seleccionar un conjunto acotado de 372 palabras de alta frecuencia, un número razonable para el posterior análisis cualitativo realizado principalmente en forma manual. Los 372 tipos de palabras de alta frecuencia representan el 3,3% del total de tipos de palabras del corpus (11.221) y tienen una cobertura del 27,5%.

#### **4.2.2 Clasificación y cobertura de las palabras de alta frecuencia**

Utilizando las listas *GSL* y *AWL*, se clasificaron las palabras de alta frecuencia previamente identificadas en palabras generales y académicas, mientras que las palabras restantes que no estaban comprendidas en una u otra lista se categorizaron como Otras palabras. En las 372 palabras de alta frecuencia, se relevaron 205 tipos de palabras generales, 61 palabras académicas y 106 correspondientes a la categoría de Otras palabras. A continuación se determinó la cobertura de cada categoría de palabras de alta frecuencia y se encontró que las palabras generales de alta frecuencia cubren el 15,4% del corpus, las palabras académicas el 4,1% y la lista de Otras palabras el 8%. En la Figura 4.5 se ilustra la relación entre los tipos de palabras de alta frecuencia y la cobertura que proveen en el corpus. El eje vertical a la izquierda indica el número de tipos de palabras en tanto que el eje vertical a la derecha señala la cobertura que estas proporcionan. Como se puede ver, las palabras generales son las predominantes dentro de las palabras de alta frecuencia y prácticamente duplican a las Otras palabras, en tanto que las palabras académicas constituyen un grupo reducido, el menos representado en las palabras de alta frecuencia del corpus. Esta tendencia también se ve reflejada en la cobertura alcanzada por cada una de las categorías de palabras de alta frecuencia, pues las palabras generales proporcionan el mayor porcentaje de cobertura del corpus, seguidas por Otras palabras y, con un porcentaje notablemente menor, por las palabras académicas.

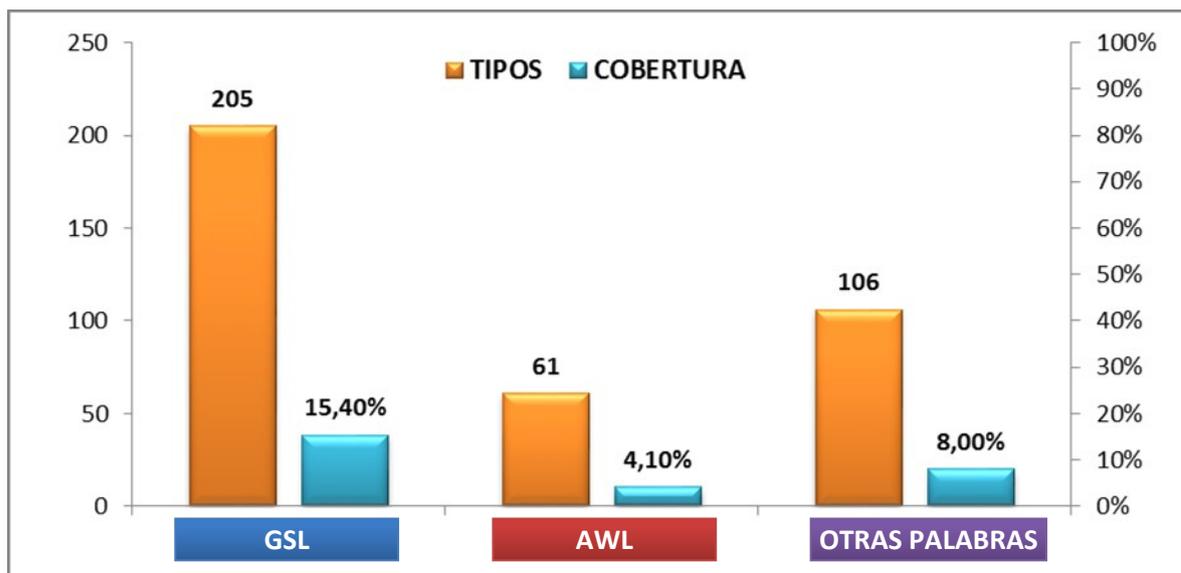


Figura 4.5. Relación entre tipos de palabras de alta frecuencia y su cobertura en el corpus

#### 4.2.3 Lemas de las palabras de alta frecuencia

Tomando el lema como unidad de análisis (Hunston, 2002; Kennedy, 1998; Nation, 2001a, 2016) y utilizando el programa *WordSmith Tools*, se lematizaron las 372 palabras de alta frecuencia del CIILTU, procedimiento que resultó en 326 lemas. Se tomó la decisión metodológica de seleccionar como palabra núcleo de cada lema a aquella con mayor frecuencia de aparición en el corpus (Sinclair, 1991). La lematización de las palabras de alta frecuencia en el CIILTU permitió reducir el número de palabras que serían sometidas a posteriori al análisis cualitativo manual.

La mayoría de los lemas de alta frecuencia (283) están compuestos por un solo miembro (por ej., *pressure* y *velocity*), y 40 de los 43 lemas integrados por más de un miembro están compuestos por dos miembros (por ej., *composite - composites*, *determine - determined*). La Tabla 4.3 muestra los diez lemas más frecuentes de palabras generales, académicas y Otras palabras, representados por sus palabras núcleo. La lista completa de los lemas de alta frecuencia se presenta en el Apéndice 1, Tabla 1.1.

Tabla 4.3.

*Los diez lemas más frecuentes de palabras generales, académicas y Otras palabras de alta frecuencia en el corpus*

Palabras generales (GSL)	Palabras académicas (AWL)	Otras palabras
1 . material	1 . equation	1 . fluid
2 . flow	2 . stress	2 . alloy
3 . temperature	3 . phase	3 . C
4 . figure	4 . section	4 . velocity
5 . pressure	5 . structure	5 . B
6 . shown	6 . energy	6 . S
7 . also	7 . volume	7 . Fig.
8 . used	8 . component	8 . atom
9 . metal	9 . obtain	9 . X
10 . surface	10 . occur	10 . polymer

Al observar de manera más detenida los lemas de alta frecuencia con mayor número de apariciones en el CIILTU, se advierte inmediatamente que un número notable de palabras categorizadas como generales y académicas, según las listas existentes en la literatura, activan en el corpus un significado directamente relacionado con la disciplina en cuestión, la ingeniería. A modo de ejemplo, palabras generales tales como *flow*, *pressure* y *metal*, y palabras académicas tales como *equation*, *stress* y *energy* expresan en los textos del corpus conceptos referidos al dominio de la ingeniería.

#### 4.2.4 Análisis cualitativo de las palabras de alta frecuencia

La identificación de numerosas palabras de alta frecuencia en el CIILTU que adquieren significados vinculados a la ingeniería, independientemente de la clasificación inicial en palabras generales, académicas y Otras palabras, llevó a plantear la necesidad de una nueva clasificación, esta vez sin recurrir a listas de palabras existentes en la literatura. Se decidió así categorizar las palabras de alta frecuencia en dos grandes grupos: técnicas (por ej., *force*, *laminar*, *load*, *psi* y *R*) y no técnicas (por ej., *assume*, *chapter*, *data*, *Fig.* y *properties*), a partir de criterios semánticos y pragmáticos propuestos por Cabré (1999), P. Meyer (1997), Muñoz (2015) y Sager (1990). Para diferenciar las palabras que activaban un significado técnico de aquellas que expresaban un significado no técnico en el corpus, se observó la palabra núcleo de cada lema en su contexto real de uso mediante la

herramienta *Concord* del programa de análisis lingüístico utilizado en nuestro estudio (cfr. Capítulo 3, sección 3.1.2). Se considera oportuno aquí hacer dos puntualizaciones. En primer lugar, el proceso de identificación y clasificación de las palabras técnicas, y su subsiguiente subcategorización se realizó con el asesoramiento de informantes clave<sup>55</sup>, esto es, especialistas en ambos subdominios (ciencia de los materiales y mecánica de los fluidos), tal como se especificara en la sección 3.2.3 del Capítulo 3. En segundo lugar, y a pesar de contar con la asistencia de especialistas, surgieron ciertas dificultades al momento de establecer categorías semántico-pragmáticas que fueran mutuamente excluyentes y de fijar límites precisos entre las distintas categorías. En respuesta a las dificultades advertidas, basamos nuestra propuesta de clasificación en los principios delineados por la *teoría de los prototipos* (Cruse, 2006; Geeraerts, 1990, 2010, 2016). Esta teoría establece como postulado fundamental que, una categoría se define a partir de un elemento núcleo o prototípico considerado el miembro más representativo de la misma, en tanto que el resto de los elementos exhiben distintos grados de centralidad o de semejanza en relación con dicho prototipo. La teoría reconoce además la existencia de elementos periféricos en el interior de las categorías y de límites difusos entre categorías.

#### 4.2.4.1 Palabras técnicas o términos

En el Capítulo 2 (sección 2.3), presentamos la definición de palabras técnicas que adoptamos en nuestro estudio, como un conjunto de términos que refieren a conceptos propios de una disciplina y que están semánticamente relacionados por ser miembros de una red conceptual particular y específica a cada esfera disciplinar (Adelstein, 2007; Cabré, 1999; Ciapuscio, 2003a; Kuguel, 2007). Una vez identificadas las palabras técnicas de alta frecuencia en el CIILTU, y en función de los conceptos que representan en la ingeniería, se subclasificaron en ocho grupos, a saber: 1) *Objetos y entidades*, 2) *Personas*, 3) *Procesos, operaciones y acciones*, 4) *Propiedades, estados y cualidades de materiales, cuerpos y sistemas*, 5) *Regiones, áreas y lugares geométricos*, 6) *Nomenclaturas químicas*, 7) *Unidades de medición*, y 8) *Símbolos*. Cabe aclarar que las categorías 1, 3 y 4 se tomaron de Cabré (1999) y Sager (1990), mientras que las categorías 2, 5, 6 y 7 se adoptaron de Muñoz (2015). Resulta oportuno señalar que a dos categorías formuladas por los mencionados autores se les realizaron modificaciones: a la categoría 4 propuesta por Cabré y Sager, *Propiedades, estados y cualidades*, se le incorporó *de materiales, cuerpos y*

---

<sup>55</sup> Nuestra traducción del término *key informant* (Bloor y F. Wood, 2006; Payne y Payne, 2004; Taylor et al., 2016).

*sistemas* a fin de representar de manera más acabada las especificidades de la ingeniería en tanto dominio disciplinar, y a la categoría 5 formulada por Muñoz, *Regiones y áreas*, se le agregó la noción de *lugares geométricos* para dar cuenta de numerosas instancias de referencias geométricas presentes en el corpus. Asimismo, es necesario precisar que la categoría 8, *Símbolos*, se creó ad hoc en base a los patrones que emergían de los datos relevados.

1) El grupo *Objetos y entidades* comprende, por un lado, palabras que hacen referencia a objetos tales como materiales, sustancias, dispositivos, artefactos y, por otro lado, a abstracciones como disciplinas, tipos de fuerzas y expresiones matemáticas<sup>56</sup>. Este grupo está representado en los textos mediante nombres concretos y abstractos. Véanse los ejemplos *engineering*, *equation*, *fluid*, *force* y *steel*.

- (1.a) *Sometimes it is useful to subdivide the discipline of materials science and **engineering** into ... / In contrast, materials **engineering** is [...] designing or engineering the structure of a material to ...*
- (1.b) ***Equation** 10.13 is termed the Manning **equation**, and the parameter *n* is ...*
- (1.c) *Consider a **fluid** flowing radially outward from a line through the origin perpendicular to the *x*-*y* plane ... / The flow of a **fluid** in a pipe may be laminar flow or it may be turbulent flow.*
- (1.d) *For symmetrical objects, this **force** will be in the direction of the free stream –a drag, *D*. / The inviscid fluid merely slides along the vane without putting any **force** on it.*
- (1.e) *The hardenability of the plain carbon 1040 **steel** is low because the hardness drops ... / For example, a 1060 **steel** is a plain carbon **steel** containing 0.60 wt% C.*

2) El grupo *Personas* incluye nombres propios e iniciales del primer y/o segundo nombre<sup>57</sup> de ingenieros, físicos y matemáticos que han realizado contribuciones muy importantes en el campo de la ingeniería y disciplinas relacionadas, como se advierte en los siguientes ejemplos: *G. G. (Stokes)*, *J. L. M. (Poiseuille)*, *L. M. H. (Navier)* y *Reynolds*.

- (2.a) *Equation 3 is called Stokes's law in honor of **G. G. Stokes**, a British mathematician and a physicist.*
- (2.b) *It was developed in 1842 by a French physician, **J. L. M. Poiseuille**, who was interested in the flow of blood in capillaries.*

<sup>56</sup> En consonancia con el tratamiento de distintas expresiones matemáticas como entidades, y en base a lo aconsejado por los informantes clave, los elementos constitutivos de dichas expresiones se analizaron como subentidades (por ej., los términos de una ecuación) y, por lo tanto, se incluyeron en esta categoría.

<sup>57</sup> Nation (2001a) señala que, en determinados textos, los nombres propios pueden ser considerados palabras técnicas y formar parte de las palabras de alta frecuencia.

- (2.c) *These equations are commonly called the Navier-Stokes equations, named in honor of the French mathematician **L. M. H. Navier** and the English mechanician Sir **G. G. Stokes** ...*
- (2.d) *It is named in honor of Osborne **Reynolds**, a British engineer, who first demonstrated that this combination of variables ...*

3) El grupo *Procesos, operaciones y acciones* abarca términos que hacen referencia a eventos o fenómenos naturales y acciones realizadas tanto por personas como por una diversidad de máquinas, dispositivos e instrumentos. Dichos procesos, operaciones y acciones se expresan mediante verbos y nominalizaciones. Vale aclarar que, en ciertos casos, algunos nombres admiten una doble lectura, ya que pueden interpretarse en función de su contexto de aparición como actos y procesos o como el efecto o producto de un proceso (por ej., *corrosion* y *reaction*). A continuación se presentan algunos ejemplos:

- (3.a) *For the continuous **cooling** of a steel alloy, there exists a critical quenching rate, ... / ... attempts are made to avoid thermal stresses, which may be accomplished by **cooling** the piece at a sufficiently slow rate.*
- (3.b) *Any **corrosion** of the zinc coating will proceed at an extremely slow rate because ... / **Corrosion** processes are occasionally used to advantage.*
- (3.c) *When a **flow** is steady, all field properties (i.e., properties at any specified point), including density, remain constant with time ... / Thus for this bias, large numbers of charge carriers **flow** across the semiconductor and to the junction, ...*
- (3.d) *The bonding action is accomplished by **melting** the solder material, allowing it to flow... / ... the introduction of copper reduces the temperature of complete **melting** along the other liquidus line ...*
- (3.e) *An overall electrochemical **reaction** must consist of at least one oxidation and one reduction reaction, ... / The eutectic **reaction**, upon cooling, is similar to solidification for pure components in that ...*

4) El grupo *Propiedades, estados y cualidades de materiales, cuerpos y sistemas* consiste de palabras técnicas utilizadas para describir objetos y entidades en la disciplina. Según lo observado en nuestros datos y como se advierte en los ejemplos debajo, los adjetivos y nombres son las categorías gramaticales empleadas para señalar características intrínsecas o cualidades atribuidas a diferentes materiales, cuerpos y sistemas, como así también estados en que puedan encontrarse.

- (4.a) *Although **applied** stresses may be tensile, compressive, shear, or torsional (or combinations of these), ... / ... the mechanical behavior of a material reflects the relationship between its response or deformation to an **applied** load or force.*
- (4.b) *A major advantage of Laplace's equation is that it is a **linear** partial differential equation. / In cross-linked polymers, adjacent **linear** chains are joined one to another at various positions by covalent bonds, ...*
- (4.c) *The **stiffness** and **strength** of semicrystalline materials are often important considerations ... / Also, the fracture **strength** of a brittle ceramic may be enhanced dramatically by ...*
- (4.d) *Consider a single crystal of BCC iron oriented such that a **tensile** stress is applied along a [010] direction. / Elastic deformation, upon application of a **tensile** load, is simply the partial uncoiling, untwisting, and straightening, ...*

5) El grupo *Regiones, áreas y lugares geométricos* está compuesto por palabras que indican referencias espaciales y geométricas. Como puede apreciarse en los ejemplos debajo, estas nociones son aportadas en los textos principalmente mediante nombres.

- (5.a) *... A is taken to be frontal **area** – the projected **area** seen by a person looking toward the object from a direction parallel to the upstream velocity, U. / The specific shape of the flow **area** in the plane of the weir plate is used to designate the type of weir ...*
- (5.b) *For example, **boundary** layers form on the surfaces of cars, in the water running down the gutter of the street, ... / Consequently, this grain **boundary** region is now highly susceptible to corrosion.*
- (5.c) *A **plane** that parallels an axis may be considered to have an infinite intercept, and, therefore, a zero index. / The projection of this area on a **plane** perpendicular to the z direction is  $A_1 - A_2$ .*

6) El grupo *Nomenclaturas químicas* se compone en su totalidad por símbolos químicos, tales como C (carbono), H (hidrógeno), N (nitrógeno), y W (tungsteno o wolframio). La presencia de símbolos de la tabla periódica de materiales en los libros de texto universitarios de ingeniería obedece a que esta esfera disciplinar se sirve de aportes provenientes de la química en su carácter de ciencia básica. Los ejemplos que siguen son evidencia de estos casos.

