



UNIVERSIDAD  
**NACIONAL**  
DE COLOMBIA

# **Metaanálisis sobre la presencia de la Lingüística Computacional en los pregrados de Ciencias del Lenguaje de Colombia**

**Cristian Nicolas Carvajal Rivera**

Universidad Nacional de Colombia

Facultad de Ciencias Humanas

Departamento de Lingüística

Bogotá, D.C.

2025



# **Metaanálisis sobre la presencia de la Lingüística Computacional en los pregrados de Ciencias del Lenguaje de Colombia**

**Cristian Nicolas Carvajal Rivera**

Trabajo monográfico presentado como requisito parcial para optar al título de:

**Licenciado en español y filología clásica**

Director:

PhD. Héctor Ramírez Cruz

Línea de Investigación:

Lingüística Computacional

Universidad Nacional de Colombia

Facultad de Ciencias Humanas

Departamento de Lingüística

Bogotá, D.C.

2025

*A mi familia y seres queridos*



## **Agradecimientos**

En primer lugar, presento mis agradecimientos a mi familia, en especial mis tías Brigitte Marcela Rivera González, Arianne Ernestina Rivera González y Lida Rosa Rivera González por su disposición en retroalimentar mi trabajo desde puntos de vista ajenos a la disciplina lingüística, así como su consejo a la hora de elaborar elementos gráficos. En segundo lugar, a mi madre, Edith Patricia Rivera González, porque en todo este trayecto me ha brindado ayuda de innumerables y vitales formas. Su ayuda ha sido indispensable. En tercer lugar, presento mis agradecimientos al profesor Héctor Ramírez Cruz por su labor desempeñada como director de mi trabajo de grado. Tanto sus retroalimentaciones como sus observaciones de forma y contenido fueron vitales en el buen desarrollo de este trabajo de grado. En cuarto lugar, me gustaría agradecer a la Universidad Nacional de Colombia, en especial al pregrado en Español y Filología Clásica. Cuando ingresé a esta carrera no tenía claridad de lo que me iba a encontrar, sin embargo, todas las asignaturas que pude cursar, las personas que pude conocer, los maestros de los que pude aprender y el aporte que estos han generado en mi vida es invaluable.

Por otra parte, me gustaría agradecer a las Instituciones de Educación Superior que colaboraron facilitando la documentación que les fue solicitada. Estas IES son la Universidad Católica de Oriente, la Universidad de Antioquia, la Universidad de la Amazonia y, por supuesto, la propia Universidad Nacional de Colombia. Es meritorio y digno de reconocimiento la buena disposición que presentaron de cara a la solicitud documental que les fue enviada. Dada la naturaleza institucional de los documentos se esperaba cierta renuencia al envío de los documentos solicitados. No obstante, los documentos fueron enviados y eso permitió que este trabajo se pudiera llevar a cabo.

## Resumen

### **Metaanálisis sobre la presencia de la Lingüística Computacional en los pregrados de Ciencias del Lenguaje de Colombia**

En esta investigación se buscó analizar la presencia de la Lingüística Computacional en los programas de pregrado del área de Ciencias del Lenguaje en Colombia. Por medio de un análisis cualitativo documental, se abordaron los siguientes documentos: Proyecto Educativo del Programa (PEP), mallas curriculares y micro currículos de seis pregrados y seis asignaturas. Esto se hizo con el fin de identificar pregrados de Ciencias del Lenguaje que incluyen asignaturas de Lingüística Computacional o contenidos transversales de esta en sus programas, comparar la presencia de la lingüística computacional en esos pregrados y, finalmente, valorar los contenidos, así como la profundidad con la que estos se abordan en los currículos revisados.

La investigación fue de tipo cualitativa exploratoria y la muestra consistió en los documentos oficiales institucionales que se les solicitaron a las universidades (PEP, malla curricular y micro currículos). La mayoría de los pregrados que fueron seleccionados para esta investigación fueron licenciaturas en lenguas, tanto en español como en lenguas extranjeras o modernas. Las asignaturas que se identificaron estaban relacionadas con el pensamiento lógico y matemático, acercamientos básicos a la programación, y la importancia, uso e implicaciones de las TIC en la educación y la sociedad en la era digital. Al final, se representó de forma gráfica y resumida el análisis realizado a los pregrados por medio de matrices DOFA para presentar debilidades, oportunidades, fortalezas y amenazas para cada programa.

**Palabras clave:** Lingüística Computacional, Procesamiento del Lenguaje Natural, Educación Superior, Tecnologías de la Información y la Comunicación, Pensamiento Lógico y Matemático.

## Abstract

### **Meta-analysis about the presence of Computational Linguistics in Language Sciences undergraduate programs in Colombia**

This research sought to analyze the presence of Computational Linguistics in undergraduate programs from the area of Language Sciences in Colombia. Through qualitative documentary analysis, the following documents were addressed: Educational Program Projects (EPP), the curricula, and the syllabi for six undergraduate programs and six courses. This was done to identify Language Sciences undergraduate programs that include Computational Linguistics courses or cross-curricular contents, compare the presence of computational linguistics in these undergraduate programs, and finally, assess the content and the depth with which it is addressed in the reviewed curricula.

This research study was exploratory and qualitative, and the sample consisted of official institutional documents requested from the universities (Educational Program Projects, Curricula, and syllabi). Most of the undergraduate programs selected for this research were language programs, both in Spanish and in foreign or modern languages. The subjects identified were related to logical and mathematical thinking, basic approaches to programming, and the importance, use, and implications of ICTs in education and society in the digital age. Ultimately, these subjects were represented graphically and summarized the analysis conducted on the undergraduate programs using SWOT matrices to show strengths, weaknesses, opportunities, and threats per program.

**Keywords: Computational Linguistics, Natural Language Processing, Higher Education, Information and Communication Technologies, Logical and Mathematical Thinking.**

# Contenido

	Pág.
<b>Resumen .....</b>	<b>VII</b>
<b>Abstract.....</b>	<b>VIII</b>
<b>Lista de figuras .....</b>	<b>XI</b>
<b>Lista de tablas.....</b>	<b>XII</b>
<b>Introducción .....</b>	<b>1</b>
<b>1 Planteamiento del Problema .....</b>	<b>4</b>
1.1 Pregunta de Investigación .....	7
1.2 Objetivos .....	7
1.2.1 General.....	7
1.2.2 Específicos .....	7
<b>2 Antecedentes de la Investigación.....</b>	<b>8</b>
<b>3 Marco Teórico .....</b>	<b>13</b>
3.1 Teorías Pedagógicas y Didácticas de la Enseñanza .....	13
3.2 El uso de las TIC en la educación .....	17
3.3 Lingüística Computacional.....	20
3.4 Pedagogía en la Lingüística Computacional.....	25
<b>4 Marco Metodológico .....</b>	<b>33</b>
4.1 Corpus.....	33
4.2 Muestra y recolección de información.....	37
4.3 Análisis de los datos.....	41
<b>5 Resultados .....</b>	<b>45</b>
5.1 Programas curriculares .....	46
5.1.1 Licenciatura en Lenguas Extranjeras con Énfasis en Inglés.....	46
5.1.2 Licenciatura en Literatura y Lengua Castellana.....	47
5.1.3 Licenciatura en lenguas extranjeras con énfasis en inglés y francés.....	49
5.1.4 Filología e idiomas: Alemán, Inglés y Francés .....	50
5.1.5 Español y Filología Clásica .....	51
5.1.6 Lingüística.....	53
5.2 Presencia de la Lingüística Computacional en los Programas Curriculares de Ciencias del Lenguaje.....	54
5.3 Valoración de los contenidos de las asignaturas seleccionadas.....	59
5.3.1 Licenciatura en Lenguas Extranjeras con Énfasis en Inglés.....	60

---

5.3.2	Licenciatura en literatura y lengua castellana .....	60
5.3.2.1	Pensamiento Lógico Matemático.....	60
5.3.2.2	Tecnologías de la Comunicación Aplicadas a la Educación .....	61
5.3.3	Licenciatura en Lenguas Extranjeras con Énfasis en Inglés y Francés	63
5.3.3.1	Uso de las TIC en la enseñanza y aprendizaje de las lenguas extranjeras.....	64
5.3.4	Pregrado en Filología e Idiomas y Pregrado en Español y Filología Clásica.....	66
5.3.5	Lingüística .....	66
5.3.5.1	Lógica Para Lingüistas .....	66
5.3.5.2	Lingüística Matemática.....	67
5.3.5.3	Informática Aplicada a la Investigación Lingüística.....	68
<b>6</b>	<b>Conclusiones.....</b>	<b>72</b>
<b>A.</b>	<b>Anexo: Script para extraer hipervínculos de las universidades .....</b>	<b>76</b>
<b>B.</b>	<b>Anexo: Script para extraer la información de las universidades .....</b>	<b>78</b>
<b>C.</b>	<b>Anexo: Script para extraer la información específica de los pregrados de ciencias del lenguaje .....</b>	<b>81</b>
<b>D.</b>	<b>Anexo: Procedimiento de recolección de información .....</b>	<b>83</b>
	<b>Bibliografía.....</b>	<b>84</b>

## Lista de figuras

Figura 1. Procedimiento de recolección de datos.....	40
Figura 2. Formato de Matriz DOFA (Debilidades, Oportunidades, Fortalezas y Amenazas) .....	42
Figura 3. Diagrama de flujo del procedimiento de análisis de datos.....	43
Figura 4. Matriz DOFA. Licenciatura en Literatura y Lengua Castellana .....	63
Figura 5. Matriz DOFA. Licenciatura en Lenguas Extranjeras con Énfasis en Inglés y Francés.....	65
Figura 6. Matriz DOFA. Lingüística. ....	69

## Lista de tablas

<b>Tabla 1. Corpus</b> .....	36
Tabla 2. Comparación sobre la presencia de la LC en los pregrados de Ciencias del Lenguaje seleccionados .....	56

---

## Introducción

La humanidad se enfrenta a una nueva revolución en los tiempos que corren. No es un secreto para cualquiera que lea este documento que la tecnología y sus avances se integran cada día más rápido a nuestras vidas de formas que no somos capaces de entender o imaginar. Todos los días aparecen enunciados y publicaciones en redes sociales y plataformas de internet en las que circulan discursos variados sobre las tecnologías digitales, su papel en la sociedad, su importancia, sus implicaciones legales y sus peligros. Es por eso por lo que los profesionales de la educación, del lenguaje y de las humanidades, en general, no pueden ser ajenos a estas tecnologías y sus avances.

Las humanidades son las ciencias más cercanas a la vida personal y social de los individuos. Dentro de sus disciplinas o áreas del conocimiento se pueden encontrar la antropología, la historia, la lingüística, la sociología, entre otras. Para este trabajo, es especialmente relevante la lingüística. Esta disciplina se encarga del estudio de una de las características humanas más elementales, más fundamentales para la especie y su desarrollo como civilización, como sociedad, el lenguaje. Es un estudio tan vital que atraviesa de forma transversal a todas las demás áreas del conocimiento que se puedan imaginar. Sin importar la profesión a la que una persona se dedique, la lingüística estará involucrada, y sus diferentes niveles de análisis harán eco dentro de ellas configurando sus prácticas discursivas.

Retomando el componente tecnológico, la subdisciplina que compete a este trabajo de grado es la *Lingüística Computacional (LC)*. Esta subdisciplina se encuentra ubicada en el punto de encuentro existente entre la lingüística, las ciencias de la computación y la inteligencia artificial. Específicamente, “La lingüística computacional (LC) es una ciencia interdisciplinaria que se ubica entre la lingüística y la informática, con énfasis en la lingüística. Su fin es la elaboración de modelos computacionales que reproduzcan uno o más aspectos del lenguaje humano” (Domínguez Burgos, 2002). Igualmente, es una disciplina que no se limita al trabajo académico y teórico puesto que tiene un gran componente orientado a las investigaciones prácticas financiadas por el sector privado.

La LC trabaja estrechamente con las herramientas computacionales que se han desarrollado para el *Procesamiento del Lenguaje Natural (PLN)*. Este es una de las principales aplicaciones y campos de estudio que se abordan dentro de las investigaciones de la LC. Su objetivo es diseñar sistemas computacionales que puedan imitar parcialmente el lenguaje humano, así como entenderlo. En sus textos principales se suelen abordar temas como lenguajes de programación, razonamiento lógico, etiquetados morfológicos y sintácticos, recuperación e información, análisis de sentimientos o emociones, entre otros.

---

Todas estas son aplicaciones de la LC con las que las personas interactúan constantemente sin saberlo. Desde el momento en el que una persona se detiene a prestar más atención a un contenido en específico, pasando por una interacción controlada, como lo puede ser un botón de reacción, y finalizando con una manifestación de una opinión o sentimiento en palabras. Todo esto es el material con el que estas áreas de trabajo dentro de la LC trabajan y al cual el sector privado le suele apostar con entusiasmo.

Poder llevar a cabo todo este tipo de análisis y mediciones es de gran importancia en el mundo de los negocios. No obstante, debemos recordar que se trata de una subdisciplina de la lingüística, por lo tanto, su componente humanístico es importante. La LC también ha contribuido en aplicaciones orientadas a la inclusión de comunidades con limitaciones físicas o cognitivas. Ha contribuido a llevar la educación a otros contextos con otro tipo de recursos. Siempre debe tenerse en cuenta que las herramientas tecnológicas, más allá del objetivo con el que son creadas, dependen de cómo las usemos. Es por eso por lo que el uso de estas debe ser consciente, crítico y reflexivo.

De ahí nace el interés que propició este trabajo de grado, de la inquietud sobre cuál es la importancia que se le da a la lingüística computacional en los pregrados de las Ciencias del Lenguaje fue lo que motivó la realización de este trabajo. Su objetivo principal es analizar la presencia de la Lingüística Computacional en los programas de pregrado del área de Ciencias del Lenguaje en Colombia. Para lograrlo se propusieron tres objetivos específicos: (1) Identificar los programas de pregrado del área de ciencias del lenguaje que incluyen la Lingüística Computacional en sus planes de estudio, (2) Comparar la presencia de la Lingüística Computacional entre diferentes programas de diversas universidades y (3) Valorar los contenidos y profundidad con que se aborda la Lingüística Computacional en los currículos revisados.

Estos objetivos tenían el propósito de orientar la reflexión que se haría sobre los pregrados seleccionados. Para este trabajo le fueron solicitados a las universidades los Proyectos Educativos del Programa (PEP) y malla curricular de cada programa, así como los micro currículos de las asignaturas que fueron elegidas de estas carreras. El alcance de esta investigación se delimitó geográficamente al país de Colombia, todas las universidades en las que se encontraran programas de Ciencias del Lenguaje que incluyeran asignaturas de LC o contenidos transversales de esta incluidos en los programas. En cuanto a la delimitación cronológica, se trabajó de forma sincrónica con la información actualmente disponible sobre estos programas curriculares. Se seleccionaron 22 universidades a las cuales se les enviaron cartas formales para la solicitud de información vía correo electrónico.

De esas veintidós universidades solo respondieron cinco, de las cuales tres facilitaron los documentos solicitados, una manifestó no poder compartir los documentos, y la última manifestó su interés por colaborar en el trabajo por medio de una reunión virtual que nunca se concretó. Los documentos que se solicitaron para construir el corpus de la investigación fueron tres: el Proyecto Educativo del Programa (PEP), la malla curricular de cada pregrado y el micro currículo de cada asignatura seleccionada. El análisis realizado

---

se representó finalmente en una matriz DOFA (Debilidades, Oportunidades, Fortalezas, Amenazas) para cada programa académico seleccionado.

En cuanto a la estructura del documento, este trabajo de grado consta de seis capítulos. En el primer capítulo se hace la presentación del problema de investigación y los objetivos establecidos para este trabajo de grado. En el segundo capítulo se presentan antecedentes investigativos correspondientes a investigaciones que se leyeron al inicio de la revisión bibliográfica y que fueron seleccionadas por ser similares a la investigación en curso. En el tercer capítulo se introdujo el marco teórico, el cual aborda en su orden de mención los siguientes temas: Teorías Pedagógicas y Didácticas de la Enseñanza, El uso de las TIC en la educación, Lingüística Computacional y la Pedagogía en la Lingüística Computacional. En el cuarto capítulo se introduce el marco metodológico conformado por el corpus, el procedimiento de recolección de datos y el proceso de análisis de datos. En el quinto capítulo se presentan los resultados de esta revisión de literatura. En el sexto capítulo se presentaron las conclusiones. A continuación, se presentan los Anexos, que son cuatro y fundamentalmente corresponden a los códigos para automatización de recolección de datos presentados en el capítulo cuatro junto al diagrama de flujo que representa el procedimiento de recolección de datos que se siguió en este trabajo de grado. Finalmente, se presentan las referencias bibliográficas utilizadas en este trabajo de grado.

Para concluir esta introducción, se hará mención sobre los principales hallazgos en el desarrollo de este trabajo de grado. Para empezar, se considera importante señalar que la mayoría de los pregrados que cumplieron con los criterios de selección fueron licenciaturas dedicadas al área de las lenguas, entre las que se encuentran el español y las lenguas extranjeras o modernas. Solo se encontró un pregrado dedicado a la lingüística pura, siendo este el pregrado en Lingüística de la Universidad Nacional de Colombia. Por otro lado, las asignaturas que fueron identificadas y seleccionadas para su análisis estaban dentro del marco de la lógica, el pensamiento lógico y matemático, el uso de las TIC en la educación, la programación básica y la informática aplicada a la investigación lingüística.

Se pudo determinar que las licenciaturas priorizan asignaturas sobre el uso de las TIC en la educación, aunque también le daban un papel relevante a la lógica. Por su parte, en el pregrado en lingüística se identificó un progreso natural entre las tres asignaturas seleccionadas. Esta organización progresiva existente entre las asignaturas les permite a los estudiantes que cursen esas tres asignaturas en el orden correcto obtener una formación básica en LC de buen nivel. Esta preparación es de gran importancia de cara a posgrados o especializaciones en LC o PLN, así como al ejercicio profesional en diferentes campos de la investigación lingüística.

# 1 Planteamiento del Problema

La *Lingüística Computacional (LC)* es un tema que se lleva investigando y desarrollando desde mediados de los años cincuenta. Esta subdisciplina de la lingüística nace a causa del interés por parte de los estadounidenses en poder descifrar comunicaciones enemigas de forma eficiente y ha generado avances en pro de facilitar la investigación lingüística, así como el manejo de datos lingüísticos. Sin embargo, su inclusión en los programas de ciencias del lenguaje no parece ir de la mano con los avances que se producen en esta área. Un ejemplo de pregrado en ciencias del lenguaje es la carrera de Español y Filología Clásica de la Universidad Nacional de Colombia. Esta carrera tiene un amplio contenido curricular dentro del cual se incluye un componente de lingüística. En su componente de lingüística se incluyen asignaturas básicas de lingüística, como los niveles de análisis lingüístico (fonética, fonología, morfología, sintaxis, semántica, pragmática, etc.). Sin embargo, el acercamiento que se tiene dentro de esos cursos con las herramientas de la LC orientadas a la investigación lingüística es poco. Este pregrado tampoco cuenta con una asignatura dedicada a la LC, pensamiento computacional, lógica u otros que puedan servir de base para un posterior acercamiento a la LC, ya sea en los componentes de fundamentación o en el componente de electivas. Lo anterior es entendible, en cierto sentido, dado que el perfil del egresado de este programa curricular no es el mismo que el de un estudiante de pregrado en lingüística pura. Esto quiere decir que para el filólogo es necesario conocer los conceptos y las aplicaciones básicas de la lingüística, ya que se podría abordar como una subdisciplina de esta última, más no es imperativo que sea ampliamente versado en las diferentes ramas y disciplinas de la lingüística; eso incluye, por supuesto, a la LC.

Pero ¿por qué es importante la LC? ¿qué hace que valga la pena preguntarse sobre su presencia en los pregrados de ciencias del lenguaje? Parte de la respuesta podría ser que ya hay presencia de la LC en el mundo de los posgrados, especializaciones y doctorados. Por esta razón, cabría imaginar que en los pregrados de las ciencias del lenguaje se estaría dando un primer acercamiento a esta de manera que los estudiantes puedan aprender conceptos básicos como algoritmos, programas, lenguajes de programación, etc., y así poder llegar mejor dispuestos al posgrado. Como ejemplos de posgrados con contenidos de LC se presenta la Universitat Pompeu Fabra, la cual tiene

una especialización en LC<sup>1</sup> o, por otro lado, en Colombia, la Universidad de Antioquia, que ofrece una Maestría en Lingüística<sup>2</sup> que incluye dentro de sus líneas de investigación y formación a la LC. Para empezar a cursar estos posgrados, es útil tener una formación previa básica sobre programación, lógica, lenguajes formales, e incluso, pensamiento computacional si se quiere tener un buen desempeño. A esto contribuye, entonces, el acercamiento previo a la LC que se podría ver dentro de los propios cursos de lingüística o en cursos específicos dedicados a la LC. No obstante, la pregunta persiste, ¿si no me interesa dedicarme a la LC, por qué debería verlo dentro de los contenidos de mi programa curricular? ¿realmente necesito adquirir conocimientos elementales en LC? ¿en qué me ayuda como estudiante de ciencias del lenguaje?

Para las preguntas anteriormente presentadas puede haber varias respuestas. En primer lugar, se puede abordar la LC a partir de sus desarrollos orientados a la investigación lingüística, los cuales facilitan la labor del investigador, así como permiten reducir los costos de tiempo y de personal tanto para la recolección como para el análisis de datos investigativos. En segundo lugar, las aplicaciones de la LC no se limitan al ámbito teórico y la academia, que es algo que puede llegar a preocupar a un estudiante de pregrado en ciencias del lenguaje en algún momento de su devenir universitario. Cuando los estudiantes se acercan al final de su carrera pueden surgir las siguientes preguntas: ¿y qué haré cuando me gradúe? ¿en qué me puedo desempeñar? La LC puede abrir nuevos campos laborales para todos estos estudiantes. Dentro de esos campos laborales u oportunidades de empleo se encuentran la traducción automática y los programas de asistencia virtual con reconocimiento de voz, aunque no son los únicos. Estas tecnologías se soportan en las investigaciones y desarrollos de la LC. En tercer lugar, y como ya se mencionó antes, existen posgrados en lingüística o humanidades digitales para los cuales una formación previa en LC o sus fundamentos es de gran ayuda o incluso un requisito.

No obstante, es claro que un estudiante se puede graduar y empezar una vida laboral con relativo grado de éxito sin haber visto contenidos de LC en toda su carrera de ciencias del lenguaje. Entonces, la duda sobre la utilidad de la LC para el estudiante de pregrado de ciencias del lenguaje permanece vigente. Es por eso por lo que también se hace necesario mencionar que dentro de los beneficios secundarios que tiene la LC para los alumnos de ciencias del lenguaje se incluye un cambio de paradigma y de razonamiento para resolver problemas y abordar los conceptos de la lingüística en aplicaciones prácticas. Aprender sobre LC supone adquirir conocimientos básicos de computación e, incluso, de programación. Para poder lograr esto, es necesario emprender un camino que lleva al estudiante hacia nuevas áreas del saber que se alejan de lo que sería usual dentro de la lingüística, que lo enfrenta con razonamientos lógicos-rationales que tienen cierto grado de lejanía con los razonamientos más cualitativos y humanos que

---

<sup>1</sup> Enlace: <https://www.upf.edu/web/masterlinguistica/linguistica-computacional>

<sup>2</sup> Enlace: <https://www.udea.edu.co/wps/portal/udea/web/inicio/unidades-academicas/comunicaciones/estudiar-facultad/posgrados/maestria-linguistica>

---

se llevan a cabo en una investigación lingüística usual. Trabajar con máquinas, con computadoras plantea un reto nuevo para el lingüista: una máquina no puede entender bien las emociones de manera empática, sino por estadísticas y marcados de las palabras o morfemas dentro de estas, una máquina no puede inferir o hacer suposiciones de la misma forma en que lo hace un humano, trabaja desde un paradigma lógico-matemático. Esta forma diferente de trabajar y de pensar le exige al investigador que se acerca por primera vez a la LC dar el paso hacia comprender y apropiarse estas lógicas en cierta medida. También, le exige al académico apropiarse más del hardware y el software que utiliza, incluso puede llevarlo a desarrollar software propio de acuerdo con sus necesidades investigativas, lo que también le brinda autonomía investigativa y enriquece a la comunidad académica con nuevos recursos para la recolección y el análisis de datos.

Ahora bien, hasta el momento hemos hablado de los beneficios que traería para los estudiantes, sin embargo, ¿qué queda para el ámbito educativo? Implementar la LC en las asignaturas de lingüística puede requerir de inversión monetaria en la forma de equipos, licencias y capacitación para el personal docente, además de que requeriría de inversión en la forma del tiempo que tomaría llevar a cabo dichos cambios. Tal gasto monetario puede significar un problema en un inicio, en especial en universidades que tengan déficit presupuestal, puesto que deben priorizar muy bien los gastos que se hacen. Además de que no se puede ignorar el factor humano, la disposición de los docentes a la hora de apropiarse herramientas y metodologías que pueden parecer ajenas para algunos de ellos o innecesariamente complejas. Por supuesto, esto requiere de trabajo por parte de la institución educativa en cuestión para la implementación de programas de capacitación docente y comités académicos que permitan identificar qué contenidos o herramientas de la LC realmente vale la pena introducir dentro de sus cursos.

Por último, también valdría la pena hacer mención sobre el trabajo interdisciplinar. En un mundo mediado por la tecnología, como lo es el nuestro, poder ir más allá de las barreras de una disciplina dada y apropiarse de los nuevos desarrollos que salen a la luz cada día también facilita el trabajo conjunto con otras disciplinas y los avances que se generan en ellas. Es por eso por lo que vale la pena considerar el trabajo conjunto en cursos, seminarios, talleres o proyectos académicos con las facultades de ingeniería y de sistemas, las cuales pueden aportar mucho a la lingüística para la comprensión de los nuevos sistemas. Además de que las áreas de la computación también necesitan ayuda considerable de la lingüística para el procesamiento del lenguaje natural. Con base en lo anterior, el aporte de la LC a la investigación lingüística, al perfil académico y profesional de los estudiantes, al ejercicio docente, a la vida laboral de los futuros profesionales en ciencias del lenguaje se puede entender por qué es importante incluir a la LC en esos pregrados bien sea como un contenido que atravesase a los cursos básicos de lingüística o como cursos especializados para que el futuro egresado pueda ser más competitivo y pueda acceder a las nuevas posibilidades que ofrece la era digital.

## **1.1 Pregunta de Investigación**

¿Cuál es la presencia de la Lingüística Computacional en los pregrados del área de Ciencias del Lenguaje de las universidades de Colombia?

## **1.2 Objetivos**

### **1.2.1 General**

Analizar la presencia de la Lingüística Computacional en los programas de pregrado del área de Ciencias del Lenguaje en Colombia.

