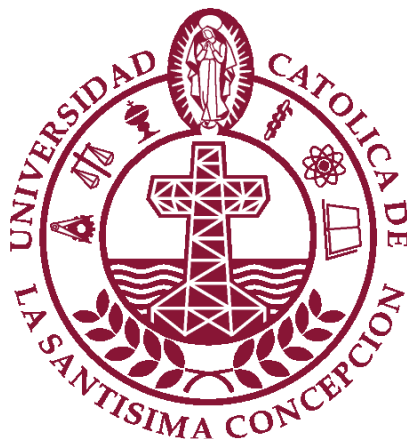


UNIVERSIDAD CATÓLICA DE LA SANTÍSIMA CONCEPCIÓN



DISPONIBILIDAD LÉXICA EN ESTUDIANTES DE TERCER Y SÉPTIMO SEMESTRE DE PEDAGOGÍA EN INGLÉS, EN LA UNIVERSIDAD CATÓLICA DEL NORTE.

ALVARO LÓPEZ BUSTAMANTE

Tesis presentada a la Facultad de Educación de la Universidad Católica de la Santísima Concepción para optar al grado académico de Magíster en Lingüística Aplicada.

DIRECTOR DE TESIS: Dr. Milko Cepeda Guerra

COMISIÓN INFORMANTE:

Concepción, Abril 2017

Dedicado a mi madre.

AGRADECIMIENTOS

A mi familia y a la hermandad.

A Omraam Mikhäel Aïvanhov.

A.M.D.G.

Tabla de Contenidos

| | |
|---|-----------|
| Agradecimientos..... | 3 |
| Tabla de contenidos..... | 4 |
| Lista de tablas..... | 7 |
| Lista de gráficos..... | 8 |
| Resumen..... | 9 |
| Abstract..... | 10 |
| Capítulo 1 – Planteamiento del problema y objetivos..... | 11 |
| 1. Planteamiento del problema..... | 12 |
| 1.1 Justificación..... | 12 |
| 1.2 Pregunta de investigación..... | 15 |
| 1.3 Objetivos..... | 15 |
| 1.3.1 Objetivo General..... | 15 |
| 1.3.2 Objetivos específicos..... | 16 |
| 1.4 Variables..... | 16 |
| 1.4.1 Variables independientes..... | 16 |
| 1.4.2 Variables dependientes..... | 16 |
| 1.5 Hipótesis..... | 17 |
| Capítulo 2 – Marco Teórico..... | 18 |
| 2.1 Relación histórica del concepto de disponibilidad léxica..... | 19 |
| 2.1.1 Antecedentes..... | 19 |
| 2.1.2. Centros de interés y disponibilidad léxica. | 21 |
| 2.2. Definiciones. | 26 |
| 2.2.1 Palabra, vocablo y unidad léxica..... | 26 |

| | |
|--|-----------|
| 2.2.2 Tipos de léxico. | 27 |
| 2.2.3 Léxico disponible y competencia léxica. | 34 |
| 2.2.4. El lexicón mental..... | 38 |
| 2.2.5. Memoria semántica..... | 43 |
| 2.3. Aplicaciones de los estudios de disponibilidad léxica. | 52 |
| 2.3.1. Sociolingüística..... | 52 |
| 2.3.2. Psicolingüística..... | 54 |
| 2.3.3. Dialectología..... | 56 |
| 2.3.4. Etnolingüística | 57 |
| 2.3.5. Didáctica de lenguas..... | 58 |
| Capítulo 3 – Marco Metodológico | 60 |
| 3. Metodología..... | 61 |
| 3.1 Muestra..... | 61 |
| 3.2 Instrumento y aplicación..... | 62 |
| 3.3 Variables..... | 63 |
| 3.3.1 Variable independiente..... | 63 |
| 3.3.2 Variables dependientes..... | 63 |
| 3.4 Determinación de centros de interés. | 64 |
| 3.5 Pilotaje de la prueba. | 66 |
| 3.6 Criterios para la tabulación de datos. | 67 |
| 3.7 Procedimientos de análisis de los datos..... | 68 |
| Capítulo 4 – Resultados | 72 |
| 4. Resultados..... | 73 |
| 4.1 Resultados Generales..... | 73 |

| | |
|---|------------|
| 4.2 Resultados por Centro de Interés: Índice de disp. léxica y frecuencias..... | 82 |
| 4.2.1 Centro de Interés: Gramática. | 82 |
| 4.2.1.1 Resultados Generales..... | 82 |
| 4.2.1.2 Número de respuestas por participante..... | 87 |
| 4.2.1.3 Índice de Disponibilidad Léxica Individual..... | 88 |
| 4.2.1.3 Clusters..... | 89 |
| 4.2.2 Centro de Interés: Morfología. | 90 |
| 4.2.2.1 Resultados Generales..... | 90 |
| 4.2.2.2 Número de respuestas por participante..... | 94 |
| 4.2.2.3 Índice de Disponibilidad Léxica Individual..... | 96 |
| 4.2.2.3 Clusters..... | 97 |
| 4.2.3 Centro de Interés: Fonología y fonética. | 98 |
| 4.2.3.1 Resultados Generales..... | 98 |
| 4.2.3.2 Número de respuestas por participante..... | 103 |
| 4.2.3.3 Índice de Disponibilidad Léxica Individual..... | 104 |
| 4.2.3.3 Clusters..... | 105 |
| 4.2.4 Centro de Interés: Discurso. | 106 |
| 4.2.4.1 Resultados Generales..... | 106 |
| 4.2.4.2 Número de respuestas por participante..... | 111 |
| 4.2.4.3 Índice de Disponibilidad Léxica Individual..... | 112 |
| 4.2.4.3 Clusters..... | 113 |
| Capítulo 5 – Discusión de resultados | 114 |
| 5.1 Análisis y discusión de resultados..... | 115 |
| Capítulo 6 – Conclusiones | 122 |

| | |
|--------------------------------------|------------|
| 6.1 Conclusiones..... | 123 |
| 6.2 Limitaciones..... | 127 |
| 6.3 Proyecciones..... | 128 |
| 7. Bibliografía..... | 131 |
| Informe de Profesores Revisores..... | 151 |
| 8. Anexos..... | 153 |
| Anexo 8.1..... | 154 |
| Anexo 8.2..... | 155 |

Lista de Tablas

| | |
|--|----|
| Tabla 4.1 Resultados Generales Tercer Semestre | 73 |
| Tabla 4.2 Resultados Generales Séptimo Semestre..... | 73 |
| Tabla 4.3 Resultados Comparativos Tercer y Séptimo Semestre..... | 74 |
| Tabla 4.4 Centros de Interés en Orden Ascendente..... | 76 |
| Tabla 4.5 Promedio de Palabras por Grupo y Centro de Interés..... | 76 |
| Tabla 4.6 Promedio de Vocablos por Grupo y Centro de Interés..... | 78 |
| Tabla 4.7 Índices de Riqueza Léxica..... | 79 |
| Tabla 4.8 Índices de Cohesión..... | 80 |
| Tabla 4.9 Índice de Disponibilidad Léxica - Gramática..... | 82 |
| Tabla 4.10 Índice de Disponibilidad Léxica – Gramática – En Español..... | 86 |
| Tabla 4.11 Número de Respuestas por participante..... | 87 |
| Tabla 4.12 Índice de Disponibilidad Léxica Individual – Gramática..... | 88 |
| Tabla 4.13 Clusters - Gramática..... | 89 |

| | |
|--|-----|
| Tabla 4.14 Índice de Disponibilidad Léxica - Morfología..... | 90 |
| Tabla 4.15 Índice de Disp.Léxica – Morfología – En Español..... | 94 |
| Tabla 4.16 Número de Respuestas por participante..... | 95 |
| Tabla 4.17 Índice de Disponibilidad Léxica Individual – Morfología..... | 96 |
| Tabla 4.18 Clusters - Morfología..... | 97 |
| Tabla 4.19 Índice de Disponibilidad Léxica – Fonética y Fonología..... | 98 |
| Tabla 4.20 Índice de Disp.Léxica – Fonet. y Fonol. – En Español..... | 102 |
| Tabla 4.21 Número de Respuestas por participante..... | 103 |
| Tabla 4.22 Índice de Disponibilidad Léxica Individual – Fonét. y Fonol. | 104 |
| Tabla 4.23 Clusters – Fonética y Fonología..... | 105 |
| Tabla 4.24 Índice de Disponibilidad Léxica – Discurso..... | 106 |
| Tabla 4.25 Índice de Disp.Léxica – Discurso – En Español..... | 110 |
| Tabla 4.26 Número de Respuestas por participante..... | 111 |
| Tabla 4.27 Índice de Disponibilidad Léxica Individual – Discurso..... | 112 |
| Tabla 4.28 Clusters – Discurso. | 113 |

Lista de Gráficos

| | |
|--|----|
| Gráfico 4.1 Total de Palabras | 74 |
| Gráfico 4.2 Número de Vocablos..... | 77 |
| Gráfico 4.3 Índice de Cohesión..... | 81 |
| Gráfico 4.4 Índice de Disponibilidad Léxica Individual Promedio..... | 81 |

RESUMEN

Se investiga la disponibilidad léxica en el tecnolecto de lingüística, de estudiantes de pedagogía en inglés. Para esto, se realizó una prueba escrita a estudiantes de tercer y séptimo semestre, en la carrera de pedagogía en inglés de la Universidad Católica del Norte, midiendo el vocabulario disponible en cuatro centros de interés relativos a la lingüística. La hipótesis, es que hay una diferencia significativa en la disponibilidad léxica de ambos grupos.

Se hizo un análisis estadístico de la información a través del software DispoGen II, para establecer la disponibilidad léxica de los estudiantes en este ámbito y encontrar las conclusiones pertinentes. Se halló que la hipótesis de trabajo se cumplía, y que el grupo de séptimo semestre posee mayor cantidad de vocablos y palabras en tres de los cuatro centros de interés, mientras que el grupo del tercer semestre le supera en el centro de interés "Morfología". Ambos grupos tenían bajos índices de cohesión, y de disponibilidad léxica en general. Asimismo, tenían prácticamente el mismo índice de riqueza léxica. Esto implica que ambos grupos están en proceso de consolidación de su léxico, y es claro a la luz de los resultados, que está conviviendo un léxico en inglés, junto con otro en español. Esta investigación se puede proyectar ampliando su alcance a otros grupos mayores y menores en edad, o investigando qué efectos tiene en las competencias comunicativas, el estar situados en una comunidad de habla hispana, mientras se aprende inglés con grandes lapsos de inmersión.

Palabras clave: Disponibilidad léxica – Lingüística – Vocabulario

ABSTRACT

Specific lexical availability on linguistics is researched in English Pedagogy's students. To achieve this, it was made a written test to students of third and seventh semester, from the English Pedagogy program at Universidad Católica del Norte, measuring available vocabulary in 4 interest centers related to linguistics. The hypothesis, is that there's a significative difference between the two groups' lexical availability.

A statistical analysis of the information was made, through the special software DispoGen II, in order to determine lexical availability on the students, and to achieve relevant conclusions. It was found that the working hypothesis was achieved, and that the seventh semester group has more terms and words en three of the four interest centers, while the third semester group surpasses them in the "Morphology" interest center. Both groups had low cohesion and lexical availability indices in general. Also, they practically had the same lexical richness index. This implies that both groups are in process of lexical consolidation, and is clear that there's an English lexicon sharing space with a Spanish one. This research can be projected opening its scope to older and younger groups, or by investigating what effects has in communication abilities, to be located inside a Spanish speaking community, while English is learnt with big immersion lapses.

Keywords: Lexical availability – Linguistics – Vocabulary.

CAPÍTULO 1 – PLANTEAMIENTO DEL PROBLEMA

1. Planteamiento del problema

1.1 Justificación

El desarrollo y la adquisición del vocabulario, es un proceso complejo donde inciden factores tanto externos como internos, pues las palabras se asocian a conceptos que inciden en la comprensión del mundo. Y dado que el uso y conocimiento de las palabras están asociados a dicha comprensión, resulta muy interesante conocer a cuántas palabras y cuáles los hablantes tienen más acceso, particularmente en el caso de los futuros profesores de lenguaje, pues su propia percepción del mundo, será lo que transmitirán a generaciones de nuevos alumnos.

Así, mediante el **índice de disponibilidad léxica**, es posible determinar qué palabras puede realizar más rápidamente un grupo de participantes, cuando se refiere a un tema determinado o se está en un contexto comunicacional dado (Valencia, 1997).

Es aquí, además, cuando entra en juego nuestro lexicón mental, que es donde se encuentra toda la información que hace posible la competencia léxica, y que se encuentra muy ligado al ambiente que nos rodea. (Palapinidi, 2012).

De hecho, uno de los aspectos esenciales para la existencia de un proceso eficiente de lectura, es el tamaño del lexicón o vocabulario interno de cada persona, que debe ser extenso, de manera que permita ampliar la capacidad de comprender y producir textos (Gallosa, 2004). Profundizando en esa misma línea, Lin (2012), plantea que resulta esencial la existencia de mecanismos que nos

permitan conocer el léxico disponible requerido por cada individuo, según el contexto comunicativo en que se encuentre, pues se trata de una competencia lingüística básica, como acabamos de mencionar.

Además, un conocimiento de ese tipo, se puede adaptar para su utilización en el contexto comunicacional de enseñanza-aprendizaje de diversas disciplinas, como por ejemplo, la lingüística.

Asimismo, el vocabulario constituye un indicador de conocimientos generales, que también influyen en la eficacia del proceso de aprendizaje, pues apoya el procesamiento cognitivo, aumentando la comprensión y velocidad lectora (Bruning, Schorow, Norby & Ronning, 2005), más aún, cuando se trata de conocimientos especializados como la lingüística. Es más, el identificar más términos en un texto, puede aumentar la posibilidad de que se use el contexto de mejor forma, para deducir el significado de otros términos desconocidos en el mismo texto (Shefelbine, 1990).

En otras palabras, poseer un mayor vocabulario implica que se puede comprender con mayor facilidad los textos donde se encuentren palabras desconocidas, pues se puede utilizar el contexto para establecer los significados que se ignoran. De esta manera, si se conoce el 90% o 95% de los términos de un texto, es posible que el individuo entienda las ideas principales contenidas en este, e infiera el significado de aquellos términos que desconozca (Hirsch, 1996, citado por Eyzaguirre & Fontaïrem, 2008). Esto se corresponde con lo que indican Lockett y Shore (2003), quienes plantean que el aprender léxico nuevo, implica una mejor delimitación del espacio semántico de todas las palabras relacionadas en el sistema, cosa que incide en una mayor riqueza de lenguaje. Es así,

entonces, que mientras se profundiza el conocimiento del lenguaje, los individuos pueden ampliar también su conocimiento de otras palabras relacionadas, que incluso no se hayan entregado dentro del programa de estudios correspondiente, generando así mayor conocimiento a partir del sólo ejercicio de la lectura.

Esto podría permitir el diseño de nuevas aproximaciones a la enseñanza del campo semántico asociado a una disciplina dada, en este caso, la lingüística.

En el contexto de Chile, se posee escasa información respecto a los índices de disponibilidad léxica que tienen los estudiantes de profesiones ligadas a la lingüística o al uso del lenguaje. Existen algunas investigaciones, como las de Ferreira y Echeverría (2010), que estudiaron la disponibilidad léxica de estudiantes de Pedagogía en Inglés, para comparar con hablantes nativos del idioma. O como la de Urzúa, Chávez y Echeverría (2006), quienes compararon la disponibilidad léxica entre alumnos universitarios y docentes, en el área de matemática, aunque sólo los segundos estudiaron el tecnolecto.

Se trata, entonces, de una información relevante, pues la comprensión que tengan los estudiantes, incidirá no sólo en ellos, sino también en las personas con quienes se relacionan, y en el caso de los profesores, en sus alumnos. Por lo tanto, es un tema muy importante, pues posee una gran incidencia en nuestra sociedad.

Esta ausencia de conocimiento incide negativamente tanto en la comprensión que se tiene de las estrategias para el perfeccionamiento de los docentes en su ámbito de trabajo, como en la forma que éstos abordan su labor de enseñanza, pues se ignora sus reales capacidades en el ámbito del lenguaje.

Corresponde, entonces, situar esta problemática de una forma concreta. Nos debemos preguntar: ¿cual es la disponibilidad léxica de los estudiantes de pedagogía?, pues se trata de una cuestión sumamente importante e interesante, dado que afecta directamente el futuro de quienes serán alumnos de los actuales estudiantes.

Es posible, incluso, relacionar la disponibilidad léxica respecto a la lingüística, con su competencia en ese ámbito, pues de ser reducido el valor de este indicador en el tecnolecto de la lingüística, implicaría posiblemente una escasa comprensión de los conceptos asociados a esta área del conocimiento (Galoso, 2004). Y al haber una comprensión limitada, resultaría hipotéticamente difícil que se desarrollen competencias adecuadas en un ámbito que no se entiende a cabalidad. Aunque eso, por supuesto, es algo que eventualmente se deberá descubrir, en alguna futura investigación.

1.2 Pregunta de investigación

¿Varía la disponibilidad léxica en el tecnolecto del area lingüística, entre los estudiantes que llevan más y los que llevan menos tiempo estudiando la carrera de Pedagogía en Inglés, de la Universidad Católica del Norte?

1.3 Objetivos

1.3.1 Objetivo General

Determinar si existen diferencias entre la disponibilidad del léxico específico en el área de la lingüística, en alumnos de tercer y de séptimo semestre de pregrado.

1.3.2 Objetivos específicos

- Determinar la disponibilidad léxica del tecnolecto en alumnos de tercer semestre de pregrado.
- Determinar la disponibilidad léxica del tecnolecto en alumnos de séptimo semestre de pregrado.
- Comparar la disponibilidad léxica del tecnolecto de los alumnos de tercer semestre de pregrado, con la de los alumnos de séptimo semestre de pregrado.

1.4 Variables

1.4.1 Variable independiente

Variable: Curso del participante.

Definición operacional: Nivel de la carrera que cursan los participantes.

1.4.2 Variables dependientes

Variable: Disponibilidad léxica del participante.

Categoría: Tecnolecto del área del lenguaje.

Definición operacional: Disponibilidad léxica reflejada a través de la prueba escrita correspondiente.

Variable: Palabras realizadas por el participante.

Categoría: Tecnolecto del área del lenguaje.

Definición operacional: Se refleja a través de la prueba escrita correspondiente.

Variable: Vocablos realizados por el participante.

Categoría: Tecnolecto del área del lenguaje.

Definición operacional: Se refleja a través de la prueba escrita correspondiente.

Variable: Riqueza léxica.

Categoría: Tecnolecto del área del lenguaje.

Definición operacional: Se establece mediante el índice de riqueza léxica.

Variable: Cohesión léxica.

Categoría: Tecnolecto del área del lenguaje.

Definición operacional: Se establece mediante el índice de cohesión léxica.

1.5 Hipótesis

1.5.1 Hipótesis

Existe una diferencia significativa entre el tecnolecto de lingüística en alumnos de tercer y de séptimo semestre de pregrado.

1.5.2 Hipótesis nula

No existe una diferencia significativa entre el tecnolecto de lingüística en alumnos de tercer y de séptimo semestre de pregrado.

CAPÍTULO 2 – MARCO TEÓRICO

2. Marco teórico

2.1 Relación histórica del concepto de disponibilidad léxica

2.1.1 Antecedentes

Uno de los primeros antecedentes del estudio de la disponibilidad léxica, aparece en Alemania, cuando F. W. Käding (1898) publica su "*Häufigkeitswörterbuch der Deutschen Sprache*", que fue el primer diccionario de lengua alemana, elaborado considerando la frecuencia de aparición de vocablos, aunque el léxico fue recopilado desde fuentes escritas y relativamente formales (documentos administrativos, comerciales, históricos, prensa, etc.), por lo que no representaba necesariamente el habla cotidiana.

Un tiempo después, Edward Thorndike (1921), publica su "*Teacher's Word Book*", que es un corpus del idioma inglés, obtenido nuevamente a partir de fuentes escritas, tales como literatura infantil, libros de texto, periódicos, y correspondencia, entre otros. Tres años después, Henmon (1924), publica "*A French word book based on a count of 400,000 running words*", un corpus de la lengua francesa, elaborado con la colaboración de profesores secundarios, y construido a partir de textos provenientes de la literatura francesa de fines del siglo 19. Este corpus compartía con el texto de Thorndike, la limitación de que se trataba de unidades léxicas descontextualizadas.

Buchanan (1927) comienza a refinar la metodología, en su "*A Graded Spanish Word Book*", que estaba basado en literatura, prensa y escritos técnicos.

Buchanan introduce el concepto de rango, esto es, acompaña cada vocablo, con un número que indica en cuántos ejemplares aparece la palabra.

El siguiente paso lo da dos años más tarde, C. K. Ogden (1930). En su texto "*Basic English*", escoge las palabras más usuales y expresivas, que aparecían en las definiciones de los diccionarios. Aunque seguía sin ser cercano al lenguaje cotidiano, se convirtió, sin embargo, en un listado muy útil para difundir la lengua inglesa durante la Segunda Guerra Mundial.

En ese mismo sentido, Hugo Bakonyi (1934) hace una compilación del léxico alemán, en su "*Die gebräuchlichsten Wörter der deutschen Sprache*". A diferencia de Ogden, utilizó como fuentes doce libros de alemán para principiantes, tres listas léxicas para niños de seis años, y una construida para sordos, así como el diccionario de F. W. Käding y el texto *Measuring Scale for Ability in Spelling*, de L. P. Ayres.

El año siguiente, Michael West y James Gareth Endicott (1935) publican "*The New Method English Dictionary*", orientado a aprendientes del inglés como idioma extranjero, buscando conseguir que cualquier texto sea comprensible a partir de las palabras utilizadas en los mismos, sin tener que recurrir a diccionarios bilingües. Ese mismo año, George van der Beke (1935) reedita el *French Word Book* de Henmon, ampliando la cantidad de datos utilizada, así como los textos usados como base, y agregando además el uso del rango y especificando los casos de polisemia y de homonimia.

Poco después, Enrique Aristizábal (1938), en su texto "*Determination expérimentale du vocabulaire écrit pour servir de base à l'enseignement de l'orthographe a l'école primaire*", introduce la innovación de crear un vocabulario, a

partir de textos espontáneos no literarios, escritos por adultos y por niños, indicando además la frecuencia de aparición y el grado de dificultad ortográfica. Ya tras la Segunda Guerra Mundial, Victor García Hoz (1953), edita “Vocabulario usual, vocabulario común y vocabulario fundamental”, orientado a la enseñanza del español como lengua materna. Para establecer el corpus, utiliza el mismo procedimiento que Buchanan y van der Beke, obteniendo entonces el vocabulario frecuente, y el vocabulario común. Las fuentes léxicas corresponden a cuatro ámbitos: privado (correspondencia personal), social indiferenciado (prensa escrita), social regulado (documentos de gobierno y eclesiásticos), y cultural (textos literarios).

Esto deja, entonces, el terreno preparado para la aparición del concepto de centro de interés, que ocurriría justamente en esa década.

2.1.2. Centros de interés y disponibilidad léxica.

Aunque la idea de léxico disponible aparece a comienzos de la década de 1950 (Michéa, 1949, 1950), es en 1956 cuando esta idea toma forma concreta, con la elaboración del listado de Le Français Élémentaire (Gougenheim, Michéa, Rivenc & Sauvageot, 1956), cuyo objetivo consistía en establecer el vocabulario que se debía utilizar para incluir en los textos de enseñanza de frances, en las colonias que Francia poseía en esos momentos.

Este listado comenzó como una iniciativa de UNESCO, y se orientaba a estudiar la frecuencia de la mención de palabras, que era la tendencia en boga de ese momento. En forma consecuente, para la elaboración de este listado, se

acudió a estudios estadísticos anteriores, que se basaban en la frecuencia de aparición de vocablos. Esto, debido a que se consideraba que aquellas palabras más mencionadas, eran las más necesarias para la comunicación.

Sin embargo, pronto se detectó que había palabras muy conocidas, que tenían una frecuencia muy baja, o que no estaban en los listados de frecuencia. A causa de ello, Michéa (1953), comienza a diferenciar entre palabras gramaticales o “atemáticas” (*mots athématiques*) y palabras significativas o “temáticas” (*mots thématiques*). Las primeras son vocablos relacionados con la situación comunicativa, y son dependientes del tema abordado en la conversación. En cambio, las segundas, corresponden a palabras que cumplen funciones gramaticales, como por ejemplo artículos, conjunciones, preposiciones, etc., y que son utilizadas sin importar el tema, lo que les otorga una gran frecuencia y regularidad.

Por lo tanto, al realizarse estos primeros estudios, se encontró que había gran frecuencia de vocablos con poco contenido semántico, en gran contraste con la escasa frecuencia relativa de verbos, adjetivos y sustantivos. Esto hizo necesario revisar el criterio de selección, basada en la frecuencia neta de aparición.

Es en este punto, que surge la idea de escoger centros de interés, para estudiar la disponibilidad léxica, pues permite estudiar de manera mucho más focalizada el léxico, considerando además otro tipo de variables: geográficas, sociológicas, de género, etc.

Esto llevó a la edición del texto *L'élaboration du Français Fondamental (1er degré)* (1964), que consistió en el estudio de la disponibilidad léxica de jóvenes

franceses, con edades entre nueve y doce años. A ellos, se les aplicó una encuesta de disponibilidad léxica con 16 centros de interés.

1. Les parties du corps. (Las partes del cuerpo).
2. Les vêtements –homme et femme-. (La ropa: hombre y mujer).
3. La maison- sans meubles-. [La casa (sin muebles)].
4. Les meubles de la maison. (Los muebles de la casa).
5. Les aliments et boissons des repas. (Alimentos y bebidas de la comida).
6. Les objets placés sur la table pour les repas. (Objetos colocados en la mesa para la comida).
7. La cuisine et ses utensiles. (La cocina y sus utensilios).
8. L'école, ses meubles et son matériel scolaire. (La escuela, muebles y material escolar).
9. Le chauffage et l' éclairage. (Calefacción e iluminación).
10. La ville. (La ciudad).
11. Le village ou le bourg. (El pueblo o la ciudad).
12. Les moyens de transport. (Los medios de transporte).
13. Les travaux des champs et jardinage. (Los trabajos del campo y del jardín).
14. Les animaux. (Los animales).
15. Les jeux et distractions. (Juegos y distracciones).
16. Les différents métiers. (Los diferentes oficios).

Este trabajo, inspiró y animó a otros para que desarrollaran sus propias investigaciones, como por ejemplo el de Dimitrijévic (1969), que investigó el inglés en Escocia, y que propuso el criterio de listas abiertas con tiempo fijo, a diferencia de los franceses, que utilizaron listas cerradas con veinte palabras. En cambio, Mackey (1971), comparó el léxico del francés en Francia, con el de las zonas francófonas de Canadá. Mackey tomó el mismo rango etéreo que los franceses, agregando más centros de interés. La diferencia con el estudio realizado en Francia, estriba en que fue más allá del establecimiento de distribuciones estadísticas, pues la información recopilada se utilizó, para establecer cuáles eran las diferencias culturales entre ambas zonas geográficas.

Aunque ya en esa época, Müller (1968) había señalado que también es importante considerar el orden en que aparecen las palabras al responder una encuesta de disponibilidad léxica, pues aunque varias palabras pueden tener la misma frecuencia de aparición, eso no quiere decir que tengan el mismo índice de disponibilidad, dado que la palabra que aparece primero, está más disponible que una que aparece después.

Sin embargo, recién el año 1983, comenzaron a aparecer fórmulas matemáticas que podían asociar la disponibilidad de un vocablo, no sólo a la frecuencia neta de aparición, sino que también con el lugar de aparición. Ese año, Roberto Lorán y Humberto López (1983), establecen que es necesario introducir un coeficiente asociado al lugar de aparición, el que se debía multiplicar por la frecuencia de la palabra. Esta fórmula fue refinada por López y Strassburger (1987), pues perdía efectividad a partir del término número 23 en una lista, lo que ha sido corregido, siendo la fórmula Strassburger-López Chávez, preferible a la

existente con anterioridad, pues es útil para cualquier posición en que se encuentren los términos (López, 1991).