- (6.a) *The influence of carbon content on hardness for Fe–C alloys is demonstrated in ...*
- (6.b) *Furthermore, the diffusion coefficient for the diffusion of **H** in Ni depends on temperature as ... / Various metal alloys [...] experience a significant reduction in*

*ductility and tensile strength when atomic hydrogen (H) penetrates into the material.*

- (6.c) *The NO<sub>x</sub> dissociates into N and O atoms, whereas the O<sub>2</sub> dissociates into its atomic species ... / Many of them are organic compounds that are chemically based on carbon, hydrogen, and other nonmetallic elements (i.e., O, N, and Si).*
- (6.d) *Other alloying elements include the refractory metals (Nb, Mo, W, Ta),... / Included in this group are niobium (Nb), molybdenum (Mo), tungsten (W), and tantalum (Ta).*

7) El grupo *Unidades de medición* está conformado por numerosas abreviaturas, siglas y acrónimos que indican diferentes unidades de medida, tales como grado Celsius (°C), grado Fahrenheit (°F), libra por pulgada cuadrada (psi), megapascal (MPa), Newton (N), pie (ft), voltio (V), entre otras. Su marcada presencia se explica por la importancia que adquiere en el campo de la ingeniería la medición de distintas magnitudes por medio de unidades fundamentales, suplementarias y derivadas, y por la coexistencia de varios sistemas de medición, entre ellos, el Sistema Internacional de Unidades (SI) y el Sistema Inglés de Unidades, también denominado Sistema Imperial. Si bien las unidades de medición se representan mayormente mediante símbolos y, por ende, podrían ser incluidas en la próxima categoría (*Símbolos*), en base a lo sugerido por los informantes clave se tomó la decisión metodológica de presentarlas como una categoría separada debido al número de miembros que la componen y su relevancia en la disciplina. A continuación presentamos algunos ejemplos:

- (7.a) *Building bricks are ordinarily fired around 900°C (1650°F) and are relatively porous. / Water at 40°F flows through the coils of the heat exchanger as shown in Fig. P8.68 at a rate of 0.9 gal/min.*
- (7.b) *A jet aircraft flies at a speed of 550 mph at an altitude of 35,000 ft, where the temperature is ... / Air at standard conditions flows through a horizontal 1 ft by 1.5 ft rectangular wooden duct at a rate of 5000 ft<sup>3</sup>/min.*
- (7.c) *However, since 1 N = 1 kg·m/s<sup>2</sup> it follows that the Reynolds number is unitless (dimensionless) –that is, ... / A vertical anchoring force of 0.1 N is required to hold the deflector in place.*
- (7.d) *Assume that the gage pressure remains at 3.06 psi, but the manometer is altered so that it contains only oil. / A polystyrene component must not fail when a tensile stress of 1.25 MPa (180 psi) is applied.*

- (7.e) *A copper wire 100 m long must experience a voltage drop of less than 1.5 V when a current of 2.5 A passes through it. / ... the lead electrode is oxidized, yielding a cell potential of 0.507 V.*

8) El grupo *Símbolos* constituye un grupo de naturaleza heterogénea integrado por signos que representan de manera sintética términos correspondientes principalmente a las categorías 1, 4 y 5, es decir, objetos y entidades; propiedades, estados y cualidades de materiales, cuerpos y sistemas; y regiones, áreas y lugares geométricos. Tomamos la decisión de presentar dichos símbolos en una categoría individual a partir de lo sugerido por los informantes clave. Además, se pudo determinar que varios de estos símbolos designan en los textos diferentes conceptos que, en algunas ocasiones, solo se pueden discernir en función del contexto. A modo de ejemplo, el símbolo R (mayúscula) puede hacer referencia a la resistencia eléctrica, a la constante universal de los gases ideales y a la reflectividad o reflectancia, y estas tres posibles acepciones señalan propiedades, estados o cualidades; en tanto que el símbolo r (minúscula) puede indicar una coordenada polar y el radio de una esfera o de un círculo, y estos posibles significados se corresponden con regiones, áreas y lugares geométricos. Obsérvense los siguientes ejemplos:

- (8.a) *... E is the electric field intensity, or the voltage difference between two points divided by the distance separating them ... / ... for some ceramic materials the magnitude of the modulus of elasticity E decreases with volume fraction porosity P according ...*
- (8.b) *The gas constant, R, which appears in Eq. 1.7, depends on the particular gas ... / The reflectivity R represents the fraction of the incident light that is reflected at the interface, or ... / We specify the position of P in terms of the generalized coordinates q, r, ...*
- (8.c) *For the gas constant R in Equation 5.8, we use Boltzmann's constant k, which has a value of... / In this expression, k, p, and cp represent the thermal conductivity, the mass density, and the specific heat at constant pressure, respectively.*

A modo de síntesis, la Tabla 4.4 muestra el esquema de categorización que proponemos en nuestro estudio para las palabras técnicas de alta frecuencia identificadas en el CIILTU sobre la base de criterios semánticos y pragmáticos. En la columna de la izquierda se presentan las ocho categorías de análisis empleadas, mientras que en la columna de la derecha se ofrecen algunos ejemplos identificados en el corpus para cada una de estas categorías (ver Apéndice 2, Tabla 2.1 para la clasificación completa de las

palabras técnicas o términos de alta frecuencia relevadas en el CIILTU). Los ejemplos se presentan discriminados según las listas en que aparecían originalmente. Esto nos permite observar con claridad que la clasificación de las palabras de alta frecuencia a partir de criterios semánticos y pragmáticos representa de manera fiel los significados y usos específicos de las palabras en libros de textos de ingeniería, en oposición a la clasificación de las palabras según listas existentes en la literatura. Vale aclarar que, en cada una de las categorías, la palabra que aparece destacada en negrita es aquella considerada el prototipo o elemento núcleo.

Tabla 4.4.

*Palabras técnicas o términos de alta frecuencia en el corpus*

Categoría	Lista según clasificación original	Palabras técnicas de alta frecuencia en el CIILTU
1) <i>Objetos y entidades</i>	AWL	bonding (n), equation, function* (n) [linear ...]
	GSL	<b>engineering*</b> (n), force* (n), pump*, steel* (n)
	OTRAS	alloy* (n), carbon*, fluid, turbine
2) <i>Personas</i>	OTRAS	<b>Reynolds</b> (n), G. G. [Stokes], J. L. M. [Poiseuille]
3) <i>Procesos, operaciones y acciones</i>	AWL	reaction* [eutectoid ...], stress*, transformation, transition
	GSL	<b>flow</b> (n/v), cooling (n/v), melting* (n/v), reduction
	OTRAS	corrosion, deformation, shear*, strain*
4) <i>Propiedades, estados y cualidades de materiales, cuerpos y sistemas</i>	AWL	chemical, constant* (adj.), energy, uniform* [... corrosion/flow]
	GSL	<b>strength</b> , applied, mechanical, steel* (adj.)
	OTRAS	incompressible, tensile, velocity, viscosity,
5) <i>Regiones, áreas y lugares geométricos</i>	AWL	area [anode ...], region [extrinsic ...], section, volume*
	GSL	<b>plane*</b> , boundary, point* [at point O], surface*
6) <i>Nomenclaturas químicas</i>	OTRAS	<b>C</b> (carbon), H (hydrogen), N (nitrogen), W (tungsten or wolfram)
7) <i>Unidades de medición</i>	OTRAS	<b>m</b> (meter), C (degree Celsius), N (Newton), V (volt)
8) <i>Símbolos</i>	OTRAS	<b>x<sup>Ⓞ</sup></b> (cartesian coordinate), F* (force), p* (pressure), R (electrical resistance, universal gas constant, reflectance)

Nota: \* = Palabras clasificadas en más de una categoría<sup>58</sup>; (n) = nombre; (adj.) = adjetivo; (v) = verbo; **negrita** = prototipo de la categoría; <sup>Ⓞ</sup> = símbolo que se usa de manera indistinta en mayúscula o minúscula

Como se aprecia en la Tabla 4.4, varias palabras técnicas fueron clasificadas en más de una categoría en base al asesoramiento experto proporcionado por los informantes clave. Aquí es necesario retomar lo señalado en la sección 4.2.4 en relación a la dificultad de establecer categorías semántico-pragmáticas mutuamente excluyentes y de fijar límites

<sup>58</sup>Algunas palabras clasificadas en más de una categoría tienen el potencial de activar tanto significados técnicos como no técnicos y por esa razón, estas palabras aparecen en ambas tablas, es decir en la tabla de palabras técnicas y en la tabla de palabras no técnicas.

de corte de navaja entre las categorías. Existen términos tales como *drag*, *force* y *stress* que admiten una doble categorización a partir de los matices de rasgos<sup>59</sup> que activan en sus contextos de aparición y de la situación comunicativa, ya que, por un lado, son entidades abstractas y como tales pertenecen a la categoría de objetos y entidades, y por otro lado, son acciones que se ejercen sobre cuerpos y materiales y por consiguiente corresponden a la categoría de procesos, operaciones y acciones.

#### 4.2.4.2 Palabras no técnicas

A los fines de esta investigación, las palabras no técnicas se conceptualizaron como aquellas palabras que se utilizan para describir y establecer relaciones entre los conceptos de una disciplina, como así también palabras relacionadas con el proceso de la investigación y la práctica profesional (ver Capítulo 2, sección 2.3). A partir de las categorías semánticas y pragmáticas formuladas por P. Meyer (1997), las palabras no técnicas se clasificaron en seis grupos: 1) *Tiempo, aspecto y modalidad*, 2) *Elementos del objeto o entidad descripta*, 3) *Relaciones entre objetos o entidades*, 4) *Elementos de la práctica académica y profesional*, 5) *El dominio del texto*, y 6) *Procesos y acciones*. Es necesario señalar aquí que la categoría 6, *Procesos y acciones*, fue creada ad hoc para el presente estudio. Siguiendo a P. Meyer y a Muñoz (2015), cada uno de los grupos se dividió a la vez en varias subcategorías. Algunas subcategorías se modificaron levemente; por ejemplo, la subcategoría *Recomendaciones e instrucciones* para la categoría *El dominio del texto*, propuesta por Muñoz, se redujo a *Instrucciones* atendiendo a los patrones relevados en el corpus y a los propósitos comunicativos del género libro de texto universitario. Al mismo tiempo, fue necesario formular subcategorías nuevas tales como *En la esfera del discurso* y *En el mundo extralingüístico* para la categoría *Procesos y acciones*.

1) El grupo *Tiempo, aspecto y modalidad* comprende palabras que expresan deixis temporal, nociones de tiempo, propiedades temporales o aspectuales de procesos, y modalidades o evaluaciones sobre el contenido proposicional de los textos. Como se observa en los ejemplos, la deixis temporal se indica mediante palabras que aparecen en sintagmas preposicionales que funcionan como adverbios (*at present*, *for the present*), adverbios de tiempo (*now*) y adjetivos (*current*).

---

<sup>59</sup> En este aspecto se realizaron consultas, por medio de correspondencia electrónica privada, a la terminóloga María Teresa Cabré.

- (1.a) *At present, the three most commonly used superconductors are ... / For the present we will assume that the fluid surface is open to the atmosphere.*
- (1.b) *In addition, LEDs are now being used for light sources. / Now, using arrows, indicate two different {111} slip directions within this plane.*
- (1.c) *Current prices for a wide variety of engineering materials are contained in Appendix C. / Development of the current generation of biodegradable polymers is frequently directed to ...*

Distintas nociones de tiempo se aportan principalmente mediante nombres (*time*) que funcionan como elemento núcleo de sintagmas nominales o como miembro constitutivo de palabras compuestas, y a través de palabras que aparecen en sintagmas preposicionales (*at this point, in time, over time*). A modo ilustrativo se ofrecen los siguientes ejemplos:

- (1.d) *Creep results are represented as a time-dependent creep modulus  $E_c(t)$ , defined by ... / ... a time-varying write magnetic flux is generated at the tip of the main pole*
- (1.e) *Also, the actual annealing time must be long enough to allow for any necessary transformation reactions. / Determine the carburizing time necessary to achieve a carbon concentration of ...*
- (1.f) *At this point, materials utilization was totally a selection process that involved deciding from a given, rather limited set of materials ... / At this point the nature of the function is unknown, and the objective of the experiments to be performed is to ...*
- (1.g) *... the air temperature within the tunnel will increase in time. / Over time, this erosion can be severe enough to require blade repair or replacement, which is very expensive.*

Las propiedades temporales o aspectuales de procesos se señalan fundamentalmente mediante adverbios de frecuencia (*sometimes, usually*) y verbos (*occur*) que tienen el potencial semántico de ubicar a los eventos en el tiempo, tal lo evidenciado en los ejemplos:

- (1.h) *Furthermore, in some materials other energy-absorptive processes occur at specific temperatures ... / For polymeric materials, discuss (a) two degradation processes that occur when they are exposed to liquid solvents and ...*
- (1.i) *Sometimes, D is also called the dielectric displacement. / Pressures are sometimes determined by measuring the height of a column of liquid in a vertical tube.*

- (1.j) For turbulent flow such calculations are **usually** accurate to within about 15%. / *Prestressed concretes, **usually** prefabricated, are commonly used for highway and railway bridges.*

Por último, las modalidades o evaluaciones realizadas por los autores sobre el contenido proposicional de los textos son señaladas esencialmente a través de adjetivos (*critical, important*) y adverbios (*only, very*). Obsérvense los ejemplos que siguen:

- (1.k) *Also **critical** to the understanding of phase diagrams is the concept of a phase. / Mechanical integrity under these conditions is **critical** ...*
- (1.l) *It is also **important** to know how many reference dimensions are required to ... / ... the ultimate disposition of the materials employed should be an **important** criterion.*
- (1.m) *It should also be noted that the aforementioned results are **only** approximate. / Our experiments would reveal that this result is **only** valid for small Reynolds numbers ...*
- (1.n) *The elastic-plastic transition is **very** well defined and occurs abruptly in ... / ... only a **very** superficial and simplified treatment is given.*

2) El grupo *Elementos del objeto o entidad descripta* engloba palabras que se emplean para describir objetos y entidades de la disciplina en términos de cantidad y de otras propiedades. Las palabras que expresan nociones de cantidad son primordialmente nombres (*magnitude, number, ratio*) y adjetivos (*high, net*) que tienen una tendencia a coocurrir con palabras técnicas. Los ejemplos a continuación ilustran estos usos:

- (2.a) *The silicone elastomers possess a **high** degree of flexibility at low temperatures ... / For some applications, such as furnace heating elements, a **high** electrical resistivity is desirable.*
- (2.b) *The **magnitude** of the modulus of elasticity is directly proportional to the density of the ... / The extent and **magnitude** of anisotropic effects in crystalline materials are functions of ...*
- (2.c) *The **net** heat transfer rate into the system is denoted with ... / However, the **net** flowrate through this section is the sum of flows through ...*
- (2.d) *The following table gives a **number** of Rockwell B hardness values that were measured ... / Similarly, a **number** of ceramic crystal structures may be considered in terms of ...*
- (2.e) *For a **ratio** greater than 2–3, the oxide coating may crack and flake off, ... / If it is known that this alloy has a Poisson's **ratio** of 0.30, compute ...*

Un número importante de los adjetivos comprendidos en este grupo se utilizan para señalar propiedades muy diversas de objetos y entidades propias de la ingeniería, como se observa en los ejemplos debajo:

- (2.f) *Unlike the absolute viscosity,  $\mu$ , which is a **known** value for a given fluid, / The hardest **known** materials are ceramics, which characteristic makes them ...*
- (2.g) *A **standard** technique for measuring pressure involves the use of liquid columns in ... / One **standard** procedure that is widely utilized to determine hardenability is ...*
- (2.h) *For a discontinuous and aligned-fiber composite having a **uniform** distribution of fibers ... / Consider one such alloy that initially has a **uniform** carbon concentration of 0.25 wt% ...*

Además, este grupo alberga un conjunto de nombres que actúan, en palabras de P. Meyer (1997, p. 12), como "clasificadores de entidades". Estos nombres representan tipos generales de entidades, tales como *conditions*, *effects* y *system*. Funcionan como el elemento núcleo de sintagmas nominales que adquieren un valor especializado debido a la presencia de palabras técnicas en la premodificación y/o en la posmodificación del elemento núcleo, como se advierte en los ejemplos:

- (2.i) *The pressure-density variations for isothermal and isentropic **conditions** are ... / Assuming **conditions** of steady state, calculate the diffusion flux of CO<sub>2</sub> through ...*
- (2.j) *The combined **effects** of strain hardening and precipitation hardening may be employed in ... / Surface tension **effects** play a role in many mechanics problems, ...*
- (2.k) *In the BG **system** the unit of length is the foot (ft), the time unit is the second (s), the force unit is the pound (lb), and ... / Determine the dimensions, in both the FLT **system** and MLT system, for ...*

Una característica distintiva de dichos nombres es que pueden emplearse además como expresiones anafóricas que remiten a elementos ya introducidos en el discurso, es decir, como recursos cohesivos (P. Meyer, 1997). Estos nombres generales tienen el potencial de referirse a objetos o entidades mencionados con anterioridad en el discurso pero, también, pueden hacer referencia a procesos y operaciones complejas que han sido desarrolladas en fragmentos extensos del texto. Los siguientes ejemplos ilustran casos de nombres generales utilizados como recursos cohesivos:

- (2.1) *This is demonstrated in Figure 10.29a, in which the tensile and yield strengths as well as the Brinell hardness number are plotted as a function of the weight percent*

*carbon [...] for steels that are composed of fine pearlite. All three **parameters** increase with increasing carbon concentration ...*

- (2.m) *Yet another fabrication technique involves the compaction of powdered metal, followed by a heat treatment to produce a more dense piece. The **process** is appropriately called powder metallurgy, frequently designated as P/M.*

3) El grupo *Relaciones entre objetos o entidades* está conformado por palabras que establecen relaciones entre conceptos de la disciplina tanto en el plano del discurso como en la realidad extralingüística. En la esfera del discurso, las relaciones entre ideas y secciones del texto se señalan principalmente mediante palabras que forman parte de expresiones que funcionan como subordinantes (*as long as*) y nexos lógicos (*in addition to*):

- (3.a) ***As long as** the gas remains dissolved in the water, the stratified lake (i.e., pure water on top, CO<sub>2</sub> water on the bottom) is stable.*
- (3.b) ***In addition to** being fewer in number, the new variables are dimensionless products of the original variables.*

Las palabras utilizadas para entablar relaciones entre objetos o entidades en el mundo extralingüístico son esencialmente verbos que describen diferentes tipos de relaciones, especialmente de causa-efecto (*influence*) y de inclusión (*include*), como así también participios (*associated, related*) que tienen la capacidad de denotar vínculos entre objetos o entidades. Estos casos se muestran en los ejemplos que siguen:

- (3.c) *These strain fields and **associated** forces are important in the strengthening mechanisms for metals. / **Associated** with each individual electron are both orbital and spin magnetic moments.*
- (3.d) *Applications **include** railway wheels and tracks, gears, crankshafts, and other machine parts ... / Factors to be considered **include** the nature of the applied load and its duration, as well as the environmental conditions.*
- (3.e) *Temperature can also **influence** the magnetic characteristics of materials. / The quenching medium will also **influence** the extent to which martensite forms.*
- (3.f) *Since there is no flow with the valve closed, the **related** efficiency is zero, ... / **Related** to this spin moment is the fourth quantum number, *m<sub>s</sub>*, for which two values are possible ...*

