

Universidad Autónoma de Querétaro  
Facultad de Psicología

## LOS VERBOS MÁS FRECUENTES EN TEXTOS ESCRITOS PARA NIÑOS

### TESIS

Que como parte de los requisitos para obtener el grado de

Maestra en Aprendizaje de la Lengua y de las Matemáticas

#### **Presenta:**

Ana Luisa Forzán De Lachica

#### **Dirigido por:**

Dra. Mónica Alvarado Castellanos

Dra. Pamela Garbus (co-directora)

Santiago de Querétaro

Octubre 2014



Universidad Autónoma de Querétaro  
 Facultad de Psicología  
 Maestría

Los Verbos Más Frecuentes en Textos Escritos para Niños

**TESIS**

Que como parte de los requisitos para obtener el grado de  
 Maestro en Aprendizaje de la Lengua y las Matemáticas

**Presenta:**  
 Ana Luisa Forzán De Lachica

**Dirigido por:**  
 Dra. Mónica Alvarado Castellanos

SINODALES

Dra. Mónica Alvarado Castellanos  
 Presidente

Mónica Alvarado C.  
 Firma

Dra. Karina Hess Zimmermann  
 Secretario

Karina Hess  
 Firma

Dra. Gabriela Calderón Guerrero  
 Vocal

Gabriela Calderón  
 Firma

Dra. Pamela Garbus  
 Suplente

Pamela Garbus  
 Firma

Mtra. Norma Fernández Ortega  
 Suplente

Norma Fernández  
 Firma

[Firma]  
 Nombre y Firma  
 Director de la Facultad

[Firma]  
 Nombre y Firma  
 Director de Investigación y  
 Posgrado

Centro Universitario  
 Querétaro, Qro.  
 Octubre 2014  
 México

## RESUMEN

El propósito de esta investigación fue indagar cuáles y cómo son los verbos (elementos formales y clases semánticas) más frecuentes en los textos proveídos a los niños mexicanos en momentos iniciales de alfabetización. Con base en la metodología de la lexicografía, identificamos el vocabulario fundamental de verbos que estuvo integrado por 155 lemas verbales, todos ellos extraídos de 604 textos literarios contenidos en los libros de lectura de la Secretaría de Educación Pública (para 1o y 2o años de primaria). Encontramos que desde la perspectiva formal los verbos más frecuentes son bisilábicos, con vocal temática *-er*, en presente del modo indicativo, 3ª persona del singular. Por su clasificación semántica, la mayoría de los verbos son de procesos materiales y mentales. Este trabajo es de naturaleza instrumental, de manera que su utilidad está relacionada con decisiones metodológicas para futuros estudios psicolingüísticos.

(Palabras clave: Verbos escritos, análisis de corpus, morfología flexiva, clasificación semántica)

## ABSTRACT

The purpose of this research was to investigate what and how are (formal elements and semantic classes) more frequent verbs that belong to written texts provided to Mexican children in initial stages of literacy. Based on lexicography methodology, we identify the fundamental vocabulary of verbs that consisted of 155 verbal slogans, all extracted from 604 literary texts from reading books provided by the Ministry of Public Education (for 1st and 2nd year of primary school). We found that from formal perspective the most common verbs are bisilabic with thematic vowel *-er*, in present indicative, 3rd person singular. By their semantic classification, most verbs referred to material and mental processes. This work is instrumental, so that its utility is related to methodological decisions from psycholinguistic future studies.

(Keywords: Written verbs, corpus analysis, inflectional morphology, semantic classification)

A Luis, mi hijo.

## **AGRADECIMIENTOS**

A la Dra. Mónica Alvarado, por su paciencia, apoyo y enseñanzas.

A la Dra. Gabriela Calderón, la Dra. Karina Hess, la Dra. Pamela Garbus y la Mtra. Norma Fernández por leer mi trabajo y por sus contribuciones.

A Silvia, Beti, Gaby y Margarita, amigas entrañables que me han acompañado y apoyado en este camino.

A Cinthia, Laura, Mary y Beta, compañeras con quienes compartí el proceso de nuestro trabajo de tesis, por su apoyo, escucha y acompañamiento.

A Ale Aldeco, Cristina Arteaga, Norma y Cristina Morales por el apoyo para la realización de este trabajo.

A mis padres y abuelos, ya que gracias a ellos y su ejemplo de vocación y trabajo logré llegar hasta este punto. A Vivit y Mano por ser parte importante en mi vida.

A mis hermanos, en especial a Alberto, por su apoyo, ejemplo y enseñanzas.

A mi hijo Luis, mi sentido y mi vida.

## TABLA DE CONTENIDO

INTRODUCCIÓN .....	11
CAPÍTULO 1 .....	13
ANTECEDENTES .....	13
Revisión de la Construcción Conceptual del Verbo a través de la Historia.....	13
La noción de verbo en la Antigüedad. ....	14
La noción de verbo en el siglo XII.....	16
La noción del verbo en el siglo XV.....	17
La noción de verbo en el siglo XVII. ....	19
La noción de verbo en el siglo XVIII. ....	20
La noción de verbo en los siglos XIX y XX.....	20
La noción contemporánea de verbo. ....	22
Niveles de significación de los verbos.....	24
El verbo, criterio morfológico. ....	24
El verbo, criterio sintáctico – semántico.....	29
Lexicografía: Estudio de los Patrones del Lenguaje.....	41
Objeto de estudio de la lexicografía y la lingüística de corpus. ....	41
Descripción General del Corpus de Textos Escritos para Niños (C-TEN) .....	44
Los textos del C-TEN. ....	45
CAPÍTULO 2.....	48
METODOLOGÍA .....	48
Planteamiento del Problema .....	48
Diseño del Análisis .....	49
A) Descripción General de los 100 Primeros Verbos del C-TEN.....	49
B) Identificación del vocabulario fundamental.....	51
b.1) Proceso de lematización. ....	51
b.2) Procedimiento para el análisis sintáctico-semántico. ....	53
CAPÍTULO 3.....	59

ANÁLISIS Y DISCUSIÓN DE RESULTADOS .....	59
A) Descripción General del Corpus de Verbos .....	59
B) Vocabulario Fundamental del Sub-corpus de Verbos.....	63
b.1) Lemas que conforman el vocabulario fundamental y su distribución por cuartiles.....	63
b.2.) Vocabulario fundamental del corpus de verbos.....	67
b.3.) Clasificación y análisis sintáctico-semántico del vocabulario fundamental.....	69
CAPÍTULO 4.....	86
CONCLUSIONES Y CONSIDERACIONES FINALES.....	86
BIBLIOGRAFÍA .....	92

## INDICE DE TABLAS

Tabla		Página
1	Aspectos de la Flexión Verbal para el Análisis Formal	22
2	Accidentes Gramaticales	23
3	Tipos de Perífrasis Verbales	25
4	Géneros Textuales que componen en C-TEN	43
5	Estructura de la Clasificación de ADESSE	68
6	Verbos No Personales Excluidos	69
7	Distribución de Lemas por Cuartil y Rango de Ocurrencias	71
8	Verbos del Primer Cuartil	75
9	Inventario de los Verbos del Vocabulario Fundamental	78
10	Tipos de Procesos, Número de Lemas, Frecuencia y Porcentaje Acumulado	80
11	Clases de la Macro-clase Material: No. Lemas, Frecuencia y Porcentaje	82
12	Clases de la Macro-clase Mental: No. Lemas, Frecuencia y Porcentaje	83
13	Clases de la Macro-clase Verbal: No. Lemas, Frecuencia y Porcentaje	84
14	Clases de la Macro-clase Existencial: No. Lemas, Frecuencias y Porcentaje	85
15	Clases de la Macro-clase Relacional: No. Lemas, Frecuencia y Porcentaje	86

16	Clases de la Macro-clase de Modulación: No. Lemas, Frecuencia y Porcentaje	87
17	Resumen de Macro-clases y Clases: No. Lemas, Frecuencia y Porcentaje Relativo y Absoluta	88
18	Los Diez Verbos Más Frecuentes del C-TEN y de ADESSE	

### ÍNDICE DE FIGURAS

Figura		Página
1	Esquema Morfotáctico Verbal	24
2	Pantalla del Analizador Léxico Word Smith Tools (6.0) Programa Concord	52
3	Entrada del Verbo Saber en la Base de Consulta ADESSE	53
4	Porcentaje de los Verbos Más Frecuentes Según Tipo de Conjugación	57
5	Representación Semiótica de la Gramática de la Experiencia del Vocabulario Fundamental	87
6	Comparación entre los Verbos del Vocabulario Fundamental del C-TEN y los Verbos Clasificados por ADESSE	89

## INTRODUCCIÓN

El propósito de la presente investigación fue indagar cuáles y cómo son los verbos (morfología flexiva y clases semánticas) más frecuentes en los textos previstos para los niños mexicanos en momentos iniciales de alfabetización.

El problema de investigación responde a un planteamiento de la lingüística aplicada asociado al desarrollo de un recurso instrumental descriptivo de la lengua escrita a la que están expuestos los niños mexicanos recién alfabetizados que cursan regularmente los grados de 1° y 2° de primaria. La psicolingüística y la psicología no pueden trabajar al margen de las descripciones de la lengua oral y escrita. Por lo que, las descripciones de este estudio eventualmente pueden servir para realizar estudios cognitivos (psicolingüísticos) sobre todo en lo referente al reconocimiento visual de palabra, lectura de palabras en contexto y fluidez lectora y para la toma de decisiones desde la psicología educativa relacionadas a la adquisición de la lengua escrita.

En el primer capítulo se presenta una aproximación al verbo. La primera parte la dedicamos a la revisión histórica de los criterios que han servido para delimitar al verbo. Si bien, nuestro estudio no fue de carácter histórico, la revisión del contexto de origen del concepto del verbo tuvo el propósito de mostrar la complejidad, histórica y actual, para definir el análisis al que son susceptibles estas unidades gramaticales, además facilitar trabajos futuros. También, la evolución de los aspectos exclusivos del verbo nos sirvió para decidir los criterios a considerar para el análisis de los verbos más frecuentes de los textos escritos para niños.

En la segunda parte del capítulo inicial expone una descripción breve de los niveles de significación del verbo: formal (morfológico) y sintáctico-semántico. Por último, abordamos los

aspectos lexicográficos y de lingüística de corpus que contribuyeron para el desarrollo de la investigación.

En el segundo capítulo se describe la metodología que sirvió para el diseño del análisis del núcleo léxico verbal. Como parte del diseño utilizamos el Corpus de Textos Escritos para Niños (C-TEN) (García, 2013) como corpus de referencia. El diseño constó de la definición del corpus de verbos, la identificación del vocabulario fundamental, la realización del análisis formal y semántico verbal.

En el tercer capítulo se muestran los resultados y la discusión del estudio. El análisis de los verbos incluyó la descripción formal (longitud (por número de sílabas)), vocal temática, tipo de conjugación y los accidentes gramaticales). La descripción semántica contiene los tipos de procesos, dominios y marcos conceptuales del núcleo léxico verbal. Adicionalmente, en los anexos ponemos a disposición los distintos inventarios que resultaron del análisis.

El cuarto, y último capítulo, lo destinamos a presentar las conclusiones de este estudio. También en éste señalamos algunas consideraciones que eventualmente pueden servir para futuras investigaciones.

# CAPÍTULO 1

## ANTECEDENTES

Como mencionamos en la introducción, el objetivo de esta investigación fue indagar cuáles y cómo son los verbos (accidentes gramaticales y clases semánticas) más frecuentes en los textos previstos para los niños mexicanos en momentos iniciales de alfabetización. En consecuencia en este capítulo, primero, presentamos una revisión histórica de la evolución del concepto del verbo. Como segundo punto, presentamos una breve descripción de los niveles de significación (morfológico y sintáctico-semántico) del verbo que nos sirvieron para el análisis. En el tercer y último punto, revisamos los planteamientos lexicográficos que contribuyeron para el análisis formal y semántico de los verbos más frecuentes. En este apartado, también, describimos el corpus de referencia de donde extrajimos el corpus de verbos. El primero y segundo punto dan cuenta de las características de los verbos que nos interesaron estudiar; y el tercer punto, aborda la parte metodológica para el trabajo con córpora y la delimitación de los criterios de frecuencia de uso para el diseño del análisis.

### **Revisión de la Construcción Conceptual del Verbo a través de la Historia**

La definición de verbo ha generado grandes debates a través de la historia. Desde los gramáticos grecolatinos en la época clásica hasta nuestros días no hay un acuerdo para el criterio a considerar para su definición. Se ha definido considerando distintos criterios dando como resultado una gran gama de intentos por delimitarlo. En este apartado trataremos de esbozar cómo se han construido distintas definiciones de verbo, desde una perspectiva histórica, a través de algunos representantes de la lengua, para comprender la complejidad del concepto y de su representación.

La revisión histórica que esbozamos a continuación, estuvo basada en los trabajos de González (1978; 2011) de quien extrajimos ideas, épocas, representantes y definiciones que nos sirvieron para advertir la evolución y criterios preponderantes utilizados para caracterizar a la palabra verbo.

#### La noción de verbo en la Antigüedad.

Desde los primeros intentos que conocemos para definir al verbo se han considerado criterios semánticos, morfológicos y sintácticos. Según González (1978) el debate de la definición de verbo comenzó desde la antigüedad con Platón quien distinguió dos partículas que opuso entre sí: 'ónoma' y 'rema'. El término 'rema' se empleó para aludir al sentido de predicado y cuyo elemento central era el verbo.

Aristóteles coincidió con Platón en que el verbo es un elemento central de la predicación y agregó que guarda relación con la temporalidad de la significación en la emisión. Posteriormente, González (1978) informa que Robins asegura que los representantes del estoicismo (cuyo movimiento filosófico en el periodo helenístico tuvo gran difusión en el mundo grecorromano) regresaron a la perspectiva del verbo como predicado de Platón. Además de coincidir con Aristóteles en la dimensión de temporalidad, identificaron el aspecto acabado o inconcluso (p.67).

A Dionisio de Tracia (170 – 90 a.C.), González (1978) afirma que a éste gramático griego que perteneció a la escuela alejandrina, le debemos la definición que se hizo clásica del verbo: parte de la oración sin inflexiones de caso, que admite las de tiempo, persona y número y que expresa una actividad o un estado afectivo (p.67).

González (1978) encontró que Panini (siglo IV a.C.), gramático de la India antigua, propuso una definición coincidente con la de Dionisio de Tracia en cuanto a la flexión verbal (tiempo, persona y número) y agregó la consideración de que el verbo es el núcleo del predicado (p. 67). Tanto Dionisio de Tracia como Panini, coincidieron en incluir los criterios morfológico, sintáctico y semántico para delimitar al verbo. Además, Panini estableció la relación funcional sintáctico-semántica del verbo como parte de una unidad mayor, el enunciado.

En esta misma época, los latinos Varrón (116-27 a.C.), Donato (siglo IV d.C) y Prisciano (siglos V-VI d.C.) consideraron otros elementos para la definición de verbo. Varrón, quien fue nombrado por Julio César como el director de la primera biblioteca pública de Roma, distinguió entre nombre y verbo considerando los accidentes gramaticales. El nombre o sustantivo cuenta con caso pero no con tiempo, al contrario del verbo, que presenta tiempo pero no caso<sup>1</sup>. Por su parte, Donato gramático latino reconocido, consideró además del criterio morfológico, el semántico para describir al verbo. En González (1978) encontramos la definición de verbo de Donato: *pars orationis, cum tempore et persona, sine casu, aut agere aliquid aut pati aut neutrum significans* (p. 70) (parte de la oración, con tiempo y persona, sin caso, que se refiere a hacer cualquier cosa o que se actúa sobre ellos o ningún hacer pero que indica estado). Por último, cabe mencionar que Prisciano elaboró una descripción exclusivamente morfológica del verbo latino.

Los primeros intentos por definir al verbo consideraron, en primera instancia, la oposición con la sustancia, es decir, entre lo dinámico y lo estático, por lo que el criterio

---

<sup>1</sup> Estas propiedades del sustantivo y del verbo sólo aplica a lenguas de casos como latín y el alemán. En español, según Bosque (2010), las diversas relaciones sintácticas se ponen de manifiesto únicamente a través del paradigma de los pronombres personales.

semántico prevaleció en su delimitación. Este rasgo semántico era general y no consideraba la relación con lo formal. Así mismo, los gramáticos grecolatinos consideraron en estas primeras distinciones los aspectos morfológicos: temporales, aspectuales, y también el comportamiento sintáctico del verbo como parte de una unidad mayor. A continuación presentamos las aportaciones de los gramáticos medievales.

#### La noción de verbo en el siglo XII.

En el siglo XII los gramáticos escolásticos enfatizaron el criterio exclusivamente semántico de los verbos, ligado al carácter racional del discurso, como lo muestra Breda (2005) en la siguiente cita:

Escaligero intenta explicar la causación del lenguaje a partir de la naturaleza. Por eso, en el uso busca lo que las palabras, o sea sus marcas formales en cuanto poseedoras de significado, representan con relación al ser. Así, admite en el nombre seis accidentes (como reflejo de las categorías relativas de la naturaleza, que dependen de una categoría absoluta) y da cuenta del proceso de causación de los accidentes a partir de la realidad. Estudia las causas de los nombres y de los verbos (ambos como categorías absolutas de la realidad) y, en mucho menor grado, de las construcciones sintácticas. En su modelo, encajan perfectamente la distinción que efectúa entre caso y declinación; el primero lo utiliza para referirse a la captación y comprensión del fenómeno natural por el entendimiento, mientras que la declinación es la marca o signo de dicha operación mental representada en el lenguaje (p. 309).

Adicionalmente, González (1978) señala los criterios de Tomás de Erfurt para quien “el modo esencial de significar del verbo indica la cosa por su ser y distancia de la sustancia (*verbum est parsoratoris significans per modum esse distantis a substantia*); el modo de

significar por el ser se origina de lo que fluye y de la sucesión; y el modo por la distancia se origina de una propiedad de la esencia determinada” (García, 1960, p.113 en González, 1978). Es decir, el concepto de verbo se define en contraste o con relación a la substancia o al nombre bajo el criterio semántico.

