

ILCL
INSTITUTO DE
LITERATURA Y
CIENCIAS DEL
LENGUAJE



PONTIFICIA
UNIVERSIDAD
CATÓLICA DE
VALPARAÍSO

COMPARACIÓN LÉXICO-GRAMATICAL EN INFORMES DE ALUMNOS DE TERCER Y QUINTO AÑO DE INGENIERÍA CIVIL INFORMÁTICA EN LA PUCV

TESIS para optar al grado de
LICENCIADO EN LENGUA Y LITERATURA
MENCIÓN LINGÜÍSTICA APLICADA

Alumna: Melissa Valdés Araneda

Becario proyecto FONDECYT 1190639

Profesor guía: Dr. René Venegas Velásquez

Viña del Mar, junio de 2019

En memoria de **Arturo Segundo Araneda Bahamondes:**

Por su gran dedicación y apoyo,

por su compañía en cada paso,

este logro fue su gran anhelo.

“Beyond the door there's peace, I'm sure,

And I know there'll be no more tears in heaven”

Índice

1. INTRODUCCIÓN	6
2. MARCO TEÓRICO	10
2.1 Discurso académico y profesional.....	10
2.2 Comunidad discursiva	13
2.3 Géneros discursivos: Académicos y profesionales.....	16
2.4 Género: Informe de ingeniería.....	19
2.5 Macropropósito comunicativo.....	22
2.6 Macrogénero e informe: Macrogénero Académico Evaluativo	26
3. METODOLOGÍA.....	29
3.1 Tipo de Investigación	29
3.2 Preguntas de Investigación	30
3.3 Objetivos.....	30
3.3.1 Objetivo general	31
3.3.2 Objetivos específicos.....	31
3.4 Corpus.....	31
3.6 Variables.....	34
3.6.1 Rasgos léxico-gramaticales	34
3.6.1.1 Índices básicos de información textual.....	34
3.6.1.2. Rasgos discursivos.....	35
3.6.1.3. Vocabulario especializado.....	36
3.6.1.4. Voz pasiva	36
3.6.1.5. Trigramas de lemas.....	36
3.6.1.6. Trigramas de POS.....	37

4. RESULTADOS	38
4.1 Índices básicos de información textual.....	38
4.2 Rasgos discursivos.....	40
4.3 Vocabulario especializado	44
4.4 Voz pasiva	48
4.5 Trigramas de lemas.....	52
4.5 Trigramas de POS.....	55
6. REFERENCIAS BIBLIOGRÁFICAS	61
7. ANEXOS	66

Índice de tablas, gráficos y figuras

Tablas

Tabla 1. Cantidad de palabras del corpus	31
Tabla 2. Informes de tercer año de ingeniería civil informática en la PUCV	32
Tabla 3. Informes de quinto año de ingeniería civil informática en la PUCV.	32
Tabla 4. Promedio relativo de estadísticas descriptivas de índice de información textual .	38
Tabla 5. Resultados de índice de información textual del test de Mann-Whitney	39
Tabla 6. Rasgos discursivos y promedio de estadísticas descriptivas	41
Tabla 7. Resultados de rasgos discursivos de la prueba de Mann-Whitney para rasgos discursivos	43
Tabla 8. Frecuencias de términos en tercer y quinto año de Ingeniería Civil Informática..	45
Tabla 9. Términos con mayor frecuencia en tercer año	46
Tabla 10. Términos con mayor frecuencia en quinto año	47
Tabla 11. Promedios del uso de pasiva.	48

Tabla 12. Resultados de test Mann-Whitney para voz pasiva.....	49
Tabla 13. Frecuencias de voces pasivas en tercer y quinto año de Ingeniería Civil Informática.	49
Tabla 14. Frecuencias de ejemplares únicos de voz pasiva en tercer año.....	50
Tabla 15. Frecuencias de ejemplares únicos de voz pasiva en quinto año.....	51
Tabla 16. Trigramas de lemas en común para tercer y quinto año.....	52
Tabla 17. Trigramas de lema únicos de tercer año.....	53
Tabla 18. Trigramas de lemas exclusivos de quinto año.....	54
Tabla 19. Trigramas de POS en común para tercer y quinto año.....	56
Tabla 20. Trigramas de POS exclusivos de tercer año.....	56
Tabla 21. Trigramas de POS exclusivos de quinto año.....	57

Gráficos

Gráfico 1. Índices básicos de operación textual.....	39
Gráfico 2. Rasgos discursivos.....	42

Figuras

Figura 1. Relación entre participantes (Parodi, Ibáñez y Venegas, 2009, p.85).....	16
Figura 2. Cuadro de denominaciones recurrentes (Harvey y Muñoz, 2006).....	25
Figura 3. Opciones disponibles en la herramienta DescApp de la plataforma PACTE	34
Figura 4. Opciones para índices básicos de información disponibles en la versión en español de SINLP de la herramienta DescApp.	36
Figura 5. Marcadores discursivos y tipos de palabras disponibles en la versión al español de SINLP de la herramienta DescApp.....	36
Figura 6. Trigramas de Lema extraídos de la herramienta DescApp.....	38
Figura 7. Trigramas de POS extraídos de la herramienta DescApp.....	38

1. INTRODUCCIÓN

El informe escrito “es una exposición escrita relativa a un tema, problema o actividad con propósitos formales de comunicación” (Amieva, 2001). En cualquiera de sus variedades, cumple un rol clave en la formación de futuros ingenieros, cobrando especial relevancia al tratarse de un ejercicio constante, como una actividad evaluativa en el desarrollo de las carreras y el ejercicio profesional. El cumplimiento de los aspectos formales del informe, al igual que el de los aspectos que atañen a la redacción, además de la presentación adecuada de un informe, son habilidades que se adquieren a lo largo de la carrera universitaria. Para el desarrollo de estas habilidades, es necesaria una alfabetización adecuada de los alumnos sobre sus implicancias, tanto académicas como también profesionales, pues son destrezas necesarias para lograr un desempeño adecuado al ingresar al mundo laboral. No obstante, una problemática latente respecto a la implementación del género en ingeniería es la escasa alfabetización respecto al mismo, donde existe poca dedicación para inducir correctamente a los alumnos en su producción, pues se asume que ya tienen dichos conocimientos (Amieva, 2001). De la misma manera, las mallas curriculares pocas veces contemplan la implementación de cursos relativos a la redacción y corrección de informes, lo que dificulta aún más la tarea, especialmente a los alumnos novatos.

Paulatinamente, los estudiantes de ingeniería van ingresando a la comunidad discursiva que les pertenece, en este caso, la de los Ingenieros Civiles en Informática. Una comunidad discursiva, desde la propuesta de Swales (1990), se entiende como un conjunto de miembros intercomunicados, agrupados en torno a una ocupación o interés similar, que comparten objetivos, reglas y prácticas asociadas al lenguaje, al igual que géneros y léxico específico, lo que les proporciona tener convenciones propias acordadas por los integrantes, que permiten la existencia de géneros característicos de estas comunidades (Parodi, Venegas, Ibáñez y Gutiérrez, 2008). Como fue explicado con anterioridad, la importancia de la producción del género informe durante la carrera del ingeniero civil informático es clave, dado que corresponde a una práctica arraigada profundamente en la comunidad discursiva y que es habitual en el ejercicio profesional. Por tanto, luego de

egresar, es necesario tener un léxico especializado y adecuado a la exigencia laboral. En este sentido, resulta imprescindible la labor del docente de ingeniería en la alfabetización, pues su contacto directo con el contexto profesional y el académico mediará el ingreso adecuado del alumno a la comunidad discursiva, correspondiente a la carrera en que se está formando.

Bajo esta premisa, es pertinente señalar el estrecho vínculo de la comunidad discursiva de la ingeniería civil informática con el discurso profesional y el académico. Referente a esto, entendemos como discurso académico al conjunto de todas las producciones lingüísticas elaboradas por los distintos miembros de múltiples comunidades discursivas, con el fin de generar, transmitir o reproducir conocimiento científico (Espejo, 2006). El discurso académico, situado dentro del espectro del discurso especializado (Parodi, 2005), suele confundirse con el discurso profesional que son “aquellos textos en los cuales se plasma un conocimiento específico del mundo, en el que, a su vez, se constituyen los propósitos de la comunidad profesional (Berkenkotter, Huckin y Ackerman, 1991; Doheny-Farina, 1991; Bhatia, 1993; Christie y Martin, 1997)” (Parodi, 2007, *en línea*). Situándonos en lo que compete a la ingeniería civil informática, es común ver confluir ambos discursos, debido a la constante preparación y simulación de los alumnos para el mundo laboral, debido a ello, no es extraño encontrar géneros importados, que corresponden a géneros que se posicionan y crean originalmente en un discurso profesional y que se utilizan en un contexto universitario (Parodi y Vásquez-Rocca, 2015).

En cuanto a la clasificación del género informe, este es incluido dentro del Macrogénero Académico Evaluativo (Jarpa, 2015), pues comparte el macropropósito comunicativo de “evaluar las competencias, aprendizajes, habilidades y dominio disciplinar adquirido por los estudiantes durante un proceso formativo en un determinado campo disciplinar (Jarpa, 2012)” (Jarpa, 2015, p.230), compartido con otros géneros pertenecientes a esta clasificación (Jarpa, 2015). Al referirnos al concepto macrogénero, se entiende que es un término útil para agrupar a un conjunto de géneros (Martin, 2002). El Macrogénero Académico Evaluativo incluiría diversos géneros que buscan evaluar los aprendizajes en el ámbito académico, como los controles, informes, pruebas, ensayos y hasta tesis de grado

(Venegas, 2010), que son escritos por estudiantes (semilegos) y son dirigidos a un profesor (experto) en el marco de una situación comunicativa de carácter didáctico-pedagógico (Jarpa, 2015).

En lo que respecta al informe escrito, estudios como el proyecto institucional DIPUC 2003 en Chile indican su importancia como medio para registrar los conocimientos adquiridos en la universidad (Harvey y Muñoz, 2006). Igualmente, se hace énfasis en la importancia del género informe en cuanto a la inserción de nuevos miembros dentro de la comunidad discursiva de la ingeniería, como parte de un proceso formativo continuo de alfabetización disciplinar, que se liga a géneros particulares, en este caso, el informe (Harvey y Muñoz, 2006). Sin embargo, los estudios sobre los rasgos léxico-gramaticales del género informe son escasos. Propuestas como la de Harvey y Muñoz (2006) estudian el género informe y su representación en el discurso de los académicos, pero no se centran en una disciplina en específico o comparan los patrones léxico-gramaticales de estos en diferentes años.

De acuerdo con la problemática, la pregunta de investigación es: ¿Cuáles son los rasgos léxico gramaticales de los informes de los alumnos de ingeniería informática de tercer y quinto año en la PUCV? A partir de esta pregunta, el objetivo general consiste en comparar los rasgos léxico gramaticales en los informes de los alumnos de tercer y quinto año de ingeniería informática en la PUCV, con el fin de observar si existe variación en el uso de tales rasgos atribuibles a la inmersión en la comunidad discursiva. Para esta investigación se analizarán los rasgos léxico-gramaticales en informes producidos por estudiantes de tercer y quinto año, tomando en consideración como variables el vocabulario especializado, los rasgos discursivos, la voz pasiva, la frecuencia de los trigramas de lemas y de categorías morfológicas (Pos, del inglés Part of Speech), además de los índices básicos de información textual. En concreto, los índices básicos de información textual, los trigramas de lemas y los trigramas de PoS fueron procesados mediante la herramienta DescApp de la plataforma PACTE, desarrollada por el grupo Redilegra, que incluye la detección de estos rasgos entre sus funcionalidades (Lillo, 2018). De igual manera, en esta plataforma se ha incluido una versión adaptada al español de SiNLP y algunos scripts propios que permiten determinar la voz pasiva en español e identificar diversos conectores

discursivos, en función de identificar los rasgos discursivos en el corpus (Lillo, 2018). Con relación al vocabulario especializado, este fue procesado con la herramienta Termout.

La relevancia de esta investigación radica en que, por medio de la comparación, se podrá ratificar si existe una variación efectiva en los patrones léxico-gramaticales a medida que los alumnos van progresando en la carrera Ingeniería Civil Informática. A raíz de esto, se podrá comprobar si los alumnos, a medida que avanzan en sus estudios, van integrándose en su comunidad discursiva, adoptando los rasgos y el vocabulario pertinente del conocimiento informático.

2. MARCO TEÓRICO

En el siguiente apartado, procederemos a definir los conceptos teóricos que sientan las bases fundamentales para esta investigación. Concretamente, daremos cuenta de los conceptos centrales que adquieren relevancia a lo largo del trabajo. De esta manera, el capítulo se articula en seis apartados. En primera instancia, se presentarán los conceptos discurso académico y discurso profesional, para continuar en los siguientes apartados con los conceptos de comunidad discursiva y géneros académicos y profesionales. Vinculado a la noción de género, se introducirá el género informe, para luego abordar la discusión sobre el macrogénero académico evaluativo, que es donde se sitúa el género informe dentro del ámbito académico. Posteriormente, se entregarán las bases sobre el concepto de macropropósito comunicativo, del cual se desprende la organización retórica-funcional del género informe, para determinar si se incluye dentro del análisis.

2.1 Discurso académico y profesional

Antes de referirnos por completo a la noción de discurso académico y profesional, debemos situarnos en la discusión del discurso especializado, que es donde se han posicionado ambos discursos (Parodi, 2005) y, además, ha sido un enorme foco de discusión en los estudios del lenguaje en estos últimos años. De acuerdo con esto, podemos referirnos al discurso especializado como un conjunto de varios tipos de textos que tienen rasgos muy prototípicos. Igualmente, Parodi (2005), define al discurso especializado como un conjunto de textos que se agrupan por tener un sistema de rasgos lingüísticos particulares, en torno a temáticas específicas no cotidianas vinculadas al conocimiento previo respecto a la disciplina de sus participantes; a su vez, el discurso especializado cumple con una función comunicativa meramente referencial y circula en contextos situacionales específicos, con un léxico altamente especializado.

En cuanto a la idea inicial de Parodi (2005), se concibe al discurso especializado desde una mirada amplia y globalizadora, postulando la idea de una heterogeneidad de textos que están insertos en una escala de gradación, dando paso a la concepción del discurso especializado como un *continuum* donde “se alinean los textos a lo largo de una gradiente diversificada que va desde un alto hasta un bajo grado de especificación” (Parodi, 2007,

online), donde coexisten desde textos escolares hasta académicos, universitarios y profesionales (Parodi, Venegas, Ibáñez y Gutiérrez, 2008). De esta manera, la propuesta de Parodi concibe al discurso especializado como un hiperónimo, que antecede a las nociones de discurso académico y discurso profesional (Parodi 2007), posicionándose dentro del *continuum* del discurso especializado como eje articulador, que media a los géneros profesionales y a otros géneros especializados, cumpliendo un rol fundamental en la construcción de un sello disciplinar especializado (Parodi et al.2008).

En primer lugar, en referencia al estudio del discurso académico, Silver (2006) sostiene que los esfuerzos iniciales por estudiar el discurso académico se realizaron separando su estudio del paradigma lingüístico imperante a mediados de los años setenta, en pleno cambio de paradigma del sistema al uso. Su principal propósito era enfocarse en un modelo de lenguaje más basado en la comunicación. Asimismo, se vincula directamente con las crecientes necesidades del inglés con propósitos específicos, con el fin de crear propuestas didácticas destinadas al ingreso a segundas lenguas (Silver, 2006), lo que también llevaría a un interés en estudiar las particularidades de los textos que circulaban en el contexto académico. La caracterización de dichas peculiaridades, como rasgos léxico-gramaticales y discursivos, posibilitó el establecimiento de géneros discursivos asociados al discurso académico, que, a su vez, posibilitan el acceso al conocimiento y a las prácticas especializadas escritas (Bhatia, 2002; Silver, 2006), por tanto, donde un determinado sujeto, a través de los géneros asociados al discurso académico, se enfrentará a un proceso progresivo de alfabetización permanente (Parodi et al. 2008).

De acuerdo a lo anterior, adoptamos la propuesta de Espejo (2006) que define al discurso académico como el conjunto de todas las producciones lingüísticas elaboradas por los distintos miembros de múltiples comunidades discursivas, con el fin de generar, transmitir o reproducir conocimiento científico. Igualmente, Parodi (2007) propone que el discurso académico, en términos comunicativo-funcionales, se caracteriza por la predominancia de la descripción con propósitos persuasivos y didácticos, que, a su vez, se posiciona como un discurso autorizado, que adquiere cierto prestigio y credibilidad en su contexto de circulación, que es preferentemente académico y que tiene propósitos fundamentalmente

académicos (Flowerdew, 2002). Además, tiene como propósito “cristalizar y difundir un conocimiento disciplinar y apoyar a la formación de los miembros en formación de esa comunidad de especialistas” (Parodi, 2007, *online*). No obstante, el discurso académico llega a tener límites difusos, por lo cual puede confundirse o asimilarse con otros discursos que se le asemejan, como es el discurso técnico científico, el discurso profesional, el discurso pedagógico o el discurso institucional (López, 2002; Flowerdew, 2004). A grandes rasgos, el discurso académico apunta a un conocimiento autorizado respecto al mundo y como este es visto, que actúa para reforzar las convicciones teóricas de la disciplina y su derecho a validar el conocimiento (Hyland, 2000), caracterizándose por tener un canal preferente en el discurso escrito, en favor de las necesidades de la academia (Parodi, 2007).

En segundo lugar, en lo que compete al discurso profesional, situado en el *continuum* del discurso especializado, presenta una problemática similar a la revisada en el discurso académico, pues ambos términos tienden a confundirse entre sí, dada a la dificultad de una caracterización estricta (Parodi, 2007). De acuerdo a esta afirmación, es preciso distinguir al discurso profesional del académico de acuerdo a su contexto de uso y circulación, que es meramente profesional. Bajo esa premisa, Bazerman y Paradis (1991a) postulan que el discurso profesional se estructura a partir de una dinámica textual, la cual entrega forma a una profesión, que, a su vez, conforma comunidades profesionales que se definen a través de los textos (Bazerman y Paradis, 1991b). En síntesis, los discursos profesionales son “aquellos textos en los cuales se plasma un conocimiento específico del mundo, en el que, a su vez, se constituyen los propósitos de la comunidad profesional” (Berkenkotter, Huckin & Ackerman, 1991; Doheny-Farina, 1991; Bhatia, 1993; Christie & Martin, 1997)” (Parodi, 2007, *online*). De igual manera, Linell (1998) propone que al tratar los discursos profesionales se debe distinguir entre:

“a) un discurso intraprofesional, es decir, entre los integrantes de una misma profesión

b) un discurso interprofesional, entre integrantes de profesiones distintas

- c) un discurso profesional-lego, entre los participantes de una comunidad profesional y un público lego”.

En los tres casos mencionados se apunta a discursos ligados a instituciones profesionales, que a su vez generan interacciones entre los miembros que pertenecen a la institución y miembros de otras instituciones, o con clientes y destinatarios ajenos a dicha institución (López, 2002). En referencia a esta postura, para la caracterización de los discursos profesionales es imprescindible considerar el contexto de producción y recepción de los discursos producidos, el proceso de construcción textual y las estructuras lingüísticas que la situación específica de comunicación requiere (López, 2002).

Finalmente, referente a la carrera ingeniería civil informática, que es fuente de estudio para esta investigación, es común ver confluir ambos discursos, académico y profesional, debido a la constante preparación y simulación de los alumnos sobre el mundo laboral, debido a ello, no es extraño encontrar géneros importados, que corresponden a géneros que se posicionan y crean originalmente en un discurso profesional y que se utilizan en un contexto universitario (Parodi, 2014). La noción de género ‘importado’ se centra en los criterios de producción del género y del lector ideal, así como en el contexto de circulación efectivo (Parodi y Vásquez-Rocca, 2015). Los géneros importados resultan tremendamente relevantes para la inserción laboral de los estudiantes “(Hyland, 2002; Candlin y Hyland, 2014; Cassany, 2006; Cassany y Morales, 2008; Cassany y López, 2010; Bach y López, 2011)” (Parodi y Vásquez-Rocca, 2015, p. 392), ya que, al lograr dominar dichos géneros, en conjunto con otras habilidades ligadas al ejercicio profesional, le permitirá al estudiante ingresar progresivamente a la comunidad discursiva (Parodi y Vásquez-Rocca, 2015).

2.2 Comunidad discursiva

Durante la alfabetización en el nivel universitario, un sujeto novicio ingresa a una comunidad discursiva, accediendo a un conocimiento disciplinar a través de la lectura de textos vinculados a su currículo (Parodi et al. 2008), al igual que con la escritura de géneros característicos en la disciplina, como son los informes. En torno al proceso de alfabetización:

“la formación del grado universitario no consiste solamente en la adquisición de contenidos sino, sobre todo, en la enculturación de los estudiantes en las prácticas académicas que configuran las esferas disciplinares. Se trata de que los estudiantes se apropien de las formas socialmente consensuadas de construir, negociar y comunicar el conocimiento en la universidad (Hyland, 2004, p. 11).

