

© Copyright by

Burcu Mutlu

May, 2016

LA EVALUACIÓN DEL VOCABULARIO EN LA ENSEÑANZA Y EL
APRENDIZAJE DEL ESPAÑOL COMO SEGUNDA LENGUA

A Dissertation

Presented to

The Faculty of the Department

of Hispanic Studies

University of Houston

In Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

By

Burcu Mutlu

May, 2016

LA EVALUACIÓN DEL VOCABULARIO EN LA ENSEÑANZA Y EL
APRENDIZAJE DEL ESPAÑOL COMO SEGUNDA LENGUA

Burcu Mutlu

APPROVED:

Marta Fairclough, Ph.D.
Committee Chair

Manuel J. Gutiérrez, Ph.D.

Alejandra González-Pérez, Ph.D.

Encarna Bermejo, Ph.D.
Department of English and Modern Languages
Houston Baptist University

Steven G. Craig, Ph.D.
Interim Dean, College of Liberal Arts and Social Sciences
Department of Economics

LA EVALUACIÓN DEL VOCABULARIO EN LA ENSEÑANZA Y EL
APRENDIZAJE DEL ESPAÑOL COMO SEGUNDA LENGUA

An Abstract of a Dissertation

Presented to

The Faculty of the Department

of Hispanic Studies

University of Houston

In Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

By

Burcu Mutlu

May, 2016

Abstract

The principal purpose of this study is to suggest a valid, reliable and practical assessment tool, which can evaluate learners' vocabulary knowledge in Spanish as a Second Language. Two Vocabulary Levels Tests were designed based on a lexical corpus in order to evaluate students' receptive and productive vocabulary. The corpus consists of the active vocabulary presented in the first- and second-year textbooks used by the language program at the University of Houston. According to the data obtained from 410 students, although the receptive and productive vocabulary knowledge of students increases as they progress through the levels, the low frequency of high scores indicates the difficulty of these tests for the participants. It is observed that while the low and high beginners differ significantly from each other, these differences are not evident among low and high intermediates. On the other hand, it is noted that the emphasis placed on vocabulary as a component of language learning decreases gradually in the curricula and in the textbooks throughout the first 4 levels, which can account for the gaps, lack of recycling and retention of words studied in the previous levels. Moreover, strong correlations were found between the average of the two tests and students' final grades. It can be concluded that these vocabulary tests are good indicators of language proficiency in general and can be used as diagnostic tools. This study also presents what kind of learning strategies beginner and intermediate students prefer to use for studying and practicing vocabulary.

The results of the comparative frequency analysis supports the findings of previous studies (Davies & Face, 2006; Godev, 2009) in the sense that the textbooks only partially reflect the vocabulary most frequently used by native speakers of Spanish in

“real world” contexts and that these high frequency words are not presented systematically. Finally, pedagogical implications for expanding vocabulary and increasing lexical performance are offered for learners of Spanish as a Second Language.

Resumen

El objetivo principal de este estudio es proponer una herramienta fiable, válida y práctica que pueda evaluar la competencia léxica de los estudiantes de español como segunda lengua (EL2). Con el fin de evaluar el vocabulario receptivo y productivo de los estudiantes, se diseñaron dos Pruebas de Niveles de Vocabulario basadas en un corpus léxico formado por el vocabulario activo de los libros de textos usados en los primeros dos años del programa en la Universidad de Houston. Según los datos obtenidos de 410 alumnos, aunque existe un aumento en el léxico receptivo y productivo de los estudiantes a medida que avanzan de nivel dentro del programa, la baja frecuencia de puntajes altos indicó la dificultad de estas pruebas. Se observó que mientras los niveles principiantes difieren entre sí de manera estadísticamente significativa en su conocimiento léxico, no se encontraron diferencias significativas entre los niveles intermedios. Por otra parte, se notó que el espacio que ocupa el vocabulario como un componente del aprendizaje va bajando gradualmente en el plan de estudios y en los libros de texto a lo largo de los primeros 4 niveles. Esto indicó los vacíos, la falta de reciclaje y retención de las palabras estudiadas en los niveles anteriores. Además, se encontraron correlaciones fuertes entre los promedios de las dos pruebas y las notas finales de los participantes. Se puede concluir que estas dos pruebas son buenos predictores del rendimiento académico de EL2 y podrían funcionar como herramientas diagnósticas. Este estudio también presenta qué tipo de estrategias prefieren usar los estudiantes principiantes e intermedios a la hora de aprender y practicar el vocabulario.

Los resultados del análisis comparativo de frecuencia apoyaron los hallazgos encontrados en los estudios previos (Davies & Face, 2006; Godev, 2009) en el sentido de

que los libros de textos sólo parcialmente reflejan las palabras frecuentes usadas por los hablantes nativos en la vida real y en la presentación de éstas no se sigue un patrón sistemático. Al final, se sugirieron algunas implicaciones pedagógicas para expandir el vocabulario y aumentar el rendimiento léxico de los estudiantes de EL2.

Acknowledgements

My sincerest thanks go to my advisor, Dr. Marta Fairclough, for always sharing her experience and wisdom with me. Her linguistic expertise not only enlightened me, but also crystalize my thinking and shape my future studies. I have been amazingly fortunate to have an advisor who gave me the freedom to explore on my own, and at the same time the guidance to recover when my steps faltered. Also, a special thanks goes to my committee members, Dr. Manuel J. Gutiérrez, Dr. Alejandra González-Pérez and Dr. Encarna Bermejo whose constructive criticism and different perspectives have improved the final version, and helped to make it more complete. I have learnt so much from each and every one of them.

My heartfelt thanks and appreciation goes to my director Dr. Flavia Belpoliti and my TA colleagues for so willingly sharing their students with me and letting me collect my data from the Spanish as a Second Language Program at the University of Houston. I owe my gratitude to all these people who have made this dissertation possible.

Most importantly, none of this would have been possible without the love and support of my family. I would like to express my heart-felt gratitude to my parents, who have always loved me unconditionally and whose good examples have taught me to work hard for the things that I aspire to achieve. Without them; I would never have been able to complete this thesis project. I am grateful to their continuous encouragement and care.

My husband and my son, to whom this dissertation is dedicated to, were my biggest motivation for completing this study. My husband has been a constant source of

love, concern, support and strength throughout this endeavor. I deeply appreciate his patience and belief in me. I promise, I'll make up for the time we didn't spend together.

Índice General

1. Introducción	1
1.1. La competencia léxica y sus componentes desde diferentes perspectivas.....	2
1.1.1. La competencia léxica referencial/interferencial.....	8
1.1.2. Vocabulario receptivo y productivo desde diferentes perspectivas.....	9
1.1.3. Las dimensiones del conocimiento léxico desde diferentes perspectivas....	11
1.2. Evolución e importancia de la enseñanza del léxico en segundas lenguas.....	16
2. Estudios Previos	26
2.1. Los estudios previos del conocimiento léxico y las destrezas lingüísticas.....	26
2.2. Los estudios previos del tamaño y la profundidad del vocabulario.....	36
2.3. Estudios previos de frecuencia léxica y su utilidad en L2.....	41
2.4. Los estudios previos de las estrategias del vocabulario.....	47
3. Marco teórico	52
3.1. La competencia léxica desde el punto de vista psicolingüístico.....	52
3.2. ¿Qué significa saber una palabra?.....	55
3.2.1. Forma escrita.....	55
3.2.2. Partes de una palabra.....	56
3.2.3. Conexión entre forma y significado.....	57
3.2.4. Concepto y referentes.....	57
3.2.5. Asociaciones.....	58
3.2.6. Funciones gramaticales.....	59
3.2.7. Colocaciones.....	60
3.2.8. Restricciones de uso.....	61

3.3. La concepción del léxico de las palabras: unidades léxicas, tipos, lemas, familia de palabras.....	62
3.4. Las variables lingüísticas del estudio: el conocimiento receptivo y productivo de las palabras.....	63
4. Metodología.....	69
4.1. Participantes.....	69
4.2. Diseño del corpus.....	79
4.3. Instrumentos.....	89
4.3.1. Cuestionario sociolingüístico y de reflexión sobre el vocabulario.....	89
4.3.2. Prueba receptiva de niveles de vocabulario.....	90
4.3.3. Prueba productiva de niveles de vocabulario.....	97
4.4.4. Procedimiento.....	99
5. Análisis de Resultados.....	103
5.1. Prueba Receptiva de Niveles de Vocabulario (PR-NV).....	104
5.1.1. Estadísticas Descriptivas de la Prueba Receptiva de Niveles de Vocabulario (PR-NV).....	104
5.1.2. Fiabilidad y validez de la prueba receptiva de niveles de vocabulario (PR-NV).....	111
5.1.3. Análisis de varianza unidireccional de la prueba receptiva de niveles de vocabulario (PR-NV).....	113
5.2. Prueba Productiva de Niveles de Vocabulario (PP-NV).....	116

5.2.1. Estadísticas descriptivas de la prueba productiva de niveles de vocabulario (PP-NV).....	116
5.2.2. Fiabilidad y validez de la prueba productiva de niveles de vocabulario (PP-NV).....	122
5.2.3. Análisis de varianza unidireccional de la prueba productiva de niveles de vocabulario (PP-NV).....	124
5.3. Contrastes de medias entre las dos pruebas de niveles de vocabulario a través de Prueba <i>t</i> bilateral.....	126
5.4. Correlaciones.....	128
5.4.1. Correlaciones de Pearson entre la prueba productiva (PP) y receptiva (PR) de niveles de vocabulario (NV).....	128
5.4.2. Correlaciones de Pearson entre los promedios de las dos pruebas y las notas finales (NF).....	128
5.5. Análisis del cuestionario de reflexión sobre el vocabulario.....	130
5.6. Análisis comparativo de frecuencia entre los libros de texto del programa y <i>A Frequency Dictionary of Spanish (FDS)</i>	141
6. Conclusiones e implicaciones pedagógicas.....	147
Apéndices.....	161
Referencias.....	182

Lista de Tablas

Tabla 1. ¿Qué significa conocer una palabra?.....	7
Tabla 2. Tipos de conocimiento léxico.....	15
Tabla 3. Tamaños de vocabulario necesarios para una cobertura de 98% (incluyendo los nombres propios) de diferentes tipos de textos.....	28
Tabla 4. Grados y fuerza del conocimiento léxico.....	32
Tabla 5. Composición del corpus español moderno de 20 millones de palabras.....	43
Table 6. What is involved in knowing a word.....	66
Tabla 7. Distribución de los participantes por niveles y grupos.....	72
Tabla 8. Distribución de participantes por estudio formal de español.....	73
Tabla 9. Distribución de participantes según el promedio total de años de estudio de EL2.....	74
Tabla 10. Número y distribución de los participantes según sus antecedentes lingüísticos.....	77
Tabla 11. Distribución de los grupos por estudio de otra LE y estancia en un país Hispanohablante.....	78
Tabla 12. Los libros de textos y capítulos que forman la base del corpus de este estudio.....	80
Tabla 13. Algunos ejemplos de las palabras incluidas y excluidas del corpus.....	81
Tabla 14. Las palabras repetidas con diferentes significados semánticos en el corpus....	82
Tabla 15. Distribución de las categorías gramáticas según los niveles.....	88
Tabla 16. Distribución de categorías gramaticales según los niveles en la prueba receptiva.....	95

Tabla 17. Distribución de categorías gramaticales según los niveles en la prueba productiva.....	98
Tabla 18. Promedios de la PR-NV.....	106
Tabla 19. Estadísticas descriptivas de la PR-NV según los grupos.....	107
Tabla 20. Comparación de puntajes esperados con promedios en la PR-NV.....	109
Tabla 21. Distribución de los ítems léxicos según los niveles en la PR-NV.....	109
Tabla 22. Medidas de consistencia interna de fiabilidad para la PR-NV.....	111
Tabla 23. Estadísticas de fiabilidad según mitades partidas para PR-NV.....	112
Tabla 24. Resultados de <i>ANOVA</i> unidireccional para la PR-NV.....	113
Tabla 25. Pruebas post hoc de comparaciones múltiples con HSD Tukey para la PR-NV.....	114
Tabla 26. Promedios de la PP-NV.....	119
Tabla 27. Estadísticas descriptivas de la PP-NV según los grupos.....	120
Tabla 28. Comparación de puntajes esperados con promedios en la PP-NV.....	121
Tabla 29. Distribución de los ítems léxicos según los niveles en la PP-NV.....	121
Tabla 30. Medidas de consistencia interna de fiabilidad para la prueba productiva.....	123
Tabla 31. Estadísticas de fiabilidad según mitades partidas para la prueba productiva..	123
Tabla 32. Resultados de <i>ANOVA</i> para la PP-NV.....	124
Tabla 33. Pruebas post hoc de comparaciones múltiples con HSD Tukey para la PP-NV.....	125
Tabla 34. Resultados de los <i>ítems</i> 21-25 sobre la presentación del vocabulario en el libro de texto por el Grupo 2, 3, 4, 5.....	136

Tabla 35. Distribución de la familia de palabras presentadas en el programa según *A*
Frequency Dictionary of Spanish (FDS) de Mark Davies (2006).....142

Lista de Figuras

Figura 1. El modelo de frecuencia del aprendizaje de vocabulario de Meara (1992).....	37
Figura 2. Las instrucciones para el proyecto.....	101
Figura 3. Distribución de los promedios en la PR-NV por niveles.....	106
Figura 4. Distribución de los promedios en la PP-NV por niveles.....	118
Figura 5. Distribución de las respuestas dadas por grupos al ítem 18: ¿Qué tan importante es el vocabulario en el aprendizaje del vocabulario en EL2?.....	134
Figura 6. Las estrategias para aprender el vocabulario seleccionadas más y menos útiles entre los ítems 26 y 30 por los 6 grupos.....	137
Figura 7. Las estrategias para practicar el vocabulario seleccionadas más y menos útiles entre los ítems 32 y 39 por los 6 grupos.....	139
Figura 8. Frecuencia de familia de palabras presentadas en el nivel principiante-bajo según <i>A Frequency Dictionary of Spanish (FDS)</i> de Mark Davies (2006).....	144
Figura 9. Frecuencia de familia de palabras presentadas en el nivel principiante-alto según <i>A Frequency Dictionary of Spanish (FDS)</i> de Mark Davies (2006).....	144
Figura 10. Frecuencia de familia de palabras presentadas en el nivel intermedio-bajo según <i>A Frequency Dictionary of Spanish (FDS)</i> de Mark Davies (2006).....	145
Figura 11. Frecuencia de familia de palabras presentadas en el nivel intermedio-alto según <i>A Frequency Dictionary of Spanish (FDS)</i> de Mark Davies (2006).....	145

Dedication

To my household of love and inspiration:

My husband, Uno, who is my one and only superhero.

My son, Dalyan, who is the meaning and joy of my life.

Bulut and Darwin, whose paws and purrs kept me warm during lonely late nights of typing.

Capítulo 1

Introducción

If vocabulary knowledge is accepted as a fundamental component of second language proficiency, it is natural to accept that one of the primary goals of language testing will be to assess whether learners know the meaning of the words they need to communicate successfully in the second language. (Read, 1997, p. 303)

De acuerdo con algunos lingüistas (Gass & Selinker, 2008; Long & Richards, 1997; Milton, 2009), el léxico constituye el componente más importante en la adquisición de segundas lenguas¹ (L2), ya que puede determinar la calidad de competencia comunicativa de forma total, tanto de la expresión oral y escrita, como de la comprensión auditiva y lectora en un determinado contexto social y cultural (Larrotta 2011; Mokhtar, Rawian, Yahaya, et al., 2010). De hecho, la falta de dominio activo de vocabulario receptivo o productivo suele ocasionar más interrupciones, bloqueos y malentendidos en la comunicación que el desconocimiento o falsa aplicación de reglas gramaticales (Cervero & Pichardo, 2000).

Tradicionalmente se ha observado una relación estrecha en el ámbito escolar entre el conocimiento del vocabulario y dominio de la lengua en general. Según Núñez Delgado y Del Moral Barrigüete (2010), una de las frases más habituales de los profesores, para sintetizar las carencias en el dominio de la lengua propia o extranjera, suele ser la de que “Tienen muy poco vocabulario”, que viene a expresar la creencia de

¹ En este estudio el término *segunda lengua* (L2) se utilizará para referirse a una lengua que se aprende después de la materna (L1) sea cual sea el contexto del aprendizaje.

que un manejo amplio y preciso del léxico garantiza la posesión de destrezas comunicativas suficientes.

En este sentido, el concepto de competencia léxica es de gran utilidad para los profesores de idiomas, pues les ayuda a evaluar el aprendizaje, los manuales de L2, las pruebas y también les permite diagnosticar las dificultades de aprendizaje de sus alumnos (Casso, 2010).

En este capítulo introductorio, primero, se define el concepto de la competencia léxica, sus componentes y dimensiones desde diferentes perspectivas. Luego, se presenta brevemente la evolución e importancia de la enseñanza del léxico en segundas lenguas. Al final, se explican los objetivos y las preguntas de este estudio junto con un resumen del contenido de los siguientes capítulos.

1.1. La competencia léxica y sus componentes desde diferentes perspectivas

Jiménez Catalán (2002) apunta la falta de un consenso en cuanto al uso del término “competencia léxica”, puesto que utilizan términos diferentes para referirse a dicha competencia: *conocimiento de una palabra, conocimiento de vocabulario, conocimiento léxico, competencia semántica y competencia léxica*².

De acuerdo con la definición del Marco Europeo de Referencia para las Lenguas (MCERL, 2002), la competencia léxica es el conocimiento del vocabulario de una lengua y la capacidad para usarlo, y se compone de elementos léxicos y elementos gramaticales:

1) Los *elementos léxicos* o *unidades léxicas* incluyen:

A) Unidades fraseológicas o unidades pluriverbales, que se componen de varias palabras que se utilizan y se aprenden como un todo. Incluyen:

² En este estudio todos estos términos se utilizan de manera intercambiable, sin ninguna diferencia semántica.

- Fórmulas fijas: a) exponentes directos de funciones comunicativas (p. ej., *Encantado de conocerle, Buenos días*, etc.) ; b) refranes, proverbios, etc. ; c) arcaísmos residuales (p. ej., *Desfacen entuertos; Válgame Dios*)
- Modismos: a) metáforas lexicalizadas, semánticamente opacas (p. ej., *Estiró la pata* (murió). *Se quedó de piedra* (se quedó asombrado). *Estaba en las nubes* (no prestaba atención); b) intensificadores, ponderativos o epítetos (p. ej., *blanco como la nieve* (puro), *blanco como la pared* (pálido)); c) estructuras fijas aprendidas y utilizadas como conjuntos no analizados (p. ej., *Por favor, sería tan amable de + infinitivo...?*).
- Otras frases hechas como: a) verbos con régimen preposicional (p. ej., *convencerse de, alinearse con, atreverse a*; b) locuciones prepositivas (p. ej., *delante de, por medio de*).
- Régimen semántico: expresiones que se componen de palabras que habitualmente se utilizan juntas (p. ej., *cometer un crimen/error, ser culpable de (algo malo), disfrutar de (algo bueno)*).

B) Palabras: una palabra puede tener varios significados distintos (polisemia) (p. ej., *tanque*, un recipiente de líquido, o un vehículo armado y blindado; *banco*, lugar para sentarse o entidad financiera. Estas palabras incluyen miembros de las clases abiertas de palabras: sustantivo, verbo, adjetivo, adverbio, aunque éstas pueden incluir conjuntos léxicos cerrados (p. ej., días de la semana, meses del año, pesos y medidas, etc.). Se pueden establecer otros conjuntos léxicos con fines gramaticales y semánticos.

2) Los *elementos gramaticales* están compuestos por clases cerradas de palabras: artículos, cuantificadores, demostrativos, pronombres personales y relativos, adverbios interrogativos, posesivos, preposiciones, verbos auxiliares y conjunciones.

En el campo de estudios de adquisición/aprendizaje/enseñanza de L2, uno de los primeros intentos por definir la *competencia léxica* es el de Richards (1976). A través de su artículo titulado “The Role of Vocabulary Teaching”, el autor establece ocho supuestos, los cuales han servido de marco general para los estudios posteriores sobre la competencia léxica. Desde este planteamiento, saber una palabra significa:

- 1) The native speaker of a language continues to expand his vocabulary in adulthood, whereas there is comparatively little development of syntax in adult life.
- 2) Knowing a word means knowing the degree of probability of encountering that word in speech or print. For many words we also know the sort of words most likely to be found associated with the word.
- 3) Knowing a word implies knowing the limitations imposed on the use of the word according to variations of function and situation.
- 4) Knowing a word means knowing the syntactic behavior associated with the word.
- 5) Knowing a word entails knowledge of the underlying form of a word and the derivations that can be made from it.
- 6) Knowing a word entails knowledge of the network of associations between that word and other words in the language.

- 7) Knowing a word means knowing the semantic value of a word.
- 8) Knowing a word means knowing many of the different meanings associated with a word. (Richards, 1976, pp.78-83)

De estos supuestos propuestos por Richards (1976) se deduce que saber una palabra engloba una variedad de aspectos como conocer las probabilidades de encontrarse la palabra en un discurso oral o escrito, saber las limitaciones que distintos contextos imponen sobre ella, saber el comportamiento sintáctico asociado con la palabra, la forma de la raíz y los derivados que a partir de ella se generan, saber las asociaciones entre ella y otras palabras de la lengua, saber su valor semántico y saber los significados asociados con la palabra. De acuerdo con Pérez Basanta (1999), la descripción que hizo Richards sobre los componentes de la competencia léxica se considera como punto de partida para una nueva forma de entender la enseñanza del vocabulario.

Por otra parte, Jiménez Catalán (2002) proporciona una doble definición del término, según lo cual, por una parte se entiende por *competencia léxica* “el conocimiento que se debe poseer para poder utilizar la palabra con propiedad” y, por otra, “la capacidad de reconocer, aprender, recuperar y relacionar las distintas palabras a nivel oral y escrito”.

Las diferentes aportaciones que se han venido produciendo han apuntado a que la adquisición léxica en el aprendizaje de una segunda lengua no es un proceso lineal sino más bien multidimensional y lleno de muy diversas facetas. Esto indudablemente ha afectado no sólo al aprendizaje sino también a la concepción multidimensional de los contenidos de los syllabus (Pérez Basanta, 1998), y por supuesto a su evaluación (López-

Mezquita, 2007).

Un conocimiento adecuado de las palabras es un requisito previo al uso efectivo de la lengua (Read, 2000), y cuanto mayor es el vocabulario de un individuo mayor es su competencia comunicativa (Meara, 1996).

Jiménez Catalán (2002) analiza el concepto de competencia léxica en los estudios de aprendizaje y enseñanza de segundas lenguas y recopila definiciones de la competencia léxica a partir de los trabajos existentes desde 1975 hasta 2000. En la Tabla 1 el investigador presenta el esquema del modelo común de competencia léxica esbozado en los distintos estudios recopilados. En esta tabla se hace referencia no sólo a autores que han abordado la competencia léxica desde una perspectiva teórica y/o empírica sino también aquellos otros que o bien han intentado demostrar la utilidad de un marco referencial sobre la competencia léxica o bien se han referido a la misma en monografías didácticas.

Tabla 1. ¿Qué significa conocer una palabra?

A	B
Conocer su gramática, su pronunciación y su ortografía.	Blum-Kulka, Ellis & Sinclair, Laufer, Nation, Wallace.
Conocer su morfología.	Beheydt, Blum-Kulka, Faerch et al.; Laufer, Nation, Richards, Wallace.
Conocer su colocabilidad.	Beheydt, Blum-Kulka, Faerch et al.; Laufer, Nation, Richards.
Conocer sus restricciones sintácticas.	Beheydt, Blum-Kulka, Carter, Faerch et al.; Laufer, Nation, Richards.
Conocer su frecuencia en lengua oral y lengua escrita.	Nation, Laufer, Richards.
Conocer en qué contextos se puede utilizar.	Beheydt, Carter, Laufer, Faerch et al.; Richards, Robinson, Nation, Wallace.
Conocer sus relaciones semánticas y sintácticas con otras palabras.	Beheydt, Blum-Kulka, Carter, Ellis & Sinclair, Faerch et al.; Laufer, Nation, Richards, Wallace.
Reconocer la palabra en su forma oral o escrita.	Ellis & Sinclair, Laufer, Nation, Wallace.
Recuperar la palabra cuando se necesita	Ellis & Sinclair, Nation, Wallace.
Conocer su significado conceptual y referencial.	Blum-Kulka, Beheydt, Laufer, Nation, Richards, Robinson, Wallace.
Conocer los sentidos que connota.	Blum-Kulka, Laufer, Nation, Richards, Robinson, Wallace.
Conocer el sentido pragmático.	Blum-Kulka, Carter, Trampe.

Nota. Adaptado de “El concepto de competencia léxica en los estudios de aprendizaje y enseñanza de segundas lenguas,” por Jiménez Catalán, 2002, *Atlantis*, 24 (2), p.155.

El mismo autor clasifica los estudios relacionados con la definición de las dimensiones de la competencia léxica en dos tendencias principales.

- a) El enfoque lingüístico manifestado a través del análisis de la complejidad del significado y de la estructura de las palabras (Blum-Kulka, 1981; Trampe, 1983)
- b) El enfoque psicolingüístico relegado tanto en el análisis de la dificultad de adquirir la competencia léxica como en los distintos tipos de conocimiento y destrezas activadas en el aprendizaje (Laufer, 1991; Nation, 1990).

Por otro lado, la definición de competencia léxica de Chapelle (1994) se basa en el marco de competencia lingüística general propuesto por Bachman (1990) e incluye “both knowledge of language and the ability to put language to use in context”. Para Chapelle (1994) la competencia léxica está formada por tres componentes:

1. el conocimiento del vocabulario y sus procesos,
2. el contexto,
3. y las estrategias cognitivas para su uso (p. 163)

En síntesis, la identificación del concepto de competencia léxica es de gran utilidad para los investigadores del vocabulario en las segundas lenguas, considerando el hecho de que la evaluación del proceso de enseñanza/aprendizaje, el diseño de los materiales y las actividades de aprendizaje, las herramientas que permiten diagnosticar las dificultades y necesidades de estos procesos, todos dependen de ésta.

1.1.1. La competencia léxica referencial/interferencial. De acuerdo con Marconi (1997) existen dos tipos de competencia léxico-semántica. Mientras la competencia inferencial consiste en el uso hipotético de las palabras en la oración, la competencia referencial se refiere a la aplicación de la palabra en el mundo real. La competencia inferencial se entiende como la habilidad de manejar una red de conexiones entre palabras y realizaciones o procesos, tales como la inferencia semántica, la paráfrasis, la definición, encontrar un sinónimo, etc. Por el contrario, la competencia referencial es la capacidad para situar las unidades semánticas en el mundo real, es decir para aplicar las palabras a las situaciones reales de uso de la lengua tales como nombrar objetos, personas, hacer preguntas, obedecer órdenes o seguir instrucciones.

A pesar de que estos dos aspectos son independientes existe una cierta interacción

entre ellos. Con frecuencia los hablantes utilizan la competencia inferencial para llevar a cabo tareas referenciales cuando carecen de la competencia referencial directa especialmente en el caso de palabras no muy frecuentes. De la misma manera, la competencia inferencial se puede ampliar y reforzar cuando observamos los casos concretos en los que una determinada palabra se usa en el lenguaje real, y extraemos información referente al conjunto de todos los casos observados.

Por otra parte, existen estudios neuropsicológicos que aportan evidencias de que la separación entre ambas competencias es real, y así se ha comprobado en casos en los que una de las dos habilidades se pierde o daña, a consecuencia de alguna lesión cerebral, mientras que la otra puede no ser afectada en el mismo grado o incluso permanecer perfectamente intacta (López-Mezquita, 2007, p.66).

1.1.2. Vocabulario receptivo y productivo desde diferentes perspectivas. De acuerdo con Nation (2013), la distinción entre el conocimiento receptivo y el productivo del vocabulario en muchos casos depende de la distinción entre las destrezas receptoras de comprensión auditiva y lectora, y las destrezas productivas de expresión oral y escrita. El término *receptivo* transmite la idea de una entrada de lengua que se recibe desde otros individuos a través de la lectura o la escucha, en otras palabras, el individuo recibe la entrada e intenta comprenderla. Se usa el término *productivo* cuando es el propio individuo que produce entradas de una lengua al hablar o escribir. Los términos *pasivo* (para la comprensión auditiva y lectora) y *activo* (para la expresión oral y escrita) se usan como sinónimos de *receptivo* y *productivo* (Laufer, 1998; Meara, 1990; Nation, 2001). Sin embargo, aquí cabe señalar que algunos se oponen al uso intercambiable de estos términos en el sentido de que también existen habilidades productivas en las destrezas

receptivas (comprensión auditiva y lectora), mientras se escucha o se lee, se produce significado y uno no es realmente pasivo.

Con el fin de clarificar mejor estos conceptos, Schmitt (2010) prefiere usar los términos *reconocimiento del significado (meaning recognition)* y *recuperación del significado (meaning recall)* para el conocimiento receptivo, y *reconocimiento de forma (form recognition)* y *recuperación de forma (form recall)* para el conocimiento productivo. En consecuencia, su distinción entre vocabulario activo y pasivo se basa más en el uso del lenguaje que en el grado de conocimiento que tiene el individuo acerca de la unidad léxica en cuestión.

Por otra parte, Melka (1982) muestra el uso inconsistente de los términos *receptivo* y *productivo* con respecto a los ítems de prueba y grados de saber una palabra y considera que existe un continuo en cuyos extremos se sitúan ambas modalidades de conocimiento léxico, esta distinción debería considerarse como una escala de conocimiento.

Meara (1990) ve la distinción entre vocabulario receptivo y productivo como el resultado de distintos tipos de asociación entre las palabras. Según su interpretación, el vocabulario productivo se activa mediante la intervención de otras palabras, mientras que en el funcionamiento del vocabulario pasivo intervienen estímulos externos, es decir, se activa cuando se oye o se ven las formas de las palabras pero no a través de enlaces asociativos con otras palabras.

Otro punto de vista es el de Corson (1995), para el cual el vocabulario receptivo incluye todo el vocabulario productivo más otros tres tipos de vocabulario: las palabras que sólo se saben parcialmente, las palabras de baja frecuencia y las palabras que se

evitan voluntariamente en el uso activo. Este punto de vista de Corson se basa más en el uso de la lengua que en el grado de conocimiento de la unidad léxica en cuestión.

1.1.3. Las dimensiones del conocimiento léxico desde diferentes perspectivas.

En cuanto a las dimensiones del conocimiento léxico fundamentalmente se toman en cuenta dos aspectos: (1) el número de las palabras que los hablantes conocen, es decir, el *tamaño* o la *extensión* del vocabulario, y (2) el grado de *profundidad* que ese conocimiento alcanza, en otras palabras, su amplitud. Algunos investigadores como Read (1993), Nagy y Herman (1987) y Anderson y Freebody (1981) también usaron los términos *breadth* cuando se refiere al tamaño del vocabulario del estudiante, es decir, ¿cuántas palabras sabe el estudiante? y *depth* cuando se trata del nivel del conocimiento del vocabulario, en otras palabras, ¿con qué profundidad los estudiantes conocen la palabra?

The first (type of vocabulary knowledge) may be called *breadth* of knowledge, by which we mean the number of words for which the person knows at least some of the significant aspect of meaning... (There) is a second dimension of vocabulary knowledge, namely the quality of *depth* of understanding. We shall assume that, for most purposes, a person has a sufficiently deep understanding of a word if it conveys to him or her all of the distinctions that would be understood by an ordinary adult under normal circumstances (Anderson & Freebody, 1981, pp. 92-93).

Por otra parte, Laufer et al. (2004) propusieron otra dimensión denominada *fuerza del conocimiento léxico*, relacionada exclusivamente con el significado de la palabra y no con la profundidad de su conocimiento y que incluye aspectos como las propiedades

gramaticales, las colocaciones o la pronunciación.

Según Meara (1996), el tamaño del vocabulario es probablemente la única dimensión que tiene importancia en vocabularios pequeños que se componen de unas cinco o seis mil palabras:

The basic dimension of lexical competence is size. All other things being equal, learners with big vocabularies are more proficient in the wide range of skills than others with smaller vocabularies and there is some evidence to support the view that vocabulary skills make a significant contribution to almost all aspects of L2 proficiency. (Meara, 1996:37)

Una vez que se ha alcanzado cinco o seis mil palabras, el tamaño es menos importante y cobra importancia la cuestión de la profundidad con que se conocen las palabras. En la misma línea, Laufer (1998) considera que el tamaño es más relevante que la profundidad y un vocabulario extenso es crucial para el alumno, puesto que existen investigaciones válidas que justificaron que el tamaño del vocabulario es un buen indicador del nivel lingüístico en términos generales:

We realize that progress in vocabulary learning is not only a quantitative issue. Knowledge of words may progress from superficial to deep at various stages of learning. Yet we consider vocabulary size rather than depth to be crucial importance to learners. Vocabulary size was found to be a good predictor of reading comprehension (Anderson & Freebody, 1981; Koda, 1989, Coady et al., 1993) and to correlate well with writing quality (Linnarud, 1986; Astika, 1993). It is an important component of fluency in speech and learners themselves associate progress in language learning with an increase in the number of words they know

(Laufer, 1998, pp. 255-256).

Read (1993) también opina que el conocimiento del vocabulario es a menudo una cuestión de grado del conocimiento, e incluso los hablantes nativos en muchas ocasiones tienen sólo un conocimiento parcial de muchas de las palabras que saben.

Por otro lado Chapelle (1994) propone dos aspectos más a los tradicionalmente tratados y distingue en total cuatro dimensiones del conocimiento léxico:

1. El tamaño del vocabulario: el número de palabras que sabe una persona.
Desde un punto de vista comunicativo, Chapelle aboga por una medida del conocimiento del vocabulario dentro de un contexto y no aisladamente:
Vocabulary ability will differ qualitatively depending on the context in which it is used, and therefore must be specified with reference to that context.
2. El conocimiento de las características de la palabra que incluyen los componentes fonémicos, grafémicos, morfémicos, sintácticos, semánticos, pragmáticos, y rasgos colocacionales.
3. La organización del léxico referida a la forma en la que las unidades léxicas se almacenan en el cerebro: The metacognitive strategies used by developing second language learners to retrieve and manage vocabulary knowledge and processes are typically compensatory in nature. Those strategies are used in order to achieve communicative goals despite limited vocabulary knowledge.

Blum-Kulka and Levenston (1983) have identified the following strategies: circumlocution, paraphrase, language switch, appeal to authority, change of topic, and semantic avoidance. As ability increases and the learner has access to a larger number of words, these compensatory plans become less essential, but

metacognitive strategies remain important because any linguistic performance is the result of learners' successful metacognitive control (p. 167).

4. Los procesos fundamentales del vocabulario a través de los cuales los hablantes acceden a su léxico mental para la comprensión y para la expresión. Dichos procesos son mas rápidos y automáticos en los hablantes nativos que en estudiantes de L2 incluso de nivel avanzado, cuyo léxico mental está menos organizado que el de un nativo.

Según Read (2000), los dos últimos componentes a los que Chapelle se refiere tienen una aplicación menos directa en la evaluación de la competencia léxica; sin embargo, su importancia es mayor en las áreas de la investigación sobre la adquisición de una segunda lengua. Del mismo modo que Chapelle (1994), Qian y Schedl (2004) proponen un marco multidimensional para el conocimiento del vocabulario: 1) el tamaño del vocabulario, referido al número de palabras que el hablante conoce; 2) la profundidad del conocimiento léxico, que incluye todas las características léxicas mencionadas anteriormente ; 3) la organización léxica, referida al almacenamiento, conexión y representación de las palabras en el léxico mental; y 4) el automatismo del conocimiento receptivo y productivo, referido a los procesos fundamentales a través de los cuales se accede a las palabras con propósitos receptivos y por otra parte productivos.

La tercera dimensión del conocimiento léxico fue propuesta por Laufer et al. (2004) y denominada como la *fuerza del conocimiento léxico (strength of vocabulary knowledge)*. La *fuerza del conocimiento léxico* está relacionada exclusivamente con el significado de la palabra y no con la profundidad en el conocimiento, que incluye aspecto como las propiedades gramaticales, las colocaciones o la pronunciación. De acuerdo con

Laufer et al. (2004), las palabras son ante todo unidades de significado y el componente más importante del conocimiento léxico es la capacidad de establecer un vínculo entre la forma y el significado de una palabra y, así proponen cuatro grados de conocimiento del significado de una palabra basados en dos distinciones dicotómicas: 1) proveer la forma para un concepto dado vs. proveer el significado de una forma dada y 2) recuperación o reconocimiento (de forma o significado).

Tabla 2. Tipos de conocimiento léxico

	Recuperación	Reconocimiento
Activo (Productivo) (proveer la forma)	Recuperación activa	Reconocimiento activo
Pasivo (Receptivo) (proveer el significado)	Recuperación pasiva	Reconocimiento pasivo

Nota. Adaptado de “Size and strength: Do we need both to measure vocabulary knowledge?” Laufer, B., Elder, C., Hill, K., & Congdon, P., 2004, *Language Testing*, 21(2), p. 206.

A continuación cada tipo de conocimiento léxico ofrecido por Laufer, et al. 2004 (pp. 206-207) es ejemplificado a través de la palabra meta *melt* (=derretir):

1. Recuperación activa: el estudiante debe proveer la palabra buscada.

Turn into: m_____

2. Recuperación pasiva: el estudiante debe demostrar que conoce el significado de la palabra *melt* que aparece dentro de un contexto. En este caso existen varias respuestas aceptables como *water, fluid, liquid*.

When something *melts* it turns into _____

3. Reconocimiento activo: el estudiante debe elegir la palabra entre varias opciones.

Turn into water

a. elect b. blame c. Melt d. threaten

4. Reconocimiento pasivo: el estudiante debe elegir el significado apropiado de la palabra dada entre 4 opciones que se proporcionan.

Melt

a. choose b. accuse c. make threats d. turn into water

1.2. Evolución e importancia de la enseñanza del léxico en segundas lenguas

La enseñanza y aprendizaje del vocabulario nunca antes habían despertado el mismo nivel de interés de la competencia gramatical, el análisis contrastivo, la lectura o la comunicación escrita, los cuales durante mucho tiempo han recibido considerable atención de investigadores e instructores en la enseñanza de segundas lenguas (Richards, 1976). Meara (1996), en su ensayo “The dimensions of lexical competence”, critica a Canale y Swain (1980) por mencionar la competencia léxica simplemente en dos ocasiones en su artículo “Theoretical bases of communicative approaches to second language teaching and testing” al definir la competencia comunicativa, lo cual en la opinión de Meara (1980) tuvo un efecto negativo en el estudio sistemático de la competencia léxica.

Después de un periodo de negligencia con respecto a otros aspectos lingüísticos, en los últimos veinte años la investigación de los procesos de aprendizaje y enseñanza del vocabulario en segundas lenguas (L2) ha pasado de ocupar una posición marginal a una posición privilegiada dentro de los estudios de adquisición, aprendizaje y enseñanza de L2 (Nation, 2001; Jiménez, 2002; Gass & Selinker 2008). Según Kadota (2003), este cambio se puede atribuir al reconocimiento del papel fundamental que juega el vocabulario en la adquisición y aprendizaje de una lengua. Como Richards apunta en

1976 esta nueva visión va a tener importantes repercusiones pedagógicas sobre todo en el diseño de los sílabos:

The role of vocabulary teaching, which in the light of the findings concerning a deeper understanding of the nature of vocabulary in both the fields of linguistics and applied linguistics, raised the issue that the teaching of vocabulary should be redefined to fit into the way we design our syllabi. (Richards, 1976, p. 77)

Paul Meara fue una de las primeras voces que abogó por la necesidad de evaluar la competencia léxica en los años 70, con la finalidad no sólo de indagar en la cantidad y en la profundidad de la dimensión léxica de los individuos, sino también de establecer umbrales léxicos para los diferentes estadios del aprendizaje (López-Mezquita, 2007).

Long y Richards (1997) notaron que a partir de mediados de la década del 80 ha surgido un interés renovado en el rol del vocabulario en el aprendizaje de L2 y éste ha sido observado en el creciente número de estudios empíricos en temas como la naturaleza del lexicón bilingüe, la adquisición del vocabulario, el almacén léxico, la recuperación léxica y el uso del vocabulario por los estudiantes de L2. Krashen y Terrel (1983) también adoptaron la postura de que la competencia léxica constituye el núcleo esencial de la competencia comunicativa y señalaron que:

Vocabulary is also very important for the acquisition process. The popular belief is that one uses form and grammar to understand meaning. The truth is probably closer to the opposite: we acquire morphology and syntax because we understand the meaning of utterances. (Krashen & Terrel, 1983, p. 155)

Más adelante, la aparición del *Enfoque Léxico* a principios de los 90 puso por primera vez énfasis en el aspecto práctico de la enseñanza y del aprendizaje del

vocabulario de una segunda lengua, defendiendo el postulado de que: “Words carry more meaning than grammar, so words determine grammar” (Lewis, 1993, p. 38). Este enfoque léxico se alinea con lo que postula Peñalver en cuanto al aspecto léxico-semántico (1991, p.92):

Reconocer la importancia de la semántica y ponerla como punto de partida en lugar de punto de llegada es algo que nunca se había hecho por influencia de la gramática tradicional, que, como sabemos, relegaba el plano semántico y concedía toda la prioridad a la gramática (morfología y sintaxis), sobre la primera. Esta medida evidentemente ha beneficiado a la enseñanza de la lengua materna y hoy los especialistas no dejan de insistir en la gran importancia del plano léxico-semántico.