#### La noción del verbo en el siglo XV.

El nacimiento de la gramática de la lengua española se asocia con la conceptualización más compleja del verbo considerando varios criterios interrelacionados. Las primeras gramáticas del español se publicaron en el siglo XV, por lo que se conoce a esta época como el nacimiento de la lingüística española. Comienza con la publicación de la *Gramática castellana* de Nebrija (1492) y concluye con el *Arte de la lengua castellana* de Gonzalo Correas (1626). La gramática de Nebrija incluye en su libro tercero (*La etimología y dicción*) varios capítulos relacionados con la definición del verbo: En el primer capítulo, sobre las diez partes de la oración de la lengua castellana, incluye como una de ellas al verbo; el capítulo décimo lo dedica exclusivamente al verbo y continúa con la descripción en el capítulo undécimo *‘de los circunloquios del verbo’*, en los capítulos duodécimo, decimotercero y decimocuarto expone respectivamente temas *‘del gerundio del castellano’*, *‘del participio’*, *‘del nombre participial infinito’*<sup>2</sup>. Nebrija ya señala el comportamiento atípico de las formas no personales del verbo y le dedicó un capítulo exclusivo a cada una. Las formas no personales han sido parte del debate de los gramáticos sobre la pertinencia de considerarlos dentro de la categoría verbal ya que pueden funcionar como sustantivos, adjetivos o adverbios.

---

<sup>2</sup> Consultado en <http://www.antoniodenebrija.org/libro3.html#10> en marzo de 2014

En su Gramática, Nebrija, resalta la cualidad predicativa del verbo y le otorga centralidad ya que por sí solo puede construir una oración. González (1978) presenta la definición del verbo propuesta por Nebrija:

*Verbo es una de las diez partes de la oración: el cual se declina por modos y tiempos sin casos. E llámese verbo que en castellano quiere dezir palabra: no porque las otras partes de la oración no sean palabras: mas porque las otras sin ésta no hacen sentencia alguna: ésta por excelencia llamóse palabra (p. 68).*

El verbo, pues, es el núcleo del predicado. También en sánscrito, así como en latín y en griego, el verbo podría formar una oración por sí solo (González, 1978, p. 68).

La *Gramática Castellana* de Villalón fue publicada en Amberes en 1558. En ella se señaló que el rasgo semántico de los verbos está atado a los accidentes gramaticales. Gonzáles (1978) reporta que Villalón, siguiendo a Donato, define al verbo de la siguiente manera:

Verbo es una boz que significa hazer, padecer (que llama el Latino action o passion) alguna obra en alguna diferencia de tiempo (...) que hay solas tres diferencias de tiempo por donde se varía el verbo por la lengua castellana (...) segundo es de notar que se varía el verbo comúnmente por cada cual de estas tres diferencias de tiempo por tres personas (...) tercero es de notar que juntamente se varía el verbo por dos números (p. 69).

La definición que aporta Gonzalo Correas, gramático representante del nacimiento de la lingüística española, sigue la línea de Dionisio de Tarcia y los gramáticos latinos. Según Hernando (2011) de quien recibe la mayor influencia es de Nebrija pero a diferencia de él trata el fenómeno verbal con mayor detalle y presta atención a aspectos de la lengua o vocabulario histórico. En la definición que presenta González (1978) de Correas sobresale también el criterio

semántico ligado al morfológico. La definición que propone Correas a mediados del siglo pasado, la presenta González (1978) fue la siguiente: *verbo es aquella palabra que significa el hazer i obrar, i dezir las cosas i ser hechas, i obradas, i decirse: i tiene boz y coxugación. La conxugación se reparte en tiempos, los tiempos en personas, i números, i modos* (p. 70).

Siguiendo a González (1978), la gramática española de Lovaina (1555) definió al verbo recurriendo nuevamente a la diferencia platónica de lo dinámico frente a lo estático. Bartolomé Jiménez Patón (1569-1640), gramático español, dice *que el verbo será aquella parte de la oración cuya naturaleza dura en quanto se exercita la tal cosa, como corre, ama* (González, 1978, p. 70). Nuevamente recurre únicamente al criterio morfológico.

Los gramáticos renacentistas que aquí presentamos compartían el interés de conseguir para la lengua castellana un lugar a la altura del latín, del griego o del hebreo; por lo que se preocuparon por describir exhaustivamente los elementos y las reglas que la rigen. En el caso del verbo, las descripciones partieron de la oposición estática-dinámica de la antigüedad para considerar todos los aspectos que contiene el verbo. También, se les atribuyó a los gramáticos de esta época, la reducción de los tiempos verbales en tres.

#### **La noción de verbo en el siglo XVII.**

Para González (1978) las gramáticas racionalistas en el siglo XVII se caracterizaron por su orientación filosófica, sobre todo de las escuelas francesas de Port-Royal. Los gramáticos racionalistas tuvieron una gran influencia de los trabajos de Descartes para realizar estudios comparativos entre lenguas con el fin de distinguir la unidad de la gramática profunda (p. 70). Los gramáticos racionalistas propusieron una clasificación de las palabras alternativa a las clases gramaticales clásicas. El criterio semántico fue el que utilizaron para la reagrupación. En este sentido, opusieron las palabras que significan objetos del pensamiento y las palabras que

significan maneras del pensamiento, por ejemplo, los sustantivos y pronombres como representantes de los objetos del pensamiento y el verbo como una de las clases de palabras que significan maneras del pensamiento. De acuerdo a lo anterior, González (1978) encontró que definen el verbo como una palabra cuyo principal uso es significar una afirmación *´vox significans affirmationan´* (p. 70).

El criterio semántico-lógico racionalista, por sí solo, no caracteriza suficientemente al verbo. La clasificación propuesta por los gramáticos racionalistas aportó la posibilidad de diferenciar la actividad cognitiva necesaria para concebir las cosas, distinta del juicio sobre las cosas. La clasificación contribuyó a las reflexiones sobre la actividad cognitiva diferenciada para significar las clases de palabras. Pero, como afirma González (1978), esta caracterización es vaga ya que no tiene sustento morfológico, ni sintáctico ni funcional, es decir, el criterio semántico-lógico por sí solo no caracteriza suficientemente el verbo.

#### La noción de verbo en el siglo XVIII.

La gramática académica del siglo XVIII, como lo reporta González (1978), la primera edición de la *Gramática de la lengua castellana* (1771) define al verbo como *“la parte principal de la oración que sirve para significar la esencia, la existencia, la acción, pasión y afirmación de todas las cosas animadas e inanimadas, y el ejercicio de qualquiera facultad que tienen estos casos, o se les atribuye”* (p. 70). Es decir, caracterizaron al verbo desde las perspectivas sintáctica y semántica. Resulta interesante que integraron clases semánticas y consideraron cualidades de la sustancia.

#### La noción de verbo en los siglos XIX y XX.

La gramática funcional del siglo XIX otorgó especial atención al comportamiento sintáctico del verbo y aportó diferentes puntos de vista sobre el criterio semántico. Andrés Bello,

considerado el más importante gramático y lingüista hispano de este siglo, resalta la importancia del verbo al considerarlo como la palabra esencial y primaria de sujeto esto según la Academia Española en 1973 encontrado por González (1978). Además, de acuerdo con González (2011-2012), Bello fue el primero en observar la importancia del comportamiento sintáctico del verbo al reconocerlo como un atributo de la proposición. Además, consideró la relevancia para la lengua de la dinámica de los sintagmas en general y del sintagma verbal en particular. También rechazó el criterio semántico-lógico pues no consideró la forma ni el funcionamiento sintáctico. En la misma línea, Rodolfo Lenz, gramático hispanoamericano, coincidió con Bello en el criterio funcional del verbo y en no incluir la perspectiva semántica al definirlo como fenómeno del sujeto de la siguiente manera *‘es una palabra que añadida a un sujeto, expresa con él un juicio completo e independiente y forma una oración’* (González, 1978, p. 71).

Alonso y Henríquez Ureña aportaron un nuevo punto de vista semántico-lógico al concepto de verbo al definirlo como *‘formas especiales del lenguaje con las que pensamos la realidad como un comportamiento del sujeto’* (González, 1978, p. 71). Es decir, lo caracterizaron bajo la consideración de una actividad cognitiva o forma de pensar la realidad, como un comportamiento que puede ser acción, estado, etc. En este sentido, la definición no incluyó a los verbos copulativos, auxiliares ni no finitos o no personales.

Antoine Meillet (1866-1936), reconocido como uno de los mayores representantes de la lingüística francesa de principio del siglo XX, modificó el punto de vista semántico-funcional al oponer el nombre y el verbo como una característica esencial de las lenguas indoeuropeas. Reportado por González (1978), Meillet en 1948 afirmó que el nombre, indica las cosas, ya se trate de objetos concretos o de nociones abstractas, de seres reales o de especies. Y el verbo, indica los procesos, ya se trate de acciones, de estados o de paso de un estado a otro” (p. 72).

Nuevamente, excluyó a los verbos copulativos que carecen de significado pleno. Varios autores (Muñoz Cortés, 1958; Criado de Val, 1958; Galichet, 1949; Secheyay, 1950) coinciden con la caracterización del verbo como proceso apoyando una caracterización semántica funcional.

Por su parte, Bernard Pottier y Émile Benveniste, dos lingüistas europeos reconocidos del siglo pasado, coincidieron en destacar la importancia del verbo al señalar que es la única clase de palabra que puede realizar un enunciado coherente. Martin (1971) se refiere a que la “teoría nodal del verbo” expuesta por estos dos autores reconoce al “verbo como el signo básico de la oración realizada” (pp. 34-35).

#### La noción contemporánea de verbo.

Los gramáticos y lingüistas contemporáneos coinciden en definir al verbo como el núcleo del sintagma verbal (SV) y observan también los rasgos formales (morfológicos). Así, el hispanista Gómez Torrego (2007), en *La Gramática didáctica del español*, define al verbo de la siguiente manera:

El verbo es una clase de palabra –o categoría léxica- que desde el punto de vista formal se compone de raíz y desinencias, y desde el punto de vista funcional es siempre el núcleo sintáctico del predicado de la oración. En relación con la naturaleza gramatical, el verbo puede llevar o no complementos (p. 137).

Por su parte, Marín (2002), en la Gramática española, coincide con Alarcos y otros lingüistas, en la siguiente definición: *‘Entendemos por verbo aquella clase de palabra susceptible de tener flexión de tiempo, modo, aspecto, número y persona, y que desempeña la función de núcleo del sintagma verbal’*. También afirman que es el verbo el que selecciona el sujeto, y que el sujeto es uno de los actantes del verbo (p. 196).

Según González (2011-2012) en la *Gramática descriptiva de la lengua española*, dirigida por Bosque y Demonte (1999), y la *Nueva gramática de la lengua española* –de RAE y AALE, 2009-, “se describe con profusión la morfología y el comportamiento sintáctico del verbo – con el inevitable uso de criterios semánticos- sin que aparezca en ninguna parte la definición o caracterización de esa clase sintáctica de palabra, categoría gramatical o simplemente clase de palabra” (pp. 185-186), dato que corroboramos en la revisión que realizamos de estas gramáticas.

Como se puede observar en las definiciones contemporáneas del verbo la *Nueva gramática de la lengua española* (2010), Gómez Torrego (2007), Marín (2002), Bosque y Delmonte (1999) coinciden en considerar los criterios formales (morfológicos) y funcionales (comportamiento sintáctico con relación al semántico) del mismo. Más adelante en este capítulo, en el apartado de niveles de significación de los verbos, profundizaremos sobre los criterios funcionales.

La revisión de la evolución de la construcción del concepto del verbo nos aportó pistas para el análisis de los verbos más frecuentes de los textos escritos para niños. Primero, nos permitió darnos cuenta de que debíamos considerar el aspecto morfológico de estos verbos: los accidentes gramaticales y, sobre todo, las variantes flexivas exclusivas de las formas verbales. Lo anterior fue con el propósito de destacar que las unidades mínimas de significado ligadas al esquema verbal presentan una sistematicidad paradigmática que puede ser interesante para construir conceptualizaciones del verbo para estudios psicolingüísticos.

Segundo, llamaron nuestra atención las formas no personales por su funcionamiento: sobre todo las perífrasis verbales que se conforman con un verbo auxiliar conjugado y un verbo no personal (en algunos casos con un nexos preposicional) pero constituyen una sola unidad

léxica y el verbo no personal presenta una función diferente (como adjetivo, adverbio o sustantivo). Tercero, atendimos la categoría semántica para determinar si los verbos frecuentes tienen o no significado pleno. Por último, notamos la necesidad de considerar al verbo como núcleo del sintagma verbal. Esta consideración coincidió con las tendencias de los planteamientos de la lingüística contemporánea. La consideración del verbo como núcleo del sintagma verbal implica una complejidad y extensión importante para su análisis; por lo que en esta tesis observamos al verbo en su comportamiento de núcleo del sintagma verbal (SV) en el nivel semántico (léxico).

Hasta este punto, la lingüística nos aportó elementos para delimitar el tipo de análisis de los verbos más frecuentes de los textos escritos para niños. A continuación retomaremos la lingüística para aclarar qué entendemos por cada rasgo o criterio que nos sirvió para el análisis de los verbos con base en los distintos niveles de significación asociados a las unidades léxicas y sus partículas.

### **Niveles de significación de los verbos.**

#### **El verbo, criterio morfológico.**

Como lo hemos señalado, los verbos son contextos privilegiados para el análisis morfológico, por lo que dedicaremos este apartado para describir la información y significados que se derivan de la estructura morfológica de estos.

La morfología es la parte de la gramática que estudia la estructura de las palabras, las variantes que éstas presentan y el papel gramatical que desempeña cada segmento en relación con los demás elementos que las componen (Bosque, 2010). En este sentido, el corpus de verbos lo realizamos de acuerdo con la morfología sincrónica ya que estudiamos el fenómeno lingüístico del verbo tal como se presenta en los textos escritos para niños que se utilizan en la

alfabetización inicial en México. También se llevó a cabo desde la morfología descriptiva pues presentamos las características y propiedades morfológicas.

El verbo contiene dos tipos de información morfológica: léxica y flexiva. La información léxica, según Bosque (2010) es la que se refiere a la estructura de las palabras y las pautas de construcción y derivación de las mismas (p.6). A diferencia de la información léxica, la flexiva se ocupa de las variaciones de las palabras. En el caso de los verbos, se refiere a los patrones de variación de la raíz y las desinencias que describen los paradigmas flexivos. De estas dos informaciones, nos interesó describir el componente flexivo.

De acuerdo con Alcoba (1999) la flexión verbal puede describirse en categorías según su productividad: formas conjugadas (incluyen los accidentes gramaticales) y formas no conjugadas o no personales (dentro de estas las perífrasis verbales). De acuerdo a las formas de conjugación en: regulares e irregulares. Relacionadas con las dos anteriores, según la vocal temática en: *-ar*, *-er*, *-ir*. Agregamos a estos puntos, un criterio función para describir según su significación en: copulativos y predicativos por el interés de caracterizar los verbos desde la perspectiva semántica. En la siguiente tabla presentamos estos aspectos:

**Tabla 1**  
*Aspectos de la Flexión Verbal Considerados para el Análisis Formal*

Aspectos	Elementos	Términos clave
Categorías productivas	Verbos conjugados	Accidentes gramaticales (tiempo, aspecto, modo, número y persona)
	Verbos no conjugados o no personales	Infinitivo, participio y gerundio. Verbos compuestos y perífrasis verbales
Formas de conjugación Vocal temática	Regulares	
	Irregulares	
	-ar	
	-er -ir	
Según su significación	Copulativos	
	Predicativos	

### Categorías productivas

Las formas conjugadas son todas las formas verbales que aceptan flexiones. Las flexiones que aceptan son los accidentes gramaticales. Los accidentes gramaticales del verbo es la información flexiva que contienen los verbos en español dan cuenta de las variaciones que implican cambio en el contenido. Las variantes que son exclusivas de los verbos son las flexiones de tiempo, modo y aspecto; y las que comparte con los sustantivos y pronombres son el número y la persona. En el siguiente cuadro presentamos una breve descripción de los accidentes gramaticales:

**Tabla 2**  
**Los Accidentes Gramaticales**

Accidente gramatical	Definición
El tiempo	Constituye una información deíctica, en sentido de que se ubica una determinada situación con respecto al momento en que se emite el enunciado (Bosque, 2010, p.6). Los tiempos verbales son presente, pasado y futuro.
El aspecto	Es la categoría -no deíctica- que expresa la estructura interna de las situación y diferencias, por ejemplo, la situación que se inicia ( <i>empezar a estudiar</i> ) de la que se repite ( <i>seguir estudiando</i> ) (Bosque, 2010, p. 7). El aspecto verbal puede ser perfecto (acción acabada como <i>Yo amé</i> ) o imperfecto (acción inconclusa como <i>Yo amaría</i> ).
El modo	Pone de manifiesto en la flexión verbal la actitud del hablante hacia la información que se enuncia (Bosque, 2010, p. 7). El modo puede ser indicativo (acción verbal real y objetiva), subjuntivo (expresa suposición, deseo, voluntad, etc.) e imperativo (formula órdenes).
El número	En el caso de los verbos proporcionan información cuantitativa pero sólo están presentes por la exigencia de concordancia (Bosque, 2010, p. 6) ( <i>los niños corren</i> ). El número puede hacer referencia a singular o plural.
La persona	Da cuenta de la persona, animal o cosa en la que recae la acción. Puede ser 1 <sup>a</sup> , 2 <sup>a</sup> y 3 <sup>a</sup> .

### Patrones de los Esquemas Verbales

En la morfología del verbo se distinguen dos elementos constitutivos: la raíz y los morfemas flexivos. La raíz es la partícula que expresa su significado. Los morfemas flexivos contienen dos sufijos con distinta información: el primero, expresa el tiempo, aspecto, modo, y la segunda, define el número y la persona.