Respecto a los textos que circulan dentro de una comunidad, se debe considerar que “quien escribe (o habla) no lo hace fuera de una comunidad sino al interior suyo siguiendo patrones y normas; hecho que, en definitiva, definirá el propósito o el contenido de aquello que se quiere comunicar” (Martínez, 2012, p.23). En relación a este hecho, una comunidad discursiva, desde la propuesta de Swales (1990), se entiende como un conjunto de miembros intercomunicados, agrupados en torno a una ocupación o interés similar, que comparten objetivos, reglas y prácticas asociadas al lenguaje, al igual que géneros y léxico específico, lo que les proporciona tener convenciones propias acordadas por los integrantes, que permiten la existencia de géneros característicos de estas comunidades (Parodi et al., 2008). Cabe destacar que Swales (1990) propone que la noción de comunidad discursiva ha sido estudiada desde distintas perspectivas sociales, en especial a los estudios sobre escritura. De acuerdo con ello, establece seis características para reconocer una comunidad discursiva:

- “1. Has a broadly agreed set of common public goals
2. Has mechanisms of intercommunication among its members,
3. Uses its participatory mechanisms primarily to provide information and feedback.
4. Utilizes and hence possesses one or more genres in the communicative furtherance of its aims.
5. In addition to owning genres, a discourse community has acquired some specific lexis.
6. Has a threshold level of members with a suitable degree of relevant content and discourse expertise” (Swales, 1990, p. 24-25)

En cuanto a la relación entre los participantes, Swales (1990) y Bhatia (2004) sostienen que las comunidades discursivas existen y se sustentan por medio de una “membresía dinámica en términos de experticia”, debido a que los miembros que componen dichas comunidades ingresan a ellas siendo aprendices y se transforman paulatinamente en expertos, a partir de su interacción discursiva en dicha comunidad (Parodi, Ibáñez y Venegas, 2009). De la relación entre los participantes se desprenden cuatro roles centrales: escritor experto, lector experto, lector semilego, lector lego. Los dos primeros apuntan a un alto grado de experticia, lo que les permite tanto conocer cómo producir los géneros que circulan dentro de la comunidad. El lector semilego se posiciona en una experticia intermedia, lo que lo lleva a tener conocimientos sobre la comunidad discursiva pero no lo suficiente como para interactuar eficazmente en la totalidad de los géneros de dicha comunidad. Finalmente, el lector lego es el miembro con menor grado de experticia. Asimismo, las relaciones discursivas prototípicas corresponden a las entre un escritor experto y un lector experto, entre un escritor experto y un lector semilego y entre un escritor experto y un lector lego; sin embargo, aunque los géneros se posicionan como elementos identitarios de las comunidades discursivas, algunas veces su circulación no se limita a una sola comunidad, por tanto, puede llegar a existir la relación entre un escritor experto de una comunidad discursiva y un lector miembro de otra comunidad (Parodi, Ibáñez y Venegas, 2009). Las relaciones discursivas se encuentran ilustradas en la siguiente figura de Parodi, Ibáñez y Venegas (2009):

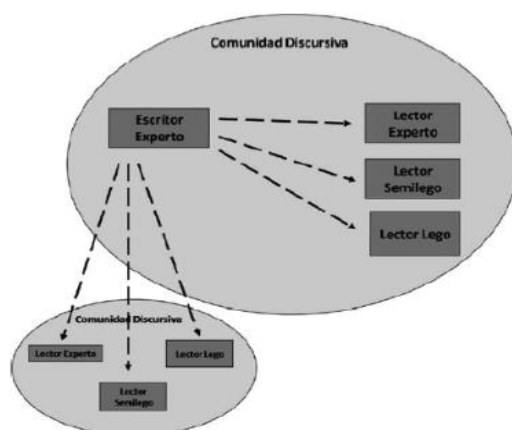


Figura 1. Relación entre participantes (Parodi, Ibáñez y Venegas, 2009, p.85)

De acuerdo con la información anterior, a grandes rasgos, una comunidad discursiva:

“Es un grupo humano que comparte unas prácticas comunicativas particulares, usando unos textos particulares, para conseguir unas finalidades específicas, entre autores y lectores que asumen roles predeterminados. A través de esas prácticas y de sus textos, los miembros de la comunidad llegan a compartir conocimientos específicos, unas habilidades cognitivas para procesarlos, unas actitudes y una determinada mirada sobre la realidad, que son exclusivas del grupo y que acaban constituyéndose como una marca de identidad frente a profesionales de otros ámbitos y frente a comunidades de otros ámbitos” (Cassany, 2008, p.2)

De esta idea se desprende que una comunidad discursiva incluye a miembros de una misma disciplina, que, en favor de integrarse correctamente a ella, deberán utilizar el lenguaje dentro de los márgenes particulares que determina la comunidad, incluyéndose, en primera instancia, como lectores, para luego pasar a convertirse en futuros escritores.

2.3 Géneros discursivos: Académicos y profesionales.

Cuando nos referimos a género, apuntamos a un concepto que ha sido estudiado desde diversos enfoques dentro del estudio del lenguaje, pasando por múltiples campos y significados, por tanto, resulta complejo delimitar el concepto a una definición estricta. En primera instancia, Aristóteles (1998), se posiciona como el primer precedente de interés al estudio de los géneros, específicamente en el año IV a.C. Su propuesta dividía a los discursos en tres grupos: los forenses o jurídicos, los deliberativos o políticos y los epidícticos o de ocasión. La clasificación toma en cuenta las prácticas discursivas de la época, tomando en consideración sus actantes, finalidades, temas y formas verbales y no verbales. El legado de Aristóteles trascendió a otros autores, que retomaron sus aportes (Quintiliano, 1942; Perelman y Olbrecht Tyteca, 2009); no obstante, la propuesta permaneció sin muchas variaciones hasta la incursión de la Lingüística textual en los años setenta (Ciapusio, 1994).

Hacia los años 30, la propuesta de Bajtín (1999) en cuanto a género se sustentó en el escaso estudio respecto a ellos, limitado a ámbitos literarios y artísticos, dejando un vacío

considerable respecto a los enunciados que se distinguen de otros por debido a su naturaleza verbal común. En referencia a su propuesta, Bajtín (1999) sienta las bases del estudio de los géneros discursivos desarrollados posteriormente. De acuerdo con sus postulados, Bajtín (1999) afirma que los géneros discursivos nacen a partir de enunciados, ya sean orales o escritos, relativamente estables, elaborados en cada esfera del uso de la lengua en base a la interacción humana, lo que significa una gran riqueza de enunciados. De esta manera, los géneros discursivos, de acuerdo con la concepción Bajtiniana (1999) son manifestaciones diversas y heterogéneas, en donde su producción se vincula al desarrollo de las esferas de la praxis o comunidades discursivas.

En concreto, Bajtín (1999) clasifica a los géneros discursivos en dos grupos de acuerdo a su grado de complejidad: géneros discursivos primarios y géneros discursivos secundarios. Por un lado, los géneros primarios son aquellos que se sitúan en la esfera de la vida cotidiana, por tanto, son enunciados reales que surgen de la comunicación inmediata, como un diálogo íntimo o familiar. Por otro lado, los géneros secundarios surgen en condiciones de comunicación cultural más organizada y compleja, generalmente escrita, como “novelas, dramas, investigaciones científicas de toda clase, grandes géneros periodísticos, etc.” (Bajtín, 1999, p.250).

Remontándonos a la perspectiva del Inglés con propósitos específicos (ESP por sus iniciales en inglés), Swales (1990) define al género como un vehículo que permite el logro de propósitos comunicativos, definidos e identificados por los miembros de una comunidad de hablantes, que construyen los géneros a partir de convenciones, formas y contenidos particulares. Asimismo, concuerda con la propuesta de Bajtín (1999) en que los géneros se identifican a partir de su forma, estructura, contenido y expectativas de la audiencia. En torno a ello, Swales (1990):

“A genre comprises a class of communicative events, the members of which share some set of communicative purposes. These purposes are recognized by the expert members of the parent discourse community, and thereby constitute the rationale for the genre. This rationale shapes the schematic structure of the discourse and influences and constrains choice of content and style. Communicative purpose is

both a privileged criterion and one that operates to keep the scope of a genre as here conceived narrowly focused on comparable rhetorical action. In addition to purpose, exemplars of a genre exhibit various patterns of similarity in terms of structure, style, content and intended audience.” (Swales, 1990, p. 58).

En los años posteriores, desde una perspectiva psicosociodiscursiva, Parodi (2008) concibe al género desde una mirada más integradora, en función a tres dimensiones: Social, Lingüística y Cognitiva. Estas dimensiones son esenciales y están relacionadas, pero el autor se focaliza en la dimensión cognitiva, debido a que es la que ha recibido menos atención en los estudios del género. En concreto, de acuerdo con los tres ámbitos mencionados, Parodi (2008) define género discursivo como:

“(…) una constelación de potencialidades de convenciones discursivas, sustentada por lo conocimientos previos de los hablantes/escritores y oyentes /lectores (almacenados en la memoria de cada sujeto), a partir de constricciones y parámetros contextuales, sociales y cognitivos...En su manifestación concreta, los géneros son variedades de una lengua que operan a través de un conjunto de rasgos lingüístico-textuales co-ocurrentes sistemáticamente a través de las tramas de un texto y que se circunscriben lingüísticamente en virtud de propósitos comunicativos, participantes implicados (escritores y comprendedores), contextos de producción, ámbitos de uso, modos de organización discursiva, soportes y medios, etc.” (Parodi, 2008:26).

De esta forma, el género discursivo se conforma a través de la interacción de la dimensión cognitiva, social y lingüística, tomando en especial consideración los procesos cognitivos de los sujetos, capaces de configurar géneros a partir de elementos externos que son almacenados en la memoria, tales como el contexto (Parodi, 2008). Finalmente, es pertinente mencionar que la noción de género discursivo, en los últimos años, ha ocupado de forma creciente un rol fundamental en las investigaciones y propuestas ligadas al discurso académico (Navarro, 2014). Algunas de estas investigaciones han proporcionado volúmenes colectivos, como los de Parodi (2010), que buscan ofrecer un panorama completo de géneros discursivos “relevantes para la formación científico-académica y profesional” (Navarro, 2014, p. 34).

Recapitulando, los géneros que circulan en el ámbito académico y profesional se sitúan dentro de un *continuum* (Parodi, 2007), donde los sujetos novatos progresan desde los géneros menos especializados, ligados a su alfabetización disciplinar, hasta llegar a los géneros más especializados, como los profesionales. En relación con concepto de género académico, este variará de acuerdo a la disciplina en la que circule y a la comunidad discursiva en la que se sitúe, lo que determinará su carácter académico (Bhatia, 2002). Asimismo, es importante retomar la noción de géneros importados, que situará a géneros de carácter profesional dentro del contexto universitario (Parodi, 2014), práctica habitual en el contexto de las ingenierías, debido a su constante contacto con el mundo laboral (Amieva, 2001).

2.4 Género: Informe de ingeniería

El informe escrito ha sido estudiado en conjunto con las investigaciones del discurso académico en sus diferentes manifestaciones. Estudios como el proyecto institucional DIPUC 2003 en Chile indican su importancia como medio para registrar los conocimientos adquiridos en la universidad (Harvey y Muñoz, 2006). De la misma manera, algunos de estos documentos “permiten aproximarse al nivel de manejo conceptual, textual y lingüístico que se despliega en el ámbito académico, y otros muestran la importancia de los informes como instrumentos para evaluar logros” (Harvey y Muñoz, 2006). Igualmente, se hace énfasis en la importancia del género informe en cuanto a la inserción de nuevos miembros dentro de la comunidad discursiva de la ingeniería, formando parte de un proceso formativo continuo de alfabetización disciplinar, que se liga a géneros particulares, en este caso, el informe (Harvey y Muñoz, 2006).

Según Amieva (2001) “un informe es una exposición escrita relativa a un tema, problema o actividad con propósitos formales de comunicación”, de esta forma, se posicionan como herramientas habituales, en ámbitos académicos como profesionales. De acuerdo a ello, los informes en el ámbito profesional comunican resultados, análisis situacionales, resúmenes de actividades, exposición de planes o propuestas, etc. (Amieva, 2001). En relación a la alfabetización académica, la elaboración de informes se vincula con la realización de trabajos prácticos de campo o laboratorio, la resolución de problemas, el desarrollo de

actividades de investigación, el registro de una actividad técnica o profesional como las realizadas en pasantías, etc., con el propósito de regular las prácticas de los alumnos, o medir sus conocimientos (Amieva, 2001). Bajo este punto, la solicitud de elaboración de informes dentro de la enseñanza de la ingeniería responde a dos tipos de motivaciones: profesionales y académicas. Por una parte, las investigaciones desarrolladas en el ámbito empresarial sugieren a la enseñanza promover la solicitud de informes claros y concisos, para desarrollar competencias fundamentales a nivel laboral, como la implementación de propuestas o situaciones en específico (Amieva, 2001). Igualmente, los informes ayudarían a los estudiantes a desarrollar habilidades no técnicas -o *soft skills*-, que les permitiría trabajar en grupo de manera más orgánica, comprender variables o implicancias básicas relacionadas con el marketing, los asuntos comerciales y financieros, comunicar ideas y sugerencias (Amieva, 2001).

En referencia al estudio de las características del género informe, Núñez y Espejo (2005), a través de un estudio realizado en torno a un cuestionario aplicado a docentes y estudiantes de la región Metropolitana de Chile respecto a la conceptualización del género informe, dieron cuenta de los rasgos y funciones que asignan a este género. En sus conclusiones, “las autoras plantean criterios para profundizar en el estudio de una posible correlación entre rasgos, clases de informes y áreas disciplinarias y sugieren la necesidad de incorporar la dimensión cognitiva a la investigación” (Harvey y Muñoz, 2006). Asimismo, los estudios de Oyanedel (2005) y Harvey (2005) en torno a informes producidos por estudiantes, han comprobado, sobre la base de evidencia funcional y textual, la existencia de al menos cuatro artefactos textuales llamados genéricamente informes: el informe de investigación, el estudio de caso, el informe bibliográfico y el informe de diagnóstico. Otro antecedente de estudio del género informe es el trabajo de Muñoz (2004), sobre estructura y patrones léxicos de las subcategorías anteriormente mencionadas, e identifica ciertas características asignables, en principio, a perspectivas disciplinarias (Harvey y Muñoz, 2006).

En el marco del estudio del informe, es pertinente mencionar la propuesta de Harvey y Muñoz (2006), que se basó en un estudio del género informe en torno a sus aspectos psico-

socio-pragmático, de acuerdo a las percepciones de los docentes respecto a las características retórico-funcionales del género. De acuerdo a sus resultados, los docentes encuestados consideran al informe como un tipo de texto expositivo breve (reporte o relación) que da cuenta de actividades, lecturas o temas particulares, con algunos componentes argumentativos (Harvey y Muñoz, 2006). En cuanto a estructura, se vincula directamente al esquema canónico “introducción-desarrollo-conclusión” y se posicionan como secciones esenciales del informe escrito; no obstante, la variación o denominación, o características de cada una de estas secciones varían de acuerdo a las distintas carreras y docentes. Por otra parte, según la investigación de Harvey y Muñoz (2006), los textos escritos que se excluyen del género informe, según los docentes encuestados, son los que:

- “Se estructuran en torno a una serie de preguntas, como es el caso de controles, pruebas de desarrollo y cuestionarios;
- Adoptan un estilo de escritura "libre" o, al parecer, menos regulado, como Ensayos, Comentarios, Escritura Creativa, y/o tratan temas "acotados", como es el caso de las Tareas” (Harvey y Muñoz, 2006).

En relación a la cita, es pertinente mencionar que existe una tendencia a diferenciar el informe escrito de otros textos académicos a partir de su función (Harvey y Muñoz, 2006). En específico, en referencia al informe en ingeniería, algunos docentes lo caracterizaron como “la presentación, modelación computacional y resolución de un problema, junto a las conclusiones del proceso” (Harvey y Muñoz, 2006). Cabe destacar que el género informe se caracteriza por ceñirse a aspectos formales, basados en formatos preestablecidos, como convenciones científicas, de redacción y elementos como bibliografía. Asimismo, se valora la claridad de un informe, de acuerdo a aspectos ligados a la redacción y organización de ideas (Harvey y Muñoz, 2006). Finalmente, se considera que serían informes conceptual y lingüísticamente apropiados, así como comunicativamente efectivos, aquellos que expliciten los vínculos entre representación mental y social, estructuración textual y descripción de la realidad (Harvey y Muñoz, 2006).

2.5 Macropropósito comunicativo

Previo a situarnos en la definición de macropropósito comunicativo, es pertinente entender la noción de propósito comunicativo, desde la perspectiva del análisis de género. En el marco del inglés con propósitos específicos (ESP por sus siglas en inglés), el análisis de género fue pensado como un método útil para la organización de propósitos comunicativos, que orientan al género a describir el uso del lenguaje en distintos contextos académicos y profesionales (Swales 1990; Bhatia 2002). A grandes rasgos, el propósito comunicativo es la función que cumple un género al interior de cierta comunidad discursiva (Sabaj y González, 2013). A su vez, es un concepto complejo, que explica a intención con la que el escritor produce un texto, según su contexto (Swales, 1990; Askehave y Swales, 2001). Cabe destacar, que la noción de propósito comunicativo se asemeja a lo que Charaudeau (2006) denomina “finalidad” de un texto, que depende de algunos factores, como el estatus o la identidad de los participantes en un intercambio comunicativo. Según Zamora y Venegas (2013), el propósito comunicativo de un texto permite diferenciar un género de otro, debido a que “la intención comunicativa del género es la que le da forma y le brinda una estructura esquemática interna” (Zamora y Venegas, 2013, p. 206). En base a esto, Swales (1990) propone que el propósito comunicativo es indispensable para llevar a cabo producciones o actividades lingüísticas dentro de una comunidad:

“A strong thread that binds the three key elements together is that a communicative purpose. It is communicative purpose that drives the language activities of the discourse community; it is communicative purpose that is the prototypical criterion for genre identity [...]” (Swales, 1990, p.10)

En concordancia con la cita anterior, el propósito comunicativo estará determinado de acuerdo a la comunidad discursiva en que se sitúe, donde los géneros discursivos funcionarán como medio para llevar a cabo dichos propósitos discursivos. En este sentido, la noción de propósito discursivo resulta fundamental para la definición y constitución de diversos géneros, académicos o no académicos (Askehave y Swales, 2001).

Al interior de un género, pueden existir más de un propósito comunicativo o coexistir una serie de “propósitos comunicativos menores” (Parodi, et al., 2015). Bajo este punto, un género discursivo integra un propósito comunicativo general, que apunta hacia un objetivo mayor, entendido bajo la noción de macropropósito comunicativo, que, a grandes rasgos, es el propósito de un género (Zamora y Venegas, 2013). De esta manera, el estudio del género permite organizar textos concretos en función del género, de acuerdo a su macropropósito comunicativo, contexto de circulación ideal, modo de organización del discurso, relación entre participantes y sistema semiótico en uso (Parodi, Venegas, Ibáñez y Gutiérrez, 2008). El estudio de género, en síntesis, ha posibilitado la descripción funcional de los géneros académicos, donde el Artículo de Investigación Científica (AIC) ha sido uno de los más estudiados (por ejemplo: Swales, 1990; Samraj, 2008; Sabaj, 2012).

En base a la descripción funcional de los géneros académicos, Swales (1990, 2004), con su modelo CARS (*create a Research Space*), propone una metodología de análisis de los propósitos comunicativos del AIC, mediante el reconocimiento de movidas y pasos presentes en la introducción. En este sentido, el texto se describe en términos de una secuencia de movidas, donde cada una de ellas está asociada a una unidad discursiva, que tiene un propósito comunicativo en particular (Ibáñez, 2010). Las movidas y pasos, como categorías, surgen a partir de la caracterización funcional de los textos. Una movida es “una unidad retórico-funcional que realiza un propósito comunicativo en un texto y que, a su vez, se realiza en una unidad textual (Askehave y Swales, 2001; Venegas, Zamora y Galdames, 2016) y un paso es un mecanismo discursivo, cuya presencia léxico-gramatical posibilita el desarrollo de una movida (Swales, 2004).

Como vimos en las propuestas abordadas en apartados anteriores, sobre propósitos y macropropósitos comunicativos y cómo estos se ven reflejados a nivel funcional y estructural en el texto, resulta imperante la necesidad de tener una base del comportamiento funcional correspondiente al género informe. En relación a los propósitos de esta investigación, nos centraremos en la propuesta de Harvey y Muñoz (2006) que consiste en un estudio acerca de las percepciones de los docentes respecto a las características retórico-funcionales del género informe, donde propone, a grandes rasgos,

que el género informe posee una estructuración tripartita, compuesta por introducción, desarrollo y conclusión, correspondiente a un esquema canónico del texto académico (Harvey y Muñoz, 2006). Además, propone que existe una representación social del género informe, en base a una concepción común de la situación retórica, que se origina a raíz de un propósito social compartido que, a su vez, se refleja en una estructuración representativa de la organización de la información y ciertas propiedades asociadas al género (Harvey y Muñoz, 2006). No obstante, la propuesta de Harvey y Muñoz (2006) no da cuenta de propósitos comunicativos explícitos en el género informe, ni tampoco se establece el macropropósito de género. En consecuencia, el modelo de Harvey y Muñoz (2006) solo da pistas de su estructuración y organización retórica-funcional, pero no lleva a cabo un análisis retórico discursivo, lo que constituye un vacío en el estudio del género informe.

En relación a la investigación, tomando en consideración las percepciones de los docentes, existen tres grandes divisiones en un informe escrito, lo que coincide con la noción de recorrido textual propuesta para la descripción por Reuter (2000, p. 79 -81):

- (i) Una sección inicial, destinada a *presentar y justificar* el tema o problema a tratar, presentando sus antecedentes teóricos (el *estado del arte*).
- (ii) Una sección media, que abarca el *procedimiento* para abordar el tema o problema, es decir, la metodología a utilizar; el tratamiento propiamente tal, caracterizado frecuentemente como *análisis*, o como los *resultados* del mismo. Esta sección se caracteriza por la aplicación de los antecedentes presentados en la sección anterior al asunto concreto tratado.
- (iii) Una sección final, orientada a extraer *conclusiones* del análisis realizado y, en la mayor parte de los casos, *comentarios* personales y/o aplicaciones de lo concluido.