En Hispanoamérica, sin duda, las investigaciones realizadas en Francia, sientan las bases para los estudios de disponibilidad léxica que se han realizado en español. Aquí es innegable el gran aporte realizado tanto por Humberto López Morales, como por Max Echeverría. El primero fue un pionero en los estudios léxicos, quien ya desde 1973 realizaba análisis sociolingüísticos basados en disponibilidad léxica (López, 1973). Asimismo, es responsable del ambicioso Proyecto Panhispánico de Estudios sobre Disponibilidad Léxica, cuya finalidad es elaborar un corpus del léxico disponible de los hablantes hispanos, que aún no hayan asistido a la universidad. Como producto de ello, se busca construir un diccionario de léxico disponible tanto para España, como para Hispanoamérica.

El doctor Max Echeverría (1987), en cambio, introdujo un concepto fundamental: el índice de cohesión léxica. En 1987, Echeverría, Herrera, Moreno y Prádenas, publicaron un estudio de disponibilidad léxica, de estudiantes de educación media en Concepción, utilizando a 400 participantes, y es aquí cuando el índice de cohesión es utilizado por primera vez. En 1991, el Dr. Echeverría publica una recopilación que ofrece una mirada de conjunto respecto a la disponibilidad léxica en Chile. Asimismo, se han sumado cada vez más investigadores al estudio de la disponibilidad léxica en nuestro país. Ya ese año 1991, Vargas (1991) estudia 20 centros de interés en la ciudad de Valparaíso. En 1999, Valencia y Echeverría estudian la disponibilidad léxica de 2.052 alumnos de cuarto año de educación media, considerando elemento extralingüísticos como género, nivel sociocultural, educación, y lugar de procedencia. Otros estudios

dignos de mención, son los de Valencia (1997), que investigó la disponibilidad léxica en estudiantes de educación básica y media, los de Urzúa, Chávez y Echeverría (2006), que compara el léxico especializado en matemáticas, entre alumnos universitarios y docentes. También Ferreira y Echeverría (2010), realizaron un estudio comparativo entre la disponibilidad léxica de estudiantes de Traducción y Pedagogía en Inglés, con la de hablantes nativos del Idioma, y Ferreira, Salcedo y del Valle (2014), han vuelto a abordar la disponibilidad léxica de estudiantes de educación básica y media. Asimismo, Cepeda, Granada y Pomes (2014), también han estudiado la disponibilidad léxica de estudiantes, esta vez de primer año de educación básica. Sin embargo, no parecen haber muchos estudios sobre disponibilidad léxica en el ámbito de la lingüística, que es lo que se aborda en la presente investigación.

2.2. Definiciones.

Tras revisar la evolución de los estudios de disponibilidad léxica, resulta necesario determinar algunas definiciones.

2.2.1 Palabra, vocablo y unidad léxica

Ante todo, debemos distinguir entre los conceptos de palabra y de vocablo. Según Luna, Viguera y Báez (2005), el **vocablo** es la “unidad lexicográfica de una lengua [...] tanto *niño* como *niños* son dos formas de un mismo vocablo; al igual que *estudia*, *estudiará*, *estudiaba* son formas del vocablo *estudiar*.”. Es decir,

un vocablo es cada una de las unidades que componen el vocabulario, y de las que pueden derivar una o varias palabras. Y así, la **palabra** es cada una de las unidades elementales que integran un corpus lingüístico, sin importar si deriva o no de otra unidad.

Sin embargo, Dimitrijévic (1969) plantea que en vez de vocablo, o palabra, un concepto más adecuado es el de **unidad léxica**, que corresponde a una unidad de significado en el lexicón mental. Esta unidad léxica, dado que es semántica, puede perfectamente contar de una o más palabras, como por ejemplo “subir las escaleras”, “intestino delgado”, etc. Varios autores (Nattinger & DeCarrico, 1992; Lewis, 2000) han señalado también la necesidad de utilizar el concepto de unidad léxica para analizar la producción léxica. Gómez (2004), indica: “parece más adecuado utilizar el término ‘unidad léxica’ para el recuento de las producciones de los hablantes, dado que son unidades independientes que permiten a los usuarios de cualquier lengua organizar y procesar la información mediante la categorización en signos lingüísticos, que almacenan en su lexicón mental”.

2.2.2 Tipos de léxico.

Se pueden distinguir distintos tipos de léxico, atendiendo a su frecuencia, y a su disponibilidad.

Léxico frecuente: Es aquél cuyos vocablos tienen una mayor frecuencia de repetición (López, 1984). Siguiendo a Galisson (1979), el concepto de frecuencia se liga al de utilidad y al de disponibilidad, pues según Galisson, las palabras más

frecuentes, son aquellas que se producen más rápido. De todas maneras, resulta natural que, al tratarse de los vocablos más producidos, estos siempre aparezcan al tope de todos los diccionarios de frecuencias.

Léxico disponible: Según Michéa (1953), “una palabra disponible es una palabra que, sin ser particularmente frecuente, está sin embargo lista para ser empleada y se presenta inmediata y naturalmente al espíritu en el momento en que se la necesita”. O sea, el léxico disponible está constituido por las palabras a las que un hablante puede acceder para su producción lingüística. De acuerdo a Dubois et al. (1998), léxico disponible es un “conjunto de palabras con una frecuencia baja y poco estable, pero usuales y útiles, que están a disposición del locutor”. Según Humberto López Morales, es el “caudal léxico utilizable en una situación comunicativa dada”, esto es, “conjunto de unidades léxicas de contenido semántico concreto” (López, 1984), en otras palabras, se trata de los vocablos que llegan antes a la conciencia del hablante, debido a un estímulo temático en especial.

Léxico activo: Para López (1984), es “aquel que además de comprender, manejamos en nuestra expresión oral o escrita.”. Según García (1953), se trata de “el conjunto de palabras que un hombre emplea corrientemente en su conversación o escritura espontánea”. Por lo tanto, son las palabras utilizadas efectivamente en la producción lingüística de un hablante.

Léxico pasivo: De acuerdo a García (1953), son “las palabras que, sin ser usadas de un modo espontáneo son, sin embargo, comprendidas cuando se leen o se oyen”. López (1984) plantea de forma parecida, que se trata de “aquel que sólo se comprende, pero que nunca usamos en nuestra actuación lingüística” (1984).

Asimismo, Pastora (1990), agrega que sería un error identificar el léxico activo con el léxico frecuente, y el léxico pasivo con el vocabulario disponible, pues el vocabulario disponible es tan activo como el léxico frecuente, dado que puede ser activado en forma instantánea por el hablante.

Léxico básico: Corresponde a los vocablos utilizados cotidianamente, en las situaciones comunicativas del diario vivir. De acuerdo a López (1984), se trata de los vocablos más estables de la lengua, por lo que aparecen con mucha frecuencia, sin importar el contexto semántico. Inicialmente, para establecer la pertenencia de un vocablo al léxico básico, se utilizaba el criterio de frecuencia de aparición. Ahora, además, se pondera la frecuencia de aparición, con el índice de dispersión de cada unidad léxica.

Léxico fundamental: De acuerdo a López (1984), el léxico fundamental es la unión del léxico básico con el léxico disponible. Se trata de conceptos complementarios, pues si bien el léxico básico muestra la proporción real de uso para cada palabra, el léxico disponible nos muestra el léxico específico a que el hablante podría acceder en una situación comunicativa determinada. En otras palabras, “los léxicos básicos recogen las clases de palabras más estables en la lengua mientras que los disponibles muestran las palabras más inestables” (Gómez, 2004).

Desde el punto de vista estadístico, en la década de 1980 se consideraba que había cuatro métodos para estudiar el léxico de una comunidad lingüística dada: diccionarios normativos, frecuencia léxica, léxico básico, y disponibilidad léxica (López, 1984). Sin embargo, tras décadas de investigación, resulta necesario superar el concepto de diccionario normativo, “pues constituye un depósito de todo, o casi todo, el caudal léxico de una lengua”, para ser reemplazado por el concepto de léxico fundamental, mucho más completo y flexible (Gómez, 2004).

Y dado que nos interesa estudiar el léxico específico del área de la lingüística, es necesario también definir el concepto de **tecnolecto**.

Tcnolecto: De acuerdo a Dorta (2015), se trata del "código empleado para la comunicación por parte de miembros de un gremio determinado de la ciencia, la técnica, la política, la cultura, entre otras esferas laborales y sociales". Se trata de un registro o una variante específica de lengua, a la que tiene acceso una comunidad determinada, y que es producto de la visión profesional de un gremio determinado. Según Pérez (2008), son “variedades sectoriales, especializadas... [propias] de determinados grupos profesionales”, las que han surgido “debido a la especialización profesional de los usuarios”. Cordero (2009) distingue entre tecnolecto y lenguaje "común" o "lenguaje general". El tecnolecto es la variante especializada, utilizada por una profesión o gremio específico, en tanto el "lenguaje general" es aquel utilizado por una comunidad de habla con fines comunicativos general, sin necesidad de poseer algún grado de especialización

para comprender la variante en cuestión. De acuerdo a Cabré (2015), existe una estrecha relación entre el habla general, y el tecnolecto, debido a que “el corpus de un tecnolecto se obtiene del lenguaje común, de cuyo campo semántico no suele separarse, aunque pueda adquirir matices especiales”.

Asimismo, existen ciertos indicadores estadísticos esenciales, que permiten caracterizar de mejor manera el corpus obtenido, a través de las pruebas de disponibilidad léxica. Se trata de los siguientes indicadores:

Índice de Disponibilidad Léxica: Para establecer la disponibilidad de una unidad léxica, se utiliza el índice de disponibilidad léxica (IDL). De acuerdo a Callealta y Gallego (2016), este índice “es un valor mediante el cual es posible ordenar las palabras según la frecuencia y el orden de aparición –el grado de “inmediatez”– con que los hablantes actualizan dichas palabras.”. Inicialmente, se calculaba a partir de la frecuencia de aparición de cada vocablo (Gougenheim et al., 1956; Dimitrijévic, 1969). Sin embargo, a partir de la década de 1980, se establecieron fórmulas más refinadas, siempre expresando este índice como un valor entre 1 (disponibilidad total) y 0 (nula disponibilidad). Lorán y López (1983), introducen un coeficiente que denota la posición relativa en que aparece un vocablo determinado, el que se combina con la frecuencia neta de aparición del vocablo en cuestión. Esta fórmula fue perfeccionada por López y Strassburger (1987), corrigiendo un defecto de la fórmula propuesta por Lorán y López (1983), que distorsionaba los resultados a partir del término 23 en un listado. Para ello, consideraron los siguientes factores:

- a) Frecuencia absoluta de cada palabra en cada posición
- b) Número de participantes en la encuesta de disponibilidad léxica;
- c) Posición máxima alcanzada por cada palabra, en el centro de interés correspondiente.
- d) La posición en que fue mencionada la palabra dentro del centro de interés correspondiente, por parte de cada participante.

Esto permitió corregir las distorsiones de la fórmula original, manteniendo un valor entre 0 y 1 para expresar la disponibilidad léxica.

Índice de disponibilidad léxica individual: Si el índice de disponibilidad léxica nos indica la disponibilidad de cada vocablo para toda la muestra, en cambio, el índice de disponibilidad léxica individual (IDLI), nos indica la disponibilidad léxica de cada vocablo, para cada participante (López & Strassburger, 1987). Consiste en una adición ponderada de los IDL de cada palabra empleada por un participante, en un centro de interés dado, la que es dividida por un valor constante que permite realizar ajustes. Por lo tanto, para determinar la disponibilidad léxica de un participante en todos los centros de interés se calcula la IDLI para cada centro, y luego se suman todos los resultados. De todas maneras, esto puede llevar a algunas apreciaciones erróneas, pues mientras mayor cantidad de centros de interés, mayor será el IDLI. Así, para corregir esta distorsión, Callealta y Gallego (2016) proponen tomar la fórmula para el IDLI de Lopez y Strassburger, y ponderarla respecto a una curva de distribución estándar,

de manera que esta distorsión se corrija cuando se trata de una gran cantidad de centros de interés.

Índice de dispersión: El concepto de dispersión o rango, corresponde a la distribución de un vocablo en distintas fuentes. (Juilland & Chang-Rodríguez, 1964). Por lo tanto, el índice de dispersión, como su nombre lo dice, es un valor que indica cuánta dispersión posee un vocablo según su área temática. Este valor, corresponde estadísticamente a la desviación estándar existente entre frecuencias de un vocablo en distintas fuentes, utilizando subcorpus de tamaño parecido. Luego Carroll (1970) derivó a partir de ello, una fórmula que sirve para corpus con particiones de tamaño distinto, la que fue revisada por Gries (2008) posteriormente, y que ha sido vuelta a replantear muy recientemente (Campillos & Ueda, 2015; Callealta & Gallego, 2016).

Índice de cohesión: Este índice señala el grado de homogeneidad en el uso del léxico en los hablantes. Se encuentra dividiendo el promedio de respuestas de cada participante, entre la cantidad de palabras distintas encontradas en el centro de interés correspondiente (Echeverría, 1991). A medida que aumenten las coincidencias, el valor del índice de cohesión se acercará a 1, lo que indicará que el centro de interés es cerrado - o compacto -, en tanto si disminuyen, esta cifra se acercará a 0, y el centro de interés será abierto o difuso. Para Echeverría, mientras mayor el índice de cohesión, más estables son las relaciones entre unidades léxicas dentro del centro de interés.

2.2.3 Léxico disponible y competencia léxica.

No obstante el avance de los estudios sobre disponibilidad léxica, todavía persisten bastantes imprecisiones al momento de definir el concepto de léxico disponible. Tal como dice Hernández (2006), la producción de palabras disponibles es “una tarea cognitiva compleja que conlleva la puesta en funcionamiento de diferentes estrategias lingüístico-psicológicas”, las que se deben estudiar para establecer una comprensión cabal de qué es el léxico disponible.

Asimismo, López (1999) manifiesta lo esencial que es el aspecto psicológico para establecer la disponibilidad léxica: “Esperemos que pronto se pueda contar no sólo con una teoría psicológica de la disponibilidad léxica que conjugue factores como memoria semántica, familiaridad, etc..., sino con toda una teoría integrada de la disponibilidad”.

Hernández (2006) indica que los procesos psicológicos de cada participante, influyen grandemente en la producción lingüística. Y aunque se sigue la definición de López (1984), de que el léxico disponible, es “el caudal léxico utilizable en una situación comunicativa dada”, sin embargo, esto no nos entrega mayor información respecto a la capacidad de utilizar realmente los términos mencionados (Hernández, 2006). Porque tener la capacidad de nombrar una palabra, no implica que un hablante puede utilizarla de forma autónoma en una situación dada. En efecto, las definiciones de léxico disponible (Michea, 1953;

Galisson, 1976), en general no se refieren al habla cotidiana, sino al léxico producido en condiciones experimentales, al responder una encuesta de disponibilidad léxica, por lo que los vocabularios disponibles, aunque nos ofrecen información respecto al uso del léxico por parte de los hablantes, dicho uso no es necesariamente el mismo que un hablante realiza al estar en una situación comunicativa distinta al de contestar una encuesta de disponibilidad léxica. Por lo tanto, se trata de un reflejo necesariamente parcial de las competencias léxicas en una situación comunicativa real.

Por lo tanto, es necesario considerar las limitaciones metodológicas de las pruebas de disponibilidad léxica, las que, como indica Hernández (2006), son “una manera de evaluar las capacidades lingüísticas de un sujeto de manera aislada y experimental”.

En este ámbito, son importantes las indagaciones de Marconi (2000), respecto a la competencia léxica, quien la considera esencialmente como la capacidad de utilizar las palabras en un contexto comunicativo dado. Según este autor, la competencia léxica incluye dos aspectos:

Competencia inferencial: Consiste en la capacidad de desplazarse en la red de conexiones semánticas que surge entre una palabra y el resto de vocablos y expresiones que pertenecen a los mismos (y también otros) campos semánticos, como en la paráfrasis y la búsqueda de sinónimos.

Competencia referencial: Capacidad de expresar las palabras en el mundo concreto, real, tanto al nombrarlas, esto es, a partir de un objeto, encontrar la

palabra correspondiente que la describe, como el mecanismo inverso, que es la aplicación de la palabra, esto es, a partir de la palabra, encontrar el objeto que es descrito. Esto es, “nombrar” el objeto.

De acuerdo a Navarro y San Martín (2009), estas competencias semánticas no poseen una correlación específica, por lo que se puede ser competente en un área, y no serlo en la otra. Según Marconi (2000), existen dos tipos de léxicos: **léxico semántico**, que corresponde a la información semántica, conceptual de las palabras, y **léxico de educto**, que corresponde a la representación gráfica y fonética de cada palabra.

Ambos léxicos intervienen en las competencias inferenciales y referenciales. Así, en la competencia inferencial, se reconoce la forma de la palabra en el léxico de educto, para luego utilizar el léxico semántico e identificarla en el mundo exterior. En la competencia referencial, se utiliza el léxico semántico para identificar el vocablo en el objeto, y luego se busca la palabra en el léxico de educto, “nombrando” el objeto.

En las pruebas de disponibilidad léxica, se utilizan en primer lugar las competencias inferenciales, pues se trata de un listado de términos ligados semánticamente. Por lo tanto, cuando un hablante produce términos al completar una prueba de disponibilidad léxica, al hacer uso de las competencias inferenciales y no de las referenciales, implica que su conocimiento de los términos le permite saber que están relacionados semánticamente, pero no que

sabe como aplicar dichos términos dentro de un contexto comunicacional apropiado.

Dicho de otra manera: no por tener la capacidad de nombrar una palabra, un hablante será capaz de comprender su significado, por lo que la producción lingüística, no implica necesariamente una comprensión cabal de todo lo producido. Por lo tanto, en un caso así, el uso de los términos en un contexto comunicativo real, sería mucho más limitado que si hubiera una comprensión previa de la palabra en cuestión.

Y además, esto hace surgir la problemática semántica respecto al idiolecto. Pues incluso aunque el vocablo sea el mismo, su carga semántica puede ser distinta para cada individuo, pues resulta que en el uso de la lengua, “existen variaciones no sólo de un país a otro, de una región a otra, de un pueblo a otro, de una clase social a otra, sino también de una persona a otra” (Sánchez, 2006). Esta variación interpersonal, de acuerdo a Hernández (2006), “no se refleja en ningún estándar semántico social, y menos en los diccionarios tradicionales.”. Por lo tanto, es necesario determinar qué es exactamente entonces el corpus obtenido a partir de las pruebas de disponibilidad léxica.

Un diccionario de léxico disponible, y un diccionario de consulta común y corriente, se diferencian tanto en su organización, como en el hecho de que los términos incluidos en el diccionario tradicional, poseen una dimensión semántica plana, pues se trata de un valor semántico asignado en función de una norma socialmente aceptada. En cambio, el vocabulario disponible se obtiene a través de

un “ejercicio de la competencia léxica individual y computa, por medio de cálculos estadísticos, una generalización de dicho ejercicio lingüístico” (Hernández, 2006).

En este punto, tanto Marconi (2000), como Hernández (2006), coinciden en que la competencia léxica expresada a través de las pruebas de disponibilidad, deben ser entendidas como un proceso que abarca todos los niveles cognitivos, pues existen factores sociológicos, psicológicos, antropológicos y múltiples otras índoles, que inciden en la producción lingüística y en las asociaciones semánticas asociadas a ella. Esto nos lleva, entonces, al concepto de lexicón mental.

2.2.4. El lexicón mental

Según Aitchison (1994) y Altmann (1995), el lexicón mental es el almacén de representaciones mentales asociadas a palabras, que alberga tanto la información semántica de las mismas, como la información sintáctica y las unidades léxicas asociadas. De acuerdo a otros autores (Marslen-Wilson, 1993), consistiría en un centro que integraría todos los procesos relacionados con el lenguaje, pues contendría las unidades que relacionan la información acústico-fonética, con la interpretación semántico-sintáctica.

En los estudios en lengua inglesa (Cutler, 2006), se distingue entre *lexical access*, que se refiere a la activación de las formas de las palabras, y *meaning*, que es la activación semántica de las mismas. De esto podemos seguir, entonces, que lexicón mental (o *mental lexicon*), es el almacén mental de las formas de las palabras, mientras que la memoria semántica (*semantic memory*), es el almacén de los conceptos asociados a palabras. Así, para autores como Levelt, Roelofs y

Meyer (1999), el lexicón mental es el conjunto de unidades léxicas que un hablante puede reconocer o bien utilizar al querer establecer una comunicación.

De acuerdo al generativismo (Chomsky, 1995), el lexicón mental tiene una estructura interna que le permite integrar una amplia gama de fenómenos. Aquí, el lexicón es considerado como un conjunto de unidades léxicas que se asimilan a las unidades mínimas sintácticas. Dentro de esta línea, según Yalcin (2014), el lexicón es parte de un sistema mayor, que incluye en primer lugar al lexicón, o sea, las unidades léxicas ligadas a sus valores semánticos correspondientes, en segundo lugar, un conjunto finito de operaciones sintácticas, que permiten construir elementos mayores a partir de las unidades básicas, y en tercer lugar, un conjunto finito de reglas de composición, que permiten determinar los valores semánticos de expresiones complejas, partiendo de sus elementos básicos y de su estructura sintáctica.

Ahora bien, las combinaciones léxicas que pueden ocurrir al hacer uso del lexicón, esto es, al realizar la lengua, produce una infinidad de efectos, y nos permite realizar metafunciones léxicas, como determinar la coherencia o corrección de la producción lingüística, respecto a otras producciones, establecer los paralelos necesarios para traducir las expresiones a otras lenguas, etc. Esto implica para Baralo (2001) - reafirmando lo que indica Yalcín (2014) -, que el lexicón mental no es un mero listado de palabras, sino “que ese repertorio también estará organizado como un conjunto finito de primitivos mentales y un conjunto finito de principios combinatorios que describen el conjunto indefinido de conceptos posibles que las oraciones expresan”. Esto, desde el punto de vista

chomskiano, implicaría que ese conjunto de principios organizativos, sería un aspecto de la gramática universal aplicada a la coordinación semántica.

Otros autores, como Altmann (1999), plantean que el lexicón mental es un conjunto muy sofisticado de circuitos neurológicos. Así, un estímulo lingüístico, activa distintos puntos del circuito. Por lo tanto, el significado de una palabra, desde esta perspectiva, sólo sería la activación de circuitos neurológicos asociados a un determinado sonido, imagen o experiencia, por lo que para dos hablantes, se activarán neuronas distintas, acordes a lo que cada cual asocie neurológicamente con cada palabra. En ese sentido, los significados comunes, corresponderían a un reforzamiento de ciertas neuronas en contextos específicos, que terminan por converger en hablantes distintos (Hernández, 2006).

De todas maneras, al revisar la literatura, nos damos cuenta que el concepto del lexicón mental como una red de conexiones léxicas, parece ser la más aceptada. Siguiendo entonces a Aitchison (1994), las características más notables de esta red son:

- Como es un sistema organizado, permite rápido acceso a cada término.
- Utiliza un sistema asociativo, aunque es posible establecer distintos ámbitos de asociación: fonético, semántico, gráfico, situacional, etc.
- Este sistema es dinámico, y se encuentra en cambio constante, debido a los nuevos flujos de información que son recibidos y que modifican tanto la información como sus relaciones internas.
- Cada palabra posee un número distinto de asociaciones.
- Esto, porque aunque algunas relaciones entre palabras están dadas por ciertas convenciones y arbitrariedades (por ejemplo, aquellas que

pertenecen a distintos campos semánticos, morfológicos, temáticos, etc.), otras relaciones son netamente personales, y son dadas tanto por el contexto comunicacional en que fueron creadas, como por el aspecto situacional y vivencial del hablante.

Atienza et al. (2008), agregan las siguientes características:

- Establecen claramente que hay relaciones lingüísticas y convencionales (aspectos gramaticales por ejemplo), como subjetivas y personales (recuerdos, información sensorial, situaciones vividas, etc.)
- Que se encuentre una unidad léxica en el lexicón mental, no implica un conocimiento de todos los aspectos de dicha unidad por parte del hablante, ni tampoco que esta unidad posea el mismo valor, forma o uso tanto para el hablante como para su comunidad de habla.
- Se puede aumentar el conocimiento sobre cada unidad que es parte de la red. Incide tanto la repetición de la unidad al ser vista, oída o leída, como la naturaleza de la misma. Este grado de conocimiento, es lo que distingue al vocabulario receptivo, de aquel que produce el hablante.

De acuerdo a Gómez (2004), desde el punto de vista lingüístico, las relaciones más destacadas al interior de esta red, son:

a) Relaciones paradigmáticas:

- Relaciones asociativas según significante: De acuerdo a la misma raíz (cartón, acartonar, cartonero), o por rima (institución, organización, asociación).

- Relaciones asociativas entre significados: Analógicas (fútbol, deporte) y ontológicas (pelota, jugador, camiseta, cancha)
- Relaciones semánticas: Polisemia, sinonimia, antonimia, etc.

b) Relaciones sintagmáticas: Se producen cuando las unidades relacionadas, son parte de categorías gramaticales distintas.

Por ejemplo:

comprar/dar/recibir - regalo/obsequio/presente - cumpleaños/aniversario/bautizo

Aquí, entonces, es importante también conocer el índice de cohesión de cada centro de interés, pues de esta forma se puede saber qué relación hay entre los términos asociados mentalmente. Tanto los campos compactos como los difusos son de interés, pues si bien los primeros denotan una relación más estrecha entre unidades léxicas, los segundos nos ofrecen información sobre la relación mental entre las palabras obtenidas.

De todas formas, el concepto de lexicón mental, se encuentra estrechamente unido al de memoria semántica, afectando los avances teóricos en la segunda, en la concepción del primero, por un asunto de consistencia teórica intrínseca, pues la descripción de la memoria semántica, tiene implícita la descripción del comportamiento del lexicón mental, al momento de la producción lingüística, pues para producir una unidad léxica en una instancia como las pruebas de disponibilidad léxica, se debe recurrir precisamente a la memoria semántica.

2.2.5. Memoria semántica

De acuerdo a Hernández (2006), cuando un hablante produce una palabra, se envía una señal a un sistema de almacenamiento cognitivo, que no es totalmente lingüístico, y que define como memoria semántica. Según Urrutia (2001), la evocación de léxico disponible, es un ejercicio complejo que implica el uso de la memoria a largo plazo, lo que le distingue de otras operaciones lingüísticas que utilizan sólo la memoria de trabajo o la memoria a corto plazo. Esto, porque la memoria se codificaría de manera jerárquica, de acuerdo a su nivel de complejidad, donde los niveles más básicos sirven para codificar los más elevados. De esta manera, mientras más compleja la codificación, hay una mayor cantidad de niveles de memoria relacionados con el recuerdo, por lo que el trazo en la memoria es más duradero. Dicho de otra forma, “puede que las palabras más disponibles sean aquellas cuya unidad mnémica tenga unos trazos más profundos, conseguidos a través de experiencias más intensas o de mayor número de entradas de información.” (Hernández, 2006).

Por lo tanto, cuando se trata de palabras que tienen abundantes experiencias asociadas a ellas (como “madre”, o “escuela”, por ejemplo), habría una representación más profunda en la memoria semántica, respecto a otras con una codificación más superficial.

Respecto a ello, ya Michéa (1953), señalaba que los vocablos con elevado índice de disponibilidad, en general ofrecían mayores asociaciones verbales y afectivas, entre otras. Carcedo González (1998) complementa esto indicando que la recuperación de contenidos en la memoria semántica, “se realiza a través de

diferentes canales, no sólo del lingüístico, y se ve condicionada directamente por parámetros situacionales o contextuales”, por lo que existirían elementos extralingüísticos, vivenciales, que también incidirían en el sentido o significado que un hablante puede asociar a un vocablo dado.