Hacia los finales de los años 90 y principios del 2000 aportaciones de investigadores como Jean Aitchison, Batia Laufer, Paul Meara, Paul Nation, John Read, Norbert Schmitt y Stuart Webb continuaron contribuyendo a que los estudios sobre el léxico alcanzaran el estatus del que goza hoy en día. De acuerdo con López-Mezquita (2007), otro importantísimo factor en esta tendencia ascendente ha sido la recopilación de corpus lingüísticos que cuentan con millones de palabras y la generación de listados de frecuencias de palabras a partir de ellos; la información que proporcionan dichos corpus ha contribuido al surgimiento de nuevas aproximaciones al estudio del vocabulario, con importantes repercusiones tanto en el campo de la lexicografía como en la metodología de la enseñanza de las lenguas, y más específicamente en la enseñanza del léxico.

Según Nation (2013), aproximadamente más de treinta por ciento de la investigación realizada en los últimos 110 años sobre el aprendizaje del vocabulario de

L1 y L2, se ha publicado en los últimos 11 años. Sin embargo, la mayoría de estos estudios se han llevado a cabo con datos recopilados principalmente de estudiantes de inglés como L2 (Collentine, Lafford & Karp, 2003; Carter & McCarthy, 1989; Coady & Huckin, 1997; French Allen, 1983; Nation, 1990; Schmitt & McCarthy, 1997). En este sentido, los estudios sobre la enseñanza y aprendizaje del vocabulario de español como segunda lengua (EL2) no han recibido la atención merecida y constituyen un área que hay que dinamizar.

Por otro lado, los implicados en el proceso de enseñanza/aprendizaje del EL2 (profesores, estudiantes, autores de manuales y materiales y editores) se preocupan fundamentalmente de cómo enseñar el vocabulario y pocas veces de qué vocabulario enseñar, cuestión que, al juicio de Benítez Pérez (2009), debe ser prioritaria. Dado que el hablante nativo no utiliza ni conoce todo el caudal léxico de su lengua, sino sólo un conjunto limitado de palabras en su comunicación diaria, resulta fundamental realizar una eficiente selección del vocabulario que se propondrá como meta en cada nivel.

En el Vocabulario del español hablado dirigido por Luis Márquez Villegas y publicado por la editorial SGEL en 1975, diccionario del español hablado que recoge las aproximadamente 4000 palabras más útiles, se especifica que en los cursos de enseñanza de lenguas extranjeras se deben introducir 1018 palabras en el nivel uno, 1198 en el dos y 1905 en el tres, lo que supone que un curso completo de idioma. Incluye 4121 palabras diferentes. Vale la pena señalar que según los estudiosos de frecuencia léxicas, las primeras 1000 palabras más usadas en español abarcan cerca del 85% del contenido del léxico de cualquier texto no especializado. De aquí la importancia de escoger correctamente el vocabulario

que debemos enseñar y el que debe aparecer en los manuales y materiales de apoyo. (Benítez Pérez, 2009, p.10).

Greve y Van Passel (1971) enumeran tres métodos principales para seleccionar el vocabulario que se debe enseñar en el aula: 1) la limitación subjetiva, basada en criterio personales; 2) limitación objetiva, por la que se seleccionan las unidades más frecuentes, 3) la limitación objetiva corregida, que añade a la frecuencia de aparición del vocablo su eficacia y su productividad. La primera es la que se ha venido utilizado tradicionalmente y así lo han puesto de manifiesto algunos análisis de manuales de lengua española, que incluyen, en ocasiones, palabras de dudosa utilidad para el estudiante dado que el hecho de que una forma se aprenda pero luego no se actualice en ejercicios y lecturas no favorece su aprendizaje. Esto evidencia que, en la enseñanza primaria, no se puede hablar *stricto sensu* de necesidades de los alumnos, sino más bien de intereses inmediatos, por lo que en un primer nivel de aprendizaje el niño debe llegar a conocer las unidades más usuales para que, más tarde, aprenda lo que necesite según para qué use la lengua que estudia (Salaberri, 1990). De la misma manera, en los niveles iniciales las unidades poco frecuentes o poco usadas deberían eliminarse y aquí es donde los estudios léxico-estadísticos ofrecen gran utilidad como herramienta de planificación de los contenidos léxicos pues permiten utilizar como criterios de selección tanto la frecuencia como la disponibilidad. De acuerdo con Sinclair y Renouf (1988), el criterio más importante para decidir si y cuándo un ítem debería estar incluido en un syllabus es frecuencia, no simplemente frecuencia de formas de palabras, sino también la frecuencia de los varios usos de estas formas y sus relacionadas formas flexionadas. La investigación basada en corpus es el procedimiento que subraya la especificación del contenido léxico y secuencia

léxica del syllabus.

En síntesis, según Núñez Delgado y Del Moral Barrigüete (2010), los principales objetivos de la investigación y estudio para una didáctica renovada del léxico deben estructurarse en dos grandes núcleos:

- a) Referidos al ámbito científico de la didáctica de la lengua: 1) reivindicar y poner de manifiesto la importancia del componente léxico como aspecto básico del dominio lingüístico; 2) elaborar fundamentaciones teóricas sobre el desarrollo tomando como referencia la aportaciones de las nuevas ciencias del lenguaje, de la ciencia cognitiva y de la psicología evolutiva y de la educación; 4) aportar evidencias empíricas sobre el hecho de que el aumento de la competencia léxica tiene repercusiones positivas en el dominio de la competencia comunicativa global; 5) comprobar la validez de nuevos enfoques para la enseñanza del léxico; 6) realizar análisis científicos sobre la precisión e imprecisión léxicas, sobre el léxico compartido y el léxico particular de cada género y sobre léxico disponible en cada centro de interés o campo semántico.
- b) Referidos al ámbito de la práctica de la enseñanza: 1) superar el modo tradicional descontextualizado de trabajar el léxico mediante métodos que conceden relevancia al contexto; 2) diseñar y aplicar programas para trabajar el léxico de modo específico y sistemático y con base en tareas de integración de destrezas. (p. 93)

A luz de estos principios sugeridos por Núñez Delgado y Del Moral Barrigüete (2010), la necesidad de prestar atención a la competencia léxica como parte esencial de la

comunicación, y de ofrecer algunas bases para desarrollarla, se presenta como una tarea urgente. Tomando en cuenta las ausencias y necesidades anteriormente mencionadas en el aprendizaje y la enseñanza de vocabulario en EL2, el objetivo principal de este estudio es proponer una herramienta fiable, válida y práctica que pueda evaluar la competencia léxica de los estudiantes de EL2. Más específicamente, el propósito de esta disertación es arrojar luz sobre las siguientes preguntas de investigación:

- 1) ¿Qué tipo de vocabulario se enseña en el primer y segundo año de un programa universitario de EL2 en Estados Unidos? ¿Cuántas palabras se enseñan en cada nivel?
- 2) En el orden de enseñanza de estas palabras, ¿se toman en cuenta sus frecuencias basadas en la producción de hablantes nativos de español en la vida real?
- 3) ¿Qué tanto éxito tiene el programa/los estudiantes en cuanto a la enseñanza /el aprendizaje de este vocabulario? ¿Existe un aumento en el conocimiento receptivo y productivo del vocabulario por parte de los estudiantes del programa EL2 en la Universidad de Houston a medida que avanzan de nivel dentro del programa? ¿Se puede hablar de diferencias estadísticamente significativas entre los primeros cuatro niveles del programa?
- 4) ¿Existe una correlación positiva entre el conocimiento léxico de los estudiantes y su competencia lingüística general en la lengua meta?
- 5) ¿Qué tipo de estrategias prefieren usar los estudiantes a la hora de aprender y practicar el vocabulario?

Para responder estos interrogantes de investigación, el estudio se organiza de la

siguiente manera:

En el Capítulo 2 se presentan algunos antecedentes dentro del campo de la adquisición del vocabulario en las segundas lenguas. En la sección 2.1 se exponen algunos estudios que han recalcado la importancia del conocimiento léxico en las destrezas lingüísticas. En la siguiente sección 2.2, se presentan estudios previos relacionados con el tamaño y la profundidad del vocabulario a través de diferentes herramientas. La frecuencia léxica y su utilidad en las segundas lenguas se acentúa a través de algunos estudios clave en el apartado 2.3 y finalmente, en la sección 2.4, se exponen algunos trabajos realizados sobre las estrategias de aprendizaje del vocabulario.

En el Capítulo 3, se establece el marco teórico a través de algunos conceptos clave que sirven de base a este estudio. Se define en detalle qué se entiende por conocer una palabra en el aspecto productivo y receptivo junto con la terminología básica que sirve como referente para la elaboración de los instrumentos. En primer lugar, se presenta (3.1.) la competencia léxica desde el punto de vista psicolingüístico. Enseguida, se explica (3.2) qué significa saber una palabra según el modelo de Nation (2013) a través de sus siguientes aspectos multifacéticos: (3.2.1) forma escrita, (3.2.2) partes de una palabra, (3.2.3) conexión entre forma y significado, (3.2.4.) concepto y referentes, (3.2.5) asociaciones, (3.2.6) funciones gramaticales, (3.2.7) colocaciones, (3.2.8) restricciones de uso. En el tercer apartado de este capítulo se define alguna terminología básica como (3.3) la concepción del léxico de las palabras a través de componentes léxicos, tipos, lemas, familia de palabras. En el cuarto apartado del Capítulo 3 se presentan en detalle (3.4) las variables lingüísticas del estudio: el conocimiento receptivo y productivo de las palabras.

En el Capítulo 4 se describe la metodología utilizada para llevar a cabo la investigación. La sección 4.1 presenta los participantes y sus características. En la sección 4.2. se detalla la formación del corpus léxico que sirve de base al diseño de los instrumentos del estudio. A ésta le sigue la sección 4.3 donde se presentan los tres instrumentos que se utilizaron para lograr los objetivos de este estudio: (4.3.1) el cuestionario sociolingüístico, (4.3.2) la prueba receptiva de niveles de vocabulario (*Receptive Vocabulary Levels Test*) y (4.3.3) la prueba productiva de niveles de vocabulario (*Productive Vocabulary Levels Test*) incluyendo la descripción de los mismos. Por último, la sección 4.4. resume los procedimientos aplicados en su implementación y evaluación.

El quinto capítulo está dedicado a la presentación de los resultados. Los datos obtenidos de 410 alumnos a través de los tres instrumentos: 1) la Prueba Receptiva de Niveles de Vocabulario (PR-NV), 2) la Prueba Productiva de Niveles de Vocabulario (PP-NV), y 3) el cuestionario de las estrategias de aprendizaje de vocabulario (EAV) fueron analizados utilizando el programa informático SPSS (Paquete Estadístico para las Ciencias Sociales). Este capítulo se encuentra dividido en siete apartados: en las primeras dos secciones (5.1. y 5.2) se presentan los datos individuales de la prueba receptiva y productiva de niveles de vocabulario respectivamente, para cada una se exponen : (5.1.1 y 5.2.2) el comportamiento global de las pruebas a través de las estadísticas descriptivas; (5.1.2 y 5.2.2) la fiabilidad y validez de la prueba receptiva y productiva; (5.1.3 y 5.2.3) los resultados de análisis de varianza unidireccional (ANOVA) con el fin de observar si los 6 grupos analizados en este estudio difieren significativamente entre sí en cuanto a sus medias y varianzas. En el siguiente apartado (5.3) se analizan los contrastes de medidas

entre las dos pruebas de niveles de vocabulario a través de Prueba *t* emparejada (*T-test*) para cada grupo; con el fin de corroborar la covariación, en la cuarta subsección (5.4.1), primero se exponen las Correlaciones de Pearson-r entre las dos pruebas de vocabulario, luego (5.4.2) se presentan las Correlaciones de Pearson-r entre el promedio de las dos pruebas y el promedio total obtenido del curso del español al final del semestre por cada participante; en seguida (5.5) se presentan los resultados del cuestionario de reflexión sobre las estrategias de aprender y practicar el vocabulario; finalmente, en el último apartado de este capítulo (5.6) se realiza un análisis comparativo entre las palabras que forman el corpus de este estudio y *A Frequency Dictionary of Spanish* de Mark Davies (2006).

Finalmente, en el Capítulo 6, Conclusiones e implicaciones pedagógicas, se exponen las conclusiones que se desprenden de esta investigación junto con las implicaciones pedagógicas, así como también se presentan algunas limitaciones del estudio y recomendaciones para futuras investigaciones.

Capítulo 2

Estudios Previos

Teachers and researchers have long recognized that lexical competence is at the heart of communicative competence (Meara, 1996, p. 35) and that vocabulary knowledge is a reliable predictor of learners proficiency in a second or foreign language. (Staeher, 2009, p.577)

En este capítulo se presentan algunos antecedentes dentro del campo de la adquisición del vocabulario en las segundas lenguas. En la sección 2.1 se exponen algunos estudios que han recalcado la importancia del conocimiento léxico en las destrezas lingüísticas³. En la siguiente sección 2.2, se presenta una serie de estudios previos relacionados con el tamaño y la profundidad del vocabulario a través de diferentes herramientas. La frecuencia léxica y su utilidad en las segundas lenguas se acentúan a través de algunos estudios clave en el apartado 2.3 y finalmente, en la sección 2.4, se exponen algunos trabajos realizados sobre las estrategias de aprendizaje del vocabulario.

2.1. Los estudios previos del conocimiento léxico y las destrezas lingüísticas

Algunos estudios han encontrado correlaciones estrechas entre el conocimiento léxico y diferentes destrezas lingüísticas. Por ejemplo, Laufer (1992) en su estudio analizó la relación entre el tamaño del vocabulario pasivo y la comprensión de textos académicos. 92 estudiantes de inglés para fines académicos, cuya primera lengua era

³ De acuerdo con el diccionario de términos clave del Centro Virtual Cervantes, la expresión *destrezas lingüísticas* hace referencia a las formas en que se activa el uso de la lengua. Tradicionalmente la didáctica ha clasificado atendiendo al modo de transmisión (orales y escritas) y al papel que desempeñan en la comunicación (productivas y receptivas). Así, se han establecido cuatro: expresión oral, expresión escrita, comprensión auditiva y comprensión lectora (para estas dos últimas se usan a veces también los términos de comprensión oral y escrita).

hebreo o árabe participaron en este estudio. Según los resultados de este estudio, se encontraron correlaciones altamente significativas entre la comprensión de lectura y los puntajes obtenidos en el vocabulario: 0.5 ($p < 0.0001$) en el caso de *Vocabulary Levels Test* y 0.75 ($p < 0.0001$) en el caso de *Eurocentres Vocabulary Levels Test*. Las diferencias más significativas se observaron en los puntajes de comprensión de lectura entre los grupos que conocían las primeras 2.000 y 3.000 familias de palabras más frecuentes. Las diferencias en los puntajes de comprensión de lectura entre los grupos de 3.000 y 4.000, 4.000 y 5.000, 3.000 y 5.000 niveles de vocabulario no fueron significativas. Laufer (1991, 1992 y 1998), a través de sus estudios, concluyó que un umbral de 3.000 familias de palabras o 5.000 *ítems* léxicos es necesario para una “comprensión mínima” de lectura, lo cual cubría 90-95% de cualquier texto; por debajo de ese umbral, las estrategias de lectura serían inefectivas. Para poder hacer consistentemente buenas suposiciones sobre el significado de las palabras desconocidas, uno debería saber 98% de las palabras en un texto. Para este tipo de cobertura, uno generalmente debe saber aproximadamente 5.000 familias de palabras (8.000 *ítems* léxicos) en inglés.

Nation y Beglar (2007) diseñaron El *Vocabulary Size Test* con el fin de medir el tamaño de vocabulario de los estudiantes de inglés como L2. Esta herramienta fue desarrollada para dar una estimación fiable y válida sobre el tamaño léxico de las primeras 1.000 a las 14.000 familias de palabras más frecuentes de los estudiantes de inglés. Según los diseñadores de esta herramienta, el tamaño de vocabulario de un estudiante de L2 puede tener un papel importante en determinar si el estudiante posee un repertorio de vocabulario suficiente para cumplir con eficacia actividades como leer una

novela, leer periódicos, ver películas, escuchar conversaciones cotidianas, etc. Nation (2006) propone diferentes tamaños de vocabulario necesarios para realizar estas tareas receptivas e indica que los estudiantes deben tener un repertorio léxico de cerca de 8.000 familias de palabras para realizar estas actividades. La siguiente tabla resume los tamaños de vocabulario necesarios para alcanzar 98% cobertura en diferentes tipos de textos.

Tabla 3. Tamaños de vocabulario necesarios para una cobertura de 98% (incluyendo los nombres propios) de diferentes tipos de textos

Textos	98% de cobertura	Nombres propios
Novelas	9.000 familias de palabras	1-2%
Periódicos	8.000 familias de palabras	5-6%
Películas para niños	6.000 familias de palabras	1.5%
Inglés oral	7.000 familias de palabras	1.3%

Note. Adaptado de “A vocabulary size test”, by Nation, I.S.P. & Beglar, D., 2007. *The Language Teacher*, p. 9.

Según Nation y Beglar (2007), alcanzar la meta de 8.000 familias de palabras es crucial para los estudiantes que quieren tratar con los textos escritos y hablados no simplificados. Otro motivo importante para medir el tamaño léxico de los hablantes no nativos es ser capaz de observar su desarrollo léxico. Medir este tamaño requiere una prueba que sea capaz de evaluar un rango amplio de tamaños de vocabulario tanto longitudinalmente como entre diferentes grupos de estudiantes. El tercer motivo fundamental para medir el tamaño léxico de los que aprenden una L2 es para poder comparar el repertorio léxico de los hablantes no nativos con los hablantes nativos. Esta comparación podría ser útil para comparar las tasas de crecimiento en los contextos del aprendizaje de segundas lenguas y determinar si un incremento adicional en el umbral de 98% de cobertura, como fue sugerido por Nation (2006), no produce efectos notables en la comprensión lectora, la lectura por placer, o la velocidad con que se lee.

En cuanto a la producción oral, Koizumi (2005) en su disertación doctoral encontró relaciones estrechas entre el conocimiento léxico y la producción oral. Más específicamente, en su estudio la investigadora analizó cómo el tamaño y la profundidad del vocabulario productivo están relacionados con la producción oral. Según Koizumi (2005), explorar el impacto del vocabulario productivo sobre la producción oral es fundamental tanto para propósitos teóricos como prácticos. En este estudio se encontraron correlaciones directas entre el conocimiento léxico y la fluidez, exactitud, complejidad sintáctica y léxica en las producciones orales de los estudiantes japoneses de inglés como segunda lengua en el nivel principiante. En este estudio, dos pruebas se usaron principalmente: (1) el *Productive Vocabulary Knowledge Test* que consta de la sección del tamaño y la sección de profundidad que tiene tres subsecciones: (a) derivaciones, (b) antónimos, (c) colocaciones y (2) la prueba de producción oral que incluyó cinco tareas grabadas de descripción, comparación y explicación que fueron todos monólogos con y sin planificación previa. Según los resultados de este estudio, los que tuvieron un conocimiento léxico más amplio y profundo tuvieron una tendencia más moderada o fuerte (dependiendo del tipo de la tarea productiva) de producir una mayor cantidad de palabras (*tokens*) y diferentes tipos (*types*) de ítems léxicos en la producción oral y mejor rendimiento productivo relacionado con el vocabulario. Más específicamente, los elementos relacionados con el conocimiento léxico en la producción oral fueron: (a) fluidez, (b) exactitud en la tarea, (c) la longitud de la unidad de *AS* (*Analysis of Speech Unit*) que es un elemento de complejidad sintáctica y (d) el índice de *Guiraud* que se usó como un elemento de complejidad léxica. Los hallazgos de este estudio también mostraron la importancia tanto del tamaño como de la profundidad

(especialmente los antónimos) del conocimiento léxico en la producción oral entre los estudiantes principiantes, lo cual indicó la necesidad de incluir los aspectos de profundidad en la enseñanza y evaluación.

Staeher (2009) en su estudio investigó el papel del conocimiento léxico en la comprensión auditiva de 115 estudiantes daneses del inglés como lengua extranjera que estaban en el nivel avanzado. Particularmente, los objetivos de ese estudio fueron descubrir hasta qué punto el tamaño y la profundidad del conocimiento léxico contribuían a la comprensión auditiva y cuánto vocabulario era necesario para la comprensión auditiva adecuada. Las dimensiones de tamaño y profundidad del conocimiento léxico fueron evaluadas a través del *Vocabulary Levels Test* (versión 2 desarrollada por Schmitt y Clapham, 2001) y el *Word Associates Test* (Read, 1993; Read, 1998) respectivamente, y las dos pruebas estaban significativamente correlacionadas con la prueba de comprensión auditiva que fue evaluada por una prueba estandarizada del Certificado de Competencia en Inglés de Cambridge. Los resultados del análisis de regresión múltiple mostraron que las dos dimensiones del vocabulario juntas representan 51% de la varianza en los puntajes de comprensión auditiva y también sugirieron que la profundidad del vocabulario no tiene tanto impacto como el tamaño del conocimiento léxico sobre estos puntajes. En general, estos resultados indicaron que el tamaño léxico, produciendo una covarianza de 49%, fue el componente básico más importante del conocimiento léxico en la comprensión auditiva y que la profundidad léxica no jugó un papel separado. En resumen, este estudio llama la atención al enfoque explícito de expandir tamaño léxico de los estudiantes en el salón de L2. Además, proporciona evidencia empírica de que, al igual que para la comprensión de lectura, un 98% de cobertura léxica es necesaria para

manejar los textos orales y para los estudiantes avanzados, un tamaño léxico de un mínimo de 5.000 familias de palabras sería útil para conseguir dicha meta.

Morris y Cobb (2003), en busca de una herramienta fiable, válida y eficiente en costos, examinaron la utilidad de una herramienta llamada *Vocabprofiler* (Cobb, 2002) como predictor del rendimiento académico en un programa de pregrado de enseñanza del inglés como segunda lengua. Con este propósito se recopilaron las muestras de escritura de 122 estudiantes que ingresaban a una universidad canadiense y en estas muestras de 300 palabras se analizaron los perfiles léxicos. Los textos escritos por los participantes de este estudio se examinaron a través del programa de *Vocabprofiler* (Cobb, 2002). En cada texto se calcularon el *type/token ratio* y las palabras en cada texto se dividieron según las siguientes categorías: las primeras 1.000 a 2.000 palabras más frecuentes, la lista de palabras académicas, palabras funcionales y palabras que no caen bajo estas categorías. Después, los puntajes de los perfiles léxicos fueron correlacionados a través de *Pearson-r* con el promedio de dos cursos de gramática que tomaron los estudiantes en el mismo programa. El primer curso fue un curso de introducción a la gramática inglesa y el segundo fue un curso de gramática pedagógica del inglés como segundo idioma. Según los investigadores del estudio, éstos son los cursos que reflejan claramente el rendimiento académico de los participantes en el programa. Los resultados de esta investigación revelaron que los perfiles léxicos de los estudiantes se correlacionaban significativamente con sus notas finales de los dos cursos. La correlación más alta ($r = 0.37$) se encontró entre la categoría de la lista de palabras académicas y el curso de gramática pedagógica. Según Morris y Cobb (2003), los perfiles léxicos, si se usaran junto con otras pruebas tradicionales de entrada como exámenes de gramática, calificaciones académicas, etc.,

podrían contribuir como predictores potenciales del rendimiento académico y reflejar las habilidades lingüísticas de los hablantes no nativos avanzados que están en estos tipos de programas.

En su estudio Laufer y Goldstein (2004) propusieron una herramienta, la denominada *Computer Adaptive Test of Size and Strength* (CATSS), que combina el tamaño del vocabulario con otra dimensión, la “fuerza” del conocimiento léxico con el fin de evaluar el conocimiento léxico de los estudiantes bilingües. Esta prueba de vocabulario descontextualizado tiene como objetivo medir el tamaño del vocabulario y la “fuerza”, entendida por los autores como los diferentes grados en los que se manifiesta el conocimiento del significado de las palabras (véase Tabla 4).

Tabla 4. Grados y fuerza del conocimiento léxico

	Recuperación	Reconocimiento
Activo (proveer la forma)	Proveer la palabra en L2, el más fuerte (1)	Seleccionar la palabra en L2,(3)
Pasivo (proveer el significado)	Proveer la palabra en L1, (2)	Seleccionar la palabra en L1 (4) el más débil

Note. Adaptado de “Testing Vocabulary Knowledge: Size, Strength, and Computer Adaptiveness.” by B. Laufer and Z. Goldstein, 2004, *Language Learning*. p.407-408.

En este estudio participaron 435 estudiantes de inglés como segunda lengua: 278 hablantes nativos de hebreo, 140 hablantes nativos de árabe y 140 hablantes nativos de ruso y se analizaron según cuatro modalidades propuestas por los autores con el fin de observar si hay una conexión entre éstas y su rendimiento lingüístico general en clase. Los hablantes de hebreo y ruso tomaron la versión inglesa-hebrea, los hablantes de árabe tomaron la versión inglesa-árabe de la prueba. Los ítems de la prueba se eligieron de los cinco niveles de frecuencia establecidos por Nation (1983) que constan de las primeras 2.000, 3.000, 5.000 y 10.000 palabras, más la *Academic Vocabulary List* (AWL). Los

participantes fueron estudiantes de inglés en diferentes niveles, se les presentaron 30 ítems del nivel de frecuencia que les correspondía según su nivel lingüístico. Cada participante fue evaluado en cuatro sub-pruebas diferentes. Según los resultados de este estudio, la jerarquía propuesta por los autores estuvo presente en todos los niveles de frecuencia de vocabulario: la recuperación activa dio los puntajes más bajos, y le siguieron en orden ascendente, la recuperación pasiva, el reconocimiento activo y el reconocimiento pasivo. También se encontraron diferencias altamente significativas ($p < 0.001$) entre las cuatro modalidades en cada nivel de frecuencia. En otras palabras, el crecimiento léxico fue diferente para las 4 modalidades del conocimiento léxico. El conocimiento léxico más avanzado, que se adquiere último, es la recuperación activa. Otro hallazgo importante de este estudio fueron las correlaciones altas entre cada modalidad y las calificaciones de clase de los participantes que reflejaban el rendimiento lingüístico general de cada participante. Los resultados mostraron que de las cuatro modalidades, la recuperación pasiva, o sea, la habilidad de demostrar el entendimiento del significado de una palabra en que los participantes tuvieron que proporcionar la traducción en L1 de una palabra dada en L2, fue el mejor indicador lingüístico del rendimiento escolar de los participantes.

Otro estudio que encontró correlaciones directas entre la competencia léxica y el conocimiento lingüístico general de los estudiantes principiantes e intermedios tanto de español como L2 como de herencia es el de Fairclough (2011). Los resultados de esta investigación mostraron que la prueba de reconocimiento léxico de tipo lista *sí/no* (Meara, 1996) se puede usar como una herramienta fiable y práctica en los exámenes de ubicación, dado que sus resultados correlacionen con otras medidas de competencia

lingüística. La prueba de reconocimiento léxico de este estudio fue diseñada en base al diccionario de 5.000 palabras más frecuentes en español de Mark Davies que se publicó en el 2006. En total 330 estudiantes universitarios: 183 estudiantes de herencia y 147 estudiantes de L2 matriculados en 4 niveles diferentes de estos dos programas tomaron esta prueba. La prueba de reconocimiento léxico contenía 120 palabras en español y 80 pseudo-palabras. La fiabilidad de esta herramienta se verificó a través del valor de alfa para 120 palabras (.972) y el método de división por mitades de *Spearman-Brown* (.981) y *Guttman* (.979). Además de la prueba de reconocimiento léxico, mientras la mitad de los participantes completaron un *cloze test* en 2007, la otra mitad completó una prueba de tareas múltiples en 2008 que tenía por objetivo evaluar la competencia gramatical. La última prueba consistía en las siguientes 4 partes: (1) traducción parcial, (2) dictado, (3) rellenar los huecos con la forma correcta del verbo y (4) una tarea de selección múltiple. La fiabilidad de estos instrumentos se confirmó por el *alfa de Cronbach* (.819 para el *cloze test*; .806 para la última prueba de tareas múltiples).

Según los resultados de este estudio, *el coeficiente de Pearson* (Pearson (330) $r \approx .796$, $p < .01$) indicó correlaciones fuertes entre la prueba de reconocimiento léxico y las otras dos pruebas (el *cloze test* y la prueba de tareas múltiples). Un *t-test* ($t(328) = 17.005$, $p < .001$) confirmó las diferencias significativas entre el grupo de L2 y de herencia. Mientras que diferencias significativas se encontraron entre los diferentes niveles del grupo de L2, no se observaron diferencias significativas entre los diferentes niveles del grupo de herencia, lo cual indica que un rango más amplio de palabras hubiera sido necesario para evitar el efecto del límite superior y hacer una mejor distinción entre los estudiantes de herencia.

Los hallazgos más importantes de este estudio indican que la prueba de reconocimiento léxico de tipo lista *sí/no* (Meara, 1996) parece ser una herramienta práctica y fiable que podría funcionar bien para evaluar el conocimiento léxico pasivo de los estudiantes de L2, especialmente de los que están en los niveles principiantes e intermedios. Aquí cabe mencionar el problema de que el conocimiento léxico de los estudiantes de L2 generalmente se basa en los libros de texto, es decir, estos estudiantes, conocen principalmente las palabras que aparecen en los libros de texto y no las palabras de frecuencia (Davies & Face, 2006; Fairclough, 2011).

Algunos estudios previos (Davies & Face, 2006; Godev, 2009) también indicaron que los libros de texto de L2 en español sólo parcialmente reflejan las palabras frecuentes usadas por los hablantes nativos en la vida real. Por lo tanto, los estudios futuros deberían tomar esta realidad en cuenta a la hora de evaluar el conocimiento léxico de estos estudiantes. Tal vez, investigaciones más profundas sean necesarias, que indiquen esta negligencia en aprendizaje/enseñanza del vocabulario en español como L2 y sobre todo, estudios empíricos que puedan arrojar luz sobre cómo evaluar el conocimiento léxico de estos estudiantes con herramientas más efectivas. En este sentido, las pruebas que se basan en las listas de frecuencia y que tienen por objetivo evaluar el conocimiento léxico de los estudiantes de L2 pueden conducir a resultados equivocados.

En síntesis, el tamaño del vocabulario se ha considerado un buen indicador del nivel de competencia lingüística de un individuo (Laufer et. al., 2004; Nation, 2001; Read, 1997). Tanto Meara (1996), como Laufer (1998) consideran la extensión léxica, más relevante que la profundidad, ya que ofrece una visión general más representativa del conocimiento léxico (Read, 2000), y es probablemente la única dimensión que tiene

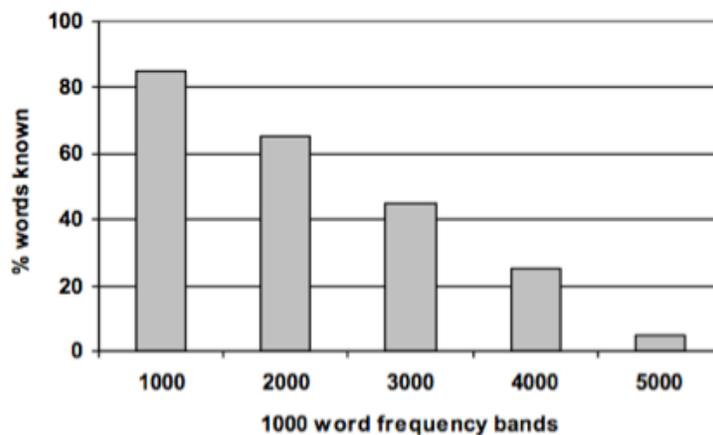
importancia en vocabularios pequeños. En este sentido, si la meta es ampliar el vocabulario de los estudiantes de segundas lenguas, sería primordial primero determinar cuál es su nivel de competencia léxica y tratar de detectar aquellos aspectos de la misma donde tienen dificultades, para después mejorar las estrategias de enseñanza y los materiales que se usan con el fin de desarrollar sus conocimientos léxicos.

2.2. Los estudios previos del tamaño y la profundidad del vocabulario

La gran parte de los estudios que analizan el tamaño y la profundidad del vocabulario tanto de hablantes nativos como de estudiantes de segundas lenguas se basan en las listas de frecuencia. El punto de partida de los estudios de L2 que siguen la perspectiva de frecuencia es que el vocabulario frecuente, tal como aparece en la lengua meta cuando los hablantes nativos la usan, debe ser el criterio principal de consideración en aprendizaje/enseñanza de vocabulario receptivo. Mientras que existen numerosas investigaciones sobre vocabulario frecuente en inglés (Nation, 1990; Nation, 2001), los estudios que siguen esta perspectiva de investigación en español como L2 son bastante limitados.

En esta línea, Meara (1992) trató de construir un modelo práctico de frecuencia a través del delineamiento de un perfil del vocabulario de los estudiantes de inglés como segunda lengua. Según Meara (2010), dividir un lexicón amplio en segmentos iguales de 1.000 palabras sería el modo más conveniente (véase la Figura 1).

Figura 1. El modelo de frecuencia del aprendizaje de vocabulario de Meara (1992)



Note. Adaptado de “A vocabulary size test” by I.S.P. Nation and D. Beglar, 2007, *The Language Teacher*, p.9.

De acuerdo con el investigador, la primera columna en la Figura 1 representa el conocimiento de las 1.000 palabras más frecuentes; la columna 2, las siguientes 1.000 palabras más frecuentes y así sucesivamente. Meara sugiere que el conocimiento de un estudiante típico es alto en las columnas más frecuentes y más bajo en las columnas menos frecuentes mostrando una inclinación distintiva hacia abajo. Mientras el conocimiento del estudiante aumenta, el perfil se mueve hacia arriba hasta llegar al tope, es decir, hasta que el estudiante conozca todas las palabras. La evidencia empírica muestra que eso es lo que ocurre con los estudiantes de L2 y con la mayoría de los individuos.

En otro estudio Milton y Alexiou (2010) analizaron el tamaño léxico de los estudiantes de griego como L2 a través de una prueba basada en un corpus de las 5.000 palabras más frecuentes sacadas del *Hellenic National Corpus*. Esta prueba fue diseñada como la versión griega de la prueba de *X-Lex* de Meara y Milton (2003) que tiene diferentes versiones en otros idiomas. En la prueba los estudiantes tienen que identificar

las palabras que conocen eligiendo *sí* o *no*. La prueba consiste en un total de 120 palabras: 20 palabras elegidas al azar de cada uno de los primeros cinco mil segmentos de palabras de frecuencia y 20 pseudo-palabras, que son palabras inventadas. La prueba se aplicó en total a 64 estudiantes universitarios de 4 niveles diferentes que estudiaban griego como L2 en la Universidad de Thessaloniki. Como se esperaría, los promedios de porcentajes (de respuestas esperadas) de cada grupo aumentaron a medida que subía el nivel, lo cual indica relaciones precisas entre el nivel del Marco Común Europeo de Referencia (MCER) y el tamaño del léxico de los estudiantes. También se encontraron diferencias significativas entre los 4 niveles diferentes que tomaron esta prueba.

En resumen, los resultados de esta herramienta sugieren que esta prueba tiene un buen nivel de validez y proporciona información fiable sobre el tamaño del léxico de los estudiantes de L2. Sin embargo, en el diseño de este tipo de prueba también hay que considerar el nivel de estimación que los estudiantes hacen para llegar a la respuesta correcta si desconocen la palabra. Para evitar este tipo de dificultad Meara y Milton (2003) sugieren que se calcule el conocimiento del vocabulario del estudiante contando sus respuestas correctas y multiplicándolas por 50 para obtener una nota pura basada en 5.000. Después se multiplican las palabras inventadas por 250, y este total se deduce de la nota pura para obtener la nota ajustada, también sobre 5.000. Según Meara y Milton (2003), de esta manera se puede compensar las estimaciones que hace el estudiante si no conoce la palabra examinada.

De acuerdo con Milton (2009), la frecuencia de palabras y el aprendizaje del vocabulario están estrechamente relacionados con la frecuencia que determina con qué palabras el estudiante se encuentra y con la frecuencia en que las encuentra. La idea de

Milton se basa en la noción universal de que las palabras más frecuentes en la lengua meta tienden a aprenderse más temprano que las palabras menos frecuentes. Por lo tanto, es primordial entender el proceso de aprendizaje del vocabulario; si se entiende qué palabras se aprenden primero y qué palabras se aprenden más tarde, es posible construir herramientas que evalúen mejor el conocimiento léxico.

Por otra parte, como menciona Meara (2010), cuando estas palabras se derivan de listas de frecuencia no representan de manera fiable la dificultad real que los estudiantes de L2 tienen con dichas palabras, y el riesgo de distorsión es sumamente alto.

Claramente, lo que se necesita es una mejor lista que podría servir como recurso, pues las listas de frecuencia generalmente no se basan en el tipo de léxico que los estudiantes de L2 están expuestos en el salón de la clase. Por lo tanto, no son recursos ideales para evaluar su conocimiento léxico. En este sentido, un recurso apropiado que refleje bien el conocimiento léxico actual de estos estudiantes y que esté, además, apoyado con una buena cantidad de estudios empíricos es tremendamente necesario en este campo.

El *Word Associates Format* (WAF), diseñado la primera vez por Read (1998), es otra de las pruebas más comunes que se usa para evaluar la profundidad del conocimiento léxico; principalmente para medir el conocimiento asociativo, el uso de colaciones y conceptos. En esta prueba la palabra meta del ítem se presenta con ocho opciones, cuatro de las cuales tienen relación con la palabra meta y las otras cuatro no. El estudiante tiene que elegir las relacionadas con la palabra meta. Schmitt, Ng y Garras (2011) en su estudio analizan la validez de esta herramienta, su estudio consta de dos partes: en el primer estudio participaron 18 estudiantes de origen japonés que estudian inglés como L2. En la primera etapa se administró la prueba WAF de 40 ítems. En una segunda etapa se

realizaron entrevistas de introspección sobre las actitudes de los participantes durante la selección de las respuestas. Esta tarea tenía el fin de comparar sus actitudes con las respuestas dadas y de esta manera poder determinar si el método utilizado reflejaba la realidad de su conocimiento léxico o no.

En la segunda parte del estudio participaron 28 estudiantes de origen chino que estudian inglés como L2 en la Universidad de Nottingham. El mismo procedimiento se repite con algunas modificaciones, especialmente en las entrevistas de introspección, pues se buscaba comparar diferentes métodos de evaluación, tipos de formatos y distractores con la distribución de las respuestas dadas. Ambos estudios indican que el WAF refleja la profundidad del conocimiento léxico de los participantes en los dos extremos de la escala de evaluación, mientras que los puntajes que están en la mitad de la escala no conducen a una interpretación fiable. Además, se encuentra una mayor posibilidad de que WAF subestime o sobreestime el conocimiento léxico.

En este sentido, parece que si el propósito de un estudio empírico es simplemente medir qué tan bien conocen los participantes las palabras examinadas, WAF no daría resultados tan obvios. Sin embargo, de acuerdo con Bogaards (citado en Schmitt, Ng & Garras, 2011) si el propósito es descubrir el conocimiento general cualitativo del lexicon, podría ser interesante observar las diferencias entre los estudiantes que pueden identificar dos o tres elementos asociados sin conocer las palabras y los que no pueden identificar ninguno. No obstante, el mayor inconveniente de las pruebas de WAF es que resultan poco prácticas, ya que la mayoría se llevan a cabo mediante entrevistas personales tal como ocurre en el estudio de Schmitt, Ng y Garras (2011), lo cual complica el número de los participantes o la cantidad de las palabras que se analizan.

2.3. Estudios previos de frecuencia léxica y su utilidad en L2

En cuanto a las investigaciones sobre la frecuencia de palabras, Mark Davies además de su corpus del español, publicó en 2006 *A frequency Dictionary of Spanish, core vocabulary for learners* (FDS) que puede ser una fuente útil para la enseñanza y el aprendizaje del vocabulario en EL2. Este diccionario de frecuencia consta de las 5.000 palabras más frecuentes basadas en un corpus amplio de textos variados. Al lado de cada entidad se indica la parte de la oración (sustantivo, verbo, adjetivo, etc.), una definición simple en inglés, y un ejemplo actual de la palabra en contexto, extraído del *Corpus del Español de 100 millones palabras* (www.corpusdelespanol.org). Las entidades también señalan si la palabra es más común en la oralidad, ficción o no-ficción para que el usuario aprenda con más exactitud cuándo y en qué contexto se usa dicha palabra.

Las palabras de este diccionario de frecuencia fueron extraídas de un Corpus de 20 millones de palabras elaborado por el mismo Davies en la Universidad de Brigham Young. La mayoría de los ítems de este corpus fueron sacados del contenido del siglo XX del *Corpus del Español de 100 millones palabras*, que a su vez consta de textos del 1200 al 1900, además de los años 2001-2002.