4) El grupo *Elementos de la práctica académica y profesional* está integrado por palabras relacionadas ya sea con la práctica académica, es decir con los métodos y procesos de investigación y sus resultados (*explain, results, shown*), o con el desempeño de la profesión (*calculate, measured*). Este grupo se manifiesta principalmente mediante dos categorías gramaticales: verbos y nombres. A continuación se presentan algunos ejemplos:

- (4.a) *To **calculate** the torque and power, we must know the tangential components of the absolute velocities at the inlet and exit. / For example, we may be asked to **calculate** the anchoring force required to hold a jet engine in place during a test.*
- (4.b) *Several theoretical mechanisms have been proposed to **explain** the creep behavior for various materials ... / ... the Burgers vector is an element of the theory that has been developed to **explain** this type of deformation.*
- (4.c) *Rather, hardnesses of this class of materials are **measured** using Vickers and Knoop techniques that ... / The viscosity of liquids can be **measured** through the use of a rotating cylinder viscometer of the type illustrated in Fig. P1.52.*
- (4.d) *They agree quite well with experimental **results** up to the point where ... / ... the interpretation of **results** obtained with this type of model is obviously more difficult than ...*
- (4.e) *However, numerous experimental and theoretical studies have **shown** that such an approach leads to .... / Recent research has **shown** that archerfish are very adept at calculating where their prey will fall.*

5) El grupo *El dominio del texto* incluye palabras que tienen como objetivo direccionar la atención de los lectores a partes específicas del texto, y palabras cuyo propósito es dar instrucciones a los lectores de libros de texto. Las palabras empleadas para dirigir la atención de los lectores a determinadas partes del texto señalan elementos tales como figuras, gráficos, imágenes y secciones puntuales del texto:

- (5.a) *This discrepancy is known as d'Alembert's paradox (see **Chapter** 9). / Flow in pipes is discussed in more detail in the next **chapter**, where it is shown that ...*
- (5.b) *This fiber-matrix bond ceases at the fiber ends, yielding a matrix deformation pattern as **shown** schematically in **Figure** 16.6 .../ A photomicrograph of a ferritic malleable iron is presented in **Figure** 11.3d.*
- (5.c) *Given the **following** data for brass, steel, aluminum, and titanium, rank them from least to greatest weight ... / Discuss the influence of the **following** factors on polymer tensile modulus and/or strength:...*

Las palabras empleadas para dar instrucciones a los lectores están representadas por verbos utilizados en modo imperativo. Estos verbos aparecen en consignas de ejercicios y problemas que se presentan a modo de práctica y de revisión de los contenidos desarrollados en los textos, tal como puede observarse en los ejemplos:

- (5.d) 3. **Calculate** the electrical conductivities of metals, semiconductors (intrinsic and extrinsic), and insulators given ... / (d) **Calculate** the heat loss per hour if steel is used and the thickness is increased to 20 mm (0.79 in.).
- (5.e) (a) **Describe** the phenomenon of dynamic equilibrium as it applies to oxidation and reduction electrochemical reactions. / 7. Briefly **describe**/diagram the spherulitic structure for a semicrystalline polymer.

6) El grupo *Procesos y acciones* contiene tanto verbos relativos a la esfera del discurso como verbos que refieren al mundo extralingüístico. El primer subconjunto está integrado fundamentalmente por verbos utilizados por los escritores para realizar actos lingüísticos, tal como se advierte en los ejemplos:

- (6.a) *For the sake of simplicity, let us first **consider** the solidification of a pure material, ... / To illustrate, let us **consider** the reduction of hydrogen ions to form bubbles of hydrogen gas on the surface of a zinc electrode ...*
- (6.b) *This chapter **discussed** various aspects of turbomachine analysis and design. / We also **discussed** how this type of diagram is employed in designing heat treatments to ...*

El segundo subgrupo consiste en verbos que no activan, en su contexto de uso, un valor especializado asociado con la disciplina, y que hacen referencia a realidades extralingüísticas, como se muestra en los ejemplos:

- (6.c) *In most engineering materials, however, there will also **exist** a time-dependent elastic strain component. / Other energy-absorptive mechanisms also **exist** that can add to the total heat capacity of a solid.*
- (6.d) *Other oxides (e.g., B<sub>2</sub>O<sub>3</sub> and GeO<sub>2</sub>) may also **form** glassy structures ... / Two water jets collide and **form** one homogeneous jet as shown in Fig. P5.94.*
- (6.e) *In order for these IC devices to **function** satisfactorily, very precise concentrations of an impurity (or impurities) must be incorporated ... / A broad-crested weir will not **function** properly under the wide range of flowrates considered in this example.*

A modo de síntesis, la Tabla 4.5 muestra el esquema de categorización que proponemos en nuestro estudio para las palabras no técnicas de alta frecuencia relevadas en el CIILTU. En la columna de la izquierda se presentan las seis categorías de análisis utilizadas; en la segunda columna, las subcategorías empleadas; en la tercer columna, las listas en que las palabras aparecían originalmente; y en la columna de la derecha, ejemplos identificados en el corpus para cada una de estas categorías (ver Apéndice 3, Tabla 3.1 para la clasificación completa). En las distintas categorías, aparece resaltada en negrita la palabra que es considerada el prototipo o elemento núcleo de cada subcategoría.

Tabla 4.5.

*Palabras no técnicas de alta frecuencia en el corpus*

Categoría	Subcategoría	Lista según clasificación original	Palabras no técnicas de alta frecuencia en el CIILTU
1) <i>Tiempo, aspecto y modalidad</i>	Deixis temporal	GSL	<b>now</b> , current*, present*
	Nociones de tiempo	GSL	<b>time</b> , point*
	Propiedades temporales o aspectuales de procesos	AWL	normally, occur
		GSL	<b>usually</b> , end* (v), often, sometimes
	Modalidades	AWL	briefly
		GSL	<b>critical*</b> , important, so* [versatile], very
2) <i>Elementos del objeto o entidad descripta</i>	Cantidades abstractas o propiedades cuantitativas de objetos o entidades	AWL	approximately, maximum, percent, range
		GSL	<b>rate</b> , average, degree, high
		OTRAS	fraction*, magnitude
	Cambios cuantitativos de objetos o entidades	GSL	<b>increase</b> , loss, reduction*
	Clasificadores de objetos o entidades	AWL	component, element, process, techniques
		GSL	<b>properties</b> , application*, conditions, effects
		OTRAS	plot* (n), specimen
	Propiedades de objetos o entidades	AWL	normal, potential, specific, uniform*
		GSL	<b>light*</b> (adj.) [... weight], common*, different, free*
		OTRAS	horizontal

3) Relaciones entre objetos o entidades	En la esfera del discurso	AWL	furthermore
		GSL	<b>example*</b> , addition*, due [... to], so* [... that, ... as to]
	Relaciones entre objetos o entidades en el mundo extralingüístico	AWL	corresponding, function*, required* [are ... for], similar
		GSL	<b>depends</b> , associated, control [has ... over], include
4) Elementos de la práctica académica y profesional		AWL	assume* [we can ...], compute* [To ... the rate of], obtain* [to ... a reliable value]
		GSL	<b>results*</b> (n), determine* [the procedure used to ...], shown*, treatment
		OTRAS	termed
5) El dominio del texto	Referencia textual	AWL	chapter, indicated* [also ... in Fig.], margin, section*
		GSL	<b>see</b> , figure, note*, shown* [as ... in Fig.], table
		OTRAS	diagram
	Deixis textual	GSL	<b>following</b> , follows
	Instrucciones	AWL	assume*, compute*, obtain*
		GSL	<b>calculate*</b> , describe*, explain*, use*
OTRAS		plot* (v)	
6) Procesos y acciones	En la esfera del discurso	AWL	illustrated*
		GSL	discussed* [this chapter ...], consider* [let us first ...], determine* [we can ...], <b>explain</b> [allows us to ...]
	En el mundo extralingüístico	AWL	function* (v) [will not ...]
		GSL	<b>exist</b> , form* (v), made*, use*

Nota: \* = Palabras clasificadas en más de una categoría<sup>60</sup>; (n) = nombre; (adj.) = adjetivo; (v) = verbo; **negrita** = prototipo de la subcategoría

Por último, y como se puede apreciar en la Tabla 4.5, un número significativo de palabras no técnicas, al igual que las técnicas, aparece en más de una categoría. A modo de ejemplo, la palabra *shown* denota significados que permiten ubicarla en tres de las seis categorías semántico-pragmáticas utilizadas en la clasificación de las palabras no técnicas.

<sup>60</sup> Algunas palabras clasificadas en más de una categoría tienen el potencial de activar tanto significados técnicos como no técnicos y por esa razón, estas palabras aparecen en ambas tablas, es decir en la tabla de palabras técnicas y en la tabla de palabras no técnicas.

Según el contexto en que se emplea puede denotar relaciones entre objetos o entidades en el mundo extralingüístico (categoría 3), actividades relacionadas con la práctica académica y profesional (categoría 4), o representar una instrucción (categoría 5). Asimismo, se advierte que ciertas palabras tienen el potencial semántico de activar significados especializados y no especializados, es decir, en algunos casos funcionan como palabras técnicas y en otros casos como no técnicas. Por ejemplo, en el CIILTU se evidencia que la palabra *point* puede activar dos significados especializados que están estrechamente relacionados con conceptos del dominio de la ingeniería, dado que puede referirse a propiedades de materiales (categoría 4 de las palabras técnicas), o puede indicar un lugar geométrico (categoría 5 de las palabras técnicas). Pero en otros contextos, *point* puede aportar una noción de tiempo (categoría 1 de las palabras no técnicas) o puede actuar como un clasificador de objetos o entidades (categoría 2 de las palabras no técnicas). Como se señalara anteriormente, esto se debe esencialmente a los distintos matices o módulos de rasgos semánticos que puede activar una unidad léxica en función de su contexto de aparición y de la situación comunicativa.

### **4.3 Síntesis del capítulo**

En este capítulo hemos mostrado los resultados que arrojó el análisis del vocabulario de libros de texto universitarios en inglés referidos al dominio de la ingeniería. En primer lugar, hemos presentado una descripción general del vocabulario del corpus. En segundo lugar, hemos detallado los hallazgos obtenidos en el análisis cuantitativo de las palabras de alta frecuencia relevadas en el corpus. Por último, y como resultante del análisis cualitativo realizado, hemos propuesto una reclasificación de las palabras de alta frecuencia identificadas en el corpus a partir del asesoramiento de dos especialistas en el campo de la ingeniería y de la combinación de criterios semánticos y pragmáticos. En el capítulo que sigue, discutiremos los resultados de la investigación y analizaremos sus posibles aplicaciones pedagógicas. Además, se reconocerán limitaciones del presente estudio, se sugerirán futuras líneas de investigación y se esbozarán algunas conclusiones finales.

## **Capítulo 5**

### **Discusión y conclusiones**

Este capítulo se estructura en cuatro secciones principales. La sección 5.1 está abocada a discutir los resultados del presente estudio a la luz de los fundamentos teóricos que sustentan la investigación, como así también de estudios previos sobre la temática. Esta sección ofrece una descripción general del vocabulario del corpus y una síntesis de los hallazgos principales en relación a las palabras de alta frecuencia relevadas en el corpus. La sección 5.2 se enfoca en las posibles aplicaciones del trabajo realizado. La sección 5.3 reconoce algunas limitaciones de este estudio y propone líneas de investigación futuras. Por último, la sección 5.4 presenta algunas conclusiones generales o finales.

#### **5.1 Discusión**

Este estudio tuvo como objetivo central describir el vocabulario de un corpus de libros de texto universitarios escritos en inglés del dominio de la ingeniería. En esta sección se discutirán los resultados obtenidos en las dos etapas principales de análisis del corpus: una descripción léxica general del corpus y un análisis cuantitativo y cualitativo detallado de las palabras de alta frecuencia identificadas en el corpus.

##### **5.1.1 Descripción general del vocabulario del corpus**

La primera etapa del análisis cuantitativo del CIILTU (Corpus de inglés de ingeniería de libros de texto universitarios) arrojó información sobre el tamaño del corpus, tamaño del vocabulario del corpus, relación estandarizada entre tipos y casos de palabras (SRTT), rango y cobertura de los tipos de palabras relevadas. Este conjunto de datos estadísticos permitió caracterizar el vocabulario de libros de texto universitarios publicados en inglés en el campo de la ingeniería.

El conteo de palabras, asistido por el programa informático *WordSmith Tools 7.0*, reveló que el corpus especializado diseñado para la presente investigación contiene 402.505 casos, número total de palabras que indica el tamaño del corpus, y 11.221 tipos de palabras, número que muestra el tamaño del vocabulario del corpus.

En base al tamaño del vocabulario y al tamaño del corpus, se calculó la relación estandarizada entre tipos y casos de palabras (SRTT), la cual arrojó un valor de 33,75%. Esto significa que hay un promedio de 34 tipos de palabras nuevas cada 100 palabras (casos) presentes en el CIILTU. La frecuencia promedio hallada indica una baja

variabilidad de tipos de palabras en el corpus o, dicho de otro modo, una alta repetición de palabras. Este resultado es muy interesante como indicador de un rasgo genérico, pues los textos que conforman el corpus son ejemplares de un único género, el libro de texto universitario. La baja variabilidad léxica puede atribuirse a una intención por parte de los autores de los textos de no introducir demasiado vocabulario especializado y de usar palabras no técnicas en distintos recursos, tales como la reformulación y tratamiento de la terminología (Ciapuscio, 1998, 2003b; Kuguel, 2007, 2009), con el objetivo de favorecer la comprensión y apropiación de conceptos y contenidos disciplinares por parte de estudiantes universitarios en su condición de lectores semilegos. En síntesis, la alta repetición de palabras puede explicarse en función del propósito comunicativo, didáctico-instructivo, del libro de texto universitario como representante prototípico de los géneros académicos.

La baja variabilidad de tipos de palabras en el corpus también podría deberse a las características e idiosincrasia propias de la ingeniería, campo disciplinar objeto de esta investigación. En efecto, la misma tendencia fue identificada por Biber (2006) en su estudio sobre el uso del vocabulario en la enseñanza universitaria, a partir del cual determinó que los libros de texto de ingeniería muestran menor diversidad en la selección de palabras que los libros de texto pertenecientes a las ciencias naturales, ciencias sociales y humanidades. Al respecto, coincidimos con Biber, quien concluye que la menor diversidad de palabras podía atribuirse, entre otras razones, a que la ingeniería se caracteriza por hacer uso de palabras de la lengua general que adquieren significados técnicos más específicos en esta esfera disciplinar.

Por otra parte, el rango de los distintos tipos de palabras en el corpus (determinado en función de la distribución de las palabras en los textos que conforman el CIILTU) reveló, tal como se esperaba, que las palabras de amplio rango en el corpus son en su gran mayoría palabras gramaticales, concentradas principalmente en los rangos más altos, mientras que las palabras léxicas predominan entre los tipos de palabras con un rango medio a bajo. Es decir, por su misma naturaleza, las palabras gramaticales se distribuyen más uniformemente en los textos del corpus, en tanto que las palabras léxicas, muchas con contenido semántico conceptualmente definido por la disciplina, tienden a aparecer y concentrarse puntualmente en algunos textos, dependiendo de las temáticas desarrolladas en los distintos capítulos de los dos libros que conforman el corpus.

En relación a la cobertura de los distintos tipos de palabras analizadas en el corpus, las palabras generales proporcionan una cobertura de alrededor del 27% de la totalidad de

palabras del corpus, las académicas de aproximadamente el 9%, las gramaticales del 45% y las palabras restantes categorizadas como Otras palabras de alrededor del 19%. Como se observa, las palabras generales y las palabras gramaticales combinadas cubren el 72% del corpus, porcentaje que se encuentra muy próximo al 75-80% de cobertura reportado en otros estudios (Coxhead y Nation, 2001; Nation, 2001a, 2001b). Asimismo, el porcentaje acumulado de las palabras generales, gramaticales y académicas proporciona una cobertura del 81% del corpus, es decir de la gran mayoría del CIILTU. Por otra parte, un dato que consideramos interesante en función del género analizado en esta investigación (el libro de texto universitario como representante prototípico de los géneros de formación en el nivel superior) es que la categoría Otras palabras, a pesar de ser la categoría con mayor número de tipos de palabras (6.437 tipos de palabras) que duplica la categoría de palabras generales (3.183 tipos de palabras), tiene una cobertura significativamente más baja que la de las palabras generales. La cobertura proporcionada por la categoría Otras palabras en este estudio, cercana al 19%, es levemente superior a la cobertura del 17% relevada para esta categoría de palabras en artículos de semi-divulgación de agronomía (Muñoz, 2015).

Con respecto a los distintos tipos de palabras identificadas en esta primera etapa de análisis, nos interesa enfocarnos puntualmente en las palabras académicas. Los resultados mostraron que las palabras académicas proporcionan una cobertura ligeramente superior al 9% de la totalidad de palabras de nuestro corpus de libros de texto, un valor muy cercano al 10% de cobertura hallado por Coxhead (2000, 2011) en un corpus de diferentes géneros académicos. Sin embargo, la cobertura proporcionada por las palabras académicas en los libros de texto de ingeniería analizados en nuestro estudio es considerablemente mayor al 6,5% reportado por Hajiyeva (2015) en libros de texto universitarios utilizados en especializaciones en inglés<sup>61</sup>. Por otra parte, el 9% de cobertura identificado para las palabras académicas en nuestro corpus es semejante a las coberturas informadas en artículos de investigación de distintas disciplinas: 9% (Martínez et al., 2009), 10,7% (Chen y Ge, 2007), 11% (Vongpumivitch et al., 2009), 12% (Khani y Tazik, 2013), 12% (J. Wang et al., 2008), pero es superior al 6% informado en artículos de semi-divulgación (Muñoz, 2015). Los distintos porcentajes relevados reflejan variación en el uso de palabras académicas en géneros académicos y científicos, y evidencian una relación entre la distribución de las palabras académicas y factores como la audiencia, propósitos y nivel de especialización de los contenidos de los géneros (Hajiyeva, 2015).

---

<sup>61</sup> Cabe aclarar que empleamos la palabra “especializaciones” como un equivalente aproximado de *majors* en inglés.