### **1.2.2 Específicos**

- 1) Identificar los programas de pregrado del área de ciencias del lenguaje que incluyen la Lingüística Computacional en sus planes de estudio.
- 2) Comparar la presencia de la Lingüística Computacional entre diferentes programas de diversas universidades.
- 3) Valorar los contenidos y profundidad con que se aborda la Lingüística Computacional en los currículos revisados.

---

## 2 Antecedentes de la Investigación

En el trascurso de esta investigación se ha encontrado gran variedad de información sobre lingüística computacional, pensamiento computacional y humanidades digitales. Sin embargo, textos con relación directa a su papel en la educación superior han sido más escasos. No obstante, se encontraron documentos que describen aplicaciones, modelos de enseñanza y experiencias de cursos que brindan un panorama útil sobre cómo se está implementando a nivel global en diferentes niveles de enseñanza.

En primer lugar, presentamos un artículo enfocado en qué se realiza en el mundo profesional de la LC: *Análisis de contenido y lingüística computacional: su rapidez, confiabilidad y perspectivas* de Brenda Lía Chávez y Jorge Martín Yamamoto. En este texto se busca comparar el análisis de unos datos realizado por un programa basado en LC y el análisis realizado por una serie de investigadores lingüistas. Como metodología se utiliza el análisis de contenido y la investigación émica, como la denominan los investigadores. Esto es un proceso deductivo a partir de los propios datos en el que se recoge información a partir de entrevistas en profundidad y cuestionarios de respuesta abierta (Chávez & Martín Yamamoto, 2014). Por otro lado, la herramienta utilizada para el procesamiento de esta información es el análisis de contenido. Mientras que, como principales hallazgos se encontró que con el software de LC el tiempo y costo requeridos para procesar los datos fue menor, además de que la confiabilidad de los resultados en lo referente a la forma en la que se presentan y el contenido de estos fue mayor dado que el software siempre usaba la misma plantilla y disposición de los datos en la representación gráfica que arrojaba al final del proceso. Contrario a como ocurre con el análisis manual realizado por los investigadores humanos, puesto que cada cual trabajaba de forma diferente y presentaba los datos con ligeras variaciones, lo cual añadía un paso más: normalizar los datos, algo que el software de LC ya hace antes de arrojar los resultados. También, se llega a la conclusión de que el software de LC ofrece ventajas en cuanto a la visualización tanto de los datos como de los resultados, una visualización gráfica y cómoda de analizar al momento de reunir la información y de sintetizar los datos finales. Adicionalmente, el software de LC se puede seguir configurando y parametrizando con cada uso, de manera que se perfeccione con el paso del tiempo. Además, una vez perfeccionado, el programa utilizado se puede reutilizar en otras investigaciones, en las cuales también se podrían hacer todo tipo de cambios o adecuaciones que sean requeridos sin ningún problema. Al final del trabajo las plantillas refinadas quedaron como recurso para futuras investigaciones.

---

En segundo lugar, se presenta un artículo en el que se aborda la situación conflictiva que se dio en la educación universitaria a partir de la pandemia del SARS COV2 (también llamado COVID-19) y el paso repentino a la virtualidad por el que se vieron afectados universidades, profesores y alumnos: el texto llamado *Estudiantes de pregrado frente a la educación mediada por TIC: percepciones en contextos de pandemia* de Gerson Rueda Vera et al. En él se aborda el siguiente objetivo: "...analizar las percepciones de los estudiantes de pregrado en torno a la educación mediada por TIC en contexto de pandemia – Covid-19" (Rueda Vera et al., 2021, p. 321). La investigación fue de corte empírico-analítica y sustentada en la investigación no experimental que ellos denominan *transeccional*. Dentro de los resultados de su trabajo investigativo se halló que dar el paso a una educación mediada por TIC de forma tan abrupta implicó problemas para los alumnos y los docentes. Los autores hacen énfasis en que la educación virtual y la educación mediada por TIC requieren de mayor autonomía por parte del alumno y de formación tecnológica y didáctica que no tenían muchos docentes. También, destacan que las poblaciones más jóvenes y que no pertenecen a ninguna población especial o víctima del conflicto fueron las que tuvieron una postura más pesimista sobre esta modalidad educativa, esto reflejaba la importancia que tiene la situación socioeconómica de los estudiantes en su postura frente a la educación formal y las herramientas que se dejan a su disposición para su formación. Es comprensible que en zonas remotas del país a las que no llegan las instituciones formales de manera presencial, pero sí los servicios de energía e internet, la educación virtual sea la mejor opción para recibir educación profesional. No se dice literalmente, pero se da a entender que las personas que trabajan y tienen menos tiempo, así como las personas que tienen condiciones socioeconómicas más complejas, suelen adoptar posturas más asertivas y optimistas cuando se trata de educación virtual. Finalmente, aunque la muestra fue de poco más de 500 estudiantes en una muestra muy heterogénea, permitió identificar puntos críticos que generaron abandono escolar y otra serie de problemas que se dieron en la pandemia; esto puede ayudar a la hora de tomar acciones preventivas orientadas a situaciones de anormalidad social extrema.

En tercer lugar, se presenta un documento que la Red Colombiana de Humanidades Digitales (RCHD) denomina "cadáver exquisito" y que presenta el recorrido que ha llevado a cabo la RCHD desde 2016 en su labor de llevar la palabra de las Humanidades Digitales a las instituciones de educación superior del país. El texto *Humanidades Digitales "a lo colombiche": cadáver exquisito de la Red Colombiana de Humanidades Digitales* presenta la función de la RCHD y los diferentes cursos, conferencias, grupos de trabajo, e incluso un programa curricular de posgrado que ayudaron a crear en conjunto con la Universidad de los Andes, en los que la RCHD ha participado, fundado o contribuido significativamente y que constituyen el conglomerado de su trabajo desde su creación. En el texto se presentan diferentes recursos que la RCHD ha generado en colaboración con la Biblioteca Nacional y el Departamento de Historia de la Universidad Nacional de Colombia para la formación en HD y, sobre todo, la integración de profesionales de todas las carreras en torno a las HD. Si bien este documento brinda una vista global de las Humanidades Digitales puestas en práctica y no tanto un extenso desarrollo teórico sobre estas, su

---

contenido, la labor que realizan en la RCHD, es de gran importancia para este trabajo porque es un ejemplo real, y no solo teórico, de cómo las ciencias humanas se pueden acercar con los desarrollos que la tecnología trae día tras día para sus respectivas disciplinas, así de cómo se pueden integrar con otras áreas del conocimiento más allá de las propias humanidades.

En cuarto lugar, presentamos una investigación mucho más acorde con el presente trabajo de grado. El nombre del artículo en cuestión es *Teaching computational linguistics to a large, diverse student body: courses, tools, and interdepartmental interaction* de Jason Baldrige y Katrin Erk. En este texto se describe una serie de cursos que, a lo largo del tiempo, fueron adaptados para poder enseñar LC a estudiantes universitarios de la mejor manera posible tomando en cuenta el gran número de alumnos y sus diferencias tanto demográficas como formativas. Se trató de un trabajo extenso que requirió de constante reelaboración por parte de los investigadores ya que la heterogeneidad no se limitaba a la edad y el sexo biológico, sino que también a la formación previa de los alumnos, es decir al programa académico del cuál provenían. No solo ese aspecto fue relevante en el desarrollo de su investigación puesto que los conocimientos previos en materias de ciencias de informática y computación de cada alumno también representaron un reto considerable que debía ser abordado. Por otro lado, se presentan los cursos que se fueron realizando: Computational Linguistics I y II, Working with Corpora, una serie de seminarios, Introduction to Computational Linguistics, etc. También, presentan las herramientas de enseñanza y los tutoriales que utilizaron, siendo algunos de ellos privativos, y otros de creación propia y código libre. De hecho, se hace mención especial de un estudiante, Ben Wing, quien creó su propia versión modificada de uno de los programas que utilizaban en uno de los cursos, versión de la cual, otro estudiante sacaría otra versión modificada y mejorada tiempo después. El trabajo realizado por el estudiante Wing es un ejemplo de cómo la LC abre nuevas posibilidades a los estudiantes y a sus investigaciones, de igual manera, permite evidenciar que esas oportunidades para el crecimiento profesional no están limitadas a la vida profesional del egresado, sino que podrían empezar a dar resultados mientras se está estudiando el pregrado. Por lo tanto, el principal aporte de este documento no son solo las diferentes herramientas de software que se generaron en los cursos, sino también el precedente que dejan para el desarrollo curricular en forma de una investigación que toma en cuenta diversos factores más allá de temas, número de alumnos y docentes. En cuanto a los resultados, en primer lugar, se señala que se deben estratificar los cursos de acuerdo con el trasfondo de los estudiantes, puesto que el nivel de conocimiento o dominio sobre herramientas informáticas y computacionales es muy heterogéneo; en segundo lugar, se indica que la formalización del curso dentro de los calendarios académicos aumenta la tasa de inscripciones; y finalmente, se llama la atención para que en el nivel de posgrado no se descuiden el contexto y la estructura de los cursos.

En quinto lugar, se presenta al trabajo de Cecilia Ovesdotter Alm et al. del Instituto de Tecnología Rochester en su artículo *A Pedagogical Model for Computational Linguistics across Curricular Boundaries*. En su documento declara:

“Entre las principales motivaciones del taller estaban fomentar la interacción interdisciplinaria, exponer a los estudiantes al pensamiento crítico y la curiosidad intelectual más allá de sus respectivos campos de estudio y ayudarlos a apreciar desde el principio de sus carreras lo que se puede obtener de un compromiso más amplio con las formas computacionales de modelado” (Alm et al., 2016).

Este taller fue una experiencia interdisciplinaria que pretendía exponer a los estudiantes al pensamiento crítico y otras perspectivas epistemológicas mediante las actividades propuestas dentro de este. Sin embargo, no pretendía que los estudiantes adquirieran dominio avanzado sobre LC o PLN, sino que buscaba un acercamiento holístico a la disciplina. Durante el desarrollo de los talleres, muchos de los estudiantes apreciaron los cuestionamientos éticos, técnicos e intelectuales a los que fueron expuestos. Muchos de estos desafíos éticos que se mencionan en el documento surgieron a partir de la interacción entre las diferentes disciplinas. Además, la experiencia permitió visualizar en la práctica el papel fundamental de la nivelación de los estudiantes, dado que los contenidos fueron adecuados tomando en cuenta el trasfondo, es decir, el conocimiento previo de los estudiantes. Adicionalmente, el taller implementó tres etapas de retroalimentación por parte de los alumnos. La primera fue antes del curso y sirvió para determinar el nivel de conocimiento y dominio que los participantes tenían previo a que participaran del taller. La segunda se dio al finalizar el curso y permitió comparar los conocimientos que tenían los estudiantes al finalizar el taller con los conocimientos que tenían antes de tomarlo. Finalmente, se dio la oportunidad de hacer una tercera y última retroalimentación meses después del curso con el fin de determinar el impacto del curso en los participantes. Esa tercera retroalimentación no se tenía contemplada en un inicio, sin embargo, los propios alumnos fueron quienes manifestaron el interés en llevarla a cabo. Se señala la importancia de analizar el impacto que los contenidos de un taller como este tienen en los estudiantes a través del tiempo. Asimismo, los autores destacan el gran interés que generaron los participantes por el PLN.

Finalmente, dentro del libro *Innovación educativa y gestión del conocimiento* (2020), producido por la editorial de la Universidad Santo Tomás, encontramos el capítulo titulado *Enseñanza del pensamiento computacional: metodologías y tecnologías educativas* de Xabier Basogain Olabe, Miguel Ángel Olabe Basogain y Juan Carlos Olabe Basogain. En este capítulo se recopilan experiencias educativas que se consideran importantes como antecedentes a lo que se podría llevar a cabo en una implementación de la LC dentro de los pregrados de ciencias del lenguaje. A lo largo del texto se presenta una serie de proyectos de formación en *pensamiento computacional (PC)* orientados principalmente a la población preadolescente, adolescente y sus profesores. Dentro del texto se detallan las experiencias y recursos tecnológicos que se utilizaron, dejando pie a futuras investigaciones sobre una educación más práctica que permita unificar *hardware* y *software* para aprender. Lo que más destaca de este capítulo es el uso de proyectos, sobre todo de la programación, como herramienta de enseñanza para la apropiación de las lógicas y herramientas para programar que se ponen a disposición de los niños. Esta metodología es importante porque las habilidades en programación y computación se

---

adquieren con la práctica, más que con la teoría. Este enfoque es altamente aplicable en los pregrados de ciencias del lenguaje ya que lo que se buscaría en cursos de LC o en las aplicaciones de la LC en cursos de lingüística de corte más tradicional, sería justamente que los alumnos apropiaran esos conceptos y esa lógica computacional, ese *pensamiento computacional*, por medio de la praxis.

Resumiendo, a lo largo de las investigaciones presentadas anteriormente se puede empezar a entender por qué es importante la LC, la facilidad que ofrecen sus herramientas tecnológicas para la investigación lingüística y las oportunidades que generan estas herramientas para la recopilación y análisis de datos. Asimismo, se destacan las experiencias en la implementación de cursos dedicados a la LC y al PC, se considera que son valiosas y deben ser tomadas en cuenta como antecedente a la hora de pensar cursos sobre LC o de modificar cursos ya existentes para integrar los avances que la LC provee a las propias disciplinas clásicas de la lingüística, como pueden ser la fonética, la fonología, la morfología o la sintaxis. Sin embargo, también se han podido reconocer los desafíos técnicos, económicos y metodológicos que plantea la implementación de la LC en los ya mencionados pregrados. Sobre esto es importante tener en cuenta que no basta con enfrentar a los estudiantes con un computador y esperar que sepan manejar los programas a la perfección. Por el contrario, para que los estudiantes puedan sacarle el máximo provecho a la LC e integrarla en su formación académica, es necesario nivelarlos en cuanto a las competencias digitales que se requiere aprender para el manejo de los programas y dispositivos a utilizar e, incluso, propiciar que modifiquen los ya existentes, o bien, que generen unos nuevos. Aun así, no se ha podido encontrar un texto que analice los contenidos de los pregrados en ciencias del lenguaje en Colombia, así como la presencia o ausencia de la LC en estos. Es por eso por lo que la labor investigativa de revisar los programas curriculares en busca de asignaturas de LC o que puedan contener LC, la revisión de asignaturas de LC o que puedan contener LC, y la revisión de demás documentación institucional relacionada, será competencia del presente trabajo de grado.

## 3 Marco Teórico

El presente trabajo de grado se ubica dentro del marco de la lingüística computacional en la educación superior y dado su carácter exploratorio pretende aportar una visión general del panorama de la LC en la educación superior, específicamente en los pregrados de ciencias del lenguaje en Colombia. Es por eso por lo que en este marco teórico es necesario hacer una revisión tanto de la pedagogía como de la propia Lingüística Computacional. Para poder construir este marco teórico, se partirá de una visión más generalizada de la pedagogía y de la Lingüística Computacional para, posteriormente, continuar con una visión más específica de la pedagogía orientada a la LC en la educación superior. Finalmente, se cerrará el capítulo con un breve resumen.

### 3.1 Teorías Pedagógicas y Didácticas de la Enseñanza

Comenzamos el capítulo con una revisión orientada a la pedagogía, un concepto amplio, pero del cual solo tomaremos al *Aprendizaje Basado en Proyectos* y la *Pedagogía por Proyectos*. Ambas son pedagogías relacionadas, pero con ciertas diferencias claves que se mencionarán más adelante. Por su parte, la didáctica se ve fuertemente delimitada cuando se parte de estos enfoques pedagógicos y es por eso por lo que será abordada a lo largo del apartado pedagógico.

Empezaremos con la pedagogía basada en proyectos (o pedagogía del proyecto). La PDP se caracteriza por estar atravesada por proyectos integrales previamente establecidos que se desarrollan en periodos de tiempo medianamente extensos. Con relación al tiempo es importante tener en cuenta que para que un proyecto sea efectivo no debe ser ni muy corto, de manera que no se puedan abordar apropiadamente los temas, pero tampoco demasiado extenso, lo cual puede llegar a agotar a los alumnos. Lo anterior aplica también para la PPP. Estos proyectos no suelen tener muchas variaciones, y si bien buscan que el aprendizaje sea más participativo por parte de los estudiantes, y más dinámico, al ser preestablecido no es del todo flexible. Los contenidos y competencias se plantean previamente por parte del docente de acuerdo con los objetivos que se hayan establecido para el curso, programa o unidad de contenido. Se busca abordar diferentes contenidos de forma integral, por lo que no es extraño que se aborden varios temas y competencias, es por eso por lo que este enfoque se puede adoptar para prácticamente cualquier materia, asignatura o curso. Además, la PDP, al estar determinada por los intereses de los docentes, instituciones educativas o grupos de trabajo que los diseñan y

ejecutan, son fácilmente trasladables a otras poblaciones similares con mínimos cambios, lo que aumenta su utilidad y permite una continua mejora de los mismos proyectos. Es decir, al poder aplicar un mismo proyecto en diferentes instituciones con grupos variados de alumnos y realizando ajustes mínimos el proyecto se va depurando, se mejora continuamente, lo que reduce la inversión necesaria para la planificación, diseño y ejecución de un proyecto. No se trata de tener una plantilla y convertir el proceso formador en el diligenciamiento de formularios y planificaciones esquemáticas que se repiten continuamente, sino en el mejoramiento continuo a partir de experiencias previas.

Por su parte, Zambrano Briones et al. (2022) señalan que la PDP conlleva la construcción del conocimiento a través de la interacción con la realidad. Esto significa que para el diseño de los proyectos que se desarrollan en esta pedagogía se tienen en cuenta las necesidades inmediatas de los alumnos y del entorno, del contexto en el que ellos habitan. Este rasgo es parcialmente compartido con la PPP, pero con un enfoque ligeramente diferente que se abordará más adelante. Por su parte, si bien el docente es quien mayormente crea el proyecto, su labor en la ejecución de las actividades no es la de ofrecerle a sus alumnos la tradicional educación magistral a la que estamos acostumbrados:

“En este proceso de enseñanza el docente será apoyo desde su papel de guía u orientador de las actividades que ejecute el estudiante, interactuando con el estudiante cada vez que lo requiera y estableciendo una orientación educativa más horizontal y dinámica” (Zambrano Briones et al., 2022).

También, es importante mencionar las habilidades que se desarrollan en este tipo de proyectos. Dentro de las competencias que los estudiantes refuerzan constantemente en los proyectos se encuentran: el trabajo autónomo, el trabajo en equipo, la habilidad para investigar, la capacidad para plantear hipótesis y discutir las, la argumentación, la planificación, la organización, el uso de los recursos tecnológicos, entre otras habilidades relacionadas con el aprendizaje y que son esenciales para la vida en general. La formación que se recibe en la PBP no está limitada a los contenidos contemplados en un plan de estudios dados, sino que son aprendizajes para la vida. Es por eso por lo que los modelos pedagógicos que trabajan proyectos son considerados como integrales y constructores de conocimiento.

Por otro lado, presentamos la *Pedagogía por Proyectos (PPP)*. La PPP se parece en principio a la PDP, pero tiene marcadas diferencias en el enfoque mediante el cual se implementa en las aulas. Se trata de una pedagogía horizontal y no vertical, como la del PDP. La forma en la que se trabaja en la PPP es la siguiente: se busca abordar una pregunta problema a través de un proyecto, de forma similar, si se quiere, a como se elabora un trabajo investigativo. Es decir, la **pedagogía por proyectos** se puede entender como

“...un proceso en el que las interacciones están mediadas por un problema o un centro de interés que empuja hacia la conjetura, es decir, hacia hipótesis

interpretativas, que presuponen actitudes hacia la investigación y la pregunta, que desemboca en múltiples preguntas” (Jurado Valencia, 2014).

Su principal característica es que nace de los intereses de los alumnos y sus dudas. Para poder realizar un proyecto dentro de la PPP, se debe preguntar a los alumnos indagando por sus intereses, de manera que el proyecto nazca de los propios deseos de los estudiantes. Una vez que se tiene claro el tema del proyecto y las preguntas que pretende responder, es deber del docente construir por su cuenta el esquema del proyecto y planear las sesiones de este tomando como base las preguntas que los estudiantes tienen en torno al tema escogido. Los proyectos de la PPP son integrales, al igual que los de la PDP, pero la decisión sobre las demás disciplinas, materias, asignaturas o áreas del conocimiento que atravesarán al proyecto se definen, de igual manera, una vez se empieza a construir el proyecto de acuerdo con las necesidades particulares del mismo y de las facilidades con las que se cuente en el lugar de la realización del proyecto. Al momento de determinar los contenidos del proyecto, las preguntas hechas por los alumnos son vitales, determinantes. Por su parte, al no tratarse de un proyecto único, diseñado desde cero con base en los intereses del docente, es más flexible y adaptable, de hecho, puede ir cambiando en el camino de acuerdo con cómo evolucionen, o bien el proyecto, o bien los alumnos. Por supuesto, también permite la integración de las TIC's en su elaboración, sin embargo, se debe tener en cuenta que la PPP se basa en el inductismo a través de una metodología mayéutica, lo que hace que no se dé prioridad a la exposición de conceptos y conocimientos por parte del docente, sino al descubrimiento de estos por parte de los alumnos. Dentro de la metodología que se emplea en la PPP, es usual hacer preguntas a los estudiantes que indaguen por sus conocimientos previos y que les permitan acceder al conocimiento por sí mismos y no a través de una fuente de conocimiento externa, elevada e inalcanzable en forma de docente. En efecto, dentro de las lógicas del PPP, el docente es un facilitador, un orientador, más no un director.

Por otro lado, cabe señalar que la PPP supone un alto costo en tiempo, recursos humanos y tecnológicos, además de que tiene un bajo grado de replicabilidad. A diferencia de la PDP, al tener que realizar un proyecto nuevo en cada curso, en cada grupo de alumnos, el tiempo necesario para su elaboración y desarrollo aumenta, además de que no se puede calcar y adaptar fácilmente de un grupo a otro. Sin embargo, esto no elimina la posibilidad de hacer proyectos en conjunto, permitiendo la semi-integración de dos cursos o grupos que sean cercanos en su progreso académico, por ejemplo, se podría plantear un proyecto con una misma temática y actividades similares para cuarto y quinto de primaria, pero tomando en cuenta que los Estándares Básicos de Competencias<sup>3</sup> varían para ambos cursos. Por lo tanto, las actividades que se crean en los proyectos se pueden adaptar y reciclar con relativa facilidad, pero la totalidad de un proyecto ya se vuelve algo

---

<sup>3</sup> Enlace de los Estándares Básicos de Competencias del Ministerio de Educación Nacional: <https://www.mineducacion.gov.co/portal/men/Publicaciones/Guias/116042:Estandares-Basicos-de-Competencias-en-Lenguaje-Matematicas-Ciencias-y-Ciudadanas>

---

más difícil de abarcar con la intención de replicarlo en otro grupo completamente diferente aun si coinciden en algunas características.

A su vez, es importante aclarar la diferencia entre un proyecto y, por ejemplo, un programa puesto que “Un proyecto está abierto, nunca se cierra; el programa al contrario tiene siempre una clausura” (Jurado Valencia, 2014). Sin embargo, esto no significa que un proyecto no tenga una secuencia de sesiones y, por ende, una sesión inicial, así como una sesión final. Lo que significa es que el trabajo de preguntarse, de hacer conjeturas y plantear hipótesis a las preguntas que van surgiendo a lo largo del proceso de formación que se da durante el paso por una institución educativa no llega a su fin con el proyecto en sí mismo, sino que se sigue desarrollando en ese camino, incluso permitiendo que se retomen elementos de un proyecto y a partir de ellos se genere un proyecto nuevo completamente diferente, pero relacionado con el anterior. También es importante señalar que la evaluación que se da en los proyectos, indistintamente de que sea PDP o PPP, es una evaluación de tipo formativa en la que se hace un balance entre lo que se ha aprendido y lo que falta por aprender. Un estudiante que trabaja mediante proyectos no tiene solo una serie de números que indican su posición en una competencia desenfadada e inmisericorde por ocupar una posición dentro de un podio de excelencia académica asumida, sino que recibe una retroalimentación más profunda y específica, personalizada dependiendo del caso. Esto favorece la significación del aprendizaje, lo que le permite al estudiante aprender de mejor manera los contenidos y apropiarlos.

Es por eso por lo que ambas pedagogías sirven como precedentes valiosos a la manera en la que se podría abordar la LC dentro de las aulas universitarias. El carácter progresivo de los proyectos permite ir desde conceptos generales y sencillos a conceptos más específicos y complejos de acuerdo con las necesidades del tema a tratar. En el caso de la LC, lo que primero se necesitaría en una asignatura orientada a la LC sería una nivelación de conocimientos acerca de las herramientas y el pensamiento computacionales, muy importantes a la hora de enfrentarse a los lenguajes de programación, los cuales se usan ampliamente dentro de la LC para sus diferentes aplicaciones. Después de esa nivelación, sería necesario ir profundizando en la programación y en el uso de los algoritmos para el procesamiento del lenguaje natural, de donde desprenden las demás aplicaciones de la LC como pueden ser la traducción automática, la recuperación de información o el análisis morfosintáctico de corpora. Para este fin cualquiera de las dos pedagogías podría ser útil, bien sea una PDP ya establecida donde los estudiantes tengan todas las herramientas necesarias para llevar a cabo un simple programa de análisis textual que permita procesar textos cortos, y que dicho trabajo constituya el proyecto en cuestión, o bien sea un PPP en el que los estudiantes decidan qué quieren hacer y con base en eso se adapte el proyecto a sus necesidades, enmarcado en la LC y las competencias que se necesita que los alumnos desarrollen. Es por eso por lo que es importante hacer un repaso por estas pedagogías, puesto que se consideran de gran relevancia para este fin.

## 3.2 El uso de las TIC en la educación

Siguiendo con la educación, también se considera importante hacer mención del papel que cumplen las TIC's en el panorama actual de la educación, en especial teniendo en cuenta que el tema general de este trabajo es lingüística computacional y, por ende, implica el uso de herramientas computacionales tanto en la formación como en la praxis. Con relación a las TIC, se puede tomar diferentes enfoques, de entre los cuales elegiremos tres. En primer lugar, las TIC pueden ser adoptadas desde un punto de vista instrumental en el cual la metodología didáctica del docente en su práctica pedagógica no se ve afectada, es decir, el profesor seguirá trabajando exactamente igual y su uso de las TIC será poco, este uso puede verse limitado a la proyección de material audiovisual, por poner un ejemplo. En segundo lugar, pueden ser vistas como una herramienta vital dentro del modelo educativo en cuestión, requiriendo de la implementación de plataformas virtuales para la enseñanza y de otro tipo de recursos digitales formativos. Y, por último, se puede ver como el centro de la enseñanza, en este caso las TIC se tratan como el centro de la formación y los demás conocimientos se articulan tomando en cuenta las tendencias digitales del momento, como puede llegar a pasar con algunas plataformas de formación virtual que se orientan a la formación en línea abordando distintas áreas del conocimiento, pero en las cuales se plantea la vida laboral ligada fuertemente al trabajo virtual.