Existen distintos modelos teóricos que describen la memoria semántica, los que describiremos a continuación:

Modelo de red jerárquica: Propuesto por Quillian (1968), fue el modelo que inicialmente tuvo más éxito. En él, se indica que existen dos tipos de relaciones: vínculos *isa*, esto es, relaciones jerárquicas entre conceptos, y vínculos de *propiedades*, que representan las características particular asociadas a conceptos específicos. Ambas formas de relación son direccionales, por lo que es importante cómo se relaciona un concepto con un elemento semántico. Este modelo teórico operaba con dos restricciones:

- Cada nodo tiene un único vínculo *isa* asociado.
- Cada vínculo de *propiedades*, debe estar tan elevado en la jerarquía como fuese posible.

Así, dos conceptos están relacionados de forma semántica, si están próximos en esta red. Y de esta manera, se podría medir la “proximidad semántica” entre ambos, como una distancia literalmente medible sobre un papel, pues correspondería a la longitud del camino que separa a ambos conceptos.

Este modelo funcionó para algunas pruebas experimentales, pero recibió una profunda crítica por parte de Conrad (1972), quien encontró que el tiempo de respuesta pareciera depender más de la frecuencia de uso, que de la cercanía o la posición en la jerarquía semántica. Esto se relaciona con una falla sistemática del modelo, llamada “Efecto Típico”, y que consistía en que se encontró que - por ejemplo - a pesar de estar en el misma jerarquía semántica de “aves”, es mucho más fácil producir “canario”, que “avestruz”.

Acá aparece, entonces, el concepto de *priming*. De acuerdo a Vivas (2009), *priming* es “[el] mejoramiento, respecto de una línea de base apropiada, en el desempeño en una tarea cognitiva, como producto del contexto o de una experiencia anterior.”. Por ejemplo, cuando se trata de *priming* semántico, consiste en que un término es reconocido de manera más rápida, si le precede un término relacionado semánticamente, que si le precede uno que no tiene relación. El estímulo sobre el que ocurre la respuesta, es denominado *objetivo* o *blanco*, mientras que el estímulo que le precede, es llamado *prime*.

Esto también puede ocurrir a través del denominado priming de repetición, que se manifiesta cuando se incrementa la rapidez o exactitud en las respuestas, cuando se repite un estímulo que se presentó anteriormente.

Modelos de propagación de activación: Collins y Loftus (1975), proponen modificar el modelo de Quillian, dejando atrás las jerarquías, y considerando únicamente la cercanía semántica. Este modelo se llamó Teoría Extendida de Propagación de la Activación, y proponía que la memoria semántica se iba activando a medida que los nodos se activaban uno a uno, dentro de la red

semántica. Posteriormente, Anderson (1983) propone su modelo ACT (Adaptative Character of Thought), que tiene mucho en común con el de Collins y Loftus, diferenciándose en lo siguiente:

- Aunque ambos modelos asumen que la activación demora en propagarse, el modelo ACT considera un tiempo muy corto, de menos de 50ms.
- Collins y Loftus plantean que la activación se mantiene un tiempo, mientras que ACT predice que ésta decae muy rápido, en menos de 500ms, cuando un concepto deja de ser activado.
- Collins y Loftus plantean que sólo un concepto puede ser activado a la vez, en cambio el modelo ACT plantea que el límite sólo está en la capacidad de atención de la persona.

Con el tiempo, se encontró evidencia de que el modelo ACT de Anderson, responde mejor ante las pruebas experimentales, pues la alta velocidad de decaimiento del priming en algunos casos, es consistente con ese modelo, pero no con el de Collins y Loftus (Ratcliff & McKoon, 1988).

Un perfeccionamiento a estos modelos, fue aportado justamente por Ratcliff y McKoon (1988), que proponen el **Modelo de Clave Compuesta**, planteando que existe mayor probabilidad de que haya dos palabras coexistiendo en forma relacionada dentro de la red semántica, a que haya palabras independientes. Esto explicaría la existencia del *priming* para relaciones asociativas que no son siempre semánticas. Sin embargo, este modelo no se ha aplicado mayormente a contextos comunicativos complejos, y según sus mismos autores, necesitaba considerar

también los mecanismos relacionales implícitos en la red semántica (McKoon & Ratcliff, 1995).

Modelos Conexionistas Distribuidos: Plaut (1995), junto con McRae y Boisvert (1988), plantean que la memoria se organiza en forma de red neuronal, funcionando de manera distribuida y en paralelo. Aquí, los conceptos no consisten en una estructura semántica “única”, sino que se representan como un patrón de activación de una gran cantidad de neuronas o unidades de proceso. Conceptos parecidos, se representan entonces, por patrones parecidos de activación. Cada nodo de la red se puede considerar que codifica un rasgo semántico específico, el que puede ser parte de múltiples conceptos, y estos rasgos que no necesariamente son verbalizables. Estas unidades se organizan en módulos, cuya organización varía cada vez que una unidad es activada por otras, lo que paulatinamente, a su vez, puede crear un patrón de activación. Este patrón, se determina a partir del peso relativo de cada conexión entre unidad. Los pesos relativos constituirían entonces, la memoria a largo plazo en la red semántica. De acuerdo a este modelo, el significado de una palabra se construye a través de la cooperación y competencia de las distintas unidades de proceso. Una vez que el sistema alcanza estabilidad, es posible decir que se ha activado el significado de una palabra. Si a partir de ese estado estable, se procesa una segunda palabra, se alcanzará más rápidamente un nuevo estado de estabilidad, si esta nueva palabra está relacionada semánticamente con la anterior, que si no lo está. Esto se debe a que algunas unidades que participan del significado de la segunda palabra, ya se encuentran activas con la primera.

Una ventaja de este modelo, es que permiten explicar de manera muy natural el proceso de aprendizaje, a través de un ajuste progresivo de los pesos en la red, de modo que se admitan nuevos vocablos y conceptos. Asimismo, una característica particular de este modelo, es que si se produce un daño en la red, ésta aún puede funcionar debido a su diseño (Farah & McClelland, 1991).

Modelos espaciales de alta dimensión: Se les llama así, porque modelan la memoria semántica, como un espacio vectorial de n dimensiones, donde se ubicarían los nodos de la memoria semántica. Cada dimensión, entonces, correspondería a un aspecto faceta del concepto descrito por la red n -dimensional. Los modelos más conocidos son HAL (Hiperespacio análogo al lenguaje) y LSA (Latent Semantic Analysis).

Modelo HAL (Hiperespacio análogo al lenguaje). Este modelo, planteado por Burgess y Lund (2000), y considera cada concepto como una serie de puntos distribuidos en un espacio con n dimensiones. Así, la semejanza semántica, es representada por la distancia entre cada punto. Por lo tanto, cada concepto se refleja como un vector consistente en la suma de la posición de cada punto. Para construir este modelo, se utilizaron 70 mil símbolos tomados de la red Usenet, elegidos entre más de 300 millones de palabras, por lo que se utiliza una matriz de co-ocurrencia de 70.000×70.000 , para establecer la base del espacio vectorial. Naturalmente, las definiciones se han realizado con la ayuda de computadoras. Esto es una extensión del modelo de Clave Compuesta de Ratcliff y McKoon, al basarse también en la co-ocurrencia de términos. Cada lado de la matriz, indica el

peso de la frecuencia de co-ocurrencia del término específico. De esta manera, cada palabra tiene un valor específico de co-ocurrencia, señalado como valores vectoriales, y la semejanza entre dos términos, está definida como la distancia entre los puntos que les representan dentro de la matriz.

Algo interesante en este modelo, es que dos términos pueden tener significados similares al ocurrir en contextos parecidos, por lo que en forma consecuente, tendrán vectores semánticos parecidos, y sin embargo, esto no implica necesariamente que se produzcan simultáneamente con frecuencia en las mismas oraciones.

Modelo LSA (Latent Semantic Analysis). Este modelo, propuesto por Landauer y Dumais (1997), busca solucionar el llamado “Problema de Platón”, que consiste en que a pesar de que somos expuestos a una cantidad limitada de estímulos y de información, los seres humanos parecemos conocer mucho más que el input lingüístico o comunicacional que hemos recibido. La idea básica tras este modelo, es que si poseemos toda la información, respecto a cada contexto en que una palabra aparece o no, esto permitiría generar un sistema de restricciones mutuas, que podría establecer de forma efectiva las semejanzas que existen entre los significados de cada palabra y oración de una producción lingüística dada.

En la práctica, el LSA funciona construyendo una matriz sobre un corpus de textos dado, donde las filas corresponden los acontecimientos o tipos de palabras, y las columnas los contextos o muestras de texto donde estas palabras aparecen. Cada casilla de la matriz se llena con la cantidad de veces que un tipo de palabra aparece en un contexto específico. A través de un análisis matemático

(descomposición en valores singulares, que se utiliza en el álgebra lineal), se consigue representar cada tipo de evento y contexto como puntos en un espacio vectorial de dimensiones múltiples.

La relación entre términos, se determina calculando el coseno del ángulo entre los vectores que definen a cada término, de modo que un valor 1 corresponde a un sinónimo exacto, y el 0 a términos no relacionados. De esta manera, términos aparentemente no relacionados, como “gato” y “perro”, pueden servir incluso como indicios de activación de la otra, aunque no exista una concurrencia de ambas en el mismo contexto. En suma, dos palabras estarán relacionadas semánticamente, si aparecen en contextos semejantes pero no necesariamente idénticos.

Este modelo ha sido utilizado por Jackson y Graesser (2006), para crear un sistema informático inteligente, que apoya a los aprendices en su adquisición lingüística, a través de un diálogo en lenguaje natural. De acuerdo al análisis estadístico de los resultados, se encontró que la herramienta informática “crea y utiliza una evaluación adecuada del conocimiento de los alumnos. Todas las correlaciones siguieron las direcciones hipotetizadas, lo que muestra que los estudiantes con altos puntajes reciben una mayor proporción de retroalimentación positiva, mientras que los alumnos de bajos puntajes reciben una alta proporción de retroalimentación negativa.” Por lo tanto, concluyen Jackson y Graesser (2006), el sistema “evalúa adecuadamente y responde al conocimiento de dominio de cada estudiante.”. Asimismo, se ha cotejado la capacidad predictiva del modelo, respecto a las asociaciones de palabras por estudiantes no anglófonos (Wandmacher, 2005). Se sometió a los estudiantes a una prueba en que se les

presentaba una palabra tomada del test TOEFL, con cuatro alternativas donde debían escoger aquella con un significado más semejante al de la palabra presentada. Hubo una diferencia mínima, pues mientras los estudiantes tuvieron un 64,4% de aciertos, el modelo predijo respuestas muy semejantes, logrando un 64,5% de aciertos.

Landauer y Dumais (1997) también pusieron a prueba el modelo, comparando la adquisición léxica de escolares de séptimo grado de EE.UU. (de aproximadamente 10 palabras al día), con la que se podía simular utilizando este modelo. El resultado fue que se conseguían resultados muy similares, y esto se debía al uso de relaciones débiles entre palabras que no estaban necesariamente en el mismo campo semántico. Por lo tanto, mucho del aumento léxico simulado por el modelo LSA, radicó en términos que no estaban incluidas en el texto, lo que ayudaría a explicar, por qué las personas parecen tener más conocimiento del que “deberían tener”, si sólo consideramos el input directo al que son expuestas.

Tanto LSA como HAL, están orientados a la investigación de grandes corpus léxicos, por lo que apuntan más que al uso de un individuo en particular, al uso por parte de una comunidad de habla. Por lo tanto, su uso parece ser más adecuado para hablantes adultos, sin ninguna característica especial, y con algún grado de desarrollo cultural. Si se trata de investigar en poblaciones con características más específicas, como personas con deterioro cognitivo, niños en años de desarrollo, o poblaciones marginales o sin escolarizar, parece ser mejor la aplicación de técnicas más específicas, orientadas al reconocimiento de la configuración de la memoria semántica, y su cambio, tanto hacia la ampliación como el deterioro. Así, resultan más adecuados para el análisis de situaciones

clínicas o especiales, el sistema DistSem (Vivas, 2004, 2007), y las técnicas Natural y Definition Finder (Vivas et al., 2008) o bien, la captura de redes semánticas naturales, que proponen Figueroa et al. (1982).

2.3. Aplicaciones de los estudios de disponibilidad léxica

Resulta necesario, además, mencionar las múltiples aplicaciones que tienen los estudios de disponibilidad léxica, en distintas disciplinas. Como hemos visto, se ha pasado desde sus orígenes, arraigados en la planificación de vocabulario para la enseñanza de una lengua, a investigaciones de tipo sociolingüístico y en muchas otras áreas. En efecto, existe una línea de investigación iniciada por Mackey (1971), que utiliza los corpus de disponibilidad léxica, para realizar comparaciones culturales entre distintas comunidades separadas geográficamente, los que también se han realizado en español (Lopez, 1995; Carcedo, 1999). Asimismo hay aplicaciones en el área etnolingüística, psicolingüística, en la lexicografía, dialectología, etc. (Carcedo 1998). A continuación, entonces, se expondrán los campos de aplicación más notables de los estudios de disponibilidad léxica.

2.3.1. Sociolingüística

Los estudios de disponibilidad léxica son muy útiles pues permiten establecer indicadores muy valiosos de variaciones en el ámbito sociolingüístico. De acuerdo a Gómez (2004), el análisis cuantitativo de la disponibilidad léxica en

una comunidad de habla dada, entrega “una excelente descripción de su norma léxica reflejada en la matriz social de dicha comunidad.”, lo que para López (1999), facilita realizar estudios comparativos interdialectales.

Así, la comparación de formas alternantes, permite cotejar la disponibilidad de vocablos de distinta procedencia, en una misma comunidad de habla, así como aquellas asociadas a distintos estilos y niveles socioeconómicos. (Semper, 2006). En ese sentido, los estudios realizados en Hispanoamérica, han sido pioneros en incorporar la variable socioeconómica, que ni Gougenheim et al. (1967) ni Dimitrijevic (1969) habían considerado en sus estudios. Así, se ha encontrado una fuerte correlación entre el uso de ciertos vocablos, y el nivel socioeconómico de los hablantes, y también en la cantidad de vocablos disponibles, disminuyendo en forma proporcional al menor nivel socioeconómico de los hablantes (Semper, 2006). También resulta importante considerar el vocabulario compartido y no compartido entre niveles socioeconómicos y educativos, sin olvidar naturalmente los fenómenos de bilingüismo, lo que puede entregar mayores luces sobre estos fenómenos.

Y en otro ámbito, según Drummond (2013), el deseo de identificación con un grupo social específico, incide en la elección de vocablos al momento de producir el habla, suprimiendo o privilegiando ciertas palabras. En esa misma línea, Potowski (2011), ha encontrado que existe una fuerte correlación entre el sentimiento de identidad hacia un determinado grupo social, y el vocabulario utilizado.

Por ejemplo, Alba (1999), encontró que la densidad de anglicismos en el español dominicano, se había incrementado respecto a estudios anteriores, en la

producción léxica de las clases socioeconómicas más altas, lo que permite establecer que los estudios de disponibilidad léxica permiten generar aportes e indicadores muy interesantes para establecer fenómenos de habla atingentes a la sociolingüística, pues de acuerdo a Gómez (2004), el aspecto léxico es esencial para caracterizar a las comunidades de habla que se estudien, y no tan sólo por las correlaciones entre disponibilidad léxica y nivel sociocultural que ya hemos mencionado, sino que además, es necesario “para establecer una estratificación sociolingüística de la comunidad objeto de estudio, producto de la variación y de la heterogeneidad.” (Gómez, 2004). Sin duda, el estudio de la correlación entre los aspectos sociales, extralingüísticos, y la disponibilidad léxica, parece ser una necesidad creciente no sólo de este campo, sino que de todos los estudios lingüísticos.

2.3.2. Psicolingüística

Una forma esencial en que la disponibilidad léxica se relaciona con la psicolingüística es a través de los estudios del lexicón mental. De acuerdo a Luque (2000), se utiliza el lexicón mental para establecer tipologías léxico-semánticas, a través de estudios comparativos del mismo fenómeno, en lenguas distintas. Desde esta perspectiva, el lexicón está constituido por unidades léxicas que conforman redes asociativas, “en función de las características biológicas y psicológicas del ser humano” (Gómez, 2004). Estas redes asociativas funcionan en múltiples niveles, por lo que una unidad léxica puede pertenecer a distintas redes, según la relación que tenga con otras unidades (parecido, oposición, sustitución, etc.).

Ahora bien, al estudiar la estructura que constituye al lexicón mental, también se analiza la relación entre estímulo y reacción léxica generada. Para ello, resulta muy útil tanto medir la densidad léxica, como determinar el índice de cohesión (Echeverría, 1991), que permite establecer si el centro de interés analizado es cerrado o abierto (esto es, compacto o difuso). Para la psicolingüística, resultan de interés los centros de interés compactos, pues implican una relación entre las distintas palabras y también entre su orden de aparición (Gómez, 2004), lo que permite estudiar el acceso léxico, y a su vez, es una herramienta muy útil para ayudar a analizar el *priming* semántico, esto es, la preponderancia de un sentido u otro en los vocablos utilizados (Barsalou, 1992). Esto lleva a estudios más profundos, dentro de distintos modelos existentes, como por ejemplo el de búsqueda autónoma (Forster, 1976), el de logogen (Morton, 1969), el conexionista (Seidenberg & McClelland, 1989), etc.

Incluso, la disponibilidad léxica en sí, también posee alcances en la neuropsicología, pues es un elemento de diagnóstico para lesiones cerebrales específicas (Moulin et al., 2002).

En suma, el problema del léxico, su producción e identidad, de acuerdo a Hernández (2006), implica una aproximación interdisciplinaria, “que considere el sistema lingüístico como la herramienta fundamental de la cognición humana.”. Hernández, entonces, enmarca el análisis de la disponibilidad léxica, de la siguiente forma: “el paradigma más afortunado para encuadrar el estudio del léxico disponible en su búsqueda de una teoría integradora son las Ciencias Cognitivas” (Hernández, 2006), por lo que hay un amplio campo de exploración sobre esta faceta particular de la disponibilidad léxica.

2.3.3. Dialectología

Una de las aplicaciones más tradicionales de los estudios de disponibilidad léxica, consiste en el análisis dialectológico comparativo, que deriva del trabajo clásico de Mackey (1971), que cotejó el léxico disponible en la zona francoparlante de Canadá, versus el francés hablado en Francia. En efecto, López (1993) señala que, respecto a los análisis de disponibilidad léxica, “donde estos estudios ofrecen su mejor y más valiosa contribución es en las comparaciones interdialectales”. Entonces, a través de la investigación de la disponibilidad léxica, es posible establecer la presencia (o no) de localismos, dialectalismos, etc., y analizar su vigencia, arraigo, vitalidad o dominio de los mismos, en el área geográfica que se estudia.

En ese sentido, por ejemplo, se ha podido determinar qué léxico se comparte en un área geográfica donde existe una lengua oficial, y varios dialectos, como lo realizó Carcedo (2001), en su investigación sobre el bable en la zona de Asturias, en España.

Siguiendo a Siegel (2010), podemos establecer que a través del estudio del vocabulario, es posible determinar aspectos del comportamiento de dialectos en contacto, lo cual se puede combinar con otros factores, como el nivel educativo y formalidad de la adquisición del dialecto.

Asimismo, los estudios de disponibilidad léxica pueden contribuir en la determinación de la validez de ciertas hipótesis, al combinar los corpus obtenidos,

con datos extralingüísticos, como la edad, o el género de las personas encuestadas (Otheguy & Zentella, 2012; Tagliamonte & Molfenter, 2007).

2.3.4. Etnolingüística

El estudio realizado por Mackey (1971), también fue fundamental para establecer la importancia que tienen los estudios de disponibilidad léxica, para la investigación de características culturales comparadas. En efecto, en ese mismo estudio, Mackey plantea que la disponibilidad léxica puede ser una medida de las diferencias culturales, por ejemplo López y García (1995) introducen centros de interés de acuerdo a las características de su muestra (habitantes metropolitanos de Madrid). Asimismo, Samper y Hernández (1997) han encontrado que factores culturales afectan la producción léxica en habitantes de la isla Gran Canaria, donde por ejemplo existen menos vocablos asociados al centro 'iluminación, calefacción y medios de airear un recinto', lo que se explicaría debido a las condiciones climáticas del lugar, que no requiere calefacción.

Incluso el aspecto fonológico permite averiguar detalles que facilitan el análisis contrastivo de variantes lingüísticas, los que a su vez, nos hablan de aspectos culturales en las comunidades de habla involucradas (Gehrmann, 2016). Asimismo, de acuerdo a Carcedo (1998), el léxico se encuentra estrechamente unido a la cultura del hablante, y permite establecer un contexto comunicativo, desde el cual el hablante mismo se comunica con los otros, y de evaluar sus actos e intenciones. Si sólo se tratara del vocabulario, nos indica Carcedo, entonces esto no sería posible de manera cabal, por lo que es necesario el conocimiento del

vocabulario, para tener una idea integral de la cultura donde nos encontramos. En otras palabras: “la lengua es el vehículo de expresión de una comunidad concreta, con unas formas de vivir determinadas, en un entorno específico, y, como tal, amoldará una parte importante de aquélla a sus peculiares costumbres y a sus propias necesidades” (Carcedo, 1998).

2.3.5. Didáctica de lenguas

Otra aplicación, es aquella que originalmente motivó la aparición de los estudios de disponibilidad léxica, y es la enseñanza de una lengua. Recordemos que para la elaboración de vocabularios para la enseñanza de una lengua, se elaboraron inicialmente listas de las palabras más utilizadas, en las que posteriormente se pondría énfasis al momento de enseñar una lengua. Y, sin duda, eso no ha variado, pues según Carcedo (1998), existe una estrecha relación entre los estudios de disponibilidad léxica y la enseñanza de un idioma, debido a la utilidad que tienen estos estudios para planificar y enseñar los vocabularios, así como para la elaboración de manuales y materiales pertinentes. Para ello, el corpus producto de los estudios de disponibilidad léxica es esencial, pues de acuerdo a Bartol (2010), es esencial para la enseñanza de una lengua, considerar lo siguiente: El vocabulario en sí, su complejidad, la enseñanza explícita del léxico, y la selección de dicho léxico. En particular la selección reviste un carácter muy especial, pues conjuga no sólo elementos netamente lingüísticos, sino culturales y sociales, dado que se debe considerar el contexto comunicativo del aprendiz, y tener en cuenta el avance progresivo en el conocimiento de la lengua, para lo cual,

se le debe entregar en forma armónica, un léxico adecuado para la etapa de aprendizaje en la que se encuentra.

Respecto específicamente a la enseñanza del español como lengua extranjera, resulta claro que la información obtenida de los estudios de disponibilidad léxica, permiten actualizar el léxico de los manuales de español, acercándolos al uso cotidiano de la lengua, permitiendo que los aprendices tengan una comunicación más efectiva en el uso diario de la lengua.

Asimismo, los estudios de disponibilidad léxica, son una herramienta muy útil para determinar las carencias y debilidades en el diseño del currículo, de acuerdo a Carcedo (2000), y además, deben servir para establecer los aspectos culturales que se mostrarán al aprendiz. Y además, a medida que los hallazgos en psicolingüística respecto al lexicón mental, naturalmente sería posible realizar aplicaciones de ello, en la enseñanza de segundas lenguas. Como indican Carrera de la Red y Bradley (2004): “Nos interesa ver los procesos de adquisición de un lengua; sería oportuno saber qué recursos hay que emplear para que un alumno incorpore una palabra a su lexicón mental”.

A modo de cierre de este marco teórico, no queda más que concordar con Carcedo (1999), respecto a “la enorme utilidad que pueden reportar las pruebas de disponibilidad léxica.”. Sin duda, permiten arrojar luces sobre múltiples áreas del conocimiento, con grandes aplicaciones prácticas, como la enseñanza de primera y segundas lenguas, la sociolingüística, la psicolingüística, la dialectología, etc. Y es indudable que su valor irá en aumento, a medida que se extiendan estos estudios en ámbitos distintos, que permitirán incrementar aún más su valor.

CAPÍTULO 3 – MARCO METODOLÓGICO

3. Metodología

En la presente investigación, se han utilizado en el aspecto metodológico los criterios aportados tanto por García (2013), y Trujillano (2015), como por Echeverría, Herrera y Segure (1980) en su manual TEVI-R. Dado que el interés principal de la investigación, es averiguar la evolución de la disponibilidad léxica respecto al habla específica relacionada con lingüística en estudiantes de pregrado, se consideraron factores afines al tipo de investigación que se deseaba realizar.

Se trata de un estudio con enfoque investigativo, y cuyo diseño es longitudinal y descriptivo-comparativo. La metodología a utilizar, para el análisis y tabulación de los datos, será cuantitativa, dada la naturaleza de la información.

3.1 Muestra

De una población total de 273 estudiantes de la carrera de pedagogía en inglés de la Universidad Católica del Norte en Antofagasta, se utilizó una muestra de 20 participantes de tercer semestre, y 20 participantes, del séptimo semestre.

Este número se consiguió, yendo a dos clases de cada grupo, y aplicando el instrumento de medición, por lo que no se trató de un número preestablecido, sino de la muestra a la que se pudo acceder en el momento.

3.2 Instrumento y aplicación

Se utilizó una encuesta, estructurada en torno a centros de interés. Se entregó a cada participante dos hojas: una para verificar que perteneciera al rango de exposición al lenguaje especializado (3 semestres y 7 semestres, respectivamente), y otra, con centros de interés relacionados con la lingüística, organizados en columnas.

Se siguieron las recomendaciones de López (1999), en cuanto a la aplicación de la encuesta:

Se consideró un tiempo de dos minutos para cada centro de interés. Se les comunicó a los participantes, que la encuesta no tenía asociada ningún tipo de calificación. Si hay copia entre compañeros, las encuestas se consideran inválidas. Asimismo, al administrar la encuesta, no se le entregó a los participantes aclaración alguna respecto a los centros de interés, para no contaminar el resultado de la prueba. La encuesta, además, fue administrada personalmente por el investigador.

Para caracterizar de mejor manera a los participantes, además, se incorporó, junto a las instrucciones para responder la encuesta de disponibilidad léxica, una breve encuesta previa, donde se preguntaba a los participantes si habían estudiado antes la carrera, en qué curso iban, y cuánto tiempo llevaban estudiando la carrera. De esta forma, se tuvo una herramienta para discriminar aquellos participantes que habían estado más expuestos al léxico especializado, de aquellos que lo habían estado menos, dentro de un mismo grupo de participantes.

3.3 Variables

3.3.1 Variable independiente

Curso del participante: definido operacionalmente como el nivel de la carrera que cursan los participantes.

3.3.2 Variables dependientes

Disponibilidad léxica de los participantes: Categorizada dentro del léxico específico del área de la lingüística. Definida operacionalmente como la disponibilidad léxica reflejada a través de la prueba escrita correspondiente.

Palabras realizadas por el participante: Categorizada dentro del léxico específico del área de la lingüística. Definida operacionalmente como la cantidad de palabras obtenidas a través de la prueba escrita correspondiente.

Vocablos realizados por el participante: Categorizado dentro del léxico específico del área de la lingüística. Definido operacionalmente como la cantidad de vocablos obtenidas a través de la prueba escrita correspondiente.

Riqueza léxica: Categorizada dentro del léxico específico del área de la lingüística. Definido operacionalmente como la cantidad de vocablos versus la cantidad de palabras obtenidas a través de la prueba escrita correspondiente.

Cohesión léxica: Categorizada dentro del léxico específico del área de la lingüística. Definido operacionalmente como el índice de cohesión determinado a partir de los términos obtenidos mediante la prueba escrita correspondiente.

3.4 Determinación de centros de interés.

Inicialmente, al comienzo de las investigaciones sobre disponibilidad léxica, los centros de interés fueron creados para distinguir distintos campos semánticos dentro del habla común de una comunidad determinada. Sin embargo, estos campos semánticos no son los mismos que abarca el habla especializada en un área dada del conocimiento.

Es aquí, cuando resulta necesario crear centros de interés específicos para el área en cuestión. Para ello, se utilizaron los siguientes criterios, siguiendo la categorización esbozada por Paredes (2014), y que sistematizamos en forma más concreta aquí:

Amplitud de los centros de interés: Es importante que los centros de interés cubran la mayor amplitud posible del campo semántico total, del área del conocimiento que se busca estudiar.