A continuación, se presenta un cuadro con las denominaciones recurrentes estipuladas por los docentes encuestados respecto al género informe, donde se mencionan objetivos o términos asociados a cada sección del informe:

Sección		Nº Menciones
Inicial	<i>Introducción</i>	30
	<i>Objetivos</i>	24
	<i>Presentación del tema o problema</i>	23
	<i>Antecedentes</i>	20
	<i>Metodología</i>	17
	<i>Marco teórico</i>	16
	<i>Otras</i>	14
	<i>No se menciona</i>	18
	Total	162
Media	<i>Desarrollo</i>	46
	<i>Resultados</i>	37
	<i>Análisis</i>	25
	<i>Discusión</i>	23
	<i>Otras</i>	17
	<i>No se menciona</i>	14
	Total	162
Final	<i>Conclusiones</i>	39
	<i>Proyecciones/Aplicaciones</i>	36
	<i>Bibliografía / Referencias</i>	21
	<i>Comentarios personales</i>	19
	<i>Correlato teórico</i>	16
	<i>Otras</i>	13
	<i>No se menciona</i>	18
	Total	162

Figura 2. Cuadro de denominaciones recurrentes (Harvey y Muñoz, 2006).

En relación con el cuadro anterior, podemos concluir que:

“Existe consenso en situar al comienzo del informe todo aquello que sirva como base, fundamento o premisa para su desarrollo; en su parte media, el desarrollo propiamente tal, incluyendo sus resultados; hacia el final, las consecuencias que se deriven de dichos resultados, además de las referencias correspondientes. Sin embargo, también se observan variaciones en cuanto al tipo y número de denominaciones específicas usadas para describir las distintas secciones de un informe escrito, diferencias que son asignables, prioritariamente, a la naturaleza de la tarea solicitada y a su carácter más o menos profesionalizante, así como a la metodología de trabajo esperada, aspectos que ciertamente se relacionan también con el ámbito disciplinar, asunto que ilustramos en los apartados siguientes” (Harvey y Muñoz, 2006).

No obstante, es pertinente mencionar que de acuerdo con el análisis llevado a cabo en la propuesta de Harvey y Muñoz (2006), los informes de ingeniería civil no ponen gran énfasis en la estructuración del informe, centrándose mayoritariamente en la claridad de la exposición y en que el informe sea escrito de manera breve o sintética.

2.6 Macrogénero e informe: Macrogénero Académico Evaluativo

Durante años, la idea de identificar un principio para la clasificación de géneros ha sido recurrente dentro de los estudios del lenguaje (Jarpa, 2015). En base a dicho interés, Bazerman (1994) se refiere a sistemas de géneros, como un concepto para explicar el comportamiento de un conjunto de actos de habla insertos en marcos de actividades sociales predefinidas. Su objetivo principal es avanzar en una teoría de escritura que se base en explicar cómo estudiantes y profesores al interior de las aulas utilizan las herramientas discursivas de los géneros para interactuar con prácticas sociales que subyacen a ellas, como otras esferas de comportamiento social (Jarpa, 2015). Por su parte, bajo el mismo interés de clasificar géneros, Bhatia (2004) propuso el concepto de colonias de géneros, donde apuntaba a un tipo de relaciones genéricas entre géneros que tienen similares propósitos comunicativos, donde los miembros de la colonia comparten propósitos comunicativos, al igual que convenciones retóricas, contextos y características léxico-gramaticales; no obstante, pueden diferenciarse en cuanto a la afiliación personal, contexto de uso o relaciones entre participantes (Jarpa, 2015).

En relación al concepto de macrogénero, se entiende que es un término útil para agrupar a un conjunto de géneros (Martin, 2002). En ese sentido, Grabe (2002) y Berkenkotter (2002) utilizaron el concepto de macrogénero para “explicar la organización y agrupación de los textos que utilizan los profesores y estudiantes para construir y comunicar estructuras de pensamiento que permitan potenciar el desarrollo de los aprendizajes” (Jarpa, 2015). Asimismo, se concluyó que los textos predominantemente informativos (expositivos) son formas de comunicación prototípicas de las ciencias sociales, las ciencias y la historia, en contraposición con los textos narrativos (Grabe, 2002) y dentro de este ámbito, es posible encontrar múltiples tipos de géneros, como textos escolares, manuales, enciclopedias, etc., que son utilizadas habitualmente en la alfabetización de los alumnos sobre determinado tema (Jarpa, 2015). A su vez, la propuesta de Grabe (2002) y Martin (2002) sobre macrogénero, se basa en el principio de que mediante la organización de la información de la estructura textual es posible reconocer estructuras de conocimiento y, por lo tanto, modelar y potenciar el aprendizaje de los estudiantes (Jarpa, 2015).

Finalmente, Gabe (2002, p.255) señala: “One goal to promote learning from texts would be to explore the more generic frames or discourse structures that are used by writers and readers to convey and explain knowledge”.

Desde otra perspectiva, por parte de la Lingüística Sistémica Funcional la noción de macrogénero se vincula directamente al discurso pedagógico (Bernstein, 1990), que otorga unidades semánticas propias al registro escolar, dando a paso a géneros curriculares, que, a su vez, desembocan en unidades mayores llamadas macrogéneros curriculares (Mangui, 2009).

Centrándonos en el género informe y su pertenencia a un macrogénero, lo incluiremos en el Macrogénero Académico Evaluativo (Jarpa, 2015), dado el macropropósito comunicativo de “evaluar las competencias, aprendizajes, habilidades y dominio disciplinar adquirido por los estudiantes durante un proceso formativo en un determinado campo disciplinar” (Jarpa, 2012), compartido con otros géneros pertenecientes a esta clasificación (Jarpa, 2015). El Macrogénero Académico Evaluativo incluiría diversos géneros que buscan evaluar los aprendizajes en el ámbito académico, como los controles, informes, pruebas, ensayos y hasta tesis de grado (Venegas, 2010). En este sentido, el Macrogénero Académico Evaluativo consiste en:

“una colección de géneros textuales escritos por estudiantes (semilegos) y dirigidos a un profesor (experto) en el marco de una situación comunicativa de carácter didáctico-pedagógico, la que tiene como propósito evaluar la adquisición y dominio de conocimiento disciplinar y comunicativo durante el proceso de enseñanza-aprendizaje que se desarrolla al interior de un sistema de actividad académico” (Jarpa, 2015, p. 230-231).

Igualmente, en relación a la propuesta de Bazerman y Pior (2003), constituyen una colección de géneros, que reúne diferentes tipos de textos escritos, como pruebas, controles, reportes, etc., aplicados en el margen de la enseñanza, con el propósito de evaluar el avance en el desarrollo de habilidades psicolingüísticas y en el dominio de conocimiento especializado requerido por la comunidad académica a la que pertenecen (Jarpa, 2015). De igual forma, existen otros estudios que se centran en el análisis de estos

géneros académicos (Carter, 2007; Parodi, 2008; Gardner y Nesi, 2013), que son utilizados para medir o dar cuenta del conocimiento aprendido de los aprendizajes en su etapa formativa, cuya característica principal es la función didáctica que ocupan en su comunidad discursiva (Jarpa, 2015).

En base a la discusión llevada en este apartado, es pertinente situar al género informe en esta categoría, debido al rol evaluativo que juegan en torno a la formación del estudiante, en este caso, del futuro ingeniero civil informático. De esta manera, a través de la implementación de los informes dentro de la formación disciplinar, el docente podrá evaluar o medir los conocimientos, habilidades o dominio disciplinar de su pupilo, que, en caso de lograr un buen dominio de las competencias mencionadas anteriormente, podrá ser reconocido por los miembros expertos de su comunidad, logrando ser aceptado en ella (Jarpa, 2015)

A modo de observación, es imperante destacar que pese a que la clasificación de Jarpa (2015) sobre el Macrogénero Académico Evaluativo se emplea en el contexto del postgrado (Bermúdez, 2009); no obstante, sigue siendo útil para los propósitos de nuestra investigación, pues el propósito comunicativo de los textos del macrogénero, especificado anteriormente, igual aplica a un contexto de pregrado, que es desde donde nos situaremos para nuestro análisis.

3. METODOLOGÍA

En el siguiente apartado se especificarán los aspectos metodológicos empleados durante la investigación. Con este fin, se presentará el enfoque, alcance y diseño ejecutado en esta investigación. A continuación, se darán a conocer las preguntas de investigación que serán respondidas en el presente estudio, junto con el objetivo general y los objetivos específicos. Finalmente, presentará el corpus utilizado en el estudio y las herramientas que permitieron recolectar y clasificar los datos.

3.1 Tipo de Investigación

Esta investigación, denominada “Comparación léxico-gramatical en informes de tercer y quinto año de ingeniería civil informática en la PUCV”, se caracteriza por tener un enfoque metodológico de tipo cuantitativo, pues se busca caracterizar los informes de ingeniería civil informática en dos momentos de la carrera, en base a sus rasgos léxico-gramaticales. Específicamente, se describen y comparan los informes escritos por alumnos de tercer y quinto año, tomando en consideración el vocabulario especializado, los rasgos discursivos, la voz pasiva, la frecuencia de los trigramas de lemas y de categorías morfológicas (Pos, del inglés Part of Speech), además de los índices básicos de información textual.

El alcance de nuestra investigación es de tipo descriptivo, debido a que busca describir los rasgos léxico-gramaticales presentes en los informes de ingeniería civil informática en tercer y quinto año de la carrera, realizando una comparación entre los datos obtenidos de ambos niveles.

El objetivo general de la investigación es comparar los rasgos léxico-gramaticales en los informes de los alumnos de tercer y quinto año de ingeniería civil informática en la PUCV. Para ello, es fundamental presentar antecedentes conceptuales relevantes a la investigación, exponer los procedimientos metodológicos que guiaron la investigación, dar cuenta de los resultados y de su interpretación en el contexto de la investigación.

Refiriéndonos al diseño de la investigación, este será de tipo no experimental, puesto que no se manipulará ninguna variable y se analizarán informes que fueron escritos en su contexto natural de producción, sin ningún tipo de intervención por parte de los

investigadores. Según Hernández, Fernández y Baptista (2000) en este tipo de investigaciones no se hace variar en forma intencional las variables independientes para ver su efecto en otras variables, sino que se observa el fenómeno en su contexto natural, tal como es, para luego analizarlo (Lillo, 2018).

En esta investigación, los índices básicos de información textual, los trigramas de lemas y los trigramas de PoS fueron procesados mediante la herramienta DescApp de la plataforma PACTE, desarrollada por el grupo Redilegra, que incluye la detección de estos rasgos entre sus funcionalidades (Lillo, 2018). De igual manera, en esta plataforma se ha incluido una versión adaptada al español de SiNLP y algunos scripts propios que permiten determinar la voz pasiva en español e identificar diversos conectores discursivos, en función de identificar los rasgos discursivos en el corpus (Lillo, 2018). En relación con el vocabulario especializado, este fue procesado con la herramienta Termout.

3.2 Preguntas de Investigación

Las preguntas que se responderán con esta investigación son las siguientes:

¿Cuáles son los rasgos léxico-gramaticales de los informes de alumnos de tercer año en comparación con los de alumnos de quinto año de ingeniería civil informática en la PUCV?

- ¿Cuáles son los rasgos léxico-gramaticales más frecuentes en los informes de alumnos de tercer y quinto año de ingeniería civil informática en la PUCV?
- ¿Qué rasgos léxico-gramaticales tienen en común los informes de tercer año de ingeniería civil informática en la PUCV y los de quinto año?
- ¿Qué diferencias léxico-gramaticales tienen los informes de alumnos de tercer y quinto año de ingeniería civil informática en la PUCV?

3.3 Objetivos

La presente investigación tuvo los siguientes objetivos:

3.3.1 Objetivo general

Comparar los rasgos léxico-gramaticales en los informes de los alumnos de tercer y quinto año de ingeniería civil informática en la PUCV

3.3.2 Objetivos específicos

- Identificar los rasgos léxico-gramaticales en los informes de alumnos de tercer y quinto año de ingeniería civil informática en la PUCV.
- Caracterizar, a través de las particularidades léxico-gramaticales, a los informes de tercer año y quinto año de ingeniería civil informática en la PUCV.
- Comparar el uso de vocabulario especializado y los rasgos discursivos en los informes de tercer año y quinto año de ingeniería civil informática en la PUCV.

3.4 Corpus

La muestra de investigación está compuesta por un total de 30 informes de alumnos que cursan ingeniería civil informática en la PUCV. De ellos, 10 corresponden a informes de tercer año y 20 a alumnos de quinto año, realizados entre los años 2017 y 2019. Los informes de tercer año fueron realizados en el marco de la asignatura “Programación Avanzada” y los de quinto año fueron realizados en el ramo “Formulación de proyectos”, del noveno semestre de la carrera. En la siguiente tabla se mostrará la cantidad de palabras correspondientes a los informes de cada nivel, y la cantidad total de palabras del corpus:

Nivel	Cantidad de informes	Cantidad de Palabras
Tercer año	10	22075
Quinto año	20	49038
Total	30	71113

Tabla 1. Cantidad de palabras del corpus

El corpus se presentará en dos tablas, cada una con el código utilizado para distinguir y procesar el informe, acompañado con el título asignado por el alumno o grupo de trabajo. A continuación, se presentará en la Tabla 2 los informes correspondientes a tercer año de ingeniería civil informática.

Código	Título
INF_TERCER1	Informe Procesamiento datos medidor eléctrico “ElecMed”
INF_TERCER2	Informe FoodAD
INF_TERCER3	Informe Sistema de tutorías
INF_TERCER4	Informe LOKTAR
INF_TERCER5	Informe sistema de gestión en gastos de condominios
INF_TERCER6	Informe aplicación vida saludable
INF_TERCER7	Informe WEEBREADRER
INF_TERCER8	Informe Banana Task Enterprise
INF_TERCER9	Informe School Moodle
INF_TERCER10	Informe Sistema de gestión de constructoras LOCCET

Tabla 2. Informes de tercer año de ingeniería civil informática en la PUCV.

En la Tabla 3 se presentan los informes correspondientes al ramo “Formulación de proyectos”, que se imparte en quinto año de la carrera:

Código	Título
INF_QUINTO1	Informe WebScaping
INF_QUINTO2	Informe visualización de servidores
INF_QUINTO3	Informe solución basada en Docker
INF_QUINTO4	Informe Motores de base de datos
INF_QUINTO5	Informe BPMN, Bus de servicio, Capas transaccionales
INF_QUINTO6	Informe Firewal, Proxys y Switch
INF_QUINTO7	Informe Desarrollo Soluciones Móviles
INF_QUINTO8	Informe Framework para desarrollos Java
INF_QUINTO9	Informe Solución .Net vs Java
INF_QUINTO10	Informe Solución basado en node.js y express.js
INF_QUINTO11	Informe AngularJS
INF_QUINTO12	Informe Sistemas distribuidos
INF_QUINTO13	Informe API REST
INF_QUINTO14	Informe ataque DoS
INF_QUINTO15	Informe ERP
INF_QUINTO16	Informe Google Drive
INF_QUINTO17	Informe Comunicación VolP
INF_QUINTO18	Informe Business Intelligence
INF_QUINTO19	Informe Conducción Autónoma
INF_QUINTO20	Informe Peer to Peer

Tabla 3. Informes de quinto año de ingeniería civil informática en la PUCV.

3.5 Procedimientos

Para la recolección del corpus, se acudió personalmente a los profesores que imparten las asignaturas “Programación avanzada” de tercer año y “Formulación de proyectos” de quinto año en la Escuela de Ingeniería Civil Informática en la PUCV. Los docentes nos facilitaron los informes en formato .pdf y .docx. Posteriormente, se guardaron los documentos en archivo .txt, para poder ingresarlos a las diferentes herramientas de análisis: La herramienta DescApp de la plataforma PACTE y la herramienta Termout. A continuación, se seleccionaron las variables a investigar en el corpus, a saber: Índices básicos de información textual, vocabulario especializado, rasgos discursivos, voz pasiva, trigramas de lemas y trigramas de POS.

Una vez obtenido el corpus, se le asignaron códigos a cada uno de los informes, divididos en los etiquetados “INF_TERCER”, para los informes de tercer año e “INF_QUINTO” para los informes de quinto año. De acuerdo con esta clasificación, el análisis léxico-gramatical de estos documentos se realizó ingresándolos a la herramienta DescApp y a la herramienta Termout para el vocabulario especializado. En primer lugar, se ingresó cada informe de manera individual a la herramienta DescApp, pasando por cada una de las funciones pertinentes a las variables seleccionadas. En segundo lugar, se crearon dos documentos: El primero, con código “INF_TERCER_TOTAL” contiene todos los informes de tercer año en un sólo archivo .txt; el segundo, con código “INF_QUINTO_TOTAL” contiene todos los informes de quinto año en un archivo conjunto en formato .txt. Ambos documentos de texto fueron procesados por la herramienta DescApp, para obtener información general respecto a los dos niveles comparados en esta investigación. En este punto, los archivos también fueron procesados por la herramienta Termout, para obtener el vocabulario especializado de cada nivel.

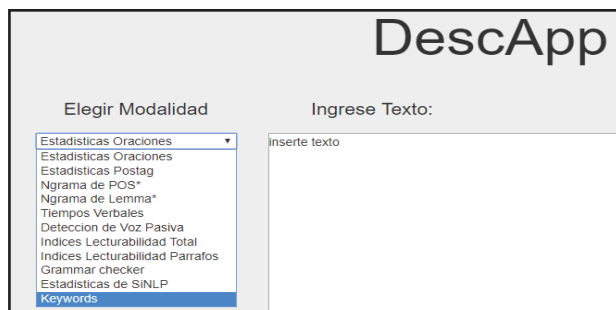


Figura 3. Opciones disponibles en la herramienta DescApp de la plataforma PACTE

Luego de la obtención de los resultados de DescApp y Termout, los datos se organizaron en tablas de Excel, para su organización. Posterior a ello, en cuanto a los índices de información textual, rasgos discursivos y voz pasiva, se comprobó la normalidad de datos con Shapiro-Wills y se realizó una prueba de muestras independientes con el test de Mann-Whitney, utilizando la herramienta Jasp. Para las otras variables, a saber, vocabulario especializado, trigramas de lemas y trigramas de POS, se establecieron las similitudes, en base a los elementos que aparecían en ambos niveles y diferencias, a partir de los elementos distintivos de cada año de estudio.

3.6 Variables

Las variables a investigar son las siguientes:

3.6.1 Rasgos léxico-gramaticales

Los rasgos léxico-gramaticales corresponden a los elementos que componen la estructura de los textos. Se define como la combinación de la gramática y el vocabulario, es decir, aquellos recursos empleados para construir la expresión del contenido que un escritor desea plasmar en un texto (Matthiessen, Teruya y Lam, 2010). En esta investigación los rasgos se operacionalizarán de la siguiente manera:

3.6.1.1 Índices básicos de información textual

Para esta investigación, se entenderá como índices básicos de información textual el número de palabras, oraciones y párrafos (Lillo, 2018). De igual manera, se incluirán al análisis los tipos de palabras, número de letras por palabra y número de palabras por oración. Con este fin, se utilizará una variación en español de SINLP dentro de la herramienta DescApp de la plataforma PACTE, desarrollada por el grupo Redilegra, para calcular la mayoría de los aspectos descriptivos del texto.

SINLP	"number words": 49038, "number types": 8697, "TTR": 0.17735225743301114, "Letters per word": 5.267506831436845, "Number paragraphs": 1369, "Number of sentences": 2152, "Number of words per sentence": 22.78717472118959,
-------	--

Figura 4. Opciones para índices básicos de información disponibles en la versión en español de SINLP de la herramienta DescApp.

3.6.1.2. Rasgos discursivos

Los rasgos discursivos son unidades lingüísticas que cumplen una función discursiva, permiten el cumplimiento de un propósito comunicativo, pues permiten instanciar un propósito específico en el discurso mismo. En este caso, se utilizará una variación en español de SINLP dentro de la herramienta DescApp de la plataforma PACTE. Particularmente, a la plataforma se le agregará un diccionario en español, donde se incluirán listas con marcadores discursivos, pronombres, negaciones y conectores empleados en el español, para medir su empleo en los informes de Ingeniería Civil Informática analizados.

SINLP	"SINLP_Demostrativos": 389, "SINLP_Pronombres": 8032, "SINLP_Pronombres_de_primera_persona": 31, "SINLP_Pronombres_de_segunda_persona": 8, "SINLP_Pronombre_de_tercera_persona": 934, "SINLP_Marcadores_discursivos": 307, "SINLP_Conectores": 1859, "SINLP_Negaciones": 320, "MD_Modalizadores_epistemicos_asertivos": 92, "MD_Modalizadores_epistemicos_no_asertivos": 9, "MDT_Adición": 88, "MDT_Aditivos_Indefinidos": 1, "MDT_Aditivos_Semejantes": 129, "MDT_Argumentativos": 48, "MDT_Conclusivos": 25, "MDT_Condicionales": 235, "MDT_Consecutivos": 207, "MDT_Contraargumentativos": 181, "MDT_Demanda": 5, "MDT_Disgresores": 0, "Distribuidores": 36, "MDT_Ejemplificadores": 358, "MDT_Finalizativos": 17, "MDT_Modalizadores": 40, "MDT_Ordenadores": 11, "Ordenadores_Temporales": 148, "MDT_Reformativos_con_inclusión": 81, "MDT_Sintetizadores": 0, "MDT_Tematizadores": 15
-------	--

Figura 5. Marcadores discursivos y tipos de palabras disponibles en la versión al español de SINLP de la herramienta DescApp

3.6.1.3. Vocabulario especializado

La variable vocabulario especializado, en esta investigación, será entendida como la utilización de palabras propias del discurso académico y profesional ligado a la comunidad discursiva en cuestión, en este caso, Ingeniería Civil en Informática. El uso de vocabulario especializado se vincula directamente a la alfabetización del estudiante en el transcurso de la carrera, que implica que el alumno se vaya apropiando de las características léxico-gramaticales o el registro de la comunidad discursiva. Para esta investigación, se calculará la frecuencia absoluta de cada palabra a través de la búsqueda de vocabulario especializado con la herramienta Termout disponible en <http://www.termout.org> (Nazar, 2016).