### **5.1.2 Análisis cuantitativo y cualitativo de las palabras de alta frecuencia en el CIILTU**

La segunda etapa de análisis del CIILTU consistió, en una primera instancia, en la identificación de las palabras de alta frecuencia en base a cálculos estadísticos, su clasificación de acuerdo con listas existentes en la literatura y el cálculo de la cobertura proporcionada por los distintos tipos de palabras de alta frecuencia, y en una segunda instancia, en la reclasificación de las palabras de alta frecuencia basada en criterios semánticos y pragmáticos y el asesoramiento brindado por los especialistas en el dominio de la ingeniería.

El cálculo de la mediana permitió identificar las 372 palabras de alta frecuencia en el CIILTU, que tienen una frecuencia de ocurrencia igual o superior a 129 casos y que representan alrededor del 3% del total de tipos de palabras del corpus. Al igual que Muñoz (2015), consideramos que la decisión metodológica de emplear el cálculo de la mediana, en lugar de la media, para la selección de las palabras de alta frecuencia fue acertada en función de las características y objetivos de la presente investigación. Primero, el cálculo de la mediana es un indicador confiable de tendencia central que permite sintetizar datos con valores muy dispares (Baroni, 2009; Levine y Stephan, 2010), como las frecuencias extremas que se registran al comienzo y al final de las listas de palabras de un corpus. Segundo, este cálculo arroja, a diferencia de la media, un conjunto más acotado de palabras, que en este estudio resulta especialmente apropiado para el análisis cualitativo de tipo manual de las palabras de alta frecuencia.

Respecto de la clasificación de las palabras de alta frecuencia de acuerdo a listas existentes en la literatura y su cobertura, identificamos que las palabras generales de alta frecuencia cubren alrededor del 15,5% del corpus, las académicas aproximadamente el 4% y la categoría de Otras palabras el 8%. Un resultado interesante para destacar es que las palabras de alta frecuencia en conjunto tienen una cobertura del 27,5% de la totalidad de las palabras del corpus. Si a las 372 palabras de alta frecuencia se le suman las 195 palabras gramaticales del corpus, se obtiene una cobertura del 72,5%. Esto demuestra que un grupo relativamente reducido de vocabulario, 567 tipos de palabras, proporciona una alta cobertura del CIILTU. Este hallazgo es muy similar al 70% de cobertura proporcionada por la suma de las palabras de alta frecuencia y palabras gramaticales en un corpus especializado de artículos de semi-divulgación del campo de la agronomía (Muñoz, 2015).

A partir de la revisión de la clasificación inicial de las 372 palabras de alta frecuencia en generales, académicas y Otras palabras, advertimos que este agrupamiento, basado en listas existentes en la literatura, no reflejaba completa y fehacientemente los significados especializados de algunas palabras. De las 205 palabras de alta frecuencia identificadas como pertenecientes a la categoría de palabras generales y de las 61 palabras de alta frecuencia categorizadas como académicas, se observó que muchas de ellas activan, en los libros de texto que conforman el CIILTU, significados vinculados con el campo de la ingeniería (a modo de ejemplo de palabras inicialmente categorizadas como generales podemos mencionar a *mass* y de palabras académicas podemos nombrar a *energy*). Por esta razón, decidimos reclasificar las palabras de alta frecuencia relevadas en el corpus en dos grandes grupos: palabras técnicas y no técnicas. Nuestra observación, y la propuesta de reclasificación, coincide con lo informado en un estudio previo sobre el vocabulario de artículos de semi-divulgación de las ciencias agrarias en inglés (Muñoz, 2015).

Sobre la premisa de que el vocabulario es condicionado por factores semánticos y pragmáticos, la clasificación de las palabras de alta frecuencia en técnicas y no técnicas se realizó en base a dos aspectos: significado y uso, a fin de intentar capturar, de este modo, las estrechas relaciones que existen entre léxico y texto (Kuguel, 2007). La utilización de criterios semánticos implicó ubicar las palabras en el sistema conceptual de la ingeniería, siguiendo lo propuesto por especialistas como Adelstein (2007), Cabré (1999) y Kuguel (2007); mientras que la aplicación de criterios pragmáticos supuso considerar los propósitos comunicativos del género libro de texto universitario, su grado de especialización, el rol social de los participantes y la relación de simetría o asimetría establecida entre ellos (Cabré, 1999; Ciapuscio, 2003a; Ciapuscio y Kuguel, 2002; P. Meyer, 1997; Pearson, 1998). Por ejemplo, palabras generales tales como *flow*, *force* y *strength* y palabras académicas tales como *bonding*, *constant* y *stress* se reclasificaron como palabras técnicas por representar conceptos específicos de la disciplina, relacionados semánticamente con otros conceptos propios de esta esfera disciplinar (Adelstein, 2009; Cabré, 1999, 2000; Kuguel, 2007), y por activar significados especializados en una situación comunicativa especializada (Cabré, 1999; Pearson, 1998).

La reclasificación de las palabras de alta frecuencia en técnicas y no técnicas tuvo, como objetivo principal, realizar una caracterización detallada del vocabulario de libros de texto universitarios en inglés del dominio de la ingeniería que resultara de la combinación de criterios cualitativos y cuantitativos. Al recategorizar las palabras de alta frecuencia, pudimos confirmar, tal lo propuesto por referentes en el área, que las palabras técnicas en

el CIILTU son palabras que activan valor terminológico dado que vehiculizan conocimiento especializado y representan conceptos relativos al campo disciplinar de la ingeniería (Adelstein, 2009; Cabré, 1999, 2000; Kuguel, 2007); en tanto que las palabras no técnicas son palabras que señalan relaciones entre los conceptos representados por las palabras técnicas, mediante la organización discursiva de los textos y la evaluación del contenido proposicional de los mismos (P. Meyer, 1997).

Los resultados revelaron diferencias a nivel semántico y pragmático entre las palabras técnicas y no técnicas del corpus. En cuanto al primer grupo, las palabras técnicas, y desde una perspectiva semántica, se refieren a conceptos relacionados con la ingeniería, ya sea a través de significados especializados que transmiten nociones específicas del dominio de la ingeniería, o de significados que en el continuum de especialización no son altamente técnicos pero que están estrechamente vinculados con este ámbito de especialidad. Las palabras técnicas identificadas en nuestro corpus tienen una mayor especificidad referencial, en función de la cantidad y tipo de información activada (más específica y compleja). Desde una perspectiva pragmática, las palabras técnicas que hemos relevado transmiten el contenido proposicional de los libros de texto analizados y vehiculizan conocimiento especializado de la ingeniería con una complejidad y nivel de abstracción acorde a la situación comunicativa del género libro de texto. En base a categorías sugeridas en estudios previos sobre palabras técnicas o términos (Muñoz, 2015) y trabajos señeros en la temática (Cabré, 1999; Sager, 1990) y al análisis de concordancias (asistido por el programa *WordSmith Tools* y enriquecido por el asesoramiento de especialistas en los dos subdominios de la ingeniería explorados en esta investigación), pudimos proponer un esquema de clasificación de las palabras técnicas del dominio de la ingeniería que intenta reflejar de manera más acabada sus características específicas: objetos y entidades; personas; procesos, operaciones y acciones; propiedades, estados y cualidades de materiales, cuerpos y sistemas; regiones, áreas y lugares geométricos; nomenclaturas químicas; unidades de medición; y símbolos.

En cuanto al segundo grupo, las palabras no técnicas, y desde una perspectiva semántica, no representan conceptos específicos de la ingeniería organizados en una estructura conceptual establecida por la comunidad de expertos y tienen una menor especificidad referencial debido a que activan menor cantidad de información sobre un referente dado (Adelstein, 2007, 2009). Dicho de otro modo, estas palabras tienen un potencial de referencia más general, aplicable a un rango amplio de referentes (P. Meyer, 1997). A partir de categorías desarrolladas en estudios previos sobre las palabras no

técnicas (P. Meyer, 1997; Muñoz, 2015) y al análisis de concordancias, realizamos una propuesta de clasificación de las palabras no técnicas identificadas en nuestro corpus que contempla las siguientes categorías: tiempo, aspecto y modalidad; elementos del objeto o entidad descripta; relaciones entre objetos o entidades; elementos de la práctica académica y profesional; el dominio del texto; y procesos y acciones. Desde una perspectiva pragmática, estas palabras se emplean para hablar sobre los contenidos temáticos de la ingeniería, evaluar el contenido proposicional y orientar al lector a través del texto (P. Meyer, 1997; Muñoz, 2015). Las palabras no técnicas usadas para hablar sobre contenidos temáticos de la disciplina representan conceptos relacionados de algún modo con la ingeniería que, aunque no tienen un significado especializado por sí solas, tienden a aparecer en combinación con palabras técnicas que denotan conceptos específicos de este campo. Por ejemplo, en el CIILTU la palabra no técnica *conditions* aparece premodificada por palabras técnicas tales como *isothermal* y *loading*. Del mismo modo, se observa que las palabras no técnicas *behaviour* y *number* son premodificadas por palabras técnicas como *fatigue* o *viscoelastic* y *Brinell hardness* o *Avogadro's*, respectivamente.

## 5. 2 Posibles aplicaciones del estudio

Los resultados de este estudio proporcionan información relevante y valiosa en el contexto de cursos de inglés con fines específicos diseñados para entrenar a estudiantes universitarios hispanohablantes en el desarrollo de estrategias de lectura en determinadas disciplinas y géneros, específicamente en cursos de lectura comprensiva en inglés en el área de la ingeniería, tal como es el caso del Módulo de Inglés que se dicta en la Facultad de Ciencias Exactas, Físicas y Naturales de la Universidad Nacional de Córdoba y que ha motivado la realización de este trabajo. Las aplicaciones pedagógicas de la presente investigación se pueden analizar en torno a dos ejes: el tratamiento del vocabulario en libros de texto en inglés del dominio de la ingeniería y el uso de listas de palabras.

Respecto del vocabulario de libros de texto universitarios en inglés del dominio de la ingeniería, los resultados de esta investigación, en consonancia con estudios anteriores, han mostrado la necesidad de concientizar a los estudiantes de cursos de inglés con fines específicos sobre al menos tres aspectos: la especificidad de las disciplinas (Bowker y Pearson, 2002; Cabré, 1999; Ciapuscio y Kuguel, 2002; Halliday, 1993, 1998, 2004; Hyland, 2002, 2006b, 2017; Hyland y Tse, 2009), la importancia del contexto de uso y de la situación comunicativa en la determinación de los significados de las palabras técnicas o no técnicas (Adelstein, 2009; Cabré, 1999, 2000; Cabré y Estopà, 2005), y la existencia de

un continuum de especialización de las palabras técnicas o términos (Cabré, 2002c; Ciapuscio, 2003a; Coxhead y Nation, 2001). Además, consideramos que la descripción semántico-pragmática de las palabras técnicas y no técnicas realizada en este estudio aporta información que contribuirá a la formulación de criterios pedagógicos más sólidos para el diseño de material didáctico, como así también para la enseñanza y evaluación del vocabulario en cursos de lectura comprensiva en inglés en el área de la ingeniería.

En lo concerniente al uso de listas de palabras, el presente estudio ha demostrado la relevancia de compilar corpus especializados pequeños para construir listas de palabras destinadas a grupos específicos de estudiantes en un campo disciplinar determinado, tal como fuera señalado por otros autores (Martínez et al., 2009; Muñoz, 2015; Nation, 2016; Nation y Webb, 2011). Creemos que la construcción de listas especializadas redunda en al menos dos beneficios. Por un lado, permite explorar e incrementar el conocimiento teórico sobre la naturaleza del vocabulario en una disciplina y género específico. Por otro lado, proporciona información valiosa con propósitos pedagógicos; por ejemplo, para el diseño de cursos de inglés con fines específicos, el establecimiento de objetivos de aprendizaje factibles de ser alcanzados, la selección fundada del vocabulario objeto de enseñanza explícita<sup>62</sup> y el diseño de instrumentos de evaluación del vocabulario más eficientes (Nation, 2016; Nation y Webb, 2011).

A la luz de los resultados obtenidos y la evidencia proporcionada en este estudio, se sugiere el diseño de listas de palabras que respondan no solo al criterio de frecuencia de ocurrencia, como la *GSL* y la *AWL*, sino además a criterios semánticos y pragmáticos, dado que, como se mencionara anteriormente, la carga semántica de las palabras varía en función del contexto de uso (Carter, 1987; Cruse, 2000; P. Meyer, 1997). Además, la presente investigación permite plantear la importancia de una adecuada selección de la unidad de conteo para la compilación de listas de palabras y el análisis del vocabulario en función de los destinatarios de las listas (Schmitt, 2010). Asimismo, nuestro estudio permite sugerir el lema como unidad de presentación del vocabulario en listas destinadas a estudiantes de cursos de lectura comprensiva que tienen un dominio limitado de la lengua inglesa o conocimientos muy heterogéneos en relación a dicho idioma. Creemos que el uso del lema como unidad de presentación en las listas de palabras es apropiado, ya que al

---

<sup>62</sup> En consonancia con lo señalado por Hsu (2014), quien sostiene que no todas las palabras en inglés tienen la misma importancia para estudiantes con necesidades muy específicas.

incluir todas sus formas flexionadas (por ej., *use, used, using*), la carga o esfuerzo de aprendizaje<sup>63</sup> por parte de los estudiantes se ve reducida (Nation, 2001a, 2016).

Por último, deseamos destacar que, si bien esta investigación tiene como destinatarios centrales a docentes y estudiantes de inglés con fines específicos, se estima que los resultados obtenidos son también de utilidad para distintos profesionales de las lenguas, en especial para profesores formadores de traductores y traductores profesionales, en al menos dos aspectos. Primero, la caracterización tipológica del libro de texto universitario en lengua inglesa aporta datos valiosos ya sea al momento de exponer a los futuros traductores a esta clase de ejemplares textuales o, durante la vida profesional, para poder traducirlos contemplando debidamente sus especificidades (Ezpeleta, García, y Montalt, 2005; Hurtado Albir, 2001). Segundo, la lista de palabras técnicas o términos, uno de los productos finales de esta investigación, puede ser usada como base para la confección de glosarios y diccionarios especializados en el área de ingeniería, herramientas lexicográficas de uso cotidiano para un traductor. Consideramos, tal como lo señala Cabré (2002b), que la recopilación y análisis de los términos de un ámbito especializado abre las puertas a un gran número de aplicaciones.

### **5.3 Limitaciones del estudio y recomendaciones para investigaciones futuras**

El presente estudio sobre el vocabulario de libros de texto de ingeniería en inglés se llevó a cabo a partir de la compilación y análisis de un corpus especializado pequeño que denominamos CIILTU (Corpus de inglés de ingeniería de libros de texto universitarios), el cual tiene aproximadamente 400.000 palabras. Aunque algunos especialistas sostienen la necesidad de construir corpus grandes, de más de un millón de palabras, para realizar generalizaciones sobre el significado y el uso de las palabras (Biber et al., 1998; Sinclair, 2004), otros referentes en la lingüística de corpus observan que los corpus pequeños son instrumentos de análisis valiosos y apropiados para el estudio de vocabulario de alta frecuencia (Hunston, 2002; Kennedy, 1998; Koester, 2010) y de géneros específicos (L. Flowerdew, 2004; Henry y Roseberry, 2001; Swales, 2006). Además, tal como señala Koester (2010), los corpus especializados al ser pequeños permiten establecer un vínculo más estrecho entre el corpus y los contextos de producción de los textos, y esto resulta de suma utilidad al momento del empleo de criterios semánticos y pragmáticos para el análisis del vocabulario de un campo de especialidad y la identificación de sus significados y

---

<sup>63</sup> Nuestra traducción para *learning burden*.

patrones de uso. Koester (2010) y C. Meyer (2004) destacan también que los corpus pequeños son más adecuados para el tratamiento manual de datos, procedimiento utilizado durante el análisis cualitativo de las palabras de alta frecuencia en nuestra investigación. Por lo tanto, si bien sugerimos la realización de estudios que indaguen el vocabulario de libros de texto universitarios de ingeniería en inglés a partir de la construcción de corpus más grandes, en cuanto al número de palabras y a las subdisciplinas representadas, consideramos que nuestro corpus ha proporcionado datos valiosos sobre los tipos de vocabulario que caracterizan a los libros de texto en el campo de la ingeniería, particularmente en los dos subdominios explorados.

Esta investigación se focalizó en el estudio del vocabulario del libro de texto universitario en inglés en un dominio, la ingeniería, y dos subdominios, la ciencia de los materiales y la mecánica de los fluidos. Los resultados, producto de la integración de métodos y técnicas cuantitativas y cualitativas, se circunscriben por ende a esta disciplina y, en especial, a estos dos subdominios de la ingeniería, y no podrían generalizarse a otros subdominios de la ingeniería o a otros campos disciplinares. Por lo tanto, una vía investigativa estaría dada por estudios de corpus que, con un enfoque metodológico similar, analicen el vocabulario de libros de texto en un número más grande de subdominios de la ingeniería para, así, poder arribar a mayores generalizaciones sobre el léxico característico de este género en el campo ingenieril, como así también comparar y contrastar los usos del vocabulario en los diferentes subdominios. Otra línea de investigación interesante sería estudiar, mediante la combinación de métodos y técnicas cuantitativas y cualitativas, los tipos de vocabulario en el género libro de texto universitario en otras esferas disciplinares, para profundizar en el conocimiento de las especificidades que adquiere el léxico en las distintas disciplinas (Bowker y Pearson, 2002; Cabré, 1999; Ciapusio y Kuguel, 2002; Halliday, 1993, 1998, 2004; Hyland, 2006b, 2017).