En términos de representatividad, Davies en su prólogo subraya que el corpus contiene la colección más amplia de registros y tipos de textos que cualquier diccionario previo de la lengua española. Como se puede observar de la Tabla 5, dos tercios del corpus vienen de registros escritos, mientras que un tercio (6.750.000 palabras) viene del español oral. Aproximadamente la mitad del corpus oral viene de transcripciones del habla espontánea, incluyendo 2.300.000 palabras del corpus del *Habla culta* de las conversaciones con hablantes de once países diferentes y el *Corpus Oral de Referencia*

de 1.000.000 de palabras que contiene transcripciones de conversaciones, lecturas, sermones, programas deportivos y muchos otros tipos del español oral. El corpus escrito se divide equitativamente entre textos literarios y no-literarios, incluyendo artículos de periódicos, ensayos, enciclopedias, cartas y textos humanistas (Davies, 2006, p.2).

Tabla 5. Composición del corpus español moderno de 20 millones de palabras				
	Núm. de las palabras	España	Núm. de las palabras	Latinoamérica
Oral	1, 00	España oral	2, 00	Habla culta (diez países)
	0, 35	Habla culta (Madrid, Sevilla)		
3, 35	1, 35		2, 00	
Transcripciones / Obras de teatro	1, 00	Transcripciones/ entrevistas (congresos, conferencias de la prensa, otros)	1, 00	Transcripciones/ entrevistas (congresos, conferencias de la prensa, otros)
	0, 27	Entrevistas en el periódico ABC		
	0, 40	Obras de teatro	0, 73	Obras de teatro
3, 40	1, 67		1, 73	
Literatura	0, 06	Novelas(BV2)	1, 60	Novelas(BV2)
	0, 00	Cuentos cortos (BV2)	0, 87	Cuentos cortos (BV2)
	0, 19	Tres novelas (BY3)	1, 11	Doce novelas (BY3)
	2. 17	En su mayoría novelas de LEXESP4	0, 18	Cuatro novelas de Argentina 5
			0, 20	Tres novelas de Chile 6
6,38	2, 42		3, 96	
Textos	1, 05	Periódico ABC	3, 00	Periódicos de seis países diferentes
	0, 15	Ensayos en LEXESP4	0, 07	Cartas de Argentina
	2, 00	Enciclopedia Encarta	0, 30	Textos humanistas(e. g., filosofía, historia de Argentina5)
			0, 30	Textos humanistas(e. g., filosofía, historia de Argentina6)
6,87	3, 20		3, 67	
Total	8, 64		11, 36	

Nota. Adaptado de *Frequency Dictionary of Spanish* by M. Davies, 2006, p.3.

Según Davies, además de tener una buena selección de diferentes géneros, este corpus es el primero que tiene un buen balance de textos de Latinoamérica y España;

aproximadamente 43% de los textos viene de España, mientras que el 57 % es de Latinoamérica. En cuanto al período de tiempo representado a través de este corpus, casi todos los textos son de 1970-2000, la mayoría de los cuales son de la década de 1990.

Sobre todo, este diccionario de frecuencia puede servir tanto para determinar el orden de la enseñanza de palabras en cada nivel, como para establecer criterios específicos necesarios para el diseño de los materiales de EL2. En este sentido, *FDS* es el más actual de los libros de referencia en esta línea y constituye un recurso valioso, tanto para los diseñadores de programa de estudios como editoriales de texto para determinar el vocabulario útil para enseñar a los estudiantes del EL2.

Según Collentine (2008), para entender la importancia de *FDS*, es importante entender sus orígenes. Davies, junto con Doug Biber, reunió un corpus de 20.000.000 palabras del español escrito y oral. Cada entidad se codificó según su parte del habla y su propiedad léxica. De esta manera, se produjo el primer análisis multidimensional del español, donde se detallan los tipos de las propiedades léxico-gramaticales y se caracterizan varios registros de la lengua.

Este diccionario no sólo contiene diferentes registros formales, sino también variados registros del habla, reflejando el habla espontánea y conversacional. De acuerdo con Collentine (2008), desde la perspectiva pedagógica, *FDS* puede servir en la creación de libros de texto que se usan en la enseñanza de segundas lenguas. Los autores de los libros de texto deberían consultar el *FDS* para asegurarse la relevancia del léxico que debería enseñarse y aprender en el salón de EL2.

Un estudio similar al estudio de Davies y Face (2006) es el estudio de Godev (2009). A diferencia del estudio de Davies y Face (2006), Godev (2009) en su estudio

analiza 5 libros de texto publicados en los Estados Unidos sobre el primer año de español universitario. Sin embargo, Godev no sólo incluye las listas de palabras, sino también el vocabulario que se presenta en los textos de lectura. Su análisis consiste en dos partes: 1) una comparación de las listas de vocabulario y del vocabulario de las lecturas, con la frecuencia que se da en *A frequency dictionary of Spanish: Core vocabulary for learners* de Davies (2006); y 2) una muestra de los aspectos logísticos que deben considerarse, tales como cantidad de tiempo en contacto con la materia, el conocimiento léxico, y de cómo se perfilan estos aspectos a través de las dimensiones receptivas y productivas del proceso de adquisición del vocabulario. El planteamiento de este estudio se basa principalmente en los hallazgos encontrados por Davies (2006): en español las 1.000 palabras más frecuentes proporcionan una cobertura de texto del 76% de los textos no ficcionales, el 79.6 % de los textos ficcionales y el 87.8% de los textos orales. Mientras que las siguientes 1.000 palabras cubren el 8% de las ítems léxicos en un texto no ficcional, el 6.5% de en texto ficcionales y el 5% en textos orales. De cada 5 libros de texto (*Puntos de partida, Tú dirás, Dicho y hecho, Arriba, Nexos*) se elige al azar una muestra de texto que se consta de 69-171 palabras y una lista de palabras que proporciona al final del capítulo. Aunque los porcentajes son diferentes, se encuentran resultados similares al estudio de Davies y Face (2006); la mayoría de los ítems léxicos presentados en los libros de texto de primer año universitario no coincide con el rango de las primeras 1.000 palabras más frecuentes en español que deberían, en realidad, ser el enfoque de este nivel. Por otra parte, el 24% de las palabras que aparecen en las muestras analizadas en este estudio están por encima del rango de las 5.000 palabras más frecuentes. En cuanto a los textos de lectura, la cobertura de los textos, es decir, el porcentaje de las palabras

conocidas por los estudiantes, oscila entre 73% y 89%. Sin embargo, según Nation (2001), los estudiantes de L2 necesitan un 95% de cobertura para aprender vocabulario, es decir, la información enfocada en significado y 98-100% de cobertura para el desarrollo de fluidez. Por otra parte, Godev (2009) subraya que leer textos simples puede servir para profundizar el conocimiento del vocabulario pero no aumentaría el tamaño del conocimiento léxico. En este sentido, según Godev (2009), los libros de texto pueden ayudar a la cantidad de vocabulario y con la selección de dicho vocabulario, para así conseguir resultados más realistas en cuanto al conocimiento léxico de los estudiantes del primer año de español como L2 en Estados Unidos.

Uno de los recursos más extensos y recientes sobre el análisis y revisión crítica de los materiales de evaluación de la competencia léxica es la tesis de maestría de Casso (2010). En su estudio Casso analiza los materiales que evalúan tanto el tamaño como la profundidad del vocabulario de los estudiantes de segundas lenguas. A través de este trabajo Casso no sólo analiza los distintos tipos de pruebas de léxico, sino también los corpus y las listas de frecuencia en los que se basan. Tras la revisión crítica de los materiales existentes en este campo, Casso presenta dos pruebas basadas en el Diccionario de Frecuencias (Almela, 2005) el cual se basa en los datos del corpus Cumbre, con el fin de medir el vocabulario receptivo y productivo de los alumnos estadounidenses del nivel B1⁴. Los resultados de estas dos pruebas muestran que las pruebas basadas en las listas de palabras más frecuentes presenta una serie de ventajas, pero también muchas desventajas. Entre las ventajas, el autor resalta que estas palabras son las primeras que hay que enseñar a los estudiantes de una segunda lengua, ya que su

⁴ El nivel B1 es el tercer curso del *Diplomas de Español como Lengua Extranjera* (DELE) que corresponde al nivel intermedio según el *Marco Común Europeo de Referencia para las Lenguas: aprendizaje, enseñanza, evaluación* (MCERL).

conocimiento les permitirá desenvolverse en las situaciones más básicas de la vida cotidiana. En cuanto a las desventajas, no todas estas palabras reflejan las necesidades comunicativas inmediatas de los estudiantes, como puede ser hablar de su trabajo o de sus aficiones. Además, puede haber estudiantes que aunque desconozcan ciertas palabras frecuentes, estas se ajusten más a sus necesidades.

En la misma línea otro estudio exhaustivo en este campo es la tesis doctoral de López-Mezquita (2007). La investigadora revisa en detalle la enseñanza y el aprendizaje del vocabulario, los distintos aspectos de la competencia léxica y las aportaciones de los corpus a la enseñanza de las lenguas. A partir de estos datos investiga los listados de frecuencias para obtener los contenidos de las pruebas de vocabulario. Así, López-Mezquita revisa las cuestiones importantes relacionadas con la elaboración de las pruebas: la fiabilidad, la validez, la retroactividad, la ética y los niveles en la evaluación, por otra parte, las cuestiones relacionadas con la praxis: el desarrollo y la elaboración de una prueba, tipos de pruebas y métodos de pruebas relacionadas con la evaluación del vocabulario en segundas lenguas. En su tesis, la autora subraya la importancia de las siguientes fases, las cuales se deberían tener en consideración al tratar de diseñar una “buena” prueba: (1) planificación, diseño y construcción; (2) administración y corrección; (3) análisis de resultados y conclusiones.

2.4. Los estudios previos de las estrategias del vocabulario

En general, los hallazgos de las investigaciones sobre las estrategias para el aprendizaje de vocabulario (EAVs) concuerdan con los estudios de las estrategias para el aprendizaje general de una segunda lengua. Según Nation (2013), EAVs son partes de las estrategias para el aprendizaje de una lengua, que, a su vez, son parte de las

estrategias de aprendizaje general. Aunque aún no existe un consenso en la literatura en cuanto a la clasificación de EAVs, como se entiende de la siguiente explicación de Brown y Payne (1994):

It may be said that in the classification of vocabulary learning strategies, consideration has to be given not only to the process of learning vocabulary but also to the L1 and the L2, to the characteristics of learners, and to the focus of the research concerned. In fact, no classification is perfect, and any individual strategy may fall into one category or another, depending on the aspect in focus. What is beyond dispute is that strategies may be divided between those that are “more directly related to individual learning tasks and entail direct manipulation or transformation of learning materials,” that is, the cognitive strategies, and those that are connected with “the learning process, planning for learning, monitoring of comprehension or production while it is taking place, and self-regulation after the learning activities have been completed,” that is, the metacognitive strategies (citado en Hatch y Brown, 1995, p.373).

Gu y Johnson (1996), en su estudio sobre las EAVs , aplican un cuestionario retrospectivo (*self-report questionnaire*) a 850 alumnos chinos avanzados de inglés como lengua extranjera. Los resultados obtenidos de este cuestionario se correlacionan con las dos *Pruebas de Niveles de Vocabulario* y una prueba de rendimiento (*CETBAND2*) que se usa como una medida de competencia lingüística.

Según los resultados del análisis de regresión múltiple, dos estrategias meta-cognitivas (i.e. *self-initiation* y *self attention*) son buenos indicadores del éxito de la prueba de rendimiento. Las estrategias de repetición ligadas a imágenes muestran ser

menos eficaces que la inferencia contextual, el uso del diccionario, la toma de notas o la reflexión meta-cognitiva. Este estudio como el estudio de Schmitt y McCarthy (1997) demuestra que no sólo el tipo de estrategia utilizada, sino su variedad y la posibilidad de combinarlas, es lo que determina resultados “buenos” o “malos”. Las diferencias individuales parecen estar relacionadas con factores de variedad y combinación de estrategias. Finalmente, un análisis de agrupación identifica 5 acercamientos de aprendizaje por parte de los estudiantes: 1) *lectores*: son los estudiantes “mejores”, a los que les encanta leer, no usan memorización, sino el contexto para aprender palabras; 2) *usuarios de estrategias activas*: es el segundo grupo con los mejores resultados en la prueba de vocabulario y rendimiento, usan más estrategias (p. ej., exposición natural, memorización, uso de diccionario, etc.) que los otros grupos; 3) *no-codificadores* y 4) *codificadores*: los no-codificadores y codificadores forman el 87% de los participantes y son muy similares, la única diferencia entre ellos es que los codificadores usan más estrategias de memorización y por último 5) *usuarios de estrategias pasivas*: es el grupo que tiene menos éxito en las pruebas de este estudio y que cree fuertemente en la memorización.

De acuerdo con Nation (2013), los resultados del estudio de Gu y Johnson (1996) deben interpretarse con cautela, ya que estos resultados se basan en los datos de un cuestionario retrospectivo y lo que los participantes dicen hacer no siempre representa lo que realmente hacen.

En la misma línea, Fan (2003) usa la *Prueba de Niveles de Vocabulario* y un cuestionario de EAVs con el fin de explorar las relaciones entre las estrategias usadas por parte de 1,067 alumnos chinos de inglés como segunda lengua y el nivel de vocabulario.

A los participantes se les pide que evalúen una serie de EAVs según su uso de frecuencia y utilidad para ellos. Los resultados señalan que las estrategias más usadas incluyen: adivinar, hacer uso de palabras conocidas y estrategias de diccionario. Las palabras claves y otras técnicas mnemónicas están entre las estrategias menos usadas y conocidas. La frecuencia de uso y utilidad no están fuertemente relacionadas, es probable que las estrategias que son relativamente fáciles de aplicar como adivinar, no necesariamente siempre dan buenos resultados. Las estrategias que requieren más tiempo como el uso de diccionarios pueden dar mejores resultados. Los estudiantes con alta puntuación en las pruebas de vocabulario usan más frecuentemente las estrategias de planificación y oportunidades de encontrar nuevas palabras tanto dentro como fuera del salón de clase. Como los estudios anteriormente mencionados, los estudiantes con mejores resultados usan más estrategias y las usan más frecuentemente que los estudiantes menos competentes en el aprendizaje vocabulario.

En otro estudio, Schmitt (1997) estudia una muestra representativa de 600 estudiantes japoneses de inglés como segunda lengua de diferentes niveles. Se obtienen hallazgos significativos: en primer lugar, se comparan dos juegos de datos, las estrategias más usadas y las que más ayudan en el aprendizaje. Estas revelan que hay algunas intercepciones. En segundo lugar, se descubre que los patrones de las estrategias empleadas cambian con el tiempo a medida que el alumno madura o se vuelve más eficiente en el segundo idioma. Muchas de las estrategias se vuelven más relevantes con la edad debido a que la imaginación, la asociación y el análisis comienzan a ser parte de un análisis más profundo.

Mientras existen una gran variedad de clasificaciones de las EAVs, Schmitt las agrupa en dos categorías principales: 1) estrategias de descubrimiento (*discovery strategies*): estrategias usadas para descubrir el significado de la palabra y 2) estrategias de consolidación (*consolidation strategies*): estrategias usadas para consolidar una palabra una vez que ya ha sido encontrado y propone 5 sub-categorías: 1) estrategias de determinación: que son estrategias individuales de aprendizaje; 2) estrategias sociales: que se usan a través de la interacción; 3) estrategias de memorización: son procesos mentales a través de los cuales los estudiantes asocian su conocimiento con las palabras nuevas; 4) estrategias cognitivas: que no requieren procesos mentales y son más mecánicas; 5) estrategias meta-cognitivas: son estrategias relacionadas con los procesos de supervisión y auto-evaluación.

Al revisar diferentes tendencias explicativas e investigativas con respecto al conocimiento léxico, en el siguiente capítulo se elaborará el marco teórico de referencia para este estudio. Asimismo, se definirán algunos conceptos relacionados con la construcción de la misma teoría.

Capítulo 3

Marco teórico

Wide range of aspects involved in knowing a word: the systems, which lie behind vocabulary (the affixation system, the sound system, the spelling system, collocation, the grammatical system, lexical sets) and the unique behavior of each word. (Nation, 2013, p. 85)

En este capítulo se establece el marco teórico a través de algunos conceptos claves que sirven de base a este estudio. Se define en detalle qué se entiende por conocer una palabra desde aspecto productivo y receptivo junto con la terminología básica que sirve como referente para la elaboración de los instrumentos. En primer lugar, se presenta (3.1.) la competencia léxica desde el punto de vista psicolingüística. Enseguida, se explica (3.2) qué significa saber una palabra según el modelo de Nation (2013) a través de sus siguientes aspectos multifacéticos: (3.2.1) forma escrita, (3.2.2) partes de una palabra, (3.2.3) conexión entre forma y significado, (3.2.4.) concepto y referentes, (3.2.5) asociaciones, (3.2.6) funciones gramaticales, (3.2.7) colocaciones, (3.2.8) restricciones de uso. En el tercer apartado de este capítulo se define alguna terminología básica como (3.3) la concepción del léxico de las palabras a través de componentes léxicos, tipos, lemas, familia de palabras. A continuación, en el cuarto apartado se presentan en detalle (3.4) las variables lingüísticas del estudio: el conocimiento receptivo y productivo de las palabras.

3.1. La competencia léxica desde el punto de vista psicolingüístico

Para muchos psicolingüistas, entre ellos Miller (1986; 1991) y Aitchinson (1987), el hecho de que un hablante pueda acceder en milésimas de segundo a una cantidad

ingente de vocabulario almacenado en su memoria, tanto en procesos de producción como de comprensión, es una prueba fehaciente de que el lexicón mental está organizado y estructurado de modo que posibilita el acceso inmediato. El *léxico mental, diccionario mental (the human Word-store)* (Aitchison, 1994) es el conjunto de palabras que un individuo conoce. Dicho conjunto se estructura de forma muy precisa, siguiendo un complejo y elaborado sistema de interconexiones que permiten almacenar un gran número de palabras y recuperarlas cuando las necesitamos. Se trata de un sistema activo en continuo cambio.

Según Aitchison (1994), el proceso mediante el cual se forma este léxico mental, es decir los mecanismos a través de los que los niños adquieren el vocabulario, se basa fundamentalmente en tres tipos de tareas: *etiquetado (labelling)*, que tiene lugar durante los dos primeros años de vida; *empaquetado (packing)*, que consiste en la clasificación de las entidades; y finalmente, *construcción de redes (network-building)* que tiene lugar gradualmente y continúa a lo largo de toda la vida de la persona. En este sentido, se puede concluir que el léxico mental es algo que va variando a lo largo de los años y que se encuentra en permanente construcción.

En cuanto a la L2, los estudiosos de la adquisición del léxico de las lenguas segundas o extranjeras (Laufer, 1986; Meara, 1984) consideran que en los estudios de la interlengua, es decir, de la adquisición del lenguaje no nativo, el léxico ha recibido un tratamiento superficial en la medida en que no se ha tratado en los trabajos claves sobre la interlengua o se ha tratado con poco rigor (Liceras & Carter, 2009).

La cuestión de si la L2 se parece a la estructura de la lengua materna fue estudiada por Wolter (2001), quien en su estudio compara los patrones de respuesta de los

hablantes nativos y no nativos a través de pruebas de profundidad del conocimiento léxico, con objeto de entender si el conocimiento en profundidad de las palabras determina un grado específico de integración del léxico mental. Al contrario de los estudios anteriores, Wolter (2001) demuestra la posibilidad de similitud entre el léxico mental de la lengua materna y el de una segunda lengua cuando se ha adquirido esta última a un cierto nivel de profundidad. Para Pérez Basanta (2003) existen tres procesos en la adquisición del vocabulario:

Entrada (input): las palabras se guardan en la memoria a corto plazo, dependiendo de la profundidad de la palabra.

Almacenamiento: toda la información se almacena en campos semánticos con referencias cruzadas, los mismos que se organizan en redes asociativas. Existen varias teorías acerca de cómo se almacenan las palabras que pertenecen a la lengua materna y a la segunda lengua entre ellas las más importantes:

- a) hipótesis del sistema extendido: las palabras se almacenan en un solo lugar
- b) hipótesis del sistema dual: las palabras se almacenan en dos lugares separados
- c) hipótesis tripartita: las palabras similares, como los cognados, se almacenan en el mismo lugar y las palabras específicas en lugares separados
- d) hipótesis de los subconjuntos: las palabras se almacenan en un solo lugar, pero organizadas en subconjuntos, dentro de los que existe una conexión más íntima entre las que pertenecen a la lengua materna.

Recuperación: las palabras se buscan y se recuperan de una manera rápida y eficiente. Por ejemplo, de una manera receptiva se percibe la forma y se recupera su significado, pero la parte productiva recupera la forma escrita y hablada de una

palabra. La facilidad con la que se recupera la palabra, se verá favorecida si la palabra es frecuente en la lengua y si la persona la ha empleado recientemente. Existen dos aproximaciones en cuanto a la forma de acceder al léxico: el acceso directo y el acceso indirecto, que supone dos etapas, una búsqueda seguida de un posterior acceso (pp.35-47).

3.2. ¿Qué significa saber una palabra?

Para Nation (2013), el conocimiento léxico es un proceso complejo y gradual en el que se aprende no sólo la forma y el significado, sino también una intrincada red de relaciones formales y semánticas entre ese *ítem* y otras palabras y morfemas que constituyen subsistemas de diferentes niveles. El conocimiento de una palabra es una representación mental de gran complejidad, que integra diferentes aspectos y componentes cognitivos, algunos más automáticos e inconscientes y otros más conscientes, reflexivos y experienciales. De acuerdo con Nation (1990), el principio general de la carga del aprendizaje es que cuanto más conocimiento y patrones familiares representa una palabra para los estudiantes, más fácil será su aprendizaje. Según el modelo de Nation (2013, pp. 65-86), este conocimiento multifacético de las palabras se esquematiza a través de los siguientes aspectos⁵ :

3.2.1. Forma escrita. De acuerdo con Nation (2001) la forma escrita de una palabra se refiere a la capacidad de escribirla correctamente, sin faltas de ortografía. La ortografía refleja otros aspectos relacionados con el vocabulario y el conocimiento lingüístico en general. Por ejemplo, según Koda (1997), el procesamiento ortográfico ineficiente no solo puede conducir a una recuperación léxica errónea, sino también a bajo

⁵ Aunque el aspecto de la forma oral es una parte de este modelo, fue excluido como no es un aspecto analizado en este estudio.

rendimiento en comprensión de lectura. Asimismo, las faltas de ortografía pueden afectar negativamente la escritura de los estudiantes, ya que usan estrategias para ocultar su ortografía deficiente, lo cual resulta en un vocabulario limitado y en evitar las palabras con ortografía más compleja.

La dificultad de forma escrita de una palabra en L2 está fuertemente relacionada con su forma en L1. Por ejemplo, las lenguas que poseen un sistema de escritura diferente, como el árabe o el chino, pueden plantearle grandes dificultades a un estudiante cuya lengua nativa se basa en un alfabeto latino. También las irregularidades ortográficas o la disociación que existen entre la forma escrita y la pronunciación de las palabras en algunas lenguas puede ser otro factor que ocasione problemas al estudiante.

3.2.2. Partes de una palabra. Nation (2013) considera que los rasgos morfológicos de una palabra constituyen un aspecto importante del conocimiento léxico. En otras palabras, de acuerdo con él, saber una palabra puede implicar saber que está formada por afijos y una raíz que pueden ocurrir en otras palabras. Tratar con las palabras complejas sugiere que hay patrones regulares que son potencialmente previsibles al construir una palabra en una lengua determinada. De hecho, si las reglas de derivación de una palabra son similares a las de la lengua materna del estudiante o a las de otras palabras de la L2 que éste ya conoce, entonces se facilita enormemente su aprendizaje.

Las palabras se almacenan formando redes en el lexicón mental y una de estas redes consistiría en agrupar las mismas por familias de palabras. Por lo tanto, el conocimiento de una palabra implicaría también el conocimiento de los otros miembros de la familia a la que aquella pertenece. A medida que el estudiante va aumentando su

dominio de la lengua meta, profundiza el conocimiento de estas agrupaciones familiares y se incrementa el número de miembros de cada una de ellas.

3.2.3. Conexión entre forma y significado. Para saber una palabra no es suficiente saber cómo suena la palabra (su forma oral), cómo se escribe (forma escrita), y su significado. Según Nation (2013), el estudiante tiene que ser capaz de enlazar la forma de la palabra con el concepto que expresa la misma. Cuanto más fuerte sea esta conexión, más fácil le resultará reconocerla cuando aparezca en un texto oral o escrito, o bien recuperar su forma para usarla de forma productiva. Schmitt (2010) también tiene una perspectiva similar a la de Nation (2013) en cuanto a la importancia de conexión entre forma y significado:

If a lexical form is familiar, but its meaning is not known, then this item is of no communicative use. Likewise, if a meaning is known, but not its corresponding form, then the item cannot be either recognized or produced. It is thus little wonder that most vocabulary material attempt to teach this form-meaning link, and that most test measure it in one-way or another (p.24).

Asimismo, la conexión entre la forma y el significado resulta más sencilla cuando tanto la forma de la palabra como el concepto que expresa coinciden aproximadamente en la lengua nativa y en la lengua meta, como en el caso de los cognados y de los préstamos de otras lenguas. Por otra parte, aunque es el aspecto fundamental del conocimiento léxico, no es suficiente para saber una palabra.

3.2.4. Concepto y referentes. Nation (2013) trata el problema de las palabras con múltiples significados o sentidos, que suelen ser las de alta frecuencia. Por un lado, existen palabras que tienen la misma forma, pero cuyo significado es completamente

diferente. Son los llamados *homónimos*, como en el caso de *banco* (p. ej., En el *banco* se guarda dinero./ ¡Siéntate en el *banco*!). Dentro de estos se incluyen tanto los homógrafos (palabras que se escriben de la misma forma pero con diferente significado como en el caso de *banco, café, capital*), como los homófonos (palabras que se pronuncian igual aunque se escriben diferente como en el caso de *asar/azar; hacia/Asia; casar/cazar*). Estos tipos de palabras polisémicas deberían contarse y aprenderse como palabras diferentes, preferiblemente en diferentes ocasiones. En cuanto a cómo los usuarios de una lengua tratan los diferentes sentidos de estos tipos de palabras, Nation (2013) cita a Nagy (1997) quien propone dos enfoques: (1) considerar que cada significado se almacena en la mente de forma independiente y cuando uno se encuentra con la palabra en cuestión debe elegir el significado más apropiado de la lista que tiene almacenada en la mente. Este proceso se denomina *selección de sentido (sense selection)*; (2) existe un significado subyacente que engloba todos los posibles usos o sentidos de la palabra. Por ejemplo, la palabra *rama* puede aludir a una parte del árbol o una sección de una disciplina. Cuando uno se encuentra con la palabra, se activa un proceso denominado *especificación referencial (reference specification)* que permite inferir, teniendo en cuenta el contexto, cuáles son los objetos del mundo real a los que se refiere la palabra. Según Nation (2013), este segundo enfoque es el enfoque más apropiado para el aprendizaje de palabras en segundas lenguas.

3.2.5. Asociaciones. En cuanto a la organización de la estructura del lexicón mental, Nation (2013) cita el estudio de Miller y Felbaum (1991) que describieron las relaciones semánticas en inglés entre un gran número de palabras y llegaron a la conclusión de que la más importante era la sinonimia. Así, afirmaron que para entender

las asociaciones dentro del lexicon hay que dividir las palabras en categorías gramaticales (i.e., nombres, adjetivos, adverbios y verbos).

Los nombres se agrupan en campos semánticos entre los cuales se establecen una jerarquía. Esta relación semántica y jerárquica entre los nombres se denomina *hiperonimia* y *homonimia*. El *hiperónimo* es el término más general, mientras el *hipónimo* es el más específico. Por lo tanto, en el ámbito referencial se trata de una relación de inclusión entre estos dos términos: *animal* es el hiperónimo de *perro*, *gato*, *caballo*, etc. Cuando la relación consiste en ser una parte de un todo (p. ej., *cocina-casa*), se habla de *merónimos* (i.e., *cocina* es el merónimo de *casa*). Los adjetivos predicativos deben distinguirse de los adjetivos no –predicativos. Esta es una de las muchas ocasiones de conexión entre la organización léxica y el comportamiento sintáctico. Hay antonimia basada en la oposición del significado (*frío/caliente*) y también antonimia basada en la oposición de la forma (p. ej., *ordenado/desordenado*). En cuanto a los verbos, se debe hacer una distinción básica entre los verbos que representan una acción y los que representan un estado (una distinción importante también en el comportamiento gramatical de los verbos). Miller y Fellbaum (citado en Nation, 2013) usan el término *troponimia* para expresar la idea de que algo se hace en una manera particular. Por ejemplo, *dar un paseo* es “andar en una manera particular”. También hay otros tipos de relaciones jerárquicas que Miller y Fellbaum denominan “vinculación”, lo cual significa básicamente involucrarse en una acción implica involucrarse en la otra (p. ej., *roncar/dormir*).

3.2.6. Funciones gramaticales. Para poder usar una palabra es necesario saber la categoría gramatical a la que pertenece y cuál es su comportamiento sintáctico. Nation (2013) menciona a algunos autores que ponen de manifiesto el papel relevante que juega el lexicon en la gramática. Así, el hecho de que la selección de una palabra u otra para expresar un significado, particularmente de los verbos, puede determinar la estructura gramatical del resto de la frase.

3.2.7. Colocaciones. Conocer una palabra también significa saber con qué palabras típicamente ocurre ésta. En todas las lenguas existen palabras que suelen aparecer juntas y que se comportan como un bloque unitario de significado; por ejemplo, *comida rápida* o *cometer un crimen*. Las colocaciones difieren mucho unas de otras en cuanto al número de elementos que las componen, la categoría gramatical de las palabras que aparecen juntas, etc.

Nation (2013, p. 83) cita a Pawley y Sider (1983), quienes argumentan que las colocaciones son uno de los motivos por los que los hablantes nativos de una lengua pueden hablarla de forma fluida. Al tener almacenadas en nuestro cerebro un gran número de secuencias de palabras que funcionan en bloque, podemos recurrir a ellas cuando las necesitamos sin tener que construirlas cada vez ítem por ítem. La investigación sobre las colocaciones demuestra que existen patrones en las lenguas y ser consciente de estos patrones baja la carga de aprendizaje de ciertas palabras en la lengua meta.

3.2.8. Restricciones de uso. En este aspecto Nation (2013) se refiere a las limitaciones en el uso de las palabras por razones sociolingüísticas y culturales. De hecho, en algunas sociedades hay serias restricciones en los términos usados para referirse a personas dependiendo a su edad, trabajo, relación con el hablante, etc.

Nation (2013) cita a Cassidy (1972) y Hartmann (1981) para llamar la atención a la importancia de saber adecuar el léxico a dichas situaciones:

There are several factors where and when certain words can be used. Failure to observe these can result in inappropriate use. One way of seeing the range of constraints on use of words and to gain information about constraints on particular words is to look “style values” or usage labels in dictionaries. (p.84)

En este sentido, el uso de frecuencia de una palabra por los hablantes nativos puede ser un factor importante en determinar sus restricciones de uso. El uso de palabras de baja frecuencia tendría un efecto poco usual en la comunicación. Discusiones explícitas y comparaciones culturales resultarían más productivas en algunos casos con el fin de clarificar estas restricciones .

En síntesis, según el modelo de Nation (2013), conocer una palabra requiere saber distintos aspectos asociados a ella, además de su forma y significado. Ese conocimiento se encuentra en interfaz con otros componentes de la mente, interrelacionados entre sí en los procesos de reconocimiento y de recuperación del ítem que aparece en el acto comunicativo. Esa interfaz está constituida por la memoria, el procesamiento, el conocimiento y la experiencia del mundo, el significado, la función sintáctica, la forma morfológica y la forma fonética.

3.3. La concepción del léxico de las palabras: unidades léxicas, tipos, lemas, familia de palabras

Por cuestiones metodológicas es de gran importancia determinar las unidades para contar las palabras. Nation (2013) plantea tres preguntas básicas al medir el tamaño del vocabulario: 1) ¿Qué se debe contar como una palabra?; 2) ¿Cómo debemos elegir las palabras que se van a incluir en una prueba?; 3) ¿Cómo debemos medir si el estudiante sabe una palabra? (p.527). El mismo investigador (2013), ofrece una clasificación de las unidades de medida léxica en cuatro categorías:

1. **Unidades léxicas (tokens):** una manera de contar las palabras en un texto escrito u oral es simplemente contar cada palabra, en otras palabras, es la unidad de ocurrencias. Aunque las mismas palabras aparezcan dos veces o más, se cuentan como unidades separadas. Por ejemplo, en la oración “*It’s not easy to say it correctly*”, hay 8 *tokens* en total, aunque dos de éstas son las mismas palabras. Esta unidad sirve para contar o medir cuántas palabras hay en una página o en una línea, a qué velocidad uno lee, o cuántas palabras puede un hablante medio producir en un minuto.
2. **Tipos (types):** En la oración anterior, “*It is not easy to say it correctly*” hay 7 *tipos*. Cuando aparece la misma palabra, no se cuenta otra vez. Es la unidad que se usa para responder a preguntas como cuántas palabras se necesitan para leer determinado libro, o cuántas palabras hay en un diccionario.
3. **Lemas (lemmas):** Se entiende por *lema* el conjunto de formas léxicas que tienen la misma base y pertenecen a la misma categoría gramatical (p. ej., *libro, libros*), aunque puedan tener inflexiones u ortografía diferentes (p. ej.

mouse, mice). Un *lema* es una palabra base (*headword*) más algunas de sus inflexiones; generalmente todos los ítems que se incluyen en un lema son de la misma categoría gramatical (p. ej., *ir* es el lema de *voy, vas, ibamos, fueron* y el resto de sus formas conjugadas). Para el Diccionario de la Real Academia de la Lengua Española el término *lema* equivale a entrada de un diccionario o enciclopedia que aparece representado en letra negrita.

4. Familia de palabras (*word families*): una familia de palabras consta de una palabra, o sea una base (*headword*), sus inflexiones y sus derivados más cercanos (los formados con afijos tales como *-ly, -ness* y *un-* en inglés) (pp. 9-11).

En este estudio se decidió usar esta última categoría como la unidad de medida léxica de los participantes. En el Capítulo 4, dentro de la sección del diseño del corpus, se explica en detalle qué se incluye en las familias de palabras de este estudio.

Esta terminología puede ser sin duda de gran ayuda para establecer medidas, evaluar y describir con claridad el conocimiento léxico de tanto los hablantes nativos como de los estudiantes de segundas lenguas. Para medir el conocimiento léxico Nation (2013) propone una prueba de vocabulario con *ítems* de vocabulario organizados en grupos según la frecuencia que tienen en listados elaborados previamente a partir de un corpus.

3.4. Las variables lingüísticas del estudio: el conocimiento receptivo y productivo de las palabras

De acuerdo con Nation (2013), la distinción entre el conocimiento receptivo y el productivo del vocabulario en muchos casos depende de la distinción entre las destrezas

receptivas de comprensión auditiva y lectora, y las destrezas productivas de expresión oral y escrita. El término *receptivo* transmite la idea de una entrada de lengua que se recibe desde otros individuos a través de la lectura o la escucha, en otras palabras, el individuo recibe la entrada e intenta comprenderla. A través de la palabra *underdeveloped*, Nation (2013) explica qué implica saber la palabra desde el punto de vista (del uso y conocimiento) receptivo de la siguiente manera:

- being able to recognize the word when it is heard;
- being familiar with its written form so that it is recognized when it is met in reading;
- recognizing that it is made up of the parts *under-*, *develop-* and *-ed* and being able to relate these parts to its meaning;
- knowing that *underdeveloped* signals a particular meaning;
- knowing what the word means in the particular context in which it has just occurred;
- knowing the concept behind the word which will allow understanding in a variety of contexts;
- knowing that there are related words like *overdeveloped*, *backward*, *challenged*;
- being able to recognize that *underdeveloped* has been used correctly in the sentence in which occurs;
- being able to recognize that words such as *territories* and *areas* are typical collocations; and
- knowing that *underdeveloped* is not an uncommon word and a pejorative word (pp.48-49).

Y, según el mismo autor (2013), desde el punto de vista (del uso y conocimiento) productivo, saber la palabra *underdeveloped* implica:

- being able to say it with correct pronunciation including stress;
- being able to write it with correct spelling;
- being able construct it using the right word parts in their appropriate forms;
- being able to produce the word to express the meaning underdeveloped;
- being able to produce the word in different contexts to express the range of meaning of *underdeveloped*
- being able to produce synonyms and opposites for underdeveloped;
- being able to use the word correctly in an original sentence;
- being able to produce words that commonly occur with it; and
- being able to decide to use or not use the word to suit the degree of formality of the situation. (p. 50)

La distinción entre el conocimiento receptivo y productivo de las palabras también se refleja en la siguiente tabla analítica de Nation (2013), a través de la cual el investigador amplía los componentes propuestos por Richards (1976), que supone saber una palabra : la forma, el significado y la forma, y el uso desde el aspecto receptivo (R) y productivo (P).

Table 6. What is involved in knowing a word

Form	spoken	R	What does the word sound like?
		P	How is the word pronounced?
	written word parts	R	What does the word look like?
		P	How is the word written and spelled?
		R	What parts are recognizable in this word?
		P	What word parts are needed to express the meaning?
Meaning	form and meaning	R	What meaning does this word form signal?
		P	What word form can be used to express this meaning?
	concept and referents associations	R	What is included in the concept?
		P	What items can the concept refer to?
		R	What other words does this make us think of?
		P	What other words could we use instead of this one?
Use	grammatical functions	R	In what patterns does the word occur?
		P	In what patterns must we use this word?
	collocations	R	What words or types of words occur with this one?
		P	What words or types of words must we use with this one?
	constraints on use (register, frequency...)	R	Where, when, and how often would we expect to meet this word?
		P	Where, when, and how often can we use this word?

Note ¹. R = receptive knowledge, P = productive knowledge

Note ². Adapted from *Learning Vocabulary in Another Language (2nd. Ed.)* by P. Nation, P., 2013, p. 49.

De acuerdo con Schmitt (2010), todas estas dimensiones mencionadas en la Tabla 6 están interrelacionadas y holísticamente conectadas. Sin embargo, no es posible evaluar todas éstas individualmente en ningún tipo de prueba. Aunque si fuera posible, la prueba sería muy larga y compleja para los fines de investigación y usos pedagógicos. Por lo tanto, los investigadores de vocabulario tienen que considerar qué aspecto limitado del conocimiento léxico van a evaluar en su estudio, y considerar cuidadosamente las

limitaciones y conclusiones de sus elecciones (p.18). El presente estudio se enfoca en el aspecto receptivo y productivo de la forma escrita del conocimiento léxico.

Particularmente, se decidió analizar el reconocimiento del significado (*meaning recognition*) y recuperación de la forma (*form recall*) escrita por los siguientes motivos:

Some aspects of vocabulary knowledge are relatively easy to measure in research (e.g. written form, meaning), while some are extremely difficult to capture (register, collocation). In addition, although all of the word types are learned concurrently, some are mastered sooner than others (Schmitt, 1998). This has implications for research, as different vocabulary measures might be appropriate at the different stages of acquisition of an item. At the beginning of the incremental learning process, measuring the meaning-form link is probably most appropriate, but as the word becomes more established, it might be better to measure some of the contextual types of word knowledge (e.g. collocation) to determine the degree of higher-level mastery of a lexical item (Schmitt, 2010, p.16).

Así, es necesario desarrollar instrumentos que sean sensibles a los diferentes grados de adquisición léxica para medir tanto el conocimiento receptivo como productivo. De hecho, los resultados de los estudios que combinan múltiples medidas de conocimiento léxico darán resultados mejores y más completos en comparación con los estudios que miden simplemente un aspecto del conocimiento léxico (Schmitt, 2010; Webb, 2005).

De acuerdo con Nation (2013) y Laufer et al. (2004), parece que en general el conocimiento receptivo de una palabra es más fácil y anterior al conocimiento

productivo. La dificultad que entraña el uso productivo supone que el aprendizaje exige más tiempo y esfuerzo (p. ej., la cantidad de conocimiento, la práctica, el acceso y la motivación) que el aprendizaje receptivo de una palabra. Además, el uso productivo de una palabra requiere el aprendizaje de nuevas formas habladas y escritas, mientras que para el uso receptivo de la palabra el estudiante sólo tiene que conocer unos cuantos rasgos distintivos de la palabra; para el uso productivo el conocimiento tiene que ser más preciso.

En resumen, el conocimiento productivo supone el receptivo y como consecuencia el vocabulario receptivo se considera de mayor tamaño que el productivo aunque no existe una medida para determinar hasta qué extremo uno es mayor que el otro.

Capítulo 4

Metodología

It is better to establish a norm in which multiple measures of vocabulary are used in studies to paint a more complete picture of vocabulary knowledge and acquisition. This could be in terms of receptive/productive mastery, different types of word knowledge, degree of mastery of an individual word knowledge aspect, contexts of use, etc. or some combination of these. (Schmitt, 2010, p.22)

En este capítulo se describe la metodología utilizada para llevar a cabo la investigación. La sección 4.1 presenta los participantes y sus características. En la sección 4.2. se detalla la formación del corpus léxico que sirve de base al diseño de los instrumentos del estudio. A ésta le sigue la sección 4.3 donde se presentan los tres instrumentos que se utilizaron para lograr los objetivos de este estudio: (4.3.1) el cuestionario sociolingüístico, (4.3.2) la prueba receptiva de niveles de vocabulario (*Receptive Vocabulary Levels Test*) y (4.3.3) la prueba productiva de niveles de vocabulario (*Productive Vocabulary Levels Test*) incluyendo la descripción de los mismos. Por último, la sección 4.4. resume los procedimientos aplicados en su implementación y evaluación.