3.6.1.4. Voz pasiva

En general, se ha visto un incremento del uso de las formas pasivas en el lenguaje divulgativo informativo (Parodi, 2005). Según Ciapuscio (2002), las formas pasivas con “se” aparecen con frecuencia en textos de divulgación científica, para acentuar el carácter impersonal del texto (Parodi, 2005). De acuerdo a esta información, el uso de voz pasiva como índice de despersonalización del texto corresponde a la inserción en la comunidad científica y la producción del género para los estudiantes. En concreto, la tendencia a la omisión del agente es característica del lenguaje de la ciencia que se mantiene en la divulgación (Ciapuscio, 2002). Para el análisis, se incluirán pasivas con “se”, pasivas con “ser” sin agente, pasivas con “ser” con agente y pasivas con “estar”. Para la detección de las formas pasivas se utilizará la herramienta DescApp de la plataforma PACTE, que entre sus funciones contempla el procesamiento de la voz pasiva en los textos.

3.6.1.5. Trigramas de lemas

En la teoría lexicográfica, un lema corresponde a la forma canónica a la que se reduce un paradigma flexivo y que se toma como representación de todas las variantes morfológicas de la palabra (Alcaraz y Martinez, 1993). Los trigramas de lemas corresponden a n-grams compuestos por tres unidades léxicas, que dan cuenta de la probabilidad de ocurrencia de una combinación prototípica de elementos en los textos del corpus, en este caso, los

informes de Ingeniería Civil Informática. En lo que refiere a esta investigación, obtendremos los trigramas de lemas con la herramienta DescApp de la plataforma PACTE.

Trigrama de lema	[["basar de dato", "40", "0.17928376137331362"], ["de lo aplicación", "31", "0.13894491506431805"], ["de uno proyectar", "28", "0.12549863296131952"], ["el nombrar de", "26", "0.11653444489265385"]
------------------	--

Figura 6. Trigramas de Lema extraídos de la herramienta DescApp

3.6.1.6. Trigramas de POS

Los trigramas de POS, al igual que los trigramas de lemas, son n-grams compuestos por tres unidades. En este caso, corresponden a tres unidades de las distintas categorías gramaticales de las palabras presentes en el corpus de informes de Ingeniería Civil Informática. Las categorías gramaticales pueden ser: sustantivo, adjetivo, artículo, pronombre, verbo, adverbio, preposición, conjunción. En concreto, a través de los trigramas de POS se estudian las combinaciones más prototípicas de las categorías gramaticales presentes en los textos del corpus. De acuerdo con esta investigación, al igual que con los trigramas de lemas, obtendremos los trigramas de POS con la herramienta DescApp de la plataforma PACTE.

Trigrama de POS	[["(ADP__AdpType=Prep) (DET__Definite=DefGender=Fem Number=Sing PronType=Art) (NOUN__Gender=Fem Number=Sing) ", "362", "1.441599299111943"], ["(DET__Definite=DefGender=Fem Number=Sing PronType=Art) (NOUN__Gender=Fem Number=Sing) (ADP__AdpType=Prep) ", "226", "0.9000039823185059"], ["(DET__Definite=DefGender=Masc Number=Sing PronType=Art) (NOUN__Gender=Masc Number=Sing) (ADP__AdpType=Prep) ", "169", "0.6730118274859623"]
-----------------	---

Figura 7. Trigramas de POS extraídos de la herramienta DescApp

4. RESULTADOS

En esta sección, se presentarán, interpretarán y compararán los resultados obtenidos en la investigación realizada. Bajo esta premisa, cada uno de los subapartados, organizados en razón de las variables definidas en el apartado anterior, darán cuenta de los rasgos léxico-gramaticales que caracterizan a tercer y quinto año de Ingeniería Civil Informática, señalando los puntos en común y diferencias entre cada nivel, en favor de la comparación.

4.1 Índices básicos de información textual

Para calcular los índices básicos de información textual de los informes de tercer y quinto año, se empleó un análisis estadístico a los datos obtenidos de la versión en español de SINLP disponible en la herramienta DescApp. En primer lugar, se comprobó la normalidad de los datos con Shapiro-Wilks, utilizando la herramienta JASP. Debido a que los datos no se distribuían de manera normal de acuerdo al test Shapiro-Wilks, se compararon las frecuencias relativas de cada categoría de análisis de los informes para ver si existían o no diferencias estadísticas, aplicando el test de Mann-Whitney, también disponible en la herramienta JASP. En la siguiente tabla, se presentan los promedios correspondientes a cada año en los índices de información textual:

Índices de información textual	Inf_quinto	Inf_tercero
Número de palabras	0,034	0,031
Número de oraciones	0,031	0,037
Número de palabras por oración	0,036	0,027
Número de letras por palabra	0,033	0,033
Número de párrafos	0,031	0,039
Número de tipos de palabras	0,035	0,029

Tabla 4. Promedio relativo de estadísticas descriptivas de índice de información textual.

En relación con la tabla anterior, en primera instancia, se puede evidenciar que existen diferencias entre los promedios de cada nivel, quinto y tercero. A continuación, se presenta un gráfico con el comportamiento de los datos en ambos niveles.

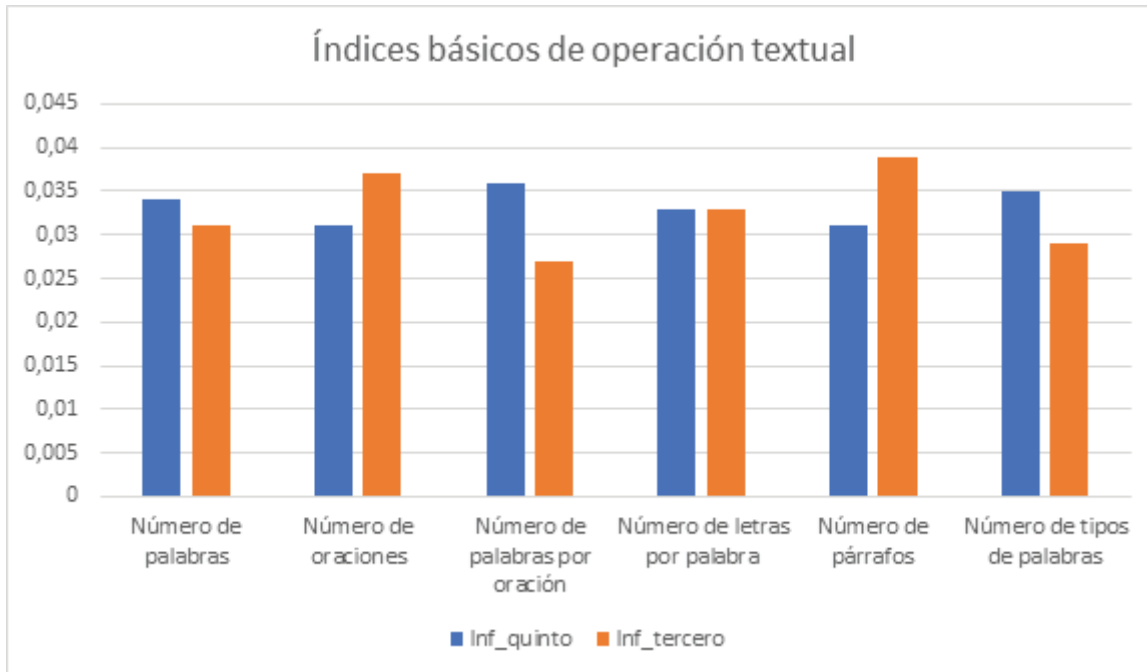


Gráfico 1. Índices básicos de operación textual.

Como se mencionó anteriormente, para poder ratificar si estas diferencias eran significativas en términos estadísticos, se realizó una prueba de muestras independientes utilizando el test de Mann-Whitney. En la tabla 5 se presentan los datos obtenidos por cada índice de información textual, donde, si P-value es menor a 0,05, significa que existen diferencias estadísticas entre los niveles comparados:

Índices de información textual	W	P-value	Resultado
Número de palabras	121.500	0.355	No hay diferencias
Número de oraciones	81.500	0.428	No hay diferencias
Número de palabras por oración	171.000	0.002	Sí hay diferencias
Número de letras por palabra	108.500	0.708	No hay diferencias
Número de párrafos	88.000	0.612	No hay diferencias
Número de tipos de palabras	129.000	0.209	No hay diferencias

Tabla 5. Resultados de índice de información textual del test de Mann-Whitney.

Tal como indica la tabla anterior, la única categoría en donde existen diferencias estadísticas es en el número de palabras por oración, donde, tomando en consideración los promedios de la tabla 4, quinto año tiene un mayor número de palabras por oración, con un

promedio de 0,036, por sobre el 0,027 de tercer año. Bajo este punto, los resultados de la prueba indican que, si bien, no existe diferencia estadística en el número de oraciones entre ambos niveles, las oraciones de quinto tienden a contener un mayor número de palabras y, por tanto, ser más extensas.

4.2 Rasgos discursivos

Con el fin de determinar los rasgos discursivos presentes en los informes de tercer y quinto año, al igual que con los índices de información textual, se realizó un análisis estadístico a partir de los datos obtenidos de la versión en español de SINLP, disponible en la herramienta DescApp, que, entre sus funcionalidades, detectaba una gran cantidad de marcadores discursivos, conectores, negaciones y pronombres. Para comenzar, se comprobó la normalidad de los datos con Shapiro-Wilks, utilizando la herramienta JASP. Al igual que en el apartado anterior, los datos no se distribuían de manera normal en base al test Shapiro-Wilks, por lo que se procedió a comparar las frecuencias relativas de cada rasgo discursivo, para ver si existían o no diferencias estadísticas, aplicando el test de Mann-Whitney. A continuación, en la tabla 6, se presentan los rasgos discursivos seleccionados con sus promedios correspondientes.

Rasgo discursivo	Inf_quinto	Inf_tercero
Pronombres 1era persona	0,032	0,035
Pronombres 2da persona	0,05	0
Pronombres 3ra persona	0,033	0,033
Marcadores discursivos	0,037	0,027
Conectores	0,035	0,03
Negaciones	0,036	0,028
MD Modalizadores epistémicos asertivos	0,034	0,031
MD Modalizadores epistémicos no asertivos	0,036	0,027
MDT Adición	0,028	0,043
MDT Aditivos indefinidos	0,05	0
MDT Aditivos semejantes	0,036	0,028

MDT Argumentativos	0,038	0,025
MDT Conclusivos	0,03	0,041
MDT Condicionales	0,037	0,026
MDT Consecutivos	0,037	0,027
MDT Contraargumentativos	0,039	0,021
MDT Demanda	0,044	0,013
MDT Distribuidores	0,045	0,01
MDT Ejemplificadores	0,039	0,022
MDT Finalizativos	0,025	0,05
MDT Modalizadores	0,036	0,029
MDT Ordenadores	0,043	0,013
MDT Ordenadores temporales	0,033	0,034
MDT Reformulativos con inclusión	0,04	0,021
MDT Tematizadores	0,042	0,017
Demostrativos	0,033	0,035
Pronombres	0,033	0,033

Tabla 6. Rasgos discursivos y promedio de estadísticas descriptivas.

Los resultados de la tabla anterior, en primera instancia, evidencian que existen diferencias entre los promedios de cada nivel, quinto y tercero. En algunos casos es el mismo promedio, a saber, en los pronombres de tercera persona y en los pronombres en general. A continuación, se presenta un gráfico que ilustra el comportamiento de los rasgos discursivos en ambos niveles.

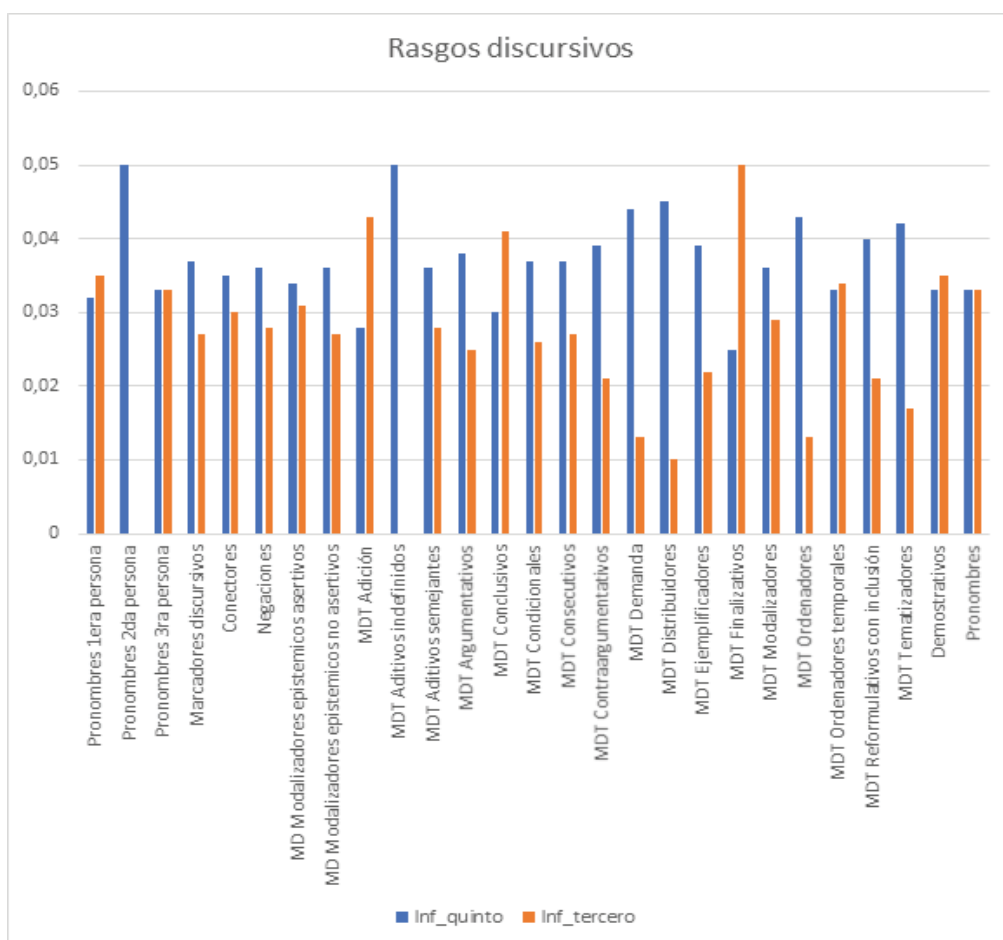


Gráfico 2. Rasgos discursivos.

Como se mencionó anteriormente, para poder ratificar si estas diferencias eran significativas en términos estadísticos, se realizó una prueba de muestras independientes utilizando el test de Mann-Whitney. En la tabla 7 se presentan los resultados de cada rasgo discursivo, donde, si P-value es menor a 0,05, significa que existen diferencias estadísticas entre los niveles comparados.

Rasgo discursivo	W	P-value	Resultados
Pronombres 1era persona	93000	0.768	No hay diferencias
Pronombres 2da persona	NAN	-	Sí hay diferencias
Pronombres 3ra persona	109000	0.708	No hay diferencias
Marcadores discursivos	133500	0.145	No hay diferencias
Conectores	126000	0.262	No hay diferencias
Negaciones	138500	0.094	No hay diferencias

MD Modalizadores epistémicos asertivos	88500	0.624	No hay diferencias
MD Modalizadores epistémicos no asertivos	128500	0.213	No hay diferencias
MDT Adición	72000	0.221	No hay diferencias
MDT Aditivos indefinidos	Nan	-	Sí hay diferencias
MDT Aditivos semejantes	121000	0.364	No hay diferencias
MDT Argumentativos	122000	0.335	No hay diferencias
MDT Conclusivos	80000	0.373	No hay diferencias
MDT Condicionales	133000	0.152	No hay diferencias
MDT Consecutivos	138000	0.098	No hay diferencias
MDT Contraargumentativos	159500	0.009	Sí hay diferencias
MDT Demanda	120500	0.232	No hay diferencias
MDT Distribuidores	141000	0.059	No hay diferencias
MDT Ejemplificadores	162000	0.006	Sí hay diferencias
MDT Finalizativos	70500	0.183	No hay diferencias
MDT Modalizadores	108500	0.706	No hay diferencias
MDT Ordenadores	128.00	0.134	No hay diferencias
MDT Ordenadores temporales	110500	0.657	No hay diferencias
MDT Reformulativos con inclusión	151000	0.024	Sí hay diferencias
MDT Tematizadores	124000	0.227	No hay diferencias
Demostrativos	105000	0.843	No hay diferencias
Pronombres	113500	0.567	No hay diferencias

Tabla 7. Resultados de rasgos discursivos de la prueba de Mann-Whitney para rasgos discursivos.

En base a los resultados de la tabla anterior, podemos afirmar que existen 5 rasgos que presentan diferencias significativas en términos estadísticos: Pronombres en segunda persona, marcadores discursivos aditivos indefinidos, marcadores discursivos contraargumentativos, marcadores discursivos ejemplificadores y marcadores discursivos reformulativos. En primer lugar, en cuanto a los pronombres en segunda persona, el promedio de quinto año es de 0,05 en contraste con el promedio de tercer año, que corresponde a cero, convirtiéndose en un rasgo que está ausente en ese año. La presencia de pronombres de segunda persona en quinto año es, en concreto, a causa de la orientación hacia el cliente que caracteriza a estos informes, donde el trato es dirigido a un interlocutor al cual debe describir una propuesta o solución informática y cuáles son las ventajas y desventajas de usarla. En segundo lugar, los marcadores discursivos aditivos indefinidos se presentan en quinto año con un promedio de 0,05, mientras que en tercer año la cifra

corresponde a cero. La presencia de aditivos indefinidos como “entre otros aspectos”, del corpus de informes de quinto año, marca la diferencia entre los niveles comparados.

Igualmente, en tercer lugar, los marcadores discursivos contraargumentativos se presentan en mayor cantidad en quinto año que en tercer año, en términos estadísticos, con un promedio de 0.039 para quinto y 0.021 para tercero. La mayor presencia de contraargumentativos se basa en la naturaleza del informe de quinto año, orientada a indicar las ventajas y desventajas de un servicio o producto informático, por tanto, es común el uso de este tipo de marcadores discursivos. En cuarto lugar, en referencia a los marcadores discursivos ejemplificadores, estos presentan un mayor promedio en quinto año, con un 0.039, en oposición con el 0,022 de tercer año. En relación a las particularidades de los informes de quinto año, explicadas con anterioridad en este apartado, se suele entregar al cliente una gran cantidad de ejemplos sobre el uso del servicio o producto informático que se habla en el informe, dando cuenta de las empresas que lo utilizan y de las múltiples funciones que cumplen dentro de la disciplina. Finalmente, en quinto lugar, los marcadores discursivos reformulativos con inclusión presentan un promedio de 0.04 en quinto año, superior al 0,021 de tercer año. Esto se debe a la fuerte presencia de reformulativos con inclusión como “incluso”, “de todas maneras” y “en algunos casos”, presentes en gran medida en el corpus de informes de quinto año. Para ver ejemplos de los rasgos discursivos que estadísticamente se presentan más en quinto año, revisar anexo 7.2.

4.3 Vocabulario especializado

Para dar cuenta de la presencia de vocabulario especializado en el corpus, se procesó el total de informes de tercer y quinto año utilizando la herramienta Termout (Nazar, 2016), para calcular la frecuencia absoluta de cada término. En concreto, tercer año presentó un total de 296 términos con frecuencia igual o superior a 2, en cambio, quinto año presentó 556 términos. De todos los datos obtenidos, se seleccionaron los 20 primeros que se presentaban en ambos informes. En la siguiente tabla se presentarán los términos que se presentaban con mayor frecuencia en ambos niveles:

Término	5to	3ro
proyecto	47	227
Método	27	176
Usuario	155	157
Dato	349	133
Aplicación	219	97
Información	184	60
Archivo	73	56
Base	117	55
Plataforma	86	34
Java	118	33
Interfaz	30	30
Software	91	27
Empresa	109	26
Objeto	33	24
Diseño	24	24
Función	35	22
Solución	35	22
Gestión	22	22
Sistema	329	21
Cliente	82	21

Tabla 8. Frecuencias de términos en tercer y quinto año de Ingeniería Civil Informática.

En la tabla 8 se muestran los términos que aparecen en ambos niveles, tercer y quinto año de Ingeniería Civil informática. De estos datos, es posible apreciar que la mayoría son relativos del dominio de la informática, como “archivo”, “base”, “dato”, “aplicación”, “plataforma”, “java”, “interfaz”, y “software”. La presencia de vocabulario especializado ligado a la informática en ambos niveles implica la apropiación del discurso de la comunidad discursiva, arraigada al ejercicio académico de producir un informe, que, tanto en tercero como en quinto año, implica dominio de la disciplina, y conocimiento sobre el diseño y la implementación de propuestas, aplicaciones o servicios informáticos. De igual manera, subyacen términos como “proyecto”, “método”, “empresa”, “función”, “solución” y “tarea”; que están ligados a la práctica de la redacción del informe, ya sea el desarrollo de la programación de una aplicación, como es el caso de los informes de tercero, o la descripción de una herramienta o solución informática, en el caso de quinto.

A continuación, para dar cuenta del vocabulario especializado que distingue a cada nivel, se presentarán dos tablas, la primera contiene los 20 términos con mayor frecuencia de tercer año y la segunda contiene los 20 términos con mayor frecuencia de quinto año.

Término INF TERCER	Frecuencia
Fase	119
String	72
Programa	63
Alumno	55
Dieta	51
Gantt	42
carta Gantt	39
Diagrama	27
Jardín	26
Alimento	24
Contraseña	24
Identificador	23
Correo	21
Menú	21
Ejercicio	20
Tutor	20
Boolean	18
Actividad	16
Administrador	15

Tabla 9. Términos con mayor frecuencia en tercer año.