Para poder abordar el uso de TIC's en la educación, se considera valiosa la distinción que Rueda Vera et al. hacen con relación a la *educación mediada por TIC's* y la *educación virtual*. Esto es importante porque el papel de las TIC's en la educación no puede verse reducido a utilizar computadores u otros dispositivos electrónicos en los espacios de formación o utilizar de vez en cuando plataformas virtuales para asignar trabajos y calificar, por poner dos ejemplos. Entonces, es necesario llegar a un acuerdo sobre cuál debería ser el papel de las TIC's en la educación y, sobre todo, en la educación superior. Para empezar, los autores señalan que:

“...la educación mediada por TIC, se fundamenta en el encuentro sincrónico con el docente desde plataformas especializadas (por ejemplo, Teams y Zoom), las cuales permiten la transmisión de información y contenidos, grabación de las clases, descarga de asistencias, entre otras acciones” (Islas Torres, 2017).

En este orden de ideas, se parte de la concepción de las TIC como herramientas facilitadoras que ayudan a agilizar o automatizar procesos que normalmente suelen suponer un gasto de tiempo considerable para los docentes, como lo pueden ser las listas de asistencia, las entregas de notas, la revisión de plagio, entre otros. Con esto se busca reducir la carga que esas actividades le derivan al docente, esenciales dentro de la visión macro de su labor educativa, pero que ocupan tiempo que el maestro puede dedicar a labores más específicas como lo pueden ser la retroalimentación personalizada a los estudiantes, las tutorías personalizadas, la creación de material didáctico o de actividades, etc. El uso de las TIC se podría considerar misceláneo, en cierta medida, lo cual no tiene por qué ser algo negativo ya que se ajusta al modelo en cuestión. Sin embargo, obvia

---

factores como la accesibilidad a los recursos digitales, la alfabetización digital de los estudiantes, los inconvenientes de conectividad, etc.

Aun así, para los estudiantes también resulta de gran utilidad en este tipo de formación el uso de las TIC ya que les permite tener un entorno de aprendizaje virtual más organizado. Esto es porque dentro de estas mecánicas se pueden implementar espacios virtuales para compartir información, fuentes, trabajos, opiniones, etc. Lo anterior es de gran ayuda para los alumnos ya que significa que pueden tener acceso a los diferentes recursos en cualquier momento, aunque las sesiones sean sincrónicas, ya que por lo general los contenidos quedan grabados en registros audiovisuales o de texto. También les permite a los alumnos llevar registros más precisos de sus avances y sus progresos académicos en la forma de la calificación. Sin embargo, también es cierto que algunos estudiantes pueden sentirse agobiados por una sobrecarga de recursos y llegar a frustrarse en su esfuerzo académico, de igual manera es cierto que la sincronía es solo en cuanto a las interacciones con los profesores durante clases o tutorías, razón por la cual en algunos momentos los estudiantes se pueden sentir abandonados frente a máquinas y sistemas que ofrecen una retroalimentación escueta.

Ahora bien, Introduciremos el concepto de educación virtual de forma breve. Este se diferencia de la educación mediada por TIC ya que parte del tercer enfoque mencionado anteriormente, un enfoque que se fundamenta en la virtualidad y en modelos pensados para lo virtual. Se busca la integración de las TIC en todo el proceso de formación e, incluso, en su ejercicio profesional futuro. Se puede decir que:

“En la educación virtual no se requiere de manera particular el encuentro con el docente y se cuentan con recursos tecnológicos que facilitan la organización de la enseñanza y los medios educativos y didácticos para que el estudiante de manera autónoma lleve adelante su proceso formativo” (Islas Torres, 2017).

De este enunciado podemos deducir que la educación virtual se concibe como un modelo complejo en el que gran variedad de recursos se articulan y se disponen para el estudiante con el fin de que pueda gestionar su aprendizaje de manera autónoma. La autonomía es un pilar fundamental de la educación virtual, al igual que en la PBP o la PPP, y es uno de los mayores retos con los que se encuentra el alumno. No obstante, la autonomía es un reto en cualquier modalidad de educación superior, por lo que ¿qué hace que la educación virtual sea especial? Su particularidad radica en el ecosistema que se crea como entorno educativo. Ofrecer educación virtual supone una inversión considerable en dinero, tiempo y capital humano para poder construir la infraestructura necesaria que sostenga en el tiempo los recursos materiales y digitales que componen el modelo educativo en cuestión. Esta infraestructura, cuando está bien implementada, permite el desarrollo de plataformas virtuales para que el alumnado y los maestros puedan ingresar a desarrollar labores propias de sus roles. Por lo anterior, para poder alojar y alimentar estas plataformas se requiere de la implementación de diferentes repositorios virtuales con material textual, audiovisual, gráfico, etc., que sirva de insumo para los usuarios. Asimismo, exige la capacitación constante del personal docente y administrativo con el fin de que

tengan dominio de las plataformas y de las didácticas que se deben emplear en dicha modalidad. Por su parte, también requiere de inversión en el área de sistemas puesto que es necesario tener un soporte que pueda responder a las necesidades de los usuarios de forma eficiente. Esto quiere decir que se debe implementar un departamento de soporte para atender las necesidades de docentes, alumnos y administrativos, ya sea de forma presencial o remota. Además, es innegable que, en la actualidad, la educación virtual se encuentra en auge con el establecimiento de plataformas como Coursera<sup>4</sup>, Udemy<sup>5</sup>, Domestika<sup>6</sup>, Open English<sup>7</sup>, el SENA Virtual<sup>8</sup>, entre otras. Este modelo educativo es válido y aplicable para formación complementaria, cursos; y también para formación de nivel superior tanto en pregrado como en posgrado.

No obstante, cuando se trata de educación virtual en grados superiores de estudios, se puede tender a generalizar el nivel de dominio informático-computacional que los alumnos tienen al ingresar a un pregrado o un posgrado. Se asume, por ejemplo, que las nuevas generaciones que han sido criadas en una era digital tienen dominio innato de la tecnología, lo cual no es necesariamente así.

“...la promesa de las tecnologías informáticas en red dista mucho de haberse cumplido con la amplitud prevista. Lo anterior, debido a que entre las habilidades tecnológicas de los llamados nativos digitales se interpone la práctica educativa obsoleta, los principios organizativos, las culturas obsoletas de la enseñanza y las dificultades inherentes a la creación de nuevos aprendizajes” (Islas Torres, 2017).

La implementación de TIC, y en general el uso de tecnología, no garantiza una solución mágica a todos los problemas. De hecho, para la educación virtual no es ideal que el dominio del estudiante se vea limitado a prender un computador e ingresar al navegador para consumir material audiovisual con fines de entretenimiento, usos de la tecnología que pueden limitarse a los dispositivos móviles como celulares, tabletas, computadores portátiles, computadores portátiles de bolsillo, consolas de videojuegos, consolas de videojuegos portátiles, etc. Un conocimiento, una competencia académica y laboral, que para generaciones anteriores al 2010 puede parecer lógico, como aprender a manejar Office y, en la actualidad, aprender a manejar programas para la edición de imágenes, audio y video, son habilidades que deberían ser adquiridas en la educación media. Sin embargo, el dominio que los estudiantes tienen de estas habilidades puede ser deficiente en algunos casos al ingresar a la universidad. Por lo anterior, no conviene simplemente enfrentar a los alumnos, por lo menos en los primeros niveles, con las herramientas y plataformas tecnológicas sin una nivelación previa. Los cursos nivelatorios orientados a la ofimática son una buena opción, por ejemplo. Pero esto es con el objetivo

---

<sup>4</sup> <https://www.coursera.org/?msocid=3fee4133bb40662e066354dfbaef6786>

<sup>5</sup> Enlace recortado: <https://n9.cl/r8c0m>

<sup>6</sup> <https://www.domestika.org/es>

<sup>7</sup> Enlace recortado: <https://n9.cl/lpxnr3>

<sup>8</sup> Enlace de carreras virtuales: <https://n9.cl/aiot7>, y enlace de cursos cortos: <https://n9.cl/bsypa>

de gestar un ambiente en el que haya una cierta igualdad en cuanto a la capacidad de los alumnos para aprovechar el entorno de formación virtual ya que, por sí mismos, cada estudiante va adquiriendo habilidades y destrezas digitales de acuerdo con sus intereses.

### 3.3 Lingüística Computacional

La Lingüística Computacional es una disciplina, o subdisciplina dependiendo del autor, de reciente creación. Sin embargo, su evolución y su desarrollo desde su surgimiento en los años 50 ha sido considerable y ha seguido un camino relativamente claro hasta la actualidad. Para poder tratar el tema del surgimiento de la LC, es importante entender algo que puede parecer obvio en un primer lugar, la historia de la LC va de la mano con la de la computación y la informática. Aun así, aunque la LC cuenta con alrededor de setenta años a sus espaldas, no hay acuerdo sobre una definición única de esta disciplina, lo cual genera problemas a la hora de abordarla.

Para empezar, se puede entender la Lingüística Computacional como una subdisciplina de la lingüística que se encarga del estudio del *lenguaje natural (NL)* y la creación de modelos computacionales para comprender y generar LN. Es decir, el núcleo a partir del cual nace la LC son las computadoras y el lenguaje natural, siendo el reto de los lingüistas computacionales hacer posible que las computadoras entiendan a las personas en LN. Por lo tanto, "...la LC trata de la construcción de sistemas informáticos que procesen estructura lingüística y cuyo objetivo sea la simulación parcial de la capacidad lingüística de los hablantes de una lengua, independientemente de su carácter comercial o de investigación básica" (Moreno Sandoval, 1998). Es importante tener claridad sobre a qué se refiere la *simulación de la capacidad lingüística humana*, puesto que la capacidad lingüística en esencia es capacidad cognitiva. No se debe incurrir en el error de tratar el cerebro humano como el procesador de una computadora, como si fueran lo mismo, dado que por más que se use esta analogía, se hace solo con fines prácticos. Tampoco se debe caer en la ilusión de inteligencia que provee la *Inteligencia Artificial (IA)*.

"... 'simular' no implica 'equivaler'. Es decir, la mente humana y el ordenador no son lo mismo. Aunque la mente humana y el ordenador se caracterizan por codificar, retener y operar con símbolos y representaciones internas, cabe advertir que el soporte físico o el hardware no es el mismo" (Yllescas, 2011).

La simulación, justamente, se refiere a imitar los resultados o las tareas que lleva a cabo el cerebro, pero partiendo del hecho de que no se comprende en su totalidad cómo funciona el cerebro humano y cómo entendemos, aprendemos o comprendemos. A través de sus propias formas de interpretar la información y de trabajar con ella, el computador busca llegar a resultados similares a los que obtienen los seres humanos. Sin embargo, el soporte físico es una gran limitante en este caso. El hecho de que para un computador la información se clasifique con relación a su peso, la cantidad de datos que contiene, mientras que los humanos la clasifiquemos a partir de relaciones con referentes anteriores, ya representa una barrera que le es difícil de superar a la computadora. No es el propósito

de este trabajo abordar la capacidad de memoria o de procesamiento neta del cerebro humano, puesto que ya se ha advertido que es algo que no se tiene claro en la actualidad y que escapa a nuestro tema de investigación, pero es necesario hacer la salvedad. Continuando con los modelos, para poder construir esos sistemas de procesamiento de LN, la tarea que deben emprender los lingüistas computacionales es la de "...reflejar la organización y funcionamiento de las estructuras y procesos lingüísticos, por una parte, y de las estructuras y procesos cognitivos, por otra" (Moreno Sandoval, 1998).

Otra de las definiciones posibles para la LC la aborda como una subdisciplina de la *Inteligencia Artificial (IA)*. "Esta Lingüística Computacional se entiende como una de las áreas de la Inteligencia Artificial, cuyo objetivo es el tratamiento del lenguaje como el instrumento fundamental de la comunicación humana, directamente ligado a la cognición" (González Fernández, 2018). Otros autores, como Gómez Guinovart también la han abordado bajo este panorama:

"Desde el punto de vista de su conexión con la informática, y también por razones históricas, la lingüística computacional se considera una subdisciplina de la *inteligencia artificial*, una especialidad de la informática que se ocupa de la comprensión de la inteligencia y del diseño de máquinas inteligentes, es decir, de máquinas y programas que presentan características asociadas a la comprensión humana, como el razonamiento, la comprensión del lenguaje hablado y escrito, el aprendizaje o la toma de decisiones" (2000).

No obstante, para poder abordar el lenguaje natural con programas y modelos computacionales, es necesario el trabajo lingüístico (gramático) tradicional primero. Es decir, se debe trabajar con los distintos niveles de análisis de la lingüística para poder establecer categorías y clasificar los diferentes elementos constitutivos de la lengua que se requieran para la investigación a realizar. La materia prima que se utiliza para trabajar en LC son los corpora, los cuales son grandes grupos de datos, ya sean en bruto o preprocesados, y que son empleados para construir los modelos o entrenarlos, en el caso de ser modelos generativos o de IA. Es por eso por lo que es importante hacer mención, brevemente de la *Lingüística de Corpus*, ya que como "La Lingüística de Corpus está... íntimamente ligada a la Lingüística Computacional, pues aporta grandes cantidades de ejemplos reales del lenguaje almacenados en bases de datos en forma de corpus de manera sistemática" (González Fernández, 2018). En algunos casos se confunden la Lingüística de corpus con la LC, como si fueran una misma o indivisibles, inseparables. Si bien es cierto que, gracias a la Lingüística de Corpus, se puede contar con grandes corpora especializados, lo cual facilita la elaboración de los modelos y los programas computacionales, esto debe servir de ejemplo al trabajo interdisciplinar que se puede llegar a realizar dentro del marco de la LC.

Ya hemos presentado algunos conceptos importantes a la hora de hablar acerca de LC, lo que sigue es presentar las formas en las que se puede abordar esta subdisciplina. Nuevamente, Gómez Guinovart nos presenta tres enfoques, principalmente: informática aplicada a la investigación lingüística, implementación de teorías lingüística, y aplicaciones

lingüísticas de la informática. De estas tres aproximaciones, González Adela nos aporta las siguientes claridades: el primero, incluye los etiquetados morfológicos y sintácticos; el segundo, incluye la elaboración, descripción y comprobación de teorías lingüísticas; y por último el tercero, enfocado en el *Procesamiento del Lenguaje Natural (PLN)*, incluye tecnologías del habla (reconocimiento y síntesis), traducción automática y extracción o recuperación de información.

A su vez, esta distinción nos permite abordar las diferentes áreas de investigación de la LC, las cuales presentaremos siguiendo a Domínguez: En primer lugar, el **Etiquetado Morfológico** brinda información acerca de la morfología de la palabra y es al que gran parte de los libros sobre LC dedican sus primeras secciones. Para esta labor se ha diseñado un gran número de algoritmos con el fin de etiquetar las palabras y aislar sus componentes constitutivos nucleares. En segundo lugar, el **Etiquetado sintáctico**, que, permite abordar oraciones complejas, y, por ende, párrafos completos posteriormente. En este proceso intervienen componentes más semánticos, puesto que a partir de esas raíces se infieren significados. El resultado del etiquetado morfológico puede ser *subpalabras*, y estas subpalabras son fragmentos que pueden o no tener significado, como por ejemplo una raíz o un afijo, etc. A partir de estos pequeños elementos, se infieren significados en el etiquetado sintáctico.

En tercer lugar, se encuentran las **Técnicas de reconocimiento de voz y conversión de texto a voz**. En las primeras se utilizan técnicas probabilísticas para transcribir el audio en datos que el computador pueda interpretar; en las segundas, se busca generar sonidos que simulen el habla humana a partir de un texto. Como ejemplo práctico no académico, pero que se considera útil, se recomienda la visualización de videos en los que por medio de IA se creen asistentes de voz o bots que sirvan de interfaz para brindar información al usuario, como la que presenta el youtuber Nate Gentile en el video en el que crea un computador para el streamer Illojuan<sup>9</sup> y como parte del concepto crea un bot de IA que funciona por reconocimiento de voz y responde generando voz a partir de un texto plano<sup>10</sup>. En cuarto lugar, la **Recuperación inteligente de información**, "...este campo incluye todos los sistemas automáticos de obtención y análisis de información para su utilización posterior por los usuarios. Estos sistemas son los que se utilizan en los corpus actuales y en los buscadores de internet" (2018). En quinto lugar, los **Sistemas de diálogo y sistemas expertos**, que "...consisten en la transmisión de información entre los usuarios y el ordenador, gracias al almacenamiento digital previo del conocimiento de expertos en un área determinada" (2018).

---

<sup>9</sup> Enlace del video: <https://www.youtube.com/watch?v=U1VamJIZqe4>

<sup>10</sup> El texto plano hace referencia a texto sin formato, por ejemplo, el texto que se utiliza en las terminales de programación o el texto del bloc de notas. El formato de archivo más conocido que contiene este tipo de texto es el .txt.

Finalmente, la **Traducción automática**, la cual es uno de los campos en los que más se trabaja dentro de la LC y es la que dio lugar al nacimiento de esta subdisciplina en primer lugar. Busca traducir dos o más lenguas en lenguaje natural a través de un programa de computador. Esta traducción puede ser completamente automática o asistida. Todavía dista mucho de ser perfecta, pero ha experimentado grandes avances en los últimos años. Google Translate<sup>11</sup> es un buen ejemplo de traductor automático.

Adicionalmente, es necesario mencionar que la LC ha desarrollado gran número de programas para la investigación lingüística. Dentro de esta línea se encuentran los etiquetadores morfológicos y sintácticos, los cuales son muy útiles para los académicos, tanto estudiantes como profesionales, quienes los usan en sus investigaciones lingüísticas para procesamiento de textos en distintos formatos. También, las herramientas de LC sirven para ayudar al investigador a reducir la inversión económica y de tiempo, estos programas hacen que el proceso de recopilación y sistematización de la información sea un proceso más eficiente, lo que le permite al investigador dedicar más tiempo al análisis de la información y de los resultados. Es importante tener en cuenta que la labor de realizar el análisis lingüístico final de los datos le corresponde al lingüista en cuestión, es decir, a un ser humano. Por otro lado, dentro de la investigación realizada con las herramientas de la LC también se hacen presentes voces discordantes que señalan que la anotación que se realiza con los etiquetadores contamina los datos y los pervierte, lo cual hace que el posterior análisis lingüístico sea erróneo al trabajar con una versión corrupta de los datos originales. Estos investigadores abogan por trabajar con los corpora en bruto o con muy poca marcación, para que no se pierdan los rasgos distintivos del texto. Por supuesto, este último enfoque no está completamente equivocado, sin embargo, es una forma de abordar el problema desde una perspectiva ligeramente desviada y sesgada, puesto que para trabajar con los corpora se pueden modificar los algoritmos o programas de acuerdo con las necesidades investigativas del momento, o incluso crear programas nuevos ya sea desde cero o a partir de otros ya existentes. Estas decisiones van a ser determinadas por la propia investigación y, por supuesto, la formación del lingüista en LC.

En relación con los usos de la LC, es inevitable traer a colación el *Procesamiento de Lenguaje Natural (PLN)*, el cual puede ser visto como una de las aplicaciones de la LC o como un campo de investigación análogo, pero separado. Por su parte, hay autores que lo plantean como parte de la LC y otros que lo plantean como parte de la IA al igual que lo hacen con la LC, a la cual también hay autores que la señalan como una subdisciplina de la IA. Sin embargo, el PLN coincide con la LC en sus aplicaciones o en su objeto de estudio, el lenguaje natural:

“El PLN consiste en la utilización de un lenguaje natural para comunicarnos con una computadora, debiendo ésta entender las oraciones que le sean proporcionadas, el uso de estos lenguajes naturales, facilita el desarrollo de

---

<sup>11</sup> Enlace: <https://translate.google.com/>

programas que realicen tareas relacionadas con el lenguaje o bien, desarrollar modelos que ayuden a comprender los mecanismos humanos relacionados con el lenguaje” (Cortes Vásquez et al., 2009).

De cierta manera, se puede ver al PLN como una disciplina hermana de la LC que le provee de soporte para la creación de los modelos de lenguaje y los programas para el abordaje del lenguaje natural. Si bien ambas pueden ser similares, la LC tiene un componente más teórico, mientras que el PLN tiene un enfoque o una aproximación mucho más práctica, e incluso comercial. Las investigaciones y los desarrollos en PLN suelen estar dirigidos a tecnologías que pueden tener un aprovechamiento comercial efectivo. Los chats bots, los asistentes virtuales y las tecnologías de inteligencia artificial que unifican esos desarrollos en el reconocimiento del lenguaje natural, su procesamiento y la producción de lenguaje natural son buenos ejemplos de aplicaciones en productos y servicios comerciales del PLN. Aunque estos desarrollos pueden parecer todavía un poco primitivos en algunos aspectos o, más bien, poco naturales, los avances son mucho más visibles y tangibles en la actualidad, por lo que estas tecnologías son prometedoras para múltiples áreas y aspectos de la vida humana.

No obstante, los desarrollos de estas tecnologías suelen estar orientados a las lenguas hegemónicas, las lenguas mayoritarias e institucionales, como es el caso de Colombia con el español. Esto implica que lenguas menos propagadas (habladas) y con menos soporte institucional ven sus desarrollos en materia de LC limitados; el acercamiento que la LC y el PLN han realizado a esas lenguas es menor y, por lo tanto, no cuentan con las mismas herramientas a su disposición. Cuando se hace referencia a la disponibilidad de las herramientas no solo se refiere al idioma en el que está escrita la interfaz de un programa en específico, sino a las lenguas que se tienen como objeto de estudio al momento de crear el programa. Es ahí donde se encuentra uno de los mayores retos, y por ende oportunidades, para los lingüistas computacionales. Tomando en cuenta que ya se cuenta con tecnologías, modelos y corpora de tamaños considerables en lenguas mayoritarias si de verdad se quiere poder llevar a cabo un estudio más complejo y profundo del lenguaje natural que utilizan los seres humanos para comunicarse, dicho estudio no puede limitarse a lenguas mayoritarias. Es decir que no todas las lenguas del mundo trabajan bajo las mismas lógicas ni funcionan de la misma manera; por supuesto, todas las lenguas comunican, pero los mecanismos con los que lo hacen pueden variar considerablemente dependiendo de la lengua en cuestión.

En resumen, se podría decir que el PLN concentra todas esas aplicaciones, aparentemente separadas, de la LC que se vieron anteriormente en este apartado y las emplea para abordar el lenguaje natural de forma más integral. Por supuesto, no se trata solo de aplicaciones comerciales, su uso en el mundo académico es de gran relevancia, puesto que es allí donde primero nacen estos avances tecnológicos de los que la comercialización es una etapa posterior. La LC es una disciplina importante para el progreso de la humanidad y para la investigación científica porque acerca la tecnología a los humanos en su cotidianidad. Su relación cercana con la inteligencia artificial también

supone que quienes se dediquen a su estudio pueden ver ampliado su campo laboral. No se puede ignorar su importancia para el uso de internet y de las nuevas tecnologías que las personas tienen a su alcance en la actualidad. A medida que la inteligencia artificial avanza y se integra más con los sistemas que forman parte de la vida diaria de las personas, nuevos problemas irán surgiendo, nuevos efectos inesperados aparecerán en las interacciones entre los usuarios y las interfaces. El lenguaje natural, al ser tan intrínsecamente humano, no es del todo predecible ni modelable, es susceptible de ser abordado utilizando la computación, pero no es completamente computarizable. Se pueden crear modelos y programar sistemas, pero siempre será necesaria la intervención de lingüistas bien preparados para retroalimentar dichos sistemas, para responder a las necesidades del momento, para abordar desde diferentes perspectivas los retos de la era digital con relación al lenguaje, y, por ende, con relación a sus usuarios, las personas.

### 3.4 Pedagogía en la Lingüística Computacional

Introducimos un apartado nuclear para este trabajo, pero del cual no se ha conseguido literatura numerosa. Aun así, lo que se ha logrado reunir ha servido para orientar sobre este tema, la pedagogía en la LC. El concepto de pedagogía es amplio, sin embargo, su elaboración teórica casi siempre gira en torno a la enseñanza, lo cual de por sí no ayuda mucho dado que discutir sobre *qué es la enseñanza* supondría una infinidad de textos, seminarios, cursos, etc., que desbordaría el alcance de este trabajo. Con el fin de establecer un punto de partida se puede decir que la pedagogía se preocupa por el qué, el cómo, el dónde, el cuándo, de la enseñanza dentro la educación como sistema para la formación de los ciudadanos.

“Si se trata de concretar un nivel de epistemologización que permita concebir la pedagogía como disciplina, y aun como disciplina científica, habría que precisar su objeto; en esta propuesta, el objeto más amplio sería el proceso educativo en toda su complejidad, y el objeto mayor o articulador sería la enseñanza” (Vasco Uribe et al., 2008).

Este proceso educativo se da en un nivel formal generalmente institucionalizado. El sistema educativo en Colombia, a grandes rasgos, consta de educación primaria, secundaria, técnica, tecnológica, universitaria y de posgrado. Esta educación formal en Colombia también incluye tanto a instituciones del sector público como del privado y está regulada por el Ministerio de Educación Nacional, cuya misión es:

“Liderar la formulación, implementación y evaluación de políticas públicas educativas, para cerrar las brechas que existen en la garantía del derecho a la educación, y en la prestación de un servicio educativo con calidad, esto en el marco de la atención integral que reconoce e integra la diferencia, los territorios y sus contextos, para permitir trayectorias educativas completas que impulsan el desarrollo integral de los individuos y la sociedad” (*Portal MEN - Presentación - Ministerio*, 2025).

---

Por supuesto, no existe solo la educación formal, sin embargo, es esta la que nos compete en este trabajo. La educación formal gira en torno a la especialización, a veces llegando al absurdo de que dos especialistas dentro de una misma área del conocimiento no se entiendan entre ellos, aunque en esencia ambos sean profesionales del mismo sector o con la misma formación de base. Si bien la lingüística no es ajena a esta problemática, los límites entre sus subdisciplinas son menos claros. Aunque los niveles de análisis esenciales de la lingüística están establecidos como disciplinas independientes, se pueden hallar puntos de encuentro entre ellas como lo puede ser la morfosintaxis, o incluso con otras disciplinas como puede ser la sociolingüística o la lingüística computacional, tema de interés para este trabajo.

Esta especialización se da principalmente en la educación superior, en especial en la formación de nivel de posgrado, la cual incluye, por lo menos en Colombia, las especializaciones, maestrías y doctorados, por mencionar las principales y más importantes. No es usual ver tanta especialización en los niveles de pregrado. Sin embargo, contenidos transversales sí que pueden ser incluidos en los currículos de diferentes programas académicos pertenecientes a una misma facultad o área del conocimiento. Esto también ocurre en el caso de las ciencias del lenguaje, dentro de las cuales incluimos la lingüística, la filología, las lenguas, e incluso, licenciaturas en lenguas. Por otro lado, la participación de la LC en la educación superior no se limita a su inclusión dentro de programas curriculares de pregrado o de posgrado; también incluye el uso de sus herramientas en los ambientes de formación, ya sea como herramientas de estudio, herramientas para investigar o como herramientas para evaluar. Esto quiere decir que no solo nos referiremos a cursos orientados a la LC, sino a cómo se utilizan sus herramientas y desarrollos dentro de la educación universitaria en las carreras de ciencias del lenguaje.