4.1. Participantes

En total 783 estudiantes de español como segunda lengua (EL2) participaron en este estudio. Sin embargo, no todos estos participantes se incluyeron en la investigación. Según los resultados del cuestionario sociolingüístico, 373 participantes fueron eliminados por diversos motivos con respecto a los propósitos del estudio. Se excluyeron los siguientes participantes: (1) aquellos que no completaron los tres instrumentos del

estudio; algunos completaron sólo el cuestionario, mientras que otros completaron sólo la prueba receptiva o productiva del vocabulario. (2) También se eliminaron los participantes que fueron estudiantes de español como lengua de herencia⁶ ya que el propósito principal de este estudio fue diseñar una herramienta fiable y válida que pueda evaluar la competencia léxica de los estudiantes de EL2. En realidad, los datos de este estudio fueron recopilados del *track* que fue diseñado exclusivamente para los estudiantes de EL2. Sin embargo, cada semestre hay estudiantes de español como lengua de herencia (ELH) en el programa que deben inscribirse en el *track* de EL2 por diferentes motivos personales o académicos. (3) De la misma manera se excluyeron los estudiantes filipinos que hablan filipino considerando el hecho que esa lengua contiene vocablos que tienen su origen en la lengua castellana. (4) Otro objetivo de esta investigación es observar qué tanto éxito tiene el programa en cuanto a la enseñanza del vocabulario, por lo tanto, se decidió excluir los estudiantes de transferencia que anteriormente estudiaron EL2 en una institución diferente. (5) Asimismo, los estudiantes que tomaron el examen de ubicación para los EL2 y recibieron créditos por el/los nivel/es anterior/es no se incluyeron en el estudio. (6) No se incluyeron tampoco los datos de los estudiantes que anteriormente habían tomado SPAN 1505. Este es un curso básico intensivo de un semestre que es el equivalente de los dos cursos de primer año (SPAN 1501 y SPAN 1502) y en este nivel se utiliza otro libro de texto que no se usó en la formación del corpus léxico de este

⁶ Según la descripción de Valdés (2000:1), el estudiante de herencia se define como un individuo “que creció en un hogar donde se habla una lengua diferente al inglés, que habla o al menos entiende tal lengua que es hasta cierto nivel bilingüe en tal lengua y en inglés”. Beaudrie y Fairclough (2012) prefieren adoptar la definición más amplia e inclusiva de Fishman (2001) que lo define como un individuo que tiene conexión individual o familiar con la lengua no mayoritaria.

estudio, por la cual no sería justo incluir a los estudiantes que tomaron anteriormente este curso. (7) Por último, también se excluyeron aquellas personas que repitieron un curso en el programa, ya que algunos de estos participantes estudiaron con otros libros de textos en diferentes semestres.

Al final, la muestra final de este estudio se redujo a 410 estudiantes, 262 mujeres y 148 hombres, de una universidad pública del Suroeste de los Estados Unidos. Todos los participantes completaron un cuestionario acerca de sus experiencias y antecedentes lingüísticos con respecto a la lengua española. Según los resultados de este cuestionario, los participantes resultaron ser en su gran mayoría jóvenes universitarios entre 18 y 26 (media= 23 años; moda= 21 años) años de edad, todos son estudiantes del español como segunda lengua⁷ (ESL) en el Departamento de Estudios Hispánicos. En la siguiente tabla se presenta la distribución de los participantes por niveles y grupos junto con algunos datos sociolingüísticos básicos que caracterizan la muestra final de este estudio.

⁷ Según Gass y Selinker (2008: 7), *Second Language Acquisition (SLA)* is the common term used for the name of the discipline. In general, *second language learner* refers to the person who learns another language after his native language (L1). Sometimes the term refers to the learner of a third or fourth language. The important aspect is that it refers to the learner of a nonnative language *after* his L1. The second language is commonly referred to as the L2. As with the phrase second language, L2 can refer to any language learned *after* learning the L1, regardless of whether it is the second, third, fourth, or fifth language.

Tabla 7. Distribución de los participantes por niveles y grupos

Grupo	Nivel	N (%)	Sexo F	Sexo M
G1	1501- Preprueba	51 (12.4%)	27 (6.6%)	24 (5.9%)
G2	1501	95 (23.2%)	57 (13.9%)	38 (9.3)
G3	1502	90 (22.0%)	55 (21.0%)	35 (8.5)
G4	2301	66 (16.1%)	45 (17.2%)	21 (5.1)
G5	2302	90 (22.0%)	63 (24.0%)	27 (6.6)
G6	3000	18 (4.4%)	15 (5.7%)	3 (0.7)
Total		410 (100%)	262(63.9%)	148 (36.1%)

La selección y categorización de los estudiantes según niveles fue hecha mediante el cuestionario sociolingüístico. Como se observa en la Tabla 7, la muestra final fue clasificada en 6 grupos considerando el nivel de español basado en la cantidad de semestres que han estudiado esta lengua en un contexto académico universitario.

Todos los participantes menos aquellos que completaron la pre-prueba cursaban clases de español durante la recopilación de los datos, que se realizó en el semestre de la Primavera del 2014. Los datos del grupo de pre-prueba fueron recopilados al principio del semestre académico de Otoño del 2014. La Facultad de Artes Liberales y Ciencias Sociales en la Universidad de Houston requiere que todos los estudiantes en las humanidades completen un mínimo de dos cursos (6 horas de crédito) del nivel 2000 o superior en una de las lenguas extranjeras ofrecidas por la universidad. En algunos casos, la carrera misma se centra en el estudio del español como especialización o sub-especialización, y por consiguiente, los estudiantes deben tomar clases de niveles 3000 y más avanzados.

Tabla 8. Distribución de participantes por estudio formal de español

Grupo/N	Nivel	Con estudios previos al nivel universitario	Sin estudios previos al nivel universitario
G1/51	1501-Preprueba	36 (8.8%)	15 (3.7%)
G2/95	1501	58 (14.1%)	37 (9.0%)
G3/90	1502	76 (18.5%)	14 (3.4%)
G4/66	2301	58 (14.1%)	8 (2.0%)
G5/90	2302	69 (16.8%)	21 (5.1%)
G6/18	3000	18 (4.4%)	0 (0.0%)
Total/410 (100%)		315 (76.8%)	95 (23.2%)

La Tabla 8 muestra la distribución por nivel de los participantes que tienen estudios previos y los que no tienen educación formal en EL2. Según los resultados de esta tabla, de un total de 410 participantes, 95 (23.2%) nunca han tomado clases de español previo a nivel universitario, mientras 315 (76.8%) han estudiado este idioma con un promedio de 2.4 años entre el sexto y el duodécimo grado. La edad promedio de la primera exposición al español de los participantes en el ámbito escolar es de 16.6 años. Como se mencionó anteriormente, la muestra final ($n=410$) fue clasificada en 6 grupos considerando el nivel de español basado en la cantidad de semestres que han estudiado esta lengua en un contexto académico universitario. La siguiente tabla refleja en detalle el promedio total de años de estudio de EL2 en el nivel universitario y secundario por nivel.

Tabla 9. Distribución de participantes según el promedio total de años de estudio de EL2

Grupo/N	Nivel	En el nivel universitario	En el nivel secundario	Promedio total de años de estudio
G1/51	1501-Preprueba	1 semana	2.6 años	2.6 años
G2/95	1501	0.5 año ¹	2.3 años	2.9 años
G3/90	1502	1 año	2.2 años	3 años
G4/66	2301	1.5 año	2.2 años	3.4 años
G5/90	2302	2 años	2.6 años	4 años
G6/18	3000	3.7 años	2.6 años	6.2 años
Total/410			2.4 años	3.7 años

Nota: ¹Un semestre corresponde a 0.5 año, 2 semestres corresponden a 1 año, etc.

El Grupo 1 (G1) está compuesto por 51 estudiantes de EL2 que están en el nivel inicial del programa universitario en que se realiza el estudio. Empezaron a aprender este idioma a la edad promedio de 17.2 años. 15 (30%) de estos estudiantes nunca habían estudiado español antes de ingresar al programa y 36 (70%) de ellos han estudiado este idioma con un promedio de 2.6 años en el nivel secundario. Los datos de este grupo fueron recopilados al principio del semestre, durante la primera semana del semestre académico de Otoño 2014, por lo tanto, los participantes de este grupo sirven como punto de partida para esta investigación.

El Grupo 2 (G2) está constituido por 95 participantes. Al igual que el grupo anterior, los participantes de este grupo también son estudiantes que están en el primer semestre del programa. Sin embargo, a diferencia del G1, los datos de este grupo fueron recopilados al final del primer semestre del programa. 37 (39%) de ellos nunca habían estudiado español antes y 58 (61%) de ellos habían estudiado EL2 con un promedio de 2.3 años previo al ingreso al programa universitario.

El Grupo 3 (G3) está constituido por 90 estudiantes de EL2 que están en el segundo semestre del programa en que se realizó el estudio, 14 (16%) de ellos nunca han

estudiado español antes de ingresar al programa y de manera similar al grupo anterior 76 (84%) de ellos han estudio esta lengua con un promedio de 2.2 años en el nivel secundario.

El Grupo 4 (G4) está compuesto por 66 estudiantes de ESL que han cursado 3 semestres, 1.5 año, en el programa. 8 (11%) de ellos nunca habían estudiado español previo al nivel universitario y 58 (89%) de ellos, al igual que el Grupo 3, habían estudiado esta lengua con un promedio de 2.2 años en el nivel secundario. Estos participantes han tenido su primera exposición a la lengua español a la edad promedio de 16.5 años.

En el caso de los participantes de G1, G2, G3 y G4 de este estudio, nos encontramos por lo general con estudiantes con la habilidad de repetir, reciclar y producir frases sueltas o palabras aprendidas en el aula. El estudiante promedio de estos grupos no puede usar estas palabras o frases de forma espontánea en una conversación y más bien se limita a contestar preguntas de forma bastante limitada. Su vocabulario es mínimo y se ubica en el contexto de lo aprendido en el aula. Frecuentemente a una pregunta responde reciclando la pregunta recibida y a menudo repite lo que dice. En cuanto a la precisión gramatical de su producción, se observa una serie de errores e imprecisiones, típicas del interlenguaje de los niveles más básicos, y su mensaje es difícil de comprender por un interlocutor no familiarizado con hablantes no-nativos (Mir, 2007).

Sin embargo, aquí cabe mencionar, el nivel de competencia comunicativa de los participantes depende también a los factores individuales como educación formal en el nivel secundario, edad de primera exposición a la lengua española, estudio de otras lenguas extranjeras, experiencia con la lengua española fuera del salón de la clase, etc....

Por lo tanto, estas descripciones de competencia comunicativa para los grupos se intentaron dar tomando en cuenta un estudiante promedio y los objetivos de los cursos diseñados para los diferentes niveles en el programa.

El Grupo 5 (G5) está constituido por 90 participantes que están en el cuarto semestre del programa de EL2. La mayoría de los estudiantes de este grupo están por terminar los requisitos de lengua extranjera que exige su especialidad. 69 (77%) de ellos habían estudiado la lengua española con un promedio de 2.6 años en el nivel secundario y el resto (23%) nunca estudió español antes de ingresar a la universidad. Al igual que el grupo anterior, estos participantes han tenido su primera exposición a esta lengua a la edad promedio de 16.5 años.

El Grupo 6 (G6) lo conforman 18 estudiantes de EL2 que ya habían terminado la segunda parte de la secuencia de nivel intermedio (SPAN 2302-Intermediate Spanish II) y que habían acumulado un mínimo de 16 horas de crédito (4 cursos) de primero y segundo año. Estos participantes comenzaron el estudio formal del español a la edad promedio de 13.6 años y han estudiado esta lengua por 3 años y 7 meses años en el nivel universitario. Los estudiantes de este grupo han seleccionado el español como carrera, principal o secundaria. Los datos de este grupo fueron recopilados de los estudiantes que estaban tomando los siguientes cursos avanzados: SPAN 3301 (Spanish Oral Communication for Critical Thinking), SPAN 3305 (Spanish Grammar Review) y SPAN 3306 (Introduction to the Study of Spanish Language). Aunque estas clases no solamente incluyen estudiantes de EL2, sino también estudiantes de ELH, se eliminaron los estudiantes que pertenecen a la segunda categoría al no ser el enfoque de esta investigación.

Los estudiantes promedios de este grupo (G6) se pueden caracterizar principalmente por la habilidad de crear con el lenguaje. Pueden hacer preguntas y contestarlas sin problemas siempre que los temas sean sobre su entorno personal. Pueden llevar a cabo con éxito transacciones sociales que ocurren diariamente como por ejemplo, invitar a un amigo al cine, comprar billetes de tren, hacer una reserva de hotel, etc. El vocabulario se centra principalmente en el contexto familiar y concreto. Saben defenderse al hablar sobre temas concretos y personales como la familia, el hogar, el trabajo, el ocio. Su producción tiene errores gramaticales, sintácticos y de pronunciación que a veces interfieren con la comprensión del mensaje pero un interlocutor sensible a las limitaciones de un hablante no-nativo puede entenderlo sin problemas. A este nivel el hablante funciona principalmente al nivel de la descripción simple con oraciones conectadas de forma sencilla. Puede enumerar eventos ocurridos en el pasado pero no puede mantener una narración completa y detallada en el pasado.

Tabla 10. Número y distribución de los participantes según sus antecedentes lingüísticos

Grupo/N	Nivel	Hablan inglés como L1	Hablan otras lenguas como L1	Hablan sólo inglés	Hablan otras lenguas
G1/51	Preprueba 1501	43 (10.5%)	8 (2.0%)	43 (10.5%)	8 (2.0%)
G2/95	1501	73 (17.8%)	22 (5.4%)	67 (16.3%)	28 (6.8%)
G3/90	1502	81 (19.8%)	9 (2.2%)	69 (16.8%)	21 (5.1%)
G4/66	2301	61 (14.9%)	5 (1.2%)	61 (14.9%)	5 (1.2%)
G5/90	2302	85 (20.7%)	5 (1.2%)	82 (20.0%)	8 (2.0%)
G6/18	3000	17 (4.1%)	1 (0.2%)	14 (3.4%)	4 (1.0%)
Total	410	360 (87.8%)	50 (12.2%)	336 (82.0%)	74 (18%)

En cuanto a los antecedentes lingüísticos de la muestra final, como se observa en la Tabla 10, mientras 87.8% de los participantes habla inglés, 12.2% habla otros idiomas

(p.ej., vietnamita, chino, coreano, urdu, turco y árabe) como su primera lengua. De manera similar, 82% puede hablar sólo inglés y 18% habla otras lenguas diferente al inglés. En cuanto al uso de la lengua en casa, 86% de los participantes reportaron hablar sólo inglés en casa.

Tabla 11. Distribución de los grupos por estudio de otra LE y estancia en un país hispanohablante

Grupo/N	Nivel	Estudio de otra LE		Estancia en un país hispanohablante	
		Sí	No	Sí	No
G1/51	1501- Preprueba	11 (2.7%)	40 (9.8%)	25 (6.1%)	26
G2/95	1501	41 (10.0%)	54 (13.2%)	31 (7.6%)	64
G3/90	1502	22 (5.4%)	68 (16.6%)	34 (8.3%)	56
G4/66	2301	11 (2.7%)	55 (13.4%)	28 (6.8%)	38
G5/90	2302	29 (7.1%)	61(14.9%)	40 (9.8%)	50
G6/18	3000	5 (1.2%)	13(3.2%)	6 (1.5%)	12
Total/410		119(29.0%)	291(71.0%)	164(40.0%)	246(60%)

En la Tabla 11 se presenta se presenta la distribución de los grupos por estudio de otra lengua extranjera (LE)⁸ y estancia en un país hispanohablante. Según los resultados de esta tabla, el 29% de los participantes de la muestra final estudiaron otra lengua extranjera diferente a la lengua española (p.ej., francés, latín, alemán, chino, japonés, árabe, etc.) antes. En el cuestionario sociolingüístico se les pidió a los participantes que dieran más información acerca de sus experiencias con la lengua española. Sólo 1% indicó haber estudiado este idioma en un país hispanohablante, mientras 40% ha estado

⁸ De acuerdo con el Diccionario de términos clave de ELE, formado por Centro Virtual de Cervantes, aquí se prefirió usar el término *lengua extranjera* en vez de *segunda lengua* con el fin de distinguir las situaciones y circunstancias en que se aprende la lengua meta (LM), se suele distinguir entre lengua extranjera (LE) y lengua segunda (L2). Cuando la LM se aprende en un país donde no es ni oficial ni autóctona, se considera una LE; p. ej., el español en Brasil. Cuando la LM se aprende en un país donde coexiste como oficial y/o autóctona con otra(s) lengua(s), se considera una L2; p. ej., el guaraní para aquellos niños paraguayos cuya L1 es el español, o bien el español para aquellos niños paraguayos cuya L1 es el guaraní.

en un país hispanohablante (México, Costa Rica, España, República Dominicana, etc.) entre 1 día y 1 mes de vacaciones, 60% reportó nunca haber estado en un país hispanohablante. También se les preguntó a los participantes de este estudio si habían tenido otro tipo de experiencia con la lengua española fuera del salón de clases. Más de la mitad (57%) respondió afirmativamente a esta pregunta informando sobre las experiencias que han tenido con los hablantes nativos como por ejemplo, haber tenido una niñera hispanohablante durante la infancia, trabajar con gente hispanohablante, haberse casado con un/a hispanohablante etc. Por otra parte, 43% respondió negativamente a esta pregunta reportando no haber tenido experiencias con la lengua española fuera del salón de clases.

4.2. Diseño del corpus

Con el fin de formar un corpus léxico para poder evaluar el vocabulario receptivo y productivo de los estudiantes de EL2 se analizaron las listas de palabras que aparecen al final de cada capítulo en los libros de textos usados en el primer y segundo año del programa de EL2 en la UH durante el semestre de otoño de 2013. Se decidió enfocarse principalmente en estas listas que se consideran como el vocabulario activo que los estudiantes tienen que aprender para poder cumplir los requisitos del nivel en cuanto al conocimiento léxico. De hecho, tanto en las pruebas que toman regularmente a lo largo del semestre como en las pruebas de producción oral y escrita se evalúa este vocabulario activo que han estudiado en el salón de la clase. En la Tabla 12 se puede ver la distribución de los capítulos estudiados según los libros de texto y los cuatro niveles del primer y segundo año de este programa para el semestre de primavera de 2014.

Tabla 12. Los libros de textos y capítulos que forman la base del corpus de este estudio

Año, semestre	Nivel	Libro de texto	Capítulos
1 año, 1er. semestre	SPAN 1501	<i>Puntos de Partida</i> , 9Ed	Cap. 1-6
1 año, 2do. semestre	SPAN 1502	<i>Puntos de Partida</i> , 9Ed	Cap. 7-12
2 año, 1er. semestre	SPAN 2301	<i>Más</i> , 2 Ed.	Cap.1-5
2 año, 2do. semestre	SPAN 2302	<i>Más</i> , 2 Ed.	Cap.6-9

Las listas de palabras fueron sacadas de 21 capítulos en total. Después, cada palabra fue etiquetada (*tagged*) según su categoría gramatical, es decir, a cada palabra le fue asignada una parte del habla. Se incluyó sólo las palabras de contenido o categorías léxicas (*content words or lexical words*) que incluyen nombres, verbos, adjetivos y adverbios y si la misma palabra apareció con la misma categoría gramatical en diferentes capítulos, se incluyó la palabra que apareció por primera vez y se excluyeron sus repeticiones en los capítulos subsiguientes. A continuación en la Tabla 13 se presentaron algunos ejemplos de estas palabras incluidas y excluidas:

Tabla 13. Algunos ejemplos de las palabras incluidas y excluidas del corpus

	Incluida	Excluida
<i>bibliotecario</i>	(n) Cap. 2 En la Universidad (1501)	Cap.4. Con el sudor de tu frente (2301)
<i>bienestar</i>	(n) Cap.11 La salud (1502)	Cap. 6 La buena vida (2302)
<i>pasatiempo</i>	(n) Cap.10 El tiempo libre (1502)	Cap. 6. La buena vida (2302)
<i>buscar</i>	(v) Cap.2 En la Universidad (1501)	Cap.5 El mundo al alcance de un clic (2301)
<i>faltar</i>	(v) Cap.9 Los días festivos (1502)	Cap. 2 Yo soy yo y mis circunstancias (2301)
<i>llorar</i>	(v) Cap.9 Los días festivos (1502)	Cap.3 Raíces (2301)
<i>mandar</i>	(v) Cap. 12 ¡Conectados! (1502)	Cap.3 Raíces (2301)

En cuanto a las palabras repetidas que aparecieron con diferentes significados semánticos se tomó en cuenta el significado semántico presentado en el capítulo correspondiente. Por lo tanto, como se puede ver en la Tabla 14, en el corpus de este estudio algunas palabras se incluyeron dos veces pero con diferentes significados semánticos.

Tabla 14. Las palabras repetidas con diferentes significados semánticos en el corpus

	Significado semántico	Nivel		Significado semántico	Nivel
la bolsa	<i>bag</i>	1501	la bolsa	<i>stock exchange</i>	2302
la nota	<i>grade</i>	2301	las notas	<i>(class) notes</i>	2301
la noticia	<i>piece of news</i>	2301	las noticias	<i>news</i>	2302
el plato	<i>plate</i>	1501	el plato	<i>course</i>	1502
la tierra	<i>soil</i>	2302	la Tierra	<i>Earth</i>	2302

Como la meta principal de este estudio es medir el vocabulario receptivo y productivo de los estudiantes de EL2, se decidió excluir las categorías funcionales que constituyen paradigmas cerrados y que tienen un contenido semántico limitado. A continuación se presentan las categorías funcionales excluidos del corpus con algunos ejemplos:

1. **Preposiciones:** *a, de, en, con, para, según, sobre, sin, a la derecha de, a la izquierda de, al lado de, cerca de, debajo de, delante de, detrás de, encima de, entre, lejos de.*
2. **Conjunciones:** *pero, y, o, porque, que, si, mientras*
3. **Las palabras transitivas:** *por eso, por fin, sin embargo, con respecto a, en cuanto a, sobre (el tema de)*
4. **Los pronombres:**
 - a) **Personales:** *yo, tú, usted (Ud.), él/ella, nosotros/ nosotras, vosotros /vosotras, ustedes (Uds.), ellos/ellas*
 - b) **Pronombres preposicionales:** *conmigo, mí, contigo, ti*
 - c) **Indefinidos:** *Algo, alguien, nada, nadie, alguna/o (s), ninguna/o(s)*
 - d) **Pronombres y adjetivos demostrativas:** *este/a, estos/as, esto, aquel, aquella, aquellos/as, ese/a, esos/as, eso, aquello*

e) **Pronombres interrogativas:** *¿cómo?, ¿dónde?, ¿qué?, ¿quién (es)? , ¿cuál (es)?, ¿cuándo?, ¿cuánto/a?, ¿cuántos/as? , ¿adónde?, ¿a qué hora?, ¿de dónde?, ¿de quién(es)?, ¿por qué?*

5. **Los adjetivos posesivos:** *mi(s), tu(s), nuestro/a(s), vuestro/a(s), su(s)*

Las unidades léxicas que se forman por las secuencias de palabras y/o cuyo significado no es compositivo, es decir, cuyo significado no se deriva de sus componentes fueron excluidas del corpus también, considerando el hecho de que el conocimiento de estas unidades abarcan aspectos sintáctico-semánticos y pragmáticos que requieren ir más allá del mero significado de la palabra. Como por ejemplo, los saludos, las expresiones idiomáticas, las frases de cortesía, los exclamativos, los nombres propios, los nombres y adverbios compuestos de más de una palabra que aparecieron en las listas de vocabulario en los libros de textos analizados fueron excluidos del corpus. Los siguientes ejemplos ilustran este tipo de unidades léxicas excluidas del corpus:

1. **Saludos, expresiones de cortesía y exclamativos:** *¿Qué tal?/ ¿Cómo está?/ Buenos días. / Buenos tardes. / Hasta mañana. / Hasta luego. / Nos vemos. / Por favor. /Perdón. / (Con) permiso. /Mucho gusto. / Igualmente. / Encantado/a./ Me gustaría (mucho).../ ¡Buen apetito/provecho!/¡Qué rico/sabroso/malo!*

2. **Expresiones idiomáticas :**

a) **Expresiones con estar:** *estar de moda, está nublado, estar bien, está bien, estar a dieta, estar a dieta, estar de vacaciones, estar unidos/distanciado.*

b) **Expresiones con dar:** *dar una caminata, dar un paseo, darse la mano.*

c) **Expresiones con hacer:** *hacer ejercicio, hacer un viaje, hacer una pregunta, hace buen/mal tiempo /calor/ fresco/ frío/ sol/ viento, hacer cola, hacer*

escalas/paradas, hacer las maletas, hacer camping, hacer una fiesta, hacer un picnic, hacer planes, hacer la cama, hacer ejercicios aeróbicos, hacer una búsqueda.

- d) Expresiones con ir:** *ir de compras, ir al extranjero, ir en autobús/avión/ barco/ tren, ir de vacaciones.*
 - e) Expresiones con llevar:** *llevar brackets, llevar lentes de contacto, llevarse bien/mal, llevar una vida sana/tranquila*
 - f) Expresiones con tener:** *tener ganas de + inf., tener miedo de, tener prisa, que + inf., tener razón, tener sueño, tener calor /frío, tener hambre, tener sed, tener dolor de, tener buen/mal carácter; tener complejo (de, por), tener sentido del humor, tener papeles.*
 - g) Expresiones con salir:** *salir bien/mal, salir de, salir con, salir para.*
 - h) Expresiones con otros verbos:** *facturar el equipaje, pasar por el control de seguridad/la aduana, sacar fotos, salir de vacaciones, tomar el sol, tomar unas vacaciones, cumplir años, dar una fiesta, pasarlo bien/mal, portarse bien/mal, dejar de + inf., guardar cama, levantar pesas, ponerle una inyección, sacar la lengua, sacarle un diente/una muela, tocarle a uno.*
- 3. Nombres Propios:** *La Pascua Florida, La Pascua Judía, La Navidad, El Día de Acción de Gracias, El Fondo Monetario Internacional (FMI), El Producto Interno Bruto (PIB).*
- 4. Nombres compuestos de más de una palabra:** *el salón de clase, la administración de empresas, los pantalones cortos, la ropa interior, el traje de baño, el centro comercial, el color café, la tarjeta de crédito, el asistente*

de vuelo, el control de seguridad, el medio de transporte, la puerta de embarque, la sala de espera, la sala de fumar, la tarjeta de embarque, el estado afectivo, los lentes de contacto, el buzón de voz, el correo electrónico, el tiempo libre, los rotos libres.

- 5. Adverbios compuestos de más de una palabra:** *con frecuencia, el fin de semana, por la mañana, por la tarde, por la noche, todos los días, por lo general, en seguida, ahora mismo, el fin de año.*

Sin embargo, como se pueden ver en los siguientes ejemplos las palabras compuestas que se conforman de dos palabras simples pero que dan un significado fueron incluidas en el corpus: *anteayer, el cumpleaños, la medianoche, el mediodía; el pasatiempo, el quehacer, la quinceañera, el bienestar.*

Con respecto al género biológico y el número de los sustantivos, aunque en el vocabulario activo de los libros de texto estas palabras se presentaron en las dos formas de género (masculina y femenina, singular/plural), sólo se incluyó la forma masculina singular por ser la forma menos marcada, más usada y básica:

el señor, el abuelo, el tío, el amigo, el compañero, el pasajero, el invitado, el paciente, el dueño, el inquilino, el vecino, el ahijado, el bisabuelo, el cuñado, el hijastro, el sobrino, el hermanastro, el empleado, el empleador.

Por otro lado, los sustantivos que se presentaron sólo en plural se incluyeron tal como aparecen en los libros de texto considerando el hecho de que éstas son las palabras que suelen aparecer en forma plural en la vida cotidiana y generalmente los estudiantes de ESL las asimilan de esa manera:

los pantalones, las chanclas, los jeans, los mariscos, la vacaciones, las tapas, el lavaplatos, el microondas, los pulmones, los anteojos, los lentes, los patines, los ojos, las canas, las pecas, los anteojos, los impuestos, las noticias, etc.

Con relación a *las profesiones* se incluyó sólo la forma masculina singular también por ser la forma más usada y menos marcada:

el profesor, el bibliotecario, el piloto, el médico, el consejero, el enfermero, el farmacéutico, el dentista, el médico, el jefe, el portero, el niño, el agricultor, el albañil, el basurero, el jardinero, el pintor, el gerente

En cuanto a *los adjetivos*, de la misma manera, se incluyó sólo la forma masculina aunque se presentaron en las dos formas de género (masculina y femenina) en los libros de texto:

alto, antipático, nuevo, abierto, aburrido, limpio, asado, fresco, ligero, mareado, pasado, resfriado, cariñoso, callado, chistoso, tacaño, ciego, mudo, sordo.

Con respecto a los cognados, según algunos lingüistas (Martin, 2012), por su parecido y facilidad en el reconocimiento, no deben ser incluidos en las secciones específicas de léxico, a no ser que se trate de falsos amigos. En el caso contrario, con la validez de la prueba, es decir, el grado en que una prueba realmente mide lo que dice medir, se va seriamente comprometida. En este sentido, las palabras que aparecieron en las listas de vocabulario como cognados cuyas equivalencias son tan transparentes en inglés fueron excluidas del corpus de este estudio.

A continuación se ejemplificaron algunos de estos cognados excluidos del corpus:
el apartamento, la cafetería, la oficina, la residencia, la universidad, el arte, las comunicaciones, la economía, la filosofía, la historia, las humanidades, la

literatura, las matemáticas, hispano, inteligente, necesario, posible, el sofá, la banana, la barbacoa, el cereal, la ensalada, la fruta, la hamburguesa, el salmón, el sándwich, el tomate, el yogur, el menú, camping, el bar, el antibiótico, la medicina, la temperatura, el blog, el e-mail, el escáner, el fax, la fotocopia, la fotocopidora, el basquetbol, el béisbol, el golf, el hockey, el tenis, el voleibol, honesto/a, irresponsable, responsable, modesto/a, el aniversario, la adopción, la educación, el/la mecánico/a, el/la arquitecto/a, el/la piloto/a, el/la profesor(a), el objetivo, el artículo, el satélite, el teléfono/ celular /móvil, calcular, copiar, escanear, fotocopiar, el bar, el carnaval, la discoteca, etc....

Sin embargo, fuera de estos cognados, en las listas de vocabulario se encontraron más palabras cuyas ortografías son muy parecidas en inglés y español. Estas palabras a causa de su semejanza ortográfica con el inglés podrían ser fácilmente reconocibles por los estudiantes de EL2. Por lo tanto, tomando en cuenta la falta del consenso sobre la inclusión o exclusión de los cognados en las pruebas de conocimiento léxico, se decidió dar una pre-prueba (véase Apéndice A para la prueba de cognados) de 100 *ítems* a los 45 estudiantes (de Span 1501) que estaban en el nivel principiante con el fin de ver si pueden conocer estas palabras con la ayuda de su primera lengua aunque todavía no habían aprendido estas palabras en el salón de la clase. La pre-prueba se les dio a los estudiantes al principio del semestre, tuvieron en total 30 minutos para completarla; se les pidió que escribieran la traducción en inglés de las palabras.

Según los resultados de esta pre-prueba, si 24 alumnos no pudieron dar la respuesta esperada, se decidió incluir el cognado en el corpus de este estudio, considerando el hecho de que estas palabras no son tan reconocibles a pesar de su

similitud ortográfica con el inglés por más de la mitad de los estudiantes. De los 100 ítems se incluyeron solamente los siguientes 12: *carrera, población, estrés, apariencia, mestizo, química, esquí, síntoma, batería, musulmán, informática, artículo*.

Tabla 15. Distribución de las categorías gramáticas según los niveles

	Span 1501	Span 1502	Span 2301	Span 2302	Total
Nombres	165	177	142	87	571
Verbos	82	85	34	31	232
Adjetivos	132	22	23	11	188
Adverbios	26	7	0	0	33
Total	405	291	199	129	1024

En resumen, al excluir todas las unidades léxicas anteriormente detalladas, al final, el corpus de este estudio se redujo a 1024 palabras. En la Tabla 15 se puede ver la distribución de categorías gramaticales según los 4 niveles en el corpus de este estudio. Considerando el hecho de que las fuentes que formaron la base del corpus de este estudio son algunos de los libros de textos más usados por los programas universitarios de EL2 en los Estados Unidos, se puede concluir que este corpus constituye una muestra muy representativa de lo que se enseña en cuanto al vocabulario en los cursos de español como L2 en los programas universitarios en este país.

4.3. Instrumentos

4.3.1. Cuestionario sociolingüístico y de reflexión sobre el vocabulario. El cuestionario se organizó en dos apartados: (1) el cuestionario del entorno sociolingüístico y (2) el cuestionario de reflexión sobre el vocabulario (véase Apéndice B). El objetivo principal del primer apartado de este cuestionario fue recoger datos personales (p.ej., edad, sexo, etc.) y lingüísticos (p.ej., primera lengua, estudio formal de la lengua española en el nivel secundario y universitario, conocimiento de otras lenguas, uso del español afuera del salón de la clase, etc.) de los encuestados.

El propósito del segundo apartado de este cuestionario fue hacer reflexionar a los participantes sobre las estrategias que prefieren usar a la hora de aprender y practicar el vocabulario en EL2. De acuerdo con una escala tipo Likert (1932) a los estudiantes se les pidió que evaluaran una serie de estrategias para aprender (p.ej., usar un diccionario bilingüe/monolingüe, usar el contexto para adivinar el significado, analizar la categoría gramatical, raíz/sufijo o prefijo de una palabra desconocida) y practicar el vocabulario (p.ej., usar *flashcards*, estudiar una palabra con su dibujo, asociar una palabra con una experiencia personal, con sus sinónimos y antónimos, usar nuevas palabras en oraciones, repetir una palabra en voz alta, escribir una palabra repetidamente y escuchar una palabra repetidamente) según su utilidad para sí mismos. Se solicitó a los sujetos que externalizaran su reacción hacia estas estrategias eligiendo uno de los 5 grados (1. *not at all useful*, 2. *slightly useful*, 3. *useful*, 4. *very useful*, 5. *extremely useful*) de la escala presentada. De la misma manera, también se les pidió a los encuestados que evaluaran la presentación del vocabulario en los libros de textos que usan actualmente y que habían usado en los niveles anteriores.

El cuestionario se administró en inglés, en la lengua dominante de los estudiantes, para facilitar al máximo su comprensión y reflexión. Esta herramienta incluye en total 40 preguntas de índole variada, desde preguntas cerradas como preguntas dicotómicas de *sí* o *no*, escalas del tipo de la Likert y de respuestas cortas a preguntas abiertas que requieren profundizar sus preferencias y comportamientos a cerca del aprendizaje y estudio de vocabulario. Todos los participantes respondieron este cuestionario en computadora y se les dio quince minutos para completarlo.

4.3.2. Prueba receptiva de niveles de vocabulario. Después de una revisión exhaustiva de los diferentes tipos de ítems y de las pruebas que se usan para la evaluación del vocabulario en segundas lenguas se decidió diseñar una prueba receptiva de niveles de vocabulario (*Receptive Vocabulary Levels Test*) que se basa en el corpus anteriormente presentado y detallado. Este tipo de prueba se consideró lo más apropiado y eficiente para evaluar el aprendizaje léxico receptivo de los EL2 por varias razones. La prioridad era diseñar una prueba descontextualizada considerando el hecho de que estos tipos de pruebas evalúan el reconocimiento receptivo de una manera más fiel y válida que los otros tipos de pruebas de vocabulario. Laufer et. al. (2004) manifiestan que no queda todo del claro si a través de los ítems contextualizados se comprueba el auténtico conocimiento de la palabra o las destrezas diferenciales que posee el examinado. En este sentido Schmitt (1999), encontró que algunos candidatos sabían los significados de ciertas palabras dentro de un determinado contexto pero no así cuando aparecían dentro de otro, o si las palabras se presentaban aisladamente. Otra ventaja que ofrecen las pruebas descontextualizadas es que existen multitud de contextos educativos en los cuales dichas pruebas constituyen sin duda una herramienta útil para establecer un

diagnóstico sobre la estimación del vocabulario receptivo de los estudiantes. Según Meara (1996), a pesar de que el objetivo final del aprendizaje del vocabulario de una L2 sea su uso comunicativo en contextos reales, las pruebas descontextualizadas pueden constituir una medida eficaz que permite comprobar un conocimiento real y objetivo del significado de las palabras sin ayuda del contexto. Así pues parece evidente que no siempre es conveniente o eficaz el uso de medidas contextualizadas.

La Prueba de Niveles de Vocabulario (The Vocabulary Levels Test) fue diseñado por Paul Nation (2001) a comienzos de los 80, como un instrumento de clase que ayudara a desarrollar un programa de aprendizaje para los alumnos de inglés. Es decir, esta prueba principalmente tiene el propósito de detectar carencias y mejorar los contenidos curriculares en el ámbito léxico. La prueba original consta de cinco partes, que se corresponden con cinco niveles de frecuencia. El primer nivel se refiere a las dos mil palabras más frecuentes, el segundo a las tres mil, el tercero a las cinco mil, el cuarto al nivel universitario (por encima de las cinco mil) y, finalmente, el quinto contiene las 10.000 palabras más frecuentes. Las listas de frecuencia en inglés son tomadas de las listas de Thorndike y Lorge (1944), Kucera y Francis (1967) y General Service List (West, 1953). Las utilizadas en el nivel académico se tomaron de la University Word List (Xue & Nation, 1984). En este sentido, de acuerdo con Nation (2001), cada nivel de frecuencia corresponde a diferentes objetivos en el aprendizaje de vocabulario. Las dos mil y tres mil palabras más frecuentes son las que debe dominar cualquier estudiante que quiera desenvolverse en inglés sin grandes problemas en las situaciones más corrientes; el nivel de las cinco mil palabras representa el tope máximo del vocabulario que puede ser objeto de instrucción en clase; las palabras de nivel universitario son las que sirven para

comprender los libros de texto y los escritos académicos; por último, el grupo de las 10.000 palabras corresponde a las menos frecuentes de la lengua.

Sin embargo, con el fin de mejorar algunos de los inconvenientes que surgieron en el uso de las listas de frecuencia en estudios previos, considerando el hecho de que estas palabras no siempre coinciden con las palabras que se enseñan en el salón de clases y además para poder evaluar el aprendizaje/conocimiento léxico receptivo/productivo de los estudiantes de ESL en el programa donde se realizó este estudio, se decidió diseñar una herramienta siguiendo el modelo de la Prueba Receptiva de Niveles de Vocabulario (véase Apéndice C) de 120 ítems basados en las palabras que se enseñan en este programa. En cuanto a la selección de las palabras, de 1024 palabras que forman el corpus se tomó, al azar, cada octava palabra para tener 120 *items* en total y una representación completa de todas las palabras del corpus.

La prueba combinó el método del emparejamiento con el formato de opción múltiple. Así, en cada ítem se les dieron a los participantes seis palabras en la parte izquierda y tres definiciones en la parte derecha. La tarea que tuvo que realizar el estudiante fue establecer una conexión entre las tres palabras y sus correspondientes definiciones dadas, seleccionando las mejores combinaciones. Los candidatos debieron escribir en cada uno de los tres huecos de la derecha el número de la palabra que da la mejor combinación. Por lo tanto, se trató de una tarea de reconocimiento para la cual el candidato tuvo que poseer un conocimiento del significado básico de la palabra en cuestión. A continuación, se presenta un ejemplo del primer nivel de esta prueba receptiva:

Ejemplo 1: *1. cuarto*

2. edificio

3. extranjero

4. lápiz

5. librería

6. palabra

_____ 4 _____ *sirve para escribir*

_____ 3 _____ *persona que es de otro país*

_____ 5 _____ *lugar donde se venden libros*

Como se pueden observar en el Ejemplo 1, estos ítems se pueden caracterizar por ser selectivos, independientes y descontextualizados. Serían selectivos porque están elaborados a partir de un conjunto de palabras escogidas según un criterio, en el caso de este estudio es el orden de la enseñanza de las palabras (i.e., según los niveles del programa); independientes porque miden el vocabulario de manera aislada sin integrarlo dentro de otras destrezas (p.ej., para la identificación de la definición correspondiente no se requiere ni comprensión de lectura ni conocimiento gramatical). Finalmente, al ser descontextualizados, no se presentaron en un contexto que pueda ayudar con la “predicción” del significado de las palabras. El fin era medir simplemente el conocimiento receptivo de la palabra por parte de los participantes.

De acuerdo con Read (2000, p.172), aquí cabe mencionar algunos aspectos importantes que se consideraron a la hora de diseñar los ítems de esta prueba:

Como el propósito de la prueba es evaluar el conocimiento del significado de la palabra objeto del estudio, las definiciones deberían ser claras, concisas y estar redactadas en un estilo que resulte fácil de entender con el fin de minimizar el tiempo de lectura y aumentar la fiabilidad y validez de la prueba. Además, en cada ítem se presentaron las palabras que pertenecen al mismo nivel que las palabras que se querían analizar. Para redactar las definiciones se utilizaron

principalmente los libros de textos o diccionarios que utilizaron los estudiantes en los niveles correspondientes.