Tal como indica la tabla anterior, el término con mayor frecuencia que aparece de manera exclusiva en tercer año es “fase”, con una frecuencia de 119. Esto se debe, en un principio, a que los informes de tercer año están orientados a explicar la creación y diseño de una aplicación, por tanto, deben dar cuenta de las fases de esta, y de sus diferentes procesos. Debido a esta premisa, una gran cantidad de los términos recurrentes en tercer año son “programa”, “alumno”, “carta gantt”, “ejercicio”, “tutor” y “actividad”. El vocabulario utilizado está orientado al discurso académico y al ejercicio de producir un informe, ligado a las implicancias de diseñar y trabajar en una aplicación, así como la distribución de tareas entre los integrantes del grupo. A continuación, se presenta la segunda tabla, con los 20 términos de mayor frecuencia en quinto año.

Término INF_QUINTO	Frecuencia
Red	146
Servidor	139
Seguridad	94
Ventaja	92
Máquina	75
Negocio	75
Framework	70
Comunicación	69
Desventaja	69
Docker	66
Nivel	60
Google	55
Contenedor	50
sistema operativo	49
Vehículo	48
Almacenamiento	45
máquina virtual	44
Motor	44
Proxy	44

Tabla 10. Términos con mayor frecuencia en quinto año.

De acuerdo con la información de la tabla 10, el término más recurrente de quinto año es “red” con 146 coincidencias en la totalidad de informes de quinto. Este término está ligado a otros recurrentes en el nivel, como “servidor”, “máquina”, “máquina virtual”, “framework”, “docker”, “contenedor”, “sistema operativo”, “proxy” y “dispositivo”, todos pertenecientes al vocabulario especializado de la informática y que no aparecen en el corpus de informes de tercer año. El uso de este vocabulario en los informes de quinto año se basa en que estos se orientan a describir servicios, códigos, aplicaciones o herramientas informáticas avanzadas, destinadas a ser implementadas en el ámbito empresarial, en el contexto de una propuesta, de ahí el detalle a los complementos del servicio informático y sus vinculaciones con la empresa. De igual forma, es frecuente el uso de términos como “seguridad”, “ventaja”, “negocio”, “comunicación” y “desventaja”, que se vinculan al ámbito de la propuesta de una solución informática, donde es necesario conocer las ventajas y desventajas de utilizar un servicio, cuál es su costo y si es factible a la hora de ser implementado.

En síntesis, en lo que se refiere a vocabulario especializado, los términos más frecuentes utilizados en tercer año se orientan al trabajo universitario, a la organización del grupo y al diseño de la aplicación como una tarea en su formación como informáticos, por lo cual se orienta al discurso académico y a la práctica evaluativa. En contraste con los informes de tercer año, los de quinto están orientados preferentemente al ámbito empresarial, dando cuenta de las características esenciales de un servicio o herramienta informática, para posteriormente ser implementados en el contexto de una empresa, es decir, en quinto año el propósito es técnico comercial y se vincula al discurso profesional. Es importante destacar que ambos niveles tienen en común términos ligados al vocabulario especializado del conocimiento de la disciplina informática, dando cuenta de su inmersión en la comunidad discursiva. Para más ejemplos revisar anexo 7.3.

4.4 Voz pasiva

Para identificar la voz pasiva en el corpus de informes de tercer y quinto año de Ingeniería Civil Informática, se utilizó la herramienta DescApp de la plataforma PACTE. Estos datos fueron sometidos a un análisis estadístico en la herramienta JASP, donde, si comparamos la tasa de ocurrencias de pasivas, divididas por el total del texto, en quinto queda en 24,6 y en tercero 18,1. En cuanto a la razón, si dividimos la cantidad de pasivas por el número total de palabras de palabras de cada nivel, en quinto queda un 0,010 y tercero queda en 0,008. En la tabla 11 se presentan los promedios del uso de pasiva correspondientes a cada nivel.

Voz Pasiva	Promedio
INF_TERCERO	1.33
INF_QUINTO	3.61

Tabla 11. Promedios del uso de pasiva.

En concreto, las estadísticas descriptivas indican que existen diferencias entre los promedios de quinto y tercero. Para determinar si estas diferencias son estadísticamente significativas, considerando la brecha en el número de palabras del corpus de cada nivel, se aplicó el test de Mann-whitney, donde si el valor de p-value es inferior a 0,05 denota que

existen diferencias estadísticas. En la tabla 12 se presentan los resultados del test de Mann-Whitney a los datos obtenidos.

Variable	W	P-value	Resultado
Voz Pasiva	318.280	< 2.2e-16	Sí hay diferencias

Tabla 12. Resultados de test Mann-Whitney para voz pasiva.

Los resultados de la prueba indican que el valor de p-value es < 2.2e-16, por tanto, tomando en cuenta los promedios ambos años, mencionados anteriormente, quinto año tiene más pasivas distintas y también mayor frecuencia en el uso de estas, lo que en relación con el carácter impersonal que brindan las voces pasivas, corresponde a una mayor despersonalización en los informes de quinto, por sobre los de tercero. Para dar cuenta de las voces pasivas que aparecen en los informes de tercer y quinto año, a continuación, se mostrarán 20 ejemplos de las voces pasivas más frecuentes en ambos años.

Voz Pasiva	Q	T
"se encuentran",	22	3
"se encuentra",	20	8
"se utiliza",	18	6
"se encarga",	10	9
"se hace",	10	3
"se necesita",	10	1
"se utilizan",	9	1
"se ve",	7	2
"se busca",	6	2
"se trabaja",	6	2
"se tiene",	5	7
"se presentarán",	5	1
"se procederá",	5	1
"se muestra",	4	7
"se logra",	4	4
"se espera",	4	2
"se quiere",	4	1
"se presenta",	3	8
"se almacenan",	3	3
"se comunican",	3	2

Tabla 13. Frecuencias de voces pasivas en tercer y quinto año de Ingeniería Civil Informática.

Tal como indica la tabla anterior, las voces más pasivas que se registran en ambos niveles dan cuenta de la despersonalización al referirse a tareas o procedimientos respecto de la programación de una aplicación o la descripción de los componentes de una herramienta o servicio informático, como “se procederá”, “se hace”, “se encarga”, “se utilizan”, “se busca”, “se tiene”. De la misma forma, ejemplos de voz pasiva como “se logra”, “se espera”, “se quiere” apuntan a los objetivos del desarrollo o implementación de una aplicación, o a las diferentes funciones y potencialidades de los servicios, propuestas o herramientas informáticas. A continuación, se muestran los 20 ejemplos con mayor frecuencia de voz pasiva propios de tercer año de Ingeniería Civil Informática.

Voz pasiva INF_TERCER	Frecuencia
"se inicia",	5
"se especifica",	4
"es utilizado",	3
"ha elaborado",	3
"se agrega",	3
"se decide",	3
"se editan",	3
"se elimina",	3
"es sabido",	2
"ha dicho",	2
"ha propuesto",	2
"se acceda",	2
"se asigna",	2
"se asocia",	2
"se encargará",	2
"se guardarán",	2
"se hablará",	2
"se hizo",	2
"se implementan",	2

Tabla 14. Frecuencias de ejemplares únicos de voz pasiva en tercer año.

La voz pasiva en tercer año de Ingeniería Civil informática, de acuerdo con los 20 ejemplos con mayor frecuencia dentro del total del corpus, indican un enfoque hacia las

decisiones y procedimientos del grupo, dichas de manera impersonal, como en “se inicia”, “se asigna”, “se asocia”, “se encargará”, “se hizo”, “se implementan”. En la tabla 15 se presentan las voces pasivas propias de quinto año.

Voz Pasiva INF QUINTO	Frecuencia
"se requiere",	12
"se recomienda",	9
"se basa",	8
"se ejecuta",	7
"se mantienen",	7
"se realiza",	7
"se trata",	7
"se realizan",	6
"se refiere",	5
"han desarrollado",	4
"se aplica",	4
"se conoce",	4
"se crean",	4
"se diferencia",	4
"se dividen",	4
"se integran",	4
"se presentan",	4
"se tenga",	4
"se ubica",	4
"ha logrado",	3

Tabla 15. Frecuencias de ejemplares únicos de voz pasiva en quinto año.

Los resultados de los 20 ejemplares con mayor frecuencia propios de quinto año están relacionados con el propósito de los informes de este año, que es describir un servicio, herramienta o solución informática, dando cuenta de sus ventajas y desventajas, de su funcionamiento y de sus características. Vinculado a esta labor, es congruente que ejemplos de voz pasiva como “se requiere”, “se recomienda”, “se trata”, “se diferencia” o “se presentan” existan con alta frecuencia dentro de la totalidad de informes.

En síntesis, quinto año presenta una diferencia considerable en términos estadísticos respecto a la cantidad de voces pasivas distintas y la frecuencia del uso de estas, en contraste con tercer año. No obstante, el carácter impersonal del texto se articula en ambos niveles, que usan la voz pasiva para referirse a las utilidades de un servicio informático, como también para dar cuenta de los procedimientos realizados al programar una

aplicación. La voz pasiva en tercer año sigue la línea de las decisiones y procedimientos del grupo, patrón léxico-gramatical que también se vio reflejado en apartados anteriores. En contraste con el carácter de la voz pasiva en tercer año, quinto año suele orientar su uso a describir el servicio, integrando observaciones e implicancias de su implementación en el ámbito empresarial, aspecto también revisado en apartados anteriores. Para visualizar ejemplos respecto a esta categoría de análisis, ver Anexo 7.4.

4.5 Trigramas de lemas

Para identificar los trigramas de lemas dentro del corpus de informes, se utilizó la herramienta DescApp de la plataforma PACTE, donde se procesaron los dos archivos que contenían el corpus de informes correspondiente a tercer y quinto año. En concreto, se seleccionaron los 20 trigramas de lema con mayor frecuencia en cada nivel. A continuación, se presentarán los trigramas de lemas que quinto y tercero tienen en común.

Trigramas de lemas	5to	3ro
base de datos	99	40
la base de	29	24
a través de	48	19
por el que	40	14

Tabla 16. Trigramas de lemas en común para tercer y quinto año.

De acuerdo con la información ilustrada en la tabla anterior, tercer y quinto año sólo coinciden en cuatro trigramas de lema. En cuanto a los primeros dos, “base de datos” y la base de”, están estrechamente relacionados, dado que las bases de datos son tópicos comunes en ambos niveles de análisis. En el caso de “a través de”, corresponde a una locución preposicional, frecuente en la escritura académica y que puede equipararse a “por medio de”, que también coincide con el último trigramas común “por el que”. Con el objetivo de comparar los trigramas de lemas en ambos niveles, a continuación, se presentan los trigramas de lemas exclusivos para tercer año.

Trigrama de lema INF_TERCER	Frecuencia
de lo aplicación	31
de uno proyectar	28
el nombrar de	49
recibir el nombrar	25
de lo clase	23
en casar de	21
recibir uno ir	18
nombrar de uno	18
diagrama de clase	17
la cantidad de	16
proyectar recibir el	16
retornar true si	16
de proyectar recibir	14
de lo fase	14
uno ir y	14
método recibir uno	14

Tabla 17. Trigramas de lema únicos de tercer año.

Sobre los trigramas de lemas exclusivos de tercer año, como combinación prototípica de tres elementos recurrentes dentro del texto, el que tiene mayor frecuencia es “de lo aplicación”, con 40 coincidencias dentro del corpus, que en su forma léxica corresponde a “de la aplicación”. La recurrencia de este trigrama, en consideración de la temática de la programación de una aplicación, nos permite inferir que se habla de manera frecuente de la aplicación creada, detallando los pasos y las características de esta. De igual manera, algunos trigramas recurrentes son “de uno proyectar”, “proyectar recibir el” y “de proyectar recibir”, que en su forma léxica corresponden a “de un proyecto”, “proyecto

recibe el” y “de proyecto recibe”. Los tres trigramas mencionados se vinculan al proyecto del informe, que como hemos mencionado, es crear una aplicación. A continuación, en la tabla 18 se presentan los trigramas exclusivos de quinto año.

Trigramas de lema INF_QUINTO	Frecuencia
el usar de	41
ventaja y desventaja	29
hoy en día	25
de lo red	25
deber a que	25
este tipo de	24
a continuación se	23
lenguaje de programación	22
de lo sistema	21
la mayoría de	21
de base de	21
el que se	20
la posibilidad de	19
ancho de banda	19
lo sistema distribuir	19
Lo capa de	19

Tabla 18. Trigramas de lemas exclusivos de quinto año.

En cuanto a los trigramas de lema exclusivos de quinto año, la frecuencia mayor de trigramas es de 41 ocurrencias, correspondiendo a “el usar de”, que en su forma léxica equivale a “el uso de”. Este trigramas está ligado directamente al enfoque de los informes

de quinto año, orientado a la descripción de la implementación o uso de diversas herramientas o soluciones informáticas. De igual manera, un trigramas de lema “ventaja y desventaja” apunta a la descripción de las implicancias del uso cierto servicio o herramienta en una empresa, dando cuenta de los beneficios y las pérdidas al respecto de su implementación. Otros trigramas como “de lo red”, “lenguaje de programación”, “de lo sistema”, “ancho de banda” y “los sistema distribuir”, apuntan a conceptos desarrollados en los informes; mientras que “por el que”, “hoy en día”, “deber a que”, “este tipo de”, “a continuación se”, “la mayoría de” y “la posibilidad de” se orientan a aspectos ligados a la redacción de los apartados del informe, además de la descripción de los sistemas informáticos. A modo de síntesis, los trigramas de lemas correspondientes a cada nivel nos permiten interpretar que su uso está ligado directamente con los temas y conceptos abordados en los informes, que varían en relación con los propósitos de cada año analizado. Para ver más ejemplos, ver Anexo 7.5.

4.5 Trigramas de POS

Los trigramas de POS corresponden a n-grams formados por tres unidades de categorías gramaticales distintas. Tal como en apartados anteriores, los Trigramas de POS fueron extraídos del corpus de tercero y quinto año con la herramienta DescApp de la plataforma PACTE. En concreto, se seleccionaron los 10 con mayor frecuencia de cada nivel. A continuación, se presentan los trigramas que aparecen en ambos niveles.

Trigramas de POS	5TO	3RO
(ADP__AdpType=Prep) (DET__Definite=Def Gender=Fem Number=Sing PronType=Art) (NOUN__Gender=Fem Number=Sing)	562	362
(DET__Definite=Def Gender=Fem Number=Sing PronType=Art) (NOUN__Gender=Fem Number=Sing) (ADP__AdpType=Prep)	561	226
(DET__Definite=Def Gender=Masc Number=Sing PronType=Art) (NOUN__Gender=Masc Number=Sing) (ADP__AdpType=Prep)	386	169
(ADP__AdpType=Prep) (DET__Definite=Def Gender=Masc Number=Plur PronType=Art) (NOUN__Gender=Masc Number=Plur)	300	109

(ADP__AdpType=Prep) (DET__Definite=Def Gender=Masc Number=Sing PronType=Art) (NOUN__Gender=Masc Number=Sing)	292	132
(ADP__AdpType=Prep) (DET__Definite=Ind Gender=Masc Number=Sing PronType=Art) (NOUN__Gender=Masc Number=Sing)	217	123
(DET__Definite=Def Gender=Masc Number=Plur PronType=Art) (NOUN__Gender=Masc Number=Plur) (ADP__AdpType=Prep)	210	92

Tabla 19. Trigramas de POS en común para tercer y quinto año.

De acuerdo con la tabla anterior, el trigramas de POS con mayor recurrencia corresponde a “Preposición + Artículo definido femenino singular + Sustantivo femenino singular”. Este trigramas de POS es el más recurrente en ambos niveles, tercer y quinto año. De igual manera ocurre con los trigramas de POS “Artículo definido femenino singular + Sustantivo femenino singular + Preposición” y “Artículo definido masculino singular + Sustantivo masculino singular + Preposición”. La alta recurrencia de estos trigramas nos permite inferir que las combinaciones gramaticales en ambos niveles son similares. Con el propósito de delimitar las combinaciones gramaticales propias de cada nivel, a continuación, se presentarán los trigramas de POS únicos en tercer año.

Trigramas de POS INF_TERCER	FREC
(DET__Definite=Ind Gender=Masc Number=Sing PronType=Art) (NOUN__Gender=Masc Number=Sing) (ADP__AdpType=Prep) ",	139
(VERB__Mood=Ind Number=Sing Person=3 Tense=Pres VerbForm=Fin) (DET__Definite=Def Gender=Masc Number=Sing PronType=Art) (NOUN__Gender=Masc Number=Sing) ",	100
(VERB__Mood=Ind Number=Sing Person=3 Tense=Pres VerbForm=Fin) (DET__Definite=Ind Gender=Masc Number=Sing PronType=Art) (NOUN__Gender=Masc Number=Sing)	93

Tabla 20. Trigramas de POS exclusivos de tercer año.

En relación con las combinaciones gramaticales exclusivas de tercer año, la que tiene mayor frecuencia de ocurrencia es “Artículo definido masculino singular + Sustantivo masculino singular + Preposición” (139). De igual manera, los trigramas de POS “Verbo indicativo presente singular tercera persona finito + artículo definido masculino singular + sustantivo masculino singular” y “Verbo indicativo presente tercera persona singular finito

+ artículo indefinido masculino singular + sustantivo masculino singular” sólo varían en un elemento: El primero tiene un artículo definido, mientras que el segundo tiene uno indefinido. Estas configuraciones gramaticales, caracterizadas por contener un verbo, denotan el carácter de los informes de tercer año, donde abundan las acciones, fases y tareas metodológicas. En la tabla 21 se presentan los trigramas exclusivos de quinto año.

Trigrama de Pos INF_ QUINTO	Frec
(NOUN__Gender=Masc Number=Sing) (ADP__AdpType=Prep) (NOUN__Gender=Fem Number=Sing)	251
(ADP__AdpType=Prep) (NOUN__Gender=Fem Number=Sing) (ADP__AdpType=Prep)	236
(NOUN__Gender=Fem Number=Sing) (ADP__AdpType=Prep) (NOUN__Gender=Masc Number=Plur)	216

Tabla 21. Trigramas de POS exclusivos de quinto año.

Estos resultados indican que la combinación gramatical más recurrente en quinto año corresponde a “Sustantivo masculino singular + preposición + sustantivo femenino singular” con 251 ocurrencias. Igualmente, “Preposición + sustantivo femenino singular + preposición” y “Sustantivo femenino singular + preposición + sustantivo masculino plural” son combinaciones frecuentes en este nivel. En este sentido, solo interactúan tres categorías gramaticales en estos tres trigramas de POS (pronombre, artículo y sustantivo).

En síntesis, tercer y quinto año comparten 7 trigramas de POS con alta frecuencia, lo que corresponde a la mayoría de los trigramas analizados, que en este caso corresponden a 10 por nivel. La alta recurrencia de estos trigramas de POS nos permite inferir que existen combinaciones gramaticales similares en ambos niveles, sin embargo, al detenernos en los trigramas de POS exclusivos de cada año, nos encontramos con que tercero presenta dos trigramas de POS compuestos por verbos, además de las preposiciones, sustantivos y artículos, lo que lo diferencia de los trigramas de POS exclusivos de quinto año, que no poseen esa configuración y se limitan a las categorías gramaticales mencionadas anteriormente. Para ver más ejemplos, revisar anexo 7.6.

5. CONCLUSIONES

Para finalizar nuestra investigación, se abordará la problemática, junto a los objetivos que guiaron al estudio. De igual manera, se realizará una síntesis de los resultados obtenidos a partir del análisis efectuado. Por último, expondremos las principales limitaciones de nuestra investigación y propondremos algunas proyecciones para futuras investigaciones.

El estudio del informe escrito ha ido de la mano con las investigaciones sobre el discurso académico en sus diferentes manifestaciones, en relación con su pertinencia para la formación académica y profesional. No obstante, los estudios sobre los rasgos léxico-gramaticales del género informe son escasos. Propuestas como Harvey y Muñoz (2006) estudian el género informe y su representación en el discurso de los académicos, pero dista de detalles sobre una disciplina en específico. La relevancia de la presente investigación está en que, por medio de la comparación, se podrá ratificar si existe una variación efectiva en los patrones léxico-gramaticales a medida que los alumnos van progresando en la carrera Ingeniería Civil Informática. A raíz de esto, se podrá comprobar si los alumnos, a medida que avanzan en sus estudios, van integrándose en su comunidad discursiva, adoptando los rasgos y el vocabulario pertinente del conocimiento informático.

El objetivo general que guio el presente estudio fue “Comparar los rasgos léxico-gramaticales en los informes de los alumnos de tercer y quinto año de Ingeniería Civil Informática en la PUCV”. A través del desarrollo de la presente investigación se compararon los informes de ambos años, tercero y quinto, a partir de sus rasgos léxico-gramaticales, operacionalizados mediante las variables, a saber, índices básicos de información textual, rasgos discursivos, vocabulario especializado, voz pasiva, trigramas de lemas y trigramas de POS. Para extraer los índices básicos de información textual y rasgos discursivos se utilizó una versión al español de SINLP disponible en la herramienta DescApp de la plataforma PACTE, desarrollada por el grupo Redilegra. De igual manera, se utilizó esta herramienta para la detección de voz pasiva, trigramas de lemas y trigramas de POS. Para la detección del vocabulario especializado se recurrió a la herramienta Termout. Con respecto a los índices básicos de operación textual, voz pasiva y rasgos discursivos, se utilizó la herramienta JASP para verificar la normalidad de datos y

posteriormente se realizó una prueba de muestras independientes para determinar si existían diferencias estadísticas significativas entre las variables dispuestas para cada nivel. Para las demás categorías de análisis, vocabulario especializado, trigramas de lemas y trigramas de POS, se establecieron las similitudes en base a los elementos en común de ambos niveles, y las diferencias a partir de los elementos distintivos de cada año de estudio.

Con respecto a los resultados obtenidos en la comparación, en primer lugar, de acuerdo con los índices de información textual, en cuanto a número de palabras, número de oraciones, número de párrafos, número de letras por palabra y número de tipos de palabras no existen diferencias estadísticamente considerables. La diferencia radica en el número de palabras por oración, donde quinto se antepone a tercero, esto significa que, pese a que el número de oraciones entre ambos niveles no se diferencia de manera considerable, las oraciones de quinto tienden a contener un mayor número de palabras y a ser más extensas.