El objeto de estudio de la presente investigación fueron las unidades léxicas simples o monoléxicas (es decir ítems que tienen la extensión de una palabra) y sus significados y usos. No obstante, en el análisis cualitativo de las palabras de alta frecuencia identificadas en el corpus fue posible advertir la importancia que la coocurrencia de palabras tiene en la activación de distintos significados técnicos y no técnicos de una misma palabra. Se pudo apreciar que, algunas combinaciones recurrentes de palabras adquieren un valor sintagmático y terminológico porque denotan significados y conceptos específicos de un ámbito de especialidad (Cabré, 1999; McCarthy, 1990), como por ejemplo *grain boundary*

*grooves, metal matrix composites y stress relief annealing*. La existencia de dichas unidades terminológicas poliléxicas o polilexemáticas supone desafíos e interrogantes en cuanto a su reconocimiento y delimitación en la extracción automática (Cabré et al., 1996) y al cálculo de frecuencias. En relación a esto último, Hoey (2009) sostiene que las unidades léxicas compuestas por más de una palabra y que señalan conceptos y significados especializados en una disciplina deberían ser computadas como un solo ítem, sumando las frecuencias individuales de cada uno de los elementos constitutivos. Al respecto, en nuestro análisis cualitativo también se pudo observar que algunas de las palabras de alta frecuencia en el corpus aparecían en combinación con palabras no técnicas para representar conceptos y activar significados específicos del dominio de la ingeniería, como en el caso de *dielectric parameters, cation-to-anion radius ratio y shape memory alloys*. Por lo tanto, podemos concluir que el análisis de combinaciones recurrentes de palabras puede proporcionar información valiosa para la descripción del vocabulario especializado en géneros específicos de campos disciplinares determinados, como los libros de texto de ingeniería. Un área muy productiva de investigación estaría constituida por el estudio detallado de combinaciones de palabras con distintos grados de fijación léxica, como ser, unidades terminológicas poliléxicas o polilexemáticas (Adelstein, 2007; Cabré et al., 1996) y distintos tipos de unidades o expresiones fraseológicas especializadas (Cabré et al., 1996), en géneros y disciplinas específicas. Los datos relevados por este tipo de trabajos ciertamente contribuirían a caracterizar con mayor grado de detalle el vocabulario de libros de texto de ingeniería y de otras disciplinas.

#### **5.4 Conclusiones finales**

Esta investigación ha intentado realizar un aporte a la literatura creciente sobre el género libro de texto universitario y a los estudios de vocabulario basados en corpus en inglés.

En relación al libro de texto universitario, consideramos que los resultados obtenidos en el presente trabajo pueden contribuir a ahondar en la caracterización de este género académico en el dominio de la ingeniería. En particular, creemos que esta investigación, que intentó atender a las particularidades de una comunidad y cultura disciplinar sobre un género específico y reconocer el impacto de la especificidad disciplinar sobre el vocabulario, puede informar el diseño de materiales y las prácticas áulicas en cursos de inglés con fines específicos en el campo ingenieril.

Respecto del vocabulario, el presente estudio, que defiende la articulación de métodos cuantitativos y cualitativos para una descripción más adecuada de la lengua en contextos específicos de uso, espera haber aportado a profundizar el conocimiento sobre la naturaleza de los distintos tipos de vocabulario empleados en los libros de textos universitarios del dominio de la ingeniería. Esta temática no ha sido ampliamente explorada todavía, pues la mayoría de las investigaciones realizadas en este campo disciplinar han adoptado métodos y técnicas eminentemente cuantitativas (Hsu, 2014; Mudraya, 2006; Ward, 2009) para la indagación del vocabulario, especialmente de palabras académicas, y la producción de listas de palabras construidas principalmente en base al criterio de frecuencia, es decir, sin considerar los significados y usos que adquieren las palabras en géneros y disciplinas específicas.

A partir del análisis de un corpus especializado de una disciplina determinada y mediante la aplicación de criterios semánticos y pragmáticos, nos propusimos recolectar datos que proporcionaran evidencias sobre el hecho de que los significados de las palabras se activan según el contexto de uso. Por ende, podemos decir que la frecuencia constituye uno de los criterios, aunque no el único, para el análisis y la selección de palabras con fines aplicados. Consideramos que los resultados obtenidos proporcionan evidencia sólida sobre la importancia de examinar el vocabulario especializado en su contexto natural de aparición, los textos especializados (Cabré, 1999, 2002a, 2003), y la relevancia de estudiar el vocabulario para la caracterización de los géneros, como los libros de texto de ingeniería analizados en nuestro estudio.

En relación a la clasificación de las palabras de un corpus especializado mediante el uso de listas existentes, en nuestro caso la *GSL* y la *AWL*, los resultados de este trabajo corroboran investigaciones previas donde se señala que estas listas pueden resultar demasiado generales o poco específicas para estudiantes de distintas disciplinas (Hyland, 2002; Hyland y Tse, 2007; Martínez et al., 2009; Muñoz, 2015). En la presente investigación pudimos establecer que un alto porcentaje de las palabras más frecuentes en el corpus, inicialmente categorizadas como generales, académicas y Otras palabras, en efecto denotan conceptos relacionados con la ingeniería. Estos resultados hacen reflexionar sobre la efectividad y el valor pedagógico de listas preestablecidas en la literatura para la descripción del vocabulario de géneros y disciplinas específicas. Además, ponen de manifiesto la necesidad de compilar corpus especializados y diseñar listas de palabras que muestren los diferentes significados y usos de las palabras en las distintas áreas disciplinares a fin de reflejar la especificidad de las disciplinas.

Para concluir, en el marco de un interés creciente de distintas disciplinas por el uso de la lengua en contextos especializados, esperamos que esta investigación represente un aporte en la descripción de las estrechas relaciones existentes entre vocabulario y género condicionadas por un campo disciplinar determinado.

## Referencias bibliográficas

- Adam, C. y Artemeva, N. (2002). Writing instruction in English for academic purposes (EAP) classes: Introducing second language learners to the academic community. En A. M. Johns (Ed.), *Genre in the classroom. Multiple perspectives* (pp. 197-207). Mahwah: Lawrence Erlbaum Associates.
- Adelstein, A. (2007). *Unidad léxica y significado especializado: modelo de representación a partir del nombre relacional madre* (Tesis doctoral). Universitat Pompeu Fabra, Barcelona. Recuperada de:  
<https://www.tdx.cat/bitstream/handle/10803/7505/taa.pdf?sequence=1>.
- Adelstein, A. (2009). El significado léxico especializado. En G. E. Ciapusio (Ed.), *De la palabra al texto. Estudios lingüísticos del español* (pp. 17-44). Buenos Aires: Eudeba.
- Adelstein, A. y Cabré, M. T. (2002). The specificity of units with specialized meaning: Polysemy as explanatory factor. *D.E.L.T.A.*, 18, 1-25. doi: 10.1590/S0102-44502002000300003
- Adelstein, A. y Cabré, M. T. (2003). Representación lexicográfica y terminográfica de las unidades terminológicas. En *Terminologia e Indústrias da Língua, Actas do VII Simpósio Ibero-Americano de Terminologia RITerm* (pp. 103-116). Lisboa: ILTEC, União Latina, Fundação Calouste Gulbenkian.
- Adelstein, A. y Kuguel, I. (2005). La lectura y la escritura de textos que comunican ciencia: la experiencia del Taller de Lectoescritura del Primer Ciclo Universitario. En *II Jornadas sobre docencia. Los docentes universitarios ante los nuevos escenarios para la formación de los estudiantes* (pp. 59-66). Buenos Aires: Universidad Nacional de General Sarmiento.
- Ahmad, U. y Mehrjooseresht, M. (2012). Stance adverbials in engineering thesis abstracts. *Procedia - Social and Behavioral Sciences*, 66, 29-36. doi: 10.1016/j.sbspro.2012.11.244
- Atkins, B. T. y Rundell, M. (2008). *The Oxford guide to practical lexicography*. Oxford: Oxford University Press.
- Atkinson, D. (2004). Contrasting rhetorics/contrasting cultures: Why contrastive rhetoric needs a better conceptualization of culture. *Journal of English for Academic Purposes*, 3, 277-290. doi: 10.1016/j.jeap.2004.07.002

- Baker, M. (1988). Subtechnical vocabulary and the ESP teacher: An analysis of some rhetorical items in medical journal articles. *Reading in a Foreign Language*, 4(2), 91-105.
- Baker, P., Hardie, A., y McEnery, T. (2006). *A glossary of corpus linguistics*. Edinburgo: Edinburgh University Press.
- Baroni, M. (2009). Distributions in texts. En A. Lüdeling y M. Kytö (Eds.), *Corpus linguistics: An international handbook* (Vol. 2, pp. 803-822). Berlín: Mouton de Gruyter.
- Basturkmen, H. (2003). Specificity and ESP course design. *RELC Journal*, 34(1), 48-63. doi: 10.1177/003368820303400104
- Basturkmen, H. (2010). *Developing courses in English for specific purposes*. Nueva York: Palgrave Macmillan.
- Bawarshi, A. S. y Reiff, M. J. (2010). *Genre. An introduction to history, theory, research, and pedagogy*. Indiana: Parlor Press.
- Bazerman, C. (1988). *Shaping written knowledge. The genre and activity of the experimental article in science*. Wisconsin: The University of Wisconsin Press.
- Bazerman, C. (1994). Systems of genre and the enactment of social intentions. En A. Freedman y P. Medway (Eds.), *Genre and the new rhetoric* (pp. 79-101). Londres: Taylor & Francis.
- Benesch, S. (1996). Needs analysis and curriculum development in EAP: An example of a critical approach. *TESOL Quarterly*, 30(4), 723-738. doi: 10.2307/3587931
- Berkenkotter, C. y Huckin, T. N. (1995). *Genre knowledge in disciplinary communication: Cognition/culture/power*. Hillsdale: Lawrence Erlbaum.
- Bhatia, V. (1993). *Analyzing genre: Language use in professional settings*. Essex: Longman.
- Bhatia, V. (2001). Analysing genre: Some conceptual issues. En S. M. Hewing (Ed.), *Academic writing in context* (pp. 79-92). Birmingham: University of Birmingham Press.
- Bhatia, V. (2002). Applied genre analysis: A multi-perspective model. *Ibérica*, 4, 3-19.
- Bhatia, V. (2004). *Worlds of written discourse. A genre-based view*. Londres: Continuum.
- Bhatia, V. (2008). Genre analysis, ESP and professional practice. *English for Specific Purposes*, 27, 161-174. doi: 10.1016/j.esp.2007.07.005
- Bhatia, V. (2012). Critical reflections on genre analysis. *Ibérica*, 24, 17-28.

- Bhatia, V. (2013). A generic view of academic discourse. En J. Flowerdew (Ed.), *Academic Discourse* (pp. 21-39). Londres: Routledge.
- Biber, D. (1993). Representativeness in corpus design. *Literary and Linguistic Computing*, 8(4), 243-257. doi:10.1093/lle/8.4.243
- Biber, D. (2006). Vocabulary use in classroom teaching and textbooks. En D. Biber, *University language: A corpus-based study of spoken and written registers* (pp. 33-46). Amsterdam: John Benjamins.
- Biber, D. (2008). Representativeness in corpus design. En T. Fontenelle (Ed.), *Practical lexicography. A reader* (pp. 63-87). Oxford: Oxford University Press.
- Biber, D., Conrad, S., y Reppen, R. (1998). *Corpus linguistics. Investigating language structure and use*. Cambridge: Cambridge University Press.
- Biber, D. y Gray, B. (2016). *Grammatical complexity in academic English. Linguistic change in writing*. Cambridge: Cambridge University Press.
- Biber, D., Johansson, S., Leech, G., Conrad, S., y Finegan, E. (2000). *Longman grammar of spoken and written English*. Essex: Longman.
- Biber, D. y Reppen, R. (2015). Introduction. En D. Biber y R. Reppen (Eds.), *The Cambridge handbook of English corpus linguistics* (pp.1-8). Cambridge: Cambridge University Press.
- Björkman, B. (2011). English as a lingua franca in higher education: Implications for EAP. *Ibérica*, 22, 79-100.
- Björkman, B. (2013). *English as an academic lingua franca. An investigation of form and communicative effectiveness*. Berlín: Walter de Gruyter.
- Bloor, M. y Wood, F. (2006). *Keywords in qualitative methods: A vocabulary of research concepts*. Londres: SAGE Publications.
- Bondi, M. (1999). *English across genres: Language variation in the discourse of economics*. Modena: Il Fiorino.
- Bondi, M. (2016). Textbooks. En K. Hyland y P. Shaw (Eds.), *The Routledge handbook of English for academic purposes* (pp. 323-334). Londres: Routledge.
- Boudon, E. y Parodi, G. (2014). Artefactos multisemióticos y discurso académico de la economía: construcción de conocimientos en el género manual. *Revista Signos. Estudios de lingüística*, 47(85), 164-195. doi: 10.4067/S0718-09342014000200002
- Bowker, L. y Pearson, J. (2002). *Working with specialized language. A practical guide to using corpora*. Londres: Routledge.

- Busso, N. y Moyetta, D. (2015). El Módulo de Idioma Inglés en la Facultad de Ciencias Exactas, Físicas y Naturales de la UNC: origen, evolución, organización, prácticas didácticas y perspectivas. En N. A. Gigena y D. Waigandt (Comp.), *Aprender y enseñar en tiempos posparentéticos*. Recuperado de <http://www.bioingenieria.edu.ar/referencia/eventos/ingles-libro/>
- Cabré, M. T. (1999). *Terminology. Theory, methods and applications*. Amsterdam: John Benjamins.
- Cabré, M. T. (2000). Elements for a theory of terminology: Towards an alternative paradigm. *Terminology*, 6(1), 35-57. doi: 10.1075/term.6.1.03cab
- Cabré, M. T. (2002a). Análisis textual y terminología, factores de activación de la competencia cognitiva en la traducción. En A. Alcina Caudet y S. Gamero Pérez (Eds.), *La traducción científico-técnica y la terminología en la sociedad de la información* (pp. 87-105). Castellón: Universitat Jaume I.
- Cabré, M. T. (2002b). Terminología y lingüística: la teoría de las puertas. *Estudios de Lingüística del Español*, 16.
- Cabré, M. T. (2002c). Textos especializados y unidades de conocimiento: metodología y tipologización. En J. García Palacios y M. T. Fuentes (Eds.), *Texto, terminología y traducción* (pp. 15-36). Salamanca: Ediciones Almar.
- Cabré, M. T. (2003). Theories of terminology: Their description, prescription and explanation. *Terminology*, 9(2), 163-199. doi: 10.1075/term.9.2.03cab
- Cabré, M. T. (2008). El principio de poliedricidad: la articulación de lo discursivo, lo cognitivo y lo lingüístico en terminología. *Ibérica*, 16, 9-36.
- Cabré, M. T. (2009). La Teoría comunicativa de la terminología, una aproximación lingüística a los términos. *Revue française de linguistique appliquée*, XIV(2), 9-15.
- Cabré, M. T. y Estopà, R. (2002). El conocimiento especializado y sus unidades de representación: diversidad cognitiva. *Sendebarr*, 13, 141-153.
- Cabré, M. T. y Estopà, R. (2005). Unidades de conocimiento especializado, caracterización y tipología. En M. T. Cabré y C. Bach (Eds.), *Coneixement, llenguatge i discurs especialitzat* (pp. 69-94). Barcelona: IULA, Documenta Universitaria.
- Cabré, M. T., Estopà, R., y Lorente, M. (1996). Terminología y fraseología. En *Actas del V Simposio Iberoamericano de Terminología: Terminología, Ciencia y Tecnología*. Ciudad de México: RITerm.

- Cabré, M. T., Estopà, R., y Vivaldi, J. (2001). Automatic term detection: A review of current systems. En D. Bourigault, C. Jacquemin, y M-C. L'Homme (Eds.), *Recent advances in computational terminology* (pp. 53-87). Amsterdam: John Benjamins.
- Cabré, M. T., Feliu, J., y Tebé, C. (2001). Bases cognitivas de la terminología: hacia una visión comunicativa del concepto. *Sendebarr*, 12, 301-310.
- Cabrera Castillo, H. G. (2012). Análisis descriptivo sobre el concepto combustión en libros de texto universitarios. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 9(3), 311-328.
- Carkin, S. (2005). English for academic purposes. En E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 85-98). Mahwah: Lawrence Erlbaum Associates.
- Carter, R. (1987). *Vocabulary. Applied linguistic perspectives*. Londres: Routledge.
- Carter, R. y McCarthy, M. (2013). *Vocabulary and language teaching*. Nueva York: Routledge.
- Chen, Q. y Ge, G. (2007). A corpus-based lexical study on frequency and distribution of Coxhead's AWL word families in medical research articles (RAs). *English for Specific Purposes*, 26, 502-514. doi: 10.1016/j.esp.2007.04.003
- Cheng, W. (2012). *Exploring corpus linguistics: Language in action*. Nueva York: Routledge.
- Chipere, N., Malvern, D., y Richards, B. (2004). Using a corpus of children's writing to test a solution to the sample size problem affecting type-token ratios. En G. Aston, S. Bernardini, y D. Stewart (Eds.), *Corpora and language learners* (pp. 137-147). Amsterdam: John Benjamins.
- Cho, D. W. y Lee, K. (2016). English relative clauses in science and engineering journal papers: A comparative corpus-based study for pedagogical purposes. *Ampersand*, 3, 61-70. doi: 10.1016/j.amper.2016.03.002
- Christie, F. y Martin, J. R. (1997). *Genre and institutions: Social processes in the workplace and school*. Londres: Continuum.
- Chung, T. M. (2003). A corpus comparison approach for terminology extraction. *Terminology*, 9(2), 221-246. doi: 10.1075/term.9.2.05chu
- Chung, T. M. y Nation, P. (2003). Technical vocabulary in specialized texts. *Reading in a Foreign Language*, 15(2), 103-116.
- Chung, T. M. y Nation, P. (2004). Identifying technical vocabulary. *System*, 32(2), 251-263. doi: 10.1016/j.system.2003.11.008

- Ciapuscio, G. E. (1998). La terminología desde el punto de vista textual: selección, tratamiento y variación. *Organon*, 26, 43-65.
- Ciapuscio, G. (2003a). *Textos especializados y terminología*. Barcelona: Institut Universitari de Lingüística Aplicada.
- Ciapuscio, G. (2003b). Formulation and reformulation procedures in verbal interactions between experts and (semi-)laypersons. *Discourse Studies*, 5(2): 207-233. doi: 10.1177/1461445603005002004
- Ciapuscio, G. y Kuguel, I. (2002). Hacia una tipología del discurso especializado: aspectos teóricos y aplicados. En J. García Palacios y M. T. Fuentes (Eds.), *Entre la terminología, el texto y la traducción* (pp. 37-73). Salamanca: Almar.
- Clark, M. e Ishida, S. (2005). Vocabulary knowledge differences between placed and promoted EAP students. *Journal of English for Academic Purposes*, 4(3), 225-238. doi: 10.1016/j.jeap.2004.10.002
- Cobb, T. y Horst, M. (2015). Learner corpora and lexis. En S. Granger, G. Gilquin, y F. Meunier (Eds.), *The Cambridge handbook of learner corpus research* (pp. 185-206). Londres: Cambridge University Press.
- Coe, R. M. (1994). "An arousing and fulfillment of desires": The rhetoric of genre in the process era –and beyond. En A. Freedman y P. Medway (Eds.), *Genre and the new rethoric* (pp. 1-20). Londres: Taylor and Francis.
- Coe, R. M. (2002). The new rhetoric of genre: Writing political briefs. En A. M. Johns (Ed.), *Genre in the classroom. Multiple perspectives* (pp. 197-207). Mahwah: Lawrence Erlbaum Associates.
- Coe, R. M. y Freedman, A. (1998). Genre theory: Australian and North American approaches. En M. L. Kennedy (Ed.), *Theorizing composition: A critical sourcebook of theory and scholarship in contemporary composition studies* (pp. 136-147). Westport: Greenwood Press.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213-238. doi: 10.2307/3587951
- Coxhead, A. (2011). The academic word list 10 years on: Research and teaching implications. *TESOL Quarterly*, 45(2), 355-362. doi: 10.5054/tq.2011.254528
- Coxhead, A. (2016). Acquiring academic and disciplinary vocabulary. En K. Hyland y P. Shaw (Eds.), *The Routledge handbook of English for academic purposes* (pp. 177-190). Nueva York: Routledge.