La morfología se vale de la morfotáctica que según Lara (2006) es la característica que tienen las expresiones verbales de una lengua, de ordenar sus morfemas en secuencias determinadas o en varios esquemas secuenciales (p. 65). El esquema verbal en el español se puede representar de la siguiente manera:

<b>(RAÍZ + Vocal temática) + sufijo 1(TAM) + sufijo 2 (NP)</b>			
Ejemplo			
Am	+ á	+ ba	+ mos (pretérito imperfecto)
Am	+ a	+ rá	+ n (futuro)

*Figura 1. Esquema morfotáctico verbal. RAÍZ = base léxica, vocal temática = a, e, i, sufijo 1= informa sobre el tiempo, aspecto y modo, sufijo 2 = es el sufijo de concordancia y se refiere al número y a la persona.*

Las formas no personales o no finitas y perífrasis verbales, como lo mencionamos en la revisión histórica, son formas verbales que se pueden considerar atípicas en cuanto a las funciones que desempeñan.

Las formas no personales a veces son auxiliados por otro verbo conjugado y se consideran como verbos compuestos o como perífrasis verbales. La diferencia estriba en que las perífrasis verbales son “combinaciones sintácticas en las que un verbo auxiliar incide sobre un verbo auxiliado, principal o pleno constituido en forma no personal (infinitivo, participio o gerundio) sin dar lugar a dos predicaciones distintas (*No puedo entrar, iremos considerando cada caso en particular*). De acuerdo con Bosque (2010) el verbo auxiliar suele aparecer conjugado pero puede no estarlo (p. 529). Es decir, lo que determina si es una perífrasis y no un verbo compuesto es si las formas verbales en su conjunto se consideran una unidad léxica.

Según Pottier (1961) auxiliar es todo verbo que es ‘incidente’ de otro verbo en un mismo sintagma verbal. Explica que su función es modificante ya que no aporta una noción nueva a la expresión sino que únicamente expresa una modalidad (p. 325). Por ejemplo: El niño *estaba*

*comiendo* un helado. El primer elemento *estar* modifica al segundo elemento *comer* que aporta la noción nueva o predicativa.

No hay acuerdo entre los gramáticos de si debieran ser incluidos o no las formas personales en la categoría de los verbos. Nuestro propósito no es fijar una postura lingüística al respecto. Lo que nos interesa resaltar son las consecuencias que se podrían observar en estudios cognitivos en la delimitación de la significación con respecto a la ambigüedad funcional.

Hualde, et al. (2010) describen los tipos de perífrasis verbales y lo presentamos en la siguiente tabla:

**Tabla 3**  
*Tipos de Perífrasis Verbales*

Tipo de perífrasis	Subtipo	Fórmulas
Modales	De obligación	haber que + infinitivo haber de + infinitivo tener que + infinitivo deber + infinitivo
	De probabilidad y duda	deber de + infinitivo venir a + infinitivo
Aspectuales (aportan información sobre el desarrollo interno de la acción)	De aspecto ingresivo (acción que va a empezar)	ir a + infinitivo estar a punto de + infinitivo
	De aspecto incoativo (acción en el momento de empezar)	echarse a + infinitivo ponerse a + infinitivo romper a + infinitivo
	De aspecto durativo (acción en pleno desarrollo)	seguir + gerundio estar + gerundio andar + gerundio
	De aspecto resultativo (acción ya acabada)	estar + participio llevar + participio dejar + participio tener + participio

Al igual que los verbos no personales, las perífrasis verbales involucran funciones y significaciones no convencionales de los verbos. Estas características y la regularidad de sus estructuras pueden ser un observable interesante para estudios psicológicos en el proceso de adquisición de la lengua escrita.

### *Formas de conjugación*

El verbo se clasifica por tipo de conjugación en regulares o irregulares. Los verbos regulares son los que no presentan variación en la raíz ni en las desinencias del modelo que siguen. Los verbos irregulares sí presentan variaciones ya sea en la raíz o en las desinencias<sup>3</sup>.

### *Criterio funcional según su significación*

El verbo también se describe según su significación en copulativos y predicativos. Los verbos copulativos unen al sujeto con el atributo, expresan la cualidad del sujeto en el predicado (*Luis es inteligente*). Los verbos predicativos expresan acción o proceso (*Yo camino por las mañanas*).

La caracterización morfológica permitió mostrar la sistematicidad y regularidad de los componentes de la palabra con respecto a su significación. A continuación presentamos el segundo nivel de significación que consideramos en este estudio descrito con el criterio sintáctico-semántico.

### **El verbo, criterio sintáctico – semántico.**

El criterio sintáctico – semántico del que partimos fue que el verbo es siempre el núcleo del sintagma verbal. Como pudimos observar en las definiciones del verbo a través de la historia, la delimitación del verbo utilizando exclusivamente el criterio semántico fue criticada por ser vago pues no consideraba la forma. Por lo tanto, optamos por la perspectiva de la noción del verbo desde la gramática funcional que considera tanto la forma como el contenido (significado).

Como lo mencionamos en la revisión histórica, la lingüística funcional es la que los gramáticos y lingüistas contemporáneos han convenido en considerar como la que aporta

---

<sup>3</sup> Para mayor información sobre los modelos de conjugación revisar Saunces-Torres (2000) en su Diccionario del español, hispanoamericano y dialectal ha fijado 100 modelos de conjugación que incluyen todas las variaciones de formas del verbo en español. Lo interesante de este diccionario es que incluye los verbos semirregulares considerando aspectos ortográficos.

elementos de análisis relevantes para definir al verbo. Halliday (1985), Halliday y Matthiesen (2004, 2006) han desarrollado esta perspectiva lingüística. El modelo lingüístico de estos autores se circunscribe en el uso de la gramática semánticamente motivada a diferencia de la gramática formal. Refieren también los autores que se basa en la noción del “plano de contenido” de Hjelmslev que incorpora la gramática y la semántica. Es decir, construyen una gramática de la experiencia relacionando así el significante con el significado pero considerando también la actividad de significación (función semiótica)<sup>4</sup>.

Desde esta perspectiva de la gramática funcional, el lenguaje construye la experiencia humana. Por un lado, da sentido a nuestra experiencia nombrando fenómenos, construyendo categorías a partir de las similitudes entre los fenómenos que se designan y elaborando taxonomías, con el conjunto de categorías. Por otro lado, la construcción del lenguaje se da acompañado por actuaciones contextualizadas en las que se hace referencia a los fenómenos nombrados. En este sentido, tanto las actuaciones que acompañan el nombrar como las características de los fenómenos nombrados, pueden modificar las acepciones y términos que se

---

<sup>4</sup> Existe entre la significación, el signo y el referente dos niveles de arbitrariedad. El semiológico que es un nivel metalingüístico de los signos (Lara, 1979:175) y el lingüístico que está en el nivel del sistema lingüístico determinado. Según Lara (1979) quien ofrece la explicación más rigurosa del signo para comprender estos niveles es Louis Hjelmslev (1972) en su obra *La estratificación de la lengua*. Lara (1979) explica que Hjelmslev dividió el signo lingüístico en cuatro partes para explicar la función semiótica o metalingüística. Hjelmslev dividió el signo lingüístico en contenido (significado) y expresión (significante) que, a su vez, subdivide en sustancia y forma. Es decir, el contenido tiene dos estratos: la sustancia del contenido y la forma del contenido. Esto mismo considera para la expresión: la forma de la expresión y la sustancia de la expresión. Esto le sirve para explicar que existe una relación propiamente lingüística de la naturaleza del signo que es formal e invariable y que ésta da valor a la sustancia ya sea fonética o escrita (Lara, 1979:175). Es decir, existe una dependencia entre la forma y la sustancia. Con esto Hjelmslev, resuelve las explicaciones lógico – semánticas que son vagas porque no tienen relación con la forma. La relación de dependencia no aparece en la relación entre la forma del contenido y la forma de la expresión. Hjelmslev explica esta relación a través de la actividad significativa (o de significación) y le llama función semiótica. La función semiótica se ubica fuera del signo lingüístico, de lo contrario, afirmaríamos que el significante pudiera determinar parte del significado o viceversa. Entonces, la función semiótica que se refiere a la actividad de significación corresponde a la capacidad de representación del ser humano. La consecuencia de la función semiótica que se traduce en una actividad significativa se concreta en la noción de sentido de Hjelmslev. El sentido es un término que resulta importante para el análisis que realizamos por la clasificación de los verbos que presentamos posteriormente.

emplean. Es por esto que los sintagmas verbales no sólo aluden a una figura o representación de un proceso (algo hecho, sentido, dicho, etc.), también consideran a los participantes y las circunstancias asociadas al proceso. De ahí que Halliday y Matthiessen (2006) definan a la experiencia como la realidad que construimos por nosotros mismos a través de los significados del lenguaje (p. 5).

Halliday y Matthiessen (2006) afirman que la construcción de la experiencia, en el caso de los verbos, que usualmente se considera como un conocimiento, que responde a la forma de taxonomías conceptuales y esquemas, no responde a sus planteamientos. Para estos autores, el tratamiento de la experiencia da cuenta de un significado que es construido en el lenguaje. Conciben la construcción de la experiencia como un sistema semántico a partir de la construcción humana usando el lenguaje como la base de interpretación (Halliday y Matthiessen, 2006, p. 2). Es decir, la construcción de la experiencia que puede ser un término contextual que excede los límites del sistema lingüístico, Halliday y Matthiessen (2004, 2006) lo estudian a través de lo que ellos denominan la gramática de la experiencia que es la gramática del sintagma verbal.

La estructura del sintagma, según Halliday y Matthiessen (2004), está constituida por un elemento central y dos periféricos. El elemento central es el proceso (grupo verbal) que se desarrolla en el tiempo. El elemento que se ubica más cercano al centro son los participantes (grupo nominal haciendo que ocurra o afectados por) involucrados en el proceso. El aspecto periférico son las circunstancias (grupo adverbial o frase preposicional) asociadas con el proceso. En este sentido es como la perspectiva funcional de la gramática de la experiencia considera tanto la forma como el contenido (significado).

*El verbo como representación del tipo de proceso.*

El verbo como representación de tipo de proceso, siguiendo a Halliday y Matthiessen (2004) afirman que una de las impresiones más potentes de la experiencia consiste en la dinámica de los eventos. Estas impresiones están incluidas en el núcleo del sintagma verbal: el verbo. El verbo representa, de forma general, según Halliday y Matthiessen (2004), una figura que se representa a través de lo que sucede, se realiza, se dice, se percibe, etc., relacionado con actantes en cierta situaciones que puede incluir tiempo, espacio, causa o manera. Se puede describir el proceso que representan los verbos considerándolos como unidad léxica porque la circunstancia no está directamente relacionada con el proceso pero sí puede completar su significado. Más adelante retomaremos estas dimensiones y planteamientos como parte de los supuestos teórico-metodológicos de la clasificación sintáctico-semántica que tomamos de base para este estudio. Pero antes haremos un paréntesis para justificar por qué elegimos la clasificación sintáctico-semántica específicamente de ADESSE para realizar el análisis de los verbos más frecuentes de los textos escritos para niños.

*Selección de la clasificación sintáctico-semántica de los verbos: ADESSE.*

Como lo mencionamos anteriormente, las clasificaciones semánticas verbales son múltiples y variadas. Existen una importante cantidad de clasificaciones semánticas del verbo, entre las que destacan las de Dixon (1991), Dowty (1979), Criado de Val (1958), Faber y Marial (1999), Fellbaum (1988), Fillmore (1985), Fillmore, Johnson y Christopher y Petruck (2003), Galichet (1949), Halliday y Matthiessen (2004), Muñoz Cortés (1958), Levín (1995,1993) y Vendler (1967). La multiplicidad y variedad de clasificaciones nos enfrentó al problema de selección de base para nuestro análisis. Consideramos tres proyectos de investigación que han abordado el etiquetado semántico de verbos en español: AnCora (Annotated Corpora)<sup>5</sup>, SenSem

---

<sup>5</sup> <http://clic.ub.edu/ancora/>

(Sentence Semantics: Creación de una Base de Datos de Semántica Oracional)<sup>6</sup> y ADESSE (Alternancias de Diatesis y Esquemas Sintáctico-Semánticos del Español)<sup>7</sup>. Los proyectos formaron parte de las opciones que consideramos para el análisis del núcleo léxico verbal de nuestro corpus de referencia.

Los proyectos de investigación antes mencionados coincidieron en la tarea de la etiquetación de clases semánticas. El proyecto AnCora, desarrollado en la Universidad de Barcelona, consta de dos corpora: un corpus para el catalán (AnCora-CAT) y otro para el español (AnCora-ESP). El corpus para el español estuvo constituido por 400.000 palabras cuya fuente principal son diversos textos periodísticos y en proporción menor de otro corpus. La anotación semántica se basa en la clasificación verbal de Vendler (1967) y de Dowty (1979). La tipología prevee cuatro tipos de eventos: estados, actividades, logros y realizaciones. Estos cuatro tipos generales se dividen en 13 subclases atendiendo a la estructura argumental, las alternancias de diátesis y los papeles semánticos.

El proyecto SenSem, llevado a cabo por el *Grup de Recerca Interuniversitari en Aplicacions Lingüístiques* (GRIAL) de Cataluña, estudió los 250 verbos más frecuentes del español con base en un corpus de 700.000 palabras cuya fuente son textos del Periódico de Cataluña en su versión de internet. La clasificación semántica consta de tres tipos: evento, proceso y estado.

ADESSE es un proyecto que se está desarrollando en la Universidad de Vigo en Santiago de Compostela. El corpus del que partió fue la parte contemporánea del corpus ARTHUS<sup>8</sup> y de

---

<sup>6</sup> <http://grial.uab.es/fproj.php?id=1>

<sup>7</sup> <http://webs.uvigo.es/adesse/>

<sup>8</sup> <http://www.bds.usc.es/>

éste retoma la información sintáctica desarrollada en el corpus Base de Datos Sintácticos del Español Actual (BDS)<sup>9</sup> que contiene el análisis sintáctico y valencial de alrededor de 160.000 enunciados tomados del corpus ARTHUS. El corpus ARTHUS consta aproximadamente de 1.5 millones de palabras cuya fuente principal son textos escritos y en menor proporción orales. Los géneros textuales contenidos son variados (narrativa 45.5%, teatro 17.7%, ensayo 12.6%, prensa 8.4% y oral 15.8%). El fundamento teórico se basa en Dixon (1991), Faber y Mairal (1999), FrameNet, Halliday (1985), Levin (1993), Vázquez (2000), y WordNet. Esta clasificación y base de datos se va actualizando conforme va avanzando la revisión de contextos sintácticos de la Base de Datos Sintácticos del Español Actual (BDS). La jerarquía se presenta en un nivel de abstracción más general que son los seis tipos de procesos o macroclases (Halliday, 1985), 14 dominios conceptuales o clases (Faber et al., 1999; primera versión FrameNet: Johnson, 2001; Dixon, 1991; WordNet: Fellbaum, 1998) y 52 marcos conceptuales, frames o subclases (Levín, 1993; Vázquez, 2000).

Como se puede apreciar en las características de los tres proyectos de investigación, ADESSE presentó cualidades relevantes para el propósito de esta tesis. ADESSE, a diferencia de los otros dos proyectos, parte de un corpus más amplio y sobre todo, con una mayor variedad en los géneros textuales contenidos en su base. La variedad de fuentes del proyecto ADESSE presentó mayor similitud con el contenido del corpus de referencia de esta tesis. Por otro lado, el análisis semántico de ADESSE fue más específico ya que parte de los marcos conceptuales delimitados por los contextos de uso de la unidad léxica verbal para después determinar los niveles más abstractos como el dominio conceptual y el tipo de proceso. Aunado a lo anterior, los presupuestos teóricos y la perspectiva lingüística (funcional) correspondieron con los

---

<sup>9</sup> <http://www.bds.usc.es/corpus.html>

previstos para este estudio. Por último, los modelos teóricos de los que partieron fueron confiables. Por lo tanto, la clasificación semántica que utilizamos para determinar las categorías de este estudio fue la del proyecto ADESSE.

#### *La clasificación verbal en el proyecto ADESSE*

El objetivo general del estudio del proyecto ADESSE fue encontrar las relaciones entre la semántica y la sintaxis de la cláusula. La perspectiva teórica parte de los supuestos de la Gramática Funcional (Halliday, 1985; Halliday y Matthiessen, 2004), de la Gramática Cognitiva (Langacker, 1990) y de la Gramática de Construcciones (Golberg, 1995). García-Miguel, Costas y Martínez (2003) explican que a partir de los supuestos de estas gramáticas en el proyecto ADESSE “partieron de la independencia y compatibilidad semántica entre significado verbal y significado construccional, entendiendo que el significado global de una expresión combina el de elementos léxicos y el de las relaciones semánticas de una manera no determinista, es decir, en un proceso de composicionalidad parcial” (p. 415). Por lo tanto, la clasificación verbal de ADESSE se aparta de una visión lexicalista al considerar que el significado del verbo se completa con la relación recíproca que establece con la construcción.

#### *La clasificación semántica verbal en ADESSE*

Como habíamos mencionado la clasificación de ADESSE se compone de 6 tipos de procesos o macro clases, 14 dominios conceptuales o clases y 52 marcos conceptuales, frames o subclases. La clasificación es jerárquica y va del nivel mayor de abstracción (tipos de proceso) al nivel más concreto (marcos conceptuales). No obstante se presente de esta manera (de mayor a menor nivel de abstracción) el fundamento semántico de la clasificación son los marcos conceptuales.

El establecimiento de clases y subclases verbales se realiza identificando las áreas de la experiencia o dominios cognitivos representados en los verbos. De esta forma, se consideran tanto la perspectiva semántico-conceptual como las relaciones del verbo con entidades actanciales. Es decir, la forma de clasificar unidades verbales atiende, en primer término, a los conceptos similares (marcos conceptuales o frames) que se completan con la participación de las entidades actanciales en el proceso. Albertuz (2007) nos ofrece un ejemplo que explica la forma de clasificar en ADESSE: los verbos *decir, narrar, responder, contar* se agrupan en la clase verbal Comunicación porque suponen conceptos similares de la misma área de la experiencia (transmisión de información) asociada con la participación en el proceso de las entidades actanciales del tipo Emisor, Mensaje o Receptor o Asunto (dijo de Juan que le parece buen estudiante) o Código (habló en inglés). La estructura de la base de consulta de ADESSE presenta por separado la información semántica y los patrones de complementación y sintácticos. Lo anterior, facilitó la consulta de las clases y subclases asignadas a cada verbo que era la información que nos interesaba.