En segundo lugar, de acuerdo con los rasgos discursivos, no existen diferencias estadísticas considerables en los rasgos, a saber: pronombres de 1era persona, pronombres de 3ra persona, marcadores discursivos, pronombres, conectores, negaciones, modalizadores epistémicos asertivos, modalizadores epistémicos no asertivos y marcadores discursivos de adición, aditivos semejantes, argumentativos, conclusivos, condicionales, consecutivos, demanda, distribuidores, finalizativos, modalizadores, ordenadores, ordenadores temporales, tematizadores, demostrativos. Mientras que existen diferencias en cuanto a los pronombres en 2da persona y los marcadores discursivos aditivos indefinidos, contraargumentativos, ejemplificadores y reformulativos con inclusión, todos con un promedio superior en quinto por sobre tercero.

En tercer lugar, en lo que se refiere a vocabulario especializado, los términos más frecuentes utilizados en tercer año se orientan al trabajo universitario, por lo cual se orienta al discurso académico y a la práctica evaluativa. En contraste, los informes de quinto están orientados preferentemente al ámbito empresarial, es decir, en quinto año el propósito es técnico comercial y se vincula al discurso profesional. Es importante destacar que ambos niveles tienen en común términos ligados al vocabulario especializado del conocimiento de

la disciplina informática, dando cuenta de su inmersión en la comunidad discursiva. En cuarto lugar, en lo que respecta a la voz pasiva quinto año presenta una diferencia considerable en términos estadísticos respecto a la cantidad de voces pasivas distintas y la frecuencia del uso de estas, en contraste con tercer año. No obstante, el carácter impersonal del texto se articula en ambos niveles, que usan la voz pasiva para referirse a las utilidades de un servicio informático, como también para dar cuenta de los procedimientos realizados al programar una aplicación.

En quinto lugar, los trigramas de lemas correspondientes a cada nivel nos permiten interpretar que su uso está ligado directamente con los temas y conceptos abordados en los informes, que varían en relación con los propósitos de cada año analizado. Por último, en sexto lugar, en cuanto a los trigramas de POS, tercer y quinto año comparten 7 trigramas de POS con alta frecuencia, lo que corresponde a la mayoría de los trigramas analizados, que en este caso corresponden a 10 por nivel. La alta recurrencia de estos trigramas de POS nos permite inferir que existen combinaciones gramaticales similares en ambos niveles, sin embargo, al detenernos en los trigramas de POS exclusivos de cada año, nos encontramos con que tercero presenta dos trigramas de POS compuestos por verbos, además de las preposiciones, sustantivos y artículos, lo que lo diferencia de los trigramas de POS exclusivos de quinto año, que no poseen esa configuración.

La principal limitación de la presente investigación se relaciona con la dificultad de encontrar corpus, debido a que los informes realizados, al ser una práctica de la formación destinada a ser revisada solo por el profesor, no están publicados, por tanto, se tuvo que acudir directamente a los docentes que imparten las asignaturas para obtener los informes desde su almacenamiento personal, el que, en algunos casos, era inexistente, y en otros casos, guardaban muy pocos ejemplares. De igual manera, la falta de teoría sobre el informe escrito, principalmente sobre su organización retórica discursiva, hizo difícil la tarea de caracterizar al género. Finalmente, se considera que, de crearse un modelo retórico discursivo sobre el género informe, sería interesante ver cómo estos patrones léxico-gramaticales se reflejan en la organización retórica discursiva y los propósitos

comunicativos del informe, de esta manera, se podría obtener una caracterización completa del género.

6. REFERENCIAS BIBLIOGRÁFICAS

Alcaraz Varó, E., & Martínez Linares, M. A. (1993). *Diccionario de lingüística moderna*. Ciudad: Ariel.

Amieva, Rita. (2001) *Elaboración de informes en la enseñanza de la Ingeniería*. Río Cuarto, Argentina: Universidad Nacional de Río Cuarto

Aristóteles (1998). *La retórica*. Madrid: Alianza

Askehave, I., & Swales, J. M. (2001). Genre identification and communicative purpose: A problem and a possible solution. *Applied linguistics*, 22(2), 195-212.

Bajtín, M. (1999). *Estética de la creación verbal*. México D.F: Siglo Veintiuno.

Bazerman, C., y Prior, P. (eds.) (2003). *What writing does and how it does it: An Introduction to Analyzing Texts and Textual Practices*. Mahwah: Erlbaum.

Bazerman, Ch. & Paradis, J. (1991a). Introduction. En Ch. Bazerman & J. Paradis (Eds.) *Textual dynamics of the professions* (pp. 3-10). Wisconsin: The University of Wisconsin.

Bazerman, Ch. & Paradis, J. (Eds.) (1991b). *Textual dynamics of the professions*. Wisconsin: The University of Wisconsin.

Berkenkotter, C. (2002). Narrative and expository macro-genre. En A. Johns (ed.), *Genre in the Classroom. Multiple Perspectives* (pp. 285-288). Mahwah: Erlbaum.

Bermúdez, N. (2009). *Escritura y Producción de Conocimiento en las carreras de posgrado*. En E. Narvaja de Arnoux (dir.), *Escritura y Producción de Conocimiento en las Carreras de Posgrado* (pp.10-33). Buenos Aires: Santiago Arcos.

Bernstein, B. (1990). *Class, Codes and Control: The Structuring of Pedagogic Discourse*. Nueva York: Routledge.

Bhatia, V. (2002). Applied genre analysis: a multi-perspective model. *Ibérica: Revista de la Asociación Europea de Lenguas para fines específicos (AELFE)*, (4), 3-19.

Bhatia, V. (2004). *World of written discourse. A genre-based view*. Londres: Continuum

- Carter, M. (2007). Ways of knowing, doing, and writing in the disciplines. *College Composition and Communication*, 58: 385-418.
- Cassany, D. (2008). Metodología para trabajar con géneros discursivos. *Jornadas sobre lenguajes de especialidad y terminología; Leioa, Biscaia, 1 febrero de 2008*.
- Charaudeau, P. (2006). El contrato de comunicación en una perspectiva lingüística: Normas psicosociales y normas discursivas. *Opción*, 22(49), 38-54.
- Ciapuscio, G. E. (1994). *Tipos textuales*. Universidad de Buenos Aires, Instituto de Lingüística.
- Ciapuscio, G. E., & Kugel, I. (2002). Hacia una tipología del discurso especializado: aspectos teóricos y aplicados. *Entre la terminología, el texto y la traducción* (pp. 37-73).
- Espejo, C. (2006). La movida concluyendo en torno al tema en informes de investigación elaborados por estudiantes universitarios. *Onomázein*, 13, 35-54.
- Flowerdew, J. (Ed.) (2002). *Academic discourse*. Cambridge: Cambridge University Press
- Flowerdew, L. (2004). The argument for using English specialized corpora to understand academic and professional language. En U. Connor & T. Upton (Eds.), *Discourse in the professions. Perspectives from corpus linguistics* (pp. 11-33). Amsterdam: Benjamins.
- Gardner, S., y Nesi, H. (2013). A Classification of Genre Families in University Student Writing. *Applied Linguistics*, 34(1): 25-52.
- Grabe, W. (2002). Narrative and Expository Macro-Genre. En A. Johns (ed.), *Genre in the Classroom. Multiple Perspectives* (pp. 285-288). Hillsdale: Erlbaum.
- Harvey, A. M. (2005). La evaluación en el discurso de informes escritos por estudiantes universitarios chilenos. En M. Pilleux, (ed.) *Los contextos del discurso*. Santiago: Frasis: 215-22.
- Harvey, A. M., & Muñoz, D. (2006). El género informe y sus representaciones en el discurso de los académicos. *Estudios filológicos*, (41), 95-114.
- Hyland, K. (2000). *Disciplinary discourses: Social interactions in academic writing*. London: Longman.
- Hyland, K. (2004). *Disciplinary discourses, Michigan classics ed.: Social interactions in academic writing*. University of Michigan Press.

Ibáñez, Romualdo. (2010). El Texto Disciplinar en la transmisión del conocimiento especializado. *Estudios filológicos*, (46), 59-80. <https://dx.doi.org/10.4067/S0071-17132010000200004>

Jarpa, M. (2015). Escritura en las disciplinas: Géneros académicos evaluativos en un programa de posgrado de Biotecnología. *Leer y escribir en contextos académicos y profesionales: Géneros, corpus y métodos*, 221-255.

Lillo, F. (2018). *Relación entre calidad de escritura y rasgos lingüístico-discursivos en las introducciones de trabajos finales de grado en ingeniería civil informática*. Tesis de Magister. Valparaíso, Chile: Pontificia Universidad Católica de Valparaíso

Linell, P. (1998). Discourse across boundaries: On recontextualisation and the blending of voices in professional discourse. *Text*, 18(2), 143-157.

López, C. (2002). Aproximación al análisis de los discursos profesionales. *Revista Signos*, 35 (51-52), 195-215.

Manghi, D. (2009). *Co-utilización de recursos semióticos para la regulación del conocimiento disciplinar. Multimodalidad e intersemiosis en el Discurso Pedagógico de Matemática en 1° año de Enseñanza Media*. Tesis Doctoral. Chile: Pontificia Universidad Católica de Valparaíso.

Martin, J. (2002). A universe of meaning. How many practices?. En A. Johns (ed.), *Genre in the Classroom. Multiple Perspectives* (pp. 269-278). Hillsdale: Erlbaum.

Martínez, J. D. (2012). Descripción y variación retórico-funcional del género Tesis Doctoral: Un análisis desde dos disciplinas y dos comunidades discursivas a partir del Corpus Te DiCE-2010. *Unpublished PhD thesis, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile*.

Matthiessen, C., Teruya, K., & Lam, M. (2010). *Key terms in systemic functional linguistics*. A&C Black.

Muñoz, D. (2004). Estructura y patrones léxicos en informes escritos de estudiantes universitarios. Ponencia presentada en 3er Encuentro Nacional de Estudios del Discurso. Universidad Austral de Chile.

Navarro, F. (2014). Manual de escritura para carreras de humanidades. CABA: Editorial de la Facultad de Filosofía y Letras UBA.

Nazar, R. (2016). Distributional analysis applied to terminology extraction. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, 22(2), 141-170.

Núñez, P., & Espejo, C. (2005). Estudio exploratorio acerca de la conceptualización del informe escrito en el ámbito académico. *Ana María Harvey (Comp.). En torno al Discurso. Contribuciones de América Latina. Santiago de Chile: Ediciones Pontificia Universidad Católica*, 135-148.

Oyanedel, F. (2005). Lo descriptivo en informes escritos de estudiantes universitarios. *Onomázein*, 1(11).

Parodi, G. (2007). El discurso especializado escrito en el ámbito universitario y profesional: Constitución de un corpus de estudio. *Revista signos*, 40(63), 147-178. <https://dx.doi.org/10.4067/S0718-09342007000100008>

Parodi, G. (2008). “Géneros del discurso escrito: Hacia una concepción integral desde una perspectiva sociocognitiva”. En G. Parodi (ed.), *Géneros Académicos y Géneros Profesionales: Accesos Discursivos para Saber y Hacer* (pp. 17-37). Valparaíso: Ediciones Universitarias de Valparaíso.

Parodi, G. (2008). *Géneros académicos y profesionales: acceso discursivo para saber y hacer*. Valparaíso, Chile: Ediciones universitarias.

Parodi, G. (2010). The rhetorical organization of the textbook genre across disciplines: ‘a colony in loops’? *Discourse Studies*, 12(2), 195-222.

Parodi, G. (2014). Genre organisation in specialised discourse: Disciplinary variation across university textbooks. *Discourse Studies*, 16(1), 65-87.

Parodi, G. (Ed.). (2005). *Discurso especializado e instituciones formadoras*. Ediciones universitarias de Valparaíso.

Parodi, G. (2015). Géneros del discurso escrito: Hacia una concepción integral desde una perspectiva cognitiva. En G. Parodi (Ed.), *Géneros académicos y géneros profesionales: Accesos discursivos para saber y hacer* (pp. 21-41). Valparaíso: Ediciones Universitarias de Valparaíso.

Parodi, G., Ibáñez, R., y Venegas, R. (2009). El Corpus PUCV-2006 del español: identificación y definición de los géneros discursivos académicos y profesionales. *Literatura y lingüística*, (20), 75-101.

Parodi, G., y Vásquez-Rocca, L. (2015). Relaciones retóricas y multimodalidad en el género Informe de Política Monetaria del discurso académico de la Economía. *Caleidoscopio*, 13(3), 388-405.

Parodi, G., Venegas, R., Ibáñez, R., y Gutiérrez, R. M. (2008). Géneros del discurso en el Corpus PUCV-2006: Criterios, definiciones y ejemplos. In *Géneros académicos y géneros*

profesionales: Accesos discursivos para saber y hacer (pp. 39-74). Pontificia Universidad Católica de Valparaíso.

Perelman, Ch. y Olbrechts Tyteca, (2009). *Tratado de la argumentación: La nueva retórica*. Madrid: Gredos.

Quintiliano, M. (1942). *Instituciones oratorias*. Madrid: Hernando.

Reuter, Y. (2000). *La description: des théories à l'enseignement-apprentissage*. Ciudad: ESF éditeur.

Sabaj Meruane, O. (2012). Uso de movidas retóricas y patrones léxico-gramaticales en artículos de investigación en español: Implicancias para la enseñanza de la escritura científica. *Boletín de filología*, 47(1), 165-186.

Sabaj, O., & González, C. (2013). Seis propósitos comunicativos no discurso de editores de revistas científicas. *DELTA: Documentação e Estudos em Linguística Teórica e Aplicada*, 29(1).

Samraj, B. (2008). A discourse analysis of master's theses across disciplines with a focus on introductions. *Journal of English for academic purposes*, 7(1), 55-67.

Silver, M. (2006). *Language across disciplines: Towards a critical reading of contemporary academic discourse*. Ciudad: Universal-Publishers.

Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.

Swales, J. (2004). *Research genres: Explorations and applications*. Ernst Klett Sprachen.

Venegas, R. (2010). *Caracterización del macro-género trabajo final de grado en licenciatura y magíster: Desde los patrones léxico-gramaticales y retórico-estructurales andamiaje de la escritura académica*. Informe Proyecto FONDECYT 1101039.

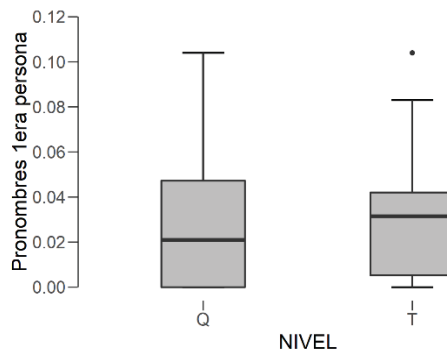
Venegas, R., Zamora, S., y Galdames, A. (2016). Hacia un modelo retórico-discursivo del macrogénero Trabajo Final de Grado en Licenciatura. *Revista signos*, 49 (Supl. 1), 247-279. <https://dx.doi.org/10.4067/S0718-09342016000400012>

Zamora, S., & Venegas, R. (2013). Estructura y propósitos comunicativos en tesis de magíster y licenciatura. *Literatura y lingüística*, (27), 201-218.

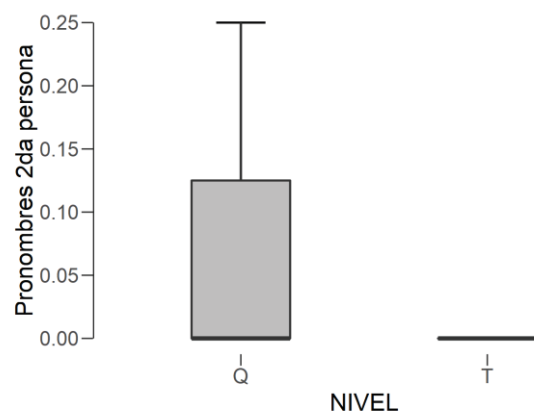
7. ANEXOS

7.1 Boxplots (gráfico de caja y bigotes) SINLP

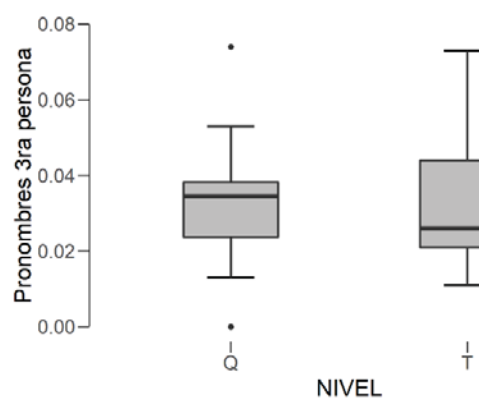
Pronombres 1era persona



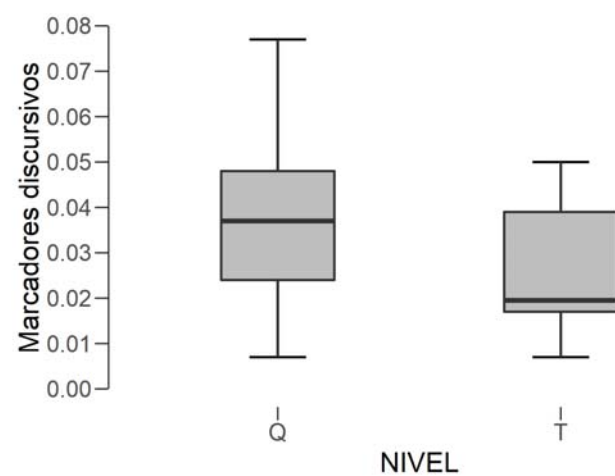
Pronombres 2da persona



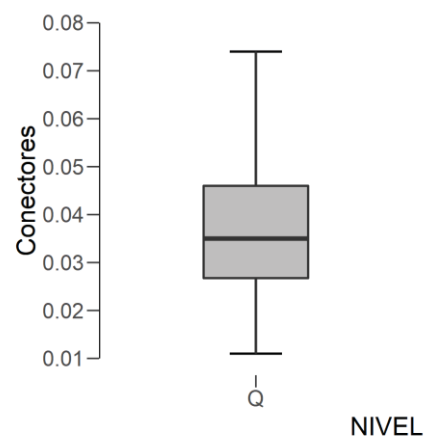
Pronombres 3ra persona



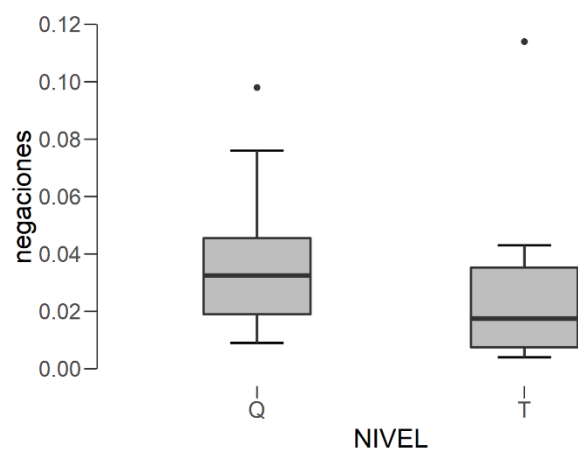
Marcadores discursivos



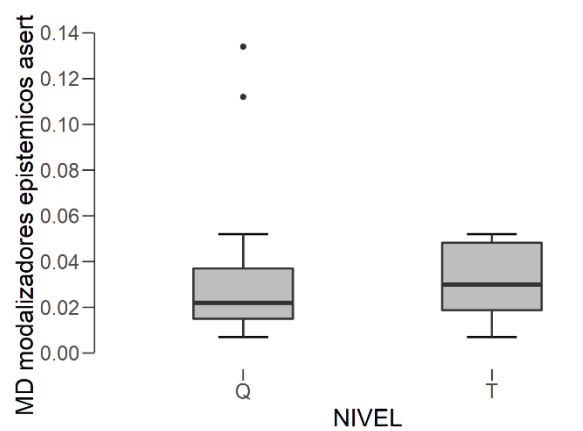
Conectores



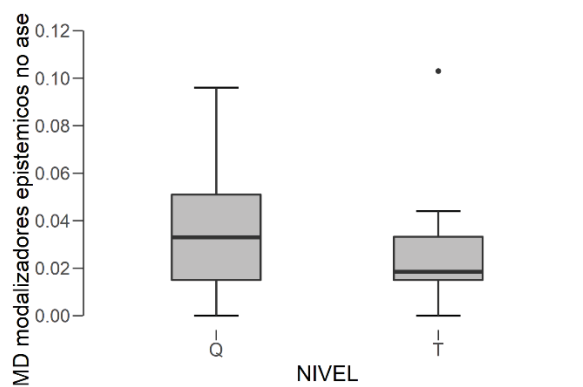
Negaciones



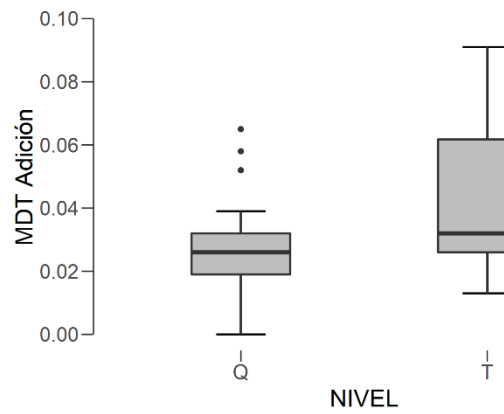
MD modalizadores epistemicos asertivos



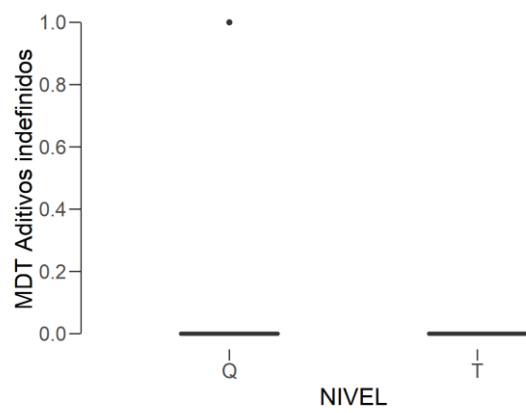
MD modalizadores epistemicos no asertivos