Producción de habla: La cantidad de vocablos expresados por centro de interés, también constituye un aspecto relevante, el que puede deberse a numerosas causas, no sólo al desconocimiento de un término en particular.

Homogeneidad en los centros de interés: Es importante que cada centro de interés abarque un campo semántico de la misma amplitud, y que idealmente no haya solapamiento entre ellos.

Por lo tanto, se buscó conjugar este criterio al momento de definir los centros de interés en la presente investigación. Así, en el caso del habla específica respecto al área de la lingüística, el procedimiento de elección de centros de interés fue de la siguiente forma:

Se utilizó la ayuda de expertos, consultando a los siete docentes que ejercían asignaturas relacionadas con la lingüística, en la carrera de Pedagogía en Inglés, de la Universidad Católica del Norte, respecto a los campos de la lingüística que cubrían en sus asignaturas (Para un listado completo de docentes y asignaturas, consultar **Anexo 1** de esta investigación). Paralelamente, se consultó la malla curricular de la carrera (ver **Anexo 2**), para conocer los campos de la lingüística que eran transmitidos a los estudiantes.

Tras la consulta, se determinó inicialmente que los centros de interés serían: Gramática, Sintaxis, Morfología, Fonética/Fonología y Discurso, lo que fue validado por los expertos. Sin embargo, durante el proceso de pilotaje de la prueba, esto cambió, como veremos. Asimismo, se permitió el uso de términos en español y en inglés.

3.5 Pilotaje de la prueba.

El pilotaje se realizó con 29 estudiantes de cuarto semestre de la carrera. Se transcribieron las respuestas de las hojas de cuestionario a una planilla Excel, la que fue traspasada al programa DispoGen II para la obtención de resultados estadísticos.

Se utilizó el indicador Alfa de Cronbach para la validación estadística de pilotaje, y los resultados fueron los siguientes:

Alfa de Cronbach: 0,4900663

Halves: 0,61715382

OddEven: 0,5129578

Para que este indicador señale una confiabilidad aceptable, debe poseer un índice de 0,7 o más. Se observó, además, que el 50% de los participantes dejó en blanco el centro de interés "Sintaxis", y el resto, contestó escasamente. Esto implica un nivel altísimo de omisión.

Al ser frecuentemente omitido, el centro de interés "Sintaxis", afecta la integridad estadística del instrumento, por lo que se evaluó la confiabilidad del instrumento, omitiendo este centro de interés. Estos fueron los resultados:

Alfa de Cronbach: 0,4871794

Halves: 0,45219491

OddEven: 0,73235318

Se observa que al utilizar el criterio OddEven en el Alfa de Cronbach, se obtiene un valor superior a 0,7 en la confiabilidad del instrumento. Por lo tanto, el instrumento es confiable, **siempre y cuando se omita el centro de interés “Sintaxis”**, y esto es lo que se hizo.

3.6 Criterios para la tabulación de datos.

Para la tabulación, se ha respetado el orden de las respuestas, así como las palabras en inglés y en español. Sin embargo, siguiendo a Echeverría (1980), se han descartado los términos no relacionados con el centro de interés en cuestión.

Además, siguiendo a López (1999), la ortografía no se consideró relevante para el estudio.

Siguiendo el criterio de García (2013) y Samper (2003), no se ha considerado el segundo de los términos repetidos por un participante, en un mismo centro de interés, se han reducido los términos a su forma singular, y se han descartado los artículos. Se ha estandarizado la ortografía y se han adaptado los términos reemplazando todos los espacios por guiones (“-”), de modo que sean legibles para DispoGen II.

El procedimiento mismo de tabulación, consistió en transcribir cada término de cada encuesta, y agregarlo a un archivo de texto plano, indicando el participante y el centro de interés mediante un código numérico.

La transcripción, implicó realizar las siguientes decisiones sobre el texto transcrito, para un mejor procesamiento por parte del software:

- Se digitó todo en minúscula
- Se anotó en singular tanto sustantivos como adjetivos.
- Se remitieron los diminutivos a su forma primitiva.
- Se remitieron todas las formas verbales al infinitivo, con la excepción del gerundio y el participio.
- Las palabras compuestas, se tipearon unidas por medio de guiones.

En total, se transcribieron 1031 términos. Este archivo de texto, quedó entonces listo para ser procesado utilizando el programa DispoGen II.

3.7 Procedimientos de análisis de los datos

Para el análisis de datos, se recurrió al programa DispoGen II (Echeverría et al., 2005). Este software consiste en una aplicación elaborada en la plataforma MatLab, que se especializa en cálculo matricial y en análisis estadísticos con multivariantes, como ocurre en el análisis de la disponibilidad léxica. Este programa, utiliza la fórmula de Strassburguer y López Chávez de 1987, para calcular los índices de disponibilidad léxica. Dicha fórmula posee la ventaja de utilizar un exponente cuya función es asíntota a cero, por lo tanto, no hay resultados iguales a 0, y de esta forma, los valores obtenidos conservan su capacidad discriminante.

Como ya se mencionó, fue utilizado el indicador Alfa de Cronbach para la validación estadística de pilotaje. En cambio, para el procesamiento de los datos en la encuesta propiamente tal, se utilizó la información ya transcrita a formato electrónico, para realizar un análisis a través del software DispoGen II, gracias al que se pudo determinar los siguientes indicadores: tanto los **índices de cohesión** para cada centro de interés, como el **índice de disponibilidad léxica individual**, por centro de interés y por años, y también la frecuencia y número de vocablos, junto con los totales y promedios de palabras, distribuido por años. Esto se ha podido expresar a través de porcentajes y de números absolutos, y es posible además observar las frecuencias relativas y absolutas de los distintos aspectos asociados a estas variables. Todos estos indicadores son útiles de la siguiente forma:

Índice de cohesión: Indica el grado de coincidencia en las respuestas dadas por cada participante, lo que da la posibilidad de interpretar de mejor forma el índice de disponibilidad léxica. Esto se debe, a que la homogeneidad en el habla, se encuentra en proporción directa al grado de disponibilidad de ese centro de interés posee para el conjunto de participantes.

Índice de disponibilidad léxica individual: A partir del índice de disponibilidad léxica individual, podemos estudiar tanto el promedio del grupo, como su mediana. Desde ahí, es posible construir una imagen de la disponibilidad léxica del grupo estudiado.

Frecuencia de vocablos: Nos muestra la cantidad de veces que un vocablo se repite en cada centro de interés, lo que permite formarnos una imagen de las

características del vocabulario al que el grupo y los participantes pueden acceder. Esta frecuencia se puede expresar como frecuencias relativas y frecuencia absoluta, lo que permite caracterizar la frecuencia de los vocablos de una manera más amplia y rica. Además, estas frecuencias de aparición también es posible exponerlas como porcentajes, para así matizar de mejor forma la información encontrada.

Total de vocablos: Este indicador da cuenta de la cantidad de vocablos distintos que cada individuo y grupo conoce, tanto en el área del conocimiento, como en cada centro de interés.

Total de palabras: Nos indica la cantidad total neta de palabras utilizadas por cada grupo e individuo, y se relaciona directamente con el total de vocablos utilizados, pues sirve para describir la riqueza léxica de cada grupo e individuo.

Promedio de palabras: Nos indica la cantidad media de veces que una palabra dada fue utilizada por cada participante, dentro de un centro de interés dado.

Promedio y mediana de vocablos: El promedio nos indica la cantidad media de vocablos que poseen los sujetos, para comunicarse en un ámbito dado del conocimiento. Ahora bien, la mediana, construida sobre el frecuencia de aparición de vocablos en cada centro de interés, da cuenta de cuanto es la cantidad de palabras tiende a disponer la mayoría de los participantes, lo que la convierte en una medida correctiva del promedio, y nos permite formarnos una idea más completa respecto a la situación comunicativa real del grupo investigado.

Clusters: Los clusters nos muestran las secuencias de vocablos más comunes dentro de cada centro de interés, entregando información que ayuda a caracterizar la realización del habla para cada grupo estudiado.

La expresión de cifras de distribución se realizará por medio de porcentajes y de números absolutos, lo que permite tener una mirada equilibrada respecto a las proporciones de cada uno de los indicadores, pues los porcentajes dan una perspectiva respecto al valor real de cada cifra, en el contexto total del resto de la información.

De acuerdo a Urzúa et al. (2006), los índices más relevantes corresponderían al índice de cohesión, el total de vocablos y el promedio de respuestas. Sin embargo, los otros indicadores mencionados, también dan información relevante respecto al comportamiento de cada grupo investigado, por lo que también se utilizarán, para formar un cuadro amplio respecto a los participantes de esta investigación.

CAPÍTULO 4 – RESULTADOS

4. Resultados

4.1 Resultados Generales

Para los participantes del tercer semestre, se encontraron los siguientes resultados generales:

Tabla 4.1. Resultados generales. tercer semestre. *Fuente: Elaboración propia.*

| Centro de Interés | Gramática | Morfología | Fonética y fonología | Discurso | Total |
|-----------------------------|-----------|------------|----------------------|----------|-------|
| Total de Palabras | 192 | 97 | 161 | 133 | 583 |
| Número de Vocablos | 88 | 45 | 77 | 94 | 304 |
| Promedio de Palabras | 9,6 | 4,85 | 8,05 | 6,65 | 7,29 |
| Índice de Cohesión | 0,11 | 0,11 | 0,1 | 0,07 | 0,1 |

En cambio, para los participantes del séptimo semestre, los resultados generales fueron los siguientes:

Tabla 4.2. Resultados generales. séptimo semestre. *Fuente: Elaboración propia.*

| Centro de Interés | Gramática | Morfología | Fonética y fonología | Discurso | Total |
|-----------------------------|-----------|------------|----------------------|----------|-------|
| Total de Palabras | 121 | 115 | 152 | 60 | 448 |
| Número de Vocablos | 68 | 54 | 62 | 42 | 226 |
| Promedio de Palabras | 6,05 | 5,75 | 7,6 | 3 | 5,6 |
| Índice de Cohesión | 0,09 | 0,11 | 0,12 | 0,07 | 0,1 |

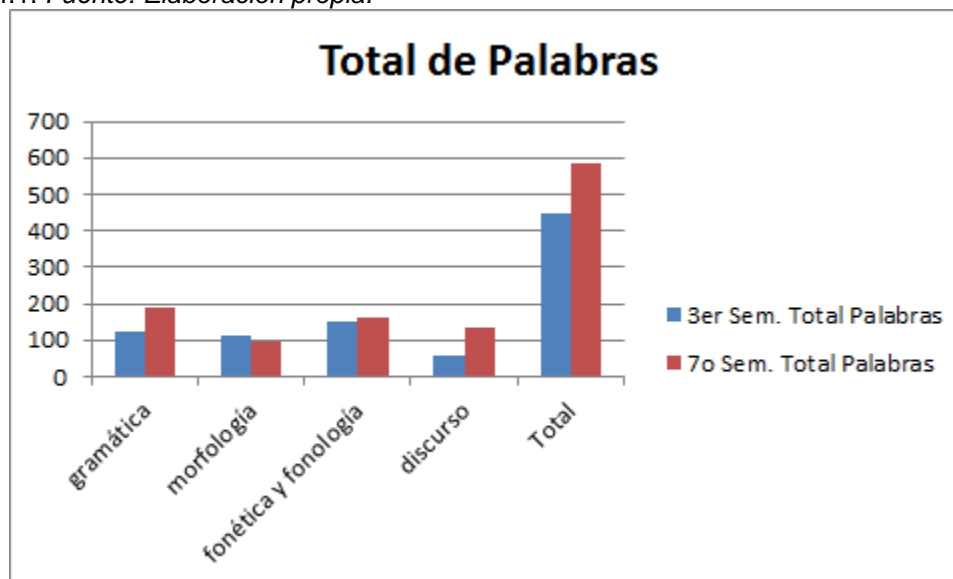
Para apreciar en forma comparativa estos resultados, se elaboró la siguiente tabla:

Tabla 4.3. Resultados comparativos, tercer y séptimo semestre. Fuente: *Elaboración propia*.

| Centro de Interés | Gramática | Morfología | Fonética y fonología | Discurso | Total |
|--|-----------|------------|----------------------|----------|-------|
| 3^{er} Sem. Total Palabras | 121 | 115 | 152 | 60 | 448 |
| 7^o Sem. Total Palabras | 192 | 97 | 161 | 133 | 583 |
| 3^{er} Sem. N° Vocablos | 68 | 54 | 62 | 42 | 226 |
| 7^o Sem. N° Vocablos | 88 | 45 | 77 | 94 | 304 |
| 3^{er} Sem. Prom. Palabras | 6,05 | 5,75 | 7,6 | 3 | 5,6 |
| 7^o Sem. Prom. Palabras | 9,6 | 4,85 | 8,05 | 6,65 | 7,29 |
| 3^{er} Sem. Índice Cohesión | 0,09 | 0,11 | 0,12 | 0,07 | 0,1 |
| 7^o Sem. Índice Cohesión | 0,11 | 0,11 | 0,1 | 0,07 | 0,1 |

Desglosando por indicador y centro de interés, y para una mayor claridad y apreciación de los resultados, se elaboraron los siguientes gráficos, que reflejan los datos de la tabla:

Gráfico 4.1. Fuente: *Elaboración propia*.



Es posible apreciar, entonces, que **el total de palabras expresadas por los participantes del tercer semestre es de 448. Esto es, 135 menos que los participantes del séptimo semestre, que le superan por un 30,1%.**

El centro de interés con menos palabras elicitadas por los participantes del tercer semestre, es el de discurso. En cambio, el centro de interés con menos palabras para los participantes del séptimo semestre, es el de morfología. Sin embargo, ese centro de interés, tiene más palabras (37, un 61,7% más) que el de discurso en los participantes del tercer semestre. Se puede observar que los centros de interés en los participantes del tercer semestre, ordenados en forma ascendente según el número de palabras contenidas, son: Discurso (60), Morfología (115), Gramática (121), Fonética y Fonología (152).

En cambio, para el séptimo semestre, son: Morfología (97), Discurso (133), Fonética y Fonología (161), y Gramática (192). Aquí es posible observar que si asignamos un número ascendente a cada centro de interés, de acuerdo al orden que le corresponde por cantidad de palabras elicitadas, vemos que desde el centro con menos palabras, hasta aquél con más palabras, los participantes del séptimo semestre tienen sistemáticamente más producción, como se observa en la tabla que sigue¹.

Tabla 4.4. Centros de interés en orden ascendente. *Fuente: Elaboración propia.*

| Centro de Interés | I | II | III | IV | Total |
|--------------------------------|-------|-------|-------|-------|-------|
| 3er Sem. Total Palabras | 60 | 115 | 121 | 152 | 448 |
| 7o Sem. Total Palabras | 97 | 133 | 161 | 192 | 583 |
| Diferencia Absoluta | 37 | 18 | 40 | 40 | 135 |
| Diferencia % | 61,7% | 15,7% | 33,1% | 26,3% | 30,1% |

Se puede observar que, aunque la diferencia absoluta de palabras entre los centros con menos producción y aquellos con más producción es casi idéntica (37

¹ La diferencia se expresa como porcentaje respecto a la cifra menor. Así, si A es 10% mayor que B, es porque constituye un 110% de B. Si es un 100% mayor, es porque es un 200% de B, etc. Los porcentajes se aproximaron en el rango de la centésima.

y 40 palabras), la diferencia porcentual varía bastante. En el primer caso, es de un 61,7%, mientras que en el segundo caso, es de un 26,3%. Sin embargo, al comparar la producción de palabras en cada centro de interés, vemos que los participantes del séptimo semestre superan a su contraparte en todos, **excepto en uno**: Morfología, donde los participantes del tercer semestre tienen 18 palabras más, esto es, un 18,6% más. Podemos apreciar otra mirada al considerar el promedio de palabras por grupo y centro de interés, lo que se expone en la siguiente tabla²:

Tabla 4.5. Promedio de palabras por grupo y centro de interés. *Fuente: Elaboración propia.*

| Centro de Interés | Gramática | Morfología | Fonética y fonología | Discurso | Total |
|--------------------------------|-----------|------------|----------------------|----------|-------|
| 3er Sem. Prom. Palabras | 6,05 | 5,75 | 7,6 | 3 | 5,6 |
| 7o Sem. Prom. Palabras | 9,6 | 4,85 | 8,05 | 6,65 | 7,29 |
| Diferencia | 3,55 | 0,9 | 0,45 | 3,65 | 1,69 |
| Diferencia % | 58,7% | 18,6% | 5,9% | 121,7% | 30,1% |

Podemos apreciar que los promedios de palabras por centro de interés, jamás subieron de 10. También que, consistentemente con los resultados ya expuestos, el grupo del tercer semestre supera al del séptimo semestre, en el centro de interés Morfología, en un 18,6%. El resultado más llamativo, es que el promedio de palabras en el centro de interés Discurso, para el séptimo semestre, más que duplica al de sus contrapartes. Además, **en el promedio de palabras totales, el grupo del séptimo semestre supera al del tercer semestre, lo que es consistente con la hipótesis inicial de esta investigación.**

² La diferencia se expresa como porcentaje respecto a la cifra menor. Así, si A es 10% mayor que B, es porque constituye un 110% de B. Si es un 100% mayor, es porque es un 200% de B, etc. Los porcentajes se aproximaron en el rango de la centésima.

Podemos ver que estas tendencias no varían, si comparamos la producción de vocablos, como es posible apreciar en el siguiente gráfico:

Gráfico 4.2. Fuente: Elaboración propia.



Según los datos de la Tabla 4.3, y guiándonos por el gráfico 4.2, podemos observar que **el total de vocablos producidos por los participantes del tercer semestre es de 226. Esto es, 78 menos que los participantes del séptimo semestre, que le superan por un 34,5%.**

Reflejando lo observado con las palabras, el centro de interés con menos vocablos producidos en el tercer semestre, es el de discurso. Para los participantes del séptimo semestre, sigue siendo el de morfología. Ambos centros de interés respectivamente tienen casi la misma cantidad de vocablos producidos, superando los participantes del séptimo semestre por un escaso margen a los del tercer semestre: 3 vocablos, un 7,1%. Los centros de interés en los participantes del tercer semestre, ordenados en forma ascendente según el número de vocablos, son: Discurso (42), Morfología (54), Gramática (62), Fonética y Fonología (68). Esto repite la misma estructura y tendencias observadas en la

producción de palabras. En cambio, para el séptimo semestre, son: Morfología (45), Fonética y Fonología (77), Gramática (88) y Discurso (94). Esto es completamente distinto de lo observado en la producción de palabras, lo que implica que el índice de riqueza léxica varía al menos para algún centro de interés.

Al igual que en el indicador anterior, si asignamos un número ascendente a cada centro de interés, según la cantidad de vocablos producidos, vemos que nuevamente los participantes del séptimo semestre tienen sistemáticamente más producción. Esto es posible verlo en la siguiente tabla³.

Tabla 4.6. . Promedio de vocablos por grupo y centro de interés *Fuente: Elaboración propia.*

| Centro de Interés | I | II | III | IV | Total |
|--------------------------------|------|-------|-------|-------|-------|
| 3er Sem. Total Vocablos | 42 | 54 | 62 | 68 | 226 |
| 7o Sem. Total Vocablos | 45 | 77 | 88 | 94 | 304 |
| Diferencia Absoluta | 3 | 23 | 26 | 26 | 78 |
| Diferencia % | 7,1% | 42,6% | 41,9% | 38,2% | 34,5% |

Aquí encontramos un comportamiento distinto al observado en la producción de palabras. Aunque el total de vocablos sigue siendo mayor en los participantes del séptimo semestre para todos los centros de interés, las diferencias se distribuyen de manera muy distinta. Así, esta es muy pequeña en el centro de interés con menos vocablos (3 vocablos, 7,1%), y luego se eleva mucho, para mantenerse muy parecida en los otros tres centros, oscilando entre 23 y 26 palabras, y alrededor del 40% de la producción léxica.

Al estudiar la producción de vocablos por centro de interés, se mantiene el mismo fenómeno observado con las palabras: Los participantes del séptimo

³ La diferencia se expresa como porcentaje respecto a la cifra menor. Así, si A es 10% mayor que B, es porque constituye un 110% de B. Si es un 100% mayor, es porque es un 200% de B, etc. Los porcentajes se aproximaron en el rango de la centésima.

semestre superan a los del tercer semestre en la producción léxica de todos los centros, excepto en Morfología. En este caso la diferencia se comporta de la siguiente forma: Les superan en 9 vocablos (versus 18 palabras), lo que constituye un 20% más, porcentaje muy parecido a la diferencia porcentual en producción de palabras.

Para tener una mirada que relacione el comportamiento tanto respecto a los vocablos, como respecto a las palabras, se elaboró una tabla⁴ con los índices de riqueza léxica (IRL).

Tabla 4.7. Índices de Riqueza Léxica. Fuente: *Elaboración propia.*

| Centro de Interés | Gramática | Morfología | Fonética y fonología | Discurso | Total |
|---------------------|-----------|------------|----------------------|----------|-------|
| 3er Sem. IRL | 0,56 | 0,47 | 0,41 | 0,7 | 0,5 |
| 7o Sem. IRL | 0,46 | 0,46 | 0,48 | 0,71 | 0,52 |
| Diferencia Absoluta | 0,1 | 0,0056 | 0,07 | 0,0068 | 0,017 |
| Diferencia % | 22,61% | 1,22% | 17,25% | 0,97% | 3,37% |

A través del índice de riqueza léxica, es posible descubrir el siguiente fenómeno: Los participantes del tercer semestre, tienen mayor riqueza léxica en gramática y en morfología, mientras que los participantes del séptimo semestre lo tienen en Discurso, y en Fonética y fonología, **así como en el total de todos los centros de interés, al sumar todas las palabras y vocablos, lo que sigue siendo consistente con la hipótesis inicial de la presente investigación.** Es de notar, sin embargo, que las diferencias porcentuales son pequeñas en tres casos: En el centro de interés Morfología, donde es de un 1,22%, en el de Discurso, con

⁴ La diferencia se expresa como porcentaje respecto a la cifra menor. Así, si A es 10% mayor que B, es porque constituye un 110% de B. Si es un 100% mayor, es porque es un 200% de B, etc. Los porcentajes esta vez se aproximaron en la milésima, debido a la presencia de diferencias muy pequeñas, que se habrían distorsionado de otra forma. Para mayor facilidad, se ha marcado en rojo, el IRL mayor entre los participantes del tercer y el séptimo semestre.

un 0,97%, y en el total acumulado, que muestra un 3,37%. En cuanto a los indicadores generales, se obtuvieron los siguientes índices de cohesión, los que se presentan a continuación, distribuidos por centro de interés y grupo de participantes⁵.

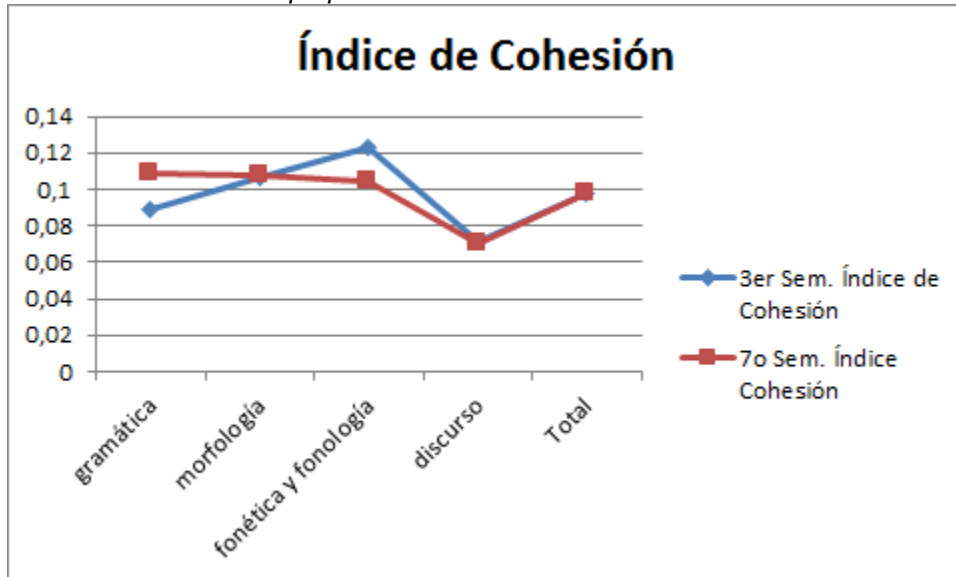
Tabla 4.8. Índices de Cohesión. *Fuente: Elaboración propia.*

| Centro de Interés | Gramática | Morfología | Fonética y fonología | Discurso | Total |
|------------------------------------|-----------|------------|----------------------|----------|--------|
| 3er Sem. Índice de Cohesión | 0,089 | 0,11 | 0,12 | 0,07 | 0,1 |
| 7o Sem. Índice de Cohesión | 0,11 | 0,11 | 0,1 | 0,07 | 0,01 |
| Diferencia | 0,02 | 0,001 | 0,02 | 0,0007 | 0,0007 |
| Diferencia % | 22,61% | 1,22% | 17,25% | 0,97% | 0,69% |

Es posible notar que, en general, los índices de cohesión son más bien bajos, no superando el valor de 0,12. Y aunque existen un par de diferencias porcentuales grandes, entre los índices para el mismo centro de interés en distintos grupos, estas diferencias (22,61% y 17,25% respectivamente), corresponden a diferencias objetivas del rango de la centésima, por lo que se trata de variaciones bastante pequeñas. Esto es coherente con la diferencia total, que es sólo 0,0007, o un 0,69%, donde el grupo de séptimo semestre supera en homogeneidad al de tercer semestre. Esto se puede apreciar más claramente en el siguiente gráfico:

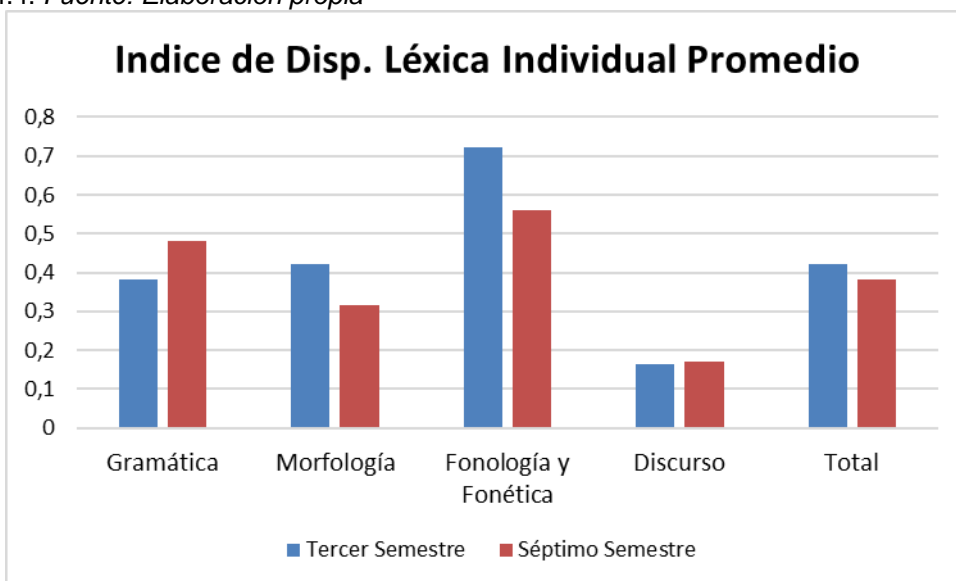
⁵ La diferencia se expresa como porcentaje respecto a la cifra menor. Así, si A es 10% mayor que B, es porque constituye un 110% de B. Si es un 100% mayor, es porque es un 200% de B, etc. Los porcentajes esta vez se aproximaron en la milésima, debido a la presencia de diferencias muy pequeñas, que se habrían distorsionado de otra forma. Para mayor facilidad, se ha marcado en rojo, **el índice de cohesión menor** entre los participantes del tercer y el séptimo semestre.