#### *Estructura de la clasificación*

El proyecto sintáctico-semántico de ADESSE está en construcción por lo que sufre modificaciones en la medida que avanzan en la revisión del corpus BDS. El punto de partida para el inicio del trabajo fueron las clasificaciones léxico-verbal de los tipos de procesos de Halliday (1985), los dominios conceptuales de Faber et al. (1999), la primera versión FrameNet: Johnson (2001), Dixon, 1991 y WordNet: Fellbaum (1998) y los marcos conceptuales o frames de Levín (1993) y Vázquez (2000) como lo habíamos mencionado anteriormente. Estas clasificaciones se han ido contrastando con los datos de los contextos de su corpus de referencia (BDS). De los planteamientos originales de los proyectos consultados han modificado y afinado

la clasificación de acuerdo a los contextos de uso. A continuación presentamos definiciones de los niveles de la clasificación semántica de ADESSE.

Según Halliday y Matthiessen (2004, 2006) los tipos de proceso o macroclases son las representaciones o figuras gramaticales de un tipo de experiencia. Todas estas figuras consisten en un proceso que se produce en el tiempo y de participantes directamente involucrados en el proceso; adicionalmente puede haber circunstancias no directamente involucradas en el proceso de tiempo, espacio, causa, modo o de otros tipos. Las siguientes definiciones de las macroclases las tomamos de Halliday y Matthiessen (2004, 2006):

- Material: es la categoría gramatical de los procesos que expresan las experiencias del mundo exterior.
- Mental: es la categoría gramatical de los procesos que expresan las experiencias del mundo interior.
- Relacional: tiene que ver con aprender a generalizar – relacionar un fragmento de la experiencia con otro: categorizar. En gramática se reconocen los procesos que expresan identificación y clasificación.
- Comportamiento: representan manifestaciones externas de elaboraciones internas. Es la actuación externa de los procesos de conciencia y de los estados físicos.
- Verbal: asociada con relaciones simbólicas construidas en la conciencia humana y comunicado en forma de lenguaje.
- Experiencial: son los procesos que dan cuenta de la existencia de los seres.

- *Modulación*<sup>10</sup>: son los procesos que dan cuenta de lo mental e implican una decisión que tiene consecuencias materiales.

El siguiente nivel de la clasificación semántica es el de los dominios conceptuales o clases. Según Albertuz (2007) los dominios conceptuales o clases son distintas conceptualizaciones en una misma área de experiencia general relacionadas con los elementos de la construcción del sintagma verbal. En este sentido, por ejemplo, del tipo de proceso material se distinguen las áreas de la experiencia o dominios conceptuales de espacio, cambio, otros hechos y comportamiento.

El último nivel de la clasificación son las subclases que Fillmore (2003) define como “marcos conceptuales”. Al respecto señala *-frames: schematic representations of the conceptual structures and patterns of beliefs, practices, institutions, images, etc. that provides a foundation for meaningful interaction in a given speech community* (p. 235). Es decir, son los marcos conceptuales que representan generalizaciones en el nivel más concreto. Por ejemplo, la clase espacio se divide en los siguientes frames: desplazamiento (*ir*), localización (*estar*), postura-posición (*sentar*), orientación (*señalar*), manera de movimiento (*mecer*) y unión (*juntar*).

En el apartado de metodología presentamos los tres niveles de la clasificación con las definiciones operativas y ejemplos de cada una. A continuación presentamos una breve explicación de cómo se asignan los dominios y marcos conceptuales.

---

<sup>10</sup> El proceso de Modulación no forma parte del conjunto de procesos descritos por Halliday y Matthiesen (2004, 2006). Este proceso fue agregado por el equipo de investigadores de ADESSE como parte de las adecuaciones realizadas en la revisión de los verbos en los contextos de su corpus de referencia (BDS).

### *La asignación de clases y subclases*

Como lo mencionamos anteriormente, el reconocimiento de clases y subclases está asociada a la identificación de dominios y marcos conceptuales que evoca el verbo como núcleo del sintagma verbal. Se designa la subclase a través de la revisión de cada contexto. Esta operación implica la diferenciación de significados o acepciones a partir de la “composicionalidad parcial”.

La asignación de clases y subclases es una tarea difícil por la naturaleza semántica de la tarea. El significado es un fenómeno complejo que escapa a límites precisos. Los significados léxicos, materia de este estudio, además de ser complejos, son multidimensionales y flexibles. Según Albertuz (2007) la complejidad tiene relación con que los significados no son datos discretos ni con fronteras precisas. La multidimensionalidad se refiere a que un verbo puede representar distintos aspectos experienciales. Por ejemplo, el verbo reír tiene que ver con los procesos fisiológicos (como llorar) pero también se relaciona con el ámbito social y comunicativo (como manifestar). Por último, el carácter flexible de los significados léxicos da cuenta de lo que le aporta el contexto al verbo para completar el sentido (acepción).

Todas las clasificaciones semánticas verbales presentan problemas para la asignación de categorías, ADESSE no es la excepción. Las categorías de ADESSE presenta dificultades como lo explican García-Miguel et al. (2003): “agrupar en la misma clase sentidos verbales entre los que se establecen relaciones de parasinonimia y co-hiponimia, incluyendo los que se oponen como negativo – positivo (*decir – callar*, ambos como comunicación; *saber – ignorar*, ambos clasificados como cognición – conocimiento)” (p.422). Otro problema del proceso de clasificación semántica en el proyecto ADESSE, según García-Miguel et al. (2003) tiene que ver con la estructura jerárquica de la clasificación. Entre otras dificultades, los límites de las clases

semánticas a veces resultan insuficientes. Por ejemplo, el verbo *encontrar* presenta rasgos de la clase de percepción pero a la vez también supone aspectos del dominio conceptual de localización. Pero, tal vez una de los problemas fundamentales es el hecho de que existen muchos verbos polisémicos que presentan una doble clasificación. Por ejemplo, el verbo enseñar se clasifica por el significado “*hacer o dejar ver, mostrar*” en verbos de percepción y también, por el significado de “*hacer saber, transmitir conocimientos para que alguien aprenda*” en verbos de conocimiento. Lo anterior, sólo se puede constatar en la actualización del significado léxico del verbo dentro del sintagma verbal.

Los supuestos teórico-metodológicos de la clasificación sintáctico-semántica de ADESSE, no obstante los problemas asociados a toda clasificación semántica, cumplió con los requerimientos necesarios para el análisis de los verbos más frecuentes de los textos escritos para niños. El criterio funcional de base nos permitió apartarnos de una visión lexicalista y nos aportó una noción de verbo relacionada con el contexto. También, los tres niveles de abstracción: tipo de proceso, dominio cognitivo y marco conceptual, sirvieron como un descriptor semántico complejo de la gramática de la experiencia. Aunado a lo anterior, resultó relevante tener la posibilidad de notar la importancia de algunas áreas de la experiencia incluidos en los textos escritos para niños en momentos iniciales de alfabetización.

Hasta el momento, hemos abordado la revisión histórica de la construcción del concepto de verbo. También los niveles de significación: morfológico y sintáctico-semántico. A continuación presentamos los aspectos lexicográficos y de lingüística de corpus que sirvieron de base para realizar el análisis verbal de este estudio.

## Lexicografía: Estudio de los Patrones del Lenguaje

Dado que el propósito del presente estudio fue indagar cuáles y cómo son los verbos más frecuentes en los textos escritos para niños en momentos iniciales de alfabetización inicial, corresponde ahora explicar el concepto de “frecuencia”. En este estudio la frecuencia tiene implicaciones teórico-metodológicas de dos áreas de estudio. La primera es la lexicografía que es la parte de la lingüística aplicada que se dedica a estudiar el léxico. La segunda, fue la lingüística de corpus que es un enfoque metodológico de la lingüística que basa sus investigaciones en datos obtenidos a partir de muestras reales de uso de la lengua.

Objeto de estudio de la lexicografía y la lingüística de corpus.

### *Lexicografía.*

Halliday (2004b), refiriéndose al inicio de la lexicografía, asegura que el estudio sistemático formal de la lengua no pudo suceder antes del surgimiento de la lengua escrita. Como refiere el mismo autor, con razón, en la tradición de los estudios de la lengua oral hay una larga tradición de estudios teóricos del habla y retóricos desde la antigüedad hasta nuestros días. Pero no fue hasta el surgimiento de la escritura que inician los estudios enfocados en la gramática y el vocabulario. El problema de base que se quiso resolver en los primeros estudios léxicos fue conservar los significados de los textos antiguos en riesgo de perderse por la misma evolución de las lenguas.

La lexicografía está asociada naturalmente, desde sus orígenes, a la preservación de los significados de los textos a través del estudio sistemático de las palabras que los componen, así como a la creación de inventarios léxicos (diccionario o thesaurus) con diferentes criterios para su ordenamiento como el fonológico, semántico (etimológico), etc. dependiendo de las características de la lengua a la que responden estas listas. En el caso de la familia de las lenguas indo-europeas, a la que pertenece el español, los criterios tradicionalmente descritos son

morfológico y semántico. Cabe mencionar que desde los primeros proyectos lexicográficos reportados por Halliday (2004b) en su mayoría fueron compilaciones de textos literarios o de lengua culta que representan la norma de la lengua. No obstante que el quehacer convencional de la lexicología es la elaboración de diccionarios, Lara (2006) asegura que también existe la vertiente que se dedica a realizar estudios de la lengua que no necesariamente concluyen en la elaboración de estos portadores léxicos.

Lara (2014) define la lexicografía como: “una parte de la lingüística aplicada. Depende de la lingüística, sobre todo, en la concepción del signo y significado que orienta el análisis semántico, en el análisis gramatical, en la interpretación de la complejidad del uso social de las palabras y los fenómenos normativos ligados a ella, en los planteamientos cuantitativos y en la manera de formular el sistema de análisis computacional; pero por sí misma no es una ciencia, sino una metodología que ofrece criterios y reglas de trabajo que, cuando se ponen en práctica le dan su existencia real y la convierten en arte, como la define el diccionario” (Lara, 2014).

#### *Lingüística de corpus.*

Como mencionamos anteriormente, la lingüística de corpus es un enfoque metodológico de la lingüística que se dedica al estudio de los patrones de la lengua con base en muestras reales del uso. Ésta ha sufrido un cambio con el surgimiento de las computadoras, tanto para la elaboración de diccionarios como para los estudios lingüísticos que se realizan a partir de corpus. Por ejemplo, en el siglo pasado se desarrollaron proyectos sobre el vocabulario, como el de Jullian y Chang-Rodríguez que elaboraron el diccionario de frecuencias del español en 1964. Las capacidades técnicas y tecnológicas eran distintas a lo que actualmente podemos hacer con el uso de la computadora. Según Leech (1992) la lingüística basada en el análisis de corpus ha tenido un resurgimiento asociado a la aparición de la computadora. Las capacidades técnicas en

1950 para conformar un corpus no se pueden comparar con la capacidad para procesar información de las computadoras aunado al banco de información digitalizada que está a disposición de los investigadores actualmente a través de internet. Las posibilidades de conformación, de manejo de información y análisis de los datos con la computadora son exponencialmente más rápidos y con amplias capacidades. De tal manera ha impactado a la lingüística de corpus y a la lexicografía el uso del ordenador que hoy se conoce una rama como lexicografía computacional.

La lexicografía computacional permitió realizar investigaciones lingüísticas sobre el comportamiento de la lengua a diferencia de los estudios normativos o generales de la lengua. Las investigaciones lingüísticas sobre el comportamiento de la lengua se realizan a partir de corpus. Torruella y Llisten (1999) definen un corpus como una recopilación de textos seleccionados según criterios lingüísticos, codificados de modo estándar y homogéneo, con la finalidad de poder ser tratados mediante procesos informáticos y destinados a reflejar el comportamiento de una o más lenguas (p. 52). El énfasis de este tipo de estudios, como lo afirman los autores, está en el reflejo del comportamiento de la lengua. Así, la definición de Sinclair (1991)<sup>11</sup>, un corpus es a collection of natural occurring language text, chosen to characterize a state or variety of a language (p. 17), abona a las restricciones de la conformación de un corpus. Para que el corpus de referencia refleje el comportamiento de la lengua se requiere que la colección sea de textos reales.

---

<sup>11</sup> John Sinclair coordinó la construcción del corpus denominado COBUILD (Collins Birmingham University International Language Database) que consta de 20 millones de palabras y dio origen al diccionario Collins Cobuild Dictionary of English Language (1987). Este proyecto fue pionero en el uso de corpus electrónico para la compilación del diccionario (Cantos y Sánchez., 2011)

A partir de la lexicografía computacional se han desarrollado c6rpora digitalizados en varias lenguas. Excede al prop6sito de esta tesis revisar la amplia gama de c6rpora conformados en la lengua en general y en el espa1ol en particular. Por ahora nos limitaremos a describir el corpus de referencia que sirvi6 para el an6lisis de este estudio.

#### Descripci6n General del Corpus de Textos Escritos para Ni1os (C-TEN)

##### *El corpus utilizado para el an6lisis de los verbos.*

El corpus que utilizamos para el an6lisis de los verbos en este trabajo fue el Corpus de Textos Escritos para Ni1os (C-TEN) que fue desarrollado para el proyecto denominado: “Las palabras escritas m6s frecuentes al alcance de ni1os mexicanos” cuyo prop6sito fue generar un listado de vocabulario fundamental del espa1ol escrito en textos dirigidos para ni1os mexicanos que sirva como recurso para el estudio de los procesos de aprendizaje de la lengua escrita” (Alvarado, 2012). El prop6sito de la formaci6n del C-TEN se relacion6 con el conocimiento del estado del l6xico para el grupo particular de textos escritos para el p6blico infantil en momentos de la alfabetizaci6n inicial. En este estudio, nos interes6 mostrar del l6xico, los verbos utilizados en estos textos. Cabe mencionar que los verbos son una categor6a fundamental de la lengua, tal vez la m6s intensa y llena de significaci6n, como lo demostramos en la revisi6n hist6rica y las descripciones anteriores.

##### *Tipolog6a del C-TEN.*

Siguiendo la tipolog6a propuesta por Sinclair (1991) y Torruella y Llisterri (1999) el C-TEN se caracteriz6 considerando distintos criterios. El C-TEN, seg6n el criterio del origen de los datos, es un corpus de lengua escrita en oposici6n al corpus oral. Seg6n la distribuci6n de los diferentes g6neros textuales incluidos es un corpus reducido. De acuerdo a la especificidad de los textos el C-TEN es un corpus especializado ya que recoge textos que puedan contribuir a la descripci6n de la lengua escrita para ni1os. Si consideramos los textos escritos para ni1os como

un género, el C-TEN puede ser un corpus genérico. Continuando con el criterio de especificidad, el C-TEN es un corpus cronológico ya que recopila textos de 1993 a 2012. Por último, según la cantidad de texto que recoge de cada documento, el C-TEN es un corpus textual (whole text corpus) pues la compilación incluye, de los textos seleccionados, todos los documentos que los conforman.

### **Los textos del C-TEN.**

Según Lara, Ham y García (1979) uno de los problemas en la conformación de los corpórea es la objetividad para la descripción del léxico, es decir, la cantidad de datos necesaria para estudios lexicográficos. En el caso del C-TEN observamos la objetividad a través de la composición del mismo. El C-TEN está conformado por los libros de texto de lectura para niños de 1º y 2º de primaria editados por la Secretaría de Educación Pública bajo la consideración de que representan un tiraje editorial de materiales escritos de amplia divulgación (con tirajes en promedio de 1,700,000 cuando los tirajes comunes son de 100,000 ejemplares). Además, son entregados al 100% de los niños que cursan educación primaria del país. A su vez, estos libros son una compilación de textos literarios de diferentes autores que corresponden con la tradición cultural universal y del país a los que están expuestos la mayoría de los niños mexicanos.

El C-TEN se compone por un conjunto de 604 textos divididos en seis libros correspondientes a cada grado (1º y 2º de primaria) en sus ediciones 1993, 1997, 2011 y 2012 y dos antologías también para cada grado en su edición de 2011.

Los textos compilados en el C-TEN son de diferentes géneros, autores diversos, de épocas y tradiciones múltiples como los libros de alta difusión descritos por Chartier (2005). En términos cuantitativos de los 604 textos la mayoría son textos literarios (79.97%) y el resto está integrado por textos informativos (19.70%) y otros (0.33%). Los géneros textuales incluidos en la compilación son variados pero coinciden con los textos que tradicionalmente se han destinado

para este momento de la alfabetización (Chartier, 2005; Hannan, 2008; Hürliman, 1982). En la siguiente tabla presentamos la distribución según el género textual de los textos del C-TEN.

**Tabla 4**  
*Géneros Textuales que Componen el C-TEN*

Total de la muestra	100%	Porcentaje por género	100%
Literarios	79.97	Cuentos	53.15
		Canciones, rimas, adivinanzas, juego de palabras, coplas	11.92
		Poemas	8.11
		Leyendas	3.31
		Fábulas	2.81
		Cuentos adaptados a obras de teatro	0.66
Informativos	19.7	Monografías	19.7
Otros	0.33	Crónica	0.16
		Instructivo	0.16

Algunos de los textos incluidos en el C-TEN cuyo origen es la tradición cultural universal y del país (cuentos, fábulas, leyendas, poemas y adaptaciones de cuentos a obra teatral) representan el 63.58% del total de textos. Los textos asociados a la musicalidad de la lengua, juegos de palabras y sabiduría popular (canciones, rimas, coplas, corridos, rondas, pregones y refranes) constituyen el 11.92%. El resto, los textos informativos (monografías), otros (una crónica y un instructivo) sólo el 20.03%.

El conjunto de textos más numeroso es el de cuentos. Algunos de los títulos incluidos están vinculados a la tradición cultural universal y se asocian con los propósitos didácticos y de entretenimiento de los orígenes de los textos escritos para niños (Chartier, 2005; Hanan, 2008): Pulgarcito, Gulliver, Jack, el cazador de gigantes, Caperucita Roja, La Cenicienta, La Bella Durmiente (del Bosque), El Gato con Botas, Robinson Crusoe; además de algunos poemas, rondas y juegos para niños.

El Corpus de Textos Escritos para Niños (C-TEN) en su conjunto lo constituyeron 141,230 palabras totales (tokens) de los que resultaron 14,572 palabras diferentes o gráficas (types).