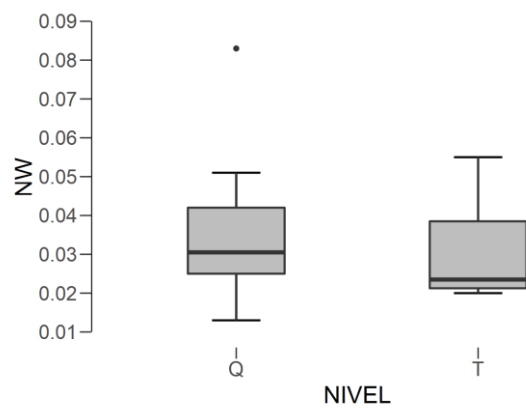
MDT Adición



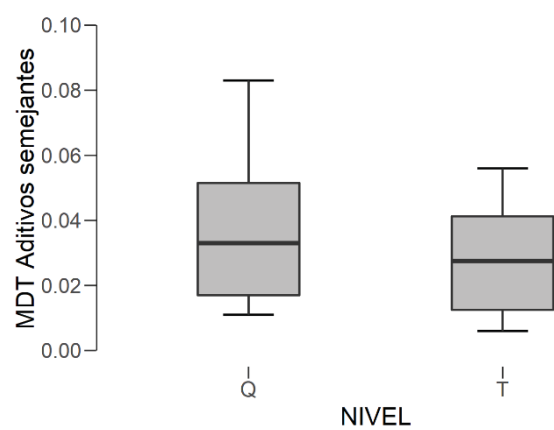
MDT Aditivos indefinidos



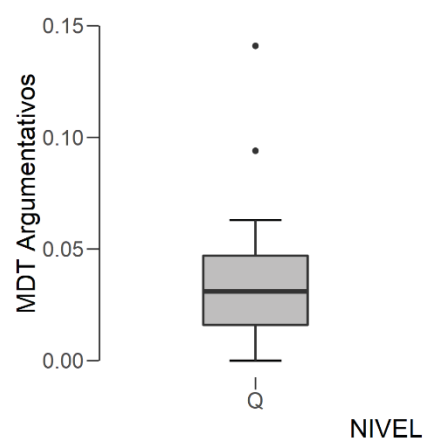
Número de palabras



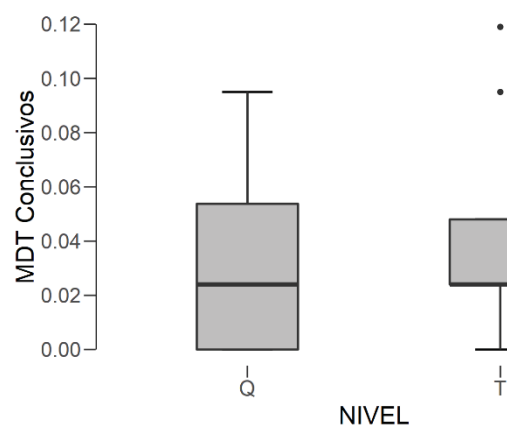
MDT Aditivos semejantes



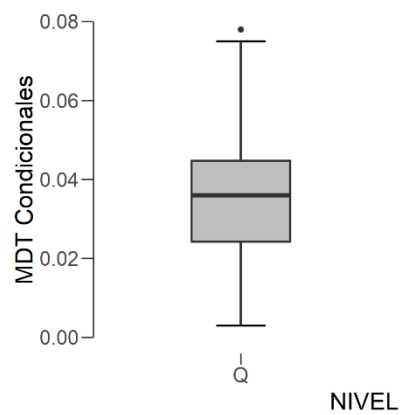
MDT Argumentativos



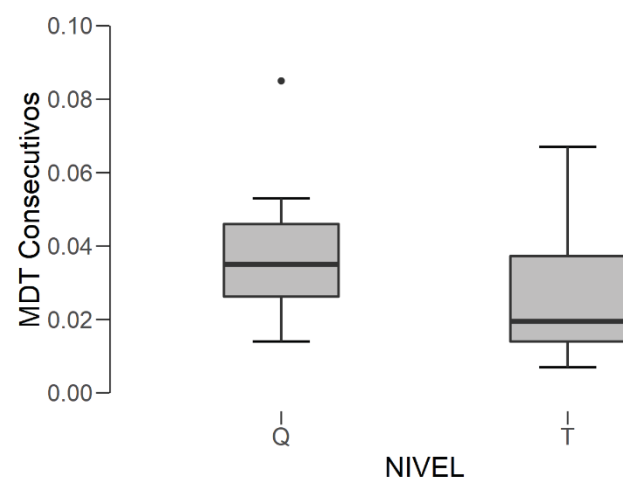
MDT Conclusivos



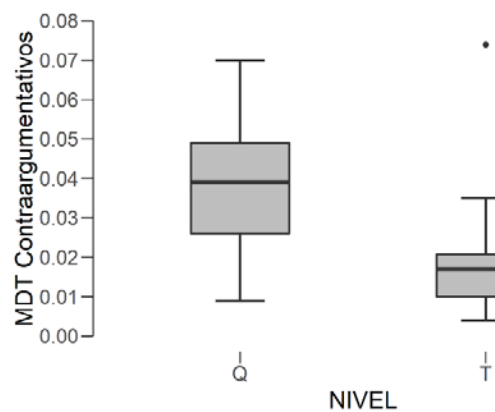
MDT Condicionales



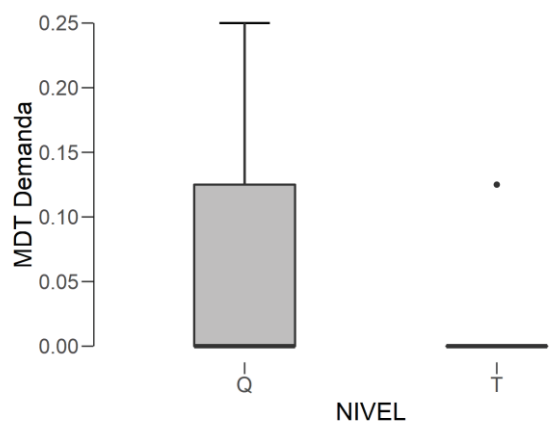
MDT Consecutivos



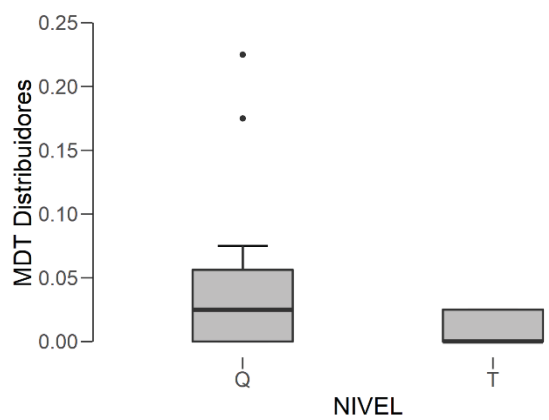
MDT Contraargumentativos



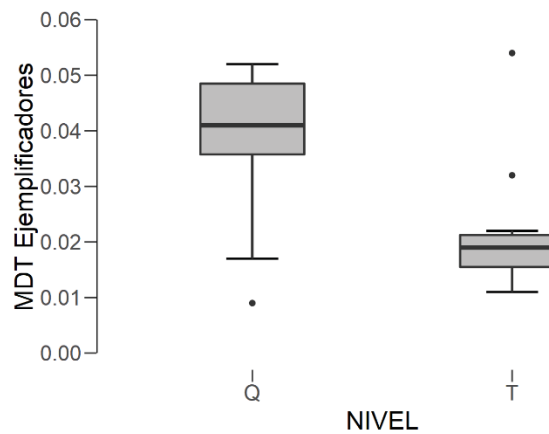
MDT Demanda



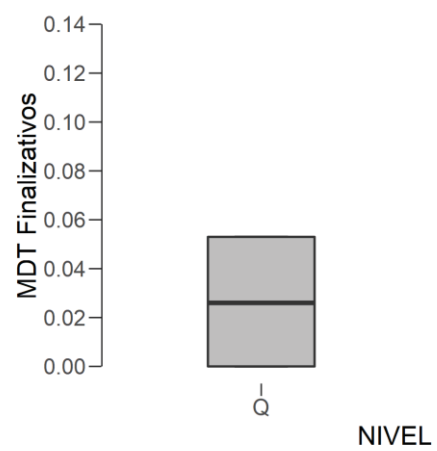
MDT Distribuidores



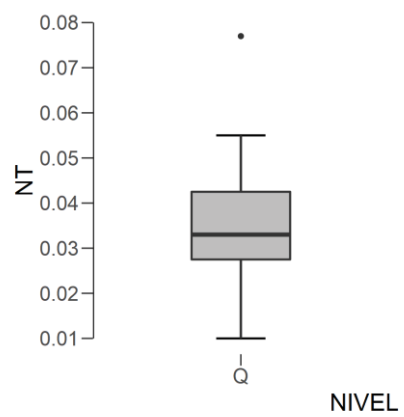
MDT Ejemplificadores



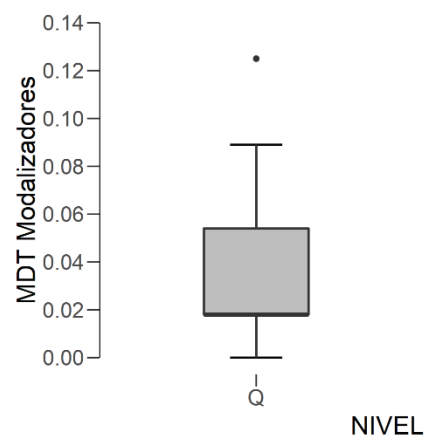
MDT Finalizativos



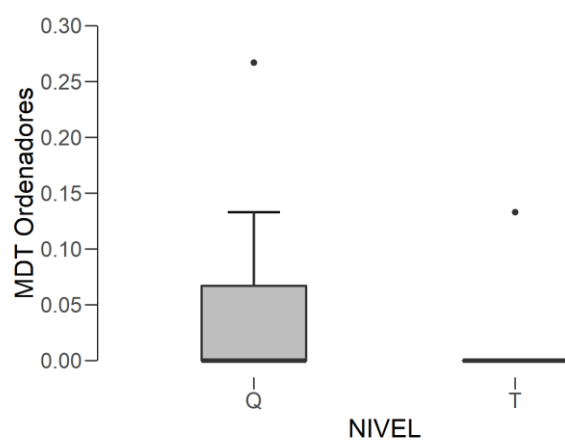
NT



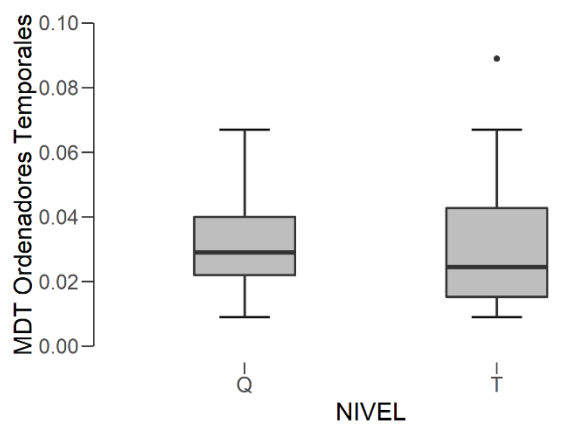
MDT Modalizadores



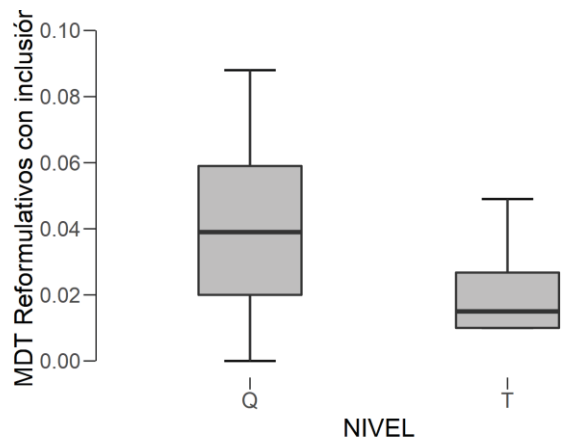
MDT Ordenadores



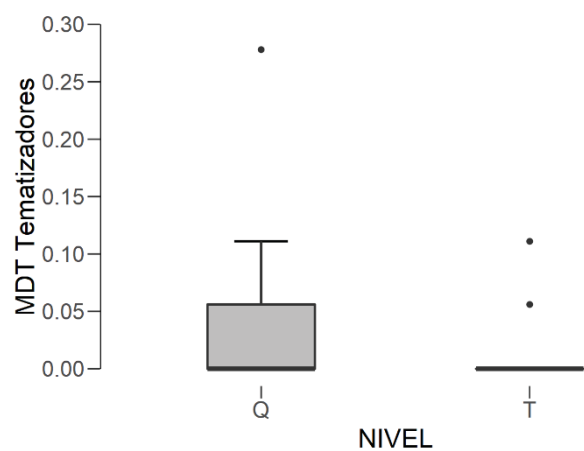
MDT Ordenadores Temporales



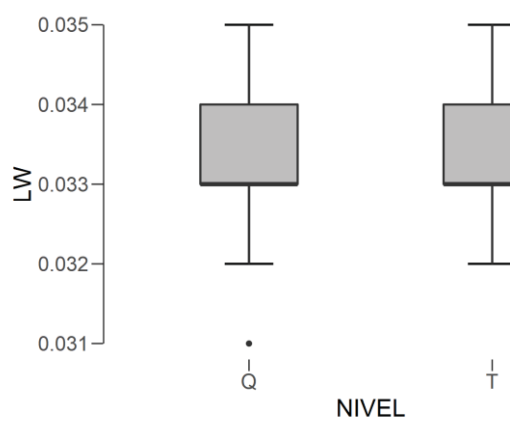
MDT Reformulativos con inclusión



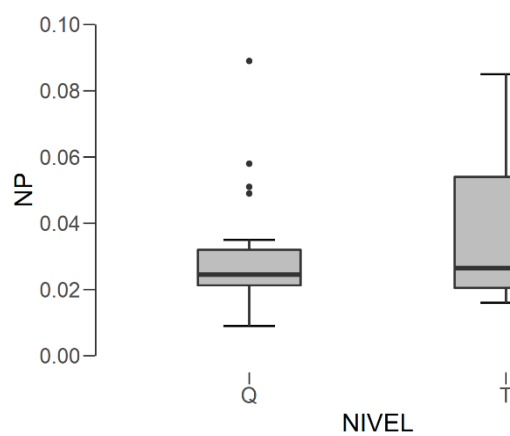
MDT Tematizadores



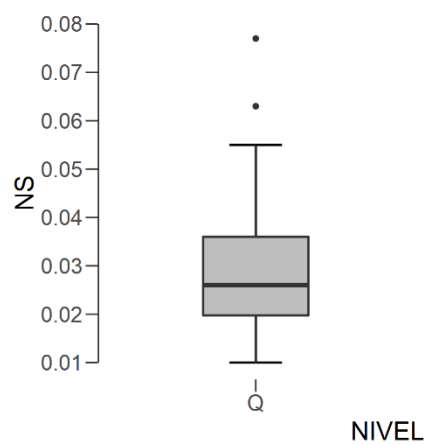
Letras por palabra



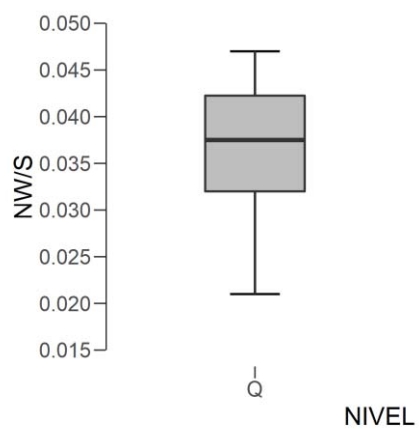
Número de párrafos



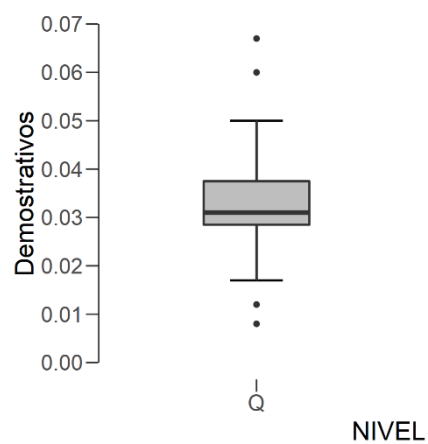
Número de oraciones



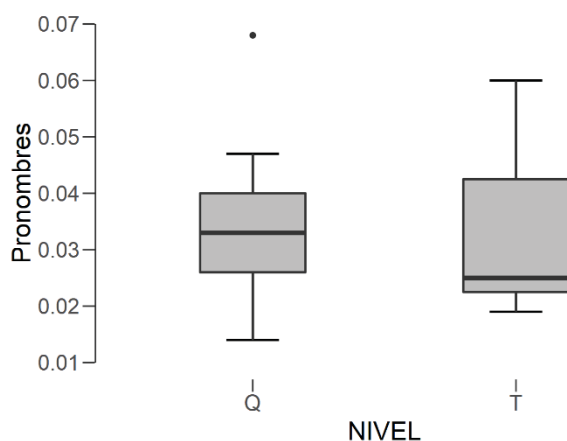
Número de palabras por oración



Demostrativos



Pronombres



7.2 Ejemplos rasgos discursivos de quinto año

Rasgo discursivo	Ejemplos del corpus de quinto año
Pronombres 2da persona	Finalmente, si no estas a gusto con ningunos de los servicios de almacenamientos de los que se habló en este documento, siempre está la posibilidad de crear una nube tú (...)
MDT Aditivos indefinidos	Durante el documento se detallarán aspectos como la descripción general de que es docker, su arquitectura, entre otros aspectos [Aditivos Indefinidos] que el equipo ha considerado relevantes para incluirlas dentro de este trabajo.
MDT Contraargumentativos	Sin embargo [Contraargumentativos]este se centra en la transformación de datos no estructurados en la web como HTML, en datos estructurados para que puedan ser almacenados y analizados.
MDT Ejemplificadores	El hecho de que el proxy represente a más de un usuario presenta muchos problemas, en concreto [Ejemplificadores] los que presuponen una comunicación directa entre 1 emisor y 1 receptor (como TCP/IP).
MDT Reformulativos con inclusión	Se pueden encontrar servidores proxy de forma gratuita fácilmente, aunque, de todas maneras [reformulativos con inclusión], tienen sus opciones de pago.

7.3 Ejemplos vocabulario especializado

Términos de quinto y tercero	Ejemplos del corpus
proyecto	Los trabajadores son de lo más importante en el proyecto , sin ellos no se podría realizar por lo cual en el software claramente deben tener una gran importancia
Método	En esta sección se analizará las clases, explicando sus métodos y funcionalidad.
Usuario	A partir de datos creados de un txt, se genera un usuario al azar.
dato	Recibe un string que actúa como nombre de la Comuna y un tipo dato Usuario.
aplicación	Nuestra aplicación busca aprovechar al máximo la programación orientada a objetos, usando uso de las distintas clases y sus métodos
información	Además, por el momento no contamos con una base de datos, por cuánto la mayoría de nuestra información se encuentra en archivos txt.
archivo	Estas clases son las que le dan vida al programa desde el inicio y permiten cargar los datos desde archivos .dat
base	Un jardín infantil es la base en la educación de todas las personas.
plataforma	Los miembros del grupo se acoplaron bien a las actividades planificadas y ofrecieron muchas horas de su tiempo para trabajar presencialmente y no a través de una plataforma online.

Java	Adquirimos mucha experiencia programando en Java , ahora tenemos una muy buena base para la continuación de este programa
interfaz	Esta aplicación además de incorporar esta interfaz para el usuario incorpora una propia para los administradores que les da las facultades necesarias para administrar de manera correcta la aplicación.
software	En la presente sección se hablará sobre el diseño planteado para WeebReader, a partir de una mirada de la arquitectura de software .
empresa	Esto se ve reflejado desde grupos de trabajos hasta en grandes empresas .
objeto	El paradigma orientado a objetos surge para solucionar los problemas que planteaban otros paradigmas, como el imperativo, con el objeto de elaborar programas y módulos más fáciles de escribir.
diseño	Para lograr un mayor entendimiento del diseño general de la aplicación se hará uso de extractos de códigos a la vez de modelos y diagramas explicativos.
función	Se usa la misma función para crear desayuno, colaciones, almuerzo y cena.
solución	Posteriormente se explicará la solución que encontramos para estas falencias, presentando un formato que no solo será fácil de utilizar para los usuarios, sino que proveerá de la información necesaria y el material visual adecuado para captar el interés de los lectores

gestión	Este documento describe el presente proyecto "Banana Task Enterprise como un software de gestión , en donde se presenta la descripción del problema que el sistema pretende solucionar
sistema	Hoy en día las grandes empresas cuentan con un sistema de gestión que las ayuda a organizar sus proyectos, tareas y los eventos que cada una de estas tiene
cliente	El programa agrega un nuevo proyecto al cual se le asocia un nombre de proyecto, un cliente , un estado

Término INF_TERCER	Ejemplos del corpus
Fase	Agregar una fase a un proyecto: Se puede entender una fase como "subtareas" de un proyecto, las cuales tienen un nombre, una descripción, un empleado encargado y un estado que define si esta hecha o no
String	Atributos private String nombreP. Private String rutP. Private String passP. Private static ListaMaterial material = new ListaMaterial().
Programa	Cuando se inicia el programa el usuario debe elegir entre iniciar sesión o registrarse.
Alumno	En este proyecto se ha propuesto una solución para la distribución del material educativo, con el fin de que los profesores puedan entregar dicho material a sus respectivos alumnos de una manera más rápida y sencilla
Dieta	En internet, podemos encontrar muchas aplicaciones para administrar los alimentos, pero no existen muchas que tomen en

	cuenta los tipos de dieta y la actividad física que puede tener una persona
Gantt	A continuación, se presentarán 3 Cartas Gantt donde se especifica el actuar del equipo de trabajo en general, así como individual desde el comienzo de la elaboración de esta aplicación
carta Gantt	--
Diagrama	Diagrama de clases Plataforma: Se presenta el diagrama de clases de plataforma y sus colecciones.
Jardín	Un jardín infantil es la base en la educación de todas las personas.
Alimento	Le dará a cada usuario las herramientas para crear sus propias dietas, una base de datos de alimentos con sus correspondientes calorías por porción
Contraseña	El método comprobarLogin verifica si el nombre de usuario y contraseña ingresado por el usuario existen en la base de datos
Identificador	Cada vez que entra alguien nuevo a la empresa, se agrega un empleado con todos sus datos: nombre, ocupación o profesión y un identificador .
Correo	En ella se almacenan todos los atributos requeridos para lograr el envío de un correo electrónico.
menú	Una vez que se inicia correctamente la sesión se muestra el menú principal
Ejercicio	Nuestra aplicación de vida saludable es una aplicación de vida saludable que promueve facilidades para comer sano y hacer

	ejercicio
Tutor	Deberá tener la habilidad de buscar un tutor a través de un match y confirmar la postulación si este es de su agrado.
Boolean	public boolean addManga(String id, Manga m) public boolean exists(String id) public boolean exists(Manga m)
Actividad	MySQLConnect es la clase utilizada para cualquier actividad que requiera una base de datos.
administrador	Esta clase posee una referencia a la clase Administrador y a la clase VentanaInicio.