Gráfico 4.3. Fuente: *Elaboración propia*



Finalmente, podemos apreciar un resumen de los distintos promedios del Índice de Disponibilidad Léxica Individual, desglosados por centro de interés, información que será explicada detalladamente al examinar el centro de interés correspondiente.

Gráfico 4.4. Fuente: *Elaboración propia*



4.2 Resultados por Centro de Interés: Índice de disponibilidad léxica y frecuencias.

4.2.1 Centro de Interés: Gramática.

4.2.1.1 Resultados Generales.

A continuación, se presentan los resultados obtenidos para cada vocablo en la muestra total, ordenados en forma descendente, de acuerdo a su Índice de Disponibilidad Léxica, y donde también se expone la frecuencia (absoluta y acumulada) de aparición de cada vocablo, y su porcentaje de aparición.

Es menester notar, que a los participantes se les dio la libertad de responder en el idioma de su elección, y respondieron abrumadoramente en inglés. En la siguiente tabla, se muestran los resultados obtenidos.

Tabla 4.9. Índice de Disponibilidad Léxica – Gramática. Fuente: *Elaboración propia.*

| Vocablo | IDL | Frecuencia | % Aparición | Frec. Acumulada |
|----------------|-------|------------|-------------|-----------------|
| verb | 0,295 | 19 | 47,5 | 0,06 |
| noun | 0,224 | 15 | 37,5 | 0,11 |
| adjective | 0,223 | 15 | 37,5 | 0,16 |
| adverb | 0,197 | 14 | 35,0 | 0,2 |
| subject | 0,127 | 7 | 17,5 | 0,22 |
| tense | 0,125 | 8 | 20,0 | 0,25 |
| article | 0,101 | 6 | 15,0 | 0,27 |
| sentence | 0,099 | 7 | 17,5 | 0,29 |
| future | 0,099 | 8 | 20,0 | 0,32 |
| pronoun | 0,095 | 10 | 25,0 | 0,35 |
| past | 0,092 | 7 | 17,5 | 0,37 |
| past-simple | 0,09 | 4 | 10,0 | 0,38 |
| modal-verb | 0,088 | 6 | 15,0 | 0,4 |
| simple-present | 0,088 | 4 | 10,0 | 0,42 |
| conjunction | 0,088 | 5 | 12,5 | 0,43 |
| gerund | 0,078 | 5 | 12,5 | 0,45 |
| present | 0,077 | 6 | 15,0 | 0,47 |
| infinitive | 0,075 | 5 | 12,5 | 0,48 |
| irregular-verb | 0,074 | 4 | 10,0 | 0,5 |
| regla | 0,072 | 3 | 7,5 | 0,5 |
| present-simple | 0,066 | 3 | 7,5 | 0,51 |
| punctuation | 0,065 | 4 | 10,0 | 0,53 |
| regular-verb | 0,064 | 4 | 10,0 | 0,54 |

| | | | | |
|----------------------|-------|---|------|------|
| word | 0,058 | 4 | 10,0 | 0,55 |
| structure | 0,056 | 3 | 7,5 | 0,56 |
| past-tense | 0,052 | 4 | 10,0 | 0,58 |
| rule | 0,051 | 3 | 7,5 | 0,58 |
| irregular | 0,049 | 4 | 10,0 | 0,6 |
| conditional | 0,046 | 5 | 12,5 | 0,61 |
| regular | 0,046 | 3 | 7,5 | 0,62 |
| phrasal-verb | 0,045 | 3 | 7,5 | 0,63 |
| present-continuous | 0,044 | 3 | 7,5 | 0,64 |
| predicate | 0,044 | 3 | 7,5 | 0,65 |
| clause | 0,044 | 2 | 5,0 | 0,66 |
| preposition | 0,043 | 3 | 7,5 | 0,67 |
| verbo | 0,04 | 2 | 5,0 | 0,67 |
| simple-past | 0,038 | 2 | 5,0 | 0,68 |
| present-tense | 0,036 | 2 | 5,0 | 0,69 |
| future-tense | 0,035 | 2 | 5,0 | 0,69 |
| present-perfect | 0,031 | 3 | 7,5 | 0,7 |
| past-perfect | 0,026 | 3 | 7,5 | 0,71 |
| vocabulary | 0,025 | 1 | 2,5 | 0,72 |
| assignment | 0,025 | 1 | 2,5 | 0,72 |
| preposición | 0,025 | 1 | 2,5 | 0,72 |
| palabra | 0,025 | 1 | 2,5 | 0,73 |
| orden-de-la-oración | 0,025 | 1 | 2,5 | 0,73 |
| presente | 0,025 | 1 | 2,5 | 0,73 |
| comma | 0,025 | 1 | 2,5 | 0,73 |
| gramatical-structure | 0,025 | 1 | 2,5 | 0,74 |
| complement | 0,022 | 2 | 5,0 | 0,74 |
| book | 0,022 | 1 | 2,5 | 0,75 |
| adverbio | 0,022 | 1 | 2,5 | 0,75 |
| intersection | 0,022 | 1 | 2,5 | 0,75 |
| capital-letter | 0,022 | 1 | 2,5 | 0,76 |
| link | 0,022 | 1 | 2,5 | 0,76 |
| pasado | 0,022 | 1 | 2,5 | 0,76 |
| past-progressive | 0,022 | 1 | 2,5 | 0,77 |
| estructura | 0,022 | 1 | 2,5 | 0,77 |
| form | 0,022 | 1 | 2,5 | 0,77 |
| organization | 0,022 | 1 | 2,5 | 0,78 |
| past-participle | 0,02 | 2 | 5,0 | 0,78 |
| railway | 0,019 | 1 | 2,5 | 0,79 |
| sustantivo | 0,019 | 1 | 2,5 | 0,79 |
| formula | 0,019 | 1 | 2,5 | 0,79 |
| future-time | 0,019 | 1 | 2,5 | 0,8 |
| tiempo | 0,019 | 1 | 2,5 | 0,8 |
| meaning | 0,019 | 1 | 2,5 | 0,8 |
| countable | 0,018 | 2 | 5,0 | 0,81 |
| progressive | 0,017 | 2 | 5,0 | 0,81 |
| auxiliary | 0,017 | 2 | 5,0 | 0,82 |
| children | 0,016 | 1 | 2,5 | 0,82 |
| artículo | 0,016 | 1 | 2,5 | 0,83 |

| | | | | |
|---------------------|-------|---|-----|------|
| writing | 0,016 | 1 | 2,5 | 0,83 |
| puntuación | 0,016 | 1 | 2,5 | 0,83 |
| presente-simple | 0,016 | 1 | 2,5 | 0,84 |
| text | 0,016 | 1 | 2,5 | 0,84 |
| object | 0,016 | 1 | 2,5 | 0,84 |
| uncountable | 0,016 | 2 | 5,0 | 0,85 |
| interjection | 0,014 | 1 | 2,5 | 0,85 |
| conector | 0,014 | 1 | 2,5 | 0,86 |
| predicative | 0,014 | 1 | 2,5 | 0,86 |
| verb-phrase | 0,014 | 1 | 2,5 | 0,86 |
| presente-continuo | 0,014 | 1 | 2,5 | 0,87 |
| simple | 0,012 | 1 | 2,5 | 0,87 |
| sujeto | 0,012 | 1 | 2,5 | 0,87 |
| pasado-simple | 0,012 | 1 | 2,5 | 0,88 |
| possesive | 0,012 | 1 | 2,5 | 0,88 |
| past-continuous | 0,012 | 1 | 2,5 | 0,88 |
| connector | 0,012 | 1 | 2,5 | 0,88 |
| modals | 0,011 | 1 | 2,5 | 0,89 |
| compound | 0,011 | 1 | 2,5 | 0,89 |
| predicado | 0,011 | 1 | 2,5 | 0,89 |
| futuro | 0,011 | 1 | 2,5 | 0,9 |
| expression-phrase | 0,011 | 1 | 2,5 | 0,9 |
| model-verb | 0,011 | 1 | 2,5 | 0,9 |
| tiempo-verbal | 0,011 | 1 | 2,5 | 0,91 |
| paragraph | 0,009 | 1 | 2,5 | 0,91 |
| lenguaje | 0,009 | 1 | 2,5 | 0,91 |
| use | 0,009 | 1 | 2,5 | 0,92 |
| superlative | 0,009 | 1 | 2,5 | 0,92 |
| possesive-pronoun | 0,009 | 1 | 2,5 | 0,92 |
| phrase | 0,008 | 1 | 2,5 | 0,93 |
| descriptive | 0,008 | 1 | 2,5 | 0,93 |
| frequency-adverb | 0,008 | 1 | 2,5 | 0,93 |
| comparative | 0,008 | 1 | 2,5 | 0,94 |
| personal-pronoun | 0,008 | 1 | 2,5 | 0,94 |
| present-progressive | 0,008 | 1 | 2,5 | 0,94 |
| ending | 0,006 | 1 | 2,5 | 0,95 |
| third-person | 0,006 | 1 | 2,5 | 0,95 |
| affirmative | 0,006 | 1 | 2,5 | 0,95 |
| syntagma | 0,006 | 1 | 2,5 | 0,96 |
| collective-noun | 0,005 | 1 | 2,5 | 0,96 |
| negative | 0,005 | 1 | 2,5 | 0,96 |
| participle | 0,004 | 1 | 2,5 | 0,96 |
| abstract-noun | 0,004 | 1 | 2,5 | 0,97 |
| question | 0,004 | 1 | 2,5 | 0,97 |
| continuous | 0,004 | 1 | 2,5 | 0,97 |
| plural | 0,004 | 1 | 2,5 | 0,98 |
| inflections | 0,003 | 1 | 2,5 | 0,98 |
| correction | 0,003 | 1 | 2,5 | 0,98 |
| singular | 0,003 | 1 | 2,5 | 0,99 |

| | | | | |
|----------------------|-------|---|-----|------|
| frequency | 0,003 | 1 | 2,5 | 0,99 |
| body | 0,003 | 1 | 2,5 | 0,99 |
| compound-word | 0,003 | 1 | 2,5 | 0,99 |
| determinant | 0,003 | 1 | 2,5 | 1 |

Es posible observar que los vocablos en español son 21, acumulando una frecuencia absoluta total de 24 y un promedio de aparición del 2,7%. Los vocablos más usados son "regla", con 3 apariciones, y "verbo", con 2. En comparación, hay 104 vocablos en inglés, con una frecuencia absoluta total de 289. Por lo tanto, los vocablos en lengua inglesa constituyen un 83,2% de la producción lingüística elicitada. Podemos apreciar, también, que si consideramos los 10 primeros vocablos con mayor disponibilidad léxica, estos concentran más de un tercio de las menciones totales (109, de un total de 313), lo que nos muestra una gran concentración de términos en ese segmento, cosa que implica una gran probabilidad de aparición de clústeres en este centro de interés. Asimismo, es posible notar que los cuatro primeros vocablos tienen un IDL de 0,19 o superior, lo que les separa del resto de vocablos, que poseen un IDL de 0,12 o inferior, con 29 vocablos con un IDL inferior a la centésima (0,01).

Ahora bien, aunque el promedio de frecuencia de aparición es 2,5, **la mediana de frecuencias es 1**, lo que es coherente con el bajo índice de cohesión hallado en toda la muestra. También, esto se conjuga con el hecho de que 47 de los vocablos aparecen más de una vez, 36 más de dos veces, 25 más de tres veces, y 17 más de cuatro veces. Asimismo, es notorio el predominio de términos asociados a la gramática prescriptiva, en particular respecto a los verbos, sus tiempos y formas de conjugación, con una muy menor presencia de otros aspectos

del estudio gramatical. Si consideramos únicamente las palabras en español, obtenemos la siguiente tabla:

Tabla 4.10. Índice de Disponibilidad Léxica – Gramática, en español. Fuente: *Elaboración propia.*

| Vocablo | IDL | Frecuencia | % Aparición | Frec. Acumulada |
|---------------------|-------|------------|-------------|-----------------|
| regla | 0,072 | 3 | 7,5 | 0,5 |
| preposición | 0,025 | 1 | 2,5 | 0,72 |
| palabra | 0,025 | 1 | 2,5 | 0,73 |
| orden-de-la-oración | 0,025 | 1 | 2,5 | 0,73 |
| presente | 0,025 | 1 | 2,5 | 0,73 |
| adverbio | 0,022 | 1 | 2,5 | 0,75 |
| pasado | 0,022 | 1 | 2,5 | 0,76 |
| estructura | 0,022 | 1 | 2,5 | 0,77 |
| sustantivo | 0,019 | 1 | 2,5 | 0,79 |
| tiempo | 0,019 | 1 | 2,5 | 0,8 |
| artículo | 0,016 | 1 | 2,5 | 0,83 |
| puntuación | 0,016 | 1 | 2,5 | 0,83 |
| presente-simple | 0,016 | 1 | 2,5 | 0,84 |
| conector | 0,014 | 1 | 2,5 | 0,86 |
| presente-continuo | 0,014 | 1 | 2,5 | 0,87 |
| sujeto | 0,012 | 1 | 2,5 | 0,87 |
| pasado-simple | 0,012 | 1 | 2,5 | 0,88 |
| predicado | 0,011 | 1 | 2,5 | 0,89 |
| futuro | 0,011 | 1 | 2,5 | 0,9 |
| tiempo-verbal | 0,011 | 1 | 2,5 | 0,91 |
| lenguaje | 0,009 | 1 | 2,5 | 0,91 |

Es posible observar que la palabra con mayor IDL, aparece en un lugar de la tabla, donde la frecuencia acumulada es ya superior a 0,5, lo que nos indica que todo el conjunto se ubica dentro de la mitad menos producida dentro de este centro de interés. Asimismo, sólo esa palabra posee una frecuencia absoluta distinta a 1, lo que nos indica que la dispersión léxica es elevada. Prácticamente todos los términos se relacionan con el funcionamiento de los verbos, y además, se enmarcan dentro de la lingüística prescriptiva.

4.2.1.2 Número de respuestas por participante

Es posible indagar un poco más en esto, si consideramos el número de respuestas por participante:

Tabla 4.11. Número de respuestas para centro de interés "Gramática". Fuente: *Elaboración propia.*

| | Tercer Semestre | Séptimo Semestre |
|-----------------|-----------------|------------------|
| 1 | 8 | 11 |
| 2 | 5 | 13 |
| 3 | 7 | 9 |
| 4 | 7 | 3 |
| 5 | 2 | 16 |
| 6 | 3 | 9 |
| 7 | 8 | 15 |
| 8 | 5 | 7 |
| 9 | 0 | 0 |
| 10 | 11 | 11 |
| 11 | 3 | 17 |
| 12 | 3 | 10 |
| 13 | 4 | 15 |
| 14 | 7 | 8 |
| 15 | 4 | 7 |
| 16 | 9 | 2 |
| 17 | 11 | 8 |
| 18 | 6 | 7 |
| 19 | 16 | 10 |
| 20 | 2 | 14 |
| Total | 121 | 192 |
| Promedio | 6,05 | 9,6 |
| Mediana | 5,5 | 9,5 |

Aquí podemos observar una diferencia sustancial en el comportamiento de los participantes, con una marcada diferencia entre quienes pertenecen al grupo del tercer semestre y aquellos del séptimo semestre. Aunque ambos grupos tienen un participante que no produjo término alguno, y poseen un máximo de producción parecido (16 en el tercer semestre, versus 17 en el séptimo), todas las cifras totales muestran una gran diferencia. En producción total, el séptimo semestre aventaja en 71 términos (58,7%) al tercero. En producción promedio, hay una diferencia de 3,55 términos, mientras que la mediana acusa aún más esta diferencia, con una diferencia de 4 términos, esto es, de un 72,2%.

4.2.1.3 Índice de Disponibilidad Léxica Individual

Es posible profundizar aún más en esto, al estudiar el IDL para cada participante, separado por grupo:

Tabla 4.12. IDL individual para centro de interés "Gramática". Fuente: *Elaboración propia*.

| | Tercer Semestre | Séptimo Semestre |
|-----------------|-----------------|------------------|
| 1 | 0,32 | 0,65 |
| 2 | 0,46 | 0,59 |
| 3 | 0,13 | 0,23 |
| 4 | 0,09 | 0,27 |
| 5 | 0,11 | 0,5 |
| 6 | 0,19 | 0,3 |
| 7 | 0,41 | 0,86 |
| 8 | 0,82 | 0,35 |
| 9 | 0 | 0 |
| 10 | 0,89 | 0,63 |
| 11 | 0,3 | 0,58 |
| 12 | 0,32 | 0,89 |
| 13 | 0,24 | 0,39 |
| 14 | 0,86 | 0,37 |
| 15 | 0,15 | 0,64 |
| 16 | 0,4 | 0,15 |
| 17 | 0,55 | 0,46 |
| 18 | 0,3 | 0,37 |
| 19 | 0,99 | 0,55 |
| 20 | 0,1 | 0,86 |
| Total | 7,63 | 9,64 |
| Promedio | 0,38 | 0,48 |
| Mediana | 0,31 | 0,48 |

De acuerdo a esta información, podemos ver que el promedio de IDL para los participantes del tercer semestre, es menor que para los participantes del séptimo semestre, **lo que se ajusta a la hipótesis inicial de esta investigación.**

La diferencia es significativa, de 0,099 o un 26,06%.

Nuevamente, podemos apreciar que esta diferencia se acentúa al considerar la mediana, subiendo esta diferencia a 0,17 o un 54,49%.

Y aunque existe un IDL individual en el tercer semestre que bordea la unidad (0,99), mientras que el máximo del séptimo semestre es de 0,89, vemos que los promedios dicen otra cosa (aprox. 0,38 versus 0,48). Y la mediana nos dice que el IDL es aún menor para los participantes del tercer semestre, en tanto esto no cambia tanto para aquellos del séptimo semestre, respecto al promedio obtenido.

4.2.1.3 Clusters

Tal como el examen de frecuencia de términos parecía indicar, se encontraron varios clusters en los datos. Son los siguientes:

Tabla 4.13. Clústers en centro de interés "Gramática". Fuente: *Elaboración propia*.

| Cantidad | Cluster |
|----------|-------------------------------------|
| 3 | verb, noun, adjective |
| 3 | verb, adjective, noun |
| 3 | adjective, noun, pronoun |
| 2 | present-simple, past-simple, future |
| 2 | noun, adjective, adverb |
| 2 | article, gerund, infinitive |
| 2 | past, present, future |
| 2 | adjective, adverb, verb |

Se trata de 8 clusters, con una repetición de 3 o 2 veces, lo que nos indica que los participantes tienden a relacionar semánticamente estos términos. Podemos observar que casi todos se relacionan o incluyen a la función verbal como parte del cluster. Todos, sin excepción, se relacionan con la lingüística

prescriptiva, lo que corrobora la información dada por la distribución de frecuencias de vocablos. Esto puede dar una idea de la forma en que esta información es relacionada por los participantes, y la forma en que se aproximan al centro de interés en cuestión.

4.2.2 Centro de Interés: Morfología.

4.2.2.1 Resultados Generales

A continuación, se presentan los resultados obtenidos para cada vocablo en la muestra total, ordenados en forma descendente, de acuerdo a su Índice de Disponibilidad Léxica, y donde también se expone la frecuencia (absoluta y acumulada) de aparición de cada vocablo, y su porcentaje de aparición.

Al igual que en el centro de interés anterior, aunque a los participantes se les dio la libertad de responder en el idioma de su preferencia, respondieron abrumadoramente en inglés. En la siguiente tabla, se presentan los resultados obtenidos.

Tabla 4.14. Índice de Disponibilidad Léxica – Morfología. *Fuente: Elaboración propia.*

| Vocablo | IDL | Frecuencia | % Aparición | Frec. Acumulada |
|-------------|-------|------------|----------------|--------------------|
| adverb | 0,243 | 13 | 32,5 | 0,06 |
| verb | 0,235 | 14 | 35,0 | 0,13 |
| word | 0,206 | 11 | 27,5 | 0,18 |
| noun | 0,182 | 11 | 27,5 | 0,23 |
| suffix | 0,167 | 11 | 27,5 | 0,28 |
| adjective | 0,152 | 8 | 20,0 | 0,32 |
| root | 0,133 | 7 | 17,5 | 0,35 |
| prefix | 0,125 | 10 | 25,0 | 0,4 |
| affix | 0,114 | 8 | 20,0 | 0,44 |
| conjunction | 0,112 | 6 | 15,0 | 0,47 |
| preposition | 0,096 | 6 | 15,0 | 0,5 |
| morpheme | 0,091 | 5 | 12,5 | 0,52 |

| | | | | |
|--------------------|-------|---|------|------|
| pronoun | 0,068 | 5 | 12,5 | 0,54 |
| structure | 0,068 | 3 | 7,5 | 0,56 |
| phrasal-verb | 0,065 | 3 | 7,5 | 0,57 |
| inflection | 0,061 | 4 | 10,0 | 0,59 |
| clause | 0,047 | 4 | 10,0 | 0,61 |
| stem | 0,047 | 3 | 7,5 | 0,62 |
| sentence | 0,047 | 4 | 10,0 | 0,64 |
| subject | 0,046 | 2 | 5,0 | 0,65 |
| predicate | 0,046 | 2 | 5,0 | 0,66 |
| oración | 0,043 | 2 | 5,0 | 0,67 |
| regular | 0,04 | 2 | 5,0 | 0,68 |
| copulative | 0,035 | 2 | 5,0 | 0,69 |
| irregular | 0,033 | 2 | 5,0 | 0,7 |
| complex-sentence | 0,031 | 2 | 5,0 | 0,71 |
| palabra | 0,025 | 1 | 2,5 | 0,71 |
| sintaxis | 0,025 | 1 | 2,5 | 0,72 |
| article | 0,025 | 1 | 2,5 | 0,72 |
| dependent-clause | 0,025 | 1 | 2,5 | 0,73 |
| análisis | 0,025 | 1 | 2,5 | 0,73 |
| recursivity | 0,025 | 1 | 2,5 | 0,74 |
| bound-word | 0,025 | 1 | 2,5 | 0,74 |
| lexical | 0,025 | 1 | 2,5 | 0,75 |
| morfema | 0,025 | 1 | 2,5 | 0,75 |
| árbol | 0,025 | 1 | 2,5 | 0,75 |
| interjection | 0,023 | 2 | 5,0 | 0,76 |
| pure | 0,023 | 2 | 5,0 | 0,77 |
| ambigüedad | 0,021 | 1 | 2,5 | 0,78 |
| intersection | 0,021 | 1 | 2,5 | 0,78 |
| independent-clause | 0,021 | 1 | 2,5 | 0,79 |
| function | 0,021 | 1 | 2,5 | 0,79 |
| transitive | 0,021 | 1 | 2,5 | 0,8 |
| estructura | 0,021 | 1 | 2,5 | 0,8 |
| descomponer | 0,021 | 1 | 2,5 | 0,81 |
| orden | 0,021 | 1 | 2,5 | 0,81 |
| form | 0,021 | 1 | 2,5 | 0,82 |
| formation | 0,021 | 1 | 2,5 | 0,82 |
| frase-adverbial | 0,018 | 1 | 2,5 | 0,83 |
| conjunction | 0,018 | 1 | 2,5 | 0,83 |
| ditransitive | 0,018 | 1 | 2,5 | 0,83 |
| núcleo | 0,018 | 1 | 2,5 | 0,84 |
| compuesta | 0,018 | 1 | 2,5 | 0,84 |

| | | | | |
|----------------------|-------|---|-----|------|
| phrase | 0,015 | 1 | 2,5 | 0,85 |
| adjetivo | 0,015 | 1 | 2,5 | 0,85 |
| rule | 0,015 | 1 | 2,5 | 0,86 |
| sustantivo | 0,012 | 1 | 2,5 | 0,86 |
| derivation | 0,012 | 1 | 2,5 | 0,87 |
| simple-sentence | 0,012 | 1 | 2,5 | 0,87 |
| etymology | 0,012 | 1 | 2,5 | 0,88 |
| paragraph | 0,01 | 1 | 2,5 | 0,88 |
| adverbio | 0,01 | 1 | 2,5 | 0,89 |
| conjunctions | 0,01 | 1 | 2,5 | 0,89 |
| syllable | 0,01 | 1 | 2,5 | 0,9 |
| linking | 0,009 | 1 | 2,5 | 0,9 |
| infix | 0,009 | 1 | 2,5 | 0,91 |
| preposición | 0,009 | 1 | 2,5 | 0,91 |
| diphthong | 0,009 | 1 | 2,5 | 0,92 |
| bound | 0,007 | 1 | 2,5 | 0,92 |
| adverb-of-manner | 0,007 | 1 | 2,5 | 0,92 |
| morphology | 0,007 | 1 | 2,5 | 0,93 |
| morfosintaxis | 0,007 | 1 | 2,5 | 0,93 |
| context | 0,007 | 1 | 2,5 | 0,94 |
| compound | 0,006 | 1 | 2,5 | 0,94 |
| adverb-of-frequency | 0,006 | 1 | 2,5 | 0,95 |
| frase-verbal | 0,006 | 1 | 2,5 | 0,95 |
| adverb-of-time | 0,005 | 1 | 2,5 | 0,96 |
| verbo | 0,005 | 1 | 2,5 | 0,96 |
| nature | 0,004 | 1 | 2,5 | 0,97 |
| linguistic-economy | 0,004 | 1 | 2,5 | 0,97 |
| simple-clause | 0,004 | 1 | 2,5 | 0,98 |
| complex-clause | 0,004 | 1 | 2,5 | 0,98 |
| disability | 0,003 | 1 | 2,5 | 0,99 |
| chunk | 0,003 | 1 | 2,5 | 0,99 |
| ambiguity | 0,003 | 1 | 2,5 | 0,99 |
| rhythm | 0,003 | 1 | 2,5 | 1 |

Es posible observar que los vocablos en español son 19, acumulando una frecuencia absoluta total de 20 y un promedio de aparición del 2,63%. El vocablo más usado es "oración", con 2 apariciones. En comparación, hay 67 vocablos en

inglés, con una frecuencia absoluta total de 192. Por lo tanto, los vocablos en lengua inglesa constituyen el 77,91% de la producción lingüística elicitada.

En el caso de este centro de interés, podemos observar que los 10 primeros vocablos con mayor disponibilidad léxica, concentran casi la mitad de las menciones totales (99, de un total de 212), lo que nos muestra una gran concentración de términos en ese segmento, cosa que al igual que en el centro de interés anterior, implica una gran probabilidad de aparición de clústeres en este centro de interés. Asimismo, estos 10 primeros vocablos posee un IDL de 0,11 o superior, mientras que el resto de vocablos posee un IDL de 0,096 o inferior.

Ahora bien, aunque el promedio de frecuencia de aparición es 2,47, **la mediana de frecuencias es 1**, lo que resulta coherente con el bajo índice de cohesión hallado en toda la muestra. También, esto se conjuga con el hecho de que 28 de los vocablos aparecen más de una vez, 19 más de dos veces, 16 más de tres veces, y 13 más de cuatro veces.

Y aunque no hay un predominio de los términos asociados a las formas verbales, de todas maneras el término “verb” tiene el segundo lugar en IDL.

Si consideramos únicamente las palabras en español, obtenemos la siguiente tabla:

Tabla 4.15. Índice de Disponibilidad Léxica – Morfología, en español. *Fuente: Elaboración propia.*

| Vocablo | IDL | Frecuencia | % Aparición | Frec. Acumulada |
|---------------|-------|------------|-------------|-----------------|
| oración | 0,043 | 2 | 5,0 | 0,67 |
| palabra | 0,025 | 1 | 2,5 | 0,71 |
| sintaxis | 0,025 | 1 | 2,5 | 0,72 |
| análisis | 0,025 | 1 | 2,5 | 0,73 |
| morfema | 0,025 | 1 | 2,5 | 0,75 |
| árbol | 0,025 | 1 | 2,5 | 0,75 |
| ambigüedad | 0,021 | 1 | 2,5 | 0,78 |
| estructura | 0,021 | 1 | 2,5 | 0,8 |
| descomponer | 0,021 | 1 | 2,5 | 0,81 |
| orden | 0,021 | 1 | 2,5 | 0,81 |
| núcleo | 0,018 | 1 | 2,5 | 0,84 |
| compuesta | 0,018 | 1 | 2,5 | 0,84 |
| adjetivo | 0,015 | 1 | 2,5 | 0,85 |
| sustantivo | 0,012 | 1 | 2,5 | 0,86 |
| adverbio | 0,01 | 1 | 2,5 | 0,89 |
| preposición | 0,009 | 1 | 2,5 | 0,91 |
| morfosintaxis | 0,007 | 1 | 2,5 | 0,93 |
| frase-verbal | 0,006 | 1 | 2,5 | 0,95 |

Podemos apreciar que la palabra con mayor IDL, aparece en un lugar de la tabla, donde la frecuencia acumulada es ya superior a 0,66, lo que nos indica que todo el conjunto se ubica dentro del tercio menos producido dentro de este centro de interés. Asimismo, sólo esa palabra posee una frecuencia absoluta distinta a 1, lo que nos indica que la dispersión léxica es muy alta.