Primero que todo, abordaremos la inclusión de la LC como programas curriculares propiamente dichos ya sea en niveles de pregrado o posgrado. En cuanto a pregrados de Lingüística Computacional en Colombia no existen, esto puede tener sentido si tomamos en cuenta la naturaleza de los pregrados. La formación que se suele dar en pregrado es una educación general y holística que busca que el estudiante tenga las bases fundamentales teóricas y epistemológicas de su disciplina en cuestión a fin de que posteriormente pueda decantarse por una especialidad dentro de su misma disciplina o relacionada con otra disciplina cercana y así construir su propio camino profesional. Bajo esta lógica, tiene más sentido la existencia de un pregrado en lingüística, o incluso en filología, que uno en lingüística computacional. Adicionalmente, para la integración de la LC en pregrados es necesario primero nivelar los conocimientos que los estudiantes tienen acerca de la computación. Esto implica que, a la vez que se les enseñan las bases de la lingüística, en sus primeros cursos habría que incluir en estos niveles iniciales de la carrera cursos sobre pensamiento lógico o computacional, informática básica o programación, de manera que los estudiantes se familiaricen con los sistemas que van a utilizar en la LC y sus interfaces. Esta formación básica busca que, una vez llegado el momento de interactuar con las herramientas, pero, sobre todo, con la teoría propia de la LC, no se encuentren con una barrera demasiado grande. Lo importante es "...to stratify courses

according to the backgrounds different populations of students have with respect to programming and formal thinking” (Baldrige & Erk, 2008). Por supuesto, no es posible diseñar un curso específico para cada estudiante y tampoco es el objetivo, pero la nivelación de conocimientos es esencial en los primeros semestres de cualquier carrera universitaria, razón por la cual es recomendable introducir cursos con contenidos transversales a la LC en estos niveles.

Por su parte, en el nivel de posgrado se pueden encontrar maestrías o especializaciones en LC, e incluso en *Humanidades Digitales (HD)*, otra rama del conocimiento más amplia incluso que la propia lingüística, que recoge las teorías, herramientas y aplicaciones tecnológicas en las diferentes áreas del conocimiento en las humanidades.

“Ubicadas en la intersección de disciplinas, las HD usan los medios digitales y la computación para producir, socializar y representar conocimiento en las Humanidades (Historia del Arte, Literatura, Historia, Filosofía, etc.) involucrando a disciplinas como el diseño, la comunicación, los medios y la programación” (Llach et al., 2020).

Por lo tanto, bajo ciertas perspectivas, la LC podría hacer parte de las HD. Sin embargo, la presencia de la LC en posgrado plantea retos en todos los niveles: económico, teórico, práctico, educativo, etc. Al igual que ocurre en el pregrado, la disparidad en saberes y destrezas de los alumnos al momento de ingresar a posgrados orientados a la LC o las HD suponen un gran reto para los coordinadores y directores de estos programas, ya que deben incluir dentro de los programas curriculares cursos destinados a dicha nivelación o, por lo menos, dentro de los cursos ya existentes espacios para realizar dicha nivelación. Como se mencionó antes, la elección de qué posgrado estudiar se hace tomando en cuenta los intereses profesionales que cada uno tiene, esto quiere decir que el alumnado puede ser heterogéneo, no todos van a tener la misma formación oficial de base, esto quiere decir que en un posgrado en lingüística computacional bien podría haber lingüistas, ingenieros, historiadores, sociólogos, filósofos, músicos, etc. Aun así, “En el peor de los casos, es más bien homogéneo, pero procedente en su inmensa mayoría de planes de estudio concebidos y/o impartidos con un escaso o nulo manejo de herramientas y técnicas informáticas” (Pareja Lora, 2024).

La problemática anteriormente vista en el pregrado se traslada a estos entornos educativos de nivel superior. Al ser tan variada la población que puede inscribirse a un posgrado es difícil asumir cuánto saben o no sobre el tema a trabajar. Por supuesto, se asume que tienen cierto dominio de la materia, ya que van a ingresar a un posgrado y en ese nivel el propósito es construir conocimiento nuevo, no solo aprender lo ya existente. Sin embargo, como lo menciona Pareja Lora, la forma en que los contenidos transversales o propios de la LC haya sido abordada en el pregrado también es determinante para una aproximación óptima del alumno en niveles superiores de estudio. La LC no puede ser vista como una herramienta más a ser usada en los modelos tradicionales de enseñanza o como una mera curiosidad. Lo ideal es que se aprovechen sus herramientas y sus

---

técnicas para abordar desde otros enfoques los contenidos. Realizar análisis morfológico y sintáctico con las herramientas propias de la LC proporcionará una base desde la cual seguir profundizando o enfrentarse a una formación más avanzada en esta área.

Siguiendo con la línea de ciencias del lenguaje es importante mencionar otra de las formas en que la LC puede ser ofertada a los estudiantes durante su trascurso por la universidad, los talleres, seminarios, semilleros, grupos de estudio, etc. Remitiendo puntualmente a los talleres, ya sean intra o intersemestrales permiten construir unidades temáticas específicas sobre las cuales trabajar sin sacrificar cursos básicos ya establecidos. No es extraño que en una asignatura regular de una universidad cualquiera queden temas sin ver o que se aborden de forma escueta. Lo anterior atiende principalmente a cuestiones de tiempo y a la ya mencionada diversidad entre los estudiantes. No todos los estudiantes aprenden de la misma manera ni a la misma velocidad, este es un reto constante y que en ocasiones parece ser omitido o menospreciado pero que es vital para la labor del docente en el aula. Se considera importante desarrollar investigaciones futuras sobre cuál es realmente el porcentaje de contenidos de los cursos que se está viendo de forma apropiada, dado que es un tema que desborda a la actual investigación.

Volviendo al tema de este trabajo, se pueden plantear talleres orientados a la formación en LC y PLN que buscan que los alumnos den sus primeros pasos en estas disciplinas sin esperar un dominio avanzado de las mismas al finalizar. Esto no significa que el aprendizaje sea escaso ni mucho menos. Sin embargo, en talleres como estos se puede permitir incluso el ingreso de estudiantes de otras carreras, lo cual contribuye al trabajo interdisciplinar y la integración entre diferentes bagajes teóricos o formas de pensar. La discusión y el trabajo en equipo son una de las mejores formas de las que disponen los académicos para construir conocimiento. Es por eso por lo que es importante fomentar la construcción de ese tipo de espacios en las universidades, para poder poner en diálogo los distintos intereses de los estudiantes y permitirles integrarlos en su formación profesional.

Por otro lado, dentro de los cursos diseñados con la LC en mente *Aprendo con NOOJ: de la Lingüística Computacional a la enseñanza de la lengua* se erige como una metodología especialmente pensada con este propósito en mente, enseñar LC. Esta metodología está orientada al trabajo textual de la mano con las herramientas de la LC. NOOJ es en sí mismo un programa dedicado al análisis textual y es por eso por lo que los alumnos desarrollan actividades sencillas con las cuales producen textos que después procesan con el programa y a través de las constantes actividades, de la praxis, van aprendiendo y dominando las herramientas digitales. Esta metodología toma en cuenta que “Las nuevas tecnologías por sí solas no cambian los métodos tradicionales basados en la absorción pasiva de contenidos, el currículum único y cerrado, la memorización y el aprendizaje fuera de contexto” (Ferramondo, 2021). Por eso busca capacitar al docente en su uso, de manera que este pueda innovar en el aula. Este tipo de metodologías son realmente útiles para poder llevar la LC a las aulas universitarias de los pregrados en

---

ciencias del lenguaje ya que, por ejemplo, en un pregrado como el de Español y Filología Clásica de la Universidad Nacional de Colombia en el que se aprenden lenguas clásicas, latín y griego, se podría utilizar para la enseñanza de dichas lenguas. Por supuesto, esto representaría un reto completamente nuevo en materia de programación, de adecuación del software a esas lenguas, pero podría hacerse. Esta metodología se considera valiosa dado que:

“Se trata de un nuevo enfoque para el estudio de la lengua que opera no con la memorización de reglas gramaticales sino con ensayos, con pruebas y tareas de carácter lúdico, que permiten al estudiante aprender la lengua en la medida que la manipula, crea textos y los analiza” (Ferramondo, 2021).

Es, por lo tanto, una aplicación real de la LC en la educación, en un curso orientado al aprendizaje de la lengua y que a su vez forma en el uso de las herramientas que provee la LC y el PLN. Para poder llevar la LC a los pregrados de ciencias del lenguaje, es importante el aprendizaje mediante la práctica. Es por eso por lo que se consideran valiosas estas experiencias más orientadas a actividades pequeñas y módulos de aprendizaje delimitados que le permiten al estudiante ir aprendiendo de su propio trabajo, sus aciertos, sus errores, etc.

Al inicio del marco teórico se presentó la Pedagogía Basada en Proyectos y la Pedagogía por Proyectos. Estos modelos pedagógicos de enseñanza son óptimos para la enseñanza de la lingüística computacional puesto que, a través de estas metodologías, “Los estudiantes deben involucrarse en un proceso sistemático de investigación, que implica la toma de decisiones en cuanto a las metas de aprendizaje, indagación en el tema y reconfiguración de su mundo interno, re significación de conocimientos” (Zambrano Briones et al., 2022). El establecimiento de cursos que utilicen la metodología de los proyectos para abordar estos conceptos de LC puede permitir que los estudiantes incursionen en esta disciplina a partir de sus propios intereses, esto podría facilitar su acercamiento a los lenguajes de programación, la sintaxis propia de los editores de texto que se utilizan para escribir código y programar. Las herramientas que se pueden utilizar para la investigación lingüística y las posibilidades que se abren a ellos con la adquisición de estos conocimientos. Por otro lado, también nos gustaría introducir la utilización de los recursos de la LC en otro contexto, o más bien, desde otro punto de vista, el del docente.

Las aplicaciones de la lingüística computacional a la docencia son muchas. Evidentemente, se le puede sacar más provecho si el modelo educativo en cuestión es mediado por TIC's o directamente educación virtual. No obstante, los docentes pueden utilizar herramientas provenientes de la LC para realizar estudios sobre las interacciones que tienen sus alumnos a través de foros entre ellos y con el personal docente, como ocurre en el artículo de Alda & Ferrero:

“Se ha aplicado una herramienta automática de análisis de textos, *Wordsmith*, para la obtención de información cuantitativamente acerca del contenido de foros en un entorno de enseñanza virtual sin requerir la intervención humana en su lectura. Se

---

ha podido establecer el carácter formal del discurso y se han identificado participantes que utilizan contenidos léxicos diferenciados. Se han creado lemas específicos para la extracción automática de elementos gramaticales y semánticos. La identificación automática de palabras clave ha permitido validar cuantitativamente los resultados obtenidos” (2007).

En este trabajo, los autores pudieron identificar de qué temas hablaban sus estudiantes, cuándo lo hacían, en qué contexto, qué interacciones eran las que más llamaban la atención y dictaban el rumbo de la conversación. Al final de su artículo proyectan la implementación de otro estudio en el que se aborde ya no solo la frecuencia de las interacciones de los alumnos, sino también la calidad de estas. Lo anterior permite visualizar otras aplicaciones que pueden llegar a tener las herramientas provenientes de la LC en entornos educativos reales. Pueden permitir realizar sondeos previos sobre el aprendizaje de los alumnos, pueden utilizarse también para dar soporte en los procesos de evaluación, controles de plagio, entre otras aplicaciones que ya son comunes hoy día. No obstante, lo que no se debe perder de vista es que más allá de las estrategias didácticas y pedagogías que se utilicen para enseñar lingüística computacional, más allá del uso que se haga para la investigación académica en el aula de las herramientas y aplicaciones que la LC provee para la comunidad lingüística, la formación en LC permite a los estudiantes ser más competentes laboralmente y estar mejor preparados para el mundo profesional del futuro cercano.

En los próximos años nos enfrentaremos a una integración agresiva de modelos de inteligencia artificial que están irrumpiendo en todas las plataformas digitales que utilizamos para estudiar, trabajar y entretenernos. Más allá del uso común, la inteligencia artificial, les permitirá a los profesionales trabajar en dichas plataformas, ya sea programándolas, reprogramándolas, modificándolas, alimentándolas, modelándolas, o utilizándolas para crear nuevo conocimiento. Lo importante es que los futuros lingüistas no sean pasivos en lo referente a las nuevas tecnologías, sino que sean activos en su creación. Esto también responde a otro propósito, uno más humanista, el de que las nuevas tecnologías no pierdan de vista el componente humano que debe mediar todo su desarrollo. Mantenernos al día nos permitirá vigilar los nuevos desarrollos. Los lingüistas son los académicos especialistas en el lenguaje, es por eso por lo que deben estar atentos a cómo se va a usar en el futuro en red, de qué forma se va a presentar la información, cómo circularan los discursos, qué uso se hará de la información lingüística que los usuarios liberan indiscretamente en internet. Ya hay herramientas que la propia LC provee para dicho propósito, por lo que es un deber mantenerse al día en esas investigaciones t

ambién. De la integración que tengamos con estas tecnologías dependerá si seremos usuarios o usados.



## 4 Marco Metodológico

Este trabajo se lleva a cabo siguiendo una metodología de investigación cualitativa. Se busca realizar un análisis de contenido a una serie de documentos oficiales proporcionados por algunas universidades de Colombia para identificar la presencia de la Lingüística Computacional o contenidos transversales fundamentales para su abordaje. Luego de realizar el análisis de contenido de los documentos se procederá a compararlos para posteriormente unificar la información. Dada la naturaleza documental de los datos con los que se va a trabajar, esta investigación se halla dentro del marco de un trabajo de archivo en cual se van a analizar las mallas curriculares de algunos programas de ciencias del lenguaje de universidades colombianas, sus respectivos Proyectos Educativos del Programa y dentro de esos programas se van a abordar los micro currículos únicamente de los cursos que puedan tener relación directa o transversal con la LC.

La naturaleza de este trabajo es exploratoria puesto que es un primer acercamiento a la presencia de la LC en los pregrados de ciencias del lenguaje en Colombia. Su propósito no es constituir una revisión teórica profunda o proponer un nuevo enfoque de trabajo dentro de esta disciplina. Asimismo, puede verse acompañado de otros trabajos posteriores que giren en torno a la implementación de la LC en los currículos de estos pregrados. Con el análisis de contenido textual que se va a realizar se espera poder identificar cómo se están abordando los contenidos seleccionados para poder reconocer sus fortalezas y oportunidades de mejoramiento.

### 4.1 Corpus

En el caso del corpus de este trabajo se trata de un corpus textual conformado por documentos oficiales institucionales que fueron solicitados mediante cartas formales dirigidas a las diferentes coordinaciones de las carreras seleccionadas. De un total de veintidós universidades contactadas, tres respondieron satisfactoriamente. La Universidad de la Amazonia<sup>12</sup> contribuyó con todos los documentos solicitados de la Licenciatura en Literatura y Lengua Castellana. La Universidad de Antioquia<sup>13</sup> contribuyó con uno de los documentos solicitados de su Licenciatura en Lenguas Extranjeras – Inglés y Francés. La

---

<sup>12</sup> <https://www.uniamazonia.edu.co/inicio/index.php/es/>

<sup>13</sup> <http://www.udea.edu.co/>

Universidad del Bosque<sup>14</sup> manifestó interés por establecer un espacio virtual de discusión en el que se pudiera ahondar en el trabajo con la intención de colaborar, aunque dicha reunión nunca tuvo lugar debido a que no se pudo coincidir en un espacio de tiempo idóneo para ambas partes. Por su parte, la Universidad Nacional Abierta y a Distancia<sup>15</sup> manifestó que, dada la naturaleza institucional de los documentos, no se había autorizado que me fueran compartidos.

Por otro lado, vía internet se consiguieron algunos documentos de la Licenciatura en Lenguas Extranjeras con Énfasis en Inglés de la Universidad Católica de Oriente<sup>16</sup> a la cual se le envió carta, pero nunca respondieron. Adicionalmente, también hacen parte del corpus los pregrados en Lingüística, Español y Filología Clásica, y Filología e Idiomas de la Universidad Nacional de Colombia<sup>17</sup>. La mayoría de sus documentos son accesibles en la red y solo se solicitaron dos documentos internamente. También es necesario mencionar que las mallas curriculares de los programas ya mencionados hacen parte del corpus, todas obtenidas en la red. A continuación, se presenta el corpus en formato de tabla con datos como su extensión, las resoluciones que lo soportan, sus componentes curriculares, entre otros.

---

<sup>14</sup> <https://www.unbosque.edu.co/>

<sup>15</sup> <https://www.unad.edu.co/>

<sup>16</sup> <https://www.uco.edu.co/>

<sup>17</sup> <https://unal.edu.co/>



Tabla 1. Corpus

Nombre de la universidad	Nombre del programa académico	Proyecto Educativo del Programa (PEP)	Malla Curricular	Asignaturas seleccionadas
<b>Universidad Católica de Oriente / Privada</b>	Licenciatura en lenguas extranjeras con énfasis en inglés / <a href="#">Enlace</a>	Acuerdo CD-010 del 30 de agosto de 2012 del Consejo Directivo. 37 páginas	Aspectos curriculares básicos establecidos de acuerdo con la normatividad vigente.	Razonamiento lógico matemático. <b>No se consiguió el documento</b>
<b>Universidad de la Amazonia / Pública</b>	Licenciatura en Literatura y Lengua Castellana / <a href="#">Enlace</a>	Resolución N.ª 11246 del MEN 135 páginas.	Resolución 18583 de 2017 del MEN.	Pensamiento lógico matemático. (9900123) y Tecnologías de la comunicación aplicadas a la educación. (6008013).
<b>Universidad de Antioquia / Pública</b>	Licenciatura en lenguas extranjeras con énfasis en inglés y francés / <a href="#">Enlace</a>	Plan de Desarrollo 2017-2027 de la Universidad. Decreto 1075 de 2015 del MEN. 121 páginas.	Plan de Desarrollo 2017-2027 de la Universidad.	Uso de las TIC en el aprendizaje y enseñanza de lenguas extranjeras. (1304037).
<b>Universidad Nacional de Colombia / Pública</b>	Filología e idiomas / <a href="#">Enlace</a>	Acuerdo 4 de 1988 del Consejo Académico. 43 páginas.	Resolución 2447 de 2008 del CSU.	<b>Asignatura no seleccionada</b>
<b>Universidad Nacional de Colombia / Pública</b>	Español y filología clásica / <a href="#">Enlace</a>	Acuerdo 111 de 1996 del Consejo Superior Universitario. 58 páginas.	Resolución 2447 de 2008 del CSU.	<b>Asignatura no seleccionada</b>
<b>Universidad Nacional de Colombia / Pública</b>	Lingüística / <a href="#">Enlace</a>	Acuerdo 10 de 1997 del Consejo Académico. 40 páginas.	Resolución 2447 de 2008 del CSU.	Lógica para lingüistas (215987), Lingüística matemática (2015988), e Informática aplicada a la investigación lingüística (215985)

Esta tabla presenta una descripción sumaria de los documentos que se recolectaron como corpus para esta investigación. Consta de seis Proyectos Educativos del Programa correspondientes a los seis pregrados de los que se logró conseguir información con sus respectivas resoluciones o disposiciones legales que los aprueban y sustentan formalmente. Adicionalmente, se indica el número de páginas de cada documento. También cuenta con seis Mallas Curriculares correspondientes a cada pregrado y que a su vez se encuentran contenidas en cada PEP. En las MC se indican también los documentos legales que las abalan o que las rigen. Finalmente, se presentan las asignaturas que fueron seleccionadas para la investigación junto con sus respectivos códigos. Es necesario señalar que de los pregrados en Español y Filología Clásica y Filología e Idiomas de la Universidad Nacional de Colombia no se eligieron asignaturas que se consideraran pudieran desarrollar contenidos de lingüística computacional.

En lo que respecta a la naturaleza de estos documentos, el PEP es un documento institucional de gran magnitud que contiene la información específica del programa académico como puede ser su misión, visión, objetivos y fundamentos teóricos que soportan su creación. Por otro lado, la malla curricular es un documento institucional de magnitud media que contiene de forma gráfica y precisa las asignaturas que se ven en la carrera, así como el orden sugerido en el que el estudiante debe cursar sus asignaturas a lo largo del tiempo. Las asignaturas se presentan organizadas por semestres y, en algunos casos, por componentes. Finalmente, el micro currículo es un documento institucional de magnitud reducida que contiene la información específica que justifica la existencia de una asignatura en particular. De forma similar a como ocurre con el PEP, la descripción de la asignatura puede contener objetivos de la asignatura, soportes teóricos, contenidos, etc. Estos documentos sirven para identificar los objetivos formativos que se establecen con los programas, lo cual es importante ya que la inclusión de asignatura está directamente relacionada con esos objetivos formativos. Es decir, si un pregrado se establece dentro de un marco de formación tradicional, es probable que no introduzca muchas asignaturas enfocadas a la vanguardia o la innovación dentro de una disciplina.

## 4.2 Muestra y recolección de información

Para la selección de la muestra, el trabajo se delimitó geográficamente en Colombia a los programas contemporáneos en el área de las ciencias del lenguaje. Los criterios que se tuvieron en cuenta para la selección de las universidades fueron los siguientes: la universidad tenía que ser colombiana, podía ser pública o privada, se buscaron pregrados de ciencias del lenguaje (lingüística, lenguas, español, filología, etc.), se filtraron aquellos programas que tuvieran en su malla curricular asignaturas de LC o contenidos fundamentales para su abordaje (programación, lógica, pensamiento lógico, pensamiento matemático, uso de TIC, etc.). Otro aspecto importante por resaltar de la muestra es que en su mayoría los pregrados que fueron escogidos son de licenciaturas en lengua castellana o lenguas modernas y extranjeras. No se halló gran número de programas sobre lingüística pura.

Para llevar a cabo la selección de universidades y programas se buscó inicialmente en la página web del Servicio Nacional de Información de la Educación Superior (SNIES)<sup>18</sup> en la cual se pudo obtener acceso a una base de datos muy completa. Sin embargo, los resultados no fueron los esperados. Algunos pregrados y universidades no aparecían, por lo que se decidió buscar otra forma de conseguir la información necesaria. La información obtenida allí se complementó con la de la página web Altillo<sup>19</sup> en la cual están organizadas las universidades por su naturaleza (pública o privada) y por departamentos. Esta última página web permite también ingresar en los enlaces para acceder a un listado de los programas académicos que ofrecen, para lo cual se procedió a utilizar un script de Python con el fin de automatizar la extracción de la información requerida.

Para la creación del script se recurrió a la ayuda del asistente de IA de Microsoft llamado Copilot<sup>20</sup>. A través de su función de búsqueda profunda se le fueron suministrando los criterios del script, así como información adicional de la página web y los datos que se necesitaban. El código se implementó en una carpeta de proyecto creada para trabajar en el editor de texto Visual Studio Code<sup>21</sup>, también propiedad de Microsoft. El proceso de recolección automática de información constó de tres etapas: la primera fue la de extraer los hipervínculos correspondientes a las universidades; la segunda fue la de extraer los datos sobre las universidades y las carreras que ofrecen y, por último, se extrajo de forma puntual la información correspondiente a las carreras de ciencias del lenguaje. A medida que se obtenían los códigos y se implementaban, se dio lugar a una etapa de ensayo y error en la que se fueron depurando los códigos hasta conseguir que cumplieran con las necesidades establecidas.

Para cumplir con la primera etapa se buscaba que el script ingresara a la página web [https://www.altillo.com/universidades/universidades\\_colombia.asp](https://www.altillo.com/universidades/universidades_colombia.asp) y extrajera los hipervínculos que estaban ligados a cada nombre de universidad del listado (Ver [Anexo A](#)). El resultado de ese script fue el archivo llamado **univ\_Col.csv** el cual contenía todos los hipervínculos con el fin de después poder acceder a ellos automáticamente. Para llevar a cabo la segunda etapa, se creó un segundo script que ingresara a los hipervínculos y extrajera la siguiente información: nombre de la universidad, nombre de la facultad, nombre de la carrera y nivel de formación (ver [Anexo B](#)). Una vez extraída la información, el script arrojaría un documento llamado **univ\_fac\_carr\_Col\_Modificado.csv**. Posteriormente, para completar la tercera etapa de la extracción automática de esta información se implementó otro script que extraería la información correspondiente a las carreras de ciencias del lenguaje (Ver [Anexo C](#)). El resultado de ese último script fue un archivo llamado **Prog\_Filtr\_Col\_Leng\_No\_Posgrado.csv**. Finalmente, se utilizó un script sencillo

---

<sup>18</sup> <https://snies.mineducacion.gov.co/portal/>

<sup>19</sup> [https://www.altillo.com/universidades/universidades\\_colombia.asp](https://www.altillo.com/universidades/universidades_colombia.asp)

<sup>20</sup> <https://copilot.microsoft.com/chats/wBXn93L35hDyDTPLfdEwJ>

<sup>21</sup> <https://code.visualstudio.com/>

---

para convertir el archivo **Prog\_Filtr\_Col\_Leng\_No\_Posgrado.csv** en un archivo .xlsx llamado **Prog\_Filtr\_Col\_Leng\_No\_Posgrado\_Excel.xlsx**.

La recolección automática de información redujo enormemente el esfuerzo humano y tiempo requeridos para la recolección de datos; sin embargo, no es un proceso perfecto y los datos arrojados contenían ruido que debió ser depurado. El ruido pueden ser datos errados, datos repetidos, datos no solicitados, entre otros. Por lo tanto, fue necesario llevar a cabo una *normalización* de los datos, el cual es un proceso manual que realiza el propio investigador. Es por eso por lo que se pasó el archivo final de formato .csv a .xlsx. El archivo final se abrió en Microsoft Excel y se procedió a revisar una por una las universidades para realizar la limpieza de los datos y verificar la información sobre los programas ofertados. Para poder verificar la información sobre los programas ofertados, hubo que recurrir a una búsqueda en línea programa por programa para validar si todavía existía, si el nombre era adecuado y después dar paso a la siguiente etapa de la recolección de los datos. En el [anexo A](#) se encontrarán los enlaces a las universidades que se eligieron al final de este proceso.