Uno de los desafíos fue que en el diseño de las definiciones no se podría usar unidades léxicas de mayor nivel que las que se evaluaron, es decir, las unidades léxicas que aparecen en las definiciones se eligieron del mismo nivel o de los niveles anteriores. De acuerdo con Nation (2001), es más conveniente ordenar los ítems según su grado de dificultad, por lo tanto, es recomendable empezar por los más sencillos y después ir aumentando el nivel. Sin embargo, durante el mes de diciembre del 2013 se llevó a cabo un estudio piloto de esta investigación a través del cual se observó que los participantes en su gran mayoría respondieron sólo las preguntas de su nivel y dejaron de responder los ítems que pertenecían a los niveles siguientes. Para compensar esta deficiencia del estudio piloto, en este estudio se decidió presentar las preguntas al azar, es decir, las preguntas no se categorizaron según los niveles y los ítems de diferentes niveles se mezclaron mediante el uso del programa RANDOM.ORG (<http://www.random.org/>).

Por otra parte, el hecho de que haya más palabras que definiciones es una estrategia para reducir las posibilidades de acertar la respuesta esperada al azar, pues el candidato podría obtener las respuestas mediante un proceso de eliminación que, en realidad, requiere el conocimiento semántico de todas las palabras presentadas en un ítem.

En resumen, de acuerdo con Nation (2001), una de las ventajas de este tipo de formato es evaluar varias palabras simultáneamente dentro de cada ítem.

Además, todas las palabras presentadas en un ítem deberían pertenecer a la misma categoría gramatical. Es decir, las palabras se agruparon y presentaron según su

Según Casso (2010), este formato resulta novedoso y práctico en el sentido de que requiere escaso tiempo de lectura y, al mismo tiempo, reduce las posibilidades de acertar al azar. Además, entre las ventajas de *la Prueba de niveles de vocabulario* se puede afirmar que es una prueba rápida y práctica, requiere una lectura mínima y es fácil de corregir y de administrar, que permite evaluar un gran número de palabras dentro de un formato muy simple.

Por otro lado, este tipo de tareas de decisión léxica ha sido criticado por evaluar únicamente el conocimiento receptivo, estar descontextualizadas y no analizar la profundidad del conocimiento (Read, 2005). En este sentido, para el investigador es crucial tener en cuenta los propósitos principales del estudio a la hora del diseño y la selección de sus instrumentos. Aquí también cabe mencionar que numerosos investigadores han encontrado correlaciones positivas de este tipo de formato con otros métodos de evaluación menos sencillos, como son las pruebas de opción múltiple para la medición de vocabulario (Meara, 1997), pruebas de evaluación de lectura de comprensión para primeras lenguas (West, Stanovic & Mitchell, 1993) y pruebas de adquisición de segundas lenguas (Kim & Krashen, 1998), entre muchos otros estudios que permiten confirmar la validez del formato.

4.3.3. Prueba productiva de niveles de vocabulario. Según López-Mezquita (2007), una buena prueba del reciente cambio tanto cualitativo como cuantitativo que se ha experimentado en la situación actual en el campo de las pruebas de vocabulario es la combinación de distintos formatos dentro de la misma prueba, que pudiesen aportar información sobre diferentes facetas (p. ej., el número de palabras que el candidato conoce, la profundidad con que las conoce y el grado de la organización del léxico). Éste es uno de los objetivos primordiales de este estudio piloto encaminado hacia la elaboración de una prueba de vocabulario que permite evaluar de manera práctica y fiable el conocimiento léxico tanto receptivo como productivo de los estudiantes de EL2. Por esta razón, con el fin de complementar la parte receptiva, se decidió diseñar una prueba productiva de niveles de vocabulario siguiendo el modelo del *Productive Vocabulary Levels Test*, creado por Laufer y Nation (1999). Es una herramienta que pretende realizar una medida del vocabulario productivo. En realidad, se trata de una versión productiva de *The Vocabulary Levels Test* cuyo formato se usó en la primera parte de este estudio también. De la misma forma que las versiones receptivas, en las versiones originales productivas se utilizan los mismos cinco niveles de frecuencia en inglés. Se diferencia en que, en lugar de emparejar palabras con definiciones, los estudiantes tienen que realizar una tarea productiva completando las palabras. Sin embargo, para cumplir con los requisitos de este estudio, al igual que la sección anterior, esta parte productiva se basa también en el corpus anteriormente mencionado. A continuación, se presentan algunos ejemplos sacados de diferentes niveles de la *Prueba productiva de niveles de vocabulario* (véase el Apéndice D) de este estudio:

Ejemplo 3: Mi padre es un hombre muy interesante. (1^{er} Nivel)

Ejemplo 4: Para llevar una vida sana es importante hacer ejercicios frecuemente. (2^o Nivel)

Ejemplo 5: Él es un chico muy tímido e introvertido; no le gusta hablar con nadie. (3^{er} Nivel)

Ejemplo 6: Es importante botar las botellas plásticas en el contenedor de reciclados. (4^o Nivel)

Como se puede ver en los ejemplos anteriores, la prueba presenta un formato de complementación. A palabra esperada se le elidió la parte final dejando un espacio en blanco. Los participantes tuvieron que completar el espacio con la palabra adecuada. El control de la producción se llevó a cabo facilitando las primeras letras de la palabra esperada, con que sólo era posible completar el espacio en blanco.

En las Tabla 17 se puede observar la distribución de la cantidad de los ítems según los niveles y las categorías gramaticales en la prueba productiva:

Tabla 17. Distribución de categorías gramaticales según los niveles en la prueba productiva

	Span 1501	Span 1502	Span 2301	Span 2302	Ítems por cat. gramatical
Nombres	8	11	8	5	32
Verbos	5	5	2	2	14
Adjetivos	6	2	1	1	10
Adverbios	2	2	0	0	4
Ítems por nivel	21	20	11	8	60

La prueba consta de 60 ítems en total y para formar los ítems de la prueba, de 1024 palabras se eligió, al azar, esta vez cada decimoséptima para tener una representación completa de todas las palabras del corpus. Al igual que la prueba receptiva los ítems son independientes y descontextualizados con el fin de proporcionar un contexto mínimo a la palabra y evitar dar pistas sobre el uso concreto de la misma. A

diferencia de la prueba receptiva, en esta prueba se trata de una tarea de recuperación y no simplemente reconocimiento, ya que los participantes tienen que acceder a la palabra en su memoria y ser capaz de usarla en el contexto que se le da.

De acuerdo con López-Mezquita (2007), los estudios que se han realizado sobre esta prueba afirman que se trata de una prueba fiable, válida y práctica para comprobar el nivel de vocabulario y su desarrollo, y por otra parte supone una medida cuantitativa adicional que permite la investigación de aspectos más importantes en la adquisición del vocabulario.

Por otra parte, Casso (2010) critica la prueba por no reproducir adecuadamente una situación comunicativa real, que refleje la producción oral o escrita porque según él, en un contexto real nadie daría pistas señalando las primeras letras de una palabra. Sin embargo, aquí cabe mencionar el propósito principal de este estudio no era medir la profundidad sino el tamaño léxico estudiado por los estudiantes de EL2. Además, conforme a López-Mezquita (2007), las pruebas que miden la profundidad del vocabulario resultan en general poco prácticas, ya que la mayoría ha de llevarse a cabo mediante entrevistas personales, lo cual ocasiona una muestra pequeña y un número limitado de palabras que se pueden incluir en un estudio. En cuanto al uso de esta herramienta es posible concluir que esta prueba tiene propósito de evaluar de manera fiable, válida y práctica la producción escrita de vocabulario complementando el instrumento receptivo propuesto en la sección anterior.

4.4.4. Procedimiento. Una de las ventajas importantes de estas dos pruebas (receptiva y productiva) de este estudio es que son susceptibles de ser virtualizadas y administradas en línea. Con el fin de aprovechar esta ventaja las dos pruebas junto con el

cuestionario fueron diseñados en *Blackboard*, que es una plataforma de aprendizaje en línea (*e-learning*) usada por la Universidad de Houston.

Durante el mes de diciembre del 2013 se llevó a cabo un estudio piloto de esta investigación. Una de las deficiencias del estudio experimental fue el número limitado de sus participantes y la variedad restringida del nivel de los grupos incluidos. Para compensar estas deficiencias del estudio piloto, este estudio se realizó con 783 participantes en 7 cursos, 37 sesiones diferentes.

En la muestra final quedó un total de 410 estudiantes de 6 niveles/grupos diferentes. Como se mencionó anteriormente, la selección y categorización de los estudiantes según niveles fue hecha mediante el cuestionario sociolingüístico, considerando el nivel de español basado en la cantidad de semestres que había estudiado esta lengua en un contexto académico universitario. La recopilación de los datos de 5 grupos se realizó durante el mes de abril del 2014 y los datos del grupo de pre-prueba que sirven como punto de partida para esta investigación fueron recolectados durante la primera semana del semestre académico de Otoño 2014.

Todos los participantes completaron este estudio en el Centro Universitario de Adquisición de Lenguas (LAC, por su sigla en inglés) en una computadora durante sus horas de clase. Se les dio quince minutos para completar el cuestionario y media hora para cada prueba. En cuanto a la evaluación, la prueba receptiva tuvo en total 120 *ítems* y cada *ítem* valía 1 punto, mientras la prueba productiva tuvo en total 60 *ítems* y cada *ítem* valía 2 puntos. En la segunda prueba si el estudiante no tuvo ningún error se le dio 2 puntos completos y en esta prueba se decidió no dar créditos parciales a las faltas de ortografía por los siguientes motivos metodológicos: 1) mejorar la validez del contenido

capturando más aspectos (p. ej., forma, funciones gramaticales, etc.) del conocimiento productivo del léxico en la evaluación; 2) observar y comparar mejor los resultados de las dos pruebas (i. e., conocimiento receptivo vs. productivo; reconocimiento vs. recuperación de las palabras); 3) hacer la distinción más clara entre diferentes niveles de competencia léxica; 4) evitar inconsistencias y subjetividad en la evaluación de las respuestas dadas (i. e., el desafío de definir qué es una respuesta “cercana” a la respuesta esperada, como por ejemplo, en los casos de *frecuentemente* en vez de *frecuente* o *relajarse* en vez de *relajarse* vs. *cielo* en vez de *ciego* o *material* en vez de *materia*).

Al inicio del estudio, como se puede ver en la Figura 2, las instrucciones del estudio se les explicaron en inglés a todos los participantes a través de una presentación de *PowerPoint* y al final, se les permitió ver los resultados de cada prueba.

Figura 2. Las instrucciones para el proyecto

**INSTRUCTIONS FOR
“PROYECTO DE VOCABULARIO”**

- Login to your **BlackBoard**
- Go to your SPAN course homepage
- Open the file “**Proyecto de Vocabulario**”
- Order of completion:
 1. Consent Letter (5 min.)
 2. Students’ Background Questionnaire (15 min.)
 3. Receptive Vocabulary Levels Test (30 min.)
 4. Productive Vocabulary Levels Test (30 min.)

Todos los participantes dieron su consentimiento por escrito para participar en la investigación; este es un requisito del Comité para la Protección de los Seres Humanos en la Universidad de Houston. Por esta razón, antes de empezar con el estudio, se les pidió a los participantes que completaran el formulario correspondiente (véase el Apéndice E).

En total, se les dio una hora y quince minutos para completar todas las etapas de este estudio. En el análisis de los resultados, a todos los participantes se les asignó un código que indica el grupo al que pertenece el estudiante, sus iniciales y su número de participante (p. ej., G3/CS/15) con el fin de mantener su anonimato.

Capítulo 5

Análisis de Resultados

En este capítulo se presenta el análisis cuantitativo descriptivo de los resultados hallados en este estudio. El mismo se encuentra dividido en seis apartados: en las primeras dos secciones (5.1. y 5.2) se presentan los datos individuales de la prueba receptiva y productiva de niveles de vocabulario respectivamente, para cada una se exponen : (5.1.1 y 5.2.2) el comportamiento global de las pruebas a través de las estadísticas descriptivas; (5.1.2 y 5.2.2) la fiabilidad y validez de la prueba receptiva y productiva; (5.1.3 y 5.2.3) los resultados de análisis de varianza unidireccional (ANOVA) con el fin de observar si los 6 grupos analizados en este estudio difieren significativamente entre sí en cuanto a sus medias y varianzas. En el siguiente apartado (5.3) se analizan los contrastes de medidas entre las dos pruebas de niveles de vocabulario a través de Prueba *t* emparejada (*T-test*) para cada grupo; con el fin de corroborar la covariación, en la cuarta subsección (5.4.1), primero se exponen las Correlaciones de Pearson-r entre las dos pruebas de vocabulario, luego (5.4.2) se presentan las Correlaciones de Pearson-r entre el promedio de las dos pruebas y el promedio total obtenido del curso del español al final del semestre por cada participante; en seguida (5.5) se presentan los resultados del cuestionario de reflexión sobre las estrategias de aprender y practicar el vocabulario; finalmente, en el último apartado de este capítulo (5.6) se realiza un análisis comparativo entre las palabras que forman el corpus de este estudio y *A Frequency Dictionary of Spanish* de Mark Davies (2006).

Los datos obtenidos de 410 alumnos a través de los tres instrumentos: 1) la Prueba Receptiva de Niveles de Vocabulario (PR-NV), 2) la Prueba Productiva de

Niveles de Vocabulario (PP-NV), y 3) el cuestionario de las estrategias de aprendizaje de vocabulario (EAV) fueron analizados utilizando el programa informático SPSS (Paquete Estadístico para las Ciencias Sociales).

5.1. Prueba Receptiva de Niveles de Vocabulario (PR-NV)

5.1.1. Estadísticas Descriptivas de la Prueba Receptiva de Niveles de Vocabulario (PR-NV). En esta sección se les dio a todos los participantes una Prueba Receptiva de Niveles de Vocabulario (PR-NV) que consta de 120 ítems basados en las palabras que se enseñaron en el programa de EL2 en la Universidad de Houston con el fin de evaluar su aprendizaje léxico receptivo. En cada ítem se les dieron a los participantes seis palabras y tres definiciones (ver Apéndice C). Como se puede ver en el siguiente ejemplo que representa uno de los ítems del nivel intermedio-alto, la tarea que tuvo que realizar el estudiante fue establecer una conexión entre las tres palabras y las correspondientes definiciones dadas, seleccionando las mejores combinaciones de forma y significado.

Ejemplo 7:

- | | |
|-------------------------|--|
| 1. <i>analfabetismo</i> | |
| 2. <i>feria</i> | _____ <i>no saber leer ni escribir</i> |
| 3. <i>ley</i> | _____ <i>déficit de lluvia por mucho tiempo</i> |
| 4. <i>orgullo</i> | _____ <i>norma establecida para regular o prohibir</i> |
| 5. <i>sequía</i> | |
| 6. <i>taza</i> | |

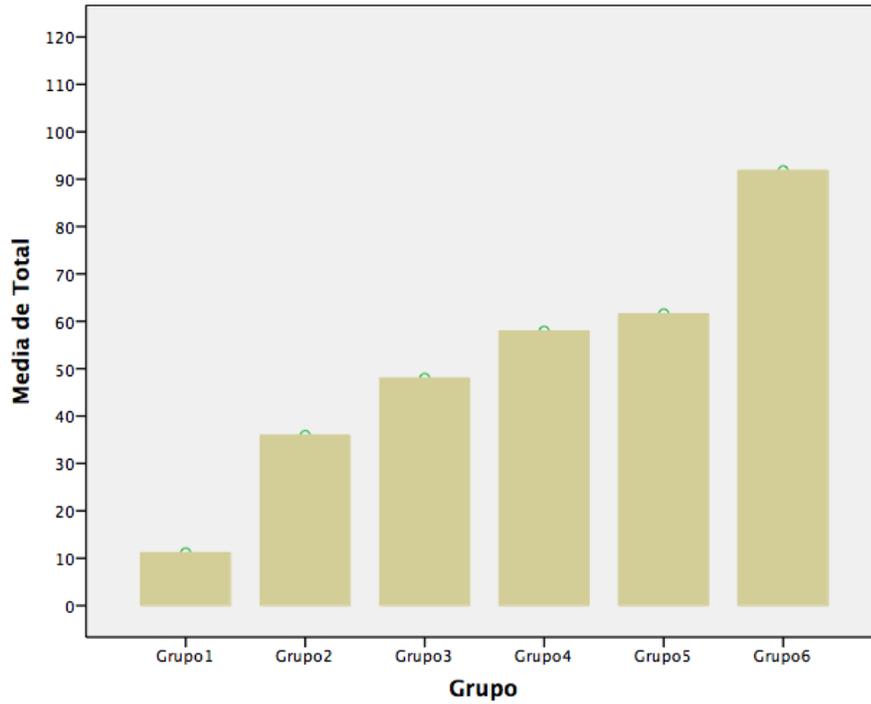
Cada *ítem* valía 1 punto y el puntaje máximo que se podría sacar en la PR-NV era de 120 puntos (100%). Si el estudiante combinaba la palabra correctamente con su definición, se le daba un punto; si no, obtenía un cero. Como se mencionó en el capítulo anterior la prueba se basaba en un corpus de 1024 palabras que se estudiaron en los primeros 4 niveles del programa de EL2 en la UH. Mientras estos 4 grupos

experimentales -1) principiante-bajo, 2) principiante-alto, 3) intermedio-bajo, 4) intermedio-alto- que aparecen como G2, G3, G4, G5 respectivamente en el análisis formaron el enfoque principal de estudio; G1 y G6 formaron los grupos de control (pre y post-evaluación) de este estudio.

Los ítems estaban compuestos de palabras léxicas: sustantivos, verbos, adjetivos y adverbios y dichas categorías estaban presentes en la prueba en una proporción que refleja su porcentaje de aparición en el corpus a partir del cual se ha elaborado. Se esperaba que los alumnos eligieran las palabras esperadas que correspondían al nivel en que se encontraban, además de las palabras correspondientes a los niveles anteriores. En este sentido, la prueba podría servir como una prueba diagnóstica⁹ del conocimiento léxico.

⁹ De acuerdo con el Diccionario de términos clave de ELE, publicado por el Centro Virtual Cervantes, se entiende por prueba diagnóstica la que tiene como finalidad determinar cuáles son los puntos fuertes y los puntos débiles del estudiante que se presenta a la misma, qué puede o no puede hacer con la lengua, y hasta qué punto se desenvuelve en las distintas habilidades. La información proporcionada por esta prueba sirve para tomar decisiones sobre la formación que debe seguir el estudiante, aunque se puede utilizar también para clasificarlo en un grupo de nivel homogéneo, con el fin de que reciba la instrucción adecuada al nivel demostrado.

Figura 3. Distribución de los promedios en la PR-NV por niveles



Según los resultados de la PR-NV, existe un aumento en el reconocimiento léxico por parte de los estudiantes del programa a medida que avanzan de nivel dentro del programa (véase la Figura 3). En la siguiente tabla se muestran en detalle los promedios de los puntajes obtenidos por cada grupo de la PR-NV.

Tabla 18. Promedios de la PR-NV

Nivel/Grupo	<i>N</i>	Promedio de puntajes
1501/G1	51	11.14 (9.3%)
1501/G2	95	35.91 (29.9%)
1502/G3	90	47.97 (39.9%)
2301/G4	66	57.92 (48.3%)
2302/G5	90	61.58 (51.3%)
3000/G6	18	91.78 (76.48%)
Total	410	51.05 (42.54%)

De acuerdo con el promedio de los puntajes del G1 (11.14 puntos), que sirve como punto de partida para esta investigación, los estudiantes principiantes empezaron el programa universitario con cierto conocimiento léxico previo, que era de esperarse considerando el hecho de que 70% de ellos estudiaron este idioma en el nivel secundario. Los resultados muestran que entre el G1 y G2, al completar el primer nivel del programa, se observa un aumento de 20.64% (24.77 puntos) en el léxico receptivo por parte de los estudiantes. Mientras, el menor aumento (3.66 puntos, 3%) se ve entre el G5 y G4 (i.e., el nivel intermedio-bajo e intermedio-alto), el mayor (30.2 puntos, 25.16%) se observa entre el G5 y G6 (i.e., el nivel intermedio-alto y el avanzado). Este último grupo está formado por estudiantes que seleccionan el español como carrera principal o secundaria, y actúan como grupo control de este estudio.

A continuación, se presentan las estadísticas fundamentales de la PR-NV a través de las medidas de tendencia central –la media, los valores mínimos y máximos–, y de variabilidad –rango, desviación estándar y varianza– con el fin de observar el comportamiento global de cada grupo en esta prueba.

Tabla 19. Estadísticas descriptivas de la PR-NV según los grupos

Grupo	N	Rango	Mínimo	Máximo	Media	Desviación estándar	Varianza	Asimetría	Curtosis
G1	51	66	0	66	11.14	12.989	168.721	2.665	8.498
G2	95	92	6	98	35.91	18.051	325.831	.844	1.120
G3	90	94	10	104	47.97	19.157	366.976	.423	.222
G4	66	93	21	114	57.92	21.335	455.179	.603	-.421
G5	90	106	8	114	61.58	24.018	576.876	.022	-.830
G6	18	38	77	115	91.78	10.344	107.007	.740	.038

Según los resultados de la Tabla 19, en general se observa un aumento gradual en los valores de tendencia central por parte de los grupos a medida que avanzan de nivel

dentro del programa. En este sentido, los valores de tendencia central indican, en general, una consistencia interna sobre cómo se agrupan los resultados de esta sección. Sin embargo, mientras todos los grupos lograron aumentar su puntaje mínimo y máximo, no se ve un progreso en estos valores del G5. Como se puede observar en la Tabla 19, la asimetría es positiva para todos los grupos. En otras palabras, la baja frecuencia de puntajes altos indica que esta prueba resulta difícil para todos grupos. Según Bachman (2004), si la asimetría o curtosis tiene un valor absoluto menor de 2 (de -2 a 2) la distribución de los puntajes es “razonablemente normal”. De acuerdo con Bachman (2004), la única excepción se observa en los valores del G1, lo cual señala, como se esperaría, el alto nivel de dificultad de esta prueba para los estudiantes que acababan de entrar al programa.

En cuanto a las medidas de la variabilidad, de acuerdo con la Tabla 19, se ve una subida gradual en los valores de rango, desviación estándar y varianza por parte de todos los grupos excepto el G6. En otras palabras, mientras los estudiantes avanzan de nivel dentro del programa, se observan más desviaciones en los puntajes individuales respecto a la media.

En resumen, estas medidas de tendencia y de dispersión muestran el alto grado de dificultad de esta prueba para los estudiantes principiantes e intermedios del programa. Considerando el hecho de que esta tarea de reconocimiento, para la cual el candidato tenía que poseer un conocimiento del significado básico de las palabras, fue diseñada en base al vocabulario estudiado en el programa, se esperaría haber encontrado medidas de tendencia más altas. En la siguiente tabla se compara el máximo puntaje esperado con los promedios obtenidos de la misma prueba por parte de los 4 grupos.

Tabla 20. Comparación de puntajes esperados con promedios en la PR-NV

Grupo	<i>N</i>	Máximo puntaje esperado	Promedios de PR-NV
G2	95	45 (37.5%)	35.91 (29.9%)
G3	90	81 (67.5%)	47.97 (39.9%)
G4	66	102 (85%)	57.92 (48.3%)
G5	90	120 (100%)	61.58 (51.3%)

Los máximos puntajes esperados se calcularon considerando la distribución de los ítems léxicos incluidos de cada nivel en la PR-NV. Como la adquisición del vocabulario es un proceso acumulativo, se esperaba que los estudiantes retuvieran el vocabulario del nivel anterior a medida que avanzan de nivel dentro del programa. Por la tanto, a los ítems léxicos de un nivel se sumaron los del siguiente para calcular el máximo puntaje esperado en cada nivel. En la siguiente tabla se ilustra en detalle la distribución de los ítems por categoría gramatical y los máximos puntajes esperados según los niveles:

Tabla 21. Distribución de los ítems léxicos según los niveles en la PR-NV

	1501	1502	2301	2302	Ítems por categoría gramatical
Nombres	18	21	15	12	66
Verbos	9	9	3	3	24
Adjetivos	15	3	3	3	24
Adverbios	3	3	0	0	6
Ítems por nivel	45	36	21	18	120 ¹
# Respuestas esperadas	45	45+36=81	81+21=102	102+18=120	120
Máximo puntaje	45	81	102	120	

Nota¹. Cada *ítem* valía 1 punto, por lo tanto, el puntaje máximo que se podría sacar en la PR-NV era de 120 puntos (100%).

Como se puede ver en la Tabla 21, entre el máximo puntaje esperado y el promedio de la PR-NV, hay una diferencia de 7.6 %, 27.6%, 36.7%, 48.7% para el G2, G3, G4, G5 respectivamente. Esta diferencia aumenta a medida que los estudiantes avanzan de nivel dentro del programa. Para explicar este hallazgo se puede argumentar que el papel que ocupa el vocabulario en el currículo va bajando gradualmente a lo largo de los primeros 4 niveles del programa. Eso es especialmente evidente en el corpus formado por los libros de textos usados durante los primeros dos años del programa; se observó que la cantidad de palabras (i. e., la cantidad de horas, el peso que tiene en la evaluación, la cantidad de actividades realizadas en el salón de clase, etc.) que se enseñan en cada nivel baja a medida que los estudiantes avanzan de nivel. Por lo tanto, en comparación con los niveles intermedios se esperaba observar un mayor aumento del conocimiento léxico en los niveles principiantes.

El papel dominante que desempeñan los libros de textos en la enseñanza de EL2 fue también criticado por López-Jiménez (2014):

Today few teachers would seriously question the central role played by L2 textbooks as a major pedagogical resource in teaching a non-native language (L2). Over the last 30 years, L2 textbooks have been characterized by their capacity to adapt to changing situations, such as the introduction of new technologies, new methodological trends in L2 teaching, or the increasingly demanding requirements of the publishing market. This flexibility is largely responsible for the dominant influence that textbooks continue to have in L2 classrooms in general and in classrooms where Spanish is taught as an L2 in particular. While the importance of the textbook as a pedagogical resource in L2 teaching has

remained constant, if not increased, over the last three decades, the language components on which textbook writers concentrate have not always received the same attention in L2 teaching. (p.163-164)

En este sentido, el espacio que ocupa el vocabulario, como un componente del aprendizaje en EL2, por los libros de texto parece ser un factor efectivo en los resultados sacados de esta prueba.

5.1.2. Fiabilidad y validez de la prueba receptiva de niveles de vocabulario

(PR-NV). El grado de fiabilidad de las dos pruebas de vocabulario se analizó a través de diferentes métodos que permiten medir la consistencia interna de cada prueba. A continuación, en la Tabla 22 se presentan los resultados del alfa de Cronbach para la PR-NV.

Tabla 22. Medidas de consistencia interna de fiabilidad para la PR-NV

Alfa de Cronbach	Alfa de Cronbach basada en elementos estandarizados	N de ítems
.967	.967	40

El valor deseable del alfa de Cronbach para una prueba de vocabulario debe estar entre 0.90 a 0.99 (Hughes, 1989; Lado, 1961). En la PR-NV, de 40 ítems, de 120 palabras, administrada a 410 alumnos, se obtuvo un valor de 0.967 para el alfa de Cronbach. Este valor indica el alto grado de fiabilidad de esta prueba. Cabe mencionar aquí que las condiciones relacionadas con el diseño, construcción y administración de la prueba se controlaron estrictamente con objeto de alcanzar la máxima fiabilidad.

Con el fin de comprobar la fiabilidad de la misma prueba se decidió también aplicar otros métodos que consisten en dividirla en dos mitades equivalentes (*split half*

reliability) y calcular la correlación entre los resultados de ambas partes. En la siguiente tabla se pueden observar los resultados obtenidos a través de *Alfa de Cronbach* de dos partes, coeficiente de Spearman-Brown y coeficiente de dos mitades de Guttman.

Tabla 23. Estadísticas de fiabilidad según mitades partidas para PR-NV

Alfa de Cronbach	Parte 1	Valor	.934
		N de ítems	20 ^a
	Parte 2	Valor	.942
		N de ítems	20 ^b
	N total de ítems		40
Correlación entre formularios			.915
Coeficiente de Spearman-Brown	Longitud igual		.956
Coeficiente de dos mitades de Guttman			.955

Nota. a. Los ítems son: ítem1, ítem2, ítem3, ítem4, ítem5, ítem6, ítem7, ítem8, ítem9, ítem10, ítem11, ítem12, ítem13, ítem14, ítem15, ítem16, ítem17, ítem18, ítem19, ítem20.

b. Los ítems son: ítem21, ítem22, ítem23, ítem24, ítem25, ítem26, ítem27, ítem28, ítem29, ítem30, ítem31, ítem32, ítem33, ítem34, ítem35, ítem36, ítem37, ítem38, ítem39, ítem40.

De acuerdo con los resultados de la Tabla 23, la fiabilidad intrínseca de la PR-NV se verificó también a través de *Alfa de Cronbach* de dos partes, coeficiente de Spearman-Brown y coeficiente de dos mitades de Guttman. En otras palabras, se puede concluir, al igual que lo hicieron algunos estudios previos (Beglar & Hunt, 1999; Read, 1998; Schmitt *et. al.*, 1999), que esta prueba cumple los requisitos de una prueba fiable como controlar bien los factores relacionados en elaboración de los ítems, influencia del factor azar, claridad de las instrucciones, relación existente entre la dificultad de la prueba y el nivel de competencia de los participantes, etc.

5.1.3. Análisis de varianza unidireccional de la prueba receptiva de niveles de vocabulario (PR-NV). Previamente se había comprobado el aumento del

conocimiento léxico receptivo por parte de los estudiantes principiantes e intermedios del programa de EL2 en la Universidad de Houston a medida que avanzan de nivel dentro del programa. Con el fin de responder si se puede hablar de diferencias estadísticamente significativas entre los resultados obtenidos para los primeros 4 niveles del programa en la PR-NV se realizó un análisis de varianza unidireccional (*one-way ANOVA*).

Según los resultados del análisis de varianzas unidireccional para muestras independientes en la Tabla 24, se observó que existe una diferencia estadísticamente significativa, $F(5, 404) = 73.47, p = .00$, entre los 6 grupos analizados para la PR-NV.

Tabla 24. Resultados de *ANOVA* unidireccional para la PR-NV

	Suma de cuadrados	Grados de libertad	Medias cuadráticas	Valor <i>F</i>	Sig.
Entre grupos	140459.716	5	28091.943	73.470	.000
Dentro de grupos	154472.774	404	382.358		
Total	294932.490	409			

Una vez que se comprobó que el nivel del estudiante determina de manera significativa la expansión del vocabulario receptivo, se realizaron pruebas posteriores (*Post-hoc Tukey*) de comparaciones múltiples, con todos los pares posibles de los 6 grupos analizados en este estudio (véase la Tabla 25) con el fin de observar entre qué grupos había diferencias.

Tabla 25. Pruebas post hoc de comparaciones múltiples con HSD Tukey para la PR-NV

(I) Grupo	(J) Grupo	Diferencia de medias (I-J)	Error estándar	Sig.	95% de intervalo de confianza	
					Límite inferior	Límite superior
G1	G2	-24.768*	3.394	.000	-34.49	-15.05
	G3	-36.829*	3.427	.000	-46.64	-27.02
	G4	-46.787*	3.646	.000	-57.23	-36.35
	G5	-50.441*	3.427	.000	-60.25	-40.63
	G6	-80.641*	5.361	.000	-95.99	-65.29
G2	G1	24.768*	3.394	.000	15.05	34.49
	G3	-12.061*	2.876	.000	-20.30	-3.83
	G4	-22.019*	3.133	.000	-30.99	-13.05
	G5	-25.673*	2.876	.000	-33.91	-17.44
	G6	-55.873*	5.027	.000	-70.27	-41.48
G3	G1	36.829*	3.427	.000	27.02	46.64
	G2	12.061*	2.876	.000	3.83	20.30
	G4	-9.958	3.169	.022	-19.03	-.88
	G5	-13.611*	2.915	.000	-21.96	-5.26
	G6	-43.811*	5.049	.000	-58.27	-29.35
G4	G1	46.787*	3.646	.000	36.35	57.23
	G2	22.019*	3.133	.000	13.05	30.99
	G3	9.958	3.169	.022	.88	19.03
	G5	-3.654	3.169	.859	-12.73	5.42
	G6	-33.854*	5.200	.000	-48.74	-18.96
G5	G1	50.441*	3.427	.000	40.63	60.25
	G2	25.673*	2.876	.000	17.44	33.91
	G3	13.611*	2.915	.000	5.26	21.96
	G4	3.654	3.169	.859	-5.42	12.73
	G6	-30.200*	5.049	.000	-44.66	-15.74
G6	G1	80.641*	5.361	.000	65.29	95.99
	G2	55.873*	5.027	.000	41.48	70.27
	G3	43.811*	5.049	.000	29.35	58.27
	G4	33.854*	5.200	.000	18.96	48.74
	G5	30.200*	5.049	.000	15.74	44.66

*. La diferencia de medias es significativa en el nivel 0.01.

Los resultados de estas pruebas posteriores mostraron diferencias estadísticamente significativas ($p < .01$) entre todos los grupos, a excepción de entre el G3 y el G4, y entre el G4 y el G5. Es decir, según la PR-NV, no se encontraron diferencias significativas entre los siguientes niveles: principiante-alto e intermedio bajo; intermedio-bajo e intermedio alto en cuanto a su conocimiento léxico. Aunque se notó una diferencia de 9.95 puntos entre los promedios de los niveles intermedio-alto (G5) e intermedio-bajo (G4), la prueba de comparaciones posteriores no estableció una diferencia significativa ($p=.022$) entre estos dos niveles. De hecho, las medias de nivel intermedio-bajo e intermedio-alto fueron muy similares (57.92 puntos para el G4 y 61.58 puntos para el G5). De acuerdo con los resultados de G3, G4, G5, la diferencia entre los promedios de los niveles aumenta a medida que los estudiantes avanzan de nivel dentro del programa. Para explicar este hallazgo se puede argumentar que los estudiantes no tienden a retener a largo plazo el vocabulario de los niveles anteriores. Esto se puede atribuir también al reciclaje limitado del vocabulario previamente estudiado por parte de los materiales usados en el programa. Además, en esta prueba se observó que mientras las palabras de los niveles principiantes dieron porcentajes de exactitud más altos (p. ej. *lápiz, extranjero, librería, quince, amable, alemán, casa, camisa, revista*); las palabras del nivel intermedio dieron porcentajes de exactitud más bajos (p. ej. *exigir, esforzarse, castaño, ondulado, tacaño, cosecha, envase, desechable, marginado, discapacitado*) para todos los grupos. Las palabras con porcentajes más bajos generalmente son las que se introducen en el nivel intermedio y reciclan menos en los materiales por razones de tiempo (i. e., horas de contacto, nivel de dificultad de la presentación de palabras concretas vs. abstractas, cantidad de estructuras gramaticales que deben cubrirse, etc.).

La memorización, revisión y repetición tienen papeles cruciales en el proceso complejo de aprendizaje de un ítem léxico. En este sentido, el reciclaje del vocabulario se convierte en un determinante clave en este proceso. De acuerdo con Nation (2011), la repetición lleva consigo beneficios cuantitativos y cualitativos al aprendizaje léxico:

Repetition is essential for vocabulary learning because there is so much to know about each word that one meeting with it is not sufficient to gain this information, and because vocabulary items must not only be known, they must be known well so that they can be fluently accessed. (p.76-77)

Con respecto a las exposiciones, Saragai, Nation y Meister (1978) estimaron que repetición representa el 20% de los factores involucrados en saber una palabra y en su estudio, concluyeron que los participantes requieren 16 o más repeticiones para agregar una palabra nueva a su lexicón. Kachroo (1962) calculó que siete o más encuentros son necesarios con la misma palabra en libros de texto para que la mayoría de estudiantes puedan aprenderla. De acuerdo con Paribakht and Wesche (1997), la probabilidad de aprender una palabra nueva después de sólo un encuentro es entre 5% y 10% (citado en López-Jiménez, 2014).

5.2. Prueba Productiva de Niveles de Vocabulario (PP-NV)

5.2.1. Estadísticas descriptivas de la prueba productiva de niveles de

vocabulario (PP-NV). Con el fin de complementar la prueba receptiva, se les dio a todos los participantes una Prueba Productiva de Niveles de Vocabulario (PP-NV) que consta de 60 ítems basados en el mismo corpus que la prueba anterior. El corpus fue diseñado utilizando palabras que se enseñaron en los primeros cuatro niveles del programa de EL2 en la UH con el fin de evaluar su aprendizaje de léxico productivo. En

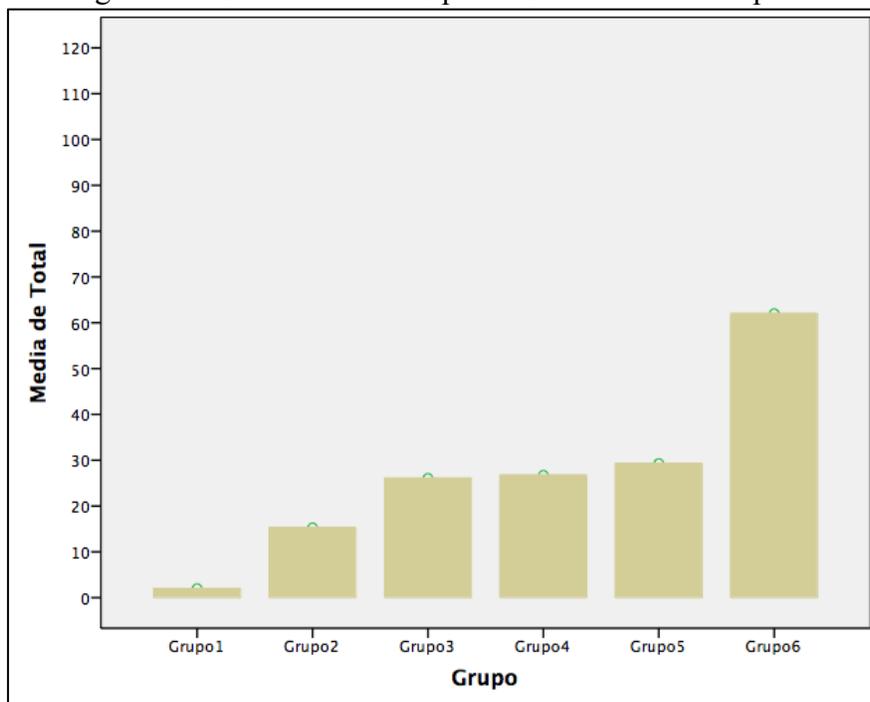
cada oración, el estudiante tenía que agregar las letras que faltaban en la palabra incompleta (Apéndice D). La producción se llevó a cabo facilitando las primeras letras de la palabra como se observa en el Ejemplo 8, extraído de los ítems del nivel principiante-alto:

Ejemplo 8: *La semana pasada mi hermana se rompió el bra___ montando su bicicleta. / La semana pasada mi hermana se rompió el brazo montando su bicicleta.*

Cada *ítem* valía 2 puntos y el puntaje máximo que se podría obtener era de 120 puntos (100%) en la PP-NV. Si el estudiante recuperaba la palabra completamente escribiendo su forma correctamente, se le daba dos puntos; si no, no recibían ningún punto. En otras palabras, las respuestas correctas contaban dos puntos; y las incorrectas y las preguntas sin contestar, cero.

Como se mencionó en el capítulo anterior, en esta prueba se decidió no dar crédito parcial cuando completaban la palabra con faltas de ortografía por los siguientes motivos metodológicos: 1) mejorar la validez del contenido capturando más aspectos (p. ej., forma, funciones gramaticales, etc.) del conocimiento productivo del léxico en la evaluación; 2) observar y comparar mejor los resultados de las dos pruebas (i.e., conocimiento receptivo vs. productivo; reconocimiento vs. recuperación de las palabras); 3) hacer una distinción más clara entre diferentes niveles de competencia léxica; 4) evitar inconsistencias y subjetividad en la evaluación de las respuestas dadas (i.e., el desafío de definir qué es una respuesta “cercana” a la respuesta esperada, como por ejemplo, en los casos de *frecuentemente* en vez de *frecuente* o *relejarse* en vez de *relajarse* vs. *cielo* en vez de *ciego* o *material* en vez de *materia*).

Figura 4. Distribución de los promedios en la PP-NV por niveles



Según los resultados de la PP-NV, al igual que en la prueba anterior existe un aumento en la recuperación léxica por parte de los estudiantes a medida que avanzan de nivel dentro del programa (véase la Figura 4). Sin embargo, en comparación con la PR-NV, se ve que el progreso es menos marcado a medida que aumenta el nivel. Los resultados obtenidos de las dos pruebas se compararán por grupos más adelante, en la sección 5.4, a través de Prueba t bilateral de diferencias.

En la siguiente tabla se presentan en detalle los promedios de los puntajes obtenidos por cada grupo en la PP-NV.