4.2.2.2 Número de respuestas por participante

Podemos indagar un poco más en esto, si consideramos el número de respuestas por participante:

Tabla 4.16. Número de respuestas por participante, para centro de interés “Morfología”.
 Fuente: *Elaboración propia.*

| | Tercer Semestre | Séptimo Semestre |
|-----------------|-----------------|------------------|
| 1 | 3 | 4 |
| 2 | 5 | 5 |
| 3 | 14 | 3 |
| 4 | 3 | 9 |
| 5 | 7 | 2 |
| 6 | 10 | 14 |
| 7 | 4 | 12 |
| 8 | 8 | 5 |
| 9 | 8 | 4 |
| 10 | 2 | 2 |
| 11 | 6 | 3 |
| 12 | 11 | 1 |
| 13 | 8 | 2 |
| 14 | 2 | 3 |
| 15 | 4 | 7 |
| 16 | 3 | 4 |
| 17 | 6 | 3 |
| 18 | 7 | 6 |
| 19 | 3 | 3 |
| 20 | 1 | 5 |
| Total | 115 | 97 |
| Promedio | 5,75 | 4,85 |
| Mediana | 5,5 | 4 |

Podemos observar que aquí el comportamiento es inverso al hallado en el centro de interés “gramática”, pues en este caso son los participantes del tercer semestre, quienes tienen una producción total mayor que las del séptimo semestre. La diferencia es de 18 términos (18,6%), mientras que la diferencia de **promedios**, es de 0,9 términos. En la mediana, esta diferencia es mucho más acusada, y aumenta a 1,5 términos, o sea a un 37,5%. Las cifras no son tan grandes como en el centro de interés anterior, pero siguen siendo bastante significativas.

4.2.2.3 Índice de Disponibilidad Léxica Individual

Es posible profundizar aún más en esto, al estudiar el IDL para cada participante, separado por grupo:

Tabla 4.17. IDL individual para centro de interés “Morfología”. Fuente: *Elaboración propia*.

| | Tercer Semestre | Séptimo Semestre |
|-----------------|-----------------|------------------|
| 1 | 0,4 | 0,36 |
| 2 | 0,47 | 0,29 |
| 3 | 0,66 | 0,07 |
| 4 | 0,05 | 0,43 |
| 5 | 0,58 | 0,09 |
| 6 | 0,52 | 0,76 |
| 7 | 0,48 | 0,68 |
| 8 | 0,69 | 0,59 |
| 9 | 0,65 | 0,48 |
| 10 | 0,32 | 0,1 |
| 11 | 0,69 | 0,21 |
| 12 | 0,1 | 0,03 |
| 13 | 0,71 | 0,1 |
| 14 | 0,22 | 0,07 |
| 15 | 0,41 | 0,41 |
| 16 | 0,48 | 0,33 |
| 17 | 0,27 | 0,08 |
| 18 | 0,58 | 0,66 |
| 19 | 0,08 | 0,28 |
| 20 | 0,06 | 0,29 |
| Total | 8,42 | 6,29 |
| Promedio | 0,42 | 0,31 |
| Mediana | 0,48 | 0,29 |

Aquí la diferencia es bastante apreciable, si se compara el promedio de IDL. Esta cifra es mayor para los participantes del tercer semestre, que para los participantes del séptimo semestre. **La diferencia es significativa, de 0,11 o un 33,87%.**

Nuevamente, podemos ver que esta diferencia se acentúa al considerar la mediana, aumentando la diferencia a 0,189 o un 65,28%, lo que indica que esta superioridad de producción es sistemática.

4.2.2.3 Clusters

Tal como el examen de frecuencia de términos parecía indicar, se encontraron varios clusters en los datos. Son los siguientes:

Tabla 4.18. Clusters para centro de interés “Morfología”. Fuente: *Elaboración propia*.

| Cantidad | Cluster |
|----------|----------------------------|
| 3 | affix, suffix, prefix |
| 2 | adverb, adjective, verb |
| 2 | phrasal-verb, adverb, noun |

En este caso, son 3 clústers, los que se repiten 3 y 2 veces, lo que nos indica que los participantes tienden a relacionar semánticamente estos términos dentro de este centro de interés.

Podemos apreciar que en 2 de los tres clústers, se menciona “adverb” y “verb” o un derivado, lo que nos puede entregar una idea de cómo esta información es relacionada por los participantes, y la forma en que estos se aproximan al centro de interés en cuestión.

4.2.3 Centro de Interés: Fonología y fonética.

4.2.3.1 Resultados Generales

En la tabla que sigue, se presentan los resultados obtenidos para cada vocablo en la muestra total, ordenados en forma descendente, de acuerdo a su Índice de Disponibilidad Léxica, y donde también se expone la frecuencia (absoluta y acumulada) de aparición de cada vocablo, y su porcentaje de aparición.

Nuevamente, aunque a los participantes se les dio la libertad de responder en el idioma de su elección, éstos respondieron abrumadoramente en inglés. A continuación, se presentan los resultados obtenidos.

Tabla 4.19. Índice de Disponibilidad Léxica – Fonética y fonología. *Fuente: Elaboración propia*

| Vocablo | IDL | Frecuencia | % Aparición | Frec. Acumulada |
|----------------|-------|------------|-------------|-----------------|
| phoneme | 0,398 | 20 | 50,0 | 0,06 |
| transcription | 0,339 | 20 | 50,0 | 0,13 |
| sound | 0,286 | 14 | 35,0 | 0,17 |
| vowel | 0,254 | 18 | 45,0 | 0,23 |
| consonant | 0,21 | 18 | 45,0 | 0,29 |
| Stress | 0,145 | 9 | 22,5 | 0,32 |
| intonation | 0,129 | 6 | 15,0 | 0,34 |
| Fonema | 0,108 | 5 | 12,5 | 0,35 |
| cluster | 0,105 | 7 | 17,5 | 0,37 |
| pronunciation | 0,097 | 6 | 15,0 | 0,39 |
| symbol | 0,091 | 6 | 15,0 | 0,41 |
| speech | 0,089 | 5 | 12,5 | 0,43 |
| suprasegmental | 0,087 | 5 | 12,5 | 0,44 |
| phonetic | 0,084 | 5 | 12,5 | 0,46 |
| phonology | 0,083 | 5 | 12,5 | 0,48 |
| allophone | 0,076 | 4 | 10,0 | 0,49 |
| phonetics | 0,068 | 3 | 7,5 | 0,5 |
| voiceless | 0,066 | 8 | 20,0 | 0,52 |
| pronunciación | 0,064 | 3 | 7,5 | 0,53 |

| | | | | |
|--------------------|-------|---|------|------|
| fricative | 0,063 | 5 | 12,5 | 0,55 |
| rhythm | 0,06 | 3 | 7,5 | 0,56 |
| nasal | 0,057 | 5 | 12,5 | 0,58 |
| schwa | 0,055 | 3 | 7,5 | 0,58 |
| diphthong | 0,054 | 3 | 7,5 | 0,59 |
| voiced | 0,053 | 6 | 15,0 | 0,61 |
| sonido | 0,05 | 2 | 5,0 | 0,62 |
| accent | 0,045 | 5 | 12,5 | 0,64 |
| listening | 0,039 | 4 | 10,0 | 0,65 |
| segmental | 0,039 | 2 | 5,0 | 0,65 |
| articulation | 0,039 | 3 | 7,5 | 0,66 |
| affricate | 0,037 | 3 | 7,5 | 0,67 |
| voice | 0,037 | 3 | 7,5 | 0,68 |
| voicethread | 0,036 | 4 | 10,0 | 0,7 |
| production | 0,034 | 3 | 7,5 | 0,71 |
| affricative | 0,032 | 2 | 5,0 | 0,71 |
| soft-palate | 0,031 | 3 | 7,5 | 0,72 |
| alveolar | 0,031 | 2 | 5,0 | 0,73 |
| glottal | 0,027 | 3 | 7,5 | 0,74 |
| glide | 0,025 | 1 | 2,5 | 0,74 |
| articulator | 0,025 | 1 | 2,5 | 0,74 |
| pitch | 0,021 | 1 | 2,5 | 0,75 |
| round | 0,021 | 1 | 2,5 | 0,75 |
| speaking | 0,021 | 1 | 2,5 | 0,75 |
| supersegmental | 0,021 | 1 | 2,5 | 0,76 |
| sibilante | 0,021 | 1 | 2,5 | 0,76 |
| linking | 0,02 | 3 | 7,5 | 0,77 |
| speak | 0,02 | 2 | 5,0 | 0,78 |
| alveolar-ridge | 0,02 | 3 | 7,5 | 0,79 |
| vocal-chord | 0,019 | 2 | 5,0 | 0,79 |
| producing | 0,018 | 1 | 2,5 | 0,8 |
| terminación | 0,018 | 1 | 2,5 | 0,8 |
| rule | 0,018 | 1 | 2,5 | 0,8 |
| vowels | 0,018 | 1 | 2,5 | 0,81 |
| unround | 0,018 | 1 | 2,5 | 0,81 |
| boca | 0,018 | 1 | 2,5 | 0,81 |
| upper-lip | 0,018 | 1 | 2,5 | 0,81 |
| abecedario | 0,018 | 1 | 2,5 | 0,82 |
| mouth | 0,018 | 1 | 2,5 | 0,82 |
| hard-palate | 0,017 | 2 | 5,0 | 0,83 |
| vocal | 0,015 | 1 | 2,5 | 0,83 |

| | | | | |
|----------------------|-------|---|-----|------|
| sílaba | 0,015 | 1 | 2,5 | 0,83 |
| dental | 0,015 | 1 | 2,5 | 0,84 |
| meaning | 0,015 | 1 | 2,5 | 0,84 |
| lower-lip | 0,015 | 1 | 2,5 | 0,84 |
| borrowed-word | 0,015 | 1 | 2,5 | 0,85 |
| articulación | 0,015 | 1 | 2,5 | 0,85 |
| pattern | 0,015 | 1 | 2,5 | 0,85 |
| prediction | 0,015 | 2 | 5,0 | 0,86 |
| transfer | 0,013 | 1 | 2,5 | 0,86 |
| africative | 0,013 | 1 | 2,5 | 0,87 |
| labio | 0,013 | 1 | 2,5 | 0,87 |
| interteeth | 0,013 | 1 | 2,5 | 0,87 |
| plosiva | 0,013 | 1 | 2,5 | 0,88 |
| tongue | 0,012 | 2 | 5,0 | 0,88 |
| strenght | 0,011 | 1 | 2,5 | 0,88 |
| speech-sound | 0,011 | 1 | 2,5 | 0,89 |
| practice | 0,011 | 1 | 2,5 | 0,89 |
| listen | 0,011 | 1 | 2,5 | 0,89 |
| velar | 0,011 | 1 | 2,5 | 0,9 |
| palatar | 0,011 | 1 | 2,5 | 0,9 |
| legible | 0,011 | 1 | 2,5 | 0,9 |
| friction | 0,011 | 1 | 2,5 | 0,91 |
| bilabial | 0,01 | 2 | 5,0 | 0,91 |
| lingua-dental | 0,009 | 1 | 2,5 | 0,92 |
| volume | 0,009 | 1 | 2,5 | 0,92 |
| semivowel | 0,009 | 1 | 2,5 | 0,92 |
| open | 0,009 | 1 | 2,5 | 0,93 |
| inteligible | 0,009 | 1 | 2,5 | 0,93 |
| soundless | 0,009 | 1 | 2,5 | 0,93 |
| mother-tongue | 0,009 | 1 | 2,5 | 0,94 |
| syllabe | 0,008 | 1 | 2,5 | 0,94 |
| alphabet | 0,008 | 1 | 2,5 | 0,94 |
| unstress | 0,008 | 1 | 2,5 | 0,95 |
| stop | 0,008 | 1 | 2,5 | 0,95 |
| labio-dental | 0,008 | 1 | 2,5 | 0,95 |
| producción | 0,008 | 1 | 2,5 | 0,96 |
| interference | 0,008 | 1 | 2,5 | 0,96 |
| jaw | 0,008 | 1 | 2,5 | 0,96 |
| tip | 0,008 | 1 | 2,5 | 0,96 |
| closed | 0,008 | 1 | 2,5 | 0,97 |
| message-unit | 0,007 | 1 | 2,5 | 0,97 |

| | | | | |
|-------------------|-------|---|-----|------|
| manner | 0,007 | 1 | 2,5 | 0,97 |
| ipa | 0,007 | 1 | 2,5 | 0,98 |
| stressed | 0,007 | 1 | 2,5 | 0,98 |
| unstressed | 0,006 | 1 | 2,5 | 0,98 |
| position | 0,006 | 1 | 2,5 | 0,99 |
| glotal | 0,005 | 1 | 2,5 | 0,99 |
| syllabbe | 0,005 | 1 | 2,5 | 0,99 |
| vibration | 0,004 | 1 | 2,5 | 0,99 |
| bridge | 0,003 | 1 | 2,5 | 1 |

Observando la tabla, podemos apreciar que si consideramos los 10 primeros vocablos con mayor disponibilidad léxica, estos suman más de un tercio de las menciones totales (123, de un total de 313), lo que nos muestra una gran concentración de términos en ese segmento, cosa que implica una gran probabilidad de aparición de clústeres en este centro de interés. Asimismo, es posible notar que los cinco primeros vocablos tienen un IDL superior a 0,2, lo que les separa del resto de vocablos, que poseen un IDL de 0,14 o inferior, con 28 vocablos con un IDL inferior a la centésima (0,01).

Ahora bien, aunque el promedio de frecuencia de aparición es 2,85, **la mediana de frecuencias es 1**, lo que es coherente con el bajo índice de cohesión hallado en toda la muestra. También, esto se conjuga con el hecho de que 47 de los vocablos aparecen más de una vez, 36 más de dos veces, 25 más de tres veces, y 17 más de cuatro veces. Y sin duda, destacan “phoneme” y “transcription” con 20 menciones, y “vowel”, y “consonant”, con 18 menciones, seguidas de “sound”, que aparece 14 veces. El resto de términos aparece 7 veces o menos.

Si consideramos únicamente las palabras en español, obtenemos la siguiente tabla:

Tabla 4.20. Índice de Disponibilidad Léxica – Fonética y fonología, en español.

Fuente: *Elaboración propia*

| Vocablo | IDL | Frecuencia | % Aparición | Frec. Acumulada |
|---------------|-------|------------|-------------|-----------------|
| fonema | 0,108 | 5 | 12,5 | 0,35 |
| pronunciación | 0,064 | 3 | 7,5 | 0,53 |
| sonido | 0,05 | 2 | 5,0 | 0,62 |
| sibilante | 0,021 | 1 | 2,5 | 0,76 |
| terminación | 0,018 | 1 | 2,5 | 0,8 |
| boca | 0,018 | 1 | 2,5 | 0,81 |
| abecedario | 0,018 | 1 | 2,5 | 0,82 |
| vocal | 0,015 | 1 | 2,5 | 0,83 |
| sílaba | 0,015 | 1 | 2,5 | 0,83 |
| articulación | 0,015 | 1 | 2,5 | 0,85 |
| labio | 0,013 | 1 | 2,5 | 0,87 |
| plosiva | 0,013 | 1 | 2,5 | 0,88 |
| palatar | 0,011 | 1 | 2,5 | 0,9 |
| legible | 0,011 | 1 | 2,5 | 0,9 |
| lingua-dental | 0,009 | 1 | 2,5 | 0,92 |
| labio-dental | 0,008 | 1 | 2,5 | 0,95 |
| producción | 0,008 | 1 | 2,5 | 0,96 |

Es posible observar que los vocablos en español son 17, acumulando una frecuencia absoluta total de 24 y un promedio de aparición del 3,5%. Los vocablos más usados son "fonema", con 5 apariciones, y "pronunciación", con 3. En comparación, hay 104 vocablos en inglés, con una frecuencia absoluta total de 289. Por lo tanto, los vocablos en lengua inglesa constituyen un 83,2% de la producción lingüística elicitada. A diferencia de los centros de interés anteriores, aquí la palabra con mayor IDL, aparece en un lugar de la tabla en que la frecuencia acumulada está en el orden de 0,35. Y luego, todo el resto de palabras

aparece cuando la frecuencia acumulada es ya superior a 0,5. Esto nos indica que a pesar de tener un comportamiento parecido, no es idéntico, y el español acá es usado de forma ligeramente más abundante que en los centros de interés anteriores.

4.2.3.2 Número de respuestas por participante

Es posible indagar un poco más en esto, si consideramos el número de respuestas por participante:

Tabla 4.21. Número de respuestas por participante, para centro de interés "Fonología y Fonética".
Fuente: *Elaboración propia*

| | Tercer Semestre | Séptimo Semestre |
|-----------------|------------------------|-------------------------|
| 1 | 9 | 11 |
| 2 | 9 | 15 |
| 3 | 9 | 8 |
| 4 | 5 | 11 |
| 5 | 7 | 4 |
| 6 | 9 | 11 |
| 7 | 6 | 12 |
| 8 | 5 | 6 |
| 9 | 7 | 8 |
| 10 | 8 | 14 |
| 11 | 7 | 5 |
| 12 | 4 | 6 |
| 13 | 6 | 4 |
| 14 | 8 | 10 |
| 15 | 9 | 3 |
| 16 | 3 | 2 |
| 17 | 13 | 7 |
| 18 | 11 | 5 |
| 19 | 13 | 10 |
| 20 | 4 | 9 |
| Total | 152 | 161 |
| Promedio | 7,6 | 8,05 |
| Mediana | 7,5 | 8 |

En esta ocasión, la diferencia entre ambos grupos no es tan marcada, a pesar de existir sistemáticamente indicadores más elevados para el grupo de séptimo semestre. Tienen una diferencia de 9 palabras totales (5,9%, lo que es estadísticamente significativo), y por lo tanto, sus promedios son parecidos (7,6 y 8,05). Asimismo, la mediana no muestra una variación demasiado grande respecto al promedio, por lo que la distribución de frecuencias es bastante pareja.

4.2.3.3 Índice de Disponibilidad Léxica Individual

Podemos profundizar aún más en esto, al estudiar el IDL para cada participante, separado por grupo:

Tabla 4.22. IDL individual para centro de interés “Fonología y Fonética”. *Fuente: Elaboración propia*

| | Tercer Semestre | Séptimo Semestre |
|-----------------|-----------------|------------------|
| 1 | 0,69 | 0,57 |
| 2 | 0,63 | 0,51 |
| 3 | 0,89 | 0,17 |
| 4 | 0,78 | 0,86 |
| 5 | 1,05 | 0,68 |
| 6 | 0,96 | 0,71 |
| 7 | 0,9 | 0,94 |
| 8 | 0,72 | 0,99 |
| 9 | 0,43 | 0,48 |
| 10 | 0,75 | 0,72 |
| 11 | 0,83 | 0,6 |
| 12 | 0,38 | 0,22 |
| 13 | 0,5 | 0,44 |
| 14 | 0,74 | 0,25 |
| 15 | 0,78 | 0,47 |
| 16 | 0,15 | 0,46 |
| 17 | 1,07 | 0,24 |
| 18 | 0,69 | 0,67 |
| 19 | 0,64 | 0,54 |
| 20 | 0,86 | 0,67 |
| Total | 14,45 | 11,2 |
| Promedio | 0,72 | 0,56 |
| Mediana | 0,75 | 0,56 |

Es posible observar que, a pesar de tener una menor cantidad neta de palabras, y un mediana menor, los participantes del tercer semestre, tienen un IDL promedio mayor que los participantes del séptimo semestre, lo que implica que utilizan palabras con mayor disponibilidad léxica, dado que la cantidad de las mismas es menor en el tercer semestre que en el séptimo.

4.2.3.3 Clusters

Tal como el examen de frecuencia de términos parecía indicar, se encontraron varios clusters en los datos. Son los siguientes:

Tabla 4.23. Clusters para centro de interés “Fonología y Fonética”. *Fuente: Elaboración propia.*

| Cantidad | Cluster |
|----------|------------------------------------|
| 2 | segmental, phonetic, transcription |
| 2 | phonetic, transcription, vowel |
| 2 | transcription, phoneme, vowel |
| 2 | suprasegmental, vowel, consonant |

Se trata de 4 clusters de tres términos cada uno, y con una repetición de 2 veces en todos. Esto nos indica que los participantes tienden a relacionar semánticamente dichos términos. El término dominante es “transcription”, apareciendo en 3 de los cuatro clusters. En esos mismos clusters, también están términos muy parecidos: “phonetic” y “phoneme”. El clúster final, es el único que no incluye ni phoneme o sus derivados, ni tampoco “transcription”. Todo esto, nos puede dar una idea de la forma en que esta información es relacionada por los participantes, y la manera en que se aproximan al centro de interés en cuestión.

4.2.4 Centro de Interés: Discurso.

4.2.4.1 Resultados Generales

En la tabla que sigue, se presentan los resultados obtenidos para cada vocablo en la muestra total, ordenados en forma descendente, de acuerdo a su Índice de Disponibilidad Léxica, y donde también se expone la frecuencia (absoluta y acumulada) de aparición de cada vocablo, y su porcentaje de aparición. Nuevamente, aunque a los participantes se les dio la libertad de responder en el idioma de su elección, éstos respondieron abrumadoramente en inglés. A continuación, se presentan los resultados obtenidos.

Tabla 4.24. Índice de Disponibilidad Léxica – Discurso. Fuente: *Elaboración propia.*

| Vocablo | IDL | Frecuencia | % Aparición | Frec. Acumulada |
|---------------|-------|------------|-------------|-----------------|
| paragraph | 0,21 | 10 | 25,0 | 0,05 |
| oral | 0,107 | 5 | 12,5 | 0,08 |
| vocabulary | 0,1 | 4 | 10,0 | 0,1 |
| talk | 0,097 | 4 | 10,0 | 0,12 |
| presentation | 0,085 | 4 | 10,0 | 0,14 |
| vocabulario | 0,072 | 3 | 7,5 | 0,16 |
| context | 0,069 | 3 | 7,5 | 0,17 |
| essay | 0,069 | 3 | 7,5 | 0,19 |
| formal | 0,067 | 4 | 10,0 | 0,21 |
| informal | 0,058 | 4 | 10,0 | 0,23 |
| message | 0,055 | 3 | 7,5 | 0,24 |
| speaker | 0,053 | 3 | 7,5 | 0,26 |
| accuracy | 0,05 | 2 | 5,0 | 0,27 |
| expositivo | 0,05 | 2 | 5,0 | 0,28 |
| speak | 0,047 | 2 | 5,0 | 0,29 |
| intention | 0,047 | 2 | 5,0 | 0,3 |
| written | 0,047 | 2 | 5,0 | 0,31 |
| speech | 0,044 | 2 | 5,0 | 0,32 |
| spoken | 0,043 | 2 | 5,0 | 0,33 |
| argumentativo | 0,043 | 2 | 5,0 | 0,34 |
| writing | 0,038 | 2 | 5,0 | 0,35 |

| | | | | |
|------------------------------|-------|---|------|------|
| argumentative | 0,036 | 3 | 7,5 | 0,37 |
| topic | 0,036 | 2 | 5,0 | 0,38 |
| utterance | 0,034 | 2 | 5,0 | 0,39 |
| conclusion | 0,034 | 3 | 7,5 | 0,4 |
| sentence | 0,032 | 2 | 5,0 | 0,41 |
| word | 0,031 | 2 | 5,0 | 0,42 |
| talking | 0,031 | 2 | 5,0 | 0,44 |
| topic-sentence | 0,03 | 2 | 5,0 | 0,45 |
| audience | 0,03 | 4 | 10,0 | 0,47 |
| introduction | 0,03 | 2 | 5,0 | 0,48 |
| descriptive | 0,03 | 3 | 7,5 | 0,49 |
| conjunction | 0,025 | 1 | 2,5 | 0,5 |
| fluidiez | 0,025 | 1 | 2,5 | 0,5 |
| paragraphs | 0,025 | 1 | 2,5 | 0,51 |
| adjective | 0,025 | 1 | 2,5 | 0,51 |
| literary-circle | 0,025 | 1 | 2,5 | 0,52 |
| academic-writing | 0,025 | 1 | 2,5 | 0,52 |
| speaking | 0,025 | 1 | 2,5 | 0,53 |
| theme | 0,025 | 1 | 2,5 | 0,53 |
| reported | 0,025 | 1 | 2,5 | 0,54 |
| indirect-speech | 0,025 | 1 | 2,5 | 0,54 |
| supportive-sentence | 0,024 | 2 | 5,0 | 0,55 |
| lesson-plan | 0,022 | 1 | 2,5 | 0,56 |
| descriptive-paragraph | 0,022 | 1 | 2,5 | 0,56 |
| parafrasear | 0,022 | 1 | 2,5 | 0,57 |
| parts-of-speech | 0,022 | 1 | 2,5 | 0,58 |
| advertising | 0,022 | 1 | 2,5 | 0,58 |
| coinage | 0,022 | 1 | 2,5 | 0,59 |
| reading | 0,022 | 1 | 2,5 | 0,59 |
| narrative-discourse | 0,022 | 1 | 2,5 | 0,6 |
| body-of-text | 0,022 | 1 | 2,5 | 0,6 |
| orden | 0,022 | 1 | 2,5 | 0,61 |
| conversation | 0,022 | 1 | 2,5 | 0,61 |
| communication | 0,022 | 1 | 2,5 | 0,62 |
| slang | 0,022 | 1 | 2,5 | 0,62 |
| idiom | 0,022 | 1 | 2,5 | 0,63 |
| eye-contact | 0,022 | 1 | 2,5 | 0,63 |
| direct-speech | 0,022 | 1 | 2,5 | 0,64 |
| information | 0,022 | 1 | 2,5 | 0,64 |
| discourse | 0,021 | 2 | 5,0 | 0,65 |
| palabra | 0,019 | 1 | 2,5 | 0,66 |

| | | | | |
|----------------------|-------|---|-----|------|
| desplante | 0,019 | 1 | 2,5 | 0,66 |
| exposition | 0,019 | 1 | 2,5 | 0,67 |
| process-paragraph | 0,019 | 1 | 2,5 | 0,67 |
| writing-skill | 0,019 | 1 | 2,5 | 0,68 |
| story | 0,019 | 1 | 2,5 | 0,68 |
| validez | 0,019 | 1 | 2,5 | 0,69 |
| time | 0,019 | 1 | 2,5 | 0,69 |
| fluency | 0,019 | 1 | 2,5 | 0,7 |
| dialect | 0,019 | 1 | 2,5 | 0,7 |
| informative | 0,018 | 2 | 5,0 | 0,72 |
| identity-theft | 0,016 | 1 | 2,5 | 0,72 |
| coherence | 0,016 | 1 | 2,5 | 0,73 |
| academic-paragraph | 0,016 | 1 | 2,5 | 0,73 |
| tense | 0,016 | 1 | 2,5 | 0,74 |
| write | 0,016 | 1 | 2,5 | 0,74 |
| estructura | 0,016 | 1 | 2,5 | 0,75 |
| argumentación | 0,016 | 1 | 2,5 | 0,75 |
| listener | 0,016 | 1 | 2,5 | 0,76 |
| academical | 0,016 | 1 | 2,5 | 0,76 |
| content | 0,016 | 1 | 2,5 | 0,77 |
| type | 0,016 | 1 | 2,5 | 0,77 |
| persuasive | 0,016 | 1 | 2,5 | 0,78 |
| expository-text | 0,016 | 1 | 2,5 | 0,78 |
| body-language | 0,016 | 1 | 2,5 | 0,79 |
| conclusion-sentence | 0,014 | 1 | 2,5 | 0,8 |
| rhythm | 0,014 | 1 | 2,5 | 0,79 |
| listening | 0,014 | 1 | 2,5 | 0,8 |
| comunicación | 0,014 | 1 | 2,5 | 0,81 |
| non-academical | 0,014 | 1 | 2,5 | 0,81 |
| interference | 0,014 | 1 | 2,5 | 0,82 |
| narrative | 0,014 | 1 | 2,5 | 0,82 |
| political | 0,014 | 1 | 2,5 | 0,83 |
| dialogue | 0,014 | 1 | 2,5 | 0,83 |
| purpose | 0,014 | 1 | 2,5 | 0,84 |
| accent | 0,014 | 2 | 5,0 | 0,85 |
| supporting-sentence | 0,012 | 1 | 2,5 | 0,85 |
| pause | 0,012 | 1 | 2,5 | 0,86 |
| transmisión | 0,012 | 1 | 2,5 | 0,87 |
| instruction | 0,012 | 1 | 2,5 | 0,87 |
| agreement | 0,012 | 1 | 2,5 | 0,88 |
| debate | 0,012 | 1 | 2,5 | 0,88 |

| | | | | |
|---------------------|-------|---|-----|------|
| target | 0,012 | 1 | 2,5 | 0,89 |
| concluding-sentence | 0,011 | 1 | 2,5 | 0,89 |
| voice | 0,011 | 1 | 2,5 | 0,9 |
| structure | 0,011 | 1 | 2,5 | 0,9 |
| social | 0,011 | 1 | 2,5 | 0,91 |
| receive | 0,011 | 1 | 2,5 | 0,91 |
| type-of-narrator | 0,011 | 1 | 2,5 | 0,92 |
| description | 0,009 | 1 | 2,5 | 0,92 |
| enviroment | 0,009 | 1 | 2,5 | 0,93 |
| mood | 0,009 | 1 | 2,5 | 0,93 |
| media | 0,009 | 1 | 2,5 | 0,94 |
| inform | 0,008 | 1 | 2,5 | 0,94 |
| text | 0,008 | 1 | 2,5 | 0,95 |
| language | 0,007 | 1 | 2,5 | 0,95 |
| funcion-of-language | 0,007 | 1 | 2,5 | 0,96 |
| pronounciation | 0,007 | 1 | 2,5 | 0,96 |
| meaning | 0,006 | 1 | 2,5 | 0,97 |
| perspective | 0,006 | 1 | 2,5 | 0,97 |
| genre | 0,005 | 1 | 2,5 | 0,98 |
| omnisciente | 0,004 | 1 | 2,5 | 0,98 |
| third-person | 0,003 | 1 | 2,5 | 0,99 |
| first-person | 0,003 | 1 | 2,5 | 0,99 |
| hook | 0,003 | 1 | 2,5 | 1 |

De acuerdo a la tabla, vemos que la palabra con mayor IDL, destaca claramente sobre todo el resto: “paragraph”, tiene un índice de disponibilidad léxica sobre 0,2, mientras la que le sigue está en el ámbito de 0,11, y la tercera palabra, tiene un IDL de 0,1, teniendo todas las palabras a continuación, un IDL menor a 0,1. Con todo, los 12 primeros términos de la lista, contienen el 25,9% de las menciones totales (50 en un total de 193), lo que nuevamente muestra una gran concentración en los términos más repetidos.