La siguiente etapa fue una fase manual. En esta etapa se revisó programa por programa buscando elegir únicamente los programas que incluían asignaturas de Lingüística Computacional o con contenidos fundamentales para la LC. De estos programas se eligieron en su mayoría licenciaturas que incluían en sus mallas curriculares asignaturas de pensamiento lógico, pensamiento lógico-matemático, lógica, programación, uso de TIC, etc. Esto condujo a que, de un total de 292 filas en el archivo original, se terminara con un total de 22 filas correspondientes a las universidades que cumplieron con los criterios. Una vez se concluyó la selección de las universidades, se recolectaron los enlaces a las páginas de esos pregrados, los correos electrónicos y nombres de los coordinadores o encargados de cada pregrado para poder enviar las solicitudes de información. Para la creación de la carta se pidió asesoría al profesor director de este trabajo de grado y, después de realizar las correcciones necesarias, el profesor las firmó para que fueran enviadas. También se procedió a redactar una plantilla de correo electrónico que conformara el cuerpo de los correos que serían enviados con las cartas. Finalmente, una vez recibidos los documentos solicitados se procedió al análisis de estos. Este procedimiento de recolección de información se ilustra en la [Figura 1](#). Para la visualización gráfica completa del procedimiento de recolección de información en un diagrama de flujo, el lector puede dirigirse al [anexo D](#).

Figura 1. Procedimiento de recolección de datos



Esta figura representa el procedimiento de recolección de datos que fue detallado en este apartado. Es una versión resumida del procedimiento y consta de cuatro etapas. La primera etapa, búsqueda en línea, es el proceso inicial de búsqueda de información que se llevó a cabo al inicio de la investigación. Fue una búsqueda inicial que pretendía determinar si la investigación era plausible, es decir, si había datos y, posteriormente, seleccionar las fuentes de las que se iban a extraer dichos datos. La segunda etapa, Web

---

Scrapping, representa un proceso de automatización de barrido de datos de página web que se implementó con el fin de realizar el trabajo de forma eficiente. El proceso se llevó a cabo con la asistencia del asistente de IA de Microsoft, Copilot y facilitó la recolección de datos. La tercera etapa, normalización de datos, fue un proceso manual. Dado que el proceso de Web Scrapping suele arrastrar datos innecesarios o repetidos, era necesario eliminar esos datos que hacían *ruido* en la muestra. Para eso se debía ingresar al documento Excel en el que estaban los datos reunidos en la etapa 2 y revisar cada columna en búsqueda de ese ruido para eliminarlas. Una vez eliminadas, se procedió a la selección de las universidades y programas que cumplieran con los requisitos establecidos. Finalmente, la cuarta etapa, recolección documental, fue aquella en la que se enviaron las cartas formales de solicitud de documentos a las universidades. Este proceso fue asistido por el profesor director del trabajo de grado. Una vez enviadas las cartas, solo quedaba esperar a que las universidades dieran sus respuestas y facilitaran los documentos solicitados para poder dar inicio al análisis de datos.

### 4.3 Análisis de los datos

Dada la naturaleza de los datos y de la metodología investigativa cualitativa que se está utilizando para este trabajo se considera como estrategia de análisis la revisión de contenido de los documentos solicitados. Esto se hace con el fin de identificar similitudes y diferencias tanto en el planteamiento de los pregrados como en el abordaje de los contenidos de las asignaturas seleccionadas. Como se muestra en la [Figura 2](#), se propuso emplear la matriz DOFA para cada pregrado seleccionado a fin de poder compararlos entre ellos. Por lo tanto, se trabajó primeramente en Excel para establecer la matriz (tabla) convencional y diligenciarla. Este Excel hace parte del material propio del investigador y no será consignado dentro del documento final. Como paso final del análisis, se establecieron los diseños de las matrices DOFA correspondientes a los pregrados de los cuales se consiguieron documentos sobre las asignaturas seleccionadas.

**Figura 2. Formato de Matriz DOFA (Debilidades, Oportunidades, Fortalezas y Amenazas)**

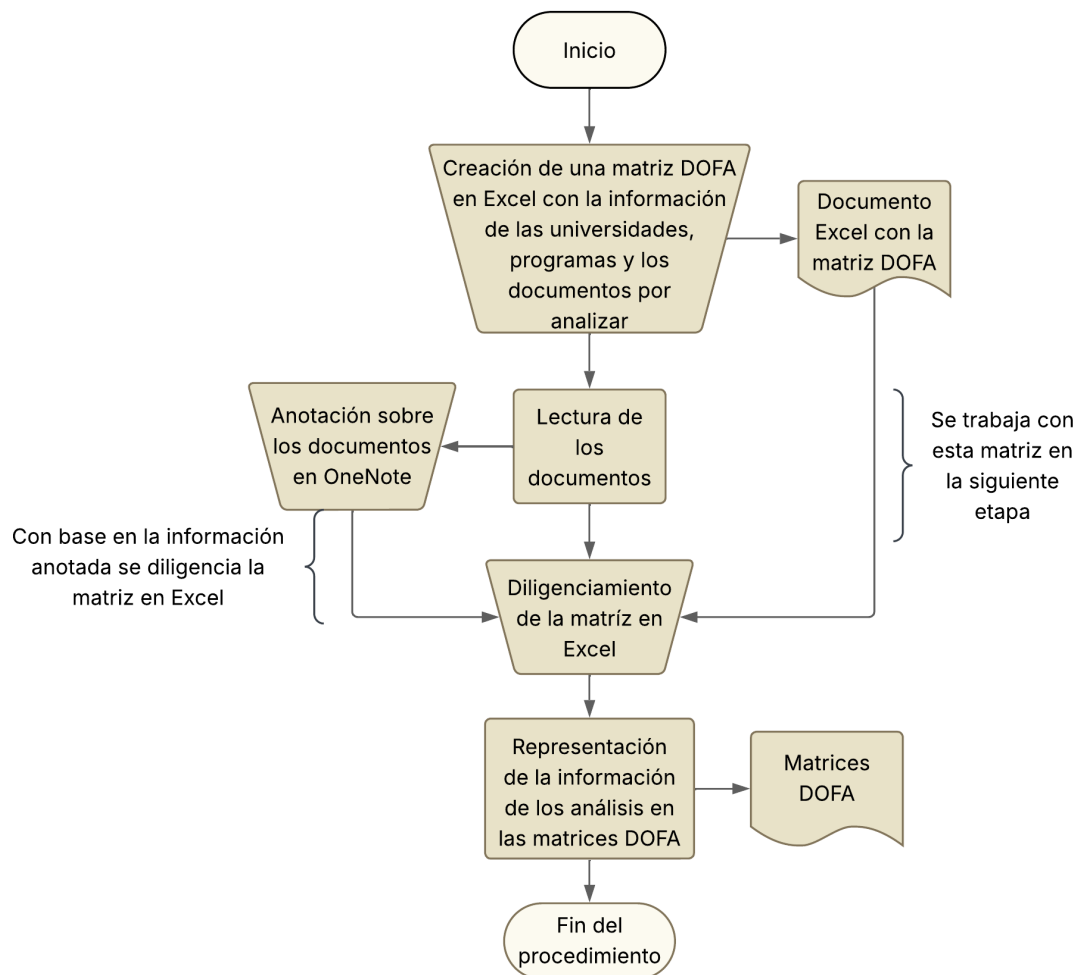


La matriz DOFA es una herramienta, u organizador gráfico, que permite representar de forma clara el estado de un proyecto, una empresa, una situación, etc. "...es una técnica de análisis que se utiliza para identificar los factores internos y externos que pueden afectar una situación o decisión. La sigla DOFA proviene de las primeras letras de cada uno de los cuatro elementos que se analizan..." (Universidad de los Andes, 2023). Estos cuatro elementos, como se puede visualizar en la matriz son: Debilidades, Oportunidades, Fortalezas y Amenazas. Las debilidades y las fortalezas son los factores internos, mientras que las oportunidades y las amenazas son los factores externos.

El procedimiento de análisis de datos consta de varias etapas. La primera etapa es la de la revisión de los documentos: en esta etapa se procedió a la lectura de los documentos para identificar los focos de interés. Esta primera etapa se complementó con el registro en OneDrive de todos esos datos que se consideraron importantes. En la segunda etapa se diligenció la matriz de Excel anotando esos datos que captaron el interés

del investigador sobre los diferentes documentos recolectados y asignándolos a alguna de las cuatro categorías de la matriz (debilidades, oportunidades, fortalezas y amenazas). Finalmente, se representó la información del análisis en las matrices DOFA correspondientes a cada uno de los pregrados en los que se aplicó a modo de recurso visual que reúna la información analizada y sirva de insumo para las conclusiones. Lo que se busca es identificar la forma en la que se abordan los contenidos para poder dar una valoración sobre ésta. La Figura 3 ilustra el procedimiento de análisis de datos mediante un diagrama de flujo.

**Figura 3. Diagrama de flujo del procedimiento de análisis de datos**



Esta figura corresponde al diagrama de flujo que representa el procedimiento de análisis documental, o análisis de datos, que se llevó a cabo en esta investigación. Básicamente, consta de cuatro etapas, aunque la primera etapa es un proceso menor. La primera etapa fue crear una matriz DOFA en Excel para registrar ahí la información de análisis y que así la diagramación fuera un proceso secundario menos exigente. La

---

segunda etapa consta del proceso más extenso en términos de tiempo, implica la lectura de todo el corpus y la toma de notas acerca de los documentos en un bloc de notas. Estas notas fueron la materia prima que se utilizó para el diligenciamiento de la matriz. La tercera etapa va de la mano con la segunda y consiste en el diligenciamiento de la matriz en Excel. Finalmente, la cuarta etapa, la diagramación, consiste en tomar los datos registrados en la tercera etapa y trasladarlos a las matrices. Se busca que las matrices sean parte del producto final porque al ser elementos visuales pueden ser más fácilmente interpretadas y menos agotadoras para el lector.

En este capítulo se abordó el corpus, junto al procedimiento de recolección de datos y el procedimiento de análisis de datos. Las descripciones precisas referentes al corpus serán suministradas en el siguiente capítulo, de resultados. El capítulo de resultados desarrollará el análisis documental en cuestión y las matrices DOFA que se mencionaron en esta metodología.

---

## 5 Resultados

Este capítulo tiene lugar después de un ejercicio de revisión documental que consistió en la lectura de los diferentes textos que se pudo reunir de las universidades seleccionadas. Para poder llegar a los resultados que se irán presentando a lo largo del capítulo, se leyó cada documento mientras se registraba su información considerada relevante en un bloc de notas unificado, que fue la base desde la cual se abordaron algunos aspectos a trabajar en esta sección. El capítulo se organizará en tres grandes secciones destinadas a abordar los tres objetivos específicos que se establecieron para este trabajo de grado en el primer capítulo: 1) Identificar los programas de pregrado del área de ciencias del lenguaje que incluyen la Lingüística Computacional en sus planes de estudio; 2) Comparar la presencia de la Lingüística Computacional entre diferentes programas de diversas universidades; y 3) Valorar los contenidos y profundidad con que se aborda la Lingüística Computacional en los currículos revisados. Por otro lado, las matrices DOFA que se expusieron en el capítulo 4, [sección 4.3](#), fueron empleadas en la tercera parte de este capítulo, sirvieron para resumir y visualizar de forma gráfica los resultados obtenidos del análisis de los documentos.

Este capítulo utilizó como fuente primaria los documentos oficiales que fueron enviados por parte de las universidades que colaboraron amablemente con este trabajo de grado. Sus respectivas referencias bibliográficas se podrán hallar en la bibliografía y fueron citados según las normas APA, 7ma edición. Los archivos que se utilizaron ya han sido detallados en la [sección 4.1](#). Por otro lado, se propuso el siguiente criterio para organizar las seis universidades que se abordaron en este apartado: en primer lugar, se ubicaría la única universidad privada de la que se recibieron los documentos solicitados y, a continuación, las demás universidades, públicas, se organizaron según su cercanía a la ciudad capital, dado que el campus principal de la Universidad Nacional de Colombia queda en Bogotá D.C. y es la universidad de la que más programas fueron seleccionados. De esta manera, el orden en que se analizará la información será el siguiente en cada sección del capítulo: 1) Universidad Católica de Oriente, privada; 2) Universidad de la Amazonia, pública; 3) Universidad de Antioquia, pública; y 4) Universidad Nacional de Colombia, pública. Por su parte, los tres pregrados del área de Ciencias del Lenguaje de la Universidad Nacional de Colombia se analizarán en el siguiente orden: 1) Filología e idiomas: inglés, francés y alemán; 2) Español y filología clásica; y 3) Lingüística. Al final se presentará un breve resumen del capítulo con recomendaciones o conclusiones.

## 5.1 Programas curriculares

Para empezar con los programas curriculares, es necesario indicar que el criterio principal que se tuvo en cuenta para seleccionar las carreras es que tuvieran relación con el área de Ciencias del Lenguaje, tales como programas profesionales y licenciaturas en lenguas, humanidades, literatura y lingüística. En segundo lugar, mediante el análisis se buscó determinar si los programas seleccionados y de los cuales se cuenta con documentación incluían asignaturas de Lingüística Computacional o contenidos de Lingüística Computacional relacionados de forma transversal. Dado que el segundo criterio solo se pudo validar tras el análisis, se encontró que no todas las carreras seleccionadas o que respondieron a la solicitud de información incluían asignaturas o contenidos de Lingüística Computacional. Tomando eso en cuenta, se encontró que algunas carreras que se seleccionaron fueron en su mayoría licenciaturas que incluían dentro de sus mallas curriculares asignaturas de lógica, pensamiento lógico, pensamiento matemático, pensamiento lógico matemático, uso de TIC, TIC aplicadas a la educación, etc. Por otro lado, uno de los documentos que se les solicitaron a los encargados de esos programas curriculares fue el Proyecto Educativo del Programa (PEP). El PEP es un documento de gran importancia porque es el que oficialmente soporta la creación y actualización de un programa. Dentro de la información que se puede encontrar en un PEP están: la legislación interna y externa que lo aprueba, la historia del programa, el marco teórico y metodológico, la malla curricular y cómo está distribuida o conformada, información adicional como puede ser información de extensión, investigación y movilidad académica, entre otros. Dada su importancia, el PEP es el primer documento que se abordó en este capítulo de resultados.

### 5.1.1 Licenciatura en Lenguas Extranjeras con Énfasis en Inglés

La Universidad Católica de Oriente se ubica en Antioquia y es de naturaleza privada. Se rige según sus Estatutos, aprobados por el Acuerdo CD-010 del 30 de agosto de 2012, del Consejo Directivo y ratificados por el Ministerio de Educación Nacional mediante resolución N.º 6770 del 31 de mayo de 2013. El programa elegido es la Licenciatura en Lenguas Extranjeras con Énfasis en Inglés. Dentro de su PEP se establece un cupo de 164 créditos y una duración de 10 semestres. Este programa se rige por la Autonomía universitaria, la Libertad de cátedra y aprendizaje, Fidelidad filosófica, Excelencia académica, Gestión del conocimiento, Evangelización y compromiso social con la región y el país, Eficiencia administrativa y Cuidado de la Casa Común. El objetivo general de esta licenciatura es:

“Formar profesionales con enfoque humanista y excelente nivel académico, con conciencia crítica e intercultural, conciencia con capacidad de acompañar procesos formativos en el idioma inglés en los diferentes ámbitos educativos y con conocimiento de la realidad educativa para la transformación de la misma” (Universidad Católica de Oriente, 2019).

Para ello se proponen formar profesionales integrales con una buena competencia en las lenguas inglés y francés, que las usen como medio de comunicación y las aborden

---

de forma consciente. También se proponen que el egresado alcance nivel C1 en inglés y A1 en francés, así como:

“Desarrollar en el estudiante una competencia pedagógica que le permita detectar, analizar y resolver adecuadamente los problemas que se le presenten, utilizar los medios de que disponga para realizar su trabajo incluyendo las tecnologías más actualizadas y ser consciente de las diferencias en las capacidades de aprendizaje de los alumnos” (Universidad Católica de Oriente, 2019).

Por último, se le da importancia a la competencia investigativa del profesional. Dentro de su metodología se encuentra el trabajo por proyectos, el aprendizaje por contenidos, el aprendizaje por tareas, el blended learning, y el aprendizaje autónomo. La metodología toma en cuenta la pedagogía del texto, pedagogía social, pedagogía crítica y pedagogía socio crítica. Además, busca que el alumno se apropie de la lengua y su funcionamiento a partir del uso práctico real. Dentro de este programa se fomenta la práctica pedagógica desde los primeros semestres dado que la UCO considera que:

“...la práctica pedagógica, la formación en investigación, el dominio de las tecnologías de la información y la comunicación, la competencia en inglés en el nivel C1 y en francés en nivel A1, posibilitan el desarrollo de competencias en el campo de las ciencias sociales (científico social) y gestor y ejecutor de procesos de investigación” (Universidad Católica de Oriente, 2019).

Además, se indica que:

“...las Prácticas Pedagógicas deben asegurar que los estudiantes comprendan y apropien las dinámicas del aula y su contexto, reconozcan las diferencias y modalidades de formación de niños, niñas, adolescentes, jóvenes y adultos, y la asocien con la disciplina que se enseña y con las situaciones, eventos o fenómenos que dicha disciplina conlleva” (Universidad Católica de Oriente, 2019).

Cabe señalar que la práctica docente también se concibe como un espacio para la investigación. Dentro del PEP se encuentra una constante reflexión sobre la importancia de la investigación en la docencia, la importancia de la interdisciplinariedad, la interculturalidad y el pensamiento crítico. El futuro profesional debe integrar estos elementos puesto que su campo de acción no se limita a la enseñanza, se puede ver entendido hacia proyectos de bilingüismo públicos o privados, traducción, investigación con entidades oficiales, etc.

### **5.1.2 Licenciatura en Literatura y Lengua Castellana**

La Universidad de la Amazonia es una universidad pública ubicada en Florencia, Caquetá, con otras sedes en Guaviare, Putumayo y Leticia. El PEP de la Licenciatura en Literatura y Lenguas Castellanas es resultado de la Resolución N.º 11246 del 02 de junio de 2017 del MEN, la Resolución 013842 del 15 de agosto de 2018 y modificado por la Resolución

---

008946 del 27 de agosto de 2019. En su PEP se establece un cupo de 165 créditos y una duración de 10 semestres. Este documento presenta una extensa elaboración teórica en la que se hace un repaso por los referentes pedagógicos que se tomaron en cuenta para la creación del programa. Se presenta que el programa tiene enfoque histórico cultural y enfoque sociocultural. Se hace revisión de las teorías de Vygotsky y Bajtín. Se aborda la característica social del lenguaje y su función primaria, la comunicación.

Como objetivo general se establece “Formar docentes con dominio del lenguaje y carácter investigativo que propicien espacios de aprendizaje autónomo regidos por principios científicos” (Universidad de la Amazonia, 2019b). Dentro de los objetivos específicos se pueden encontrar: “Brindar al futuro maestros referentes teóricos y prácticos que lo habiliten como docente de la lengua y la literatura”, “Aproximar a los estudiantes a las discusiones nacionales e internacionales sobre la enseñanza de la lengua, la literatura y otros sistemas significativos”, “Desarrollar la competencia comunicativa en términos de su desempeño como sujeto social y como futuro docente” y “Ofrecer escenarios de práctica adecuados para que los estudiantes mejoren sus desempeños como docentes de lengua y literatura” (Universidad de la Amazonia, 2019b), entre otros.

Por otro lado, dentro del PEP se presentan los tres ciclos de formación que fueron establecidos para esta carrera. El ciclo básico: “En esta etapa se ubica como principal objetivo en cada uno de los componentes la identificación, comprensión y acercamiento a los saberes propios de la Licenciatura en Literatura y Lengua Castellana” (Universidad de la Amazonia, 2019b). El ciclo de profesionalización: “En este ciclo los objetivos de cada componente serán: relacionar saberes con otras disciplinas (interdisciplina); reconocer los saberes en su contexto, utilizar los conocimientos adquiridos en su área profesional a desempeñar (práctica docente)” (Universidad de la Amazonia, 2019b). Y, finalmente, ciclo de profundización: “Destaca como principales objetivos: la reflexión y problematización sobre los usos de los conocimientos adquiridos durante los ciclos anteriores, y que tienen como fin transformar tanto sus prácticas como sugerir nuevos caminos para la comunidad académica” (Universidad de la Amazonia, 2019b).

Adicionalmente, se establecen los saberes específicos y disciplinares de la Licenciatura en Literatura y Lengua Castellana. En primer lugar, estudios sobre el lenguaje y la comunicación. En segundo lugar, estudios sobre la teoría literaria y la literatura. En tercer lugar, estudios transdisciplinarios del lenguaje. Finalmente, lectura crítica, aspectos semióticos, estéticos, sociológicos e historiográficos. Con respecto a la educabilidad del ser humano, se hace referencia a cómo configuramos una segunda realidad desde una dimensión sociocultural que va más allá de lo biológico. El ser humano tiene potencial más allá de lo que su evolución biológica ha determinado, pero para poder explorar ese potencial se necesita apoyo, se necesita educación, orientación. Por lo tanto, se habla de que:

“El hombre es un sujeto incompleto que tiene la capacidad para ser educado por otros, para autoeducarse y para educar. La educabilidad del ser humano es

activada por los procesos de instrucción y educación dados en el proceso de enseñanza aprendizaje” (Universidad de la Amazonia, 2019b).

### **5.1.3 Licenciatura en lenguas extranjeras con énfasis en inglés y francés**

La Licenciatura en lenguas extranjeras con énfasis en inglés y francés es ofertada por la universidad de Antioquia, una universidad pública ubicada en Medellín. El PEP de este programa académico se elaboró con base en el Decreto 1075 del 26 de mayo de 2015 del MEN. A su vez, se alinea con el Plan de Desarrollo 2017-2027 de la Universidad, el Proyecto Educativo Institucional 2022 y otras políticas institucionales. De forma similar a como sucede con el PEP del programa anterior, en este documento se realiza una extensa elaboración teórica que nutre y sustenta al programa en cuestión. Adicionalmente, en este documento se establece un cupo de 156 créditos y una duración de 10 semestres para este programa curricular. Los principios de formación de la Escuela de Idiomas fueron aprobados el 12 de marzo de 2021 por el acta 90 del Consejo de Escuela. Dentro de estos principios se contemplan diferentes dimensiones de formación del estudiante.

En primer lugar; “Dimensión académica: mediante la apropiación conceptual y el desarrollo de competencias del saber —saber hacer, saber aprender y saber convivir—, empoderar al sujeto para su ejercicio profesional” (Universidad de Antioquia, 2022). En segundo lugar:

“Dimensión cultural: desde el conocimiento y la exploración de las complejas relaciones entre las lenguas y las culturas, se busca desarrollar en los estudiantes (sic) de deconstruir y visibilizar las relaciones inequitativas entre individuos y grupos culturales diversos y proponer intercambios justos que beneficien a todos los actores” (Universidad de Antioquia, 2022).

En tercer lugar, la “Dimensión sociopolítica: la idoneidad de los estudiantes para formarse como agentes de cambio en capacidad de entablar relaciones con otros para la participación y la transformación social” (Universidad de Antioquia, 2022). Finalmente, la “Dimensión actitudinal: hace referencia al desarrollo de capacidades humanas que promueven el relacionamiento espontáneo del ser con su entorno, incluidas otras especies de la naturaleza” (Universidad de Antioquia, 2022).

Adicionalmente, la licenciatura se orienta hacia la formación de docentes que reflexionen sobre su identidad como maestros, promuevan la reflexión sobre temas sociales como las relaciones de poder y la diversidad, fomenten el desarrollo de una consciencia crítica e intercultural, desarrollen su competencia en inglés y francés, así como que transformen los espacios de aprendizaje. Además, se espera que el futuro maestro priorice el papel del lenguaje como una forma de construir conocimiento, no como un fin. Por otro lado, se espera que el egresado genere puentes en los que se encuentren la teoría y la práctica, se integren las TIC, se forme en autonomía, responsabilidad y compromiso,

así como el desarrollo personal. Dentro de este apartado también se presenta el concepto de lengua con el que van a trabajar y se presentan estas tres visiones de lengua: “(a) la lengua como sistema o código lingüístico, (b) la lengua como un sistema comunicativo o un discurso, y (c) la lengua como una ideología, como experiencia del mundo o práctica social” (Universidad de Antioquia, 2022).

Por otro lado, se señala la importancia del español en la formación de lenguas extranjeras dado que a través de la lengua madre se da el primer acercamiento a los contenidos lingüísticos, pedagógicos y transversales. También se conciben al inglés y al francés como lenguas internacionales, lo cual justifica en parte la existencia del programa. En cuanto a la pedagogía, se indica que se aborda el trabajo por proyectos, la literacidad crítica, el Análisis Crítico del Discurso y las perspectivas interdisciplinarias. En lo que respecta a la didáctica, se señala la necesidad existente de un docente flexible que se pueda adaptar a las situaciones que se presentan en los entornos de aprendizaje. Finalmente, en el perfil del egresado se indica como una de sus características “Una sólida formación humana, pedagógica y didáctica que le permiten promover crítica y creativamente la educación integral para sus estudiantes” (Universidad de Antioquia, 2022).

#### **5.1.4 Filología e idiomas: Alemán, Inglés y Francés**

El pregrado de Filología e Idiomas: Alemán, Inglés y Francés es ofertado por la Universidad Nacional de Colombia, cuya sede principal se ubica en la ciudad de Bogotá D.C. Este PEP se crea bajo el Acuerdo 4 de 1988 del Consejo Académico. Se estableció un cupo de 140 créditos y una duración de 8 semestres. Desde su creación sufrió varios cambios en su formalización y en cuanto al departamento al que pertenecía. Sin embargo, en la actualidad está adscrito al Departamento de Lenguas Extranjeras, el cual pertenece a la Facultad de Ciencias Humanas. El objetivo del programa se establece bajo el Acuerdo 215 de 2008: “Formar docentes en lengua extranjera que sean capaces de reflexionar sobre aspectos de la lengua, su pedagogía, su cultura y que puedan impulsar cambios a través de la investigación y el diario quehacer como investigadores” (Universidad Nacional de Colombia, 2015).

Comparte características pedagógicas con los otros tres programas que se van a presentar a continuación pertenecientes al área de Ciencias del Lenguaje de la misma universidad. Además, se adscribe a los Lineamientos para la Formación de Estudiantes de la Universidad Nacional de Colombia establecidos en el Acuerdo 033 de 2007. Estos lineamientos se fundamentan por los principios de excelencia académica, formación integral, contextualización, internacionalización, formación investigativa, interdisciplinariedad y flexibilidad. De igual manera, se establece la obligatoriedad de un componente de nivelación, el cual, dependiendo de la carrera, puede ser en matemáticas o en competencias comunicativas (lectura y escritura). La mencionada flexibilidad es muy importante para el desarrollo integral de los estudiantes puesto que garantiza que puedan abordar temas de su interés más allá de los propios de su programa de estudios. Este

componente obliga a que el 20% de los créditos del programa sean de asignaturas electivas.

Adicionalmente, en cuanto a la flexibilidad también se encuentra la opción de la doble titulación, la cual permite que el estudiante pueda ampliar su horizonte académico, profesional y laboral. Por otra parte, se señala la naturaleza interdisciplinaria del programa. Así como se señala la importancia de la formación interdisciplinaria y la formación intercultural dentro del PEP. Como metodologías se contemplan al aprendizaje cooperativo, aprendizaje basado en problemas y pedagogía por proyectos, tomando en consideración que “...estos modelos didácticos activan habilidades de pensamiento orientado al análisis, la síntesis, la inferencia, la abstracción y la creatividad, entre otros, los cuales subyacen a los procesos de interpretación, argumentación y proposición” (Universidad Nacional de Colombia, 2015).