Tabla 26. Promedios de la PP-NV

Nivel/Grupo	N	Promedio de puntajes
1501/G1	51	2.00 (1.66%)
1501/G2	95	15.26 (12.71%)
1502/G3	90	26.07 (21.72%)
2301/G4	66	26.73 (22.27%)
2302/G5	90	29.29 (24.4%)
3000/G6	18	62.00 (51.6%)
Total	410	26.89 (22.41%)

Según los resultados de la Tabla 26, los estudiantes empezaron el programa universitario sin casi ningún tipo de conocimiento léxico productivo. Aunque se observa un aumento de 13.26 puntos (11.05%) y 10.81 puntos (9%) en el aprendizaje del léxico productivo entre los niveles principiantes del programa, la subida de este aspecto del conocimiento léxico no es tan considerable entre los niveles intermedios (i.e., G4 y G5). Probablemente, el hecho de que al vocabulario se le suele dar más peso que a la gramática en los niveles principiantes en comparación a los niveles intermedios podría ser un factor que afecta estos resultados. La mayor importancia del vocabulario es también evidente por la mayor cantidad de las palabras que se presentan en los materiales usados en el programa en los niveles principiantes. En esta prueba de vocabulario, aunque el menor aumento, o sea, 0.66 puntos (0.55%), aparece entre el G3 y G4 (i.e., principiante-alto e intermedio-bajo), al igual que en la prueba anterior, el aumento más considerable, 32.71 puntos, o sea 27.25%, se observa en el G6. Esto se puede atribuir al mayor peso de las estructuras gramaticales y tiempos verbales (p. ej., el imperfecto de indicativo, el presente perfecto de indicativo, el pluscuamperfecto de indicativo y el presente de subjuntivo) que aparecen en el currículo del nivel intermedio-bajo.

A continuación, el comportamiento global de cada grupo en esta prueba se puede observar a través de las medidas de tendencia central y de variabilidad.

Tabla 27. Estadísticas descriptivas de la PP-NV según los grupos

Grupo	N	Rango	Mínimo	Máximo	Media	Desviación estándar	Varianza	Asimetría	Curtosis
G1	51	18	0	18	2.00	3.555	12.640	2.379	7.168
G2	95	56	2	58	15.26	12.103	146.473	1.415	1.768
G3	90	74	2	76	26.07	18.032	325.164	.898	.363
G4	66	88	2	90	26.73	19.281	371.771	1.120	.874
G5	90	92	2	94	29.29	20.744	430.320	.744	-.114
G6	18	66	26	92	62.00	15.537	241.412	-.353	.500

Según los resultados de la Tabla 27, mientras en general se observa un aumento gradual en los puntajes máximos, los puntajes mínimos no varían mucho especialmente en los primeros cinco grupos, lo cual también afecta la subida gradual de las medidas de variabilidad (i.e., rango, desviación estándar y varianza) entre los grupos. Es decir, se ve una mayor dispersión en el aprendizaje del vocabulario productivo entre los estudiantes del programa a medida que se sube de nivel. Como se puede observar en la Tabla 27, la asimetría positiva y la baja frecuencia de puntajes altos en los primeros cinco grupos indican el alto nivel de dificultad de esta prueba para estos grupos. Además, al igual que la prueba anterior, se observó que mientras palabras abstractas de los niveles intermedios como *demora*, *herencia*, *ahorro*, *ascender* aparecieron entre los ítems con porcentajes de exactitud más bajos; las palabras presentadas en los niveles principiantes como *hombre*, *blusa*, *jugar*, *tocar*, *vivir* se encontraban entre los ítems con los porcentajes más altos en todos los grupos en esta prueba.

En resumen, estas medidas de tendencia y de dispersión muestran un alto grado de dificultad de esta prueba productiva para los estudiantes principiantes e intermedios del programa. En la siguiente tabla se compara el máximo puntaje esperado de cada nivel

(considerando la cantidad de los ítems incluidos de cada uno de éstos) con los promedios sacados de la misma prueba por parte de los 4 grupos en la PP-NV.

Tabla 28. Comparación de puntajes esperados con promedios en la PP-NV

Grupo	<i>N</i>	Máximo puntaje esperado (%)	Promedios de PR-NV (%)
G2	95	42 (35%)	15.26 (12.71%)
G3	90	82 (68.3%)	26.07 (21.72%)
G4	66	104 (86.66%)	26.73 (22.27%)
G5	90	120 (100%)	29.29 (24.41%)

Al igual que la prueba anterior, los máximos puntajes esperados se obtuvieron calculando la cantidad de los ítems léxicos incluidos de cada nivel en la PP-NV. En la siguiente tabla se muestra la distribución de los ítems por categoría gramatical y los máximos puntajes esperados según los niveles:

Tabla 29. Distribución de los ítems léxicos según los niveles en la PP-NV

	1501	1502	2301	2302	Ítems por cat. gram.
Nombres	8	11	8	5	32
Verbos	5	5	2	2	14
Adjetivos	6	2	1	1	10
Adverbios	2	2	0	0	4
Ítems por nivel	21	20	11	8	60 ¹
# Respuestas esperadas	21	21+20= 41	41+11= 52	52+8= 60	
Máximo puntaje	42	82	104	120 ²	

Nota¹. Cada *ítem* valía 2 puntos.

Nota². El puntaje máximo que se podría sacar en la PP-NV era de 120 puntos (100%).

A diferencia de la prueba receptiva, en esta prueba ya no se trata simplemente de reconocimiento léxico sino que los participantes tienen que lograr acceso a una determinada palabra en la memoria y ser capaces de producir la parte que falta del ítem en el contexto que se les da. De acuerdo con Schmitt (2010), los resultados de PP-NV deben interpretarse con cautela y los usuarios de esta prueba deben tener claro lo qué mide y lo qué no mide esta prueba. Como señala el autor:

The Productive Vocabulary Levels Test has been promoted as an active/productive test, but according to my proposed terminology, it is closest to a form-recall test, which assesses the form-meaning link. Furthermore, it is not productive in the sense that it does not require examinees to produce lexical items in the course of their spoken or written output. (p. 205).

En este sentido, esta herramienta se comporta como la prueba complementaria de la prueba anterior. Nation (2012) prefiere el término “habilidad productiva controlada” para describir el uso de una palabra cuando así lo soliciten los instructores o investigadores sin importar si se produce en un contexto libre como en la tarea de escribir oraciones, o en un contexto restringido como en el caso de esta prueba donde se proporciona una oración y se espera un ítem léxico determinado (*target word*).

5.2.2. Fiabilidad y validez de la prueba productiva de niveles de vocabulario

(PP-NV). Para la PP-NV de 60 ítems, administrada a 410 alumnos, se calculó el alfa de Cronbach. El resultado fue 0.942 (véase la Tabla 30), lo cual indica que, al igual que la prueba anterior, esta también tiene un alto grado de fiabilidad, ya que hay un 94% probabilidad de que los resultados de esta prueba sean válidos (*true scores*), con un margen de error de sólo un 6%.

Tabla 30. Medidas de consistencia interna de fiabilidad para la prueba productiva

Alfa de Cronbach	Alfa de Cronbach basada en elementos estandarizados	N de ítems
.942	.939	60

Con el fin de comprobar la fiabilidad de esta sección se usaron los mismos métodos de dos mitades que en la sección anterior. En la siguiente tabla se muestran los resultados obtenidos a través de *Alfa de Cronbach* de dos partes, coeficiente de Spearman-Brown y coeficiente de dos mitades de Guttman para la PP-NV.

Tabla 31. Estadísticas de fiabilidad según mitades partidas para la prueba productiva

Alfa de Cronbach	Parte 1	Valor	.879
		N de ítems	30 ^a
	Parte 2	Valor	.902
		N de ítems	30 ^b
	N total de ítems		60
Correlación entre formularios			.874
Coeficiente de Spearman-Brown	Longitud igual		.933
Coeficiente de dos mitades de Guttman			.918

Nota. a. Los ítems son: ciego1, ruido2, cuchillo3, usuario4, rito5, recordar6, brazo7, picante/picoso8, seda9, pulmones10, tímido11, volar12, nivel13, aprobar14, secadora15, gato16, abogado17, dentista18, pobre19, relajarse20, alegre21, botar22, ascender23, sandalias24, armario25, demora26, piso27, materia28, frecuentemente29, tocar30.

b. Los ítems son: ducharse31, preocupado/preparado32, tranquilo33, conocer34, mil35, ser36, mantequilla/manteca37, jugar38, hombre39, herencia40, bajo41, temprano42, azúcar43, planchar44, suegro45, enero46, blusa47, noticias48, tarjeta49, mostrar50, ahorro51, éxito52, postre53, anaranjado54, diecisiete55, receta56, ajedrez57, licenciatura58, vivir59, lucha60.

Como se puede ver en la Tabla 31, se encontraron correlaciones altas entre dos partes/mitades de la PP-NV señalando el grado de precisión global de esta herramienta.

5.2.3. Análisis de varianza unidireccional de la prueba productiva de niveles de vocabulario (PP-NV). Como se había comprobado la existencia de un

aumento en el aprendizaje del conocimiento léxico productivo por parte de los estudiantes principiantes e intermedios del programa de EL2 en la Universidad de Houston a medida que avanzan de nivel dentro del programa, se realizaron las pruebas de análisis de varianza unidireccional para establecer si las diferencias encontradas entre los grupos resultaban ser estadísticamente significativas (véase la Tabla 32). Se encontró que, al igual que la prueba de reconocimiento anterior, el nivel al que pertenece el estudiante tuvo un impacto significativo, $F(5, 404) = 44.76$, $p = .00$, en la PP-NV.

Tabla 32. Resultados de *ANOVA* para la PP-NV

	Suma de cuadrados	Grados de libertad	Media cuadrática	Valor <i>F</i>	Sig.
Entre grupos	60877.775	5	12175.555	44.755	.000
Dentro de grupos	109907.601	404	272.049		
Total	170785.376	409			

Por medio de las pruebas posteriores (*Tukey Post-hoc*) de comparaciones múltiples, se logró establecer entre qué niveles se encontraban las diferencias estadísticamente significativas.

Tabla 33. Pruebas post hoc de comparaciones múltiples con HSD Tukey para la PP-NV

(I) Grupo	(J) Grupo	Diferencia de medias (I-J)	Error estándar	Sig.	95% de intervalo de confianza	
					Límite inferior	Límite superior
G1	G2	-13.263*	2.863	.000	-21.46	-5.06
	G3	-24.067*	2.891	.000	-32.34	-15.79
	G4	-24.727*	3.075	.000	-33.53	-15.92
	G5	-27.289*	2.891	.000	-35.57	-19.01
	G6	-60.000*	4.522	.000	-72.95	-47.05
G2	G1	13.263*	2.863	.000	5.06	21.46
	G3	-10.804*	2.426	.000	-17.75	-3.86
	G4	-11.464*	2.643	.000	-19.03	-3.90
	G5	-14.026*	2.426	.000	-20.97	-7.08
	G6	-46.737*	4.240	.000	-58.88	-34.60
G3	G1	24.067*	2.891	.000	15.79	32.34
	G2	10.804*	2.426	.000	3.86	17.75
	G4	-.661	2.673	1.000	-8.31	6.99
	G5	-3.222	2.459	.779	-10.26	3.82
	G6	-35.933*	4.259	.000	-48.13	-23.74
G4	G1	24.727*	3.075	.000	15.92	33.53
	G2	11.464*	2.643	.000	3.90	19.03
	G3	.661	2.673	1.000	-6.99	8.31
	G5	-2.562	2.673	.931	-10.22	5.09
	G6	-35.273*	4.386	.000	-47.83	-22.71
G5	G1	27.289*	2.891	.000	19.01	35.57
	G2	14.026*	2.426	.000	7.08	20.97
	G3	3.222	2.459	.779	-3.82	10.26
	G4	2.562	2.673	.931	-5.09	10.22
	G6	-32.711*	4.259	.000	-44.91	-20.52
G6	G1	60.000*	4.522	.000	47.05	72.95
	G2	46.737*	4.240	.000	34.60	58.88
	G3	35.933*	4.259	.000	23.74	48.13
	G4	35.273*	4.386	.000	22.71	47.83
	G5	32.711*	4.259	.000	20.52	44.91

*. La diferencia de medias es significativa en el nivel 0.01.

En la Tabla 33 se muestran los resultados de dichas pruebas para la PP-NV por grupos. Según los resultados de la Tabla 33, al igual que la prueba anterior, mientras los niveles principiantes y el nivel avanzado difieren entre sí de manera estadísticamente significativa ($p < .01$) en cuanto a su conocimiento léxico productivo, no se encontraron diferencias significativas entre los niveles intermedios. En resumen, aunque los estudiantes de EL2 demostraron un aumento numérico constante durante los primeros cuatro niveles del programa, al igual que la prueba anterior, el impacto de este aumento no es significativo para diferenciar los niveles principiante-alto, intermedio-bajo e intermedio-alto entre sí en cuanto a su habilidad productiva controlada.

5.3. Contrastes de medias entre las dos pruebas de niveles de vocabulario a través de Prueba *t* bilateral

Para corroborar la diferencia que proponen numerosos investigadores (Ellis & Beaton, 1993; Fan, 2000; Laufer, 1998; Laufer et al., 2004; Nation & Waring, 1997; Shintani, 2011; Stoddard, 1929;) entre el conocimiento receptivo y productivo del vocabulario de los estudiantes de L2, se realizaron las pruebas *t* correspondientes entre los resultados obtenidos de la prueba receptiva (PR) y productiva (PP) de niveles de vocabulario (NV) por cada grupo.

Para determinar si los 6 grupos difieren significativamente en sus resultados de dos pruebas se realizaron las pruebas *t* correspondientes. De acuerdo con los resultados de la prueba *t* para muestras independientes, se observaron diferencias estadísticamente significativas entre el conocimiento del vocabulario receptivo y productivo de cada grupo en la PR-NV y PP-NV. Los resultados de la prueba *t* fueron $t(50) = 5.345$, $p = .00$; $t(94) =$

15.417, $p = .00$; $t(89) = 18.540$, $p = .00$; $t(65) = 18.619$, $p = .00$; $t(89) = 20.657$, $p = .00$; $t(17) = 13.712$, $p = .00$ respectivamente para los 6 grupos.

Además, tal como encontró Laufer (1998) en su estudio, se observó que cuanto más alto sea el vocabulario receptivo de uno, más amplia será la diferencia entre el tamaño del vocabulario receptivo y el vocabulario activo controlado. Esto también explica por qué se observan diferencias cada vez más altas entre los dos tipos del conocimiento léxico a medida que los estudiantes avanzan de nivel dentro del programa.

En síntesis, se puede concluir que, como se esperaba, se observaron diferencias estadísticamente significativas entre el conocimiento del vocabulario receptivo y productivo de los estudiantes principiantes e intermedios del programa de EL2 en la PR-NV y PP-NV. El vocabulario receptivo de los participantes fue significativamente más amplio que su vocabulario activo controlado en este estudio. La explicación que proponen Nation (2013) y Laufer et al. (2004) para explicar esta diferencia apoya estos resultados encontrados en este estudio. Parece que, en general, el conocimiento receptivo de una palabra es más fácil y anterior al conocimiento productivo. La dificultad que entraña el uso productivo supone que el aprendizaje exige más tiempo y esfuerzo (p. ej., la cantidad de conocimiento, la práctica, el acceso y la motivación) que el aprendizaje receptivo de una palabra. Además, el uso productivo de una palabra requiere el aprendizaje de nuevas formas habladas y escritas, mientras que para el uso receptivo de la palabra el estudiante sólo tiene que conocer unos cuantos rasgos distintivos de la palabra; para el uso productivo el conocimiento tiene que ser más preciso. En resumen, tal y como se desprende de los resultados de este estudio, el conocimiento receptivo supera al

productivo y como consecuencia el vocabulario receptivo se considera de mayor tamaño que el productivo, tal como ocurrió en este estudio.

5.4. Correlaciones

5.4.1. Correlaciones de Pearson entre la prueba productiva (PP) y receptiva (PR) de niveles de vocabulario (NV). En esta sección se realiza un análisis de correlación entre los resultados de las dos pruebas. Los resultados del coeficiente de correlación Pearson entre la PR y PP reveló correlaciones para todos los grupos, a excepción del G1, $r(49) = .350$, $p > .01$, como se esperaba. Como se mencionó anteriormente, este grupo formó el punto de partida del estudio y los participantes de este grupo estaban al principio de sus estudios universitarios en EL2, por lo tanto, no se esperaba encontrar correlaciones en los resultados de G1. Por otra parte, se hallaron fuertes correlaciones entre la PP-NV y la PR-NV en el nivel principiante-bajo, $r(93) = .691$, $p < .01$; principiante-alto, $r(88) = .820$, $p < .01$; intermedio-bajo $r(64) = .780$, $p < .01$; intermedio-alto $r(88) = .790$, $p < .01$ y avanzado $r(16) = .820$, $p < .01$ del programa. En resumen se puede concluir que el reconocimiento léxico de los estudiantes de EL2 en los niveles principiantes, intermedios y avanzados se correlaciona de manera positiva con su conocimiento de léxico productivo controlado en este estudio. En este sentido, estas dos pruebas se pueden usar como herramientas complementarias para evaluar el progreso de los estudiantes de EL2 en cuanto a su conocimiento léxico.

5.4.2. Correlaciones de Pearson entre los promedios de las dos pruebas y las notas finales (NF). En este apartado se intentó ver si existía o no una correlación entre los resultados de los promedios de las dos pruebas y el promedio total obtenido del curso del español en que los estudiantes estaban inscritos. Los promedios de las dos

pruebas fueron correlacionados a través de coeficiente de *Pearson- r* con el promedio total obtenido del curso del español al final del semestre por cada participante con el fin de corroborar si existía una covariación entre el conocimiento léxico de los estudiantes y su competencia lingüística general en la lengua meta como proponen numerosos investigadores (Albrechtsen, Haastrup, & Henriksen, 2008; Alderson 2005; Fairclough, 2011; Laufer, 1992; Laufer, 1997; Laufer & Goldstein, 2004; Saviile-Troike, 1984). En este sentido, la nota obtenida al final por el grupo experimental se consideró como predictor del rendimiento académico en un programa de EL2 por el hecho de que esta nota refleja el promedio de diferentes aspectos lingüísticos¹⁰.

De acuerdo con los resultados, mientras se encontraron correlaciones estadísticamente significativas entre los promedios de las dos pruebas de vocabulario y las notas finales para los siguientes grupos: $r(49) = .385$, $p = .01$ para el G1; $r(93) = .635$, $p = .00$ para el G2; $r(88) = .635$, $p = .00$ para el G3; $r(64) = .439$, $p = .00$ para el G4 y $r(88) = .659$, $p = .00$ para el G5 respectivamente), la correlación no fue significativa, $r(16) = .387$, $p = .11$ para el G6. Cabe mencionar aquí que, las notas finales de los estudiantes de G6 reflejaron la suma ponderada del curso avanzado en que cursaban en el momento de la recopilación de los datos. Como se mencionó en el capítulo anterior, los datos de este grupo control fueron recopilados de los estudiantes que estaban tomando los siguientes cursos avanzados: SPAN 3301 *Spanish Oral Communication for Critical Thinking*, SPAN 3305 *Spanish Grammar Review* y SPAN 3306 *Introduction to the Study of Spanish Language*. Estos cursos, en realidad fueron diseñados para mejorar aspectos

¹⁰ Esta nota final fue el promedio calculado a través del siguiente sistema de puntuación: 8% de entrevista oral, 12% de actividades de escritura, 40% de 4 pruebas (que constan de diferentes secciones como comprensión auditiva, vocabulario, gramática, comprensión lectora y expresión escrita), 15% de examen final oral, 10% de participación, 15% de tareas en línea.

lingüísticos más específicos de los estudiantes, particularmente, se enfocan en contenidos específicos, lo cual exige que el estudiante adquiriera un vocabulario especializado según el curso. En vez de fomentar su competencia lingüística general como en los niveles más bajos, el énfasis cae en la terminología nueva que se aprende en el curso y que contribuye a determinar la nota final de la clase. Por lo tanto, no se esperaba encontrar una correlación entre el conocimiento léxico de los estudiantes en el nivel avanzado (i.e., G6) y sus notas finales.

Por otra parte, parece que en los niveles principiantes e intermedios del programa, las dos pruebas de niveles de vocabulario dieron correlaciones significativas con las notas finales y podrían funcionar como predictor del rendimiento académico de EL2. Por lo cual, se puede concluir que estas dos pruebas servirían como instrumentos con propósito diagnóstico para ubicar a los estudiantes en diferentes niveles del programa. Además, se podrían usar para detectar los puntos fuertes y débiles del conocimiento léxico de los estudiantes que están en los primeros dos años del programa. Si estas dos pruebas se aplican a los estudiantes del programa periódicamente, por ejemplo, al final de cada nivel, sería posible observar no simplemente el progreso del conocimiento léxico de los estudiantes individualmente, sino también permitiría obtener información acerca de la enseñanza del vocabulario en el programa. En este sentido, podría funcionar como una herramienta para la evaluación del programa y los materiales usados con respecto a la enseñanza del componente léxico.

5.5. Análisis del cuestionario de reflexión sobre el vocabulario

En este apartado del Capítulo 5 se presentan los resultados del cuestionario de reflexión sobre las estrategias de aprendizaje y práctica del vocabulario. Esta herramienta

fue diseñada con el fin de recoger la postura y las preferencias de los estudiantes respecto al aprendizaje del vocabulario en los cursos de EL2. Cabe destacar la gran variedad de los tipos de preguntas que recogió el cuestionario, se intentó incluir ítems de índole variada, desde preguntas de elección múltiple y utilizando la escala de *Likert*, a preguntas abiertas (véase el Apéndice B, preguntas 17-40).

En la primera pregunta de este cuestionario (i.e., *ítem 17*) se les pidió a los participantes que mencionaran las estrategias que consideraban útiles para su propio aprendizaje del vocabulario. Entre las respuestas más comunes (80%, N=148) de los estudiantes del nivel principiante (G2/ N=95 y G3/ N=90) se encontraron: *usar las palabras en la lengua hablada (con los hablantes nativos), hacer/usar las tarjetas de practica (flash cards), repeticiones en voz alta y escribir las palabras nuevas repetidamente construyendo oraciones con ellas*. A continuación se ejemplifican algunas respuestas dadas a esta pregunta por parte de estudiantes en diferentes niveles:

Ejemplo 9: *Writing it repeatedly, vocabulary cards, conversing with someone who is a native speaker. (G2/MC/17)*

Ejemplo 10: *Creating and using flashcards of all new vocab is very helpful to me. Repeating the out loud and practicing forming sentences out loud has proven beneficial to me. (G2/RG/40)*

Ejemplo 11: *One method is continuous practice of using the vocabulary words in sentences via written activities and oral conversations. (G3/WB/14)*

En cuanto a las estrategias preferidas (61%, N=93) por parte de los participantes del nivel intermedio (G4/ N=66 y G5/ N=90), *crear/usar las tarjetas de práctica, repeticiones tanto orales como visuales* siguieron siendo las estrategias más populares:

Ejemplo 12: *I have found it easiest to learn vocabulary simply through repetition. The same way you would do with spelling in English. I will write the word on a flashcard with the English translation on the back. I then quiz myself using the English word and I try to guess the correct Spanish word on the back. I noticed when I do this I remember them better. (G5/ME/13)*

Ejemplo 13: *I felt using the phrases in a conversational setting (no matter how discomfoting,) was one of the strongest tools for learning Spanish. I studied the most when I knew I would be speaking. (G4/BZ/10)*

De acuerdo con O'Malley y Chamot (1990), la repetición escrita y oral son una de las estrategias más comunes en muchas partes del mundo. Son estrategias tan arraigadas que los estudiantes se resisten a renunciar a ellas y probar otras estrategias.

Probablemente por su practicidad y eficacia son altamente preferidas por los estudiantes que están en los niveles más iniciales de L2. De manera similar, según Brown (1980):

Word lists and flash cards can be used for the initial exposure to a word, but most students continue to use them to review it afterwards. One main advantage of flash cards is that they can be taken almost anywhere and studied when one has a free moment (Brown, 1980). Another is that they can be arranged to create logical groupings of the target words (Gairns & Redman, 1986; Cohen, 1990).

En el nivel intermedio también hubo un 26% ($N=40$) que prefería aprender a través de interacción como por ejemplo, *actividades en parejas/grupos, preguntar al instructor o a un hablante nativo el significado de las palabras, etc.:*

Ejemplo 14: *Group assignments are helpful in learning vocabulary. Also slow repetitive reading and remembering of the words helped me remember more vocabulary.* (G4/CJ/14)

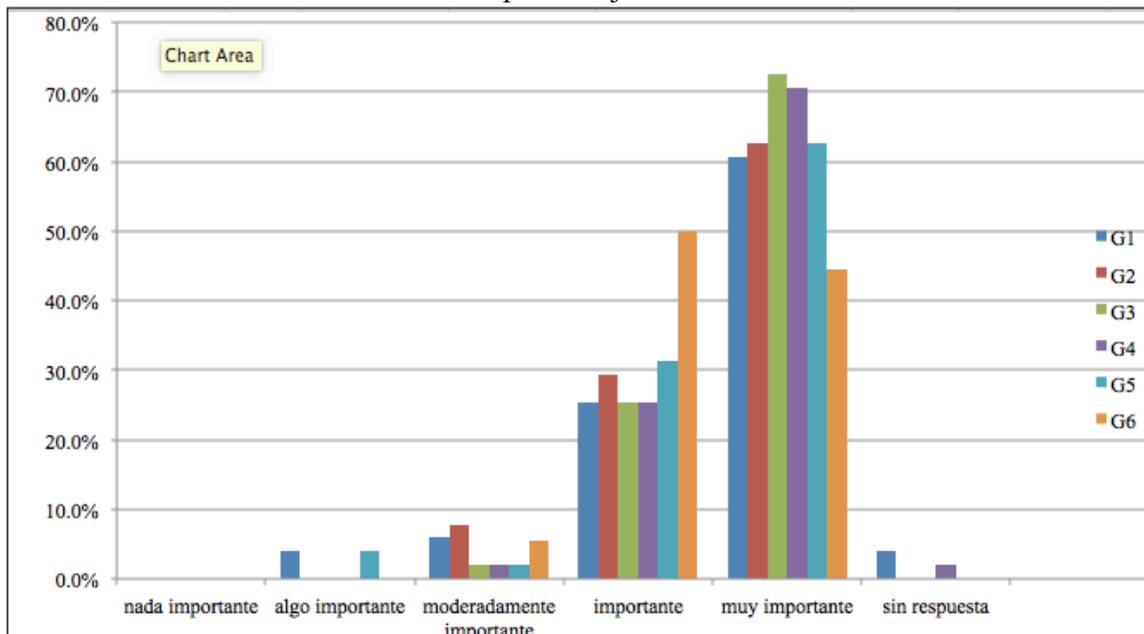
Ejemplo 15: *Having discussions in Spanish in class and having games, like having teams that race to think of similar vocabulary words and such.*
(G5/EA/78)

Ejemplo 16: *I personally liked being able to speak out loud with other students and drilling routines or activities to learn vocabulary.* (G5/BA/07)

Los demás del nivel intermedio (13%/N=20) preferían usar estrategias como *ver la tele/películas, escuchar música/la radio, practicar con materiales en línea del curso*, etc.

En el siguiente ítem de esta encuesta se les preguntó a los participantes qué tan importante resultaba ser el vocabulario en el aprendizaje del EL2 en una escala tipo *Likert* de 5 niveles: (1) “nada importante”, (2) “algo importante”, (3) “moderadamente importante”, (4) “importante”, (5) “muy importante”. En la Figura 5 se presenta la distribución de las respuestas dadas por los grupos a este ítem del cuestionario.

Figura 5. Distribución de las respuestas dadas por grupos al ítem 18: ¿Qué tan importante es el vocabulario en el aprendizaje del vocabulario en EL2?



De acuerdo con los resultados de la Figura 5, de 410 estudiantes en total, no hay ningún estudiante (0%, $N=0$) que encontrara el vocabulario “nada importante” y sólo el 0.9% ($N=4$) lo encontró algo importante. La mayoría de los estudiantes encontraron el vocabulario “importante” (19%, $N=79$) y “muy importante” (43%, $N=176$). También se observa que a medida que los estudiantes avanzan de nivel dentro del programa, aumenta la importancia del vocabulario en el aprendizaje del EL2 para ellos.

Los ítems entre el 21 y el 25 fueron diseñados para la evaluación del libro de texto¹¹ que estaban usando los participantes en el momento de la recopilación de los datos. Se les presentaron 5 declaraciones acerca de la presentación del vocabulario en su libro de texto preguntándoles en qué medida estaban de acuerdo o en desacuerdo con

¹¹ En los ítems 19 y 20 se les preguntó a los participantes que especificaran el nombre de los libros de textos que estaban usando/habían usado en el nivel principiante e intermedio del programa. Estos ítems se usaron principalmente para la eliminación de los participantes que no habían usado los libros de texto, *Puntos de Partida (9 Ed.)* en el primer año y *Más (2 Ed.)* en el segundo año del programa.

éstas. Las declaraciones se presentaron en una escala de valoración de tipo Likert: (1) totalmente en desacuerdo, (2) desacuerdo, (3) neutral, (4) de acuerdo, (5) muy de acuerdo y los participantes eligieron la respuesta que mejor representó su opinión o valoración, dentro de esta escala preestablecida.

Los resultados de las respuestas dadas a estas afirmaciones por los grupos experimentales (i.e., G2, G3, G4 y G5) de este estudio se presentan en la Tabla 34. De acuerdo con los resultados del ítem 21 (*Vocabulary presented in this textbook is useful for me to communicate with native speakers of Spanish in real -life situations*) reflejaron que un 81% de los estudiantes en el nivel principiante y un 65% de los alumnos en el nivel intermedio encontraron el vocabulario presentado en su libro de texto útil para comunicarse con los hablantes nativos en la vida real. Según los resultados del ítem 22 (*The vocabulary relates to engaging topics*), un 75.5 % de los estudiantes principiantes y un 68% de los que estaban en el nivel intermedio pensaban que el vocabulario presentado en el libro está relacionado con temas atractivos. En cuanto a los resultados del ítem 23 (*The vocabulary is presented in a meaningful context.*), mientras un 77.2% de los participantes que estaban en el primer año del programa creían que el vocabulario estaba presentado en un contexto significativo, el porcentaje resultó menor (un 60%) de los que estaban en segundo año del programa. Los resultados del ítem 24 (*The textbook provides engaging vocabulary activities.*) del cuestionario reflejó que para un 61% de los participantes principiantes y un 42% de los participantes intermedios del programa, el libro de texto ofrece actividades interesantes. Los resultados obtenidos a través del ítem 25 (*The vocabulary is recycled through the chapters so I can review the words that I learnt in the previous chapters.*) del cuestionario mostraron que mientras un 59% de los

alumnos principiantes estaban de acuerdo en que el vocabulario se recicla a lo largo de los capítulos del libro del curso, en el nivel intermedio un 46% pensaban lo mismo.

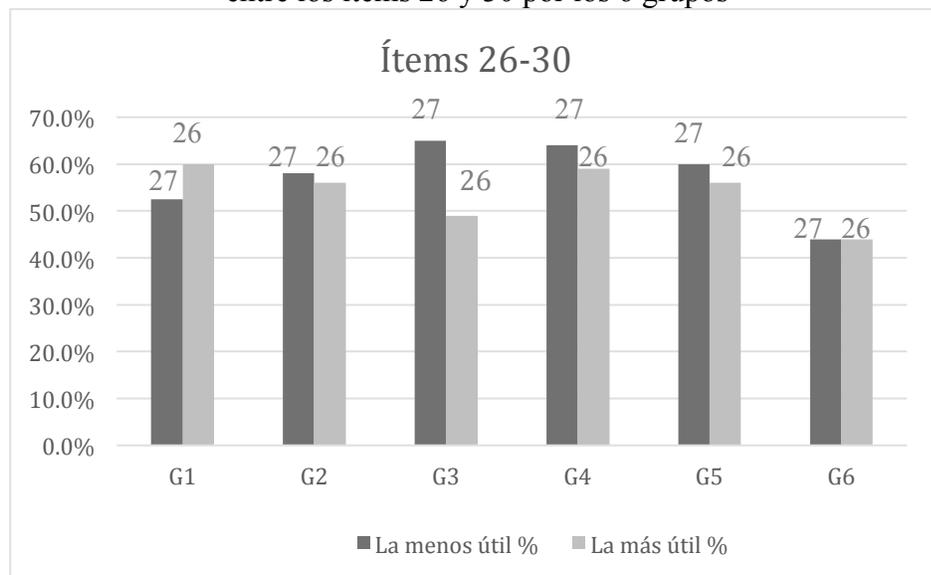
Tabla 34. Resultados de los *ítems* 21-25 sobre la presentación del vocabulario en el libro de texto por el Grupo 2, 3, 4, 5

		de acuerdo & muy de acuerdo	desacuerdo & totalmente en desacuerdo
	Grupo/N	N (%)	N (%)
<i>ítem</i> 21	G2 (95)	79 (83%)	4 (4.2%)
	G3 (90)	71 (78.8%)	3 (3.33%)
	G4 (66)	48 (72.72%)	6 (9.1%)
	G5 (90)	54 (60%)	12 (13.3%)
<i>ítem</i> 22	Grupo/N		
	G2 (95)	73 (76.84%)	3 (3.15%)
	G3 (90)	65 (72.22%)	3 (3.33%)
	G4 (66)	43 (65.15%)	6 (9.1%)
<i>ítem</i> 23	Grupo/N		
	G2 (95)	76 (80%)	4 (4.21%)
	G3 (90)	67 (74.44%)	4 (4.44%)
	G4 (66)	35 (53.03%)	14 (21.21%)
<i>ítem</i> 24	Grupo/N		
	G2 (95)	68 (71.57%)	5 (5.26%)
	G3 (90)	45 (50%)	12 (13.33%)
	G4 (66)	27 (40.9%)	17 (25.75%)
<i>ítem</i> 25	Grupo/N		
	G2 (95)	65 (68.42%)	6 (6.3%)
	G3 (90)	45 (50%)	11 (12.22%)
	G4 (66)	29 (43.93%)	20 (30.3%)
	G5 (90)	43 (47.77%)	23 (25.5%)

Entre los *ítems* 26 al 30 del cuestionario, se les presentaron a los participantes las siguientes estrategias para aprender vocabulario desconocido: (26) *usar un diccionario*

bilingüe, (27) *usar un diccionario monolingüe*, (28) *deducir el significado de una palabra por su contexto*, (29) *analizar la categoría gramatical de una palabra desconocida*, (30) *analizar los afijos y la raíz de una palabra desconocida*. A los encuestados se les pidió que eligieran el nivel de utilidad de estas estrategias para ellos en una escala de valoración de tipo Likert: (1) nada útil, (2) algo útil, (3) útil, (4) muy útil, (5) extremadamente útil. En la Figura 6 se ilustran las estrategias seleccionadas más y menos útiles entre los ítems 26 y 30 por los 6 grupos.

Figura 6. Las estrategias para aprender el vocabulario seleccionadas más y menos útiles entre los ítems 26 y 30 por los 6 grupos



De acuerdo con las estadísticas de la Figura 6, se observa una consistencia en las respuestas dadas por parte de cada grupo. Mientras la estrategia presentada en el ítem 27, *usar un diccionario monolingüe*, fue valorada como la estrategia menos útil, la estrategia presentada en el ítem 26, *usar un diccionario bilingüe*, fue evaluada como la estrategia más útil de aprender el vocabulario desconocido por parte de todos grupos. Según los

resultados del grupo experimental, mientras un 62% de los alumnos en el nivel principiante (i.e., G2 y G3, $N= 185$ en total) e intermedio (i.e., G4 y G5, $N=156$ en total) del programa encontró *usar un diccionario bilingüe* nada útil o algo útil, un 52% de los estudiantes principiantes y un 57% de los intermedios encontró *usar un diccionario monolingüe* muy útil o extremadamente útil. Estos resultados se pueden atribuir al hecho de que los diccionarios bilingües resultan prácticos y adecuados para los estudiantes que están en los niveles iniciales, ya que es muy fácil y rápido descubrir el significado de una palabra. Sin embargo, los problemas son que a veces la palabra tiene varios significados y los estudiantes no saben qué significado sería el más apropiado al contexto en que necesitan usarla y también este tipo de diccionario anima a los estudiantes a traducir más que a pensar en español. Por otra parte, los diccionarios monolingües son preferidos por estudiantes más avanzados (Azorín Fernández & Climent de Benito (2005); Rushstaller, 2004; Schmitt, 1997) ya que les anima a pensar en español y ayuda a practicar la comprensión escrita cuando buscan el significado de las palabras.

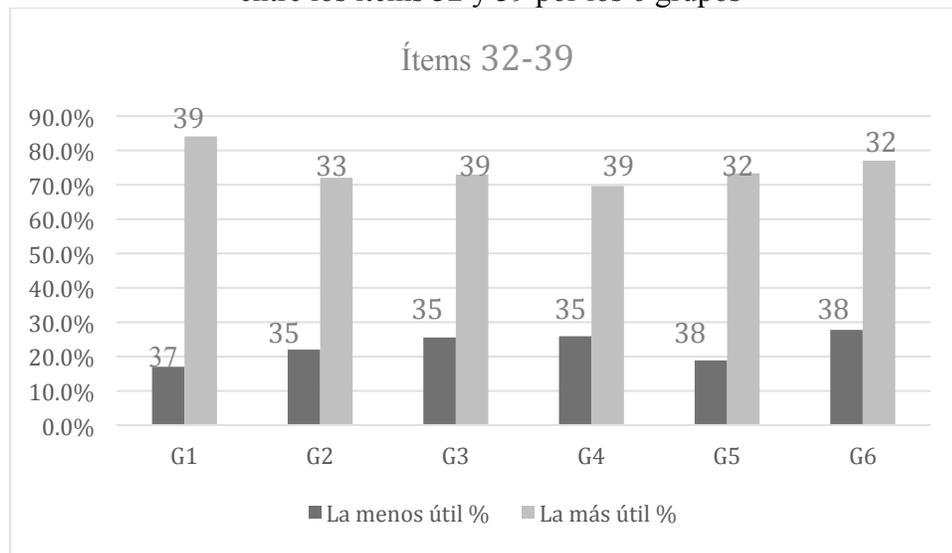
Los ítems 31 y 40 fueron preguntas abiertas donde a los participantes se les pidió que especificaran qué otras estrategias consideran útiles para aprender/practicar el vocabulario, si usan estrategias diferentes de las que fueron presentadas en el cuestionario. Entre las respuestas dadas para los ítems 31 y 40, no se observaron respuestas diferentes de las que se dieron al ítem 18 del cuestionario.

Entre los ítems 32-39 se les presentaron a los participantes las siguientes estrategias para practicar el vocabulario conocido: (32) *usar las tarjetas de practica (flash cards)*, (33) *estudiar una palabra con su imagen en vez de su significado*, (34) *relacionar una palabra con una experiencia personal*, (35) *relacionar una palabra con*

sus sinónimos y antónimos, (36) repetir la palabra en voz alta, (37) escribir una palabra repetidamente, (38) escuchar una palabra repetidamente, (39) usar una palabra en oraciones. Al igual que los ítems anteriores, a los encuestados se les pidió que eligieran el nivel de utilidad de estas estrategias para ellos en la misma escala de utilidad anterior.

En la siguiente figura se muestran las estrategias seleccionadas más y menos útiles entre los ítems 32 y 39 según los 6 grupos.

Figura 7. Las estrategias para practicar el vocabulario seleccionadas más y menos útiles entre los ítems 32 y 39 por los 6 grupos



Según los resultados de la Figura 7, en el nivel principiante bajo (22%, $N=21$), principiante alto (26%, $N=23$) e intermedio bajo (26%, $N=17$) del grupo experimental, el ítem 35, *relacionar una palabra con sus sinónimos y antónimos* fue valorada como la estrategia menos útil. Este resultado se puede atribuir al hecho de que las asociaciones de palabras dependen del significado de las palabras y este tipo de conexiones, las redes de asociaciones que se establecen entre una palabra y las demás palabras de la lengua,

requiere un conocimiento léxico más profundo (Schmitt, 2010). Por otra parte, un 72% ($N=68$) de los estudiantes en el nivel principiante-bajo (G2) encontró, el ítem 33, la estrategia de *estudiar una palabra con su imagen* muy útil o extremadamente útil. La utilidad de esta estrategia tiene sentido para los principiantes dado que los libros de textos presentan palabras más concretas y visuales en estos niveles, mientras que en los niveles más avanzados aumenta la cantidad de las palabras más abstractas. *Usar una palabra en oraciones*, el ítem 39, fue valorado como la estrategia más útil por parte de los participantes en el nivel principiante-alto (73%, $N=66$) e intermedio-bajo (70%, $N=46$). Este resultado marca la importancia del contexto en el aprendizaje del vocabulario para los niveles principiantes e intermedios.

Además, se observó una similitud entre las evaluaciones del G5 y G6 en cuanto a la utilidad de las estrategias para practicar el vocabulario. Parece que mientras que el ítem 38, la estrategia de *escuchar una palabra repetidamente*, fue la estrategia menos útil para los estudiantes que estaban en el nivel intermedio-alto (19%, $N=17$) y avanzado (28%, $N=5$), el ítem 32, *usar las tarjetas de práctica* fue la estrategia más útil para los estudiantes del G5 (73%, $N=66$) y G6 (77%, $N=14$). Estos resultados muestran que en los niveles más avanzados del programa los estudiantes son más visuales (i.e., quieren ver la forma escrita de las palabras,) que auditivos. En este sentido, es crucial considerar la importancia del formato escrito en la presentación del vocabulario para expandir el léxico y aumentar el rendimiento léxico de los estudiantes intermedios. El uso de *las tarjetas de práctica* se puede desarrollar también a través de listas, asociaciones de palabras, uso de la palabra en una oración, definiciones/descripciones en la lengua meta, etc.