También podemos observar que el promedio de frecuencia de aparición es 1,53, mientras que la mediana es 1, tal como en el centro de interés anterior, lo que sigue reflejando el bajo índice de cohesión en la muestra.

Ahora, respecto a la frecuencia de aparición misma, 36 palabras (de un total de 126) tienen 2 menciones o más, 16 tienen 3 o más menciones, y 8 tienen 4 menciones o más. La palabra que destaca es “paragraph”, con 10 apariciones. Luego sigue “oral”, que tiene 5 menciones, y luego todo el resto tiene 4 o menos.

Si consideramos únicamente las palabras en español, obtenemos la siguiente tabla:

Tabla 4.25. Índice de Disponibilidad Léxica – Discurso, en español. Fuente: *Elaboración propia.*

| Vocablo | IDL | Frecuencia | % Aparición | Frec. Acumulada |
|---------------|-------|------------|-------------|-----------------|
| vocabulario | 0,072 | 3 | 7,5 | 0,16 |
| expositivo | 0,05 | 2 | 5,0 | 0,28 |
| argumentativo | 0,043 | 2 | 5,0 | 0,34 |
| fluidez | 0,025 | 1 | 2,5 | 0,5 |
| parafrasear | 0,022 | 1 | 2,5 | 0,57 |
| orden | 0,022 | 1 | 2,5 | 0,61 |
| palabra | 0,019 | 1 | 2,5 | 0,66 |
| desplante | 0,019 | 1 | 2,5 | 0,66 |
| validez | 0,019 | 1 | 2,5 | 0,69 |
| estructura | 0,016 | 1 | 2,5 | 0,75 |
| argumentación | 0,016 | 1 | 2,5 | 0,75 |
| comunicación | 0,014 | 1 | 2,5 | 0,81 |
| transmisión | 0,012 | 1 | 2,5 | 0,87 |
| omnisciente | 0,004 | 1 | 2,5 | 0,98 |

Hay 14 vocablos en español, con una frecuencia absoluta total de 18, y un promedio de aparición del 3,21%. Los vocablos más usados son "vocabulario", con 3 apariciones, junto con “expositivo” y “argumentativo”, con 2 apariciones cada uno. En comparación, hay 112 vocablos en inglés, con una frecuencia absoluta total de 193. Por lo tanto, los vocablos en lengua inglesa constituyen un 88,89% de la producción lingüística elicitada. Aquí se acentúa la tendencia observada en el centro de interés anterior, respecto a la frecuencia acumulada. La palabra con mayor IDL, aparece en un lugar de la tabla con una frecuencia acumulada del

orden de 0,155. Las dos palabras que siguen, aparecen cuando hay una frecuencia acumulada de 0,28 y 0,34. Recién la cuarta palabra aparece con una frecuencia acumulada de 0,5, y las que siguen ya están asociadas a frecuencias acumuladas mayores. Esto indica que alguna de las primeras tres palabras, debería representar un rol relativamente destacado en algún punto del análisis.

4.2.4.2 Número de respuestas por participante

Es posible indagar un poco más en esto, si consideramos el número de respuestas por participante:

Tabla 4.26. Número de respuestas por participante, para centro de interés “Discurso”.
Fuente: *Elaboración propia.*

| | Tercer Semestre | Séptimo Semestre |
|-----------------|-----------------|------------------|
| 1 | 3 | 6 |
| 2 | 2 | 8 |
| 3 | 1 | 6 |
| 4 | 3 | 4 |
| 5 | 1 | 11 |
| 6 | 7 | 14 |
| 7 | 2 | 11 |
| 8 | 2 | 4 |
| 9 | 9 | 8 |
| 10 | 4 | 17 |
| 11 | 1 | 6 |
| 12 | 3 | 5 |
| 13 | 4 | 3 |
| 14 | 2 | 2 |
| 15 | 2 | 6 |
| 16 | 2 | 2 |
| 17 | 1 | 4 |
| 18 | 1 | 2 |
| 19 | 6 | 6 |
| 20 | 4 | 8 |
| Total | 60 | 133 |
| Promedio | 3 | 6,65 |
| Mediana | 2 | 6 |

En este centro de interés, vemos que existe una diferencia muy marcada entre ambos grupos. De hecho, los participantes del séptimo semestre, más que duplican en producción a los del tercer semestre, con una diferencia del 122,17%.

Esto mismo se ve acentuado en la mediana de respuestas, donde los participantes del séptimo semestre, triplican a los del tercer semestre.

4.2.4.3 Índice de Disponibilidad Léxica Individual

Podemos profundizar aún más en esto, al estudiar el IDL para cada participante, separado por grupo:

Tabla 4.27. IDL individual para centro de interés "Discurso". Fuente: *Elaboración propia*.

| | Tercer Semestre | Séptimo Semestre |
|-----------------|-----------------|------------------|
| 1 | 0,28 | 0,28 |
| 2 | 0,28 | 0,24 |
| 3 | 0,21 | 0,1 |
| 4 | 0,1 | 0,26 |
| 5 | 0,03 | 0,2 |
| 6 | 0,17 | 0,31 |
| 7 | 0,12 | 0,26 |
| 8 | 0,17 | 0,18 |
| 9 | 0,39 | 0,22 |
| 10 | 0,25 | 0,13 |
| 11 | 0,07 | 0,15 |
| 12 | 0,1 | 0,12 |
| 13 | 0,3 | 0,08 |
| 14 | 0,14 | 0,09 |
| 15 | 0,17 | 0,18 |
| 16 | 0,04 | 0,15 |
| 17 | 0,03 | 0,11 |
| 18 | 0,03 | 0,04 |
| 19 | 0,16 | 0,13 |
| 20 | 0,24 | 0,2 |
| Total | 3,29 | 3,43 |
| Promedio | 0,16 | 0,17 |
| Mediana | 0,17 | 0,16 |

Al igual que en el centro de interés anterior, a pesar de que los participantes del tercer semestre tienen una cantidad neta de palabras y una mediana menor

(sensiblemente menor aquí), esto no se refleja de manera tan acentuada en el IDL promedio. Aunque no es tan marcada como en el centro de interés anterior, sí es posible ver que la diferencia entre IDL neto, es de 0,14 aproximadamente, esto es, de un 4,2%. Mientras que la mediana del tercer semestre es ligeramente superior a la del séptimo semestre, en 0,003. Esto es, en un 1,79%.

4.2.4.3 Clusters

Tal como el examen de frecuencia de términos parecía indicar, se encontraron varios clusters en los datos. Son los siguientes:

Tabla 4.28. Clusters para centro de interés “Discurso”. Fuente: *Elaboración propia*.

| Cantidad | Cluster |
|----------|----------------------------|
| 4 | formal, informal |
| 2 | argumentative, descriptive |
| 2 | intention, formal |
| 2 | expositivo, argumentativo |

Al igual que con el centro de interés anterior, tenemos 4 clusters, uno con cuatro repeticiones, y el resto con dos. Todos compuestos por dos términos. Aquí, el término “formal” se repite en dos clústers, mientras que “argumentativo” y “argumentative” son parte de los otros dos clústers. Además, algo muy interesante, es que aparece un clúster en español, cosa única en todos los centros de interés observados, y que se había previsto en el análisis inicial de frecuencias, debido a la forma en que esas palabras destacaban por su frecuencia de aparición. Como en los centros de interés anteriores, todo esto nos puede dar una idea de la forma en que esta información es relacionada por los participantes, y la manera en que se aproximan al centro de interés en cuestión.

CAPÍTULO 5 – DISCUSIÓN DE RESULTADOS

5. Discusión

5.1 Análisis y discusión de resultados

El objetivo de esta investigación, consiste en averiguar si existen diferencias entre la disponibilidad léxica de los estudiantes de los dos grupos examinados (Tercer y séptimo semestre). Si observamos la tabla 4.3, dado que las muestras son de igual tamaño, podemos notar que claramente hay mayor número de palabras producidas en 3 de los 4 centros de interés, en el grupo de séptimo semestre respecto al grupo del tercer semestre. Esto implica que sin duda la hipótesis de trabajo **se cumple**, pero deja también espacio para ciertos fenómenos que claramente difieren de la hipótesis.

Dentro de ello, lo más llamativo es que la producción léxica en el centro de interés “Morfología”, es superior en los alumnos de tercer semestre respecto a los de séptimo semestre. Esto, sin embargo, se puede relacionar con el hecho de que al observar la malla curricular de la carrera (ver Anexo 2), en el tercer semestre, vemos que los estudiantes se encuentran asistiendo al curso “Análisis Morfosintáctico de la Lengua Inglesa”, lo que implica que están en pleno proceso de ajuste de su curva del olvido (Murre & Dros, 2015). A diferencia de los estudiantes de séptimo semestre, cuya exposición más reciente a la morfología fue un semestre antes, en la asignatura “Análisis Contrastivo Inglés/Español”, lo que podría justificar esta llamativa diferencia.

Desde un punto de vista diacrónico, es posible observar dos fenómenos predominantes:

- 1) Aumento sistemático de la producción léxica en 3 de 4 centros de interés, desde tercer a séptimo semestre, y
- 2) Desplazamiento en la producción léxica, adquiriendo más importancia unos centros que otros, a medida que el tiempo avanza.

El fenómeno 1) ocurre acorde a lo esperable según la literatura, pues de acuerdo a Palapinidi (2012), un mayor tiempo de exposición a un léxico determinado, implica que dicho léxico aumenta en forma acorde en el participante expuesto, afirmación que se confirmaría en la mayoría de los resultados de la presente investigación.

Respecto al fenómeno 2) podemos ver que el énfasis en cada grupo es distinto: Para el tercer semestre, el centro de interés con mayor producción es Fonética y Fonología, en cambio, para el séptimo, es Gramática. En efecto, parece haber un desplazamiento en los centros con mayor producción desde el tercer semestre al séptimo: Gramática cambia de lugar con Fonética y Fonología, pasando del segundo al primer lugar, y Discurso cambia de lugar con Morfología, pasando del cuarto al tercer lugar. Esto posiblemente, también podría implicar un mayor énfasis en la pronunciación y producción hablada durante el tercer semestre, mientras que en el séptimo semestre el énfasis pareciera estar en la producción escrita, dado el intercambio de lugares observado.

Según Chomsky (1965, en Tinajero, 2010), existen los siguientes niveles de competencia lingüística:

- Nivel fonológico: Conocimiento de las reglas de combinación de fonemas, pronunciación de los morfemas que componen las palabras.
- Nivel sintáctico: Reglas que definen la combinación de morfemas y palabras, para convertirlas en unidades mayores, como oraciones y textos.
- Nivel semántico: Reglas que asignan significado a los enunciados, y que establecen relaciones semánticas entre ellos.
- Nivel léxico: Relativo a las propiedades y significados más relevantes de las palabras que constituyen una lengua.
- Nivel morfológico: Reglas para la formación de nuevas palabras.

Es interesante observar que el desplazamiento ocurrido entre centros de interés, se corresponde con un movimiento desde el nivel fonológico al nivel sintáctico o léxico, lo que implica que se produce en los estudiantes una evolución en el campo semántico al que ponen mayor atención.

Y dado que las competencias comunicativas complementan y dan sentido a las competencias lingüísticas, vemos que de acuerdo a Hymes (1965, en Llobera et al., 1995), las competencias comunicativas incluyen las siguientes sub-competencias:

- Lingüística o gramatical: Esto es, conocimiento sobre el sistema de funcionamiento de la lengua.

- Sociolingüística: Entrega mecanismos de adaptación al contexto y a la situación.
- Discursiva: Referente a la coherencia y la cohesión de las distintas clases de discurso.
- Estratégica: Referente a la interacción, los actos de habla, y las acciones para recuperar la comunicación tras interrupciones o dificultades en la misma.

Vemos entonces, que en general, las fortalezas en la producción de habla no se encuentran en el discurso, ni en la morfología para la presente muestra. Por lo tanto, podemos considerar que en este caso, hay un mayor énfasis en las competencias gramaticales o lingüísticas, aunque nada impide que posiblemente esto siga cambiando a medida que pase el tiempo.

Sin embargo, a pesar de tales variaciones, hay indicadores que no manifiestan diferencias tan grandes. Por ejemplo, los índices de cohesión son parecidos para ambos grupos. Y no sólo eso, si observamos el gráfico 2, vemos que los índices coinciden en los centros de interés de “Morfología” y de “Discurso”, aquellos que poseen menos producción léxica en ambos grupos.

Asimismo, es posible observar cierta correspondencia, entre los campos con mayor producción, y el valor del índice de cohesión. Así, “Fonética y fonología” posee el mayor índice de cohesión de los cuatro centros de interés para el grupo del tercer semestre, mientras que “Gramática”, posee por poco el mayor índice de cohesión para los participantes del séptimo semestre.

De todas maneras, los índices de cohesión son relativamente bajos en todos los centros de interés, para ambos grupos, lo que nos indica que existe una gran variedad de palabras distintas, producidas por los participantes. Esto implica que existe un núcleo bastante laxo de palabras recurrentes, lo que se puede comprobar fácilmente verificando los índices de disponibilidad léxica para cada palabra: en ningún caso son más de cinco las palabras que tienen asociado un alto índice de disponibilidad léxica en todos los centros de interés, para ambos grupos de participantes.

Esto nos lleva a considerar también el índice de riqueza léxica, donde de acuerdo a la tabla 4.7, podemos notar que existen dos centros de interés donde un grupo de participantes manifiesta una marcada superioridad sobre el otro: Gramática (Tercer semestre sobre séptimo semestre) y Fonética y fonología (Séptimo semestre sobre tercer semestre).

Se trata nuevamente de datos coherentes con los resultados observados en los índices de cohesión y en la cantidad de producción observada. Sin embargo, hay un fenómeno muy interesante, y es que el índice de riqueza léxica total, no tiene variación significativa entre ambos grupos.

Esto, que parece paradójico, se puede explicar por el hecho de que ambos grupos se encuentran en una etapa de formación, y de adquisición de competencias lingüísticas respecto al tecnolecto que esta investigación estudia. Por lo tanto, el vocabulario de ambos grupos se encuentra en un período de crecimiento, cosa que muestran las diferencias en producción léxica entre ambos grupos.

Respecto al Índice de Disponibilidad Léxica Individual (ver gráfico 4.4), es posible apreciar que el grupo del tercer semestre tiene un IDL individual promedio, levemente mayor que el grupo del séptimo semestre (0,38 versus 0,42), lo que se debe a la gran diferencia en disponibilidad léxica promedio en el centro de interés “Fonología y fonética”, y también a la diferencia en el centro de interés “Morfología”, que son justamente los centros de interés donde la producción léxica del tercer semestre destacó.

Este resultado parece contraintuitivo, pues a pesar de que el grupo del séptimo semestre posee una mayor producción léxica neta (135 palabras, un 30% más), el IDL individual promedio, es inferior al del tercer semestre. Sin embargo, se trata de un resultado que corrobora lo ya observado por Cepeda, Granada & Pomes (2014), respecto a que “se desprende [...] que el índice de disponibilidad léxica no necesariamente equivale a la mayor cantidad de palabras emitidas”, efecto que se manifiesta claramente en los presentes resultados.

Respecto a los clusters, en el centro de interés “Gramática”, todos se relacionan con la lingüística prescriptiva, lo que hace pensar que este campo semántico, es asociado fuertemente con esa área de la lingüística, ligando la gramática con las reglas de uso para la variante de habla dominante y no con, por ejemplo, el funcionamiento del lenguaje propiamente tal.

En el centro de interés “Morfología”, al interior de los clusters predominan las palabras “adverb” y “verb”, lo que hace considerar que parece ser recurrente para los participantes, asociar la morfología con la construcción verbal y adverbial.

En cuanto al centro de interés “Fonética y fonología”, dominan las expresiones “transcription”, “phoneme” y “phonetic”, lo que hace considerar que

posiblemente se asocie esta área de la lingüística, con un aspecto netamente práctico y operativo, como es la transcripción del habla oral.

Respecto al centro de interés “Discurso”, los vocablos “argumentative”, “argumentativo” y “formal”, dominan los clusters. Esto nos muestra que existe una tendencia a considerar esta área de la lingüística, justamente desde un punto de vista netamente formal o descriptivo (de hecho la palabra con lejos mayor frecuencia, es “paragraph”), y no, por ejemplo, desde el punto de vista de la disciplina del análisis del discurso.

Aquí, además, se cuenta con la interesante novedad de que hay un cluster en lengua castellana, compuesto por las palabras “expositivo” y “argumentativo”, lo que va en la misma línea de lo indicado anteriormente, respecto a la forma en que se considera al campo semántico del discurso.

Finalmente, respecto al idioma de la producción léxica, resulta ser abrumadoramente en lengua inglesa, debido a que la instrucción en el aula es realizada casi totalmente en inglés, pues se procura someter a los estudiantes de la carrera a una inmersión en L2. Sin embargo, a pesar de ello, sí existe producción en español en todos los centros de interés, aunque bastante reducida, y limitada a una presencia más bien marginal dentro de la producción léxica.

Esta producción, además, sigue el mismo lineamiento de la producción en inglés de cada centro de interés, con la excepción quizás del centro de interés “Morfología”, que en español se enmarca más bien dentro de términos correspondientes al análisis morfosintáctico.

CAPÍTULO 6 – CONCLUSIONES

6. Conclusiones

6.1 Conclusiones

Los resultados obtenidos al realizar el presente estudio, nos muestran que la hipótesis de investigación, “Existe una diferencia significativa entre el tecnolecto de lingüística en alumnos de tercer y de séptimo semestre de pregrado”, **se comprueba como correcta.**

Asimismo, los resultados corroboran las conclusiones de la investigación de Urzúa, Sáez y Echeverría (2006), respecto a que "existe un crecimiento del léxico disponible por parte de los alumnos a medida que éstos aumentan sus años de estudio", pues resulta efectivo que hubo un aumento de un 30% en el léxico disponible entre el grupo que cursaba tercer semestre, y el que cursaba séptimo semestre.

Esto también está en línea con lo encontrado por Ferreira, Salcedo y Del Valle (2014), quienes indican que “los resultados evidencian de forma general un aumento en el promedio de palabras y palabras diferentes a medida que aumenta el nivel educativo”.

Sin embargo, este crecimiento no fue parejo, pues no ocurrió en todos los centros de interés, dado que en uno, “Morfología”, la producción léxica disminuyó. Además, se produjo un fenómeno bastante contraintuitivo: aunque el vocabulario producido fue mayor en el grupo de séptimo semestre, respecto al tercer semestre, el índice de disponibilidad léxica individual, al contrario, fue mayor en el tercer semestre, que en el séptimo semestre. Por lo tanto, no se manifestó una correlación lineal o directa entre IDL y cantidad de léxico disponible. Sin embargo,

este comportamiento que puede ser llamativo, en realidad corrobora y coincide con los resultados obtenidos por Urzúa, Sáez y Echeverría (2006), y también por Cepeda, Granada & Pomes (2014), respectivamente.

En efecto, es importante destacar que presumiblemente, un efecto ligado a la curva del olvido de Ebbinghaus (Murre & Dros, 2015), pudo haber influido en que específicamente el centro de interés de “Morfología” haya tenido mayor producción en el grupo del tercer semestre, que en el del séptimo semestre.

Por otra parte, existe escasa variación al comparar el índice de cohesión de los dos grupos estudiados, lo que se condice completamente con las conclusiones obtenidas en estudios anteriores (Ferreira, Salcedo y Del Valle, 2014; Cepeda, Granada & Pomes, 2014; Urzúa, Sáez & Echeverría, 2006), e implica que ambos grupos de participantes adquieren su léxico de forma semejante, a medida que aumenta su nivel educativo. Esto también se refleja en el hecho de que su índice de riqueza léxica es muy similar, lo que además constituye una manifestación de que en ambos grupos hay un proceso de aprendizaje de competencias no sólo léxicas, sino también comunicativas.

Al estudiar la distribución de clusters, es posible realizar una aproximación relativamente amplia a la mirada que los participantes tienen, respecto a los campos semánticos representados por cada centro de interés. La conclusión que se saca de ello, es que esta aproximación se encuentra fuertemente influida por el enfoque que se le da a la lingüística, como parte del rol pedagógico que los participantes tendrán en el futuro.

Así, el centro de interés “Gramática” está orientado hacia la lingüística prescriptiva, pues generalmente el rol de un profesor de inglés, consiste en

inculcar un habla específica en aprendices de L2, por lo que requiere – al menos como formación básica – de ese enfoque específico para realizar su labor con mayor eficacia.

El enfoque dirigido hacia la construcción verbal y adverbial en el centro de interés “Morfología”, también es consistente con la labor de enseñar a otros, la manera de construir formas verbales y adverbiales para comunicarse, una mirada práctica que también se repite en “Fonética y fonología” y en “Discurso”.

Ahora bien, esta aproximación a las competencias lingüísticas, es síntoma de una tensión bastante compleja de resolver, pues implica un problema de economía de tiempo, entendiéndolo como un recurso escaso que se ha de distribuir en la mejor forma posible. Esto, porque existen al menos tres necesidades lingüísticas fundamentales en el caso de los futuros docentes de pedagogía en inglés: La primera, es la transmisión de ciertas competencias lingüísticas a terceros, necesarias para el acto mismo de practicar la docencia en aprendices de L2. La segunda, son las habilidades lingüísticas que faciliten la labor cognitiva de comprender lo producido por la academia, respecto a su disciplina específica, así como de producir ellos mismos investigación. Y la tercera, son las competencias lingüísticas que devienen en competencias comunicativas dentro del campo semántico específico en cuestión. Dado que los estudiantes de pedagogía en inglés, deben centrarse además en el aprendizaje de otras competencias también, ligadas netamente a aspectos pedagógicos y de formación general, naturalmente no es posible dedicar íntegramente el tiempo a la satisfacción de estas tres necesidades lingüísticas fundamentales. Por lo tanto, corresponde escoger alguna. Y en este caso, naturalmente se ha escogido aquella

que complementa de manera más directa la práctica diaria de la docencia, que es la transmisión de competencias lingüísticas a terceros.

Por otra parte, el que los términos hayan pertenecido mayoritariamente a la lengua inglesa, resulta una consecuencia natural del régimen de inmersión al que están sometidos todos los estudiantes de la carrera de Pedagogía en Inglés. Ahora bien, es importante notar también, que a los estudiantes de cuarto año, se les prepara para rendir el test de competencias en inglés Cambridge English: Advanced (CAE), lo que también puede influir en la naturaleza de las competencias lingüísticas que desarrollan los participantes.

Los términos en español, de todas formas, representan claramente la orientación que se le da a los conceptos lingüísticos, acordes a una formación pedagógica, en especial en el centro de interés “Morfología”, donde se relacionan directamente con la asignatura de Análisis Morfosintático existente en la malla curricular.

Estos términos, junto con la importancia que tiene el centro de interés “Fonología y fonética” en el grupo de tercer año, también nos da cuenta de que se trata de participantes que están conviviendo entre el inglés y el español, lo cual es natural, pues habitan una comunidad donde la lengua totalmente dominante, es el español.

Finalmente, el hecho que los índices de cohesión sean bajos, implica que si bien hay una gran variedad de palabras distintas en la producción de los hablantes, también hay un vocabulario común relativamente reducido, lo que se explica por el hecho de que se encuentran en pleno período de consolidación de su léxico específico de lingüística.

De todas formas, el que los índices de cohesión y de riqueza léxica se mantengan parecidos, nos indica que existe consistencia al momento de recibir y mantener el tecnolecto de lingüística, por parte de los estudiantes, más allá de aspectos específicos que den una mayor variedad a la distribución de la producción léxica.

6.2 Limitaciones

Una de las limitaciones esenciales percibidas durante el desarrollo de la presente investigación, consistió en que la exposición al tecnolecto de lingüística no fue administrada de forma controlada, lo que generó algunas posibles distorsiones respecto a las respuestas, específicamente en el centro de interés “Morfología”.

Ello, debido a que las encuestas de disponibilidad léxica se tomaron a fines de semestre, en Junio, de manera que los estudiantes del tercer semestre ya habían sido expuestos durante todo un semestre al tecnolecto en cuestión. Por supuesto, esto implica una elección: ¿se les toma la prueba antes de ser expuestos al léxico específico, o después? Parece más lógico hacerlo después, o de otra forma, debemos considerar que al comienzo del tercer semestre, los participantes vienen llegando después de casi tres meses de vacaciones. Y dado que los estudiantes de séptimo semestre han sido expuestos por mucho más tiempo al tecnolecto, su consolidación sería mayor, por lo que resulta más adecuado al final del semestre, en condiciones más parejas de exposición al léxico.