Por su parte, en lo referente al perfil del egresado de este programa, se señala que:

“Los egresados tendrán una sólida competencia comunicativa en lengua extranjera, competencias investigativas e interculturales y formación en la pedagogía de la misma. Estas competencias les permitirán ser docentes, usuarios y analistas de la lengua extranjera. Estarán también en capacidad de entender y valorar su cultura y las foráneas, de integrar una visión investigativa y reflexiva en su ejercicio profesional y de liderar propuestas en su área de desempeño” (Universidad Nacional de Colombia, 2015).

Dentro de los espacios que se destinan a la formación interdisciplinaria de los alumnos y para fomentar el dominio de la lengua se instauraron clubes de conversación y cafés literarios. Estos espacios de encuentro, conversación y estudio no son exclusivos de los estudiantes del pregrado en Filología e Idiomas, por lo que estudiantes de otros programas académicos pueden participar activamente de ellos. Son importantes en el PEP ya que complementan la formación extracurricular de los estudiantes y fomentan la integración entre alumnos de diferentes programas curriculares.

### **5.1.5 Español y Filología Clásica**

El pregrado en Español y Filología Clásica es ofertado por la Universidad Nacional de Colombia, cuya sede principal se ubica en la ciudad de Bogotá D.C. Su PEP se creó bajo el acuerdo 111 de 1996 del Consejo Superior Universitario. Empezó a ofertarse en el año de 1997 de manera anual y desde 2011 se ofrece semestralmente, lo cual contribuyó al aumento en los estudiantes de ingreso y egresados. Se estableció un cupo de 125 créditos y una duración de 8 semestres. Este pregrado nace a partir de la necesidad de “...mantener, fortalecer y fomentar el estudio y la enseñanza de las Lenguas Clásicas en el país...” (Universidad Nacional de Colombia, 2014a) y concibe a las Lenguas Clásicas como objeto de conocimiento. Busca “...conocer con profundidad la estructura, historia, desarrollo, evolución y literaturas del Español y las Lenguas Clásicas...” (Universidad

---

Nacional de Colombia, 2014a). Se señala como uno de los argumentos por los cuales la Universidad oferta este programa el hecho de que "...la Universidad promueve el acceso con equidad al sistema educativo colombiano en las diversas ramas del saber, independientemente de las tendencias del mercado..." (Universidad Nacional de Colombia, 2014a). Adicionalmente, se reconoce al español como la segunda lengua con más hablantes nativos en el mundo y una lengua de interés cultural universal.

Por su parte, son dos los objetivos que orientan a este programa curricular. El primero es "Preparar profesionales capaces de difundir y contribuir al conocimiento y desarrollo de las competencias del Español y al conocimiento de la Filología Clásica". El segundo es "Estimular en la Universidad y en el país el estudio de los textos producidos en las Lenguas Clásicas, tanto desde el punto de vista lingüístico como desde el punto de vista cultural" (Universidad Nacional de Colombia, 2014a). Para lograrlo se promueve la formación de profesionales integrales, competentes, socialmente responsables y con perspectiva transdisciplinaria. Dichos profesionales deberán comprender que las Lenguas Clásicas son importantes dado su papel fundacional en la cultura y pensamiento occidental, además de que el latín es la lengua de la que descende el Español. Por su parte, el perfil del egresado licenciado en Español y Filología Clásica indica lo siguiente:

"El egresado estará capacitado para desempeñarse como docente de Español y Lenguas Clásicas, y para participar en la planeación, diseño y evaluación de procesos de enseñanza y aprendizaje de estas con el fin de impulsar proyectos de transformación pedagógica. Además, podrá realizar asesorías a editoriales y traducciones de textos clásicos a la lengua española e investigar con profundidad los diferentes problemas relativos a las Lenguas Clásicas y el Español" (Universidad Nacional de Colombia, 2014a).

En lo que refiere al plan de estudios, la estructura de los programas de la Universidad Nacional de Colombia se unificó en 2008 para facilitar traslados, intercambios, dobles titulaciones, homologaciones, entre otros. Es así como todos los programas académicos de la Universidad tienen los mismos cuatro componentes: nivelación, fundamentación, disciplinar o profesional, y electivo. Por otro lado, el modelo pedagógico que se aplica dentro de las estrategias pedagógicas que se implementan en este programa es el Modelo Pedagógico Humanista. Además, tomando en cuenta la cercanía con la Lingüística, se aplica extensamente el Análisis Lingüístico Descriptivo. Se asume como principal objetivo referente a las Lenguas Clásicas el "...desarrollo de la habilidad para leer y entender textos clásicos en la lengua de origen (griego o latín)" (Universidad Nacional de Colombia, 2014a). No obstante, se señala que "La lectura en contexto no se orienta a la resurrección del mundo clásico" (Universidad Nacional de Colombia, 2014a).

Sin embargo, se rescata la importancia de la traducción ya que le permite al estudiante volcar su conocimiento y le exige trabajar en su capacidad interpretativa literaria. También se plantea como enfoque la lectura de los textos y no el aprendizaje de la gramática. La gramática se concibe como un aprendizaje colateral. Se presenta una reflexión en torno a cómo enseñar las Lenguas Clásicas y se destaca la ventaja que tienen

tanto el griego como el latín para su aprendizaje: su léxico está congelado en el tiempo. Es decir que, a diferencia de una lengua moderna como puede ser el inglés o el francés, las cuales están en constante desarrollo y transformación, el léxico de las Lenguas Clásicas está congelado en los documentos escritos en esas lenguas. Estos textos son, de hecho, el objeto de análisis de la filología clásica.

En lo referente a la lengua, y para finalizar esta sección, se abordan los niveles de análisis lingüístico. Se señala la importancia de dedicar tiempo al estudio de la sintaxis, dado que es un aspecto de la lengua que requiere de un considerable trabajo para dominarla. Se busca que la aproximación a la lengua sea desde la perspectiva de la lengua viva, una lengua que comunica, que crea cultura e identidad. Se espera realizar una "...aproximación integral a la lengua, a sus historias, a sus culturas" (Universidad Nacional de Colombia, 2014a). Adicionalmente, se plantea la problemática de la identidad en América, la cual está estrechamente ligada a la lengua y al uso del lenguaje.

### **5.1.6 Lingüística**

El pregrado en Lingüística es ofertado por la Universidad Nacional de Colombia, cuya sede principal se encuentra en la ciudad de Bogotá D.C. Es un pregrado amplio y completo que profundiza en diferentes aspectos de la lingüística, por lo que puede llegar a considerarse como precursor de los dos pregrados presentados anteriormente. El PEP de este pregrado se creó bajo el Acuerdo 10 de 1997 del Consejo Académico, consta de 123 créditos y 8 periodos académicos o semestres. Los objetivos del programa según el PEP son tres. En primer lugar, "Estudiar y enriquecer el lenguaje en general y el español de Colombia, así como las lenguas indígenas y criollas como patrimonio cultural, natural y ambiental del país. Así mismo, contribuir a su preservación" (Universidad Nacional de Colombia, 2014b). En segundo lugar, formar profesionales de manera científica, ética y humanística. Y, en tercer lugar, "prestar apoyo al Estado en lo pertinente a la disciplina, siguiendo los principios de la autonomía académica e investigativa" (Universidad Nacional de Colombia, 2014b). Por otro lado, dentro del perfil del egresado, se establece que:

"Hay tres grandes campos de acción: la investigación, el diseño de políticas en la esfera la comunicación y la investigación educativa. Los egresados podrán participar en investigaciones sobre problemáticas del lenguaje en general, de la lengua española, de las lenguas indígenas y de los procesos comunicativos. Estarán capacitados para solucionar problemáticas de los procesos de producción de textos orales y escritos, así como para cualificar los procesos discursivos desde diversas manifestaciones comunicativas y culturales. Así mismo, podrán contribuir al conocimiento y preservación de las lenguas indígenas y criollas, proponer y desarrollar políticas lingüísticas en el campo de la educación y de los medios de comunicación e impulsar proyectos de transformación de la enseñanza y/o el aprendizaje de las lenguas habladas en el país" (Universidad Nacional de Colombia, 2014b).

Este programa se rige por los mismos lineamientos para la formación de los estudiantes de la Universidad Nacional de Colombia establecidos según el Acuerdo 033 de 2007 del Consejo Superior Universitario que se presentaron anteriormente. Dentro de las estrategias pedagógicas que rigen a este pregrado y que están expuestas en el PEP (Universidad Nacional de Colombia, 2014b) se puede encontrar: la apropiación de conceptos, la adquisición de destrezas de análisis, el conocimiento de información manejo de datos empíricos relativos a un tema o problema específico y, la adquisición de competencias generales para la vida académica. Se busca que el estudiante adquiera competencias transversales, así como las disciplinares de su área.

También se incluyen en el PEP aspectos sobre la movilidad académica interna y externa, con intercambios a nivel nacional e internacional, así como la importancia de las prácticas y pasantías. Con respecto a las pasantías, es necesario decir que dentro de los pregrados de la Universidad Nacional de Colombia se contemplan como una opción de las disponibles para cumplir con los requisitos de trabajo de grado. Igualmente, se incluye hacia el final un recuento con trabajos realizados por egresados de la carrera, investigaciones y trabajos de grado de gran valor. Adicionalmente, se hace un recuento de grupos de investigación junto a su información complementaria, como lo puede ser su número de integrantes, su campo de investigación, sus docentes a cargo, sus productos y su calificación por parte de Colciencias.

Para resumir, en este apartado se presentaron los datos que se recopilamos como resultados de la revisión documental que se hizo de los Proyectos Educativos del Programa pertenecientes a los programas curriculares que fueron seleccionados para esta investigación. Dentro de la información reunida y presentada hay datos generales sobre esos pregrados como puede ser su cupo de créditos, su duración en semestres, sus fundamentos teóricos, pedagógicos y metodológicos, así como las disposiciones legales internas y externas que los avalan. Se omitió la información específica sobre las mallas curriculares ya que esa información será presentada en la siguiente sección de este capítulo.

## **5.2 Presencia de la Lingüística Computacional en los Programas Curriculares de Ciencias del Lenguaje**

En esta sección se comparará la presencia de la Lingüística Computacional en los programas curriculares de Ciencias del Lenguaje que fueron seleccionados. Para poder alcanzar ese objetivo, se presentará una tabla sumaria con la información sobre las universidades, los programas y las asignaturas que fueron seleccionadas. Esta tabla contiene comentarios sobre la presencia de esas asignaturas. Posteriormente, en el resto de la sección se describirá cada una de las seis mallas curriculares con las que se trabajó; esta información procede tanto de los PEP como de las mallas curriculares. Sin embargo, es necesario aclarar que todos los PEP con los que se trabajó contienen dentro de ellos las mallas curriculares junto con descripciones de sus componentes o unidades de

---

contenido, por lo tanto, las mallas curriculares que se obtuvieron en línea en las páginas web de las diferentes carreras cumplen un papel más de apoyo e ilustración que como fuente primaria. Al final se concluirá con un breve resumen de la información presentada.

Para poder dar inicio al trabajo que se hará en esta sección, empezamos por la [Tabla 2](#) que presenta la comparación propiamente dicha. En esta tabla se incluyeron los nombres de las universidades, los nombres de las carreras, los nombres de las asignaturas, número de créditos, intensidad y distribución horaria, carácter obligatorio u optativo, y anotaciones comparativas o comentarios de ser necesarios. El criterio que se siguió para la organización de las asignaturas en la tabla fue el mismo que ya se presentó al inicio del capítulo, es decir: 1) Universidad Católica de Oriente, 2) Universidad de la Amazonia, 3) Universidad de Antioquia, 4) Universidad Nacional de Colombia (Filología e idiomas: Alemán, Inglés y Francés, Español y Filología Clásica, y Lingüística).

Tabla 2. Comparación sobre la presencia de la LC en los pregrados de Ciencias del Lenguaje seleccionados

Nombre de la Universidad	Nombre del Programa Académico	Nombre de la asignatura seleccionada	Comentarios comparativos
<b>Universidad Católica de Oriente</b>	Licenciatura en lenguas extranjeras con énfasis en inglés	Razonamiento lógico matemático	<u>No se consiguieron datos de este micro currículo y la universidad no respondió a la solicitud formal.</u>
<b>Universidad de la Amazonia</b>	Licenciatura en literatura y lengua castellana	Pensamiento lógico matemático	Octavo semestre. Área de Fundamentos Generales. 2 créditos. 48 horas de trabajo presencial y 48 horas de trabajo: 96hrs. independiente. Optativa.
		Tecnologías de la comunicación aplicadas a la educación	Octavo semestre. Área de Fundamentos Generales. 2 créditos. 48 horas de trabajo presencial y 48 horas de trabajo: 96hrs. independiente. Optativa.
<b>Universidad de Antioquia</b>	Licenciatura en lenguas extranjeras con énfasis en inglés y francés	Uso de las TIC en la enseñanza y aprendizaje de las lenguas extranjeras	Noveno semestre. Componente de Pedagogía y Didáctica. 4 créditos. 64 horas de trabajo presencial y 64 horas de trabajo independiente: 128hrs. Obligatoria.
<b>Universidad Nacional de Colombia</b>	Filología e Idiomas: Alemán, Inglés y Francés Español y Filología Clásica	<u>Asignatura no seleccionada</u>	<u>En estos pregrados no se encontró oferta de asignaturas destinadas a la LC o fundamentales para su abordaje.</u>
		Lógica para lingüistas	Quinto semestre. Componente de Fundamentación. 3 créditos. 64 horas de trabajo presencial y 80 horas de trabajo independiente: 144hrs. Optativa.
	Lingüística	Lingüística matemática	Sexto semestre. Componente de fundamentación. 3 créditos. 64 horas de trabajo presencial y 80 horas de trabajo independiente: 144 horas. Optativa
		Informática aplicada a la investigación lingüística	Sexto semestre. Componente de Fundamentación. 3 créditos. 64 horas de trabajo presencial y 80 horas de trabajo independiente: 144hrs. Optativa.

---

En esta tabla se presentaron los seis programas curriculares que fueron analizados, junto a las asignaturas que se eligieron de estos programas. Posteriormente, se elaboró una breve comparación de la presencia que tienen esas asignaturas en sus respectivos pregrados. En esta comparación se señaló el semestre en el que se ofertan, el componente al que pertenecen, su cupo de créditos y la intensidad horaria sugerida por la Universidad. Se señala también que no se logró reunir el micro currículo de la asignatura seleccionada en la Universidad Católica de Oriente, Razonamiento Lógico Matemático. Adicionalmente, se señala que en los pregrados de Filología e Idiomas y Español y Filología Clásica de la Universidad Nacional de Colombia no se seleccionaron asignaturas que se consideraran relacionadas con la Lingüística Computacional. Al contrario, en el pregrado en Lingüística sí se encontraron asignaturas relacionadas con esta área.

Para empezar con el desarrollo de esta sección se abordará la nomenclatura de estas asignaturas seleccionadas puesto que permitió realizar desde el inicio una diferenciación general. Las asignaturas seleccionadas se pudieron separar en dos grandes grupos: pensamiento o razonamiento lógico matemático y uso de TIC en la educación. Esto, por sí mismo nos permite entender la forma en que los pregrados se acercan a los contenidos, herramientas o usos de la LC. El *Pensamiento Lógico Matemático (PLM)* es importante en muchos contextos de la sociedad actual. Esta forma de pensar y abordar los problemas ayuda a tener una visión crítica y analítica de las situaciones o incógnitas que se presentan en el día a día dentro de diferentes disciplinas. No obstante, en el contexto de esta investigación su papel principal es el de servir de base para el acercamiento a los sistemas computacionales, la informática y la programación. Si lo que se busca es una apropiación de las herramientas tecnológicas de las que se dispone, empezar por el PLM es, sin duda, una decisión acertada.

Esto se puede ver especialmente en el pregrado en Lingüística de la Universidad Nacional de Colombia, se trata de tres cursos, todos interconectados y que llevan una línea de progresión en los conocimientos. En estos tres cursos se puede ver cómo se avanza desde fundamentos de lógica matemática y cómo se aplica al procesamiento del lenguaje, hasta llegar a aplicaciones más concretas como el PLN. Esto último tiene sentido si tomamos en cuenta que se trata de un pregrado en Lingüística, una carrera que podría considerarse *pura*, como lo sería la matemática, la física, la biología o la medicina, por mencionar algunas.

Por otro lado, encontramos asignaturas dedicadas a la aplicación de las TIC a la educación. Estos cursos incluyen contenidos formales y teóricos de computación e informática, pero no de pensamiento lógico o de programación. Son cursos de corte más humanístico y reflexivo en los que se busca abordar el papel, la importancia y las implicaciones que las TIC tienen en la sociedad y, específicamente, en la educación. Por supuesto, como licenciaturas que son, así como programas altamente enfocados a la investigación, no se limitan a una reflexión en clase. Estos cursos también promueven la innovación y la vanguardia educativa mediante las TIC. Son cursos que buscan que las TIC se apropien por parte del maestro de manera tal que este, a su vez, logre que sus

---

futuros alumnos las apropien de forma crítica y responsable. La responsabilidad social y el uso ético de las herramientas tecnológicas es de gran importancia en estos cursos.

Con base en lo anterior, no es de extrañar que el periodo de formación en el que se suelen ofertar estos cursos sea en las etapas finales de sus respectivos programas académicos. Esto tiene sentido si tomamos en cuenta que para poder entrar a abordar tanto conceptos como herramientas matemáticas, computacionales e informáticas aplicadas a la lingüística, es decir al lenguaje, primero debemos saber de lingüística. Es por eso por lo que en este tipo de programas curriculares encontramos una constante, los componentes fundamentales y los componentes disciplinares o profesionales. En los componentes fundamentales encontramos contenidos transversales a todo el programa y las demás asignaturas que se van a tomar. Son comunes los talleres de textos, los cursos de nivelación en matemáticas, cursos dedicados a la investigación, entre otros.

Por su parte, la Universidad de la Amazonia nos dice que “Este componente tiene como objetivo identificar, reconocer y manejar los saberes básicos del ámbito académico en cuanto a escritura, lectura, formación ética, y social” (Universidad de la Amazonia, 2019b). A su vez, la Universidad Nacional de Colombia señala en el Artículo 9 del Acuerdo 33 de 2007 del Consejo Superior Universitario lo siguiente sobre el componente de fundamentación:

“Este componente introduce y contextualiza el campo de conocimiento por el que optó el estudiante desde una perspectiva de ciudadanía, humanística, ambiental y cultural. Identifica las relaciones generales que caracterizan los saberes de las distintas disciplinas y profesiones del área, el contexto nacional e internacional de su desarrollo, el contexto institucional y los requisitos indispensables para su formación integral” (Consejo Superior Universitario - Universidad Nacional de Colombia, 2025).

Por ende, es entendible que, dependiendo de la naturaleza del programa, se vean cursos introductorios a la lingüística. Sin embargo, el componente en el que realmente se profundiza en los estudios de la lengua en este tipo de pregrados es el componente disciplinar o profesional. Este componente es el que engloba la mayoría de los saberes específicos de un área del conocimiento. Es decir, si se trata de programas curriculares de las Ciencias del Lenguaje, se esperaría encontrar en este componente la mayoría de las asignaturas dedicadas a la lingüística per se. Hablamos, entonces, de asignaturas como fonética, fonología, morfología, sintaxis, pragmática, semántica, etc. En pregrados como los de lenguas extranjeras, además, es usual encontrar esas mismas asignaturas, pero con una especialización adicional orientada a la lengua de estudio en cuestión. No ocurre de la misma forma en los pregrados en español o lingüística, en donde esas asignaturas puede que aborden específicamente el español por ser la lengua madre del estudiante, pero también abordará la disciplina en términos generales.

Como pudimos evidenciar, la disposición temporal de las asignaturas es coherente con los programas curriculares en donde se ofertan. Su presencia también tiene sentido,

---

con programas de licenciatura en los que prima el uso de TIC en entornos educativos, y con el pregrado de lingüística, que dedica tres asignaturas al trabajo de Lingüística Computacional, aún si no llevan ese nombre.

Finalmente, se procedió a abordar el tema de la obligatoriedad de la asignatura y su intensidad horaria. En este aspecto se encuentra un panorama variado, dado que los cursos en su mayoría son optativos, siendo obligatorio únicamente el de Uso de las TIC en la enseñanza y aprendizaje de lenguas extranjeras de la Licenciatura en Lenguas Extranjeras con Énfasis en Inglés y Francés de la Universidad de Antioquia. Este aspecto es relevante dado que el no establecer una asignatura como obligatoria dentro de un plan de estudios puede hacer que menos estudiantes se vean motivados a inscribirla. No es un secreto que la obligatoriedad de una asignatura es un gran motivante a la hora de inscribir asignaturas, no solo se inscribe lo que se quiere ver, sino lo que se debe ver. No obstante, eso es una generalidad, y para poder validar ese postulado haría falta, por lo menos, una encuesta que permita determinar los motivos por los que un estudiante inscribe tal o cual asignatura. Sin embargo, la naturaleza no obligatoria de una asignatura también puede implicar que no tiene un papel principal o disciplinar en sus respectivos pregrados.

Por otro lado, remitiéndonos a la intensidad horaria también encontramos diversidad. Estas asignaturas se hayan dentro de un rango que se establece entre las 96 y las 192 horas. Sin embargo, estas intensidades de tiempo suelen estar distribuidas en dos mitades, una de trabajo presencial o en clase, es decir, asesoría directa por parte del docente, y otra de trabajo autónomo extracurricular. Las asignaturas con menor intensidad horaria son las de la Universidad de la Amazonia, y las de mayor intensidad horaria son las de la Universidad Nacional de Colombia. La asignatura de la Universidad de Antioquia se acerca bastante a la intensidad horaria de la UNAL.

En esta sección se presentó inicialmente una tabla sumaria que contiene información sobre las universidades, pregrados y asignaturas seleccionadas. Se aportó más información sobre las asignaturas seleccionadas y su presencia en sus respectivos pregrados. Se abordaron los semestres en los que se ofrecen, su valor en créditos académicos, su intensidad horaria y su calidad de optativa u obligatoria. Esto se hizo con el fin de tener un amplio panorama comparativo de cómo están integradas esas asignaturas en sus respectivos programas curriculares. A continuación, se presentará la tercera, y última sección de este capítulo, la cual se dedica con mayor detalle a las asignaturas que se consideraron propias de la Lingüística Computacional o con contenidos de esta área.

### **5.3 Valoración de los contenidos de las asignaturas seleccionadas**

En esta tercera sección del capítulo se presentarán las asignaturas que fueron seleccionadas como propias de la Lingüística Computacional o con contenidos de esta área. Para poder llevar a cabo la valoración que nos exige el tercer objetivo de esta

investigación se detallaran sus contenidos, sus estrategias pedagógicas o didáctica y esto dará paso a la valoración por medio de una matriz DOFA para cada programa curricular. Los documentos que se usaron como fuentes primarias fueron los micro currículos, descripciones de asignatura o contenidos programáticos de cada asignatura seleccionada. El orden en el que se presentarán las asignaturas está directamente relacionado con el orden en el que se presentaron los pregrados en cuestión. Se presentará el pregrado y posteriormente sus asignaturas seleccionadas. La matriz DOFA de cada asignatura irá al final de cada subsección. Finalmente, se hará un resumen final del capítulo y de los resultados presentados.

### **5.3.1 Licenciatura en Lenguas Extranjeras con Énfasis en Inglés**

En este pregrado perteneciente a la Universidad Católica de Oriente, y en él se encontró una asignatura que cumplía con los requisitos de búsqueda utilizados. La asignatura en cuestión es Razonamiento lógico matemático. Después de haberla incluido en la lista de las asignaturas de interés, se envió la carta formal de solicitud documental correspondiente. No obstante, nunca se obtuvo respuesta por parte de la universidad y no se pudo conseguir su micro currículo en línea. Por lo tanto, en este apartado no fue posible registrar información sobre esta asignatura.

### **5.3.2 Licenciatura en literatura y lengua castellana**

En este pregrado perteneciente a la Universidad de la Amazonia se encontraron dos asignaturas que cumplen con los criterios de selección. Estas dos asignaturas son Pensamiento Lógico Matemático y Tecnologías de la Comunicación Aplicadas a la Educación. Ambas pertenecen al mismo componente curricular, Fundamentos Generales. A su vez, ambas son asignaturas optativas que no son obligatorias de ser cursadas para graduarse de la carrera. Comparten, también, la intensidad horaria.

#### **5.3.2.1 *Pensamiento Lógico Matemático***

El documento que soporta a esta asignatura es su Micro Currículo. Su código es el 9900123 y no tiene prerrequisitos. Está pensada para que el futuro docente pueda integrar las herramientas para la resolución de problemas y el análisis de las que las matemáticas nos proveen. Se toma en cuenta que "...el razonamiento lógico está integrado a nuestro mundo cambiante, complejo y en permanente condición de incertidumbre" (Universidad de la Amazonia, 2019). En esta asignatura se aborda la *Lógica simbólica*, y de esta se dice que: "Como 'modelo', la Lógica Simbólica permite interpretar y comprender la realidad; como 'medio' favorece el desarrollo de capacidades analíticas, de creación, de interpretación y razonamiento, condiciones indispensables de todo profesional" (Universidad de la Amazonia, 2019a).

Cabe señalar que en el documento se indica que esta asignatura:

“...ha sido elaborado con la intención de proporcionar a los estudiantes que se inician en la carrera de Licenciatura en literatura y lengua castellana de la Universidad de La Amazonia, un conjunto de conocimientos y experiencias que le permitan reestructurar cognoscitiva y conceptualmente su bagaje lógico-matemático mediante la implementación de actividades y estrategias que favorezcan el razonamiento y el desarrollo de habilidades y destrezas para afrontar con éxito su formación profesional” (Universidad de la Amazonia, 2019a).

No obstante, dentro de la malla curricular, se oferta para octavo semestre, una etapa ya avanzada de la carrera. Dentro de las competencias específicas que se desarrollan en esta asignatura se encuentran dos que llamaron la atención. En primer lugar, utiliza adecuadamente el lenguaje y sus estructuras para plantear y resolver problemas del entorno que favorezcan el desarrollo del pensamiento lógico. En segundo lugar, descompone una situación compleja en elementos simples e identifica las relaciones entre estos e identifica relaciones causales y deduce conclusiones lógicas. Por otro lado, en lo que refiere a sus objetivos, se pueden señalar apropiar el razonamiento lógico y utilizar el lenguaje matemático como un elemento de comunicación. Adicionalmente, el curso busca proporcionar herramientas básicas como mecanismos de planteamientos de problemas de investigación, así como brindar elementos de razonamiento lógico frente a situaciones cotidianas.