Las 14,572 palabras gráficas (types) se clasificaron de acuerdo con su naturaleza en léxicas (o de contenido) y de función (o gramaticales)<sup>12</sup>. Si consideramos las categorías gramaticales, las palabras léxicas (verbos, sustantivos, adjetivos y adverbios) conforman el 98.84% y las palabras de función (pronombres, artículos, preposiciones, conjunciones e interjecciones) el resto (1.16%). Esto se explica porque las palabras de función son categorías restringidas en cuanto a las palabras que las componen y las léxicas son ilimitadas. Pero si el porcentaje lo observamos a partir de las ocurrencias, la proporción cambia y resulta un mejor descriptor del C-TEN. En el caso de las palabras léxicas, bajo esta consideración, se obtuvo el 55.60% y las de función 44.40%. Cabe mencionar que los verbos fueron la categoría que resultó con el mayor número de palabras totales (tokens) del C-TEN con el 41.61%.

La frecuencia es un criterio utilizado por la lexicografía para las descripciones de la lengua. De este criterio nos valimos para realizar la descripción tanto morfológica (primer nivel de significación) como la sintáctica-semántica (segundo nivel de significación). En el capítulo de metodología abordamos los procedimientos que llevamos a cabo considerando la frecuencia. Lo anterior con el fin de desarrollar el recurso metodológico que pretende ser útil para investigaciones psicolingüísticas en español sobre todo los relacionados al análisis de la demanda cognitiva que ofrece la categoría gramatical verbal.

---

<sup>12</sup> Para conocer el procedimiento, herramientas y listado de las palabras del corpus según la categoría gramatical, consultar García (2014). En este estudio únicamente presentamos el inventario de verbos.

## CAPÍTULO 2

### METODOLOGÍA

#### **Planteamiento del Problema**

El objetivo del presente estudio fue proporcionar un análisis de los verbos contenidos en el C-TEN con la intención de facilitar datos que puedan resultar eventualmente útiles para estudios futuros en los que se empleen los verbos escritos presentes en textos de alta divulgación para niños que cursan el primer ciclo de la educación básica. En consecuencia, nuestro estudio pretende responder a dos cuestiones centrales:

¿Cuáles y cómo son los verbos más frecuentes en los textos escritos para niños mexicanos en momentos iniciales de alfabetización?

y más específicamente,

¿cuáles son las características sintáctico-semánticas más frecuentes de los verbos que componen el vocabulario fundamental contenidos en el C-TEN?

Para responder a estas preguntas nos planteamos las siguientes tareas:

- a) Elaboración de la descripción general del corpus de verbos.

A partir de las palabras gráficas contenidas en el listado general se realizó: el análisis formal general de los 100 primeros verbos del C-TEN. En este análisis se identificaron verbos no personales y perífrasis verbales; se clasificaron los verbos de acuerdo con su longitud y vocal temática; se describieron de acuerdo con el tipo de conjugación en la que se encontraron los verbos dentro de los textos, así como los accidentes gramaticales

(tiempo, aspecto, modo, número y persona) y el tipo de verbo según su significación (predicativos y copulativos).

b) Identificación del vocabulario fundamental

A partir del corpus del C-TEN, se retomó la base de verbos (tal como aparecían en los textos de base) para: b.1) lematizarlos y, a partir de ello, b.2) identificar el vocabulario fundamental del corpus de verbos.

A partir del vocabulario fundamental se realizó: b.3) el análisis sintáctico-semántico de los verbos.

## **Diseño del Análisis**

A partir de la identificación de metas y tareas, el plan de análisis que seguimos fue el siguiente:

### **A) Descripción General de los 100 Primeros Verbos del C-TEN**

Para este primer análisis lexicográfico, tomamos como base el tratamiento estadístico que diseñó el equipo de investigadores coordinado por Lara (1979) que conformó el Corpus del Español de México (CEMC) bajo los criterios de Ham (1979). Para realizar el primer acercamiento descriptivo del corpus, se identificaron las primeras 100 palabras gráficas (ordenadas de mayor a menor frecuencia). El corte de 100 responde al principio estadístico para el manejo de información masiva en el que, como afirma Ham (1979), “existen propiedades, ponderaciones y resúmenes que permiten sintetizar los datos para describir un fenómeno dentro de los propósitos y capacidades delimitadas” (p. 54).

El propósito para realizar este procedimiento fue conocer los verbos más frecuentes del C-TEN desde una perspectiva formal; es decir, consideramos la forma de aparición de los verbos

en los textos de base: según la conjugación que presentaron, los accidentes gramaticales (tiempo, aspecto, modo, persona y número) y su significación (copulativos y predicativos).

El análisis formal partió del inventario general de verbos ordenado de mayor a menor frecuencia que obtuvimos del analizador léxico WordSmithTools (6.0) con el apoyo del programa WST-WordList. Identificamos y enlistamos los 100 primeros verbos conjugados y también registramos, en otro inventario, los verbos no conjugados o no personales que encontramos en el proceso de búsqueda.

El listado de verbos no conjugados o no personales nos sirvió para determinar la importancia (en términos de número de formas gráficas y ocurrencias de las mismas) de este tipo de verbos y para elaborar la lista para consulta con las frecuencias correspondientes.

En este sentido, realizamos una búsqueda en el C-TEN de las formas típicas perifrásticas verbales. Realizamos las búsquedas a través de fórmulas que diseñamos para alimentar el analizador de concordancias WST-Concord. Estos tres inventarios (el listado de los 100 primeros verbos, los verbos no conjugados y algunos contextos de las perífrasis verbales) los organizamos para su consulta en los anexos. La figura muestra la pantalla de la búsqueda que realizamos para obtener los contextos de los verbos no personales.

N	Concordance	Set	Word #	Sen	Sen	Parz	Parz	Hea	Hea	Sec	Sec	File	Date	%
25	en peligro? ¿Cómo ha ido todo? <b>He estado preocupado</b> . —Aquí todo está		53,005	4,1	189	34	119			0	389	Compilación pai 2012-Nov-08 00	2012-11-08 00:00	37%
26	doctor. Es necesario que el <b>hueso esté colocado</b> exactamente en la		14,094	1,0	209	28	249			0	109	Compilación pai 2012-Nov-08 00	2012-11-08 00:00	10%
27	por la cerradura. El pequeño <b>Joaquín estaba acurrucado</b> debajo de las		2,222	189	159	10	609			0	29	Compilación pai 2012-Nov-08 00	2012-11-08 00:00	2%
28	llamado Jorge el Grande. El rey <b>Jorge estaba preocupado</b> porque su única		122,349	10,	389	34	819			0	879	Compilación pai 2012-Nov-08 00	2012-11-08 00:00	87%
29	alimento preferido es el Krill. El <b>krill está formado</b> por una población de		128,219	10,	259	34	879			0	919	Compilación pai 2012-Nov-08 00	2012-11-08 00:00	91%
30	el agua está embarrado. El <b>lagarto está pisado</b> , mil escamas ¡ten		11,562	899	449	28	179			0	89	Compilación pai 2012-Nov-08 00	2012-11-08 00:00	8%
31	una reverencia. El lagarto El <b>lagarto está mojado</b> , tiene cara de pantano,		11,545	899	309	28	179			0	89	Compilación pai 2012-Nov-08 00	2012-11-08 00:00	8%
32	las alas sirven para volar. El <b>lobo está cansado</b> El lobo feroz está		34,670	2,6	409	28	829			0	259	Compilación pai 2012-Nov-08 00	2012-11-08 00:00	24%
33	, bum, bum, cling clang. Pobre <b>niño, estaba asustado</b> . Necesitaba a su		101,929	8,7	809	34	609			0	729	Compilación pai 2012-Nov-08 00	2012-11-08 00:00	72%
34	tanto cariño. Luego vio que el <b>pajarrillo estaba parado</b> en el candil. El ave lo		109,629	9,3	649	34	689			0	789	Compilación pai 2012-Nov-08 00	2012-11-08 00:00	78%
35	espín: no pueden agarrarlo los <b>perros. Está relacionado</b> con los de nuestra		10,276	809	259	28	139			0	79	Compilación pai 2012-Nov-08 00	2012-11-08 00:00	7%
36	de su disfraz de elefante. <b>Poncho estaba sentado</b> con la cara entre las		77,490	6,3	219	34	369			0	559	Compilación pai 2012-Nov-08 00	2012-11-08 00:00	55%
37	tarareaba una canción. ¿ <b>No puedes estar callado?</b> dijo el papá elefante .		4,245	341	809	21	119			0	39	Compilación pai 2012-Nov-08 00	2012-11-08 00:00	3%
38	. Quiso tomar un poco de agua, <b>pues estaba cansado</b> y sediento de tanto		34,350	2,6	359	28	819			0	249	Compilación pai 2012-Nov-08 00	2012-11-08 00:00	24%
39	la libertad ni las emociones a las <b>que estaba acostumbrado</b> . Entonces		115,889	9,9	969	34	759			0	829	Compilación pai 2012-Nov-08 00	2012-11-08 00:00	82%
40	, y Gulliver, mirando de reojo, vio <b>que estaba rodeado</b> por cientos de		104,979	8,9	729	34	649			0	749	Compilación pai 2012-Nov-08 00	2012-11-08 00:00	74%
41	. Había una vez un conejo <b>que estaba amarrado</b> con un mecate. De		2,429	197	679	11	199			0	29	Compilación pai 2012-Nov-08 00	2012-11-08 00:00	2%
42	no quiere ir a la escuela. <b>No quiere estar encerrado</b> en una clase, frente a		110,729	9,4	249	34	699			0	789	Compilación pai 2012-Nov-08 00	2012-11-08 00:00	78%
43	el ambiente ¿Te das cuenta? ¡El <b>rió está contaminado!</b> ¡Aquí no debe haber		114,669	9,8	809	34	739			0	819	Compilación pai 2012-Nov-08 00	2012-11-08 00:00	81%
44	. Así estás muy bien!. <b>Romeo estaba encantado</b> : nunca le habían		40,363	3,1	339	28	989			0	299	Compilación pai 2012-Nov-08 00	2012-11-08 00:00	28%

Figura 2. Pantalla del Analizador Léxico Word Smith Tools (6.0) Programa Concord.

A partir del listado de los 100 primeros verbos conjugados analizamos la longitud silábica, la vocal temática, el tipo de conjugación (regular e irregular), los accidentes gramaticales (tiempo, aspecto, modo, persona, número) y el tipo de verbo según su significación (predicativo y copulativo). Para realizar esta caracterización nos apoyamos en la *Nueva gramática de la lengua española (Manual)* (Bosque, 2010) y el *Diccionario del verbo español, hispanoamericano y dialectal* (Suances-Torres, 2000). Elaboramos la caracterización formal de los verbos más frecuentes del C-TEN, es decir, de los 100 primeros verbos, con base en la estadística descriptiva.

## B) Identificación del vocabulario fundamental.

### b.1) Proceso de lematización.

El primer procedimiento para la identificación del vocabulario fundamental fue lematizar los vocablos del corpus. Sobre el conjunto de verbos, se estableció la forma infinitiva de los mismos como cabeza léxica o lema. Con ayuda de la herramienta WST-Lemma generamos un listado de los lemas a partir de cada uno con los contextos registrados de cada paradigma verbal.

Cada cabeza léxica o lema mantuvo las ocurrencias de las formas gráficas que las componen, de tal manera que en la determinación de la frecuencia se incluyeron los lemas

organizados de mayor y menor ocurrencia. A partir de este listado de lemas, dividimos por cuartiles el inventario. El primer cuartil nos sirvió también, para conocer desde otro criterio estadístico los verbos más frecuentes. También se realizó una descripción de la distribución de las frecuencias en el listado general.

De acuerdo con el propósito del estudio de describir los verbos más frecuentes del C-TEN, el tratamiento que dimos en un primer momento (listados ordenados de mayor a menor frecuencia, lematización, división por cuartil) respondió a la búsqueda de un criterio estadístico para establecer el vocabulario fundamental para la realización del análisis de los verbos del C-TEN. Para la conformación de los distintos proyectos lexicográficos producidos a partir del Corpus del Español de México Contemporáneo (CEMC) entre ellos el Diccionario del Español de México (DEM), Ham, propuso considerar para la conformación del vocabulario fundamental (los vocablos que incluirían en el Diccionario como representantes de la lengua del español mexicano) la acumulación de los tres primeros cuartiles de frecuencias absolutas del CEMC, que corresponden al 75% de todas las ocurrencias. Esto es bajo la consideración de que estos vocablos se utilizan en el 75% de las emisiones verbales de los mexicanos (Lara, 2007, p. 13). Bajo este criterio el vocabulario fundamental del C-TEN se conformó por 173 verbos (lemas) con frecuencia mayor o igual a 22, que agruparon 14056 formas gráficas (tokens). El otro criterio con base en la frecuencia de aparición igual o mayor a 20 ( $F \geq 20$ ) fue establecido por el *Centre de Recherche et d' Etude pour la Diffusion du Français*. Este criterio se utilizó para definir el vocabulario fundamental del francés. Bajo este criterio incluimos 185 verbos (lemas) que agrupan 14,305 formas gráficas (types). Estableciendo la diferencia de criterios optamos por utilizar el de  $\geq 20$  con el fin de desarrollar el mismo tratamiento del análisis de sustantivos (García, 2014) del C-TEN para posteriores estudios y comparaciones.

Incluimos una cláusula de exclusión para el núcleo léxico o vocabulario fundamental con el que trabajamos en el análisis sintáctico-semántico aunado al criterio de frecuencia de uso. El criterio que consideramos fue excluir los verbos que no tienen significación plena como son los copulativos y los de soporte o de apoyo porque el interés del análisis fue semántico.

### **b.2) Procedimiento para el análisis sintáctico-semántico.**

El procedimiento para la clasificación y análisis sintáctico-semántico del vocabulario fundamental incluyó el uso de la base de datos de verbos del proyecto de Alternancia de Diátesis y Esquemas Sintáctico-semánticos del Español ADESSE, misma que fue introducida en el capítulo anterior. En esta sección enfatizamos los pasos que desarrollamos para la clasificación del vocabulario fundamental.

La estructura de la clasificación de ADESSE consta de 6 macro-clases (tipo de procesos), 14 clases (dominios conceptuales) y 52 subclases (marco conceptual). Las tareas que realizamos para la clasificación fueron:

- reconocimiento y asignación de clases y subclases;
- identificación de la macro-clase; y
- agrupamiento de los niveles de concentración de las clases semánticas.

Para el reconocimiento y asignación de clases y subclases, tomamos como base los datos encontrados en el proyecto ADESSE, hicimos un rastreo de cada uno de los lemas que conforman el vocabulario fundamental de verbos del C-TEN para identificar la subclase correspondiente con los datos de ADESSE. Con ello obtuvimos una descripción general del lema.

En los casos en que el lema presentaba más de un sentido o acepción, el criterio por el que optamos fue el de preferir la acepción con mayor número de ocurrencias. Así por ejemplo,

en el caso del lema *saber* consideramos la acepción relacionada con *conocer* porque, como se puede notar en la siguiente tabla, presentó la mayor frecuencia de uso en la base de ADESSE. En el siguiente cuadro presentamos un ejemplo:

Entrada	Sentido o acepción	Ocurrencias
SABER I	Conocer: tener noticia de algo verdadero o capacidad de hacer algo	3044
SABER II	Tener sabor	11

Figura 3. Entradas del Verbo Saber en la Base de Consulta ADESSE

Una vez identificado el sentido asignado al lema, registramos la subclase reconocida para esa acepción. Con base en este registro identificamos el nivel superior de clasificación, es decir, la clase.

Uno de los problemas fundamentales a los que nos enfrentamos en este nivel de clasificación fue precisamente la multidimensionalidad del significado que se manifestó en la variedad de los componentes experienciales que pueden ser evocados simultáneamente. Por ejemplo, el verbo *encontrar* evoca una acción de percepción pero al mismo tiempo remite a una acción de localización. En estos casos abrimos una categoría que denominamos “intersección” para incluir estos casos. Las intersecciones se presentaron al interior de las clases pero también algunos casos que comparten niveles superiores de abstracción (macro-clases) como en el ejemplo del verbo *encontrar*.

Una vez asignadas las clases y subclases, procedimos a la identificación de las macro-clases.

*Identificación de la macro-clase (tipo de proceso).*

La macro-clase o tipo de proceso es el nivel máximo de abstracción en el plano de la clasificación sintáctico-semántica del aspecto léxico. La identificación de la macro-clase consistió

en la búsqueda del criterio básico de la asignación de subclase y clase en la clasificación jerárquica de base conceptual de ADESSE. Así la clasificación de mayor abstracción la definimos a través de los niveles de mayor concreción que son los marcos y dominios conceptuales a los que responden los verbos de forma más frecuente en los contextos del corpus de la base de datos consultada. La clasificación jerárquica que sirvió para realizar dicha caracterización la presentamos a continuación.