- Coxhead, A. (2018). *Vocabulary and English for specific purposes research: Quantitative and qualitative perspectives*. Londres: Routledge.
- Coxhead, A. y Nation, P. (2001). The specialized vocabulary of English for academic purposes. En J. Flowerdew y M. Peacock (Eds.), *Research perspectives on English for academic purposes* (pp. 252-267). Cambridge: Cambridge University Press.
- Cruse, A. (2000). *Meaning in language: An introduction to semantics and pragmatics*. Oxford: Oxford University Press.
- Cruse, A. (2006). *A glossary of semantics and pragmatics*. Edinburgo: Edinburgh University Press.
- Crystal, D. (2003). *English as a global language* (2da ed.). Cambridge: Cambridge University Press.
- Cubo de Severino, L. (2007). *Los textos de la ciencia. Principales clases del discurso académico-científico*. Córdoba: Comunicarte.
- Dúcculi, E., Muñoz, V., y Beck, S. (2016). Estudio del vocabulario en un corpus de artículos de investigación de ecología. *Cronía, Revista de Investigación de la Facultad de Ciencias Humanas*, 11, 189-212.
- Dudley-Evans, T. (1994). Genre analysis: An approach to text analysis for ESP. En M. Coulthard (Ed.), *Advances in written text analysis* (pp.219-228). Londres: Routledge.
- Dudley-Evans, T. y St. John, M. (1998). *Developments in English for specific purposes. A multi-disciplinary approach*. Cambridge: Cambridge University Press.
- Egins, S. y Martin, J. (2000). Géneros y registros del discurso. En T. Van Dijk (Comp.), *El discurso como estructura y proceso* (pp. 335-372). Barcelona: Gedisa.
- Eskey, D. (2005). Reading in a second language. En E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 563-579). Mahwah: Lawrence Erlbaum Associates.
- Estopà, R. (2009). Los extractores de terminología: logros y escollos. En A. Alcina, E. Valero, y E. Rambla (Eds.), *Terminología y sociedad del conocimiento* (pp. 117-146). Bern: Peter Lang.
- Evison, J. (2010). What are the basics of analysing a corpus? En A. O’Keeffe y M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 122-135). Londres: Routledge.
- Ezpeleta, P., García, C., y Montalt, V. (2005). El género textual: un concepto transversal e integrador en el diseño de asignaturas de traducción y lenguaje científico-técnicos.

- En I. García-Izquierdo (Ed.), *El género textual y la traducción. Reflexiones teóricas y aplicaciones pedagógicas* (pp. 217-240). Bern: Peter Lang.
- Fakhruddin, W. y Attan, A. (2012). Putting knowledge gained into practice in civil engineering lab reports. *Procedia - Social and Behavioral Sciences*, 70, 1501-1512. doi: 10.1016/j.sbspro.2013.01.218
- Feez, S. (2002). Heritage and innovation in second language education. En A. M. Johns (Ed.), *Genre in the classroom. Multiple perspectives* (pp. 43-69). Mahwah: Lawrence Erlbaum Associates.
- Ferguson, G. (2007). The global spread of English, scientific communication and ESP: Questions of equity, access and domain loss. *Ibérica*, 13, 7-38.
- Flowerdew, J. (1993). Concordancing as a tool in course design. *System*, 21(2), 231-244. doi: 10.1016/0346-251X(93)90044-H
- Flowerdew, J. (2000). Discourse community, legitimate peripheral participation, and the non-native-English speaking scholar. *TESOL Quarterly*, 34(1), 127-150. doi: 10.2307/3588099
- Flowerdew, J. (2011). Action, content and identity in applied genre analysis for ESP. *Language Teaching*, 44(4), 516-528. doi: 10.1017/S026144481000042X
- Flowerdew, J. y Peacock, M. (2001). Issues in EAP: A preliminary perspective. En J. Flowerdew y M. Peacock (Eds.), *Research perspectives on English for academic purposes* (pp. 8-24). Cambridge: Cambridge University Press.
- Flowerdew, J. y Wang, S. H. (2015). Identity in academic discourse. *Annual Review of Applied Linguistics*, 35, 81-99. doi: 10.1017/S026719051400021X
- Flowerdew, L. (2004). The argument for using English specialized corpora to understand academic and professional language. En U. Connor y T. A. Upton (Eds.), *Discourse in the professions: Perspectives from corpus linguistics* (pp. 11-33). Amsterdam: John Benjamins.
- Flowerdew, L. (2013). Needs analysis and curriculum development in ESP. En B. Paltridge y S. Starfield (Eds.), *The handbook of English for specific purposes* (pp. 325-346). Londres: John Wiley & Sons.
- Freedman, A. y Medway, P. (1994). Locating genre studies: Antecedents and prospects. En A. Freedman y P. Medway (Eds.), *Genre and the new rhetoric* (pp. 1-20). Londres: Taylor and Francis.
- Garofolo, A. (2012). El vocabulario de artículos de investigación de psicología: estudio de corpus. En I. V. Bosio, V. M. Castel, G. Ciapuscio, L. Cubo, y G. Müller (Eds.),

- Discurso especializado: estudios teóricos y aplicados* (pp. 71-86). Mendoza: Editorial FFyL-UN Cuyo y SAL.
- Garofolo, A. y Picchio, R. (2016). El vocabulario del artículo de investigación de botánica en inglés. *Contextos de Educación*, 20, 41-47.
- Gavioli, L. (2005). *Exploring corpora for ESP learning*. Amsterdam: John Benjamins.
- Geeraerts, D. (1990). The lexicographical treatment of prototypical polysemy. En S. L. Tsohatzidis (Ed.), *Meanings and prototypes: Studies in linguistic categorization* (pp. 195-210). Londres: Routledge.
- Geeraerts, D. (2010). *Theories of lexical semantics*. Oxford: Oxford University Press.
- Geeraerts, D. (2016). Prospects and problems of prototype theory. *Diacronía*, 3, 1-16. doi: 10.17684/i4A53en
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Nueva York: Cambridge University Press.
- Grabe, W. y Stoller, F. (2011). *Teaching and researching reading* (2da ed.). Londres: Pearson Longman.
- Guantiva, R., Cabré, M. T. y Castellà, J. M. (2008). Clasificación de textos especializados a partir de su terminología. *Íkala, Revista de lenguaje y cultura*, 13(19), 15-39.
- Ha, A. Y. H. y Hyland, K. (2017). What is technicality? A technicality analysis model for EAP vocabulary. *Journal of English for Academic Purposes*, 28, 35-49. doi: 10.1016/j.jeap.2017.06.003
- Hajiyeva, K. (2015). A corpus-based lexical analysis of subject-specific university textbooks for English majors. *Ampersand*, 2, 136-144. doi: 10.1016/j.amper.2015.10.001
- Halliday, M. A. K. (1993). On the language of physical science. En M. A. K. Halliday y J. Martin (Eds.), *Writing science. Literacy and discursive power* (pp. 54-68). Londres: The Falmer Press.
- Halliday, M. A. K. (1998). Things and relations. Regrammaticising experience as technical knowledge. En J. R. Martin y R. Veel (Eds.), *Reading science. Critical and functional perspectives on discourses of science* (pp. 185-236). Londres: Routledge.
- Halliday, M. A. K. (2004). *The language of science*. Londres: Continuum.
- Halliday, M. A. K. y Hasan, R. (1989). *Language, context and text: Aspects of language in a social-semiotic perspective* (2da ed.). Oxford: Oxford University Press.

- Hamp-Lyons, L. (2011). English for academic purposes. En E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 89-105). Londres: Routledge.
- Henry, A. y Roseberry, R. L. (2001). Using a small corpus to obtain data for teaching a genre. En M. Ghadessy, A. Henry, y R. L. Roseberry (Eds.), *Small corpus studies and ELT: Theory and practice* (pp. 93-133). Amsterdam: John Benjamins.
- Hirsh, D. y Nation, P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language*, 8(2), 689-696.
- Hoey, M. (2009). Corpus linguistics and word meaning. En A. Lüdeling y M. Kytö (Eds.), *Corpus linguistics: An international handbook* (Vol. 2, pp. 972-987). Berlín: Walter de Gruyter.
- Hoey, M., Mahlberg, M., Stubbs, M., y Teubert, W. (2007). *Text, discourse and corpora: Theory and analysis*. Londres: Continuum.
- Hsu, W. (2011). The vocabulary thresholds of business textbooks and business research articles for EFL learners. *English for Specific Purposes*, 30(4), 247-257. doi: 10.1016/j.esp.2011.04.005
- Hsu, W. (2014). Measuring the vocabulary load of engineering textbooks for EFL undergraduates. *English for Specific Purposes*, 33(1), 54-65. doi: 10.1016/j.esp.2013.07.001
- Hsueh-chao, M. H. y Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403-430.
- Huckin, T. N. (2003). Specificity in LSP. *Ibérica*, 5, 3-17.
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
- Hurtado Albir, A. (2001). *Traducción y traductología. Introducción a la traductología* (5ta ed.). Madrid: Ediciones Cátedra.
- Hutchinson, T. y Waters, A. (1987). *English for specific purposes: A learning-centred approach*. Cambridge: Cambridge University Press.
- Hüttner, J., Smit, U., y Mehlmauer-Larcher, B. (2009). ESP teacher education at the interface of theory and practice: Introducing a model of mediated corpus-based genre analysis. *System*, 37(1), 99-109. doi: 10.1016/j.system.2008.06.003
- Hyland, K. (1994). Hedging in academic writing and EAP textbooks. *English for Specific Purposes*, 13(3), 239-256. doi: 10.1016/0889-4906(94)90004-3

- Hyland, K. (1999). Talking to students: Metadiscourse in introductory coursebooks. *English for Specific Purposes*, 18(1), 3-26.
- Hyland, K. (2000). *Disciplinary discourses: Social interactions in academic writing*. Londres: Longman.
- Hyland, K. (2002). Genre: Language, context, and literacy. *Annual Review of Applied Linguistics*, 22, 113-135. doi: 10.1017/S0267190502000065
- Hyland, K. (2004). *Genre and second language writing*. Michigan: The University of Michigan Press.
- Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. Londres: Continuum.
- Hyland, K. (2006a). *English for academic purposes. An advanced resource book*. Londres: Routledge.
- Hyland, K. (2006b). Disciplinary differences: Language variation in academic discourses. En K. Hyland y M. Bondi (Eds.), *Academic discourse across disciplines* (pp. 17-48). Frankfurt: Peter Lang.
- Hyland, K. (2015). Genre, discipline and identity. *Journal of English for Academic Purposes*, 19, 32-43. doi: 10.1016/j.jeap.2015.02.005
- Hyland, K. (2017). English in the disciplines: Arguments for specificity. *ESP Today*, 1, 5-23. doi: 10.18485/esptoday.2017.5.1.1
- Hyland, K. y Hamp-Lyons, L. (2002). EAP: Issues and directions. *Journal of English for Academic Purposes*, 1(1), 1-12. doi: 10.1016/S1475-1585(02)00002-4
- Hyland, K. y Tse, P. (2007). Is there an 'academic vocabulary'? *TESOL Quarterly*, 41(2), 235-254. doi: 10.1002/j.1545-7249.2007.tb00058.x
- Hyland, K. y Tse, P. (2009). Academic lexis and disciplinary practice: Corpus evidence for specificity. *International Journal of English Studies*, 9(2), 111-129.
- Hyon, S. (1996). Genre in three traditions. *TESOL Quarterly*, 30(4), 693-719. doi: 10.2307/3587930
- Hyon, S. (2002). Genre and ESL reading. A classroom study. En A. M. Johns (Ed.), *Genre in the classroom. Multiple perspectives* (pp. 121-141). Mahwah: Lawrence Erlbaum Associates.
- Johns, A. M. (1997). *Text, role and context: Developing academic literacies*. Cambridge: Cambridge University Press.
- Johns, A. M. (2002). Introduction: Genre in the classroom. En A. M. Johns (Ed.), *Genres in the classroom. Multiple perspectives* (pp. 3-13). Mahwah: Lawrence Erlbaum Associates Publishers.

- Johns, A. M. (2013). The history of English for specific purposes research. En B. Paltridge y S. Starfield (Eds.), *The handbook of English for specific purposes* (pp. 5-30). Malden: Wiley-Blackwell.
- Johns, A. M. y Dudley-Evans, T. (1991). English for specific purposes: International in scope, specific in purpose. *TESOL Quarterly*, 5(2), 297-314. doi: 10.2307/3587465
- Kanoksilapatham, B. (2015). Distinguishing textual features characterizing structural variation in research articles across three engineering sub-discipline corpora. *English for Specific Purposes*, 37, 74-86. doi: 10.1016/j.esp.2014.06.008
- Kaplan, R. B. (1993). The hegemony of English in science and technology. *Journal of Multilingual and Multicultural Development*, 14(1-2), 151-172. doi: 10.1080/01434632.1993.9994526
- Kennedy, G. (1998). *An introduction to corpus linguistics*. Londres: Longman.
- Khani, R. y Tazik, K. (2013). Towards the development of an academic list for applied linguistics research articles. *RELC Journal*, 44(2), 209-232. doi: 10.1177/0033688213488432
- Koester, A. (2010). Building small specialised corpora. En A. O'Keeffe y M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 66-79). Londres: Routledge.
- Kostina, I. (2011). Clasificación de la variación conceptual de los términos basada en la modulación semántica discursiva. *Íkala, Revista de lenguaje y cultura*, 16(27), 35-73.
- Koutsantoni, D. (2006). Rhetorical strategies in engineering research articles and research theses: Advanced academic literacy and relations of power. *Journal of English for Academic Purposes*, 5(1), 19-36. doi: 10.1016/j.jeap.2005.11.002
- Kuguel, I. (2006). La generación del significado en contexto. *Revista de Lengua y Literatura*, 34, 65-80.
- Kuguel, I. (2007). *La semántica del léxico especializado*. (Tesis doctoral). Facultad de Filosofía y Letras, Universidad Nacional de Buenos Aires. Recuperada de <http://repositorio.filo.uba.ar/handle/filodigital/1571>
- Kuguel, I. (2009). Géneros académicos y grados de especialización. En G. E. Ciapuscio (Ed.), *De la palabra al texto. Estudios lingüísticos del español* (pp. 95-127). Buenos Aires: Eudeba.
- Kuhn, T. S. (1971). *La estructura de las revoluciones científicas* (Agustín Contín, Trad.). México D.F.: Fondo de Cultura Económica.

- Laufer, B. (1992). How much lexis is necessary for reading comprehension? En H. Bejoint y P. Arnaud (Eds.), *Vocabulary and applied linguistics* (pp.126-132). Londres: Macmillan.
- Laufer, B. y Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307-322. doi: 10.1093/applin/16.3.307
- Laufer, B. y Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, 22(1), 15-30.
- Lee, J. J. y Casal, J. E. (2014). Metadiscourse in results and discussion chapters: A cross-linguistic analysis of English and Spanish thesis writers in engineering. *System*, 46, 39-54. doi: 10.1016/j.system.2014.07.009
- Lei, L. y Liu, D. (2016). A new medical academic word list: A corpus-based study with enhanced methodology. *Journal of English for Academic Purposes*, 22, 42-53. doi: 10.1016/j.jeap.2016.01.008
- Levine, D. M. y Stephan, D. F. (2010). *Even you can learn statistics: A guide for everyone who has ever been afraid of statistics*. Nueva Jersey: Pearson Education.
- Liu, J. y Han, L. (2015). A corpus-based environmental academic word list building and its validity test. *English for Specific Purposes*, 39(1), 1-11. doi: 10.1016/j.esp.2015.03.001
- Love, A. M. (1991). Process and product in geology: An investigation of some discourse features of two introductory textbooks. *English for Specific Purposes*, 10, 89-109. doi: 10.1016/0889-4906(91)90003-F
- Love, A. M. (1993). Lexico-grammatical features of geology textbooks: Process and product revisited. *English for Specific Purposes*, 12, 197-218. doi: 10.1016/0889-4906(93)90002-6
- Love, A. M. (2001). Introductory textbooks and disciplinary acculturation: A case study from social anthropology. En M. Hewings (Ed.), *Academic writing in context. Implications and applications* (pp. 122-139). Birmingham: University of Birmingham Press.
- Love, A. M. (2013). Introductory concepts and 'cutting edge' theories: Can the genre of the textbook accommodate both? En J. Flowerdew (Ed.), *Academic discourse* (pp. 76-91). Londres: Routledge.