Esto, por supuesto, no pudo evitar que de todas formas influyera mucho la curva de olvido de Ebbinghaus, respecto a la consolidación del léxico en los participantes, lo que generó la distorsión ya mencionada, respecto al centro de interés “Morfología”.

Por otra parte, debo hacer acá un reconocimiento a la excelente disposición de toda la Escuela de Pedagogía en Inglés de la Universidad Católica del Norte, quienes estuvieron muy lejos de constituir una limitación dentro de mi investigación, pues tuvieron la gentileza de coordinar y facilitar sin problemas el acceso tanto a las aulas, como a los participantes.

6.3 Proyecciones

Los resultados de la presente investigación, abren las puertas a numerosos aristas que se desprenden de ellos, pues existen muchos aspectos que ofrecen un punto de partida para futuras investigaciones. Por ejemplo:

- Aumentar el alcance de la presente investigación, midiendo esta vez la disponibilidad léxica de los estudiantes recién llegados a la carrera, para de esta forma, tener una mirada aún más transversal de la evolución de los participantes.

- Organizar una investigación con posibles ramificaciones socio y psicolingüísticas: Averiguar cuál es la posible relación entre cantidad de términos producidos, y el índice de disponibilidad léxica correspondiente.
- Investigar qué efecto puede tener en las competencias lingüísticas y comunicativas, una formación en inglés en un medio que habla principalmente español.
- Estudiar la impronta que producen asignaturas como por ejemplo, Análisis Morfosintáctico, en la percepción y uso del léxico específico de lingüística.
- Determinar cuánto influye la preparación para el test CAE en el desarrollo de las competencias lingüísticas en los alumnos de cuarto año de la carrera de Pedagogía en Inglés de la Universidad Católica del Norte.
- Realizar un seguimiento, y ampliar la investigación, aplicando el instrumento de encuesta de disponibilidad léxica, a profesionales egresados de la carrera, que estén en su primer año como docentes.

También sería interesante determinar, utilizando como base la información encontrada en la presente investigación, si es verdadero el aserto de Rico & Nikleva (2016), respecto a que el grado de dominio de la lengua en los “estudiantes repercutirá indiscutiblemente en la adquisición de las competencias y

capacidades vinculadas al resto de las materias que constituyen su currículo formativo”, en particular en el ámbito lingüístico.

De todas maneras, resulta esencial profundizar en el conocimiento respecto al léxico disponible de los futuros docentes, pues eso permitiría planificar y desarrollar programas de mejoramiento de las competencias léxicas en los distintos ámbitos donde fuera necesario. (Ferreira, Salcedo & del Valle, 2014), y esto, nos permitiría también, de paso, verificar la realidad de esta bella afirmación de Liliana Borrero:

“El desarrollar el elemento léxico en los estudiantes es fundamental y debiese trabajarse durante todo el proceso académico de los estudiantes. En este sentido [...] una enseñanza explícita del léxico potencia el incremento de vocabulario y por ende de las habilidades lingüísticas vinculadas a los procesos de lectura, escritura y todos los aprendizajes académicos en general”

(Borrero, 2008).

Y expandir aún más, nuestro campo de acción.

Antofagasta, Marzo de 2017.

CAPÍTULO 7 – BIBLIOGRAFIA

BIBLIOGRAFÍA

Aitchison, J. (1994). *Words in the mind: an introduction to the mental lexicon*. Oxford: Blackwell.

Alba, O. (1999). Densidad de anglicismos en el léxico disponible de la República Dominicana. En Samper Padilla et al. (eds.). *Actas del XI Congreso Internacional de la ALFAL*. Las Palmas de Gran Canaria: Universidad de Las Palmas de Gran Canaria, Tomo II, pp.853-865

Altmann, G. (1995). Lexical Statistics and cognitive models of Speech Processing. En Altmann, G. (ed.) *Cognitive Models of Speech Processing: Psycholinguistic and Computational perspectives*. Cambridge (Massachusetts) y London (England): MIT Press, 211-235.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Aristizabal, E. (1938). *Determination expérimentale du vocabulaire écrit pour servir de base à l'enseignement de l'orthographe a l'école primaire*. [s.l.: s.n.]

Atienza, E., Castro, D., Inglés, M., López, C., Martín, E., Pueyo, S. & Vañó, A. (2008). *Diccionario de términos clave de ELE*. Madrid: SGEL.

Bakonyi, H. (1934). *Die gebräuchlichsten Wörter der deutschen Sprache*. Munich: E. Reinhardt.

Baralo, M. (1997). La organización del lexicón en la lengua extranjera, *Revista de Filología Románica*, 14 (1), 59-71.

Baralo, M. (2001). El lexicón no nativo y las reglas de la gramática. En Pastor, S. & Salazar, V. (eds.) *Tendencias y líneas de investigación en adquisición de segundas lenguas*. Alicante: Universidad de Alicante, pp. 23-38

Barsalou W. L. (1992). *Cognitive Psychology: an Overview for Cognitive Scientists*. New Jersey, Hillsdale: Lawrence Erlbaum p. 221-244.

Bartol, J. (2010). Disponibilidad léxica y selección del vocabulario. En Castañer, R. & Lagüéns, V. *De moneda nunca usada: Estudios dedicados a José M^a Enguita Utrilla*. Zaragoza: Instituto Fernando El Católico, pp. 85-107

Beke, G. (1935). *French word book: tabulated and edited*. New York: Macmillan.

Bruning, R., Schorow, S. Norby, N. & Ronning, R. (2005). *Psicología cognitiva y de la instrucción*. Madrid: Pearson Educación.

Borrero, L. (2008). *Enseñando a leer: Teoría, práctica e intervención*. Colombia: Norma.

Buchanan, M. A. (1927). *A graded Spanish book*. Toronto: University of Toronto Press.

Burgess, C. & Lund, K. (2000). The dynamics of meaning in memory. En E. Dietrich & B. Arthur (Eds.) *Cognitive dynamics: Conceptual and representational change in humans and machines*. Mahwah, New Jersey: Erlbaum. pp. 117-156.

Cabré, M. (2015) [*Curso de Lexicología*]. Santa Clara, Cuba: Universidad Central "Marta Abreu".

Callealta, F. & Gallego, D. (2016) Medidas de disponibilidad léxica: comparabilidad y normalización. *Boletín de Filología*, 51 (1)

Campillos, L. & Ueda, H. (2015). Frecuencia y dispersión léxicas en textos médicos divulgativos en español. *Ibérica*, (30), 61-84

Carcedo, A. (1998). Tradición y novedad en las aportaciones hispánicas a los estudios de disponibilidad léxica. *Lingüística*, (10), 5-68.

Carcedo, A. (1999). Análisis de errores léxicos del español en la interlengua de los finlandeses. En: Jiménez Juliá et al. *Actas del IX Congreso Internacional de ASELE: Enfoque comunicativo y gramática*. Santiago de Compostela: Universidad de Santiago de Compostela, 465-472.

Carcedo, A. (2000). Índices léxico-estadísticos y graduación del vocabulario en la enseñanza de E/LE (aspectos culturales). En: Franco Figueroa, M. et al. (eds.). *Nuevas perspectivas en la enseñanza del español como lengua extranjera*. Cádiz: Universidad de Cádiz, tomo I, 174-183.

Carcedo, A. (2001). *Léxico disponible de Asturias*. Turku, Finlandia: Universidad de Turku.

Carrera, F. & Bradley, W. E. (2004). El proyecto sobre disponibilidad léxica en el mundo hispánico. Su aplicación en Cantabria. En Casado, M., Martínez, G. & Scurfield, S. (coord.). *Encuentro de lenguas y literaturas*. Santander: Universidad de Cantabria, p. 21-30.

Carroll, J.B. (1970). An alternative to Juilland's usage coefficient for lexical frequencies, and a proposal for a Standard Frequency Index (SFI). *Computer Studies in the Humanities and Verbal Behavior*, (3), 61-65.

Cepeda, M., Granada, M. & Pomes, M. (2014). Disponibilidad léxica en estudiantes de primero básico. *Literatura y lingüística*, (30).

Chomsky, N. (1995). Some Notes on Economy of Derivation and Representation. En *The Minimalist Program*. Massachussets : MIP Press. pp. 129-160.

Collins, A. M. & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82, 407–428.

Conrad, C. (1972). Cognitive economy in semantic memory. *Journal of Experimental Psychology*, (92) 149-154.

Cordero, S. (2009). Algunas consideraciones sobre el lenguaje común y el lenguaje técnico. *Káñina*, (33 Especial), 75-80

Cutler, A. (2006). Lexical Access. En Max Planck Institute for Psycholinguistics. *Encyclopedia of Cognitive Science*. Nijmegen, Paises Bajos: Max Planck Institute.

Dimitrijévic, N. (1969). *Lexical availability. A new aspect of the lexical availability of secondary school children*. Heidelberg: Julius Gross.

Dorta, A. (2015). *El tecnolecto de los profesionales de la Botánica: discurso científico escrito en lengua española*. Santa Clara, Cuba: Universidad Central "Marta Abreu".

Drummond, R. (2013). The Manchester Polish STRUT: Dialect Acquisition in a Second Language. *Journal of English Linguistics*, 41(1), 65-93

Dubois, J., Giacomo, M., Guespin, L., Marceixesi, C., Marcellesi, J. & Mével, J. (1998). *Diccionario de Lingüística*. Madrid: Alianza.

Echeverría, M. (1991). Crecimiento de la disponibilidad léxica en estudiantes chilenos de nivel básico y medio. En H. López Morales (ed.). *La enseñanza del español como lengua materna. Actas del II Seminario Internacional sobre Aportes de la lingüística a la enseñanza del español como lengua materna*. Río Piedras: Universidad de Puerto Rico, p. 61-78.

Echeverría, M., Herrera, M., Moreno, P. & Prádenas, F. (1987). Disponibilidad léxica en Educación Media. *Revista de Lingüística Teórica y Aplicada*. (25), 55-115.

Echeverría, M., Herrera, M. & Segure, M. (1980). *TEVI-R Test de vocabulario en imágenes: manual de aplicación*. Concepción: Universidad de Concepción.

Echeverría, M., Urzúa, P. & Figueroa, I. (2005). *DispoGen. Programa computacional para el análisis de la disponibilidad léxica*. Concepción: Universidad de Concepción [Software].

Echeverría, M. & Vargas, R. (2006). *DispoGrafo. Programa computacional para el análisis de relaciones semánticas*. Concepción: Universidad de Concepción [Software].

Echeverría, M., Vargas, R., Urzúa, P. & Ferreira, R. (2008). DispoGrafo: Una nueva herramienta computacional para el análisis de relaciones semánticas en el léxico disponible. *RLA* (46), 81-91

Eyzaguirre, B. & Fontaire, L. (2008). *Las escuelas que tenemos*. Santiago: Centro de estudios públicos.

Farah, M. J. & McClelland, J. L. (1991). A computational model of semantic memory impairment: Modality specificity and emergent category specificity. *Journal of Experimental Psychology: General*, (120): 339–357.

Ferreira, A., Salcedo, P., & del Valle, M. (2014). Estudio de disponibilidad léxica en el ámbito de las matemáticas. *Estudios filológicos*, 54.

Ferreira, R. & Echeverría, M. (2010). Redes semánticas en el léxico disponible de inglés L1 e inglés LE. *Onomazéin*, 21.

Figuerola, J., Carrasco, M. & Sarmiento, C (1982). Sobre la teoría de las redes semánticas. EN: *VI Encuentro Nacional y I Latinoamericano de Psicología*. Guadalajara

Forster, K. I. (1976). Accessing the mental lexicon. En R.J.Wales & E.Walker (Eds.), *New Approaches to Language Mechanisms*. (pp. 257-287). Amsterdam: North-Holland.

Galisson, R. (1979). *Lexicologie et enseignement des langues*. Paris: Hachette.

Galoso, M. (2004). Déficit léxico y comprensión lectora. *Puertas a la lectura*, (17), 35-37.

García, M. (2013). *El léxico disponible en estudiantes de 4º de educación secundaria obligatoria en Santander*. Santander. Universidad de Cantabria.

García, V. (1953). *Vocabulario usual, vocabulario común y vocabulario fundamental: determinación y análisis de sus factores*. Madrid: Instituto San José Calasanz

Gehrmann, R. (2016). *The west katuic languages: Comparative phonology and diagnostic tools*. Chiang Mai, Tailandia: Payap University.

Gómez, M. (2004). *La disponibilidad léxica de los estudiantes preuniversitarios valencianos: reflexión metodológica, análisis sociolingüístico y aplicaciones*. Valencia: Universidad de Valencia.

Gougenheim, G., Michéa, R., Rivenc, P. & Sauvageot, A. (1956). *L'élaboration du Français élémentaire: étude sur l'établissement d'un vocabulaire et d'une grammaire de base*. Paris:Didier.

Gries, S.T. (2008). Dispersions and adjusted frequencies in corpora. *International Journal of Corpus Linguistics*,13 (4), p. 403-437.

Henmon, V.A.C. (1924). *A French word book based on a count of 400,000 running words*. Madison, Wisconsin: University of Wisconsin.

Hernández, N. (2006). *Hacia una teoría cognitiva integrada de la disponibilidad léxica: El léxico disponible de los estudiantes castellano-manchegos*. Salamanca, España: Universidad de Salamanca.

Hymes, D. (1965). Acerca de la competencia comunicativa. En Llobera et al. (1995). *Competencia comunicativa. Documentos básicos en la enseñanza de lenguas extranjeras*. Madrid: Edelsa.

Jackson, G. T. & Graesser, A. C. (2006). Aplicaciones del diálogo humano de tutoría al AutoTutor: Un sistema inteligente de tutoría. *Signos*, 39, 31-48.

Juilland, A. & Chang-Rodríguez, E. (1964). *Frequency Dictionary of Spanish Words*. La Haya: Mouton.

Kaeding, F. W. (1897). *Häufigkeitwörterbuch der Deutschen Sprache: Festgestellt Durch einen Arbeitsausschuss der Deutschen Stenographiesysteme*. Berlin: E.S. Mittler und Sohn.

Landauer, T. K. & Dumais, S. T. (1997). A solution to Plato's problem: the latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104, 211-240.

Levelt, W., Roelofs, A. & Myers, A. (1999). A Theory of Lexical Access in Speech Production, *Behavioural and Brain Sciences*, 22, 1-75.

Lewis, M. (2000). *Teaching Collocations: Further Developments in the Lexical Approach*. Londres: Language Teaching Publications.

- Lin, J. (2012). El estudio de disponibilidad léxica de los estudiantes chinos de español como lengua extranjera. *Marcoele Revista de Didáctica*, 14, 1-14
- Lockett, J., & Shore, W. (2003). A narwhal is an animal: Partial word knowledge biases adults' decisions. *Journal of Psycholinguistic Research*, 32 (4), 477-496.
- López, H. (1973): *Disponibilidad léxica de los escolares de San Juan*. San Juan, Puerto Rico: Universidad de Puerto Rico.
- López, H. (1984). *Enseñanza de la lengua materna: Lingüística para maestros de español*. Madrid: Playor.
- López, H. (1996). Los estudios de disponibilidad: pasado y presente. *Boletín de Filología de la Universidad de Chile*. (24), 245-259.
- López, H. (1999). *Léxico disponible de Puerto Rico*. Madrid: Arco Libros.
- López, H. & García, F. (1995). Disponibilidad léxica en Andalucía. Proyecto de investigación. *Revista de Estudios de Adquisición de la Lengua Española (REALE)* 3, 65-76.

López, J. (1991). La enseñanza de la lengua materna. En: *Actas del II Seminario Internacional sobre "Aportes de la lingüística a la enseñanza de la lengua materna"*. Río Piedras, Puerto Rico: Universidad de Puerto Rico.

López, J. (1993). *El léxico disponible de escolares mexicanos*. Mexico: Alhambra Mexicana.

López, J. & Strassburger, C. (1987). Otro cálculo del índice de disponibilidad léxica. En: *Actas del IV Simposio de la Asociación Mexicana de Lingüística Aplicada*. México: UNAM.

Loran, R., & López, H. (1983). *Nouveau calcul de l'indice de disponibilité*. San Juan, Puerto Rico: University of Puerto Rico.

Luque, J. (2000). Introducción a la tipología léxica. En Marconi, B. *La competencia léxica*. Madrid: Antonio Machado Libros.

Luna, E., Viguera, A. & Báez, G. (2005). *Diccionario básico de lingüística*. México: UNAM.

Mackey, W. M. (1971): *Le vocabulaire disponible du français*. Paris: Didier.

Marslen-Wilson, W. (1993). Issues of Process and Representation in Lexical Access. En Altmann y Shillock (eds.) *Cognitive Models of Speech Processing: The second superlonga meeting*. Hillsdale, California: Laurence Erlbaum. pp. 187-220.

McKoon, G. & Ratcliff, R. (1995). Conceptual combinations and relational contexts in free association and in priming in lexical decision and naming. *Psychonomic Bulletin & Review*, 2 (4), 527-33

Michéa, R. (1949). Introduction pratique a une statistique du langage. *Les Langues Modernes*, 43, 173-186.

Michéa, R. (1950). Vocabulaire et culture. *Les langues Modernes*, 44, 188-189.

Michéa, R. (1953). Mots frequents el mots disponibles. Un aspecto nouveau de la statistique du langage. *Les langues modernes*, (47), 338-344.

Morton, J. 1969. Interaction of information in word recognition. *Psychological Review*, (76), 165–178

Moulin, C., Perfect, T., Conway, M. North, A., Jones, R. & James, N. (2002). Retrieval induced forgetting in Alzheimer's disease. *Neuropsychologia*, (40), 862–867.

Müller, C. (1968). *Initiation à la statistique linguistique*. Paris: Larousse.

Murre, J. & Dros, J. (2015). Replication and Analysis of Ebbinghaus' Forgetting Curve. *PLOS ONE*, 10 (7)

Nattinger, J. & DeCarrico, J. (1992). *Lexical phrases and language teaching*. Oxford: Oxford University Press.

Navarro, M. & San Martín, C. (2009) Emergencia y desarrollo de las habilidades metalingüísticas relacionadas con la regulación de la información: un estudio transversal en niños de edades comprendidas entre los 22 meses y los 10 años. *RLA*, 47 (1), 61-82

Ogden, C.K. (1930). *Basic English: A General Introduction with Rules and Grammar*. London: Paul Treber & Co.

Otheguy, R. & Zentella, A. C. (2012). *Spanish in New York: language contact, dialect leveling, and structural continuity*. New York: Oxford University Press.

Palapanidi, K. (2012). La aplicación de la disponibilidad léxica a la didáctica del léxico de LE. *Nebrija*, 11(6), 69-77.

Paredes, F. (2014). A vueltas con la selección de “centros de interés” en los estudios de disponibilidad léxica: para una propuesta renovadora a propósito de la disponibilidad léxica en ELE. *Nebrija*, (16).

Pastora, J. F. (1990). *El vocabulario como agente de aprendizaje*. Madrid: La Muralla.

Pérez, J. (2008). *El léxico de especialidad*. La Coruña: Universidad de La Coruña.

Potowski, K. (2011). Intrafamilial Dialect Contact. *The Handbook of Hispanic Sociolinguistics*, 579-597.

Quillian, M. R. (1968). Semantic memory. En Minsky, M. (Ed.) *Semantic information processing*. Cambridge, MA: MIT Press.

Ratcliff, R. & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review*, 95, 385–408.

Rico, A. & Nikleva, D. (2016). Análisis de la competencia lingüístico-discursiva escrita de los alumnos de nuevo ingreso del Grado de Maestro en Educación Primaria. *Signos*, 49(90), 48-70.

Salcedo, P., Ferreira, A., Valle, M. & Hernández, R. (2011). *Léxico disponible de estudiantes de Concepción Chile en las áreas de datos y azar, álgebra, números y geometría*. Concepción: Universidad de Concepción.

Samper, J. (2006). Disponibilidad léxica y sociolingüística. En Blas, J., Velando, M. & Casanova, M. (coord.). *Discurso y sociedad : contribuciones al estudio de la lengua en contexto social*. Castellón de la Plana, España: Universitat Jaume I

Samper, J. & Hernández, C. (1997). El estudio de la disponibilidad léxica en Gran Canaria: datos iniciales y variación sociolingüística. En: Almeida, M. & Dorta, J. (eds.). *Contribuciones al estudio de la lingüística hispánica. 301 Homenaje a Ramón Trujillo Carreño*. La Laguna: Universidad de La Laguna, tomo II, 229-239.

Samper, M. (2003). *Evolución de la disponibilidad léxica en estudiantes grancanarios de enseñanza primaria y secundaria*. Salamanca: Universidad de Salamanca.

Sánchez, J. (2006). El idiolecto y su traducción: tres ejemplos italianos. *Revista de la Sociedad Española de Italianistas* (3), 165-184

Seidenberg, M. & McClelland, J. (1989). A Distributed, Developmental Model of Word Recognition and Naming. *Psychological Review*, 96 (4), 523-568

Shelfelbine, J. (1990). Student factors related to variability in learning word meanings from context. *Journal of Reading Behavior*, 22 (1), 71-97.

Siegel, J. (2010). *Second dialect acquisition*. Cambridge: Cambridge University Press

Tagliamonte, S. A. & Molfenter, S. (2007). How'd You Get That Accent?: Acquiring a Second Dialect of the Same Language. *Language in Society*, 36 (5), 649-675.

Thorndike, E.L. (1921). *The teacher's world book*. New York: Columbia University.

Tinajero, L. (2010) Competencia en comunicación lingüística: Competencia Lingüística vs. Competencia Comunicativa. *Temas para la Educación*, (6).

Trujillano, P. (2015). *Disponibilidad léxica en estudiantes de 1º de secundaria del curso 2014-2015 en el I.E.S. Marismas (Santoña, Cantabria)*. Comillas, España: Universidad de Cantabria.

Urrutia, M. (2001). *Edad y acceso al léxico*. Concepción: Universidad de la Concepción.

- Urzúa, P., Sáez, K., & Echeverría, M. (2006). Disponibilidad léxica matemática: análisis cuantitativo y cualitativo. *RLA*, 44 (2), 59-76.
- Valencia, A. (1997). Disponibilidad léxica: Muestreos y estadísticos. *Onomazein*, 2, 197-226.
- Valencia, A. & Echeverría, M. (1999). Disponibilidad léxica en estudiantes chilenos. Santiago de Chile: Universidad de Chile y Universidad de Concepción.
- Vargas, P. (1991). La disponibilidad léxica: un diagnóstico al hablante medio de la V Región en seis centros de interés. *Nueva Revista del Pacífico*, (33-36)
- Vivas, J. (2004). Método Distsem: procedimiento para la evaluación de distancias semánticas. *Perspectivas en Psicología*, (1), 56-62.
- Vivas, J. (2007). Análisis de redes semánticas aplicado a contenidos académicos. Métodos e instrumentos. En Kronmüller, E. & Cornejo, C. (Comps.) *Ciencias de la Mente: Aproximaciones desde Latinoamérica*. Santiago de Chile: JC Sáez Editor.
- Vivas, J., Huapaya, C., Lizarralde, F., Arona, G., Comesaña, A., Vivas, L., & García, A. (2008). Distsem e Infosem: Instrumentos para la evaluación de la Memoria Semántica. Método y aplicaciones. En Rodríguez, M. & Padilla, V. (comps.) *Cognición y memoria, sus representaciones y mediciones*. Monterrey: Universidad Autónoma de Nuevo León

Vivas, J. (2009). *Evaluación de redes semánticas : Instrumentos y aplicaciones*.
Mendoza: EUDEM.

Wandmacher, T. (2005) How semantic is Latent Semantic Analysis?. En: *Récital 2005*. Dourdan: LIMSI-CNRS.

West, M. & Endicott, J. G. (1935). *The new method English dictionary: explaining the meaning of 24,000 items within a vocabulary of 1,490 words*. Londres: Longmans, Green and Co.

Yalcin, Seth. (2014). Semantics and Metasemantics in the Context of Generative Grammar. En: Burgess, A. y Sherman, B. (ed.). *Metasemantics : new essays on the foundations of meaning*. Oxford: Oxford University Press.

Informe de Profesores Revisores

8 – ANEXOS

ANEXO 1

Listado de Docentes de la Carrera de Pedagogía en Inglés, Universidad Católica del Norte, y sus asignaturas correspondientes.

Docentes consultados

Mauricio Lacazette – Director Escuela de Inglés UCN

M. Cecilia Ávila – Coordinadora de Investigación, Escuela de Inglés

Rosa Pezoa Tudela – Docente, Mg. en Ling. Aplicada

Chinger Zapata – Académico

Lucía Ramos – Académica

Macarena Rojas – Académica

Emily Noble – Académica

Asignaturas relacionadas con la lingüística, dictadas por los docentes consultados

Taller de Lengua Inglesa IV – Mauricio Lacazette

Lingüística General, Discurso Oral y Escrito en Inglés I - M. Cecilia Ávila

Gramática Inglesa Aplicada Elemental, Discurso Oral y Escrito en Inglés II–

Lucía Ramos

Fonética y Fonología Inglesa – Macarena Rojas

Fonética Aplicada Segmental, Fonética Aplicada Suprasegmental

- Emily Noble

Análisis Morfosintáctico de la Lengua Inglesa, Discurso Oral y Escrito en

Inglés III – Rosa Pezoa T.

Gramática Inglesa Aplicada Avanzada

Análisis Contrastivo Inglés/Español – Chinger Zapata

ANEXO 2

| MALLA CURRICULAR | | | |
|--|---|--|--|
| •Primer semestre | •Tercer semestre | •Quinto semestre | •Séptimo semestre |
| <ul style="list-style-type: none"> •Discurso Oral y Escrito en Inglés ALTE 1-A •Fonética Aplicada Segmental •Gramática Inglesa Aplicada Elemental •Antropología de la Educación •Informática Educativa •Discurso Oral y Escrito Español I | <ul style="list-style-type: none"> •Discurso Oral y Escrito en Inglés ALTE 2-A •Fonética Fonología Inglesa •Análisis Morfosintáctico de la Lengua Inglesa •Historia y Cultura Anglosajona •Psicología del Desarrollo y del Aprendizaje •Teoría Curricular •Formación General Teológica | <ul style="list-style-type: none"> •Discurso Oral y Escrito en Inglés ALTE 3-A •Literatura Inglesa: Poesía y Drama •Met. y Adquisición de 2o Idioma •Recursos Didácticos Tradicionales y Digitales •Evaluación del Proceso Pedagógico •Integrador: Planif. y Ev. Práctica I •Introducción al Trabajo Científico | <ul style="list-style-type: none"> •Taller de Lengua Inglesa ALTE 4-A •Didáctica de EFL E. Media •Orientación Educacional •Seminario de Investigación •Integrador: Profesor Inglés E.G. Básica Práctica III •Formación General Teológica |
| •Segundo semestre | •Cuarto semestre | •Sexto semestre | •Octavo semestre |
| <ul style="list-style-type: none"> •Discurso Oral y Escrito en Inglés ALTE 1-B •Fonética Inglesa Aplicada Suprasegmental •Gramática Inglesa Aplicada Avanzada •Fundamentos de la Educación •Introducción a la Psicología •Discurso Oral y Escrito Español II | <ul style="list-style-type: none"> •Discurso Oral y Escrito en Inglés ALTE 2-B •Lingüística General •Literatura Inglesa: Cuento y Novela •Planificación del Proceso Pedagógico e Integrador •Psicología de la Educación •Diseño y Ev. Proyecto Educativo | <ul style="list-style-type: none"> •Discurso Oral y Escrito en Inglés ALTE 3-B •Análisis Contrastivo Inglés/Español •Literatura Inglesa Contemporánea •Didáctica de EFL E.G. básica •Administración Educativa •Integrador: Rec. Didác., Práctica II •Met. Inv. Cualitativa-Cuantitativa | <ul style="list-style-type: none"> •Taller de Lengua Inglesa ALTE 4-B •Formación General Electiva •Ética y Moral Profesional •Integrador: Profesor Jefe e Inglés E. Media Práctica IV •Seminario de Investigación |