**Tabla 5**

*Estructura de la Clasificación de ADESSE*

MACROCLASE/ clase	Descripción	Ejemplo
MATERIAL	Es la categoría gramatical de los procesos que expresan las experiencias del mundo exterior (Halliday y Matthiesen, 2004, 2006).	
Espacio	Una entidad posee una determinada localización, configuración u orientación espacial, o bien experimenta algún tipo de cambio en su localización, configuración u orientación (ADESSE, 2013).	<i>mover, llevar,</i>
Cambio	Una entidad es creada, destruida o experimenta algún tipo de alteración de sus propiedades físicas (ADESSE, 2013).	<i>pintar, romper, cambiar</i>
Otros hechos	Comprende un conjunto heterogéneo de clases de verbos que se incluyen en esta jerarquía procesos físicos en los que ninguna de las entidades participantes experimenta necesariamente cambios en su constitución, apariencia física o su localización (ADESSE, 2013).	<i>llover, brillar, trabajar</i>
Comportamiento	Un ser vivo (prototípicamente un ser humano) experimenta o realiza un proceso que se manifiesta de forma externa y que tiene que ver con su existencia psico-biológica y/o interpersonal o social (ADESSE, 2013).	<i>portar(se), llorar, casar</i>
MENTAL	Es la categoría gramatical de los procesos que expresan las experiencias del mundo interior (Halliday y Matthiesen, 2004, 2006).	
Sensación	Una entidad capacitada para tener sentimientos o emociones se ve afectada psíquicamente por algo o muestra determinada disposición subjetivo hacia algo (ADESSE, 2013).	<i>temer, atemorizar, gustar</i>
Percepción	Una entidad dotada de órganos sensoriales tiene contacto objetivo a través de éstos con alguna realidad del entorno u obtiene alguna información de este (ADESSE, 2013).	<i>ver, escuchar, mostrar</i>

Cognición	Una entidad dotada de capacidad intelectual realiza cualquier tipo de actividad cognoscitiva objetiva o subjetiva (ADESSE, 2013).	<i>pensar, entender</i>
Elección	Una entidad consciente toma a otra, seleccionándola de entre un conjunto de entidades alternativas posibles, para una determinada finalidad (puede tratarse de una función desempeñada por otra) (ADESSE, 2013).	<i>decidir, elegir, escoger</i>
VERBAL	Asociada con relaciones simbólicas construidas en la conciencia humana y es comunicado en forma de lenguaje (Halliday y Matthiessen, 2004, 2006).	
Comunicación	Una entidad dotada de capacidad comunicativa, transfiere información por medio de cualquier sistema semiótico a otra entidad (ADESSE, 2013).	<i>decir, hablar</i>
Valorar	Una entidad dotada de capacidad comunicativa y de conciencia valora verbalmente una entidad o un hecho por alguna razón o con algún argumento (ADESSE, 2013).	<i>criticar, felicitar</i>
Emisión de sonido	Una entidad (animal o persona) emite sonidos característicos, eventualmente (dependiendo del contexto o de la construcción), pueden servir para comunicar algo (ADESSE, 2013).	<i>aullar, ladrar, cantar</i>
EXISTENCIAL	Es la categoría gramatical de los procesos que dan cuenta de la existencia de los seres (ADESSE, 2013).	
Existencia	Una entidad, una propiedad o un proceso está presente o ausente, existe o carece de existencia (ADESSE, 2013).	<i>ocurrir, suceder</i>
Fase-tiempo	Clase genérica que incluye fase y tiempo (ADESSE, 2013).	<i>empezar, durar</i>
Vida	Un ser vivo (prototípicamente un animal, y de forma especial un ser humano) permanece en un estado o experimenta un proceso o un cambio de estado directamente relacionado con su existencia biológica (ADESSE, 2013).	<i>nacer, morir</i>
RELACIONAL	Es la categoría gramatical de los procesos que ver con aprender a generalizar, relacionar un fragmento de la experiencia con otro, categorizar, identificación y clasificación (Halliday y Matthiessen, 2004, 2006).	
Atribución	Una entidad se encuentra vinculada con otra, sea ésta última una entidad independiente, otra identidad de la entidad inicial, una propiedad, una función o cualquier otro tipo de hecho que permite caracterizarla (ADESSE, 2013).	<i>equivaler, incluir, costar</i>
Posesión	Una entidad tiene como propiedad, parte o pertenencia otra entidad (ADESSE, 2013).	<i>tener, necesitar, recibir, dar</i>
MODULACIÓN	Es la categoría gramáticas de los procesos que dan cuenta de lo mental e implican una decisión que tiene consecuencias materiales (ADESSE, 2013).	

Causación	Una entidad causante causa de modo general que tenga lugar una situación protagonizada por una entidad ´causada´ (ADESSE, 2013).	<i>ayudar, permitir, imponer</i>
Disposición	Una entidad, típicamente humana, muestra cierto estado o disposición en relación a la realización de un evento (ADESSE, 2013).	<i>tratar, atreverse</i>
Aceptación	Una entidad (típicamente un ser humano) acepta o rechaza un objeto, un hecho o un contenido proposicional (ADESSE, 2013).	<i>aceptar, rechazar</i>

---

*Agrupamiento de los niveles de concentración de las clases semánticas.*

A partir de la identificación de las macro-clases (tipos de procesos) con sus respectivas clases (dominio conceptual) y subclases (marco conceptual) de cada uno de los lemas que conforman el vocabulario fundamental del C-TEN procedimos a ordenarlos. El criterio de organización fue primero reunir todos los verbos que coincidieran en cada una de las seis macroclases, los enlistamos de acuerdo con la clase y subclase pero agrupados todos los que pertenecen a cada macro-clase. Con esta información, enlistamos las macro-clases ordenándolas de mayor a menor número de lemas. Cuando hubo coincidencia en el número de lemas el orden se definió de acuerdo con la macro-clase que presentó el mayor número de ocurrencias. Una vez establecido el orden descendente se obtuvo el porcentaje de frecuencia de uso con respecto al vocabulario fundamental y el porcentaje acumulado con esta misma referencia. El porcentaje acumulado nos sirvió de base para elaborar el agrupamiento de los niveles de concentración. Los niveles de concentración que establecimos fueron tres. El nivel superior de concentración representa las clases semánticas con mayor número de verbos y frecuencia de uso del núcleo léxico verbal del C-TEN. Le siguió un nivel intermedio con un número de verbos y rango de frecuencias menor que el superior. Y por último, el nivel inferior que es el de menor representación.

Los procedimientos, la identificación del vocabulario fundamental y el análisis sintáctico-semántico (léxico) nos permitió describir los verbos más frecuentes del C-TEN desde distintos ángulos que pueden ser útiles para estudios de los distintos modelos explicativos de la adquisición de la lengua escrita. En el capítulo siguiente presentamos los resultados en el mismo orden que presentamos en este apartado metodológico.

## CAPÍTULO 3

### ANÁLISIS Y DISCUSIÓN DE RESULTADOS

#### A) Descripción General del Corpus de Verbos

Dedicamos este espacio para describir los 100 primeros verbos (formas gráficas) más frecuentes del C-TEN. De este conjunto nos interesó describir las características de los verbos no conjugados o no personales y las perífrasis verbales. De los verbos conjugados nos interesaron específicamente la longitud (por número de sílaba), la vocal temática, el tipo de conjugación, los accidentes gramaticales y el tipo de verbo según su significación. Presentamos en este apartado inicial, a modo de primer acercamiento, para abonar a la descripción de los verbos del corpus desde la perspectiva formal o primer nivel de significación.

Para realizar este tratamiento estadístico consultamos el listado general del C-TEN que obtuvimos a partir del analizador léxico WST-WordList. En este primer listado aparecieron las 14,572 palabras diferentes (types) organizadas de mayor a menor frecuencia. Del conjunto de las primeras 473 palabras más frecuentes seleccionamos los 100 primeros verbos conjugados<sup>13</sup> Como parte de este conjunto (473 palabras) estaban incluidos 27 verbos no conjugados o no personales (infinitivo, participio y gerundio). Éstos fueron excluidos del listado final. Los verbos no personales que excluimos aparecen enlistados en la siguiente tabla con las ocurrencias de cada uno.

**Tabla 6**  
*Verbos No Personales Excluidos*

---

<sup>13</sup> El Anexo 1 corresponde al listado general de verbos (types) con sus respectivas frecuencias absolutas y relativas. También se pueden distinguir los 100 primeros verbos (types) conjugados y los 27 verbos (types) no conjugados o no personales.

verbos - ocurrencias					
ver	186	vivir	59	leer	36
hacer	162	saber	57	hecho	35
ser	160	dormir	54	nadar	35
comer	122	corriendo	51	cuidado	34
tener	95	estar	51	poder	34
ir	92	volar	48	escribir	33
llegar	87	hablar	44	pasar	33
jugar	69	dar	42		
decir	67	visto <sup>14</sup>	41		
salir	63	dicho	39		

La aparición de 22 verbos en infinitivo, 4 en participio y 1 en gerundio dentro de los verbos con las mayores ocurrencias, señala la presencia de perífrasis verbales que son las “combinaciones sintácticas en las que el verbo auxiliar incide en un auxiliado, principal o pleno constituido en forma no personal sin dar lugar a dos predicaciones” (Bosque, 2010). Con la herramienta WST-Concord realizamos una búsqueda de las fórmulas típicas de las perífrasis de las que resultaron 88.61% aspectuales y 11.39% modales<sup>15</sup>. Considerando que el C-TEN está constituido por 11,790 enunciados, las perífrasis tienen una presencia del 5.81%.

De los 100 primeros verbos conjugados con mayores ocurrencias del C-TEN obtuvimos los siguientes resultados sobre longitud de palabra (silábica), vocal temática, las tipos de conjugación, accidentes gramaticales y significación. Esto se presenta a continuación.

***Longitud (por número de sílabas).***

La longitud por número de sílabas de las palabras gráficas con mayor número de ocurrencias fue de los verbos bisilábicos que representan el 60% y junto con los verbos monosilábicos (15%) suman el 75%. La medida de longitud de los verbos fue difícil de delimitar.

Los cinco primeros verbos de este inventario fueron: *es, dijo, era, había y está*. De estos cinco

<sup>14</sup> *Visto* es un verbo con dos acepciones: ver y vestir. Como teníamos la duda de cuál era la que aparecía con mayor frecuencia en los contextos del corpus C-TEN realizamos una búsqueda con el programa WST-Concord de la que resultó que el total de las ocurrencias respondía a la acepción de ver (ejemplos: había visto, he visto) por lo que comprobamos que son formas no personales (participio).

<sup>15</sup> En el Anexo 2 se pueden consultar algunos contextos donde aparecen perífrasis. Los contextos inmediatos y el contexto general, es decir, el texto donde aparecen.

verbos, cuatro son de apoyo que se actualizan en los textos como verbos compuestos o perífrasis verbales. Señalamos anteriormente la importante presencia de verbos que pueden funcionar como compuestos o participar en una perífrasis verbal. En tal caso, será difícil establecer los límites para realizar el conteo de sílabas ya que, por ejemplo, los verbos y a veces preposiciones que conforman la perífrasis verbal son una sola pieza léxica. Por otro lado, el verbo es una categoría flexiva que presenta variaciones paradigmáticas, por lo que establecer una medida de longitud para un verbo considerando el paradigma al que pertenece resulta complicado.

#### ***Vocal Temática.***

Los resultados que obtuvimos con el criterio de la vocal temática de los 100 primeros verbos no coinciden con los del español general. La clase de verbos *-er* fueron los que presentaron mayor porcentaje con (49%), le siguieron los verbos *-ar* con 28% y por último los verbos *-ir* (23%). Según Alcoba (1999) la clase mayor es la de la vocal temática *-a* en el español porque es la más productiva con más del 90% de los casos.

#### ***Tipo de Conjugación.***

Los verbos más frecuentes por tipo de conjugación en nuestro corpus fueron los irregulares que representaron el 79% del total. Este dato es relevante ya que se comprueba que la irregularidad afecta principalmente a los verbos más frecuentes como en la mayoría de las lenguas (Halpethmath y Sims, 2010:274) como se puede apreciar en la gráfica que presentamos a continuación.

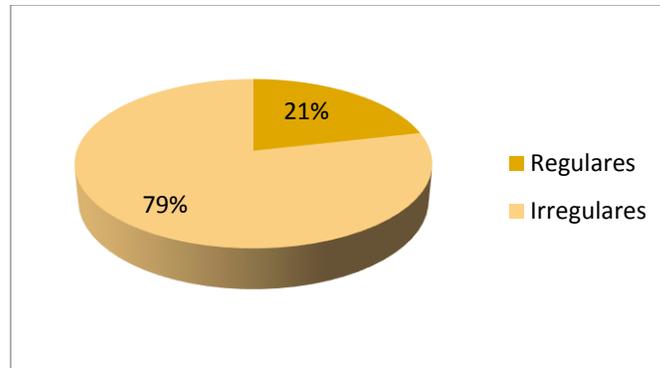


Figura 4. Porcentaje de los Verbos Más Frecuentes Según Tipo de Conjugación

### ***Accidentes Gramaticales.***

#### *Tiempo y Aspecto.*

El tiempo (tense) más recurrente en los verbos del corpus analizado fue el presente con el 58% a diferencia del pretérito que es el tiempo (tense) convencional en los discursos narrativos. Tal vez, este dato se puede explicar con el uso del recurso narrativo (presente histórico) que puede ser utilizado en los textos escritos para niños para crear sensación de mayor cercanía con el oyente o lector.

Los rasgos aspectuales que observamos se presentaron en el 19% del copretérito. El copretérito muestra una gran variedad de estos rasgos. Por ejemplo, puede indicar que una acción es de carácter duradero (*los niños reían mucho*) o que es habitual, se acostumbra o se repite (*en mi infancia nadaba diario*), también que una acción pasada comenzó pero no terminó (*leía cuando llegó Juan*), además que una acción es dudosa, posible, deseable o que sólo sucede en la fantasía (*creía que te divertirías*), por último, expresa con cortesía una acción (*quería pedirte de favor*).

*Modo.*

El modo indicativo fue el que presentó mayor frecuencia con 86%, el modo imperativo siguió con 12% y el menos usual fue el subjuntivo con 2%. Por lo que los verbos más frecuentes de esta muestra respondieron a enunciaciones de acciones verbales reales y objetivas.

*Persona y Número.*

La 3ª persona es la que presenta mayor frecuencia de uso con un 73% a diferencia del 16% de la 2ª persona. Coinciden estos datos con lo que afirman Haspelmath y Sims (2010) de que en general es más importante la 3ª persona que la segunda en la frecuencia de uso de las lenguas. En el caso del número, el singular se presenta en el 95% de las ocurrencias de estos 100 primeros verbos.

*Según su significación.*

Los verbos predicativos son los más frecuentes ya que representan el 84% del total. Los verbos predicativos responden a la definición típica del verbo en la que su significado expresa estado, acción o proceso.

## **B) Vocabulario Fundamental del Sub-corpus de Verbos**

### **b.1) Lemas que conforman el vocabulario fundamental y su distribución por cuartiles.**

Como lo explicamos en el capítulo de metodología, con el fin de tener mejor manejo de los datos y de acuerdo con los procedimientos lexicológicos usados tradicionalmente procedimos a lematizar. Del listado general de verbos de 6040 formas gráficas (types) se obtuvieron 1280 lemas o cabezas léxicas<sup>16</sup>.

---

<sup>16</sup> En el Anexo 3 se puede consultar el árbol de lemas, es decir, el listado de cabezas léxicas con las formas gráficas (types) correspondientes. Se puede enlistar en orden alfabético o por frecuencia (mayor o menor).

***Distribución de lemas por cuartil: ocurrencias y rangos de frecuencia de uso.***

A continuación presentamos los resultados de la distribución de lemas por cuartil y su relación con la frecuencia absoluta.

**Tabla 7**

*Distribución de Lemas por Cuartil y Rango de Frecuencia*

<b>Cuartil</b>	<b>N lemas</b>	<b>% lemas</b>	<b>Rango de ocurrencia</b>
Q1	7	0.56	De 717 a 1977
Q2	24	1.90	De 152 a 691
Q3	82	6.50	De 43 a 151
Q4	1147	90.96	De 1 a 41
Total	1260	100	

La correlación de número de lemas y los rangos de frecuencia en estos resultados se establece una relación progresiva inversa. Se puede notar que la mayoría (90.96%) de los lemas están agrupados en el último cuartil con las frecuencias más bajas (de entre 1 y 41 de frecuencia absoluta). En contraste con el primer cuartil que alberga únicamente siete lemas pero con frecuencias de uso elevadas (entre 717 y 1977 ocurrencias). Es decir, la concentración de lemas tiene la relación siguiente: a menor número de lemas, mayor es la frecuencia de uso y a mayor número de lemas se presentan las frecuencias más bajas en el C-TEN.

En el último cuartil, la presencia de hápax legómena (palabras que aparecen sólo una vez) fue de 411 que representó el 32% de los lemas totales. Según Ham (1979) en los estudios lexicográficos se puede afirmar que los vocablos de los tres primeros cuartiles se utilizan en el 75% del corpus (en este caso del C-TEN) y el resto de vocablos con las menores ocurrencias no son representativos del uso de la lengua que se está estudiando. El hápax legómena es una muestra del vocabulario no usual (poco frecuente) de estos textos.

*Características de los verbos del primer cuartil.*

El interés de este estudio es describir los verbos más frecuentes en los textos escritos para niños por lo que otro acercamiento a estos verbos es la descripción del primer cuartil. La frecuencia de uso de los verbos que aparecen en el primer cuartil es muy alta con respecto a los verbos que aparecen con un uso limitado. Por esta razón, analizamos las características de los 7 verbos que se agrupan en el primer cuartil. A continuación presentamos el listado de estos verbos con la descripción de algunas propiedades.