- Ma, Y. y Lin, W. (2015). A study on the relationship between English reading comprehension and English vocabulary knowledge. *Education Research International*, 1-14. doi: 10.1155/2015/209154
- Malvern, D., Richards, B., Chipere N., y Durán P. (2004). *Lexical diversity and language development: Quantification and assessment*. Hampshire: Palgrave Macmillan.
- Martin, J. R. (1992). *English text. System and structure*. Amsterdam: John Benjamins.
- Martin, J. R. (1997). Analyzing genre. Functional parameters. En F. Christie y J. R. Martin (Eds.), *Genre and institutions. Social processes in the workplace and school* (pp. 1-39). Londres: Continuum.
- Martin, J. R. y Rose, D. (2008). *Genre relations. Mapping culture*. Londres: Equinox.
- Martínez, I. (2011). Capitalizing on the advantages of the Latin American EAP situation: Using authentic and specific materials in EAP writing instruction. *Ibérica*, 21, 31-48.
- Martínez, I. A., Beck, S. C., y Panza, C. B. (2009). Academic vocabulary in agriculture research articles: A corpus-based study. *English for Specific Purposes*, 28(3), 183-198. doi: 10.1016/j.esp.2009.04.003
- Maswana, S., Kanamaru, T., y Tajino, A. (2015). Move analysis of research articles across five engineering fields: What they share and what they do not. *Ampersand*, 2, 1-11. doi: 10.1016/j.amper.2014.12.002
- Mauranen, A. (2018). Conceptualizing ELF. En J. Jenkins, W. Baker, y M. Dewey (Eds.), *The Routledge handbook of English as a lingua franca* (pp. 7-24). Nueva York: Routledge.
- Mauranen, A., Pérez-Llantada, C., y Swales, J. (2010). Academic Englishes. A standardized knowledge? En A. Kirkpatrick (Ed.), *The Routledge handbook of world Englishes* (pp. 634-651). Nueva York: Routledge.
- Mautner, G. (2009). Corpora and critical discourse analysis. En P. Baker (Ed.), *Contemporary corpus linguistics* (pp. 32-46). Londres: Continuum.
- McCarthy, M. (1990). *Vocabulary*. Oxford: Oxford University Press.
- McEnery, T. y Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge: Cambridge University Press.
- McEnery, T. y Wilson, A. (2001). *Corpus linguistics. An introduction* (2da ed.). Edinburgo: Edinburgh University Press.

- Meara, P. (1996). The dimensions of lexical competence. En G. Brown, K. Malmkjaer, y J. Williams (Eds.), *Performance and competence in second language acquisition* (pp. 35-53). Cambridge: Cambridge University Press.
- Meyer, C. (2004). *English corpus linguistics. An introduction*. Cambridge: Cambridge University Press.
- Meyer, P. G. (1997). *Coming to know: Studies in the lexical semantics and pragmatics of academic English*. Tübingen: Gunter Narr Verlag.
- Miller, C. R. (1984). Genre as social action. *Quarterly Journal of Speech*, 70(2), 151-167. doi: 10.1080/00335638409383686
- Miller, C. R. (1994). Rhetorical community: The cultural basis of genre. En A. Freedman y P. Medway (Eds.), *Genre and the new rethoric* (pp. 57-66). Londres: Taylor and Francis.
- Moini, R. e Islamizadeh, Z. (2016). Do we need discipline-specific academic word lists? Linguistics academic word list (LAWL). *Journal of Teaching Language Skills*, 35(3), 65-90.
- Moore, T. (2002). Knowledge and agency: A study of 'metaphenomenal discourse' in textbooks from three disciplines. *English for Specific Purposes*, 21(4), 347-366. doi: 10.1016/S0889-4906(01)00030-8
- Mudraya, O. (2006). Engineering English: A lexical frequency instructional model. *English for Specific Purposes*, 25(2), 235-256. doi: 10.1016/j.esp.2005.05.002
- Muñoz, V. L. (2015). The vocabulary of agriculture semi-popularization articles in English: A corpus-based study. *English for Specific Purposes*, 39, 26-44. doi: 10.1016/j.esp.2015.04.001
- Muñoz, V. L. (2017). *Los rótulos cohesivos en el artículo de investigación de las ciencias agrarias: un estudio contrastivo inglés-español* (Tesis doctoral). Facultad de Lenguas, Universidad Nacional de Córdoba. Recuperada de <https://rdu.unc.edu.ar/handle/11086/5607>
- Myers, G. A. (1992). Textbooks and the sociology of scientific knowledge. *English for Specific Purposes*, 11(1), 3-17. doi: 10.1016/0889-4906(92)90003-S
- Nagy, W. E. y Herman, P. A. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. En M. G. McKeown y M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp.19-35). Nueva Jersey: Lawrence Erlbaum Associates.

- Nation, I. S. P. (2001a). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, I. S. P. (2001b). Using small corpora to investigate learner needs. En M. Ghadessy, A. Henry, y R. Roseberry (Eds.), *Small corpus studies in ELT. Theory and practice* (pp. 31-45). Amsterdam: John Benjamins.
- Nation, I. S. P. (2005). Teaching and learning vocabulary. En E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 581-595). Mahwah: Lawrence Erlbaum Associates.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, 63(1), 59-82. doi: 10.1353/cml.2006.0049
- Nation, I. S. P. (2008). *Teaching vocabulary: Strategies and techniques*. Boston: Heinle, Cengage Learning.
- Nation, I. S. P. (2016). *Making and using word lists for language learning and testing*. Amsterdam: John Benjamins.
- Nation, I. S. P. y Kyongho, H. (1995). Where would general service vocabulary stop and special purposes vocabulary begin? *System*, 23(1), 35-41. doi: 10.1016/0346-251X(94)00050-G
- Nation, I. S. P. y Webb, S. (2011). *Researching and analyzing vocabulary*. Boston: Heinle.
- Neilson, K. J. (2016). A text analysis of how passive voice in a biology textbook impacts English language learners. (Tesis de maestría). School of Education, Hamline University, Minnesota. Recuperada de [https://digitalcommons.hamline.edu/cgi/viewcontent.cgi?article=5218&context=hse\\_all](https://digitalcommons.hamline.edu/cgi/viewcontent.cgi?article=5218&context=hse_all)
- Nhã, V. T. T. (2015). Should an ESP course be specific or general? A literature review on the specificity debate. *VNU Journal of Science: Foreign Studies*, 31(4), 37-45.
- O'Keeffe, A., McCarthy, M., y Carter, R. (2007). *From corpus to classroom. Language use and language teaching*. Cambridge: Cambridge University Press.
- Orna-Montesinos, C. (2010). Hyponymy relations in construction textbooks: A corpus-based study. En M. L. Gea-Valor, I. García-Izquierdo, y M. J. Esteve (Eds.), *Linguistic and translation studies in scientific communication* (pp. 91-114). Bern: Peter Lang.
- Orna-Montesinos, C. (2012). *Constructing professional discourse: A multi-perspective approach to domain-specific discourses*. Newcastle: Cambridge Scholars.

- Paltridge, B. (1996). Genre, text type, and the language learning classroom. *ELT Journal*, 50(3), 237-243. doi: 10.1093/elt/50.3.237
- Paltridge, B. (2001). *Genre and the language learning classroom*. Ann Arbor: University of Michigan Press.
- Parkinson, J. y Adendorff, R. D. (2004). The use of popular science articles in teaching scientific literacy. *English for Specific Purposes*, 23(4), 379-396. doi: 10.1016/j.esp.2003.11.005
- Parodi, G. (2008a). La organización retórica del género manual: ¿Una "colonia encadenada"? En G. Parodi (Ed.), *Géneros académicos y géneros profesionales: accesos discursivos para saber y hacer* (pp. 170-198). Valparaíso: Ediciones Universitarias de Valparaíso.
- Parodi, G. (2008b). Lingüística de corpus: una introducción al ámbito. *RLA, Revista de Lingüística Teórica y Aplicada*, 46(1), 93-119.
- Parodi, G. (2010). Rhetorical organization of textbooks. En G. Parodi (Ed.), *Academic and professional discourse genres in Spanish* (pp. 634-651). Amsterdam: John Benjamins.
- Parodi, G. (2014). Genre organization in specialized discourse: Disciplinary variation across university textbooks. *Discourse Studies*, 16(1), 65-87. doi: 10.1177/1461445613496355
- Payne, G. y Payne, J. (2004). *Key concepts in social research*. Londres: SAGE Publications.
- Pearson, J. (1998). *Terms in context*. Amsterdam: John Benjamins.
- Qian, D. D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review*, 56(2), 282-308. doi: 10.3138/cmlr.56.2.282
- Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language learning*, 52(3), 513-536. doi: 10.1111/1467-9922.00193
- Ragan, P. H. (2001). Classroom use of a systemic functional small learner corpus. En M. Ghadessy, A. Henry, y R. L. Roseberry (Eds.), *Small Corpus Studies and ELT: Theory and practice* (pp. 207-236). Amsterdam: John Benjamins.
- Reppen, R. (2010). Building a corpus: What are the basics? En A. O'Keefe y M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 31-38). Londres: Routledge.

- Richardson, P. W. (2004). Reading and writing from textbooks in higher education: A case study from economics. *Studies in Higher Education*, 29(4), 505-521. doi: 10.1080/0307507042000236399
- Rowley-Jolivet, E. (2015). Quantification in conference talks and proceedings articles in engineering. *English for Specific Purposes*, 38, 11-22. doi: 10.1016/j.esp.2014.10.003
- Russo, A. (2008). Dificultades prácticas en el desarrollo de la terminología económica orientada a la traducción. En M. T. Cabré, C. Bach, y C. Tebé (Eds.), *Literalidad y dinamicidad en el discurso económico*. Barcelona: Institut Universitari de Lingüística Aplicada y Documenta Universitària.
- Sager, J. C. (1990). *A practical course in terminology processing*. Amsterdam: John Benjamins.
- Sager, J. C. (1998). In search of a foundation: Towards a theory of the term. *Terminology*, 5(1), 45-57. doi: 10.1075/term.5.1.05sag
- Salager-Meyer, F., Llopis de Segura, G. M., y Ramos, R. C. G. (2016). EAP in Latin America. En K. Hyland y P. Shaw (Eds.), *The Routledge handbook of English for academic purposes* (pp. 109-124). Nueva York: Routledge.
- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Basingstoke: Palgrave Macmillan.
- Scott, M. (2001). Comparing corpora and identifying key words, collocations, and frequency distributions through the WordSmith Tools suite of computer programs. En M. Ghadessy, A. Henry, y R. L. Roseberry (Eds.), *Small corpus studies and ELT: Theory and practice* (pp. 47- 67). Amsterdam: John Benjamins.
- Scott, M. (2004). WordSmith Tools (Versión 4.0) [Software]. Oxford: Oxford University Press.
- Scott, M. (2010). What can corpus software do? En A. O’Keeffe y M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 136-151). Londres: Routledge.
- Scott, M. (2016). WordSmith Tools (Versión 7.0) [Software]. Oxford: Oxford University Press.
- Scott, M. y Tribble, C. (2006). *Textual patterns. Key words and corpus analysis in language education*. Amsterdam: John Benjamins.

- Seidlhofer, B. (2001). Closing a conceptual gap: The case for a description of English as a lingua franca. *International Journal of Applied Linguistics*, 11(2), 133-158. doi: 10.1111/1473-4192.00011
- Seidlhofer, B. (2004). Research perspectives on teaching English as a lingua franca. *Annual Review of Applied Linguistics*, 24, 209-239. doi: 10.1017/S0267190504000145
- Şen, Y. y Kuleli, M. (2015). The effect of vocabulary size and vocabulary depth on reading in EFL context. *Procedia - Social and Behavioral Sciences*, 199, 555-562. doi: 10.1016/j.sbspro.2015.07.546
- Shabani, M. B. y Tazik, K. (2014). Coxhead's AWL across ESP and Asian EFL journal research articles (RAs): A corpus-based lexical study. *Procedia - Social and Behavioral Sciences*, 98, 1722-1728. doi: 10.1016/j.sbspro.2014.03.599
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Sinclair, J. (2001). Preface. En M. Ghadessy, A. Henry, y R. Roseberry (Eds.), *Small corpus studies and ELT. Theory and practice* (pp. vii-xv). Amsterdam: John Benjamins.
- Sinclair, J. (2003). *Reading concordances: An introduction*. Londres: Pearson/Longman.
- Sinclair, J. (2004). *Trust the text. Language, corpus and discourse*. Londres: Routledge.
- Sinclair, J. (2005). Corpus and text: Basic principles. En M. Wynne (Ed.), *Developing linguistic corpora: A guide to good practice* (pp. 1-16). Oxford: Oxbow Books.
- Singleton, D. (1999). *Exploring the second language mental lexicon*. Cambridge: Cambridge University Press.
- Stubbs, M. (2004). Language corpora. En A. Davies y C. Elder (Eds.), *The handbook of applied linguistics* (pp. 106-132). Malden: Blackwell.
- Swales, J. M. (1990). *Genre analysis. English in academic and research settings*. Cambridge: Cambridge University Press.
- Swales, J. M. (1995). The role of the textbook in EAP writing research. *English for Specific Purposes*, 14(1), 3-18. doi: 10.1016/0889-4906(94)00028-C
- Swales, J. M. (2004). *Research genres: Explorations and applications*. Cambridge: Cambridge University Press.
- Swales, J. M. (2006). Corpus linguistics and English for academic purposes. En E. Arnó Macià, A. Soler Cervera, y C. Rueda Ramos (Eds.), *Information technology in languages for specific purposes. Issues and prospects* (pp.19-33). Nueva York: Springer.

- Swales, J. M. (2009). Worlds of genre – metaphors of genre. En C. Bazerman, A. Bonini, y D. Figueiredo (Eds.), *Genre in a changing world* (pp. 1-16). Colorado: The WAC Clearinghouse.
- Tardy, C. (2004). The role of English in scientific communication: *Lingua franca or Tyrannosaurus rex?* *Journal of English for Academic Purposes*, 3(3), 247-269. doi: 10.1016/j.jeap.2003.10.001
- Taylor, S. J., Bogdan, R., y DeVault, M. L. (2016). *Introduction to qualitative research methods: A guidebook and resource* (4ta ed.). Nueva Jersey: John Wiley & Sons.
- Temmerman, R. (2000). *Towards new ways of terminology description: The sociocognitive-approach*. Amsterdam: John Benjamins.
- Todd, W. R. (2017). An opaque engineering word list: Which words should a teacher focus on? *English for Specific Purposes*, 45, 31-39. doi: 10.1016/j.esp.2016.08.003
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. Amsterdam: John Benjamins.
- Trimble, L. (1985). *English for science and technology: A discourse approach*. Cambridge: Cambridge University Press.
- Valipouri, L. y Nassaji, H. (2013). A corpus-based study of academic vocabulary in chemistry research articles. *Journal of English for Academic Purposes*, 12(4), 248-263. doi: 10.1016/j.jeap.2013.07.001
- Vongpumivitch, V., Huang, J., y Chang, Y. (2009). Frequency analysis of the words in the academic word list (AWL) and non-AWL content words in applied linguistics research papers. *English for Specific Purposes*, 28(1), 33-41. doi: 10.1016/j.esp.2008.08.003
- Wang, J., Liang, S., y Ge, G. (2008). Establishment of a medical academic wordlist. *English for Specific Purposes*, 27(4), 442-458. doi: 10.1016/j.esp.2008.05.003
- Wang, M. y Koda, K. (2007). Commonalities and differences in word identification skills among learners of English as a second language. *Language Learning*, 57(1), 201-222. doi: 10.1111/j.1467-9922.2007.00416.x
- Wang, P. (2017). A corpus-based study of English vocabulary in art research articles. *Journal of Arts & Humanities*, 6(8), 47-53. doi: 10.18533/journal.v6i8.1255
- Ward, J. (2009). A basic engineering English word list for less proficient foundation engineering undergraduates. *English for Specific Purposes*, 28(3), 170-182. doi: 10.1016/j.esp.2009.04.001
- West, M. (1953). *A general service list of English words*. Londres: Longman.

- Williams, R. (1985). Teaching vocabulary recognition strategies in ESP reading. *The ESP Journal*, 4(2), 121-131. doi: 10.1016/0272-2380(85)90015-0
- Wood, A. (2001). International scientific English: The language of research scientists around the world. En J. Flowerdew y M. Peacock (Eds.), *Research perspectives on English for academic purposes* (pp. 81-83). Nueva York: Cambridge University Press.
- Wood, D. C. y Appel, R. (2014). Multiword constructions in first year business and engineering university textbooks and EAP textbooks. *Journal of English for Academic Purposes*, 15, 1-13. doi: 10.1016/j.jeap.2014.03.002
- Woodward-Kron, R. (2008). More than just jargon – the nature and role of specialist language in learning disciplinary knowledge. *Journal of English for Academic Purposes*, 7(4), 234-249. doi: 10.1016/j.jeap.2008.10.004
- Yang, M. (2015). A nursing academic word list. *English for Specific Purposes*, 37(1), 27-38. doi: 10.1016/j.esp.2014.05.003

## **Corpus**

- Callister, W. D. y Rethwisch, D. G. (2010). *Materials science and engineering: An introduction* (8va ed.). Nueva Jersey: John Wiley & Sons.
- Young, D. F., Munson, B. R., Okiishi, T. H., y Huebsch, W. W. (2011). *A brief introduction to fluid mechanics* (5ta ed.). Nueva Jersey: John Wiley & Sons.