Con el fin de discernir las diferencias entre las creencias de los estudiantes y la frecuencia del uso de las palabras por los hablantes nativos, en el siguiente estudio se decidió también analizar si el vocabulario enseñado por el programa es un vocabulario que los estudiantes encontrarán en la vida real.

5.6. Análisis comparativo de frecuencia entre los libros de texto del programa y *A Frequency Dictionary of Spanish (FDS)*

En esta sección del Capítulo 5 se analiza si en el orden de enseñanza de las palabras presentadas en el programa se toman en cuenta sus frecuencias basadas en la producción de hablantes nativos de español en la vida real. Para esto se realizó un análisis comparativo de frecuencia entre los libros de textos usados en el primer y segundo año del programa y *A Frequency Dictionary of Spanish (FDS)* de Mark Davies (2006). El corpus de este estudio, que fue formado de las listas de palabras presentadas en los libros de textos *Puntos de Partida, 9 Ed.* (2011) y *Más, 2 Ed.* (2014) como vocabulario activo, se comparó con el *FDS* de las 5.000 palabras más frecuentes. El planteamiento de este análisis se basa principalmente en:

The goal of a textbook should be to prepare students for the “real world” (as represented by data from an accurate corpus of “real world” speech); the goal should not be that of simply mastering the domains and language functions of “textbook language”. (Davies & Face, 2006)

Según el *FDS*, la distribución de la frecuencia de la familia de palabras presentadas a lo largo de los 4 niveles del programa se muestra en la siguiente tabla.

Tabla 35. Distribución de la familia de palabras presentadas en el programa según *A Frequency Dictionary of Spanish (FDS)* de Mark Davies (2006)

	1- 1000	1001- 2000	2001- 3000	3001- 4000	4001- 5000	sobre 5000	total
principiante- bajo	192	70	44	32	13	54	405
principiante- alto	83	45	38	33	11	81	291
intermedio- bajo	34	34	28	20	11	72	199
intermedio- alto	32	26	28	10	7	26	129
total	341	175	138	95	42	233	1024

Según los datos de la Tabla 35, en el primer año del programa (i.e., en el nivel principiante-bajo y principiante-alto) se observa una cobertura de 27.5 % de las primeras 1.000 palabras más frecuentes en el libro de texto *Puntos de Partida, 9 Ed.* Por otra parte, mientras que 41% del vocabulario presentado en los niveles principiantes aparece entre las primeras 1.001 y 5.000, un 19% de las palabras presentadas en el primer año del programa están por encima del rango de las primeras 5.000 más frecuentes. En cuanto al vocabulario presentado en el segundo año del programa (i.e., en el nivel intermedio-bajo e intermedio-alto), mientras un 5% de las palabras presentadas en el programa entran dentro de las primeras 5.000 familias de palabras más frecuentes, 30% de estas palabras presentadas en el nivel intermedio están por encima del rango de las primeras 5.000 más frecuentes. En resumen, según el *FDS*, sólo un 16% de las primeras 5.000 familias de palabras más frecuentes se presentan como vocabulario activo en el programa de EL2 en la Universidad de Houston durante los primeros dos años. Aquí cabe mencionar que mientras en el corpus se incluyeron sólo las categorías léxicas que incluyen nombres, verbos, adjetivos y adverbios, el *FDS* incluye también las categorías funcionales. Sin embargo, teniendo en cuenta que las categorías funcionales representan sólo un 1.3% de

las primeras 5.000 familias de palabras en el *FDS*, el porcentaje de cobertura de las palabras frecuentes por los dos libros de texto utilizados en el programa no cambia mucho.

Se analizó también si en el orden de enseñanza de estas palabras en los dos libros de texto usados en el nivel principiante e intermedio del programa toman en cuenta las frecuencias de uso en la vida real. De acuerdo con el *FDS*, la tendencia que se siguió en el orden de enseñanza del vocabulario en el programa de EL2, se ilustra en las siguientes figuras (véase Figura 8-11).

Figura 8. Frecuencia de familia de palabras presentadas en el nivel principiante-bajo según *A Frequency Dictionary of Spanish (FDS)* de Mark Davies (2006)

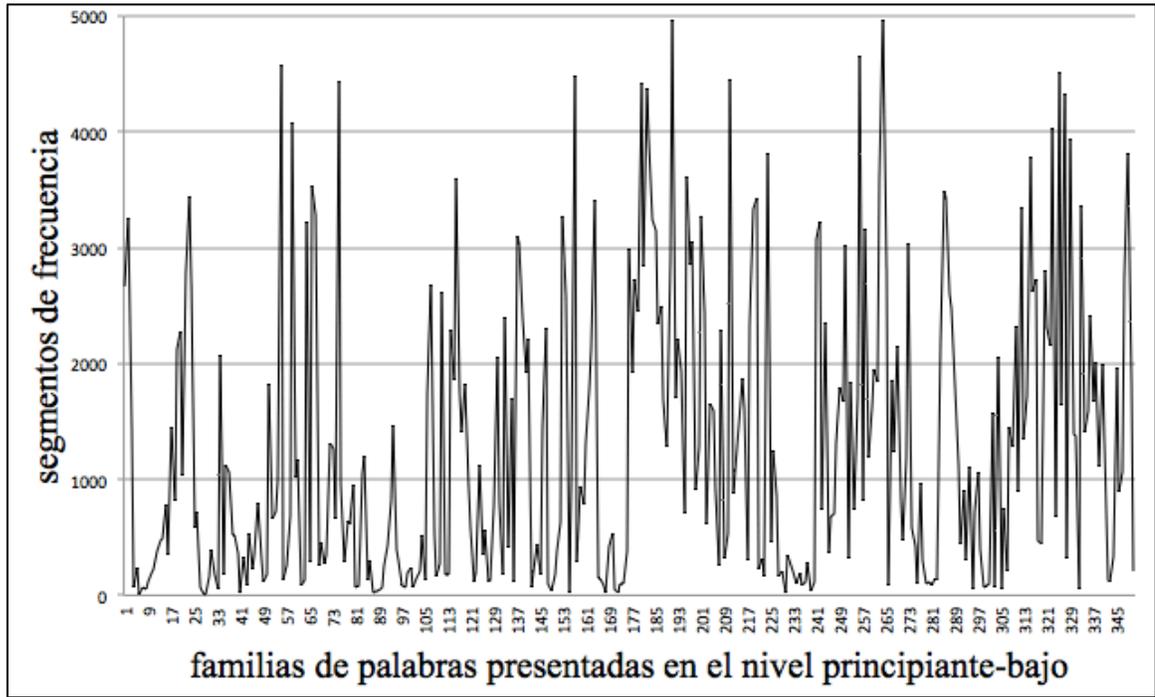


Figura 9. Frecuencia de familia de palabras presentadas en el nivel principiante-alto según *A Frequency Dictionary of Spanish (FDS)* de Mark Davies (2006)

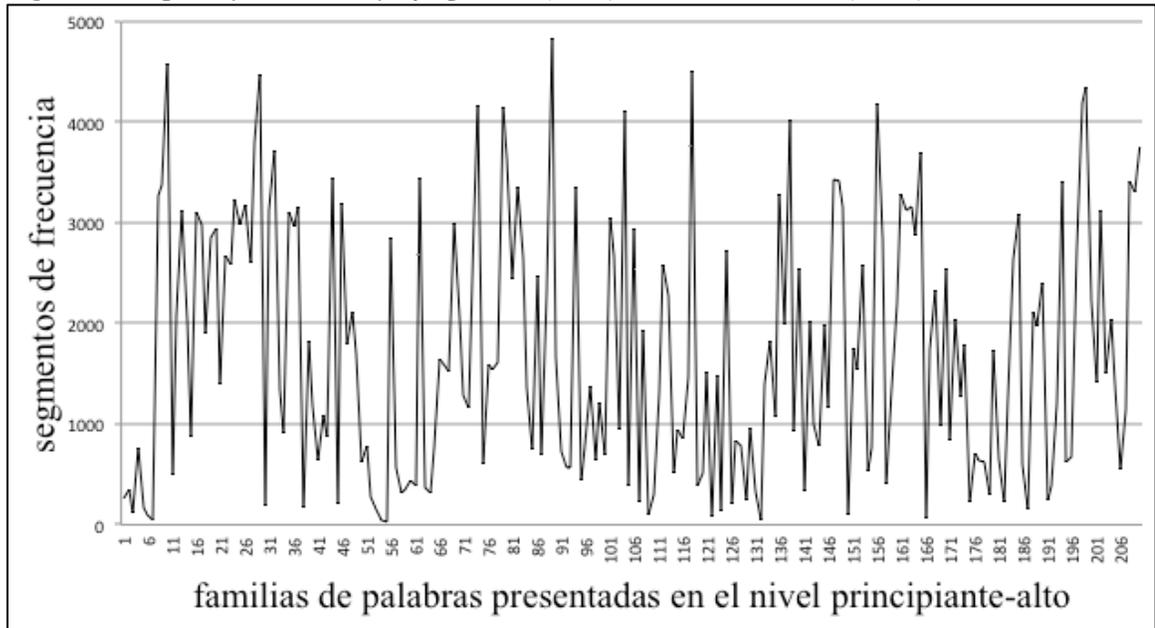


Figura 10. Frecuencia de familia de palabras presentadas en el nivel intermedio-bajo según *A Frequency Dictionary of Spanish (FDS)* de Mark Davies (2006)

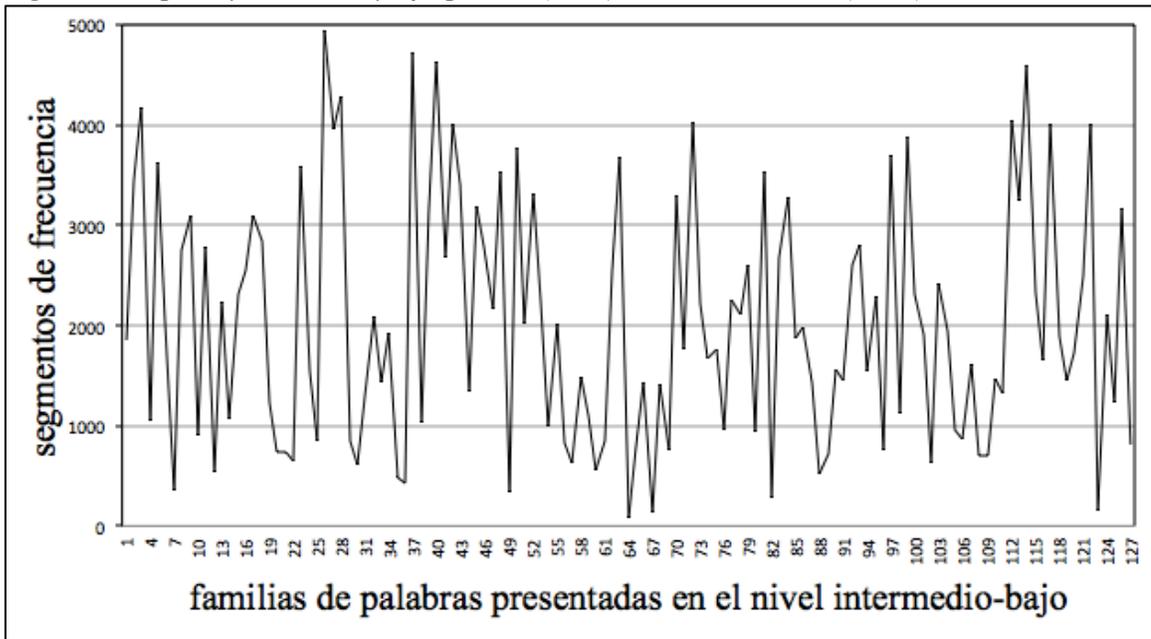
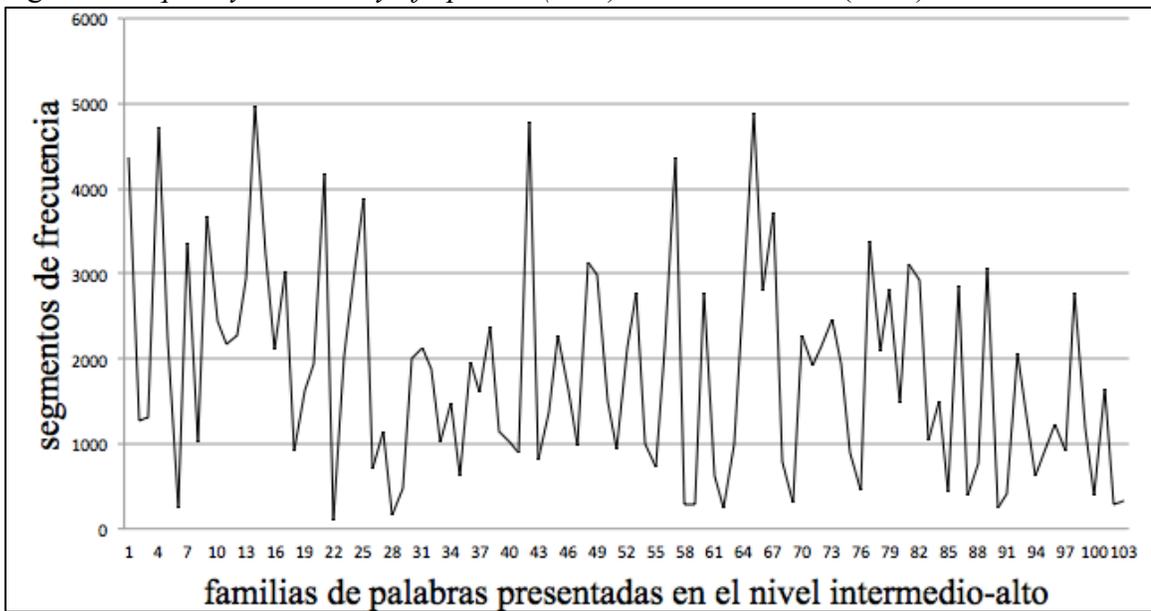


Figura 11. Frecuencia de familia de palabras presentadas en el nivel intermedio-alto según *A Frequency Dictionary of Spanish (FDS)* de Mark Davies (2006)



Los hallazgos encontrados en este análisis apoyan los resultados de los estudios previos (Davies & Face, 2006; Godev, 2009) en el sentido de que los libros de texto de L2 en español sólo parcialmente reflejan las palabras frecuentes usadas por los hablantes nativos en la vida real. Además, como ilustran las Figuras entre 8 y 11, aunque algunas de las palabras presentadas en los dos libros de textos aparecen dentro de las primeras 5.000 palabras más frecuentes, en la presentación de éstas no se sigue un patrón sistemático, en cuanto a su frecuencia de uso por parte de los hablantes nativos. Es otras palabras, la frecuencia de su uso en la vida real no se toma en cuenta en los dos libros de textos analizados en este estudio.

Desde la perspectiva pedagógica, los hallazgos encontrados en este análisis podrían resultar beneficiosos para el diseño de programas y materiales en EL2 considerando el hecho de que en español las 1.000 palabras más frecuentes proporcionan una cobertura de texto del 76% de los textos no ficcionales, el 79.6 % de los textos ficcionales y el 87.8% de los textos orales. Mientras que las siguientes 1.000 palabras cubren el 8% de las palabras en un texto no ficcional, el 6.5% de en texto ficcionales y el 5% en textos orales (Davies, 2005, p.109).

A la luz de los resultados encontrados e interpretados en esta parte del trabajo, en el siguiente capítulo, a modo de conclusión, se resumirán las respuestas dadas a las preguntas de investigación anteriormente planteadas y propondrán algunas implicaciones pedagógicas para incorporar estos aspectos didácticos en el programa de EL2 con el fin de aumentar el rendimiento léxico de los estudiantes en Universidad de Houston. Se espera que estos resultados sean de utilidad para mejorar la instrucción de los estudiantes de EL2 en el vocabulario.

Capítulo 6

Conclusiones e implicaciones pedagógicas

El estudio realizado en la presente investigación partía de la falta de instrumentos fiables, válidos y prácticos para evaluar el aprendizaje y desarrollo del conocimiento léxico de los estudiantes de EL2. El marco teórico de este trabajo se basó principalmente en el planteamiento de Nation (2013) acerca de qué significa saber una palabra, según el cual, el conocimiento de una palabra es una representación mental de gran complejidad, que integra diferentes aspectos y componentes cognitivos, algunos más automáticos e inconscientes y otros más conscientes, reflexivos y experienciales. Por consiguiente, el conocimiento léxico es un proceso complejo y gradual en el que se aprende no sólo la forma y el significado, sino también una intrincada red de relaciones formales y semánticas entre ese *ítem*, y otras palabras y morfemas que constituyen subsistemas de diferentes niveles.

Más específicamente, el propósito de esta disertación fue arrojar luz sobre las siguientes preguntas de investigación. A continuación, a modo de conclusión, se exponen las respuestas dadas a estas interrogantes junto con las implicaciones pedagógicas, así como también se presentan algunas limitaciones del estudio y recomendaciones para futuras investigaciones.

Los primeros interrogantes de esta investigación fueron: ¿Qué tipo de vocabulario se enseña en el primer y segundo año de un programa universitario de EL2 en Estados Unidos? ¿Cuántas palabras se enseñan en cada nivel? Para responder estas preguntas de investigación, en primer lugar, se construyó un corpus léxico basado en las palabras que aparecen en los libros de textos usados en el primer y segundo año del programa de EL2

en la Universidad de Houston durante el semestre de otoño de 2013. Con el fin de formar dicho corpus léxico se analizaron las listas de palabras que aparecen al final de cada capítulo como *vocabulario activo*. Las listas de palabras fueron sacadas de 21 capítulos en total. Después, cada palabra fue etiquetada (*tagged*) según su categoría gramatical. Se incluyó sólo las palabras de contenido que incluyen nombres, verbos, adjetivos y adverbios. Al final, el corpus se redujo a 1024 familias de palabras (i.e., 405 del nivel principiante-bajo, 291 del nivel principiante-alto, 199 del nivel intermedio-bajo y 129 del nivel intermedio-alto) en total. A partir de este corpus se observó que la cantidad de las palabras presentadas en cada nivel baja gradualmente a medida que los estudiantes avanzan de nivel dentro del programa. Mejor dicho, el papel que ocupa el vocabulario en el currículo va bajando gradualmente a lo largo de los primeros 4 niveles del programa por razones curriculares (i. e., menos horas de contacto, mayor cantidad de estructuras gramaticales que deben cubrirse, el peso que tiene en la evaluación, etc.). Considerando el hecho de que los recursos que formaron la base del corpus de este estudio son algunos de los libros de textos más usados por los programas universitarios de EL2 en los Estados Unidos, se puede concluir que este corpus constituye una muestra representativa de lo que se enseña como *vocabulario activo* en los cursos universitarios de EL2 en Estados Unidos.

Las segundas interrogantes planteadas para esta investigación fueron: ¿Qué tanto éxito tiene el programa/los estudiantes en cuanto a la enseñanza /el aprendizaje de este vocabulario? ¿Existe un aumento en el conocimiento receptivo y productivo del vocabulario por parte de los estudiantes del programa EL2 en la Universidad de Houston a medida que avanzan del nivel dentro del programa? ¿Se puede hablar de diferencias

estadísticamente significativas entre los primeros cuatro niveles del programa? Para responder estos interrogantes, se diseñaron dos pruebas de niveles de vocabulario basadas en dicho corpus léxico para evaluar el vocabulario receptivo y productivo de los estudiantes que estudian en el programa de EL2 en la Universidad de Houston. Las pruebas se aplicaron en línea a 410 estudiantes durante el semestre académico de Primavera del 2014.

Según los resultados de la prueba receptiva de niveles de vocabulario (PR-NV), existe un aumento en el reconocimiento léxico por parte de los estudiantes del programa a medida que avanzan de nivel dentro del programa. Sin embargo, la baja frecuencia de puntajes altos y dispersiones entre los grupos indicaron la dificultad de esta prueba, especialmente para los estudiantes del nivel intermedio. Los resultados de pruebas *post hoc* de comparaciones múltiples, por su parte, mostraron diferencias estadísticamente significativas para los grupos del nivel principiante, pero no dieron diferencias estadísticamente significativas en los niveles intermedios en cuanto a su conocimiento léxico receptivo. Considerando el hecho de que esta tarea de reconocimiento, para la cual el candidato tenía que poseer un conocimiento del significado básico de las palabras, fue diseñada en base al vocabulario estudiado en el programa, se esperaba haber encontrado medidas de tendencia más altas. Por otra parte, también se notó que el espacio que ocupa el vocabulario (i. e., la cantidad de horas, el peso que tiene en la evaluación, la cantidad de actividades realizadas en el salón de clase, etc.), como un componente del aprendizaje en EL2, va bajando gradualmente en el plan de estudios y en los libros de texto a lo largo de los primeros 4 niveles, y parece ser un factor efectivo en los resultados. Esto también indicó los vacíos y la falta de retención de las palabras estudiadas en los niveles

anteriores. Probablemente una de las causas es la falta de sistematicidad en la enseñanza del vocabulario como señala Folse (2004):

Vocabulary is not dealt with sufficiently. Yes, some teachers cover some vocabulary, but this is hardly ever done very systematically. Vocabulary is something that everyone assumes that learners will somehow pick up, much the same way everyone assumes that students will just pick up good pronunciation. (p. 131)

Esto conlleva la necesidad de darle más tiempo a la enseñanza del vocabulario en cada nivel. El programa de estudios, los libros de texto y los instructores deberían seguir reciclando continuamente el vocabulario estudiado en los cursos anteriores para garantizar que los estudiantes incorporen estas palabras en su repertorio léxico. Aunque no existe un consenso en la investigación sobre cuántas veces el estudiante tiene que encontrarse con la misma palabra para asimilarla, los estudios empíricos (Kachroo, 1962; Paribakht & Wesche, 1997; Saragai et al., 1978) han enfatizado la necesidad de varios encuentros para enriquecer el aprendizaje de las palabras. En este sentido López-Jiménez (2014) sugiere que:

Several encounters with the same ítem will not only help secure vocabulary knowledge in the learners' memory, so that lexical ítems may be accessed automatically, but also introduce additional features of the same word (e.g., lexical collocations). Lastly, learner autonomy should be promoted through the inclusion of L1 glossaries as well as sections that provide students with effective vocabulary learning strategies. (p.178)

Como parte del programa de enseñanza de lenguas, se debería enseñar a los estudiantes ciertas técnicas efectivas para estudiar y practicar el vocabulario que aumentan la autonomía y la responsabilidad de los estudiantes fuera del salón de la clase como Nation (1997) lo indica en las siguientes líneas:

Language learners see lack of vocabulary as being one of the major hindrances in being able to use another language. But, it is clear from the studies of size and text coverage that is not possible for teachers to present and give practice with all the vocabulary needed for normal language use. It is thus critically important that learners become autonomous with regards to their vocabulary growth. (p. 9)

En cuanto a los resultados de la prueba productiva de niveles de vocabulario (PP-NV), al igual que en la prueba receptiva, se observó un aumento en la recuperación léxica por parte de los estudiantes a medida que avanzan de nivel dentro del programa. Sin embargo, en comparación con la PR-NV, se vio que el progreso es menos marcado. De acuerdo con los resultados de las pruebas posteriores (*Tukey Post-hoc*) de comparaciones múltiples, mientras los niveles principiantes difieren entre sí de manera estadísticamente significativa en cuanto a su conocimiento léxico productivo, no se encontraron diferencias significativas entre los niveles intermedios. En comparación con el contenido léxico, el mayor peso de las estructuras gramaticales y tiempos verbales en el currículo del nivel intermedio explicaron el descenso del progreso en los niveles intermedios. Además, al igual que la prueba anterior, se observó que mientras las palabras abstractas de los niveles intermedios aparecieron entre los ítems con porcentajes de exactitud más bajos; las palabras presentadas en los niveles principiantes se encontraban entre los ítems con los porcentajes más altos en todos los grupos en esta prueba.

En resumen, las estadísticas descriptivas mostraron un alto grado de dificultad de esta prueba productiva para los estudiantes principiantes e intermedios del programa. A diferencia de la prueba receptiva, la dificultad que entraña el uso productivo supone que el aprendizaje exige más tiempo y esfuerzo (Nation, 2013). Por lo tanto, al igual que en los estudios previos (Ellis & Beaton, 1993; Fan, 2000; Laufer, 1998; Laufer et al., 2004; Shintani, 2011; Stoddard, 1929; Waring, 1997), la prueba productiva resultó ser más difícil para los participantes en este estudio. De hecho, los resultados de la prueba *t* para las dos pruebas, confirmaron las diferencias estadísticamente significativas entre el conocimiento del vocabulario receptivo y productivo de los estudiantes principiantes e intermedios del programa. En resumen, como se esperaba, el vocabulario receptivo de los participantes fue significativamente más amplio que su vocabulario activo controlado en este estudio.

En cuanto a la tercera pregunta de investigación: ¿Existe una correlación positiva entre el conocimiento léxico de los estudiantes y su competencia lingüística en la lengua meta? Con el fin de responder esta interrogante y corroborar la covariación que proponen numerosos investigadores (Albrechtsen, Haastrup & Henriksen, 2008; Alderson, 2005; Fairclough, 2011; Laufer, 1992; Laufer & Goldstein, 2004) se realizó un análisis estadístico entre las dos pruebas y el promedio total obtenido del curso del español al final del semestre por cada participante. Se encontraron correlaciones estadísticamente significativas entre los promedios de las dos pruebas de vocabulario y las notas finales para los estudiantes principiantes e intermedios. Además, los resultados del coeficiente de correlación Pearson entre la prueba productiva (PP) y receptiva (PR) de niveles de vocabulario (NV) también reveló correlaciones para estos estudiantes. En este sentido,

estas dos pruebas se pueden usar como herramientas complementarias para evaluar el progreso léxico de los estudiantes que están en el nivel principiante e intermedio del programa.

Por otra parte, parece que en los niveles principiantes e intermedios del programa, las dos pruebas de niveles de vocabulario dieron correlaciones significativas con las notas finales y, por lo tanto, podrían funcionar como predictores del rendimiento académico de EL2. En este sentido, se puede concluir que estas dos pruebas servirían como pruebas con propósito diagnóstico para ubicar a los estudiantes en diferentes niveles del programa. Además, se podrían usar para detectar los puntos fuertes y débiles del conocimiento léxico de los estudiantes que están en los primeros dos años del programa. Si estas dos pruebas se dan a los estudiantes periódicamente, por ejemplo, al final de cada nivel, sería posible no simplemente monitorear el progreso individual de cada estudiante, sino también permitiría obtener información acerca del éxito que tiene el programa en cuanto a la enseñanza del vocabulario.

El grado de fiabilidad de las dos pruebas de vocabulario se analizó a través de diferentes métodos que permiten medir la consistencia interna de cada prueba. Mientras, en la PR-NV, de 40 ítems de 120 palabras, administrada a 410 alumnos, se obtuvo un valor de 0.967; para la PP-NV, de 60 ítems/palabras, administrada a 410 alumnos, el alfa de Cronbach dio por resultado 0.942. Estos valores indicaron alto grado de fiabilidad para las dos pruebas. La fiabilidad de dos pruebas también se comprobó a través del *Alfa de Cronbach* de dos partes, coeficiente de Spearman-Brown y coeficiente de dos mitades de Guttman.

En síntesis, se puede concluir, al igual que lo hicieron otros estudios previos (Beglar & Hunt, 1999; Read, 1998; Schmitt *et. al.*, 1999), que esta prueba cumple los requisitos de una prueba fiable como controlar bien los factores relacionados en elaboración de los ítems, influencia del factor azar, claridad de las instrucciones, relación existente entre la dificultad de la prueba y el nivel de competencia de los participantes, etc. En este sentido, podrían funcionar como herramientas fiables y prácticas para la evaluación del programa y los materiales usados con respecto a la enseñanza del componente léxico.

Con el fin de responder la cuarta pregunta de investigación: ¿Qué tipo de estrategias prefieren usar los estudiantes a la hora de aprender y practicar el vocabulario?, se diseñó un cuestionario para hacerles reflexionar a los participantes sobre la importancia del vocabulario y las estrategias que prefieren usar a la hora de aprenderlo/practicarlo. De acuerdo con los resultados de este cuestionario, la mayoría de los estudiantes en el programa encontraron el vocabulario *importante* o *muy importante*. Se observó un aumento en la importancia dada al vocabulario por parte de los estudiantes a medida que avanzan de nivel dentro del programa. Estos resultados señalaron que mientras los estudiantes en los niveles más avanzados pensaron que las palabras son muy importantes -y por ello desean aprenderlas-, el papel que ocupa el vocabulario en el currículo va bajando gradualmente a lo largo de los primeros 4 niveles del programa. Lewis (2000), por su parte, llamaba la atención a los problemas que acarrear las deficiencias en el dominio del léxico, en especial en alumnos que ya han alcanzado un cierto nivel:

La carencia de léxico supone, en mayor medida que las deficiencias gramaticales, una gran laguna en los estudiantes de lenguas extranjeras. Esto se hace mucho más evidente en los estudiantes de nivel intermedio y avanzado, en los que parece que no hay un progreso claro. La principal diferencia entre ambos niveles no está en la complejidad de la gramática, sino en la cantidad de unidades léxicas que tienen disponibles en el lexicón mental. No reconocer esto supone condenar a los estudiantes a un eterno nivel intermedio. (p. 8)

En este sentido, es necesario darle suficiente tiempo a la enseñanza del vocabulario considerando la importancia cada vez más creciente que tiene para los estudiantes en el programa.

En cuanto a las estrategias para aprender vocabulario desconocido, mientras que usar un diccionario bilingüe fue evaluada como la estrategia más útil, usar un diccionario monolingüe fue valorada como la estrategia menos útil para aprender el vocabulario desconocido por parte de los participantes de este estudio. Como parte del programa de enseñanza de lenguas se debería enseñar a los estudiantes ciertas técnicas efectivas que ayuden a aprender, repasar y ampliar vocabulario, a la vez que la utilización de fuentes y recursos para potenciar el desarrollo individualizado de este. Por ejemplo, dentro de esas técnicas habría que incorporar la enseñanza del manejo del diccionario ya que esta es una herramienta consultada por muchos estudiantes en el programa.

De acuerdo con Borden et al. (2004), es posible diseñar actividades prácticas relacionadas con el uso del diccionario bilingüe o monolingüe dependiendo del nivel de los estudiantes. Por ejemplo, los de nivel principiante pueden familiarizarse con el uso del diccionario, practicar la búsqueda de términos, aprender la disposición de las entradas

en el diccionario, la información que contienen y las abreviaturas comunes; los estudiantes de nivel avanzado pueden tratar de interpretar las diferentes definiciones que el diccionario provee y elegir la definición más adecuada según el contexto en el que se ubican o pueden aprender diversos sinónimos y antónimos de una palabra descubriendo la riqueza léxica de la lengua española.

Con respecto a las estrategias para practicar el vocabulario, mientras que los estudiantes en el nivel principiante-bajo encontraron la estrategia de estudiar una palabra con su imagen *muy útil* dado que los libros de textos presentan palabras más concretas en estos niveles, la estrategia de usar una palabra en oraciones se seleccionó como una estrategia más útil por parte de los participantes en el nivel principiante-alto e intermedio-bajo. Lo cual, marca la importancia del contexto en el aprendizaje del vocabulario para los niveles principiantes e intermedios.

Por otra parte, se observó una similitud entre las preferencias de los estudiantes del nivel intermedio-alto y avanzado en cuanto a la utilidad de las estrategias para practicar el vocabulario. El uso de las tarjetas de práctica fue la estrategia más útil para los estudiantes avanzados, lo cual recalca la importancia del formato escrito en su presentación para expandir el vocabulario y aumentar el rendimiento léxico de los estudiantes intermedios. El uso de las tarjetas de práctica se puede desarrollar también a través de listas, asociaciones de palabras, uso de la palabra en una oración, definiciones/descripciones en la lengua meta, etc. Por otra parte, las estrategias tienen mayor eficacia cuando los estudiantes combinan diferentes métodos de aprendizaje y desarrollan su propia estrategia de aprender/practicar el vocabulario. En este sentido, el uso de las tarjetas de práctica permite la libertad de una aproximación ecléctica, siempre

que ésta sea sistemática y esté bien fundamentada, desarrollando la autonomía del aprendizaje, además de mejorar su memoria. Otra forma que resultaría eficaz para ampliar el léxico de los estudiantes en el nivel intermedio sería mediante asociaciones de palabras. Las asociaciones fomentan la memorización de palabras por parte de los alumnos. Según Tornberg (2009), es muy importante que los alumnos puedan relacionar la nueva palabra con sus conocimientos anteriores. De esa manera pueden recordar la palabra durante más tiempo. Los alumnos retienen la palabra en la memoria a largo plazo si pueden memorizarla a través de la creación de redes de asociaciones.

En síntesis, por cuestiones de tiempo e intensidad del plan curricular, en los niveles intermedios y más avanzados, es fundamental promover la autonomía en el aprendizaje, aumentando la responsabilidad y capacitando al estudiante mediante la utilización de estrategias de aprendizaje y práctica del vocabulario. Para alcanzar esta meta, se podría organizar talleres en que se muestren estrategias y técnicas útiles que ayuden a aprender, repasar y ampliar vocabulario, a la vez que la utilización de fuentes y recursos (ej. diccionarios, gramáticas, estrategias compensatorias, etc.) para potenciar el desarrollo individualizado de vocabulario.

En resumen, desde la perspectiva pedagógica, los hallazgos encontrados en este cuestionario de reflexión sobre el vocabulario podrían resultar beneficiosos para la selección de las técnicas de presentación y enseñanza del vocabulario en el programa.

Finalmente, a través de la última pregunta de investigación, se analizó si en el orden de enseñanza de las palabras presentadas en el programa se toman en cuenta sus frecuencias basadas en la producción de hablantes nativos de español en la vida real. Para esto se realizó un análisis comparativo de frecuencia entre las palabras presentadas por

los textos *Puntos de Partida, 9 Ed.* (2011) y *Más, 2 Ed.* (2014) como vocabulario activo, y las 5.000 palabras más frecuentes de *A Frequency Dictionary of Spanish (FDS)* de Mark Davies (2006).

Los resultados de este análisis apoyaron los hallazgos encontrados en los estudios previos (Davies & Face, 2006; Godev, 2009) en el sentido de que los libros de texto de L2 en español sólo parcialmente reflejan las palabras frecuentes usadas por los hablantes nativos en la vida real. De acuerdo con el *FDS*, mientras se cubre sólo el 11.2 % de las primeras 5.000 palabras más frecuentes en *Puntos de Partida, 9 Ed.* (2011), en el primer año del programa, esta cobertura baja al 5% en *Más, 2 Ed.* (2014) en el segundo año del programa. Además, se observó que, aunque algunas de las palabras presentadas en los dos libros de texto aparecen dentro de las primeras 5.000 palabras más frecuentes, en la presentación de éstas no se sigue un patrón sistemático en cuanto a su frecuencia de uso por parte de los hablantes nativos. Esto conlleva la necesidad de tomar en cuenta su frecuencia de uso como un criterio fundamental a la hora de selección de los libros de texto si la meta es lograr la competencia comunicativa. En este sentido, además del *Corpus del Español* (<http://www.corpusdelespanol.org>), el Diccionario de frecuencias de Davies (2006) publicado por el mismo autor, podría resultar útil para los autores y editores de los libros de textos en la selección y categorización de las palabras que se deben enseñar en cada nivel. Estas palabras de frecuencia presentadas en el diccionario de Davies, podrían servir en el diseño de los textos de lectura y práctica usados en el salón de EL2. En el mismo diccionario el propio autor propone treinta listas temáticas de vocabulario, lo que él llama *call-out boxes*. Estas listas presentan las palabras temáticamente relacionadas con : los animales, el cuerpo, la comida, la ropa, la familia,

los profesiones, las nacionalidades, los colores, el tiempo, los deportes, los verbos de movimiento y comunicación etc. Además, el mismo diccionario propone algunos temas gramaticales que resultan difíciles para los estudiantes principiantes e intermedios, tales como los diminutivos más comunes, los superlativos y los sufijos derivativos para formar sustantivos, los verbos y adjetivos más comunes que se usan con el subjuntivo, etc. En el diccionario también aparecen listas que comparan el uso de sustantivos, verbos, adjetivos y adverbios de acuerdo a diferentes registros y géneros. En base a los resultados de este estudio, se recomienda tomar en cuenta las palabras que aparecen dentro de las primeras 5.000 familias palabras más frecuentes antes de enseñar las que están por encima de este rango en la enseñanza del vocabulario en los niveles más principiantes. Por ejemplo, de acuerdo con el *FDS*, mientras palabras como *langosta, camarones, atún, salchicha, jamón, espárragos* no aparecen dentro de las 5.000 familias de palabras más frecuentes, bajo el tema de “la comida”, se presentan en el libro de texto del nivel principiante. Por otro lado, aunque palabras como *miel, grasa, maíz, harina, uva, chocolate, trigo, caldo* entran en el mismo rango, no aparecen dentro del vocabulario activo en los libros de textos analizados. Además, al seleccionar libros para el nivel intermedio, hay que asegurarse de que los materiales incluyan actividades de reciclaje focalizadas en el léxico de los niveles principiantes.

Para las investigaciones futuras se sigue que la cantidad de los libros de texto incorporados en la elaboración del corpus de este estudio podría expandirse para que sea mucho más representativo de lo que se enseña en Estados Unidos. De esa manera, también sería posible diseñar pruebas de niveles de vocabulario estandarizadas basadas en un corpus más representativo. Cabe aquí también mencionar, la necesidad urgente y

creciente de un programa de computadora que permita administrar las pruebas léxicas en línea, y analizar sus ítems en detalle de manera mucho más práctica y rápida. Por otra parte, sería interesante analizar y comparar los resultados de la prueba receptiva y productiva de niveles de vocabulario con la producción escrita y oral libre.

Para concluir, el trabajo desarrollado en esta disertación mostró que es imprescindible implementar una evaluación sistemática del vocabulario, lo que indicaría si el proceso de enseñanza se desarrolla según se planificó o hace falta realizar modificaciones para evitar vacíos en el repertorio léxico de los estudiantes de EL2.

Apéndice A

Prueba de cognados

¿Cómo se dice..... en inglés? Translate the following words into English. You may not have studied some of these words yet, but let's see if you can guess their equivalents in English.

1. gas:		35. ensalada:	
2. electricidad:		36. conservador:	
3. barbacoa		37. museo:	
4. millón:		38. dominó:	
5. universidad:		39. extinguir:	
6. residente:		40. química:	
7. diccionario:		41. entretenimiento:	
8. carrera:		42. parcial:	
9. diciembre:		43. completo:	
10. junio:		44. contaminación:	
11. población:		45. motocicleta:	
12. cabina:		46. tren	
13. filosofía		47. economía:	
14. foto:		48. sociología	
15. vacaciones:		49. discriminación:	
16. estrés:		50. autobús:	
17. septiembre:		51. humanidades:	
18. residencia:		52. esquí:	
19. inmigrante:		53. egoísta:	
20. zona:		54. cine:	
21. personalidad:		55. pasaporte:	
22. apariencia:		56. interés:	
23. gasolina:		57. piloto:	
24. física:		58. abuso:	
25. concierto:		59. símbolo:	
26. secretario:		60. religioso:	
27. extrovertido:		61. experiencia:	
28. mestizo:		62. valle:	
29. ciencias		63. metro	
30. desierto:		64. síntoma:	
31. arte:		65. nacionalidad:	
32. funcionar:		66. bautismo:	
33. electricista:		67. insensible:	
34. imagen:		68. batería:	

69. musulmán:		85. agosto:	
70. facultad:		86. capital:	
71. noviembre:		87. modelo:	
72. bicicleta:		88. informativo:	
73. responsabilidad		89. justo:	
74. examen:		90. objetivo:	
75. octubre:		91. cliente:	
76. referencia:		92. académico:	
77. electrónico:		93. cero:	
78. doméstico:		94. oficina	
79. calculadora:		95. océano:	
80. beneficio:		96. bilingüe:	
81. carácter:		97. indígena:	
82. generoso:		98. informática:	
83. discoteca:		99. artículo:	
84. teatro:		100. norte:	

Apéndice B

Students' Background Questionnaire

1. Name: _____
 2. E-mail: _____
 3. Age: _____
 4. Sex: Male / Female
 5. What is (are) your first language(s)? _____
 6. What language(s) do you speak? _____
 7. What language(s) do you speak at home? _____
 8. Age at which you started to learn Spanish: _____
 9. How long have you been studying Spanish? Please specify how many year(s) or semester(s). _____ year(s) or _____ semester(s)
 10. What other language(s) have you studied as a second language?
-

	a. Which Spanish classes have you taken?	b. How many years in total?	c. List the name(s) of the institution(s) where you have taken these classes:
11. In middle school/ high school			
12. In university /college			

13. a. Have you ever studied Spanish in a Spanish-speaking country? _____

b. What classes have you taken and how long was your study period? _____

c. In which country and institution? _____

14. Have you ever been in a Spanish speaking country? Yes/ No

15. If your answer is “yes” to the previous question, please explain to which country, when and how long your stay was:

16. Do you have **other experiences** with the Spanish language (e.g., working with Spanish- speaking people, having a Spanish-speaking caregiver in your early childhood, getting married with a native speaker of Spanish, etc.)?

Vocabulary Reflection Questionnaire

17. Thinking back **on your own experience of learning Spanish**, what **methods or strategies** do you consider **helpful to your vocabulary learning**?