**Tabla 8**

*Lemas, Frecuencia Absoluta y Número de Formas Diferentes de Verbos del Primer Cuartil*

<b>VERBOS MÁS FREC.</b>	<b>FREC. ABSOL.</b>	<b>% TOT AL</b>	<b>FORMAS</b>	<b>N FORM. DIF.</b>
<b>SER</b> (aux, irreg, -er)	<b>1977</b>	<b>7.13</b>	<i>es, era, fue, soy, eran, eres, será, sea, somos, sido, sería, sean, seas, serán, eras, seremos, éramos, serían, serías, fuera, sé</i>	27
<b>DECIR</b> (irreg, -ir)	<b>1032</b>	<b>3.72</b>	<i>dijo, dice, decir, decía, dicen, dijeron, dije, decían, digo, dile, diciendo, dijiste, dices, diga diré, diciéndole, dime, decirles, díganme, decirme, digan, dirían, decíamos, decimos, decirlo, decirnos, decírselo, dígame, digamos digas, dijera, dijeran, dijeras, dijese, dijimos, diles, dinos</i>	39
<b>HABER</b> (aux, irreg, -er)	<b>990</b>	<b>3.57</b>	<i>había, hay, han, habían, has, haber, hemos, hubiera, hubo, habría, habrá, habrían, hayas, hubieran, habías, habrás, habrías, hubieras, habíamos, haberse, hayan haberlo, haberme, haberte, habidas, habido, habrán, hubiéramos, hubieron</i>	30
<b>ESTAR</b> (aux, reg, -ar)	<b>943</b>	<b>3.40</b>	<i>está, estaba, están, estaban, estoy, estar, estás, estamos, estuvo, estuviera, estábamos, estaría, esté, estuvieron, estabas, estará, estaremos, estarían, estén, estés, estuvieran, estando, estarán, estaríamos, estarías, estarnos, estarse, estemos, estuve, estuviéramos, estuviese, estuvimos, estuviste</i>	33
<b>TENER</b> (aux, irreg, -er)	<b>918</b>	<b>3.31</b>	<i>tiene, tenía, tener, tengo, tienen, tienes, tuvo, tenemos, tenían, tuvieron, tengas, tendrás, tenga, tendré, tendríamos, tuviera, tendrá, tuve, tendremos, tendría, tengan, tenías, ten, tuvimos, tenerme, tenerte, tengamos, teníamos, tuviéramos, tuvieran, tuvieras, tuviese, tuviésemos</i>	36

<b>HACER</b> (irreg, -er)	<b>755</b>	<b>2.72</b>	<i>hacer, hace, hizo, hacen, hacía, haciendo, hicieron, hacerlo, haces, hago, hacían, haremos, hacemos, hice, hacerle, hará, hacerse, haré, hiciste, haría, haz, hacerla, haga, hacerles, hacerlos, hagamos, hagan, hicimos, hacernos, hacías, hazte, hiciera, hicieran, hacerme, hacerte, haciéndole, haciéndolos, haciéndome, hagáis, hagámoslo, hágase, harán, haríamos, harían, harías, hazle, hicieras</i>	<b>48</b>
<b>IR</b> (semi aux, irreg, -ir)	<b>717</b>	<b>2.58</b>	<i>va, ve, iba, ir, voy, vamos, van, iban, vas, fui, irse, ido, vaya, vengo, vámonos, fueran, fuera, íbamos, vayan, vete, iré, iría, vayas, íbas, iremos, irían, irte, yendo, irá, irán,irme, fueras, fuiste, idos, irlos, irnos, váyanse, ve</i>	<b>38</b>
<b>TOTAL DE FORMAS DIFERENTES INCLUIDAS EN EL PRIMER CUARTIL</b>				<b>251</b>

Como se puede notar en la tabla anterior, los 7 verbos que conforman el primer cuartil son verbos auxiliares o de apoyo (*ser, haber, estar, tener*) y verbos semi-auxiliares (*ir y hacer*) con la única excepción de *decir*. Podemos concluir que los verbos más frecuentes de acuerdo con el criterio de segmentación por cuartiles (Q1) con excepción de *decir*, son verbos que como asegura Pottier (1961) tienen una función modificante ya que no aportan una noción nueva a la expresión sino que únicamente expresan una modalidad. Es decir, son verbos que no tienen significado pleno.

Por otro lado, todos los verbos que se agrupan en el primer cuartil son irregulares. Esto coincide con lo que aseguran Haspelmath y Sims (2010) sobre que existe una relación entre la frecuencia de uso y la necesidad de diferenciación en las formas. En la flexión irregular en los verbos más frecuentes se presentan mayor número de formas gráficas distintas en el paradigma verbal y con ello mayor diferenciación como lo muestra el siguiente ejemplo:

1ª sgl	soy	1ª sgl	abandono
2ª sgl	eres	2ª sgl	abandonas
3ª sgl	es	3ª sgl	abandona
1ª pl	somos	1ª pl	abandonamos
2ª pl	son	2ª pl	abandonan
3ª pl	son	3ª pl	abandonan

En este ejemplo presentamos el verbo *ser* que fue el que presentó mayor frecuencia (1977) en el primer cuartil y el verbo *abandonar* que aparece en el último cuartil con 3 ocurrencias. En el verbo *ser*, aunque presenta sincretismo en la 2ª y 3ª de plural igual que en *abandonar*, se pueden apreciar las diferencias en las formas gráficas del resto del paradigma flexivo, lo que no sucede con el verbo *abandonar*.

Por tanto, los datos muestran que los verbos más frecuentes a los que están expuestos los niños en momentos iniciales de alfabetización inicial son funcionales, carecen de significación plena y son irregulares. Esto puede implicar una demanda en los niños para hacer uso de otras estrategias para su identificación, anticipación y significación en el caso de la lectura y en la representación de las variaciones en el caso de la escritura.

El interés de este estudio fue de carácter léxico por lo que las características de los verbos del primer cuartil que no tienen significación plena son relevantes para delimitar con qué verbos realizamos el análisis (sintáctico-semántico) siguiendo el criterio de exclusión de no considerar los verbos copulativos y los de soporte o de apoyo. Los verbos que excluimos del vocabulario fundamental fueron, los verbos copulativos y semi-copulativos: *ser*, *estar*, *parecer* y *seguir* y los de apoyo: *tener*, *hacer*, *dar*, *pegar*<sup>17</sup>, *poder*, *poner* y *tomar*.

#### **b.2.) Vocabulario fundamental del corpus de verbos.**

De acuerdo con el criterio establecido por la frecuencia de aparición igual o mayor a 20 ( $F \geq$ ) establecido por el *Centre de Recherche et d' Etude pour la Diffusion du Français* que se utilizó para definir el vocabulario fundamental de esta tesis incluimos 185 verbos (lemas) con

---

<sup>17</sup> El verbo *pegar* no lo encontramos como de apoyo en las gramáticas generales del Español pero la base de ADESSE lo clasifica de esta manera por lo que decidimos excluirlo.

14,305 ocurrencias. En estos datos ya se excluyeron los verbos copulativos y de apoyo. A continuación se enlista el inventario de verbos que conforma el vocabulario fundamental.

En la siguiente tabla presentamos el inventario de verbos que conforman el vocabulario fundamental<sup>18</sup> respetando el orden de mayor a menor frecuencia absoluta de los lemas.

**Tabla 9**

*Inventario de los Verbos del Vocabulario Fundamental*

Verbos									
1-20	21-40	41-60	61-80	81-90	91-100	101-120	121-140	141-160	161-185
decir	volar	creer	crecer	sentar	pasear	medir	Alcanzar	peinar	golpear
ver	empezar	comenzar	necesitar	terminar	esconder	oler	Cazar	rodear	morder
poder	buscar	caminar	reír	echar	ocurrir	quitar	preocupar	salvar	navegar
llegar	gritar	gustar	sacar	ganar	tejer	casar	Soltar	trépar	partir
saber	caer	saltar	tocar	tratar	visitar	exclamar	Sonar	asomar	repetir
comer	jugar	ayudar	conocer	cuidar	pegar	recoger	Lavar	dirigir	acompañar
salir	responder	mover	esperar	intentar	amar	suceder	Recibir	probar	arrastrar
vivir	contestar	decidir	escribir	descubrir	bailar	construir	Beber	reunir	reconocer
llevar	escuchar	llorar	aprender	formar	escapar	dibujar	Cruzar	doler	
llamar	pedir	oír	usar	meter	robar	sonreír	Detener	flotar	
pasar	volver	soñar	vender	parpadear	utilizar	cerrar	Guardar	importar	
contar	abrir	comprar	convertir	secar	cambiar	colocar	Lanzar	mandar	
venir	correr	acercar	existir	acabar	conseguir	olvidar	Estudiar	Mantener	
pensar	hablar	cortar	nadar	soplar	observar	picar	Ordenar	matar	
encontrar	entrar	cantar	andar	enseñar	permitir	producir	Pesar	quemar	
preguntar	regresar	tirar	levantar	preparar	pintar	alejarse	Pisar	regalar	
dejar	traer	despertar	asustar	viajar	agarrar	imaginar	prometer	atacar	
mirar	servir	subir	bajar	explicar	alimentar	inventar	Romper	atrapar	
sentir	deber	leer	llenar	perder	desaparecer	abrazar	Brincar	brillar	
dormir	trabajar	aparecer	nacer	lograr	entender	aceptar	Invitar	enojar	

Lara (2006) y Cantos, et al. (2011) coinciden en señalar que la frecuencia de determinadas palabras no es fortuita, sino que representa el índice de uso de los conceptos de una

<sup>18</sup> En el Anexo 4 se puede consultar el listado del conjunto de verbos que componen el vocabulario fundamental con sus respectivas ocurrencias.

comunidad. En este sentido, los lemas contenidos en el vocabulario fundamental de verbos es el reflejo del índice de uso de los procesos (Halliday y Matthiesen, 2004, 2006) en los textos escritos para niños.

**b.3.) Clasificación y análisis sintáctico-semántico del vocabulario fundamental.**

El etiquetado semántico de los verbos exigió la adopción teórica y la toma de decisiones en el proceso. Como explicamos en los antecedentes optamos por la clasificación de ADESSE. Bajo esta clasificación el nivel máximo de abstracción en el plano semántico son las macro-clases. Las macro-clases reconocidas en ADESSE (2014) son los procesos, que pueden ser: material, mental, verbal, existencial, relacional y de modulación. Esta clasificación general nos sirvió de punto de partida para organizar los resultados que obtuvimos en la clasificación de los verbos del vocabulario fundamental. El fundamento de la clasificación fueron los dominios cognitivos y los marcos conceptuales. Por lo tanto, primero analizamos las agrupaciones de los niveles inferiores (clases y subclases) que resultan mejores descriptores de los verbos más frecuentes de los textos escritos para niños. Bajo estos niveles se caracterizan los elementos que comparten de la base conceptual más general y amplia. Aun llegando a estos niveles inferiores lo que presentamos en este estudio es una clasificación semántica en un nivel general de la cual se puede partir en el futuro para realizar estudios más particulares. En la tabla 10 presentamos los resultados generales de las macro-clases.

**Tabla 10**  
*Tipo de Proceso, Número de Lemas, Frecuencia y Porcentaje Acumulado*

Etiqueta (tipo de proceso)	N lemas Frecuencia % % acumulado
Material	97 6044 42.09 42.09
Mental	33

	3208
	23.33
	64.42
Verbal	15
	2364
	16.45
	80.87
Existencial	17
	1246
	7.43
	89.51
Relacional	17
	1067
	7.43
	96.94
Modulación	6
	440
	3.06
	100.00

Como podemos observar, de las seis macro-clases el mayor porcentaje de ocurrencias (64.42%) se concentra en los procesos material y mental. El proceso material representa el 42.09% siendo el más frecuente, y el mental, que es el que sigue, abarca el 23.33% del total. A este nivel lo denominamos *superior* para desarrollar a continuación la caracterización de las clases y subclases.

En el nivel intermedio de frecuencia de uso agrupamos a los verbos que corresponden a las macro-clases verbal y existencial. En conjunto, los verbales (16.45%) y existenciales (8.64%) representan el 25.09% del total de verbos del vocabulario fundamental.

Por último, en el nivel inferior de ocurrencias se agrupan los verbos de las macro-clases relacional y de modulación que en conjunto representan alrededor del 10%. Los verbos relacionales se presentan en el 7.43% y los modulares en el 3.06%.

A continuación describimos a mayor profundidad las agrupaciones superior, intermedio e inferior de las macro-clases.

*Nivel Superior de Concentración: Macro-clases Material y Mental.*

La diferencia en cuanto al número de lemas y ocurrencias entre las macro-clases material y mental es amplia. La macro-clase material (97 lemas) supera por 64 lemas a la macro-clase mental (33 lemas). También en el número de ocurrencias la macro-clase material superó casi al doble a la macro-clase mental. Presentaron 6,044 y 3,208 ocurrencias respectivamente. Por ello podemos concluir que la macro-clase material concentra el mayor porcentaje de acumulación tanto de lemas (52.42%) como de ocurrencias (42.09%) de los verbos más frecuentes del C-TEN.

*Macro-clase material.*

En el marco conceptual amplio de la macro-clase de verbos materiales, que resultó ser el de mayor concentración del corpus, identificamos las clases de espacio, otros hechos, cambio y comportamiento. En el vocabulario fundamental del C-TEN los verbos de espacio son los que presentan, a su vez, el mayor número tanto de lemas (48) como de ocurrencias (3006). Este elevado número de ocurrencias representa el 24.91% del vocabulario fundamental, seguido de los verbos caracterizados como de espacio. En orden de mayor a menor representatividad en la macro-clase material, encontramos los verbos de comportamiento (13 lemas y 952 ocurrencias), de otros hechos (20 lemas y 819 ocurrencias) y finalmente los de cambio (16 lemas y 690 ocurrencias). En la siguiente tabla presentamos algunos ejemplos de verbos que pertenecen a cada una de las clases que se agrupan en la macro-clase material<sup>19</sup>.

**Tabla 11**

*Clases de la Macro-clase Material: Número de Lemas, Frecuencia y Porcentaje*

Clase	N lemas	Frecuencia	% de la macro-clase	Verbos (ejemplos)
-------	---------	------------	---------------------	-------------------

<sup>19</sup> En el Anexo 5 se puede consultar el inventario completo de los verbos (lemas) que componen la macro-clase o tipo de proceso Material con sus respectivas clases y subclases.

Espacio	49	3580	59.23	<i>llegar, meter, levantar, entrar, sentar</i>
Comportamiento	13	952	15.75	<i>visitar, casar, dormir, llorar, despertar</i>
Otros hechos	20	819	13.60	<i>jugar, picar, agarrar, usar, sonar</i>
Cambio	16	690	11.42	<i>abrir, cortar, formar, tejer, lavar</i>

Los verbos de la macro-clase material clase espacio, que son los de mayor ocurrencia, se agrupan a su vez en subclases que son: desplazamiento, localización, postura-posición y unión. La subclase en la que se concentró el mayor número de lemas fue en la de los verbos de desplazamiento en la que encontramos 34 lemas con 3006 ocurrencias que representan el 49.73% de la macro-clase y el 20.92% del vocabulario fundamental. Los lemas que se presentaron con mayor frecuencia fueron, por ejemplo, *llegar, salir, llevar, venir y volar*.

Las subclases de la clase denominada de comportamiento son: relaciones, fisiología y de intersección con otra clase. Entre los lemas que se presentaron con mayores ocurrencias de las subclases son los que comparten elementos del marco conceptual con respecto a la fisiología, encontramos por ejemplo: *comer, dormir, llorar, reír y parpadear*. Todavía es posible encontrar elementos comunes en estos lemas, por ejemplo, *dormir, llorar, reír y parpadear* son prototípicos de esta subclase y *comer* tiene que ver con ingestión como *beber y alimentar*.

La clase siguiente de acuerdo con el criterio de mayor a menor frecuencia es la que agrupa los otros hechos. Las subclases de los otros hechos son: actividad, contacto, control, uso, emisión y competición. En la subclase que identificamos más ocurrencias fue en la de actividad

que agrupa a 4 lemas con 312 ocurrencias. Los lemas de esta subclase son: *jugar, trabajar, preparar y bailar*.

La última clase de la macro-clase material que es la que presentó menor número de ocurrencias es la de cambio que representa el 7.61% con respecto al grupo de referencia (macro-clase). Las subclases son: de modificación y de creación. La subclase con mayor número de lemas (11) y ocurrencias (511) fue la de modificación. Algunos ejemplos de lemas que encontramos en esta subclase son: *abrir, cortar, crecer, secar y cambiar*.

El nivel de representatividad con respecto al vocabulario fundamental de las clases de la macro-clase material fue: clase de espacio representa el 24.91%, la de comportamiento el 6.62%, la clase de otros hechos el 5.72% y la de cambio 4.80%.

#### *Macro-clase mental.*

El marco conceptual de los verbos mentales<sup>20</sup> incluye las clases de percepción, de cognición, sensoriales y de elección. De éstas, las clases de percepción y cognición tienen el mismo número de lemas (12) pero con respecto a las ocurrencias los verbos perceptivos (1620) presentan el doble de los verbos que encontramos de cognición (800). Seguido de estas clases en orden de mayor a menor ocurrencia, se presentó la clase de los verbos de sensación (7 lemas y 274 ocurrencias), después la clase general o de los verbos prototípicos de la macro-clase con 2 lemas y 247 ocurrencias. A continuación la clase elección de la que encontramos sólo un lema con una ocurrencia de 85. Por último, el lema *encontrar* con un elevado nivel de ocurrencias (200), se clasifica en la intersección entre mental y material compartiendo las clases de

---

<sup>20</sup> El Anexo 6 corresponde al inventario completo de la macro-clase o proceso mental con sus respectivas clases y subclases.

percepción y localización. Estos verbos que comparten elementos de dos marcos conceptuales pueden resultar interesantes para los estudios cognitivos.

**Tabla 12**

*Clases de la Macroclase Mental: No. Lemas, Frecuencia y Porcentaje*

Clase	N lemas	Frecuencia	% de la macroclase	Verbos (ejemplos)
Percepción	12	1602	49.94	<i>ver, mirar</i>
Cognición	10	800	24.94	<i>saber, conocer, aprender</i>
Sensación	7	274	8.54	<i>gustar, asustar, amar, preocupar, doler</i>
General o prototípica	2	247	8.54	<i>pensar, entender</i>
Elección	1	85	2.65	<i>Decidir</i>
Intersección	1	200	6.23	<i>Encontrar</i>

Los verbos perceptivos representan el 49.93% de la macro-clase mental. Algunos ejemplos de esta subclase de verbos son *ver, mirar, sentir, buscar y escuchar*. Los lemas que presentamos a manera de ejemplos presentan ocurrencias mayores a 100. El lema *ver* presenta la segunda ocurrencia mayor del vocabulario fundamental con 661 después de *decir* (1032).

Los verbos de cognición se presentaron con un porcentaje de 28.94% con respecto a la macro-clase mental. Algunos ejemplos de los lemas más frecuentes que encontramos en esta clase son: *saber, creer, soñar, conocer y aprender*. De estos ejemplos a excepción de *creer* todos comparten elementos en los que la acción significa conocimiento. En el caso de *creer* tiene una significación que da cuenta de un proceso distinto: creencia.

En las clases general o prototípica y sensorial encontramos el mismo número de ocurrencias (274) con diferencia en el número de lemas. En el caso de los verbos sensoriales, el

conjunto de lemas resultó de 7 a diferencia de los verbos prototípicos entre los que únicamente distinguimos dos. Los lemas más frecuentes que se agruparon en los sensoriales, a manera de ejemplo, son: *gustar, asustar, amar, preocupar y doler*. Como se puede notar, estos verbos de actividad mental perceptiva distan mucho de ser todavía un grupo homogéneo. Excede a los propósitos de descripción general de este estudio la profundización de la caracterización de estos y otros verbos pero no dejamos de señalar la importancia de las distinciones y que resultaría interesante para estudios posteriores. Por otro lado, los dos lemas que encontramos en la clase general o prototípica son *pensar* y *entender*. El lema *pensar* tiene una frecuencia alta, aparece 200 veces en los contextos del C-TEN.