Sus contenidos están organizados en siete unidades. Unidad 1: Simbolización y proposiciones, unidad 2: Valores de verdad de las proposiciones, unidad 3: Relaciones y leyes lógicas, unidad 4: Inferencia y deducción de la lógica proposicional, unidad 5: Lógica de predicados, unidad 6: Inferencias mediatas. El silogismo categórico y, unidad 7: Conceptos básicos de conjuntos. En cuanto a la metodología, se trabaja bajo un modelo constructivista, guiado por el Aprendizaje Basado en Problemas. Se parte de problemas elegidos entre maestro y alumnos a partir de la cotidianidad. La metodología está centrada en el alumno y su aprendizaje. Se recomiendan grupos de entre 5 y 8 alumnos durante el desarrollo de la asignatura. Finalmente, las herramientas didácticas que se contemplan para la asignatura son: “Portafolio pedagógico, conferencia magistral, talleres, lecturas, salidas de campo, laboratorios y exposición de resultados de proceso investigativo” (Universidad de la Amazonia, 2019a).

### **5.3.2.2 Tecnologías de la Comunicación Aplicadas a la Educación**

El documento que soporta a esta asignatura es su Micro Currículo. Su código es el 6008013. Sus prerrequisitos son Comunicación I y II. Al momento de establecer este curso dentro de la malla curricular se tiene en cuenta la importancia de mantenerse actualizado con los últimos avances tecnológicos. “...el estudiante de literatura y lengua castellana reconoce las herramientas digitales, el rol en los procesos de comunicación y su impacto en la educación” (Universidad de la Amazonia, 2023). De esta manera, se puede entender que esta asignatura tiene un fuerte enfoque crítico y pretende que el estudiante reflexione sobre el papel de las nuevas tecnologías, su importancia e implicaciones legales. El

---

estudiante comprende que “...lo digital cruza con lo humanístico y esta intersección transforma lo que entendemos como humano, humanidades o ciencias humanas” (Universidad de la Amazonia, 2023), por lo que afronta su aprendizaje en TIC con la idea de que “...la formación en las tecnologías permite analizar su impacto, reconocer las utilidades y las implicaciones legales que conlleva el uso de estas en el aula” (Universidad de la Amazonia, 2023).

La formación de esta asignatura se plantea en tres fases: exploratoria de búsqueda en bases de datos, analítica de las plataformas digitales en la educación y contexto de las implicaciones del uso de las TIC en el aula. En cuanto a la metodología, siempre se toman en cuenta tanto al PEI como al PEP. Se contempla el trabajo acompañado por el profesor y el trabajo independiente. Se busca formar para la investigación, en este sentido se espera que el estudiante interiorice y analice las herramientas digitales y recursos de los que dispone para su propia formación, así como para su quehacer docente. Algunas de las competencias que los estudiantes desarrollan son: reconocen el papel de las TIC, sus avances e impacto, analizan las implicaciones sociopolíticas de las TIC en la educación y, sugieren el uso de diferentes herramientas en la preparación, desarrollo y evaluación de procesos académicos.

En los contenidos de la asignatura se pueden encontrar las siguientes unidades temáticas: Tecnologías de la comunicación vs de la información, La comunicación y sus cambios por la tecnología, Tecnología en la educación. Min Tic, Habeas data y ciberseguridad, La cultura digital en los procesos educativos, El rol del docente de literatura y lengua frente a las nuevas tecnologías, Inteligencia artificial, entre otros. Todos estos contenidos permiten entender el horizonte inmediato que la asignatura contempla para sus espacios de formación, el papel y la importancia de la tecnología en la sociedad y en la práctica educativa. Se preocupa por las implicaciones legales y los riesgos que las TIC pueden suponer para la sociedad, en específico para los entornos académicos, para la comunidad académica. A continuación, se presentará la [Figura 4](#), la cual contendrá la matriz DOFA correspondiente al análisis global realizado sobre la Licenciatura en Literatura y Lengua Castellana.

**Figura 4. Matriz DOFA. Licenciatura en Literatura y Lengua Castellana**



Esta matriz representa el análisis que se llevó a cabo de las dos asignaturas que fueron seleccionadas de la Licenciatura en Literatura y Lengua Castellana de la Universidad de la Amazonia. Los aspectos que se abordan en las diferentes secciones de la matriz buscan señalar aspectos en los que se puede trabajar a futuro en una posible reforma o actualización del currículo. Sin embargo, se considera que es un currículo con gran enfoque a la actualidad, se reconoce el trabajo que se lleva a cabo en esas dos asignaturas y la formación que les dan a sus futuros docentes.

### 5.3.3 Licenciatura en Lenguas Extranjeras con Énfasis en Inglés y Francés

En este pregrado perteneciente a la Universidad de Antioquia se encontró una asignatura que cumplía con los criterios de selección. La asignatura en cuestión es Uso de las TIC en la enseñanza y aprendizaje de las lenguas extranjeras. A diferencia del resto de asignaturas que se presentan en esta sección, esta sí es obligatoria, lo cual le puede

---

aumentar el grado de éxito y garantiza que todos los egresados cumplirán con las competencias establecidas para la asignatura.

### **5.3.3.1 Uso de las TIC en la enseñanza y aprendizaje de las lenguas extranjeras**

Esta asignatura está soportada en su *Programa Oficial de Curso*. Su código es 1304037 y pertenece al componente de Pedagogía y didáctica. Tiene un prerrequisito: Tendencias en Enseñanza y Evaluación II Código 1304/03 0. Dentro del documento se especifica que este curso busca que el estudiante desarrolle una alta competencia lingüística y consciencia crítica con relación a las TIC, así como una integración efectiva de la tecnología en el quehacer docente. Adicionalmente, se busca que los alumnos analicen el impacto de la alfabetización digital en los ámbitos educativo y social.

Su programa está orientado por la pregunta ¿Cómo nos integramos los maestros a las comunidades académicas? Por su parte, dentro del desarrollo del curso, el docente fomenta "...desarrollar sus competencias digitales y generar materiales y recursos que inviten a la reflexión crítica y que promuevan el aprendizaje de las lenguas extranjeras, atendiendo así al desarrollo de los componentes digitales y ciudadanas de los maestros en formación" (Universidad de Antioquia, 2024). Adicionalmente, el documento señala que la asignatura tiene especial relevancia en el contexto de la post pandemia dado que, en los tiempos de la virtualidad obligada, se pudo evidenciar el potencial que la tecnología tiene para la educación.

Dentro de los objetivos específicos de la asignatura encontramos: Reconocer enfoques teóricos y didácticos, Diseñar objetos virtuales de aprendizaje, Diseñar propuestas metodológicas de integración de las TIC a diferentes propuestas curriculares de enseñanza de la lengua extranjera, Reflexionar críticamente acerca del papel que la tecnología puede jugar en los contextos educativo y social, y Explorar las posibilidades (affordances) del rol social de la tecnología y su impacto en el ámbito de la enseñanza y aprendizaje de lenguas extranjeras. Por otro lado, la asignatura ofrece un gran aporte a la formación de los estudiantes de este pregrado: Sensibilización sobre el uso razonable de las TIC, Promoción de la inclusión y la equidad, Preparación para la realidad educativa actual, Desarrollo de una ciudadanía digital responsable, Promoción de la innovación educativa, Desarrollo del pensamiento crítico y analítico, Promoción del aprendizaje autónomo, Preparación para la elaboración de proyectos de investigación, y Promoción de la cultura investigativa.

Los contenidos de la asignatura están organizados en tres grandes unidades. En primer lugar, los objetos virtuales de aprendizaje y su utilización en la enseñanza y aprendizaje de lenguas extranjeras. En segundo lugar, la producción de información en la web 2.0. Y, en tercer lugar, la moderación de plataformas de aprendizaje y de cursos virtuales utilizados para el aprendizaje y enseñanza de lenguas extranjeras. Como metodología se emplea el Aprendizaje Basado en Problemas, Aprendizaje Basado en

Proyectos, Aprendizaje invertido. Con el uso de estas metodologías de enseñanza, se busca que el estudiante tenga un papel crítico y proactivo frente a la tecnología, para lo cual se privilegia el “...uso instrumental y el uso didáctico y crítico de las diferentes herramientas tecnológicas exploradas” (Universidad de Antioquia, 2024). Lo anterior se plantea desde una mirada crítica e intercultural para la enseñanza y el aprendizaje de lenguas a partir de la pedagogía por proyectos. El uso de las TIC es amplio en el curso, razón por la cual existen repositorios digitales para alojar trabajos y proyectos de los estudiantes, así como herramientas para que ellos utilicen en su formación. A continuación, se presentará la Figura 5 que contiene el análisis global de la Licenciatura en Lenguas Extranjeras con Énfasis en Inglés y Francés.

**Figura 5. Matriz DOFA. Licenciatura en Lenguas Extranjeras con Énfasis en Inglés y Francés**



Esta matriz representa de forma visual los aspectos principales que arrojó como resultado el análisis de los documentos sobre la Licenciatura en Lenguas Extranjeras con Énfasis en Inglés y Francés y su asignatura de Uso de las TIC en la enseñanza y

aprendizaje de las lenguas extranjeras. Tiene aspectos en común con la matriz de la Figura 4, sin embargo, el componente ético y legal es, aparentemente, más amplio y profundo en esta asignatura. No obstante, ambas asignaturas dedicadas al uso de las TIC en la educación son muy similares y buscan que el egresado de sus respectivos programas no sea un ente pasivo de cara a las nuevas tecnologías de la información y la comunicación.

### **5.3.4 Pregrado en Filología e Idiomas y Pregrado en Español y Filología Clásica**

Estos dos pregrados pertenecen a la Facultad de Ciencias Humanas de la Universidad Nacional de Colombia. Si bien se cuentan entre las Ciencias del Lenguaje y dentro de sus mallas curriculares se puede encontrar un extenso componente dedicado a la lingüística, dada su naturaleza más específica y su orientación al campo de la educación dado que entregan título de licenciados, los contenidos de lingüística son menos profundos que los del pregrado en Lingüística de la misma universidad. Por lo tanto, al momento de hacer la búsqueda de programas curriculares y asignaturas no se pudo hallar ninguna que cumpliera con los requisitos establecidos. Las tres asignaturas que se van a abordar en el pregrado de Lingüística no están ofertadas, ni siquiera como electivas u optativas, en estos pregrados. Por supuesto que un estudiante podría verlas, pero para eso tendría que pasar por un proceso administrativo de solicitud de cupo posterior a los periodos regulares de inscripción de asignaturas. Lo anterior, por sí mismo, ya representa una barrera para que los estudiantes de estos pregrados puedan cursar estas asignaturas.

### **5.3.5 Lingüística**

El pregrado de Lingüística también pertenece a la Universidad Nacional de Colombia y, dentro de ella, a la Facultad de Ciencias Humanas. En este programa curricular se encontraron tres asignaturas que cumplían los requisitos de búsqueda y selección. Estas tres asignaturas son Lógica para lingüistas, Lingüística matemática e Informática aplicada a la investigación lingüística. Las tres hacen parte del componente de fundamentación de la carrera. Se presentarán en el mismo orden en el que aparecen en la malla curricular. Las tres asignaturas son optativas, por lo que no es necesario verlas para graduarse.

#### **5.3.5.1 Lógica Para Lingüistas**

Esta asignatura también tiene su soporte en los Contenidos Programáticos. Es una asignatura optativa, su código es 215987 y hace parte del componente de fundamentación. No tiene prerrequisitos. Si bien en un inicio puede parecer que este curso es el que debería gozar de un papel primario por encima de Lingüística Matemática, como de hecho sucede en la malla curricular, se considera que este debe ir después de la anteriormente expuesta ya que en Lingüística Matemática se da la primera aproximación a la lógica. Por otro lado, en este curso se profundiza ampliamente en la lógica, es más específico. Dentro de la descripción de esta asignatura inicialmente se contempla que la lógica ayuda a fortalecer

la visión crítica y analítica de los estudiantes. Por tal razón en ella se busca abordar la lógica proposicional y lógica de predicados. De igual manera, se busca abordar otras perspectivas como la lógica modal.

Se propone que los estudiantes descubran "...cómo puede ser aplicada la lógica y sus sistemas formales en el análisis y descripción de los fenómenos presentes en las lenguas naturales" (Bautista Peña, 2024b). Por otro lado, "El objetivo general del curso es favorecer el pensamiento analítico y crítico de los estudiantes a través del estudio de las herramientas que ofrece la lógica formal" (Bautista Peña, 2024b). Con respecto a los objetivos específicos se contemplan: comprender y aplicar los conceptos básicos de la lógica, como proposiciones, símbolos, conectores, etc.; familiarizarse con los diferentes sistemas formales; evaluar y construir argumentos utilizando estrategias lógicas; y, examinar la conexión entre la lógica y la lingüística.

La metodología que se emplea en esta asignatura es de carácter teórico práctico, buscando desarrollar en los estudiantes el procesamiento analítico. Para lograrlo se ha establecido que la mitad de la clase sea teoría y la otra mitad sea práctica. Por su parte, los contenidos que se han planteado para este curso son: Introducción a la lógica, Lógica proposicional, Teoría de conjuntos, Lógica de predicados y Temas adicionales (Lógica Modal y Teoremas de Gödel).

### **5.3.5.2 Lingüística Matemática**

Esta asignatura está soportada por sus *Contenidos Programáticos*. Su código es 2015988 y hace parte del Componente de Fundamentación de la carrera. No tiene prerrequisitos dado que no es una asignatura obligatoria. Este curso es fundamental para el abordaje de la LC, así como provee de herramientas procedimentales y de análisis valiosas para los otros dos cursos analizados. El curso busca "...acercar a los estudiantes a las herramientas matemáticas más comunes para ocuparse del estudio del lenguaje y, a su vez, pretende profundizar en el papel de la lógica matemática en los desarrollos de la teoría de autómatas y las gramáticas formales en aras de comprender su influencia en los desarrollos de la teoría lingüística en el siglo XX" (Arévalo, 2025). Para ello se enmarca en la profundización de problemáticas lógicas (consistencia, decidibilidad, completud) junto a otros avances que ha tenido la lingüística desde la segunda mitad del siglo XX hasta la actualidad.

El objetivo general de la asignatura es: "Capacitar a los estudiantes de lingüística en las herramientas lógico-matemáticas fundamentales para comprender la manera en que se replantean y resuelven algunos problemas lingüísticos a la luz de ellas" (Arévalo, 2025). Por otra parte, tiene tres objetivos específicos que dictan la forma en que se desarrolla el curso. En primer lugar, profundizar en los desarrollos de la lógica matemática. En segundo lugar, evaluar de manera crítica la noción de gramaticalidad. Y, en tercer lugar, analizar la importancia de la teoría de autómatas y la computabilidad para el desarrollo de los lenguajes de programación, los estudios sobre procesamiento de lenguaje natural y la

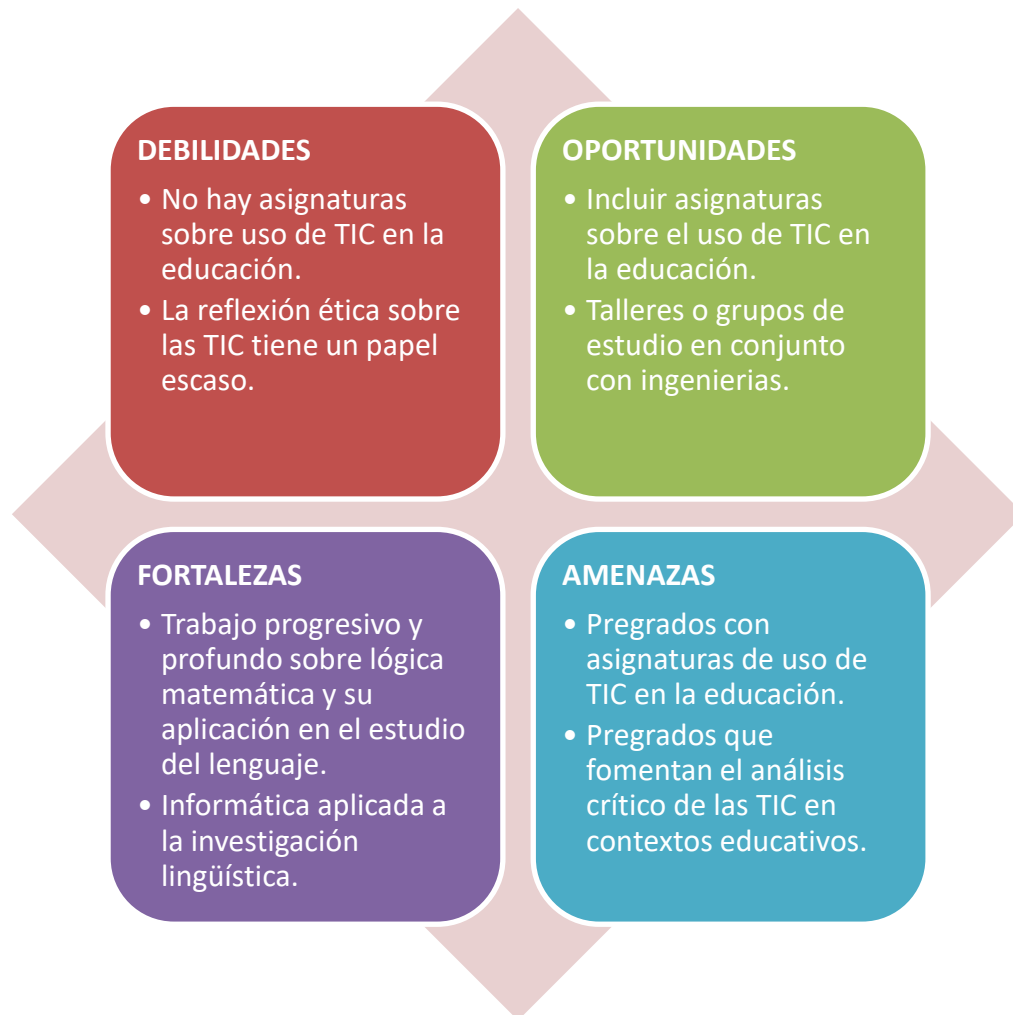
inteligencia artificial. Por otro lado, los contenidos del curso giran en torno a conceptos matemáticos fundamentales, desarrollos matemáticos del estudio del lenguaje, gramática generativa chomskiana, y otros campos de estudio.

### **5.3.5.3 *Informática Aplicada a la Investigación Lingüística***

Esta asignatura tiene como código el 215985, pertenece al componente de fundamentación del pregrado de Lingüística y es optativa. No tiene prerrequisitos, sin embargo, en este curso se concentran y se aplican los conocimientos que se adquieren en los dos cursos anteriores, razón por la cual el autor de este trabajo sí considera que se podría llegar a establecer alguno de los dos anteriores cursos como prerrequisito para esta asignatura. En esta “Se buscará que los estudiantes adquieran una comprensión sólida de los principios de la informática, enfatizando en programación, mientras que exploran sus aplicaciones en la investigación lingüística” (Bautista Peña, 2024a). Adicionalmente, se abordan contenidos como programación básica, procesamiento de textos, Procesamiento del Lenguaje Natural (PLN) y análisis de datos mediado por computadores.

En este mismo documento de Contenidos Programáticos (Bautista Peña, 2024a) se establecen cinco objetivos específicos. En primer lugar, adquirir conocimientos en los principios básicos de la informática y la programación. En segundo lugar, desarrollar habilidades en programación empleando Python y R para la manipulación de datos lingüísticos y no lingüísticos. En tercer lugar, desarrollar habilidades de pensamiento crítico y aplicado a la resolución de problemas que involucren conjuntos de datos complejos. En cuarto lugar, aplicar los conocimientos de programación para resolver problemas lingüísticos. Finalmente, en quinto lugar, fortalecer la formulación y solución de problemas de investigación enmarcados principalmente en el paradigma cuantitativo.

Por otro lado, se señala también el enfoque teórico práctico de la asignatura. Adicionalmente, se indican los contenidos de la asignatura, los cuales son seis. en primer lugar, Introducción a la informática. En segundo lugar, Lingüística de corpus. En tercer lugar, Procesamiento del lenguaje natural (PLN). En cuarto lugar, Programación para lingüistas (Python). En quinto lugar, Programación para lingüistas (R y R Studio). Y, finalmente, Informática y ética. En la [Figura 6](#) se presenta la matriz DOFA que representa el análisis realizado en esta sección.

**Figura 6. Matriz DOFA. Lingüística.**

La Figura 6 representa de forma gráfica y sintética el análisis que se realizó de las tres asignaturas que fueron elegidas dentro del pregrado de Lingüística de la Universidad Nacional de Colombia. Estas asignaturas fueron Lógica para Lingüistas, Lingüística Matemática e Informática Aplicada a la Investigación Lingüística. Se destaca el carácter progresivo que tienen las tres asignaturas dado que se complementan de excelente manera. Las dos primeras son fundamentales para el trabajo de la última y entre todas constituyen una base ideal para los estudiantes que deseen dedicarse a la Lingüística Computacional, así como para aquellos que no estén del todo interesados en este campo.

En este capítulo de resultados se hizo un repaso por la información que arrojó el análisis de los documentos reunidos y registrados en el corpus. Se abordaron por separado los pregrados incluidos en el estudio, la presencia de la Lingüística Computacional en estos y las asignaturas seleccionadas. Los resultados permiten identificar que, por lo menos, en los pregrados analizados hay dos líneas principales en las que se abordan estos contenidos transversales y fundamentales para la LC. En primer lugar, tenemos la línea

---

del pensamiento lógico y matemático, en la cual se adquieren conocimientos y formación básica en las lógicas mediante las cuales trabajan las computadoras. Es una etapa que no se puede ignorar si se quiere actualizar un currículo para que esté a la vanguardia de la LC y las TIC. Además, es un conocimiento importante a la hora de ingresar al mundo de la programación. En segundo lugar, se pudo hallar un enfoque más orientado al uso, presencia, importancia, consideraciones e implicaciones de las TIC en la sociedad y la educación. Este enfoque es de gran relevancia en las licenciaturas, por lo que no es una sorpresa encontrarlo en la mayoría de los currículos que se revisaron pertenecientes a licenciaturas. Sin embargo, con más razón, llama la atención que no haya una asignatura dedicada a este tema en los pregrados de Filología e Idiomas y de Español y Filología Clásica.

Adicionalmente, es necesario señalar que en todos los PEP de los pregrados que fueron seleccionados para formar parte de esta investigación se menciona la flexibilidad curricular como un elemento positivo de sus respectivos currículos. Con flexibilidad horaria, por lo general, hacen referencia a las asignaturas electivas. Las asignaturas electivas son concebidas como un elemento del currículo que les permite a los estudiantes orientar su formación hacia intereses personales que pueden estar relacionados con otras facultades o planes de estudio. Es por eso por lo que, por ejemplo, la Universidad Nacional de Colombia exige que el 20% de los créditos de cualquier programa curricular sean asignaturas electivas. Esta flexibilidad puede parecer muy positiva en un inicio, sin embargo, cuando se ve reflejada en una malla curricular con una ruta de progreso difícil de identificar, los estudiantes pueden sentirse frustrados y desorientados. Esto puede llevar a abandono escolar, cancelación de asignaturas o falta de cupos.

Dentro de la información recolectada como resultados se pudo evidenciar una evidente postura crítica con relación a las TIC, la inteligencia artificial y la responsabilidad cibernética. Abordar estas perspectivas es de gran relevancia para un docente en formación si se espera que no solo utilice herramientas computacionales para complementar levemente sus entornos de enseñanza. El uso de las TIC tiene implicaciones legales y éticas, más allá de lo que cabría imaginar. Por lo tanto, es de gran importancia que un licenciado en formación aborde las TIC desde diferentes ángulos, que existan espacios destinados a la discusión y reflexión en torno a estas, pero también a la creación de materiales, a la innovación educativa de la mano con las nuevas tecnologías. Para poder lograr semejante propósito se considera que la implementación de talleres, grupos de investigación, seminarios, y asignaturas electivas interdisciplinarias abiertas a otras facultades que aborden estos contenidos básicos referentes a la Lingüística Computacional son una opción que podría contribuir al trabajo interdisciplinario y a la interiorización de estas herramientas tecnológicas.

Específicamente se podría decir que los únicos cursos dedicados estrictamente a contenidos directamente relacionados con la LC son los de pensamiento lógico matemático, lingüística matemática e informática aplicada a la investigación lingüística. Este último es, de hecho, uno de los campos de acción de la LC. Estos cursos les brindan

---

a los estudiantes conocimientos y habilidades básicas para poder involucrarse posteriormente en aplicaciones más avanzadas de la LC o el PLN, aplicaciones que implican el uso de herramientas computacionales igualmente avanzadas. Adicionalmente, debe recordarse que los programas seleccionados y encontrados en la recolección de datos giran en torno al lenguaje, a la lengua española y las lenguas extranjeras o lenguas modernas. Por lo tanto, tienen un considerable componente dedicado a la lingüística en sus mallas curriculares, lo que pretende generar formación lingüística elemental en estos futuros docentes.

En efecto, un licenciado egresado de estos pregrados no puede tener, inicialmente, una formación lingüística tan profunda y detallada como un egresado de Lingüística pura. Sin embargo, al estar relacionados con la educación, la pedagogía y didáctica de las lenguas, el papel de las TIC en su quehacer como docentes investigadores, docentes investigan sobre la enseñanza y el aprendizaje de las lenguas, sus estudiantes no pueden dejar de lado las nuevas tecnologías, los avances que surgen diariamente en materia de aplicaciones, programas, plataformas, inteligencia artificial, entre otros. Tal vez haya menos presencia de la esperada o de la deseada de la LC en los pregrados de Ciencias del Lenguaje en el país, aunque esta es una investigación exploratoria cualitativa, razón por la cual no puede aportar grandes conclusiones ni cifras. Sin embargo, lo que se ha podido identificar es que los pregrados de Ciencias del Lenguaje están a la vanguardia, en constante mejora y búsqueda por no quedar atrás de cara a los avances de las nuevas tecnologías.

---

## 6 Conclusiones

Al finalizar el capítulo anterior ya hemos presentado puntos importantes que serán útiles para las conclusiones de este trabajo. De acuerdo con el objetivo general de este trabajo de grado “Analizar la presencia de la Lingüística Computacional en los programas de pregrado del área de Ciencias del Lenguaje en Colombia”, a continuación, se presentarán las conclusiones a las que se llegó en este trabajo de grado.

Con respecto al primer objetivo específico; “Identificar los programas de pregrado del área de ciencias del lenguaje que incluyen la Lingüística Computacional en sus planes de estudio”, se identificaron alrededor de veintidós programas de pregrado. En su mayoría fueron licenciaturas en literatura, lengua castellana, lenguas extranjeras y lenguas modernas. El único pregrado dedicado a la lingüística pura fue el de Lingüística de la Universidad Nacional de Colombia. De estos veintidós pregrados solo se lograron obtener los documentos (Proyecto Educativo del Programa, Malla Curricular y Micro Currículo) de seis de ellos. De los seis pregrados, tres fueron externos a la Universidad Nacional de Colombia (Licenciatura en Lenguas Extranjeras con Énfasis en Inglés, Licenciatura en Lenguas Extranjeras con Énfasis en Inglés y Francés, Licenciatura en Literatura y Lengua Castellana) y tres internos (Filología e idiomas: Alemán, Inglés y Francés, Español y Filología Clásica, y Lingüística).