Término INF_QUINTO	Ejemplos del corpus
Red	Para estos fines, existen distintos tipos de arquitecturas de red , desde la más conocida y utilizada actualmente la arquitectura Cliente-Servidor o algunas menos manipuladas como la arquitectura Cliente-Cola-Cliente.
Servidor	Este modelo contrasta con el modelo cliente- servidor , que es el modelo más común de interacción de red que se emplea en las aplicaciones de internet
Seguridad	Los sistemas P2P comparten la mayoría de los requisitos de seguridad de los sistemas distribuidos tradicionales.
Ventaja	Pero, como todo sistema posee sus ventajas y desventajas, y su utilidad dependerá del tipo de uso que se le dé.
Máquina	Existe una capa intermedia llamada hipervisor, también llamado monitor de máquina virtual,

Negocio	En la cual se detalla aspectos como la descripción de su funcionamiento considerando desde su modelo de negocio hasta su arquitectura.
Framework	Esto se debe a que ambos son entendidos por el Framework .NET.
Comunicación	Estas bibliotecas dan forma a módulos de trabajo mas grande como los son Asp.net para el trabajo de servidores, Ado.net para la comunicación con la capa de datos
Desventaja	A continuación se muestra una tabla con las ventajas y desventajas de utilizar el framework AngularJS.
Docker	Se presentará la información recabada sobre la herramienta Docker para la asignatura de formulación de proyectos informáticos.
Nivel	A nivel empresarial existen varios usos para los cuales se suele utilizar Docker, como la gestión de proyectos, en la cual se busca evitar incompatibilidades de entorno
Google	Es en este punto donde entra en escena el, por aquel entonces, recién aparecido intérprete de JavaScript de Google
contenedor	El contenedor de Docker aloja todo lo necesario para ejecutar una aplicación
sistema operativo	Mucho tiempo se especuló sobre qué sistema operativo funcionaba el software de Tesla.
vehículo	Por su parte, de manera similar a Toyota, utiliza un completo sistema de control del vehículo integrado por sensores
almacenamiento	El coste del sistema global se ve limitado, además, porque en P2P se reducen las capacidades de cálculo, almacenamiento y

	ancho de banda ociosas.
máquina virtual	Existe una capa intermedia llamada hipervisor, también llamado monitor de máquina virtual , que administra los recursos coordinando y validando las solicitudes de recursos desde las máquinas virtuales a la CPU
motor	Dichas instancias son nombradas contenedores o motores de virtualización
proxy	Los componentes que se expondrán son Firewall, Proxy , Switch y Router.

7.4 Ejemplos voz pasiva

Voz Pasiva	Ejemplos del corpus
"se encuentran",	Dichas explicaciones se encuentran en los puntos 1.1.1 y 1.1.2 respectivamente.
"se encuentra",	Menor rendimiento Dado que la máquina virtual se encuentra en una capa intermedia, no hace uso del hardware real de forma directa, obteniendo un rendimiento inferior.
"se utiliza",	A grandes rasgos Docker se utiliza para desarrollar e implementar aplicaciones que requieran estar en diferentes entornos, en los que se necesita integridad en la aplicación tanto para entornos y máquinas diferentes.
"se encarga",	Redis se encarga de actualizar en el menor tiempo posible los timelines de todos sus usuarios.
"se hace",	Cuando muchos usuarios se identifican como uno solo, se hace difícil que el recurso pueda diferenciarlo y esto puede ser un inconveniente cuando necesariamente hay que hacer una identificación
"se necesita",	Para programar se necesita del ambiente de desarrollo Visual Studio 2010 Service Pack 1
"se utilizan",	La principal funcionalidad de este modelo es diferencia y mantener registro de las distintas

	generación de sistemas distribuidos, para esto se utilizan estas 4 características
"se ve",	Este modelo se ve reflejado en la siguiente figura.
"se busca",	En cuanto a la comunicación de datos, uno de los aspectos relevantes es la interconexión de redes, donde se busca entregar o recibir información de otra entidad.
"se trabaja",	El análisis del sistema en el cual se trabaja debe tomarse en serio para invertir recursos en Hardware o software que no necesitaran.
"se tiene",	Lo complejo de estos sistemas es que se tiene que tener una documentación, ya sea extensa o simple, debido a que se necesitan entender la estructura interna del sistema para poder realizar las llamadas a los procedimientos.
"se presentarán",	A continuación se presentarán aspectos positivos y negativos al usar el almacenamiento en la nube, Google Drive
"se procederá",	Una vez que se tenga una buena modelización del negocio, se procederá a establecer el modelo de datos que mejor se acomode a las necesidades del proyecto.
"se muestra",	Cada vez que un Action termina su ejecución, se muestra un resultado al usuario.
"se logra",	Esto se logra mediante un sistema de permisos de acceso y distintos tipos o niveles de autoridad.
"se espera",	Básicamente, lo que se espera de este software, es que sirviera como ayuda para planificar; que materiales se iban a necesitar durante el proceso de producción y como se gestionaría su adquisición.
"se quiere",	En este sentido si se quiere un informe especial o distinto de los que lleva el producto, se tendría que solicitar a la empresa implantadora que añada la funcionalidad requerida.
"se presenta",	De la misma manera, a continuación se presenta una tabla con las ventajas y desventajas de utilizar el framework Angular.
"se almacenan",	Todos los documentos se almacenan en la nube, evitando de esta forma riesgos de fallos de disco duro y ahorrando la necesidad de

	hacer respaldos.
"se comunican",	Estas se comunican directamente con el sistema operativo, sin ningún programa intermediario.

Voz pasiva INF_TERCER	Ejemplos del corpus
"se inicia",	Ventana donde se inicia el programa y se pasan los datos a las ventanas
"se especifica",	A continuación, se presentarán 3 Cartas Gantt donde se especifica el actuar del equipo de trabajo en general, así como individual desde el comienzo de la elaboración de esta aplicación.
"es utilizado",	Método privado que es utilizado por getMangaInfo para conseguir el Json ubicado en la URL del manga.
"ha elaborado",	Antes de definir las clases y las colecciones necesarias, se ha elaborado el modelo de base de datos del sistema en la figura
"se agrega",	Luego de rellenar los datos, se agrega a un HashMap que contiene todos los proyectos asociados a la constructora.
"se decide",	Se decide trabajar con una base de datos relacional para tener un manejo de los datos más rápida y eficiente.
"se editan",	Esta clase es una lista que contiene todos los alumnos del sistema, se editan sus atributos nombre y observaciones, además de las funciones agregar, eliminar, verificar si existe y buscar. Esta clase también implementa a la clase interfaz Lista.
"se elimina",	Retorna true si se elimina exitosamente, false en caso de que no exista el proyecto.
"es sabido",	Como bien es sabido una constructora que se enfoque en realizar una variedad de proyectos está compuesta por una gran cantidad de trabajadores a sus espaldas
"ha dicho",	Análisis ElecMed es una aplicación que, como se ha dicho anteriormente, se enfoca en tomar los datos y procesarlos.
"ha propuesto",	En el presente trabajo, se ha propuesto una aplicación funcional, que de forma eficiente agrupa la información de los usuarios y la

	utilice de manera correcta.
"se acceda",	Recibe un nombre de fase y proyecto, de forma que se acceda a este último directamente y a la fase indicada para poder eliminarla.
"se asigna",	Métodos setNombre se asigna un nombre.
"se asocia",	A su vez tanto la tabla dieta como la tabla dietaRecomendada mantienen una relación uno a muchos (1:n) con la tabla comidas, dando a entender que una dieta (tanto propia como recomendada) se asocia a varias comidas, pero cada comida solo se asocia a una dieta.
"se encargará",	Este MangaMap se encargará de manejar una colección de mangas favoritos del usuario.
"se guardarán",	Los mangas que el usuario posea se guardarán en la base de datos, en la tabla favoritemangas.
"se hablará",	En la presente sección se hablará sobre el diseño planteado para WeebReader, a partir de una mirada de la arquitectura de software.
"se hizo",	Ventana que contiene todos los materiales previamente ingresados por el usuario, estos se mostraran en forma de lista, señalando la fecha de ingreso, el nombre de material, una breve descripción sobre estos, la cantidad que fue añadida al proyecto, y en caso de haber retirado parte de este se señala la fecha en que se hizo y la cantidad.
"se implementan",	También se implementan otros requerimientos del programa como la Conexión a la Base de Datos con sus respectivas Consultas.

Voz Pasiva INF_QUINTO	Ejemplos del corpus
"se requiere",	Pero en cambio, si se requiere utilizar otros protocolos, se requieren procedimientos muy complejos, es más recomendable utilizar SOA.

"se recomienda",	SQL Server se recomienda para sistemas que trabajen en windows.
"se basa",	La plataforma Java tiene la particularidad de que sólo se basa en software, el que se ejecuta en otras plataformas basadas en hardware.
"se ejecuta",	HANDLER es la función que se ejecuta cuando se correlaciona la ruta.
"se mantienen",	\$scope, y la vista se enlaza a esas propiedades, en todo momento la vista y el controlador se mantienen sincronizados.
"se realiza",	En un ataque DDos, el ataque se realiza desde múltiples pc, con lo que el administrador no podría saber de dónde viene el ataque por lo cual es mucho más complicado detener esta denegación de servicio distribuida.
"se trata",	El outsourcing funciona cuando se trata de un negocio de servicios donde la informática es fundamental para operar una aplicación, y así poder apoyar al cliente; un negocio de alta tecnología de bajo contacto.
"se realizan",	En un ataque Dos, se realizan muchas peticiones de un mismo pc, por lo cual el administrador del servicio atacado puede ver el lugar de procedencia de estos ataques, banear la ip y eliminar el problema.
"se refiere",	SIP (Session Initiation Protocol) SIP es un protocolo que proporciona servicios de telefonía similares a H.323, pero es menos complejo y requiere menos recursos. QoS que se refiere a la calidad de servicio
"han desarrollado",	Para satisfacer las necesidades de cada tipo de cliente, se han desarrollado sistemas que ofrecen varias maneras de guardar, organizar y manipular datos, por medio de diferentes estructuras lógicas.
"se aplica",	En caso de que no se realicen, se aplica un rollback, con el cual la transacción es cancelada por completo.
"se conoce",	Este script crea una copia de seguridad de una base de datos de SQL Server, creando una copia de uno o más archivos (o grupos de archivos) de la base de datos. A esto se conoce como Backup database.

"se crean",	Por lo general, un ESB implementa un modelo de mensajes para cumplir con la tarea de gestionar toda la comunicación, se crean mensajes normalizados que el ESB recibe y transmite, realizando además los cambios de formato necesarios para que las aplicaciones de destino entiendan el mensaje.
"se diferencia",	Una VPN se diferencia de un proxy, sobre todo, en que estas se encargan de cifrar el tráfico que pasa a través de la red, cosa que no hace el proxy.
"se dividen",	No hay separación entre especificación e implementación. Las clases en Java no se dividen en estas dos partes.
"se integran",	Es el punto donde las aplicaciones J2EE se integran con otros sistemas no J2EE o con sistemas de legado.
"se presentan",	A continuación se presentan algunos de estos frameworks:
"se tenga",	Una vez que se tenga una buena modelización del negocio, se procederá a establecer el modelo de datos que mejor se acomode a las necesidades del proyecto.
"se ubica",	Se ubica en el computador del cliente, es la interfaz gráfica del sistema e interactúa con el usuario.
"ha logrado",	Cada una de las mejoras evolutivas realizadas ha logrado mayor rapidez.

7.5 Ejemplos trigramas de lema

TRI-LEMMA 3RO	Ejemplos del corpus
base de dato	Antes de definir las clases y las colecciones necesarias, se ha elaborado el modelo de base de datos del sistema
de lo aplicación	En ella se almacenan las funciones principales de la aplicación tales como búsqueda, agregar, eliminar o editar un dato
de uno proyectar	Edita todos los datos de un proyecto : nombre, cliente e integrantes.

	Puede agregar o quitar estos últimos.
el nombrar de	Editar empleado: Recibe el nombre de un integrante del mapa para luego editarlo.
recibir el nombrar	Eliminar encargado de fase: Recibe el nombre del proyecto. Elimina a todos los encargados del proyecto.
la base de	ListaJardines donde se encuentra una lista con todos los jardines en la base de datos
el nombrar del	MangaMap tiene una instancia de TreeMap ordenado por la ID del manga, la cual contiene el nombre del manga, es decir, se utiliza un TreeMap para tener guardados los mangas ordenados por nombre.
de lo clase	La razón de seguir este patrón es para conservar una instancia de la clase Reader, teniendo acceso a sus datos siempre que se desee, evitando conflictos en los datos.
en casar de	Editar un proyecto: Un proyecto se puede editar en caso de que sea necesario modificar el nombre, el cliente o su estado
a través de	Java es un lenguaje de programación orientado a objetos, fue creado principalmente para que los desarrolladores de aplicaciones escriban y programen una vez y para que dicho programa pueda ser ejecutado a través de cualquier dispositivo.
Recibir uno id	mostrarValorizacion este método recibe un id y retorna una valorización.
Nombrar de uno	Finalizar fase: Recibe el nombre de un proyecto y de una fase.
Diagrama de clase	Se presenta el diagrama de clases de plataforma y sus colecciones.
La cantidad de	Esto se debe a la cantidad de factores que afectan e interaccionan entre sí a la hora de su funcionamiento

Proyectar recibir el	Proyecto recibe el nombre de la fase, la descripción, el estado en el que se encuentra y al encargado de la fase
Retornar true si	Retorna true si agrega a todos los empleados, false si no agrega a ninguno.
Por el que	¿Cuál es la razón por la que las organizaciones gastan importantes cantidades de recursos diseñando planes de nuevas orientaciones al cliente y nuevos posicionamientos estratégicos, pero finalmente al momento de ejecutar dichos planes la obtención de los recursos necesarios no está disponible?
De proyectar recibir	Fase de proyecto recibe un nombre de fase y proyecto, de forma que se acceda a este último directamente y a la fase indicada para poder eliminarla.
De lo fase	Recibe el nombre de un proyecto para poder acceder a él, el nombre de la fase a editar y todos los atributos nuevos.

TRI-LEMMA 5TO	Ejemplo de corpus
base de datos	El historial (Logging) mantiene el registro de cambios en la base de datos .
a través de	Si el sistema se cae, la base de datos puede ser restaurada a través del log, siendo utilizado como un punto de salvado.
por el que	Una vez que inicia su trayectoria, el vehículo compara el mapa con lo que le indican los sensores LIDAR para asegurar la ruta por la que se desplaza.
ventaja y desventaja	Ventajas y desventajas de Proxy
la base de	Si la base de datos es muy grande, se podría comprimir la salida del mysqldump.
hoy en día	La tabla a continuación corresponde a la clasificación realizada por DB-engines de los 20 motores de base de datos más populares hoy en día .
de lo red	Los HUD trabajan en la capa 1 de la red sin manejo alguno del tráfico interno usados hasta inicios del 2000 como sistemas de análisis de paquetes de red a través de copias.

deber a que	Esto se debe a que ambos son entendidos por el Framework .NET.
Este tipo de	Debido a lo precedente, este tipo de virtualización difiere de los otros modelos en que solo funciona con un solo sistema operativo
A continuación se	A continuación se presentarán algunas alternativas y se explicará brevemente cada una de ellas.
Lenguaje de programación	Una de las primeras cosas a las que nos enfrentamos después de tener una idea o proyecto a desarrollar en el área móvil, es decir, una App, es qué modelo, tecnología y lenguaje de programación usaremos para el desarrollo.
De lo sistema	Tiene un mejor rendimiento dentro de los sistemas operativos, ya que tiene acceso a todos los recursos del sistema
La mayoría de	Dado que la mayoría de Apps híbridas pueden ser visualizadas de forma web, en muchos casos, permite que usuarios sin o con poco espacio de almacenamiento en su dispositivo puedan hacer uso de la misma sin realizar una descarga e instalación.
De base de	10 M accesos de base de datos
El que se	Los vehículos que presentan esta competencia poseen la inteligencia necesaria para comprender el entorno en el que se encuentran inmersos
La posibilidad de	Al trabajar aplicaciones en contenedores con Docker existe la posibilidad de requerir del uso de varios contenedores
Ancho de banda	1 mes 8 USD: 81 servidores en 40 países, Ilimitado ancho de banda , todas las herramientas disponibles.
Lo sistema distribuir	El advenimiento de los sistemas distribuidos ha estado soportado en dos importantes innovaciones tecnológicas.

7.6 Ejemplos trigramas de POS

Trigramas de POS	Ejemplos del corpus
(ADP__AdpType=Prep) (DET__Definite=Def Gender=Fem Number=Sing PronType=Art) (NOUN__Gender=Fem Number=Sing)	Este modelo permite realizar cambios al esquema de la base de datos sin detener su funcionamiento.
(DET__Definite=Def Gender=Fem Number=Sing PronType=Art) (NOUN__Gender=Fem Number=Sing) (ADP__AdpType=Prep)	Cada marca pone a su sistema un precio distinto, el cual puede variar dependiendo del tamaño de la base de datos

(DET__Definite=Def Gender=Masc Number=Sing PronType=Art) (NOUN__Gender=Masc Number=Sing) (ADP__AdpType=Prep)	Basándose en el modelo de componentes dinámico y completo para Java OSGi, las aplicaciones y componentes de este servidor pueden ser remotamente instalados
(ADP__AdpType=Prep) (DET__Definite=Def Gender=Masc Number=Plur PronType=Art) (NOUN__Gender=Masc Number=Plur)	El primer paso para las compañías que buscan implementar Business Intelligence es hacer un inventario de los datos que la empresa produce
(ADP__AdpType=Prep) (DET__Definite=Def Gender=Masc Number=Sing PronType=Art) (NOUN__Gender=Masc Number=Sing)	En el modelo cliente-servidor, como su nombre da a entender, posee 2 principales actores (...)
(ADP__AdpType=Prep) (DET__Definite=Ind Gender=Masc Number=Sing PronType=Art) (NOUN__Gender=Masc Number=Sing)	Todo lo mencionado previamente es de orden general para cualquier modelo de interacción de un sistema distribuido
(DET__Definite=Def Gender=Masc Number=Plur PronType=Art) (NOUN__Gender=Masc Number=Plur) (ADP__AdpType=Prep)	Los sistemas de información fueron evolucionando para dar soporte a todo tipo de empresas en sus necesidades de gestión cotidianas.

Trigrama de POS INF_TERCER	Ejemplos del corpus
(DET__Definite=Ind Gender=Masc Number=Sing PronType=Art) (NOUN__Gender=Masc Number=Sing) (ADP__AdpType=Prep) ",	Este proyecto propone la implementación de un sistema de gestión de información
(VERB__Mood=Ind Number=Sing Person=3 Tense=Pres VerbForm=Fin) (DET__Definite=Def Gender=Masc Number=Sing PronType=Art) (NOUN__Gender=Masc Number=Sing) ",	(...) este método obtiene el rut ingresado por el usuario y lo busca en la lista de profesores.
(VERB__Mood=Ind Number=Sing Person=3 Tense=Pres VerbForm=Fin) (DET__Definite=Ind Gender=Masc Number=Sing PronType=Art) (NOUN__Gender=Masc Number=Sing)	Primeramente, accede a un proyecto , luego a una fase de este. Modifica. nombre, descripción y encargado de a fase.

Trigrama de Pos INF_ QUINTO	Ejemplos del corpus
(NOUN__Gender=Masc Number=Sing) (ADP__AdpType=Prep) (NOUN__Gender=Fem Number=Sing)	(...) ya que esta trabaja directamente con el motor de Docker lo que le permite manejar el balance de carga y políticas de red.
(ADP__AdpType=Prep) (NOUN__Gender=Fem Number=Sing) (ADP__AdpType=Prep)	Luego, se presentan cuatro de los motores de base de datos relacionales más utilizados y cuatro motores no relacionales (...)
(NOUN__Gender=Fem Number=Sing) (ADP__AdpType=Prep) (NOUN__Gender=Masc Number=Plur)	La tabla a continuación corresponde a la clasificación realizada por DB-engines de los 20 motores de base de datos más populares hoy en día.