18. How **important** do you think **vocabulary learning** is in learning Spanish? Please rate it on the following scale of 1 to 5.

1 Not very important	2 Somewhat important	3 Moderately important	4 Important	5 Extremely important
-----------------------------------	-----------------------------------	-------------------------------------	-----------------------	------------------------------------

19. Which **Spanish textbook(s)** have you been using /did you use in your **beginning** Spanish course(s)?

20. Which **Spanish textbook(s)** have you been using/did you use in your **intermediate** Spanish course(s)?

Textbook evaluation: Considering the **Spanish textbook** that you are **currently using** in your **Spanish class**, please **rank** the following statements on a scale of 1 to 5:

	1 strongly disagree	2 disagree	3 neutral	4 agree	5 strongly agree
21. Vocabulary presented in this textbook is useful for me to communicate with native speakers of Spanish in real -life situations.					
22. The vocabulary relates to engaging topics.					
23. The vocabulary is presented in a meaningful context.					
24. It provides engaging vocabulary activities.					
25. The vocabulary is recycled through the chapters so I can review the words that I learnt in the previous chapters.					

Please **rank** the following **learning strategies** for **new vocabulary** according to their **usefulness for you** on a scale of 1 to 5:

	1 not at all useful	2 slightly useful	3 useful	4 very useful	5 extremely useful
26. To use a bilingual dictionary (Span. to Eng. & Eng. to Span.).					
27. To use a monolingual dictionary (Span.to Span.).					
28. To guess from textual context.					
29. To analyze part of speech of an unknown word. (e.g., predicting if an unknown word is a noun, adjective, adverb, etc.) in a sentence.					
30. To analyze affixes and roots of an unknown word. [(e.g., felicidad (n), amable (adj.)]					
31. If you use other vocabulary learning strategies for new vocabulary , please specify which ones:					

Please **rank** the following **study strategies for vocabulary** according to their **usefulness** **for you** on a scale of 1 to 5:

	1 not at all useful	2 slightly useful	3 useful	4 very useful	5 extremely useful
32. To use flash cards					
33. To study a word with a picture instead of its definition					
34. To connect a word to a personal experience					
35. To connect a word to its synonyms and antonyms.					
36. To repeat a word aloud to oneself.					
37. To write a word repeatedly.					
38. To listen a word repeatedly.					
39. To use new words in sentences.					
40. If you use other studying strategies for vocabulary please specify which ones:					

Apéndice C

Receptive Vocabulary Levels Test

This is a vocabulary test. Please read these instructions **carefully**.

For each set of six words, there are three meanings. You must choose the right word to go with each meaning. Select the number of that word next to its meaning. Here is an example:

Modelo: 1. amigo
2. cocina 4 se usa para leer
3. dinero 2 parte de una casa
4. libro 5 animal con cuatro patas
5. perro
6. señor

Some words are in the test to make it more difficult. You do not have to find a meaning for these words. In the example above, these words are *amigo*, *dinero* and *señor*.

If you have no idea about the meaning of a word, do not guess. But if you think you might know the meaning, then you should try to find the answer. To get **helpful** and **effective feedback** it is necessary to try **each question** in the test. You have **30 minutes** to do this test.

1. abrazarse
2. aburrirse _____ expresar algo
3. decir _____ limpiar algo con agua u otro líquido
4. lavar _____ una manera de demostrar afecto entre amigos
5. resfriarse
6. subir

1. barato
2. cómodo _____ confortable
3. gris _____ cincuenta y veinte
4. ochenta _____ mezcla de blanco y negro
5. rosado
6. setenta

1. cuarto
2. edificio _____ sirve para escribir
3. extranjero _____ persona que es de otro país
4. lápiz _____ lugar donde se venden libros
5. librería
6. palabra

- 1. acostumbrarse
- 2. ahorrar _____ pedir imperiosamente
- 3. esforzarse _____ percibir algo por la nariz
- 4. exigir _____ trabajar con ánimo en busca de un objetivo
- 5. desperdiciar
- 6. oler

- 1. caminar
- 2. discutir _____ ir a pie
- 3. encontrarse _____ hablar sobre un tema
- 4. sentirse _____ experimentar emociones
- 5. toser
- 6. viajar

- 1. amarillo
- 2. fiel _____ poco inteligente
- 3. largo _____ color oscuro de piel o pelo
- 4. moreno _____ una de las características de buen amigo
- 5. perezoso
- 6. tonto

- 1. cansado
- 2. delgado _____ próximo
- 3. primero _____ exhausto
- 4. quinientos _____ cinco veces cien
- 5. setecientos
- 6. siguiente

- 1. amistad
- 2. calificación _____ cara de las personas
- 3. cicatriz _____ relación entre amigos
- 4. horario _____ nota obtenida de un examen
- 5. oración
- 6. rostro

- 1. caminadora
- 2. chiste _____ nivel de estudio
- 3. grado _____ máquina para correr
- 4. papa _____ verdura que se puede comer frito
- 5. pasillo
- 6. vela

- 1. aburrido
- 2. congelado _____ amado
- 3. limpio _____ no sucio
- 4. norte _____ contrario a sur
- 5. peor
- 6. querido

- 1. acera
- 2. cosecha _____ tiempo libre
- 3. deuda _____ cantidad excesiva de habitantes
- 4. humo _____ dinero que una persona tiene que pagar
- 5. ocio
- 6. sobrepoblación

- 1. adivinar
- 2. ayudar _____ no gustar
- 3. contestar _____ responder
- 4. esperar _____ predecir el futuro
- 5. merendar
- 6. odiar

- 1. asiento
- 2. arvejas _____ tipo de verdura
- 3. chuleta _____ carne de cerdo para barbacoa
- 4. cuerpo _____ conjunto del sistema orgánico de los humanos
- 5. dueño
- 6. horno

- 1. barrio
- 2. ciudadano _____ parte sólida de los árboles
- 3. envase _____ no aceptación de una idea o acción
- 4. esperanza _____ algo que sirve para conservar productos
- 5. madera
- 6. rechazo

- 1. fiebre
- 2. jugo _____ bebida de fruta
- 3. oído _____ escala que hace un avión
- 4. parada _____ órgano que percibe sonidos
- 5. patinaje
- 6. pavo

- 1. camarero
- 2. camarones _____ tipo de mariscos
- 3. garganta _____ un lugar para hacer camping
- 4. montaña _____ persona que sirve en un restaurante
- 5. tarjeta
- 6. vuelo

- 1. aduana
- 2. ciclismo _____ malestar físico
- 3. cita _____ comida típica española
- 4. resfriado _____ encuentro entre dos o más persona
- 5. tapas
- 6. zanahoria

- 1. aspiradora
- 2. billete _____ producto de la leche
- 3. invitado _____ algo necesario para viajar
- 4. letra _____ una máquina que sirve para limpiar
- 5. mar
- 6. queso

- 1. analfabetismo
- 2. feria _____ no saber leer ni escribir
- 3. ley _____ déficit de lluvia por mucho tiempo
- 4. orgullo _____ norma establecida para regular o prohibir
- 5. sequía
- 6. taza

- 1. almorzar
- 2. empezar _____ llegar a un destino
- 3. querer _____ contrario de finalizar
- 4. mirar _____ tomar lo que te dan o te mandan
- 5. recibir
- 6. venir

- 1. aceite
- 2. basura _____ control médico
- 3. chequeo _____ diversión y entretenimiento
- 4. lavadora _____ un electrodoméstico para ropa
- 5. maleta
- 6. pasatiempo

- 1. consejero
- 2. cuaderno _____ el hijo de mi tío
- 3. esposo _____ el hombre casado
- 4. fiesta _____ parte de un apartamento
- 5. primo
- 6. ventana

- 1. anuncio
- 2. barba _____ falta de trabajo
- 3. desempleo _____ información o aviso clasificado
- 4. entrevista _____ conversación entre dos personas
- 5. soledad
- 6. sonrisa

- 1. brindis
- 2. gerente _____ persona que administra
- 3. impresora _____ contrato de vida o médico
- 4. jubilación _____ aparato que sirve para reproducir textos
- 5. licenciatura
- 6. seguro

- 1. cama
- 2. diente _____ sirve para dormir
- 3. jardín _____ lugar donde hay plantas y flores
- 4. llave _____ sitio dónde se venden ropa y comida
- 5. mercado
- 6. oro

- 1. ahora
- 2. allá _____ más tarde
- 3. casi _____ eternamente
- 4. luego _____ en este momento
- 5. siempre
- 6. todavía

- 1. albañil
- 2. apodo _____ amor o afecto
- 3. bigote _____ la mujer de mi hijo
- 4. cariño _____ persona que participa en un negocio con otras
- 5. nuera
- 6. socio

1. desechable
2. discapacitado _____ de un solo uso
3. invernadero _____ persona rechazada por la sociedad
4. marginado _____ persona que tiene alguna deficiencia
5. medioambiental
6. orgulloso

1. ahijado
2. contabilidad _____ expresión de alegría o satisfacción
3. despido _____ estudio de operaciones económicas
4. entierro _____ persona que se dedica a la instalación de agua
5. felicitación
6. fontanero

1. cantar
2. comer _____ consumir comida
3. enseñar _____ volver a un lugar
4. escuchar _____ una de las responsabilidades del profesor
5. regresar
6. trabajar

1. falda
2. fecha _____ indica el día, mes y año
3. película _____ lugar para visitar cuando hace sol
4. playa _____ cosa que se usa después de bañarse
5. sillón
6. toalla

1. anoche
2. demasiado _____ nunca
3. desgraciadamente _____ infelizmente
4. jamás _____ excesivamente
5. tampoco
6. ya

1. alcoba
2. algodón _____ ropa deportiva
3. gorra _____ reducción de los precios
4. rebajas _____ sirve para protegerse del sol
5. reloj
6. sudadera

- 1. alemán
- 2. alto _____ diez y cinco
- 3. amable _____ educado, agradable
- 4. feo _____ de un país en Europa
- 5. once
- 6. quince

- 1. bandera
- 2. cuchara _____ pausa en el trabajo
- 3. descanso _____ zona intermedia entre dos países
- 4. frontera _____ utensilio que sirve para tomar sopa
- 5. gente
- 6. idioma

- 1. aumentar
- 2. mudarse _____ cambiar de casa
- 3. parecerse _____ dar un toque a una tecla
- 4. pulsar _____ recitar oraciones religiosas
- 5. rezar
- 6. suspender

- 1. camisa
- 2. casa _____ lugar para vivir
- 3. mesa _____ ropa para llevar al trabajo
- 4. mochila _____ publicación periódica para leer
- 5. pizarrón
- 6. revista

- 1. asado
- 2. deportivo _____ contento
- 3. divertido _____ de buena salud
- 4. feliz _____ estilo de cocinar carne
- 5. libre
- 6. sano

- 1. entender
- 2. despertarse _____ continuar
- 3. divertirse _____ dejar de dormir
- 4. llover _____ tener una idea clara de algo
- 5. pedir
- 6. seguir

1. ateo

2. callado

3. castaño

4. ondulado

5. tacaño

6. terco

_____ un color de pelo

_____ pelo ni liso ni rizado

_____ persona que no le gusta gastar o compartir

Apéndice D

Productive Vocabulary Levels Test

Complete the following blanks with the appropriate words. They can be **nouns, verbs, adjectives or adverbs**. You **don't need to conjugate the verbs**; put them in **infinitive** forms.

The initial letters of the words are already given; you just need to type the **rest of the word**. Here are some examples:

Modelo 1: Esta noche hay una fiesta en el centro estudiantil para los estudiantes nuevos.

Modelo 2: A mi hermana le gusta mucho hablar por teléfono.

You will have **30 minutes** to do this test.

1. Él *se quedó cie* _____ debido a un accidente durante su niñez. Por eso, no puede ver.
2. Hay mucho rui _____ en la biblioteca. Por eso, no puedo estudiar bien allá.
3. Primero se cortan las cebollas en rodajas con un cu _____.
4. Esta plataforma de red social no tiene ningún usua _____; parece que nadie la utiliza.
5. Ir a la iglesia todos los domingos es un ri _____ religioso en mi familia.
6. Mi abuelo no tiene buena memoria y le cuesta re _____ los nombres de sus nietos.
7. La semana pasada mi hermana se rompió el bra _____ montando su bicicleta.
8. La salsa que comimos en la fiesta era muy pi _____.
9. Este vestido de se _____ es muy cómoda. Por eso, me encanta la textura de ese material.
10. Los pul _____ son los órganos de la respiración que se encuentran situados en la cavidad torácica.
11. Él es un chico muy tí _____ e introvertido; no le gusta hablar con nadie.
12. A él le molesta mucho vo _____ en avión. Por eso, no viaja mucho a otros continentes.
13. Existen muchos criterios diferenciales que determinan el ni _____ económico de un país.
14. Tengo que apro _____ todos mis exámenes para graduarme en mayo.
15. Después de lavar la ropa, ella siempre la pone en la se _____.
16. Tengo un ga _____ que se llama Pompón.
17. Si quieres ser abo _____, tienes que estudiar derecho.
18. Ayer me dolía mucho un diente, por eso decidí ir a la den _____.
19. Es una familia tan po _____ que no tiene dinero para comprar comida para sus hijos.

20. Ella practica yoga dos veces a la semana para rel _____ se.
21. Hoy estoy muy ale _____ porque mis padres vienen a Houston.
22. Es importante bo _____ las botellas plásticas en el contenedor de reciclados.
23. Los empleados deben trabajar en equipo para as _____ en esta compañía.
24. Tengo traje de baño para ir a la playa pero necesito comprarme unas san _____.
25. Mi compañero de cuarto es muy desordenado, nunca organiza su ropa en el ar _____.
26. En nuestro último viaje tuvimos que esperar mucho en el aeropuerto porque hubo una de _____ de cinco horas en nuestro vuelo.
27. La biblioteca está en el tercer pi _____ de este edificio.
28. Este semestre mi mat _____ favorita es biología.
29. Para llevar una vida sana es importante hacer ejercicios fre _____.
30. ¿A ti te gusta to _____ el piano?
31. Por las mañanas después de du _____ se, mi esposo siempre se viste y desayuna.
32. El estudiante está muy pre _____ porque hoy tiene un examen muy importante de matemáticas.
33. Su primo es un bebé muy tran _____, casi nunca llora y siempre duerme.
34. Quiero co _____ el restaurante nuevo que está en la calle San Felipe.
35. Diez multiplicado por cien es m _____.
36. Mi mejor amigo estudia medicina porque quiere s _____ médico en el futuro.
37. Mi hermana siempre come pan con mant _____ en el desayuno.
38. A mi madre le encanta ju _____ al tenis los fines de semana.
39. Mi padre es un hom _____ muy interesante.
40. Al morir sus padres le dejaron una gran he _____.
41. Mi primo es un chico muy simpático, gordo y ba _____.
42. Los domingos mis padres se levantan muy tem _____ pero yo prefiero dormir hasta el mediodía.
43. A mi mejor amiga le gusta tomar su café con crema y mucho azú _____.
44. No le gusta nada plan _____ la ropa; es uno de los quehaceres domésticos que menos le gusta.
45. Mi sue _____ es un magnífico abuelo para mis hijos.
46. E _____ es el primer mes del año.
47. Para ir al trabajo, siempre me gusta llevar una blu _____ y unos pantalones.
48. ¿Has visto las no _____ ayer en la televisión? El presidente nuevo habló ante el público.
49. Yo siempre prefiero pagar con tar _____ de crédito, casi nunca tengo dinero en efectivo.
50. ¿Me puedes mos _____ las fotos que sacaron en su último viaje?

51. El aho _____ de energía debe ser un movimiento mundial para enfrentar los cambios que experimenta el medio ambiente.
52. Este estudiante necesita estudiar mucho para tener éxi _____ este semestre.
53. Me gusta comer helado como pos _____.
54. Mi color favorito es ana _____; es un color muy vivo y brillante.
55. Once y seis son die _____.
56. ¿Me puedes dar la re _____ de la salsa de guacamole que preparaste para la fiesta?
57. Mi padre juega muy bien al aje _____.
58. En 1978 terminé mi carrera universitaria y obtuve mi li _____ en historia del arte.
59. En las ciudades grandes, la mayoría de las personas prefieren vi _____ en apartamentos.
60. La lu _____ contra el racismo y la discriminación es una cuestión prioritaria para la sociedad.

Apéndice E

UNIVERSITY OF HOUSTON CONSENT TO PARTICIPATE IN RESEARCH

PROJECT TITLE: Assessment of Vocabulary Knowledge in Spanish as a Second Language (*Evaluación del vocabulario en la enseñanza y el aprendizaje de español como segunda lengua*)

You are being invited to take part in a doctoral dissertation research project conducted by Burcu Mutlu from the Hispanic Studies Department at the University of Houston under the supervision of Dr. Marta Fairclough.

NON-PARTICIPATION STATEMENT

Taking part in the research project is voluntary and you may refuse to take part or withdraw at any time without penalty or loss of benefits to which you are otherwise entitled. You may also refuse to answer any research-related questions that make you uncomfortable. A decision to participate or not or to withdraw your participation will have no effect on your academic standing in this class.

PURPOSE OF THE STUDY

The objective of this study is to evaluate the receptive and productive Spanish vocabulary knowledge of second language learners. The results of this study may show if the teaching and learning materials, strategies and methods that are being implemented are working effectively. If they are not, what changes should be made and how educators might improve courses designed for second language learners. The study will take an hour and 20 minutes, which is equal to one class meeting.

PROCEDURES

You will be one of approximately 850 subjects invited to take part in this project at the Language Acquisition Center in UH. The activities of this project will be completed during regular time class. First, you will be asked to complete a linguistic background questionnaire, which will take 20 minutes. Second, you will be asked to complete a receptive vocabulary test in Spanish that will take 30 minutes. This test consists of 40 matching questions in which you pair the words with their corresponding meanings. Third, you will be asked to complete a productive vocabulary test that will take 30 minutes as well. In this third exercise only the initial letters of words are given and you will be asked to complete the remaining letters of these words according to the context.

CONFIDENTIALITY

All the subjects' identity will be held in confidence. Every effort will be made to maintain the confidentiality of your participation in this project. Each subject's name will be paired with a code number by the principal investigator. This code number will appear on all written materials. The list pairing the subject's name to the assigned code number will be kept separate from all research materials and will be available only to the principal investigator. Confidentiality will be maintained within legal limits.

RISKS/DISCOMFORTS

There are no foreseeable risks in this study.

BENEFITS

While you will not directly benefit from participation, your participation may help investigators better understand what students know about vocabulary aspects of the Spanish language, and how educators might improve courses designed for second language learners.

ALTERNATIVES

Participation in this project is voluntary and the only alternative to this project is non-participation. If you decide not to participate in this study, you will be required to complete the regular classroom activities (worksheets, exercises from the course book, etc.) assigned by your Spanish professor.

INCENTIVES/REMUNERATION

Depending on the course you are taking some extra credit as homework grade could be offered for your participation at your professor's discretion. Furthermore, at the end of this study you will be able to see the results of the receptive and productive vocabulary levels tests and you will have an idea about their vocabulary size and knowledge.

PUBLICATION STATEMENT

The results of this study may be published in scientific journals, professional publications, or educational presentations; however, no individual subject will be identified.

SUBJECT RIGHTS

1. I understand that informed consent is required of all persons participating in this project.
2. I have been told that I may refuse to participate or to stop my participation in this project at any time before or during the project. I may also refuse to answer any question.
3. Any risks and/or discomforts have been explained to me, as have any potential benefits.
4. I understand the protections in place to safeguard any personally identifiable information related to my participation.
5. I understand that, if I have any questions, I may contact Burcu Mutlu (bmutlu@central.uh.edu) or I may also contact Marta Fairclough (mfairclo@central.uh.edu), faculty sponsor.
6. **Any questions regarding my rights as a research subject may be addressed to the University of Houston Committee for the Protection of Human Subjects (713-743-9204). All research projects that are carried out by Investigators at the University of Houston are governed by requirements of the University and the federal government.**

I have read the contents of this consent form and have been encouraged to ask questions. I have received answers to my questions to my satisfaction. I give my consent to participate in this study, and have been provided with a copy of this form for my records and in case I have questions as the research progresses

**a) I agree to continue with this study.
study.**

b) I disagree to continue with this

Referencias

- Aitchison, J. (1987). *Words in the Mind: An Introduction to the Mental Lexicon*. Oxford: Basil Blackwell.
- Albrechtsen, D., Haastrup, K., & Henriksen, B. (2008). *Vocabulary and writing in a first and second language: Processes and development*. Basingstoke: Palgrave Macmillan.
- Alderson, C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. London/New York: Continuum
- Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 77–117). Newark, DE: International Reading Association.
- Astika, G. G. (1993). Analytical assessments of foreign students' writing. *RELC Journal*, 24, 61-71.
- Azorín Fernández, D., & Climent, B. J. (2005), El diccionario como instrumento didáctico en la enseñanza/aprendizaje de E/LE , Actas del XXI Congreso de la Asociación española de Lingüística Aplicada (AESLA), Publicaciones de la Universidad de Santiago de Compostela, 671-694.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F. (2004). *Statistical analyses for language assessment*. Cambridge: Cambridge University Press.
- Beaudrie, S. M., & Fairclough, M. A. (Eds.). (2012). *Spanish as a heritage language in the United States: The state of the field*. Washington, DC: Georgetown University

Press.

- Beglar, D., & Hunt, A. (1999). Revising and validating the 2000 word level and university word level vocabulary tests. *Language Testing*, 16 (2), 131-162.
- Benítez Pérez, P. (2009). ¿Qué vocabulario hay que enseñar en las clases de español como lengua extranjera? *Monográficos MarcoELE*, 8, 9 – 12.
- Blum-Kulka, S., & Levenston, E. A. (1983). Universals of Lexical Simplification. In C. Faerch & G. Casper (Eds.) *Strategies in Inter-language Communication* (pp.119-139). London and New York: Longman.
- Brown, D. F. (1980). Eight Cs and a G. *Guidelines*, 3, 1-17.
- Brown, C., & Payne, M. E. (1994). Five essential steps of processes in vocabulary learning. TESOL Convention, Baltimore, M. D.
- Canale, M., & Swain, M. (1980). *Approaches to communicative competence*. Singapore.
- Carter, R., & MacCarthy M. (1989). *Vocabulary and language teaching*, London, UK: Longman.
- Cassidy, F. G. (1972). Toward more objective labeling in dictionaries, in J.E. Alatis (Ed.) *Studies in Honor of Albert H. Marckwardt* (pp. 49-56). Washington D.C.: TESOL
- Casso, J. (2010). *Análisis y revisión crítica de los materiales de evaluación de la competencia léxica. Elaboración de un test de vocabulario de nivel umbral*. (Mater's Thesis) . Madrid, España: Universidad de Antonio Nebrija.
- Cervero, M. J. & Pichardo, F. (2000), *Aprender y enseñar vocabulario*, Madrid: Edelsa.
- Chappelle, C. A. (1994). Are C-tests valid measures for L2 vocabulary research? *Second Language Research*, 10 (2), 157-187.

- Coady, J., Magoto, J., Hubbard, P., Graney, J., & Mokhtari, K. (1993). High frequency vocabulary and reading proficiency in ESL readers. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second Language Reading and Vocabulary Learning*, 217-228, Norwood, NJ: Ablex.
- Coady, J., & Huckin, T. (Eds.). (1997). *Second Language Vocabulary Acquisition*, Norwood, NJ: Cambridge University Press.
- Cohen, A. D. (1990). *Language learning: Insights for learners, teachers, and researchers*. Boston, Mass.: Heinle & Heinle.
- Collentine, J. G., Lafford, B. & A. Karp. (2003). The Acquisition of Lexical Meaning by Second Language Learners: An analysis of general research trends with evidence from Spanish. In B. Lafford and R. Salaberry (Eds.) *Studies in Spanish Second Language Acquisition: the State of the Science*, (pp. 130-159). Washington, DC: Georgetown University Press.
- Collentine, J. (2008). A Frequency Dictionary of Spanish: Core Vocabulary for Learners by Mark Davies. *The Modern Language Journal*, 91 (2), 344.
- Corson, D. J. (1995). *Using English Words*. Dordrecht: Kluwer Academic Publisher.
- Davies, M. (2002). *Corpus del Español: 100 million words, 1200s-1900s*. Retrieved from <http://www.corpusdelespanol.org>.
- Davies, M. (2005). Vocabulary Range and text coverage: Insights from the forthcoming Routledge Frequency Dictionary of Spanish. In D. Eddington (ed.), *Selected Proceedings of the 7 th Hispanic Linguistics Symposium* (pp. 106-115) Somerville, MA.
- Davies, M. (2006). *A frequency Dictionary of Spanish, core vocabulary for learners*.

London: Routledge.

Davies, M., & Face, T. L. (2006). Vocabulary coverage in Spanish textbook: How representative is it? In N. Sagarra & A.J. Toribio (Eds.), *Selected proceedings of the 9th Hispanic Linguistics Symposium* (pp. 132- 143). Somerville, MA.

Ellis, N. C., & Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning*, 43, 559-617.

Fairclough, M. (2011). Testing the lexical recognition task with Spanish/English bilinguals in the United States. *Language Testing*, 28(2), 273-297.

Fan, M. (2000). How big is the gap and how to narrow it? An investigation into the active and pasive vocabulary knowledge of L2 learners. *RELC Journal* 31 (2), 105-119.

Fan, Y. M. (2003). Frequency of use, perceived usefulness, and actual usefulness of second language vocabulary strategies: A study of Hong Kong learners. *The Modern Language Journal*, 87(2), 222-241.

Fishman, J. A. (2001). *Can threatened languages be saved?: Reversing language shift, revisited : a 21st century perspective*. Clevedon, England: Multilingual Matters.

Folse, K. S. (2004). *Vocabulary myths: Applying second language research to classroom teaching*. Ann Arbor: University of Michigan Press.

French Allen, V. (1983), *Techniques in teaching vocabulary*, Oxford: Oxford University Press.

Gairns, R., & Redman, S. (1986). *Working with words: A guide to teaching and learning vocabulary*. Cambridge: Cambridge University Press.

Gass, S. M., & Selinker, L. (2008). *Second language acquisition: An introductory course* (3rd ed.). New York: Routledge.

- Godev, C. B. (2009). Word-frequency and vocabulary acquisition: An analysis of elementary Spanish college textbooks in the USA. *Revista de Lingüística Teórica y Aplicada*, 47 (2), 51-68.
- Greve, M., & Van Passel, F. (1971). *Lingüística y enseñanza de lenguas extranjeras*. Madrid: Fragua.
- Gu, Y., & Johnson, P. K. (1996). Vocabulary learning strategies and language learning outcomes. *Language Learning*, 46, 643-697.
- Haahr, M. (1998). Random.Org. Retrieved from <http://www.random.org/>
- Hartmann, R. R. K. (1981). Style values: Linguistic approaches and lexicographical practice. *Applied Linguistics*, 2 (3), 263-273.
- Hatch, E. & Brown, C. (1995). *Vocabulary, semantics and language educations*. Cambridge: CUP.
- Jiménez Catalán, R. M. (2002). El concepto de competencia léxica en los estudios de aprendizaje y enseñanza de segundas lenguas. *Atlantis*, 24 (2), 149-162.
- Kachroo, J. N. (1962). Report on an investigation into the teaching of vocabulary in the first year of English. *Bulletin of the Central Institute of English*, 2: 67-72.
- Kadota, S. (2003). What is mental lexicon?. In S. Kadota (Ed.), *Eigo no mental lexicon: Goi no kakutoku, syori, gakusyuu* [The English mental lexicon] (pp. 2-11). Tokyo: Shohakusha.Academic.
- Kim, H., & Krashen, S. (1998). The Author Recognition and Magazine Recognition Tests and Free Voluntary Reading as predictors of vocabulary development in English as a Foreign Language. *System*, 26, 515-523.
- Koda, K. (1989). The effects of transferred vocabulary knowledge on the development of

- L2 reading proficiency. *Foreign Language Annals*, 22, 529-542.
- Koda, K. (1997). Orthographic knowledge in L2 lexical processing. In J. Coady & T. Huckin (Eds.) *Second Language Vocabulary Acquisition*. Cambridge: Cambridge University Press.
- Koizumi, R. (2005). *Relationships Between Productive Vocabulary Knowledge and Speaking Performance of Japanese Learners of English at the Novice Level*. (Doctoral Dissertation). Japan, Thsukuba: University of Thsukuba
- Krashen, S. D., & Terrell, T. D. (1983). *The natural approach: Language acquisition in the classroom*. Oxford: Pergamon Press.
- Kucera, H. & Francis, W. N. (1967). *Brown corpus of standard American English*. Retrieved from <http://www.hit.uib.no/icame/brown/bcm.html>
- Larrotta, C. (2011). Second language vocabulary learning and teaching: Still a hot topic! *Journal of Adult Education Information Series*, 40 (1), 1-11.
- Laufer, B. (1986). Possible changes in attitude towards vocabulary acquisition research. *International Review of Applied Linguistics in Language Teaching*, 24 (1), 69-75.
- Laufer, B. (1992). How much lexis is necessary for reading comprehension? In P.J.L. Arnaud, & H. Béjoint, (Eds), *Vocabulary and Applied Linguistics* (pp. 126-132). London: Macmillan.
- Laufer, B. (1997). The lexical plight in second language reading: words you don't know, words you think you know, and words you can't guess. In J. Coady & T. Huckin, *Second Language Vocabulary Acquisition* (pp. 20-34). Cambridge: Cambridge University Press.
- Laufer, B. (1998). The Development of Passive and Active Vocabulary in a Second

- Language: Same or Different? *Applied Linguistics*, 19 (2), 255-271.
- Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, 16 (1), 33-51.
- Laufer, B., Elder, C., Hill, K. & Congdon, P. (2004). Size and strength: do we need both to measure vocabulary knowledge? *Language Testing*, 21 (2), 202-226.
- Laufer, B., & Goldstein, Z. (2004). Testing Vocabulary Knowledge: Size, Strength, and Computer Adaptiveness. *Language Learning*, 54 (3), 399-436.
- Lewis, M. (1993). *The lexical approach: The state of ELT and a way forward*. Hove, England: Language Teaching Publications.
- Liceras, J. M., & Carter, D. (2009). La adquisición del léxico. In E. Miguel, & C. Piera (Eds.) *Panorama de Lexicología* (pp. 371-404), Barcelona: Ariel.
- Likert, R. (1932). *A technique for the measurement of attitudes*. New York: The Science Press.
- Linnarud, M. (1986). *Lexis in composition: A performance analysis of Swedish learners' written English*. Lund: CWK Gleerup.
- Long, M., & Richards, J. C. (1997). Series editors' preface. In Coady, J. & Huckin, T. (Eds.), *Second language vocabulary acquisition IX-X*, Cambridge: Cambridge University Press.
- López-Jiménez, M. D. (2014). A critical analysis of the vocabulary in L2 Spanish Textbook. *Porta Linguarum: Revista internacional de didáctica de las lenguas extranjeras*, 22, 163-182.
- López-Mezquita, M. T. (2007). *La evaluación de la competencia léxica: Tests de*

vocabulario. Su fiabilidad y validez. Madrid: Ministerio de Educación y Ciencia, Centro de Investigación y Documentación Educativa.

Marconi, D. (1997). *Lexical competence*. Cambridge, Mass: MIT Press.

Martin, M. A. Z. (2012). Los diminutivos en español: aspectos morfológicos, semánticos y pragmáticos. Los valores estilísticos de los diminutivos y la teoría de la cortesía verbal. In L. Luque Toro, J. F. Medina Montero, & R. Luque (Eds.), *Léxico Español Actual III* (pp. 123- 140.), Venecia: Cafoscarina.

Meara, P. (1980). Vocabulary acquisition: A neglected aspect of language learning. *Language teaching and linguistics*, 13 (4), 221-246.

Meara, P. (1984). The study of lexis in interlanguage. In A. Davies, C. Criper & A. Howatt (Eds.) *Interlanguage*. Edinburgh: Edinburgh University Press.

Meara, P. (1990). A note on passive vocabulary. *Second Language Research*, 6 (2), 150-154.

Meara, P. (1992). *EFL vocabulary tests*. Wales University: Swansea Centre for Applied Language Studies.

Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjaer, J. Williams (Eds.), *Performance and Competence in Second Language Acquisition* (pp. 35-53). Cambridge: Cambridge University Press.

Meara, P. (1997). Towards a new approach to modelling vocabulary acquisition. In N. Schmitt & M. Mc Carthy (Eds.) *Vocabulary: Description, Acquisition and Pedagogy* (pp. 109-121). Cambridge: Cambridge University Press.

Meara, P., & Milton, J. (2003). *X_Lex the Swansea Levels Test*. Newbury: Express.

Meara, P. M. (2010). The bibliometrics of vocabulary acquisition: An exploratory study.

- RELC Journal*, 43 (1), 7-22.
- Melka, T. F. J. (1982), Receptive versus productive vocabulary: A survey. *Interlanguage Studies Bulletin*, 6 (2), 5-33.
- Miller, G. A. (1986). Dictionaries in the Mind. *Language and Cognitive Processes*, 1(3), 171-185.
- Miller, G. A. (1991). Semantic Networks of English. *Cognition*, 41, 197-229.
- Miller, G. A., & Felbaum, C. (1991). Semantic networks of English. In Levin, B. & Pinker, S. (Eds.), *Lexical and conceptual semantics* (pp. 197-229). Oxford. Blackwell.
- Milton, J. (2009). *Measuring second language vocabulary acquisition*. Bristol: Multilingual Matters.
- Ministerio de Educación, Cultura y Deporte. (2002). *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación*. Madrid: MEC/Anaya.
- Mir, M. (2007). La entrevista de competencia oral según las directrices de ACTFL. *XVIII Congreso Internacional de la Asociación para la Enseñanza del Español como lengua Extranjera (ASELE)*, pp. 636-641, Alicante.
- Mokhtar, A. A., Rawian, R. M., Yahaya, M. F., Abdullah, A., Mansor, M., Osman, M. I., Zakaria, A. A., Murat, A., Nayan, S., & Mohamed, A. R. (2010). Vocabulary knowledge of adult ESL learners. *English Language Teaching*. 3(1), 71-80.
- Morris, L., & Cobb, T. (2003). Vocabulary profiles as predictors of the academic performance of Teaching English as a Foreign Language trainees. *System*, 32 (1), 75- 87.
- Nagy, W., & P. Herman. (1987). Breadth and depth of vocabulary knowledge:

- Implications for acquisition and instruction. In M. McKeown & M. Curtiss (Eds.) *The nature of vocabulary acquisition* (pp. 19-35), Hillsdale, NJ: Erlbaum.
- Nation, I. S. P. (1983). Teaching and testing vocabulary. *Guidelines*. 5(1), 12-25.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. New York: Newbury House.
- Nation, I. S. P. (1997). Review of Introducing Language Awareness. In Lier, L (Ed.). *Studies in Second Language Acquisition* (pp. 163-176), London: Penguin.
- Nation, I. S. P., & Waring, R. (1997). Vocabulary size, text coverage, and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6–19). Cambridge: Cambridge University Press.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. New York, NY: Cambridge University Press.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63 (1), 59-82.
- Nation, I. S. P. & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31 (7), 9-13.
- Nation, I. S. P. (2011). *Learning and Teaching Vocabulary: Collected Writings*. Shanghai: Shanghai Foreign Language Education Press.
- Nation, I. S. P. (2012) Teaching language in use: Vocabulary. In M. Eisenmann & T. Summer (Eds.) *Basic Issues in EFL Teaching and Learning* (pp.103-116). University Winter Verlag.
- Nation, I. S. P. (2013). *Learning vocabulary in another language (2nd Ed.)*. New York, NY: Cambridge University Press.
- Núñez Delgado, M.P., & Moral Barrigüete, C. (2010). Competencia léxica y competencia

comunicativa: bases para el diseño de programas didácticos en la educación escolar. *Lenguaje y Textos*, 23, 91-97.

O'Malley, J. M., & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.

Paribakht, T., & Wesche, M. (1997). Vocabulary enhancement activities and Reading for meaning in second language vocabulary acquisition, in J. Coady and T. Huckin (Eds.), *Second Language Vocabulary Acquisition* (pp. 174-200). Cambridge: Cambridge University Press.

Pawley, A. & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J.C Richards & R.W. Schmidt (Eds.). *Language and Communication*, (pp. 191-225), London: Longman.

Peñalver, M. (1991). *La Lingüística y la enseñanza de la lengua española en el Bachillerato*. Granada: Cornares.

Pérez Basanta, C. (1998). La enseñanza del vocabulario. *GRETA: Revista para Profesores de Inglés*, 6 (1), 4-8.

Pérez Basanta, C. (1999). La enseñanza del vocabulario desde una perspectiva lingüística y pedagógica. In S. Salaberri (Ed.). *Lingüística Aplicada para Segundas Lenguas* (pp. 262-306). Almería: Universidad de Almería.

Pérez Basanta, C. (2003). Psycholinguistics and Second Language Vocabulary Teaching. *Indian Journal of Applied Linguistics*. 29 (2), 35-47.

Pérez-Gironés, A. M. & Adán-Lifante, V. (2014). *Más. Español Intermedio (2nd Ed.)*. New York, NY: McGraw Hill.

Pérez-Gironés, A. M., Becher, A. & Casilde, I. (2011). *Puntos de Partida (9th Ed.)*. York,

NY: McGraw Hill.

- Qian, D.D. & Schedl, M. (2004). Evaluation of an in-depth vocabulary knowledge measure for assessing Reading performance. *Language Testing*, 21, 28-52.
- Read, J. (1988). Measuring the vocabulary knowledge of second language learners. *RELC Journal*, 19 (2), 12–25.
- Read, J. (1997). Vocabulary testing. In N. Schmitt & M. McCarthy. *Vocabulary Description, Acquisition, and Pedagogy* (pp. 303-320). Cambridge: Cambridge University Press.
- Read, J. (2005). Applying lexical statistics to IELTS listening test. *Research Notes*, 20, 12-16.
- Read, J. (2007). Second language vocabulary assessment: current practices and new directions. *International Journal of English Studies*, 7(2), 105-125.
- Richards, J. C. (1976). The role of vocabulary teaching. *TESOL Quarterly*, 10, 77-89.
- Rushstaller, S. (2004). Sobre la génesis del diccionario académico. Las Ordenanzas de Sevilla como fuente de material léxico en el Diccionario de Autoridades. In C. Polzin-Haumann & W. Schweickard. *Zeitschrift für romanische Philologie* (pp.106-127), 120 (1).
- Salaberri, A. (1990): El libro de texto: selección y explotación. In Bello et al (Eds.). *Didáctica de las segundas lenguas. Estrategias y recursos básicos*. Madrid: Santillana.
- Saragai, Y., Nation, P., & Meister, G. (1978). Vocabulary learning and reading. *System*, 6, 70-78.
- Saville-Troike, M. (1984). What really matters in second language learning for academic

- achievement? *TESOL Quarterly*, (18) 2, 199-219.
- Schmitt, N. (1997). Vocabulary learning strategies. In Schmitt, N and McCarthy, M. (eds.) *Vocabulary: Description, Acquisition, and Pedagogy*. Cambridge: Cambridge University Press.
- Schmitt, N., & McCarthy, M. (1997), *Vocabulary: description, acquisition and Pedagogy*, Cambridge, UK: Cambridge University Press.
- Schmitt, N. (1998). Tracking the incremental acquisition of second language vocabulary: A longitudinal study. *Language Learning* 48, 2: 281-317.
- Schmitt, N. (1999). The relationship between TOEFL vocabulary items and meaning, association, collocation, and word class knowledge. *Language Testing* 16, 2: 189-216.
- Schmitt, N., Schmitt, D., & Clapham., C. (2001). "Developing and exploring the behaviour of two new versions of the vocabulary levels test." *Language Testing* 18 (1): 55-88.
- Schmitt, N. (2010). *Researching Vocabulary: A Vocabulary Research Manual*. Palgrave Press.
- Schmitt, N., Ng, J.W.C. & Garras, J. (2011). The Word Associates Format: Validation evidence. *Language Testing*, 28 (1), 105-126.
- Shintani, N. (2011). A comparative study of the effects of input-based and production-based instruction on vocabulary acquisition by young EFL learners. *Language Teaching Research*, 15 (2), 137-158.
- Sinclair, J. M., & Renouf, A. (Eds.). (1988). A lexical syllabus for language learning. In

- R. Carter & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 140-158). Harlow: Longman.
- Staehr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition*, 31(4), 577- 607.
- Stoddard, G.D. (1929). An experiment in verbal learning. *Journal of Educational Psychology*, 20, 452-457.
- Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words*. New York: Teachers College, Columbia University.
- Valdés, G. (2000). *Spanish for native speakers: AATSP professional development series handbook for teachers K-16, (Vol.1)*, New York: Harcourt College.
- Webb, S. (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in Second Language Acquisition*, 27, 33-52.
- West, M. (1953). *A General Service List of English Words*. London: Longman.
- West, R. F., Stanovich, K. E., & Mitchell, H. R. (1993). Reading in the real world and its correlates. *Reading Research Quarterly*, 28, 34-50.
- Wolter, B. (2001). Comparing the L1 and L2 Mental Lexicon: A Depth of Individual Word Knowledge Model. *Studies in Second Language Acquisition*, 23. 41-68.
- Xue, G., & Nation, I.S.P. (1984). A University Word List, *Language Learning and Communication*, 3 (2), 215-229.