En su mayoría, este objetivo se cumplió adecuadamente. Tomando en cuenta que se trata de un trabajo guiado por el enfoque cualitativo, los datos recopilados no se vieron representados en la forma de un gran corpus con miles de entradas y mediciones estadísticas. Por supuesto, esto podría hacerse en un futuro trabajo investigativo que pretenda ahondar en una medición precisa de lo que se ha planteado en este trabajo de grado. Incluso se podrían implementar herramientas como la que se presenta en el artículo de Chávez y Martín Yamamoto (2014). Las universidades fueron seleccionadas a partir de criterios objetivos, más no de manera aleatoria. Se tuvo siempre claro el tipo de programa curricular que se necesitaba incluir en este trabajo de grado.

En lo que respecta al segundo objetivo específico; “Comparar la presencia de la Lingüística Computacional entre diferentes programas de diversas universidades”, de los seis pregrados analizados —Licenciatura en Lenguas Extranjeras con Énfasis en Inglés, Licenciatura en Lenguas Extranjeras con Énfasis en Inglés y Francés, Licenciatura en Literatura y Lengua Castellana, Filología e idiomas: Alemán, Inglés y Francés, Español y Filología Clásica, y Lingüística—, solo el pregrado en Lingüística de la Universidad Nacional de Colombia se puede considerar como un pregrado plenamente dedicado a la

---

lingüística. Los demás programas tienen un marcado enfoque en la docencia, razón por la cual los contenidos de lingüística que se abordan en sus currículos son elementales, básicos y con enfoque a la lengua o lenguas que se abordan en la carrera.

Por lo tanto, los contenidos de lingüística computacional que se incluyen en la mayoría de estos pregrados están orientados, en su mayoría, al uso de las TIC en la educación. En estos cursos dedicados a las TIC, se otorga relevancia a la reflexión en torno al papel, la importancia, las implicaciones, las consecuencias, los aspectos legales y el potencial de las TIC en la educación. Suelen ser asignaturas con fuertes componentes de análisis crítico, responsabilidad social, uso responsable de las TIC, innovación en el uso de las TIC en la educación y que promueven la reflexión en torno a las TIC tanto en la formación del docente como en su futuro trabajo como educador. Por otro lado, se encontraron asignaturas dedicadas a la LC en el pregrado en Lingüística de la Universidad Nacional de Colombia; en los tres cursos que fueron seleccionados de este pregrado, se puede evidenciar una clara ruta de progreso que está orientada a una formación básica, pero fundamental en los conceptos de la LC. Se considera que estos cursos pueden constituir una base formidable para los estudiantes que quieran decantarse por esta subdisciplina de la lingüística.

Este objetivo también se cumple mayormente. Dado que se pretendía comparar la presencia de la lingüística computacional en los pregrados seleccionados, y tomando en cuenta que de estos pregrados hubo dos de los que no se seleccionó ninguna asignatura, se pudo establecer una comparación apropiada. Esta comparación no se limitó a si estaba o no presente la lingüística computacional en estos pregrados, sino que también abordó la forma en la que estaba presente. De ahí que se estableciera esa comparación entre las asignaturas dedicadas a las TIC y las asignaturas dedicadas a la lingüística computacional propiamente dicha.

En lo referente al tercer objetivo específico; “Valorar los contenidos y profundidad con que se aborda la Lingüística Computacional en los currículos revisados”, se halló que los contenidos de LC se abordan con mayor profundidad en las tres asignaturas seleccionadas del pregrado en Lingüística de la Universidad Nacional de Colombia. En Lógica para Lingüistas, se sientan las bases de la lógica proposicional, constituye el punto de partida para abordar programas y algoritmos utilizados en la investigación de la LC. En Lingüística Matemática, se ponen en práctica estas bases de lógica. Se abordan aplicaciones matemáticas en el estudio del lenguaje y desarrollos matemáticos que son elaboraciones más complejas creadas a partir de esas bases lógicas. Finalmente, en Informática Aplicada a la Investigación Lingüística, se integran los conocimientos aprendidos en los dos anteriores cursos y se ponen en práctica a partir de los programas y desarrollos tecnológicos que se utilizan en el procesamiento del lenguaje natural. Se aborda desde la programación básica hasta la ética del trabajo con herramientas digitales.

Este objetivo se cumplió a cabalidad. Es importante tener en cuenta que los micro currículos, descripciones del programa o contenidos programáticos no son una descripción completa en forma de secuencia didáctica. Mientras que en una secuencia didáctica o en

---

un proyecto, como aquellos presentados en el Marco Teórico como la Pedagogía Basada en Proyectos y la Pedagogía por Proyectos, se presenta una secuencia de actividades detalladas en la que se especifican las sesiones, actividades a realizar, objetivos y competencias, en estos documentos descriptivos de cursos se hallaron principalmente los objetivos, justificaciones y listados de temas a tratar en cada asignatura.

Por otro lado, en los PEP y micro currículos revisados se pudo hallar presencia de pedagogías que promueven el trabajo autónomo a partir de los intereses de los alumnos, así como aquellos que trabajan a partir de preguntas de investigación y no solo del modelo catedrático. Dentro de estos modelos, algunos de los mencionados fueron el aprendizaje por problemas, el aprendizaje por proyectos, el aprendizaje participativo, etc. Esto coincide con lo visto cuando se presentó el artículo de Alm et al. (2016). Esto permitió evidenciar que, en efecto, para poder abordar temas referentes a la LC, que requieren de un considerable componente práctico, optar por pedagogías que le permitan al alumno ser activo en su proceso de formación es recomendable. Se considera que sería de gran ayuda incluir o adoptar elementos de modelos como Aprendiendo con Nooj, presentado en el texto de Ferramondo (2021), en estas asignaturas.

En esta investigación se descubrió que los programas de Ciencias del Lenguaje están, en efecto, preocupados por el papel y las implicaciones que las nuevas tecnologías tienen en la formación y futuro ejercicio profesional de los estudiantes. Tomando en cuenta que gran parte de los pregrados seleccionados fueron licenciaturas, no es de extrañar que se privilegie a las TIC en la educación por encima de la LC. Por otro lado, la reflexión se plantea desde el punto de vista ético y de responsabilidad social y legal sobre el uso de la información. La seguridad informática es importante y estos cursos sirven para dar lugar a esa discusión. También se evidenció que se le da importancia a la creación de Objetos Virtuales de Aprendizaje, así como de didácticas que implementen las TIC de forma activa en los entornos educativos. Por último, se considera que los pregrados de Filología e Idiomas y Español y Filología Clásica deben incluir asignaturas dedicadas o bien a las TIC en la educación o bien a la LC, pero dado que otorgan el título de licenciado, puede que lo más recomendable sea el de TIC, con adaptaciones específicas a las necesidades de cada pregrado, por ejemplo la inclusión de Humanidades Digitales en el pregrado de Español y Filología Clásica, dado su alto contenido humanístico y su interés por los textos, lenguas y culturas clásicas.

Finalmente, se hace necesario valorar las contribuciones que este trabajo de grado deja a la comunidad académica. Para lograrlo, no podemos ignorar la naturaleza investigativa de esta investigación. Tomando en cuenta que se trató de una investigación de carácter exploratorio con la que el autor pretendía dar un primer vistazo a la situación de la lingüística computacional en la educación superior en Colombia, las contribuciones no se pueden establecer desde el punto de vista teórico o metodológico. No obstante, se contemplan otro tipo de contribuciones.

Dentro de las contribuciones se identifican primero aquellas de carácter empírico puesto que como resultado quedó el análisis de los pregrados y asignaturas seleccionados

---

a partir de las fuentes oficiales con información detallada sobre estos. De la mano con lo anterior, se establecen contribuciones de conocimiento. Se recogió la información más importante de los documentos abordados y que tenía relación con la lingüística computacional. Se espera que se desarrollen futuras investigaciones de naturaleza más amplia en la que se pueda abordar mediante metodologías cuantitativas o mixtas un mayor número de instituciones educativas, pregrados y asignaturas de diversas ubicaciones geográficas para poder obtener información más abundante y precisa.

## A. Anexo: Script para extraer hipervínculos de las universidades

El propósito de este script es ingresar a la página web [https://www.althillo.com/universidades/universidades\\_colombia.asp](https://www.althillo.com/universidades/universidades_colombia.asp) y extraer los hipervínculos que están enlazados a cada nombre de universidad. Una vez hecha la extracción debe generar un archivo .csv que contenga ese listado de enlaces. El script está escrito en el lenguaje de programación Python con ayuda de Copilot, el asistente de inteligencia artificial propiedad de Microsoft.

```
import requests
from bs4 import BeautifulSoup
import csv
from urllib.parse import urljoin # Útil para construir URLs completas

# Descargar el contenido de la página
url = "https://www.althillo.com/universidades/universidades_colombia.asp"
response = requests.get(url)
soup = BeautifulSoup(response.content, "html.parser")

# Criterios de búsqueda para instituciones en Colombia
criterios = [
    "universidad", "universitaria", "institución",
    "corporación", "fundación", "politécnico",
    "escuela", "instituto", "centro"
]

# Extraer las universidades o instituciones de educación superior
instituciones = []

for link in soup.find_all("a"):
    text = link.text.strip()
    href = link.get("href")
    if text and href:
        for criterio in criterios:
            if criterio in text.lower(): # Ajusta el filtro según los
criterios
```

```
        full_href = urljoin(url, href) # Construir el enlace
completo si es relativo
        instituciones.append((text, full_href))
        break # Detiene la iteración de criterios para evitar
duplicados

# Guardar las instituciones en un archivo CSV
output_path = "Resultados/Universidades/univ_Col.csv"
with open(output_path, mode="w", newline="", encoding="utf-8") as file:
    writer = csv.writer(file)
    writer.writerow(["Nombre de la Institución", "Enlace"]) # Encabezados
de las columnas
    writer.writerows(instituciones) # Escribir las instituciones en el
archivo

print(f"Los datos han sido guardados en '{output_path}'.")
```

## B. Anexo: Script para extraer la información de las universidades

Este script tiene como propósito ingresar a cada hipervínculo proporcionado por el script del anexo B y extraer los siguientes datos: nombre de la universidad, nombre de la facultad, nombre del pregrado y nivel de formación. Cabe señalar que este script extrajo la información de todas las universidades y pregrados que contenían los hipervínculos. Es por eso que posteriormente sería necesario filtrar esa información. El script está escrito en el lenguaje de programación Python con ayuda de Copilot, el asistente de inteligencia artificial propiedad de Microsoft.

```
import requests
from bs4 import BeautifulSoup
import csv
from urllib.parse import urljoin

# Leer los enlaces de "univ_Col.csv"
input_file = "Resultados/Universidades/univ_Col.csv"
output_file = "Resultados/Univ_Fac_Carr/univ_fac_carr_Col_Modificado.csv"

# Crear un archivo CSV para almacenar los datos recopilados
with open(output_file, mode="w", encoding="utf-8", newline="") as outfile:
    writer = csv.writer(outfile)
    writer.writerow(["Universidad", "Facultad", "Carrera", "Nivel de
Formación"]) # Agregar columna de nivel de formación

# Leer los enlaces desde el archivo de entrada
with open(input_file, mode="r", encoding="utf-8") as infile:
    reader = csv.reader(infile)
    next(reader) # Saltar los encabezados

# Iterar sobre cada universidad/enlace
for row in reader:
    universidad_nombre = row[0].strip() # Nombre de la universidad
    universidad_url = row[1].strip() # Enlace de la universidad

# Realizar solicitud HTTP a la URL de la universidad
```

```
try:
    response = requests.get(universidad_url)
    soup = BeautifulSoup(response.content, "html.parser")

    # Buscar el bloque que contiene facultades y carreras
    content = soup.find("blockquote")
    current_facultad = None
    current_nivel = None # Nivel de formación actual

    if content:
        lines = content.get_text(separator="\n").split("\n") #
Dividir texto por líneas
        for line in lines:
            line = line.strip()

            # Ignorar líneas vacías
            if not line:
                continue

            # Identificar niveles de formación
            if line.lower() in [
                "pregrado", "pregrados", "especialización",
"especializaciones",
                "maestría", "maestrías", "doctorado",
"doctorados", "posgrado", "posgrados",
                "postgrados", "postgrado", "pre-grado", "pre-
grados", "post-grado",
                "post-gradados", "pre grados", "pos grados",
"post grados", "pre grado",
                "pos grado", "post grado"
            ]:
                current_nivel = line # Actualizar el nivel de
formación
                continue # Pasar a la siguiente línea para no
registrar como carrera

            # Identificar facultades
            if line.startswith("Facultad de"):
                current_facultad = line # Actualizar la
facultad
                continue # Pasar a la siguiente línea

            # Registrar carreras válidas incluso si no hay
facultades explícitas
```

---

```
        writer.writerow([universidad_nombre,
current_facultad or "No especificado", line, current_nivel])
    except Exception as e:
        print(f"Error al procesar {universidad_nombre}: {e}")

print(f"¡Datos extraídos y guardados en '{output_file}'!")
```

## C. Anexo: Script para extraer la información específica de los pregrados de ciencias del lenguaje

El propósito de este script era el de tomar los datos generados por el script del Anexo C y de ellos extraer solo la información referente a los pregrados de ciencias del lenguaje. Debía generar un archivo .csv que contuviera dicha información. Sin embargo, para poder trabajar cómodamente en Excel era necesario convertir ese archivo a .xlsx. Para eso se utilizó un pequeño script que no es necesario incluir en ningún anexo. El script está escrito en el lenguaje de programación Python con ayuda de Copilot, el asistente de inteligencia artificial propiedad de Microsoft.

```
import pandas as pd
import os

# Cargar el archivo CSV original
archivo_csv = 'Resultados/Univ_Fac_Carr/univ_fac_carr_Col.csv'

try:
    # Leer el archivo CSV
    datos = pd.read_csv(archivo_csv)

    # Limpiar espacios en las columnas relevantes
    datos['Carrera'] = datos['Carrera'].str.strip()
    datos['Nivel de Formación'] = datos['Nivel de Formación'].str.strip()

    # Definir palabras clave positivas (ciencias del lenguaje y afines)
    palabras_clave = [
        'filología', 'lingüística', 'idiomas', 'lenguaje',
        'traducción', 'literatura', 'humanidades', 'lenguas',
        'español', 'clásica', 'inglés', 'francés', 'alemán'
    ]

    # Definir palabras clave negativas (programas de posgrado y afines)
    palabras_excluir = [
```

```
'posgrado', 'postgrado', 'especialización', 'maestría',  
'doctorado', 'diplomado', 'magister', 'máster',  
'post-grado', 'pos grado', 'pregrado superior',  
'especializaciones', 'doctorados', 'masterado'  
]  
  
# Crear filtro para incluir carreras relacionadas con ciencias del lenguaje  
filtro_lenguaje = datos['Carrera'].str.contains('|'.join(palabras_clave),  
case=False, na=False, regex=True)  
  
# Crear filtro para excluir programas de posgrado  
filtro_excluir_carrera =  
~datos['Carrera'].str.contains('|'.join(palabras_excluir), case=False,  
na=False, regex=True)  
filtro_excluir_nivel = ~datos['Nivel de  
Formación'].str.contains('|'.join(palabras_excluir), case=False, na=False,  
regex=True)  
  
# Combinar filtros  
filtro_combined = filtro_lenguaje & filtro_excluir_carrera &  
filtro_excluir_nivel  
  
# Aplicar el filtro al DataFrame  
carreras_filtradas = datos[filtro_combined]  
  
# Crear la carpeta de destino si no existe  
carpeta_destino = 'Resultados/Prog_Filtr'  
os.makedirs(carpeta_destino, exist_ok=True)  
  
# Guardar los resultados en un archivo CSV  
archivo_salida = os.path.join(carpeta_destino,  
'Prog_Filtr_Col_Leng_No_Posgrado.csv')  
carreras_filtradas.to_csv(archivo_salida, index=False)  
print(f"Archivo '{archivo_salida}' generado con éxito.")  
except FileNotFoundError:  
    print(f"El archivo '{archivo_csv}' no se encontró. Verifica la ruta y el  
nombre del archivo.")  
except KeyError:  
    print("El archivo CSV no contiene las columnas requeridas ('Carrera' o  
'Nivel de Formación'). Verifica el formato del archivo.")  
except Exception as e:  
    print(f"Ocurrió un error: {e}")
```

## D. Anexo: Procedimiento de recolección de información

Este diagrama de flujo representa de forma gráfica el procedimiento que se siguió para la recolección de la información que fue utilizada para la construcción del corpus. Representa la etapa de búsqueda en línea de las universidades, la automatización de la recolección de datos sobre universidades y carreras mediante una serie de scripts en Python. Constó de un proceso cíclico de perfeccionamiento de los scripts en el que se iba alimentando a Copilot con los inputs necesarios para crear el script para luego ejecutarlo y verificar su funcionamiento. Este proceso debía repetirse hasta que el script funcionara correctamente. Posteriormente, representa la normalización y limpieza de esa información, así como la adición de nueva información de contacto de los respectivos programas para el posterior envío de cartas formales para la solicitud de documentos institucionales. Seguidamente, incluye el envío de las cartas y la recepción de la documentación solicitada a las universidades. Para poder visualizar el diagrama de flujo dirigirse al siguiente enlace: [https://unaledu-my.sharepoint.com/:b:/g/personal/crcarvajalr\\_unal\\_edu\\_co/EZQqEjipU7tNh9e4Z0YtNjgBEAx8C6hcvdfR94u6bb6Dlw?e=zjBIJe](https://unaledu-my.sharepoint.com/:b:/g/personal/crcarvajalr_unal_edu_co/EZQqEjipU7tNh9e4Z0YtNjgBEAx8C6hcvdfR94u6bb6Dlw?e=zjBIJe)

## Bibliografía

Alm, C. O., Womack, K., Haake, A., & Engström, T. (2016). A Pedagogical Model for Computational Linguistics across Curricular Boundaries. *Language and Linguistics Compass*, 10(7), 335-345. <https://doi.org/10.1111/lnc3.12195>

Arévalo, V. A. (2025). *Contenidos Programáticos. Lingüística matemática*. Universidad Nacional de Colombia.

Baldrige, J., & Erk, K. (2008). Teaching Computational Linguistics to a Large, Diverse Student Body: Courses, Tools, and Interdepartmental Interaction. En M. Palmer, C. Brew, & F. Xia (Eds.), *Proceedings of the Third Workshop on Issues in Teaching Computational Linguistics* (pp. 1-9). Association for Computational Linguistics. <https://aclanthology.org/W08-0201/>

Bautista Peña, L. F. (2024a). *Contenidos Programáticos. Informática aplicada a la investigación lingüística*. Universidad Nacional de Colombia.

Bautista Peña, L. F. (2024b). *Contenidos Programáticos. Lógica para lingüistas*. Universidad Nacional de Colombia.

Chávez, B. L., & Martín Yamamoto, J. (2014). Análisis de contenido y lingüística computacional: Su rapidez, confiabilidad y perspectivas. *Anales de Psicología*, 30(3), 1146-1150. <https://doi.org/10.6018/analesps.30.3.154931>

Consejo Superior Universitario - Universidad Nacional de Colombia. (2025). *ACUERDO 033 DE 2007 [ACUERDO 033 DE 2007]. SISTEMA DE INFORMACIÓN NORMATIVA, JURISPRUDENCIAL Y DE CONCEPTOS «RÉGIMEN LEGAL»*. [https://legal.unal.edu.co/rlunal/home/doc.jsp?d\\_i=34245](https://legal.unal.edu.co/rlunal/home/doc.jsp?d_i=34245)

Cortes Vásquez, A., Vega Huerta, H., & Pariona Quispe, J. (2009). Procesamiento del Lenguaje Natural. *Revista de Ingeniería de Sistemas e Informática*, 6(2), 45-54.

Domínguez Burgos, A. (2002). Lingüística Computacional: Un Esbozo. *Botletín de Lingüística*, 18, 104-119. <https://www.redalyc.org/articulo.oa?id=34701805>

Ferramondo, C. (2021). *La consolidación de un nuevo enfoque pedagógico en "Aprendo con NooJ: De la lingüística computacional a la enseñanza de la lengua"*. <http://hdl.handle.net/2133/20276>

Ferrero, P., & Alda, J. (2007). *Aplicación de herramientas de lingüística computacional en foros virtuales*. Editorial Complutense. <https://hdl.handle.net/20.500.14352/53730>

Gómez Guinovart, X. (2000). Lingüística Computacional. En *Manual de ciencias del lenguaje* (1. Ed.). Edicións Xerais de Galicia. <https://uwm.edu/spanish-portuguese/wp-content/uploads/sites/216/2015/04/ManualDeCienciasDaLinguaxe.pdf#page=214.29>

González Fernández, A. (2018). Aproximación a la Lingüística Computacional y sus herramientas para el trabajo con corpus. *Futhark. Revista de Investigación y Cultura*, 13, Article 13. <https://revistascientificas.us.es/index.php/futhark/article/view/15938>

González Gil, E., Aparicio Gómez, O. Y., Ostos Ortiz, O. L., Barajas Frutos, M., Alcaraz Domínguez, S., Rodríguez Illera, J. L., Kaechele Obreque, M., Forés Miravalles, A., Cano García, E., Pons Seguí, L., Olabe Basogain, M. Á., Olabe Basogain, J. C., Basogain Olabe, X., & Quintana Albalat, J. (2020). *Innovación educativa y gestión del conocimiento*. Universidad Santo Tomás. <http://hdl.handle.net/11634/23837>

Islas Torres, C. (2017). La implicación de las TIC en la educación: Alcances, Limitaciones y Prospectiva. *RIDE. Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, 8(15), 861-876. <https://doi.org/10.23913/ride.v8i15.324>

Jurado Valencia, F. (2014). La pedagogía por proyectos vs. La pedagogía según programas estandarizados. 9. <https://rutamaestra.santillana.com.co/la-pedagogia-por-proyectos-vs-la-pedagogia-segun-programas-estandarizados/>

Llach, M. J. A., Galindo, J. C. M., Rodríguez, O. G. R., Gallini, S., Mahecha, T. M., & Liévano, J. N. J. (2020). Humanidades Digitales “a lo colombiche”: Cadáver exquisito de la Red Colombiana de Humanidades Digitales. *Revista de Humanidades Digitales*, 5, 217-235. <https://doi.org/10.5944/rhd.vol.5.2020.27837>

Moreno Sandoval, A. (1998). *Lingüística Computacional. Introducción a los modelos simbólicos, estadísticos y biológicos*. Síntesis.

Olabe Basogain, M. Á., Olabe Basogain, J. C., & Basogain Olabe, X. (2020). Enseñanza del pensamiento computacional: Metodologías y tecnologías educativas. En *Innovación educativa y gestión del conocimiento*. Universidad Santo Tomás. <http://hdl.handle.net/11634/23837>

Pareja Lora, A. (2024). Educación del pensamiento computacional para alumnos de un posgrado semipresencial en Humanidades: Experiencias con clase invertida / Computational Thinking Education for Students of a Post- Graduate Degree in the Humanities: Flipped Classroom Experiences. In *Journal: Propósitos y Representaciones*, e-ISSN 2310-4635. *Propósitos y Representaciones*, 8(1). <http://dx.doi.org/10.20511/pyr2020.v8n1.439>

*Portal MEN - Presentación—Ministerio.* (2025). Portal MEN - Presentación. <https://www.mineducacion.gov.co/1780/w3-propertyvalue-51454.html>

Rueda Vera, G., Luna, H., & Avendaño Castro, W. R. (2021). Estudiantes de pregrado frente a la educación mediada por TIC: Percepciones en contextos de pandemia. *Revista Boletín Redipe*, 10(8.(2021)), 321-339. <https://doi.org/10.36260/rbr.v10i8.1408>

Universidad Católica de Oriente. (2019). *Proyecto Educativo del Programa. Licenciatura en Lenguas Extranjeras con Énfasis en Inglés*. Universidad Católica de Oriente.

Universidad de Antioquia. (2022). *Proyecto Educativo del Programa. Licenciatura en Lenguas Extranjeras con Énfasis en Inglés y Francés*. Universidad de Antioquia.

Universidad de Antioquia. (2024). *Programa Oficial del Curso. Uso de las TIC en la enseñanza y aprendizaje de las lenguas extranjeras*. Universidad de Antioquia.

Universidad de la Amazonia. (2019a). *Micro Currículo. Pensamiento lógico matemático*. Universidad de la Amazonia.

Universidad de la Amazonia. (2019b). *Proyecto Educativo del Programa. Licenciatura en Literatura y Lengua Castellana*. Universidad de la Amazonia.

Universidad de la Amazonia. (2023). *Micro Currículo. Tecnologías de la comunicación aplicadas a la educación*. Universidad de la Amazonia.

Universidad de los Andes. (2023, septiembre 21). *Matriz DOFA: ¿Qué es y como potencia tus fortalezas?* Universidad de los Andes. <https://programas.uniandes.edu.co/blog/que-es-una-matriz-dofa-descubre-como-usar-esta-herramienta-para-potenciar-tus-fortalezas>

Universidad Nacional de Colombia. (2014a). *Proyecto Educativo del Programa. Licenciatura en Español y Filología Clásica*. Universidad Nacional de Colombia.

Universidad Nacional de Colombia. (2014b). *Proyecto Educativo del Programa. Lingüística*. Universidad Nacional de Colombia.

Universidad Nacional de Colombia. (2015). *Proyecto Educativo del Programa. Licenciatura en Filología e Idiomas: Alemán, Francés e Inglés*. Universidad Nacional de Colombia.

Vasco Uribe, C. E., Martínez Boom, A., & Vasco Montoya, E. (2008). Educación, pedagogía y didáctica: Una perspectiva epistemológica. *Filosofía de la educación, 2008*, ISBN 978-84-8164-853-9, págs. 99-128, 99-128. <https://dialnet.unirioja.es/servlet/articulo?codigo=2659697>

Yllescas, J. C. T. (2011). Puentes entre la lingüística computacional y la psicolingüística. *Revista de Lingüística y Lenguas Aplicadas*, 6, 341-352. <https://doi.org/10.4995/rlyla.2011.914>

Zambrano Briones, M. A., Hernández Díaz, A., Mendoza Bravo, K. L., Zambrano Briones, M. A., Hernández Díaz, A., & Mendoza Bravo, K. L. (2022). El aprendizaje basado en proyectos como estrategia didáctica. *Conrado*, 18(84), 172-182. [http://scielo.sld.cu/scielo.php?script=sci\\_abstract&pid=S1990-86442022000100172&lng=es&nrm=iso&tlng=es](http://scielo.sld.cu/scielo.php?script=sci_abstract&pid=S1990-86442022000100172&lng=es&nrm=iso&tlng=es)