**APÉNDICES**

## Apéndice 1

**Tabla 1.1**

*Lista completa de lemas de alta frecuencia en el corpus*

GSL		AWL	OTRAS PALABRAS
1. <b>material</b> materials	11. <b>problem</b> problems	1. <b>equation</b> equations, Eq.	1. <b>fluid</b> fluids
2. <b>flow</b> flows	12. <b>type</b> types	2. <b>stress</b> stresses	2. <b>alloy</b> alloys
3. <b>temperature</b> temperatures	13. <b>unit</b> units	3. <b>phase</b> phases	3. <b>C</b>
4. <b>figure</b> Fig.	14. <b>example</b>	4. <b>section</b>	4. <b>velocity</b>
5. <b>pressure</b>	15. <b>point</b> points	5. <b>structure</b> structures	5. <b>B</b>
6. <b>shown</b>	16. <b>number</b>	6. <b>energy</b>	6. <b>S</b>
7. <b>also</b>	17. <b>determine</b> determined	7. <b>volume</b>	7. <b>atom</b> atoms
8. <b>metal</b> metals	18. <b>force</b> forces	8. <b>component</b> components	8. <b>X</b>
9. <b>surface</b>	19. <b>thus</b>	9. <b>obtain</b> obtained	9. <b>polymer</b> polymers
10. <b>value</b> values	20. <b>water</b>	10. <b>occur</b> occurs	10. <b>electron</b> electrons

GSL	
21. <b>system</b> systems	33. <b>pipe</b>
22. <b>given</b>	34. <b>terms</b>
23. <b>result</b> results	35. <b>solid</b>
24. <b>direction</b>	36. <b>low</b> lower
25. <b>properties</b>	37. <b>use</b> using, used
26. <b>time</b>	38. <b>change</b> changes
27. <b>rate</b>	39. <b>possible</b>
28. <b>strength</b>	40. <b>only</b>
29. <b>form</b>	41. <b>table</b>
30. <b>same</b>	42. <b>following</b>
31. <b>steel</b> steels	43. <b>so</b>
32. <b>field</b>	44. <b>solution</b>

AWL
11. <b>constant</b>
12. <b>process</b> processing
13. <b>specific</b>
14. <b>chapter</b>
15. <b>furthermore</b>
16. <b>function</b>
17. <b>element</b> elements
18. <b>indicated</b>
19. <b>area</b>
20. <b>data</b>
21. <b>required</b>
22. <b>concentration</b>

OTRAS PALABRAS
11. <b>E</b>
12. <b>M</b>
13. <b>V</b>
14. <b>fiber</b> fibers
15. <b>carbon</b>
16. <b>F</b>
17. <b>ceramic</b> ceramics
18. <b>density</b>
19. <b>deformation</b>
20. <b>crystal</b>
21. <b>G</b>
22. <b>diameter</b>

<b>GSL</b>		<b>AWL</b>	<b>OTRAS PALABRAS</b>
45. <b>relatively</b>	57. <b>important</b>	23. <b>maximum</b>	23. <b>D</b>
46. <b>plane</b> planes	58. <b>increase</b> increases, increasing	24. <b>design</b>	24. <b>strain</b>
47. <b>length</b>	59. <b>discussed</b>	25. <b>layer</b>	25. <b>diffusion</b>
48. <b>particle</b> particles	60. <b>addition</b>	26. <b>transformation</b>	26. <b>ft</b>
49. <b>weight</b> wt	61. <b>applied</b>	27. <b>assume</b>	27. <b>tensile</b>
50. <b>small</b>	62. <b>air</b>	28. <b>normally</b>	28. <b>P</b>
51. <b>control</b>	63. <b>heat</b>	29. <b>variables</b>	29. <b>fracture</b>
52. <b>application</b> applications	64. <b>very</b>	30. <b>compute</b>	30. <b>flowrate</b>
53. <b>consider</b> considered	65. <b>characteristics</b>	31. <b>region</b>	31. <b>specimen</b>
54. <b>large</b>	66. <b>iron</b>	32. <b>reaction</b>	32. <b>atomic</b>
55. <b>composition</b>	67. <b>mechanical</b>	33. <b>range</b>	33. <b>magnetic</b>
56. <b>mass</b>	68. <b>composite</b> composites	34. <b>ratio</b>	34. <b>Y</b>

<b>GSL</b>		<b>AWL</b>	<b>OTRAS PALABRAS</b>
69. <b>liquid</b>	81. <b>behavior</b>	35. <b>techniques</b>	35. <b>cell</b>
70. <b>then</b>	82. <b>line</b>	36. <b>bonding</b>	36. <b>mm</b>
71. <b>different</b>	83. <b>note</b>	37. <b>similar</b>	37. <b>N</b>
72. <b>boundary</b>	84. <b>copper</b>	38. <b>illustrated</b>	38. <b>H</b>
73. <b>high</b>	85. <b>follows</b>	39. <b>approximately</b>	39. <b>termed</b>
74. <b>head</b>	86. <b>speed</b>	40. <b>normal</b>	40. <b>T</b>
75. <b>respectively</b>	87. <b>drag</b>	41. <b>potential</b>	41. <b>R</b>
76. <b>glass</b>	88. <b>various</b>	42. <b>uniform</b>	42. <b>corrosion</b>
77. <b>motion</b>	89. <b>cooling</b>	43. <b>dimensions</b>	43. <b>magnitude</b>
78. <b>called</b>	90. <b>electrical</b>	44. <b>chemical</b>	44. <b>thermal</b>
79. <b>often</b>	91. <b>grain</b>	45. <b>analysis</b>	45. <b>molecular</b>
80. <b>see</b>	92. <b>pump</b>	46. <b>corresponding</b>	46. <b>coefficient</b>

<b>GSL</b>	
93. <b>relative</b>	105. <b>free</b>
94. <b>size</b>	106. <b>average</b>
95. <b>curve</b>	107. <b>therefore</b>
96. <b>power</b>	108. <b>effects</b>
97. <b>conditions</b>	109. <b>model</b>
98. <b>include</b>	110. <b>light</b>
99. <b>plate</b>	111. <b>band</b>
100. <b>represented</b>	112. <b>necessary</b>
101. <b>common</b>	113. <b>greater</b>
102. <b>loss</b>	114. <b>single</b>
103. <b>equal</b>	115. <b>treatment</b>
104. <b>shape</b>	116. <b>degree</b>

<b>AWL</b>
47. <b>parameters</b>
48. <b>margin</b>
49. <b>parallel</b>
50. <b>briefly</b>
51. <b>percent</b>
52. <b>transition</b>

<b>OTRAS PALABRAS</b>
47. <b>psi</b>
48. <b>L</b>
49. <b>U</b>
50. <b>diagram</b>
51. <b>plastic</b>
52. <b>ions</b>
53. <b>K</b>
54. <b>modulus</b>
55. <b>matrix</b>
56. <b>reynolds</b>
57. <b>shear</b>
58. <b>mpa</b>

<b>GSL</b>	
117. <b>explain</b>	129. <b>crack</b>
118. <b>associated</b>	130. <b>critical</b>
119. <b>well</b>	131. <b>yield</b>
120. <b>becomes</b>	132. <b>just</b>
121. <b>hardness</b>	133. <b>case</b>
122. <b>total</b>	134. <b>load</b>
123. <b>melting</b>	135. <b>body</b>
124. <b>depth</b>	136. <b>slip</b>
125. <b>according</b>	137. <b>elastic</b>
126. <b>general</b>	138. <b>produce</b> produced
127. <b>room</b>	139. <b>present</b>
128. <b>state</b>	140. <b>depends</b>

<b>OTRAS PALABRAS</b>
59. <b>axis</b>
60. <b>viscosity</b>
61. <b>aluminum</b>
62. <b>viscous</b>
63. <b>linear</b>
64. <b>radius</b>
65. <b>microstructure</b>
66. <b>equilibrium</b>
67. <b>shaft</b>
68. <b>laminar</b>
69. <b>cylinder</b>
70. <b>Q</b>

<b>GSL</b>	
141. <b>new</b>	153. <b>relationship</b>
142. <b>now</b>	154. <b>calculate</b>
143. <b>steady</b>	155. <b>measured</b>
144. <b>due</b>	156. <b>expression</b>
145. <b>describe</b>	157. <b>related</b>
146. <b>developed</b>	158. <b>typical</b>
147. <b>particular</b>	159. <b>chain</b>
148. <b>difference</b>	160. <b>impurity</b>
149. <b>distance</b>	161. <b>rather</b>
150. <b>engineering</b>	162. <b>side</b>
151. <b>reduction</b>	163. <b>tube</b>
152. <b>net</b>	164. <b>gas</b>

<b>OTRAS PALABRAS</b>
71. <b>horizontal</b>
72. <b>streamline</b>
73. <b>crystalline</b>
74. <b>pi</b>
75. <b>cubic</b>
76. <b>dislocation</b>
77. <b>silicon</b>
78. <b>conductivity</b>
79. <b>plot</b>
80. <b>W</b>
81. <b>fraction</b>
82. <b>momentum</b>

<b>GSL</b>	
165. <b>influence</b>	177. <b>end</b>
166. <b>standard</b>	178. <b>known</b>
167. <b>made</b>	179. <b>scale</b>
168. <b>expressed</b>	180. <b>sometimes</b>
169. <b>level</b>	
170. <b>find</b> found	
171. <b>long</b>	
172. <b>resistance</b>	
173. <b>usually</b>	
174. <b>current</b>	
175. <b>exist</b>	
176. <b>order</b>	

<b>OTRAS PALABRAS</b>
83. <b>fatigue</b>
84. <b>tank</b>
85. <b>turbulent</b>
86. <b>molecules</b>
87. <b>nozzle</b>
88. <b>pearlite</b>
89. <b>Z</b>
90. <b>brittle</b>
91. <b>incompressible</b>
92. <b>turbine</b>
93. <b>acceleration</b>
94. <b>hydrogen</b>

## Apéndice 2

Tabla 2.1.

*Lista completa de palabras técnicas o términos de alta frecuencia en el corpus*

Categoría	Lista según clasificación original	Palabras técnicas o términos de alta frecuencia en el CIILTU
<b>1) Objetos y entidades</b>	AWL	bonding* (n), constant*(n), equation, function* (n) [linear ...], layer, stress*, structure [band/crystal ...]
	GSL	air, band* [conduction ...], body, chain* [backbone ...], composite* (n), copper* (n), crack* (n), current* (n), curve [hardenability ...], drag*, engineering* (n), expression, field* [a ... is applied], force* (n), gas*, glass*, grain*, head* [the fan ...], impurity, iron*, light (n) [coherent ...], line [power ...], liquid* (n), load*, material (n), metal*(n), model*(n) [Bohr model], particle, pipe, plate [weir ...], pump*, resistance* [air ...], scale* [Mohs ...], solid* (n), solution*, steel* (n), terms* [pi ...], tube, water
	OTRAS	alloy* (n), aluminum* (n), atom, axis, carbon*, cell* (n), ceramic* (n), coefficient [Chezy/Hall ...], crystal*(n), cylinder, electron*, fiber*, fluid*, fraction*, hydrogen, ions, matrix, microstructure [martensitic ...], molecules, momentum*, nozzle, pearlite, pi, plastic*, polymer* (n), shaft, shear*, silicon*, strain*, streamline, tank, turbine
<b>2) Personas</b>	OTRAS	G. G. [Stokes], G. H. [Hagen], H. [Blasius], J. L. M. [Poiseuille], L. M. H. [Navier], Reynolds*(n)
<b>3) Procesos, operaciones y acciones</b>	AWL	analysis* [fractographic ...], design, reaction* [eutectoid ...], stress*, transformation [a martensitic ...], transition
	GSL	addition* [energy ...], application* [load ...], cooling (n/v), crack* (n/v), current* (n), drag*, flow (n/v), force*, load*, melting* (n/v), motion, pump* (v), pressure*, reduction*, resistance*, slip (n), treatment* [quenching ...]
	OTRAS	acceleration*, corrosion, deformation, diffusion, dislocation, flowrate*, fracture, shear*(n), strain*(n)

<b>4) Propiedades, estados y cualidades de materiales, cuerpos y sistemas</b>	AWL	bonding* (adj.), energy, chemical, concentration [carbon ...], constant* (adj.), reaction* (adj.) [...turbines], uniform* [... corrosion, flow], volume*
	GSL	applied*, chain* (adj.), composite* (adj.), composition, copper* (adj.), critical* [... flow], depth, elastic, electrical, engineering* (adj.), free* [electrons, energy], gas* (adj.) [... welding], glass* (adj.), grain* (adj.), hardness, head*, heat, iron* (adj.), length, liquid* (adj.), mass, mechanical, melting* (adj.), metal* (adj.), plane* (adj.), point* [annealing ...], power, pressure*, pump* (adj.), relative*, room* (adj.) [... temperature], side*, solid* (adj.), speed, standard* [atmosphere, half-cell], steady, steel* (adj.), strength, surface* (adj.), temperature, unit* (adj.) [... cells], weight, yield* [... strength]
	OTRAS	alloy* (adj.), aluminum* (adj.), atomic, brittle, carbon* (adj.), cell* (adj.), ceramic* (adj.), conductivity, crystal* (adj.), crystalline, cubic, density, electron* (adj.), equilibrium, fatigue, fiber* (adj.), flowrate*, fluid* (adj.), incompressible, laminar, linear, magnetic, modulus [... of elasticity], molecular, momentum* [angular ...], plastic* (adj.), polymer* (adj.), Reynolds* (adj.) [... numbers], silicon* (adj.), tensile, thermal, turbulent, velocity, viscosity, viscous
<b>5) Regiones, áreas y lugares geométricos</b>	AWL	area [anode ...], region [extrinsic ...], section* [a ... area], volume*
	GSL	band*, boundary, direction [coordinate ...], plane*(n), point* [at point O], surface*
	OTRAS	diameter, radius
<b>6) Nomenclaturas químicas</b>	OTRAS	B* (boron), C* (carbon), F* (fluorine), H* (hydrogen), N* (nitrogen), P* (phosphorus), W* (tungsten o wolfram)
<b>7) Unidades de medición</b>	OTRAS	C* (degree Celsius), F* (degree Fahrenheit), ft (feet), g* (gram), h* (hour), K (degree Kelvin), m* (meter), mm (milimeter), mpa (megapascal), N (Newton), P* (poise), psi (pound per square inch), R* (degree Rankine), s (second), V* (volt), W* (watt)

<b>8) Símbolos</b>	OTRAS	B* (magnetic field), C* (electrical capacitance), D (drag), d (distance, diameter), E (modulus of elasticity, electric field intensity), F* (force), f (friction factor), G (shear modulus), g* (acceleration of gravity), H* (magnetic field intensity, enthalpy), h* (Planck's constant), k (Boltzmann's constant, thermal conductivity), l (length), M (molar concentration), m* (mass), N* (Manning's coefficient), P* (polarization, power), p (pressure), Q (discharge), q (heat), R* (electrical resistance, universal gas constant), r (radius of a sphere or circle, a polar coordinate), T (absolute temperature, transmissivity), t (time), U (internal energy), V* (volume), v (velocity), w* <sup>⊙</sup> (weight), x <sup>⊙</sup> (cartesian coordinate), x (unknown variable in equations), y <sup>⊙</sup> (cartesian coordinate), y (dependent variable in equations), z <sup>⊙</sup> (cartesian coordinate), z (dependent variable in equations)
--------------------	-------	---

Nota: \* = Palabras clasificadas en más de una categoría; (n) = nombre; (adj.) = adjetivo; (v) = verbo; <sup>⊙</sup> = símbolo que se usa de manera indistinta en mayúscula o minúscula

### Apéndice 3

Tabla 3.1.

*Lista completa de palabras no técnicas de alta frecuencia en el corpus*

Categoría	Subcategoría	Lista según clasificación original	Palabras no técnicas de alta frecuencia en el CIILTU
<b>1) Tiempo, aspecto y modalidad</b>	<b>Deixis temporal</b>	GSL	current*, now, present*
	<b>Nociones de tiempo</b>	GSL	time, point* [at this ...]
	<b>Propiedades temporales o aspectuales de procesos</b>	AWL	normally, occur
		GSL	end* (v), just*, often, sometimes, then, usually
	<b>Modalidades</b>	AWL	briefly
		GSL	critical*, important, just*, necessary, only, possible, relatively, so* [versatile], very, well*
<b>2) Elementos del objeto o entidad descripta</b>	<b>Cantidades abstractas o propiedades cuantitativas de objetos o entidades</b>	AWL	approximately, dimensions, maximum, percent, range, ratio
		GSL	average (n/adj.), degree, distance, equal, greater, high, large, level, long*, low, net, number, order* [on the ... of], rate, rather* [... complex], scale*, single, size, small, total* (n), value, various
		OTRAS	fraction*, magnitude

<b>(Elementos del objeto o entidad descripta)</b>	<b>Cambios cuantitativos de objetos o entidades</b>	GSL	increase, loss, reduction*
	<b>Clasificadores de objetos o entidades</b>	AWL	component, data* [additional/difussion ...], element, parameters, phase, potential* (n), process, techniques, variables
		GSL	application*, behavior, case, characteristics, conditions, difference, effects, field*, form* (n) [a ... of failure], point* [an important ... to], problem* [becomes a ...], properties, room* (n) [storage ...], side*, standard* (n), state* (n), system, table* [the periodic ...], terms*, type, unit* [length ...]
		OTRAS	plot* (n), specimen
	<b>Propiedades de objetos o entidades</b>	AWL	normal, parallel, potential* (adj.), required* [the horsepower ...], specific, uniform*
		GSL	common*, developed* (adj.), different, end* (adj.) [... supports] , free*, general, given* [at a ... temperature], known*, light (adj.), made* (adj.), measured*, model* (adj.), new, particular, present*, same, shape, standard*, total* (adj.) [... saturation], typical
OTRAS		horizontal	
<b>3) Relaciones entre objetos o entidades</b>	<b>En la esfera del discurso</b>	AWL	furthermore
		GSL	according [... to], addition* [in ... ], also, due [... to], end* [To this ...], example* [for ...], long* [as ... as], order* [in ... to], rather* [... than], respectively, so* [... that, ... as to], therefore, thus, well* [as ... as]
	<b>Relaciones entre objetos o entidades en el mundo extralingüístico</b>	AWL	corresponding, function* [is a ... of], indicated* [by/with ...], required* [are ... for], similar
		GSL	associated, becomes*, change, common* [share a ... element], control [has ... over], depends, determine*, example*, expressed [... as/in], follows*, include, influence, produce* [... fatigue], related, relationship, relative*[.. to], result* [... from], yield* (v)

<b>4) Elementos de la práctica académica y profesional</b>		AWL	analysis*, assume* [we can ...], compute* [To ... the rate of], data*, obtain* [to ... a reliable value]
		GSL	applied* [these laws can be ...]*, calculate* [to ... the torque], called, describe*, determine* [the procedure used to ...], developed* (v), explain* [several theoretical mechanisms ...], find*, known*, measured* (v), note*, results* (n), see*, shown* [experimentally, it has been ... ], treatment* [theoretical ...]
		OTRAS	termed
<b>5) El dominio del texto</b>	<b>Referencia textual</b>	AWL	chapter, illustrated* [... in Fig.], indicated* [also ... in Fig.], margin, section*
		GSL	example*, figure, given*, note*, problem* [see Lab ... 8], represented* [... schematically in Fig], shown* [as ... in Fig.], solution*, table*
		OTRAS	diagram
	<b>Deixis textual</b>	GSL	following, follows*
	<b>Instrucciones</b>	AWL	assume*, compute*, obtain*
		GSL	calculate*, consider*, describe*, determine*, explain*, find*, note*, see*, state* (v), use*
		OTRAS	plot* (v)

<b>6) Procesos y acciones</b>	<b>En la esfera del discurso</b>	AWL	illustrated* [will be ... by considering]
		GSL	becomes* [it .... necessary to], consider* [let us first ....], describe*, determine* [we can ...], developed* (v) [are ... in the next section], discussed [this chapter ...], explain* [allows us to ...], represented* [is ... as a line], see* [as we will ...], shown* [this can be ...], treatment* [the present ...], use* [we ...]
	<b>En el mundo extralingüístico</b>	AWL	function* (v) [will not ...]
		GSL	becomes* [... solid], exist, form* (v), made*, produce* [... an efficient turbine], use*

Nota: \* = Palabras clasificadas en más de una categoría; (n) = nombre; (adj.) = adjetivo; (v) = verbo;