Encontramos un lema de la clase de verbos de elección que representa únicamente el 2.65% con respecto a la macro-clase mental. El lema que encontramos fue *decidir* con 85 ocurrencias.

Por último, cabe mencionar que aparecieron dos lemas (*encontrar* y *esconder*) que son verbos de actividad mental y material que implican dos tipos de procesos: uno de percepción y el otro de localización. Como lo mencionamos anteriormente, estos verbos pueden resultar interesantes para los estudios cognitivos porque implican dos acepciones de dos macro-clases diferentes.

*Nivel intermedio de ocurrencia: macro-clases verbal y existencial.*

Las macro-clases de nivel de ocurrencia intermedia fueron similares en cuanto a número de lemas y no así en el número de ocurrencias. De la macro-clase verbal encontramos 15 lemas con 2364 ocurrencias y de la macro-clase existencial distinguimos 17 lemas y 1246 ocurrencias. La macro-clase verbal superó a la existencial a razón del doble de ocurrencias. Si consideramos

que la mayoría de los textos contenidos en el C-TEN fueron literarios, tal vez, esto explique la relevancia de uso de los procesos verbales.

*Macro-clase verbal.*

El conjunto de verbos que encontramos en la macro-clase verbal<sup>21</sup> la mayoría (93.78%) se caracterizan por tener la significación general o prototípica de referirse a eventos de comunicación. Entre ellos distinguimos, por ejemplo, *decir, contar, preguntar, gritar y responder*. Todos los ejemplos que aquí mostramos presentan ocurrencias mayores a 120. Cabe mencionar que *decir* fue el lema con mayor número de ocurrencias del vocabulario fundamental con 1032. Lo anterior, se puede explicar porque es un marcador discursivo de uso en textos de este tipo (principalmente narrativos) y también puede ser que como son textos escrito este verbo permite señalar el discurso directo. Encontramos un lema distinto en cuanto a la significación que es *cantar* (con 77 ocurrencias) que da cuenta de emisión de sonido. También en esta clase encontramos un verbo *escribir* que comparte elementos de las macro-clases verbal y material que se incluyen en las clases comunicación y creación.

**Tabla 13**

*Clases de la Macro-clase Verbal: No. Lemas, Frecuencia y Porcentaje*

Clase	N lemas	Frecuencia	% de la macro-clase	Verbos (ejemplos)
Comunicación	13	2217	93.78	<i>decidir, contar, preguntar, gritar, responder</i>
Emisión de sonido	1	77	8.26	<i>Cantar</i>
Intersección	1	70	2.96	<i>Escribir</i>

<sup>21</sup> El Anexo 7 corresponde al listado completo de los verbos que se incluyen en la Macro-clase o Proceso Verbal con sus respectivas clases y subclases.

*Macro-clase existencial.*

La macro-clase de verbos existenciales<sup>22</sup> se constituye por tres clases: general o prototípica, fase-tiempo y vida según la clasificación que elegimos para realizar el análisis. Encontramos lemas de cada una de estas categorías con número de ocurrencias similares. La clase general o prototípica representa el 36.03%, la clase fase-tiempo 35.31% y la clase de vida el 28.65%. En la clase que encontramos más lemas fue en la general (7) y en la que menos fue en la de vida (4). Los lemas que encontramos en la clase prototípica, por ejemplo, son: *pasar, existir, ocurrir, aparecer, desaparecer*. El primer lema de este conjunto apareció con una ocurrencia alta de 230. Entre los lemas que distinguimos de fase-tiempo estuvieron los verbos aspectuales que dan cuenta, por un lado, de una fase: *empezar, comenzar, terminar, acabar* y, por otro, los que hacen referencia al tiempo: *esperar* y *mantener*. Para el caso de los lemas de la clase de vida encontramos cuatro que son: *vivir, nacer, matar* y *salvar*.

**Tabla 14**  
*Clases de la Macro-clase Existencial: No. Lemas, Frecuencia y Porcentaje*

Clase	N lemas	Ocurrencias	% de la macro-clase	Verbos (ejemplos)
General o prototípica	7	449	36.03	<i>pasar, existir, ocurrir, aparecer, desaparecer</i>
Fase-tiempo	6	440	35.31	<i>empezar, comenzar, esperar, terminar, acabar</i>
Vida	4	357	28.61	<i>vivir, nacer, matar, salvar</i>

*Nivel inferior de ocurrencia: macro-clases relacional y de modulación.*

Las macro-clases relacional y de modulación representan alrededor del 10% de las ocurrencias del vocabulario fundamental. De este 10% la macro-clase relacional representa el

<sup>22</sup> El Anexo 8 corresponde al inventario completo de verbos que componen la Macro-clase o Proceso Existencial con sus respectivas clases y subclases.

7.43% de ocurrencias y la de modulación sólo el 3.06%. La macro-clase relacional supera a la de modulación en número de lemas por 9 lemas así como en ocurrencias por 627 ocurrencias.

*Macro-clase relacional.*

En el caso de los verbos que encontramos de la macro-clase relacional<sup>23</sup> se distribuyen a razón de la mitad de acuerdo con las siguientes clases: posesión con el 53.80% y atribución con el 46.20%. A manera de ejemplo, los lemas más frecuentes que encontramos de las clases de posesión fueron: *deber, comprar, necesitar, vender, perder*. Nuevamente distinguimos que el grupo no es homogéneo: tanto *deber, vender* como *comprar* comparten elementos del marco conceptual relacional pero, en un nivel de mayor especificidad de significado, *deber* y *vender* dan cuenta de una actividad de transferencia a diferencia de *comprar* que es una actividad de adquisición. Por otro lado, los lemas más recurrentes de la clase de verbos de atribución fueron: *llamar, servir, convertir, medir* y *pesar*. En cuanto a la heterogeneidad distinguimos por ejemplo la marcada diferencia entre *llamar* que da cuenta de una significación de denominación y *medir* o *pesar* que incluyen actividades de medida.

**Tabla 15**

*Clases de la Macro-clase Relacional: No. Lemas, Frecuencia y Porcentaje*

Clase	N lemas	Frecuencia	% de la macro-clase	Verbos (ejemplos)
Posesión	11	574	53.80	<i>deber, comprar, necesitar, vender, ganar</i>
Atribución	6	493	46.20	<i>llamar, servir, convertir, medir, pesar</i>

<sup>23</sup> El Anexo 9 corresponde al inventario de los verbos que componen la Macro-clase o Proceso Relacional con sus respectivas clases y subclases.

*Macro-clase de Modulación.*

Por último, la macro-clase de modulación<sup>24</sup> es en la que encontramos el menor número de lemas (6) y mayor heterogeneidad del conjunto de lemas. La mayoría de las ocurrencias se agrupa en los verbos de la clase de causación (70.45%); de esta clase de verbos encontramos tres lemas: *ayudar*, *dejar* y *permitir*. En estos tres lemas nuevamente comparten elementos del marco conceptual de causación pero el primero *ayudar* da cuenta de una acepción que hace referencia a la inducción y los otros dos a la de permiso. De la clase de verbos de disposición distinguimos dos: *tratar* e *intentar*. Por último, sólo encontramos un lema de la clase de verbos de aceptación: *aceptar*.

**Tabla 16**

*Clases de la Macro-clase de Modulación: No. Lemas, Frecuencias y Porcentaje*

Clase	N lemas	Frecuencia	% de la macro-clase	Verbos (ejemplos)
Causación	3	310	70.45	<i>dejar, ayudar, permitir</i>
Disposición	2	102	23.18	<i>tratar, intentar</i>
Aceptación	1	28	6.36	<i>Aceptar</i>

Notamos que los verbos relacionales y de modulación que responden a la categoría gramatical de los procesos que requiere significar el signo con el referente del proceso de la experiencia que describe, además de emparejar el significado con el significante, constituyentes del signo lingüístico de la forma gráfica. En el caso de estos verbos (relacionales y de modulación) incluyen la demanda del aprendizaje de las generalizaciones, relacionar un fragmento de la experiencia con otro, la construcción de la gramática de los procesos de

<sup>24</sup> El Anexo 10 corresponde al inventario de verbos (lemas) que componen la Macro-clase o Proceso de Modulación con sus respectivas clases y subclases.

identificación y clasificación. Como se puede apreciar, el grado de complejidad es mayor y tal vez esto explique que son verbos de uso restringido en los textos escritos para niños.

En la siguiente tabla presentamos un resumen de las macro-clases y clases que encontramos en el C-TEN con mayores ocurrencias, así como algunos ejemplos de los verbos más frecuentes de cada clase.

**Tabla 17**

*Resumen de Macroclases y Clases: No. Lemas, Frecuencia y Porcentaje Relativo y Absoluto*

MACROCLASE/clase	N lemas	Frecuencias	% frecuencia macroclase	% frecuencia vocabulario fundamental
<b>MATERIAL</b>	<b>97</b>	<b>6044</b>	<b>100</b>	<b>42.09</b>
Espacio	49	3580	59.23	24.91
Comportamiento	13	952	15.75	6.62
Otros hechos	20	822	5.16	5.72
Cambio	16	690	11.42	4.80
<b>MENTAL</b>	<b>33</b>	<b>3208</b>	<b>100</b>	<b>22.33</b>
Percepción	12	1602	49.94	11.15
Cognición	10	800	24.94	5.56
Sensación	7	274	8.54	1.91
General	2	247	7.70	1.72
Elección	1	85	2.65	0.59
Intersección (material+mental)	1	200	6.23	1.39
<b>VERBAL</b>	<b>15</b>	<b>2364</b>	<b>100</b>	<b>16.45</b>
Comunicación	13	2217	93.78	15.43
Emisión de sonido	1	77	3.26	0.54
Intersección (verbal+material)	1	70	2.96	0.49

<b>EXISTENCIAL</b>	<b>17</b>	<b>1246</b>	<b>100</b>	<b>8.67</b>
General	7	449	36.03	3.12
Fase-tiempo	6	440	35.31	3.06
Vida	4	357	28.65	2.48
<b>RELACIONAL</b>	<b>17</b>	<b>1067</b>	<b>100</b>	<b>7.43</b>
Posesión	11	574	53.80	3.99
Atribución	6	493	46.20	3.43
<b>MODULACIÓN</b>	<b>6</b>	<b>440</b>	<b>100</b>	<b>3.06</b>
Causación	3	310	70.46	2.15
Disposición	2	102	23.18	0.71
Aceptación	1	28	6.36	0.19

El siguiente esquema de la gramática de la experiencia de Halliday y Matthiessen (2004, p. 172) representa los tipos de procesos como un espacio semiótico dividido en regiones o dominios cognitivos. Los dominios cognitivos representados son clases prototípicas. Como lo habíamos mencionado los límites y fronteras de estos procesos y dominios cognitivos son difusos e implican relación con sus fronteras. Tomamos este esquema para explicar los resultados que obtuvimos de los tipos de proceso precisamente como un espacio semiótico. Dentro del espacio semiótico de la gramática de la experiencia podemos observar ciertas regiones dotadas de una clara relevancia. En este sentido, los procesos y dominios cognitivos asociados al mundo físico fueron los más importantes en términos de frecuencia de uso. Le siguen los que se relacionan con el mundo de la conciencia. En la frontera entre el mundo de la conciencia y de las relaciones abstractas resultó el proceso y dominios cognitivos que le siguen en cuanto a frecuencia de uso. El mundo de las relaciones abstractas fue de uso restringido. Por lo tanto, el esquema que a continuación presentamos podemos considerarlo como la

representación semiótica de la gramática de la experiencia de los verbos más frecuentes de los textos escritos previstos para niños mexicanos en momentos iniciales de alfabetización.

La gramática de la experiencia es una representación semiótica del conocimiento del mundo. La región privilegiada dentro de esta representación en el Corpus de Textos Escritos para Niños (C-TEN) fue la de los procesos que expresan las experiencias del mundo exterior. Estos procesos del mundo exterior que hacen referencia los procesos físicos son variados, tales como localizar, cambiar, crear, destruir o experimentar algún tipo de alteración. Pero también incluyen procesos físicos que no tiene que ver con los anteriores sino con la existencia psico-biológica y/o interpersonal o social. Lo mismo sucede con las otras dos regiones de la experiencia: el mundo de la conciencia y el mundo de las relaciones abstractas los dominios y marcos conceptuales son variados en cada uno de estos mundos. Todos estos procesos, dominios y marcos conceptuales que componen la representación semiótica de la gramática de la experiencia son una representación a su vez del conocimiento del mundo en su vertiente de lo dinámico que necesariamente tiene que ver con los participantes y las circunstancias en las que suceden dichas experiencias.

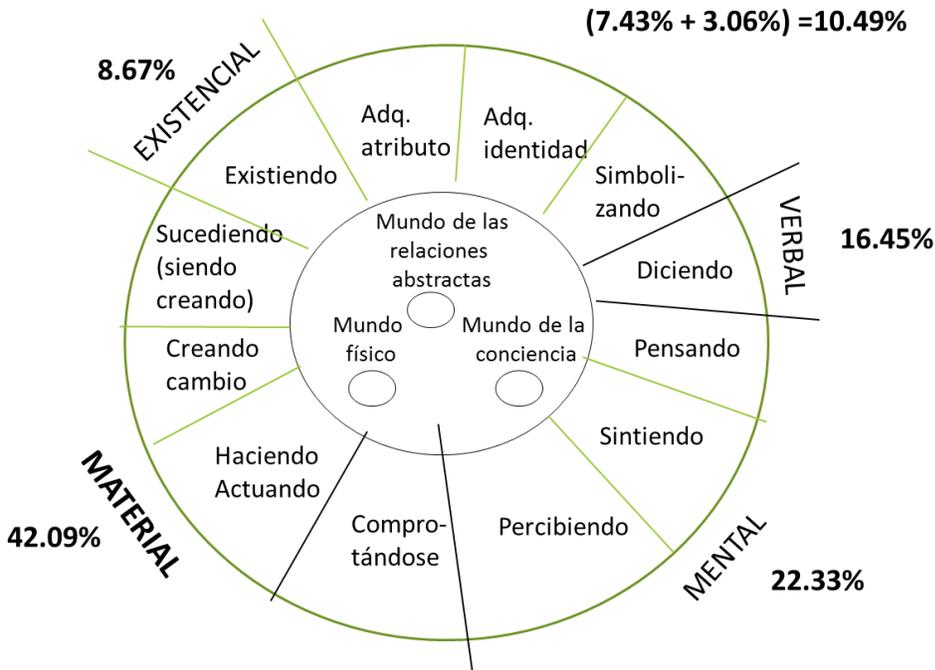


Figura 5. Representación Semiótica de la Gramática de la Experiencia del Vocabulario Fundamental Adaptación del esquema Gramática de la experiencia: tipo de proceso de Halliday, y Matthiessen (2004) para la presentación de los resultados del análisis sintáctico-semántico.

Por último, para contar con una referencia con qué comparar los resultados que obtuvimos del análisis semántico-sintáctico de los verbos del vocabulario fundamental, presentamos la siguiente tabla que establece el contraste con nuestros resultados de los verbos más frecuentes del C-TEN y los resultados reportados en el sistema de base de datos de ADESSE al 31 de marzo de 2014 de los verbos que han clasificado del corpus ARTHUS-BDS. La siguiente gráfica muestra la comparación.

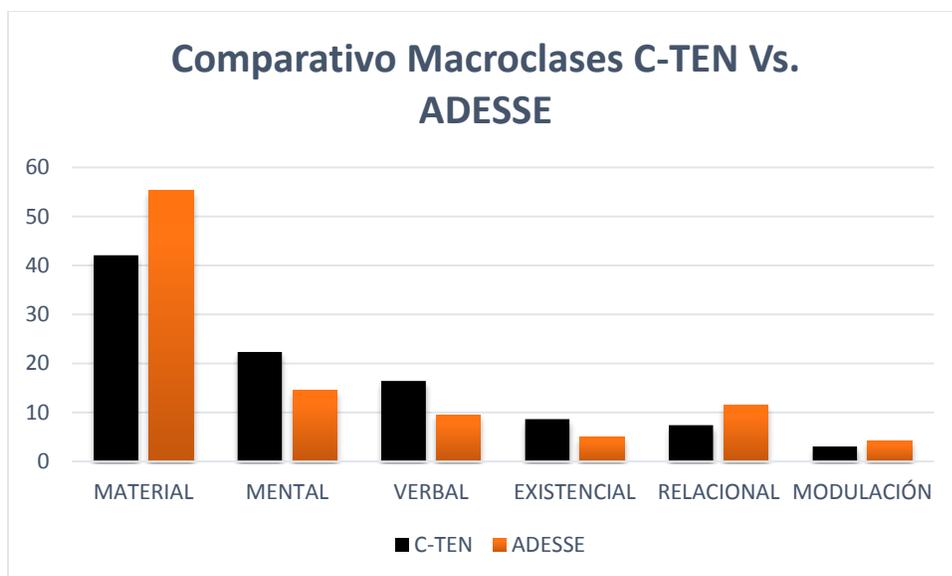


Figura 6. Comparación entre los Verbos del Vocabulario Fundamental del C-TEN y los Verbos Clasificados por ADESSE

Como lo señalamos en el capítulo de antecedentes, el límite del análisis que realizamos tuvo relación con las restricciones propias de la clasificación que utilizamos pero sobre todo con los recursos (materiales, de tiempo y humanos) con los que contamos para realizar esta tesis. Los límites de la estructura jerárquica de la clasificación se presentaron en las dificultades en el proceso de clasificación de aquellos verbos con frontera semántica difusa. Es decir, el caso de un mismo verbo, que denominamos de intersección, que respondió a diferentes procesos y/o dominios conceptuales y/o marcos conceptuales diversos como en el caso de *encontrar* y *escribir*. Otra limitación asociada a las dificultades clasificatorias fueron los verbos polisémicos en los que la misma palabra gráfica (lema) presentó doble clasificación como en el caso de *enseñar*, *llevar* y *saber*. Aunado a lo anterior, el límite de recursos precisó de optar por el sentido o acepción de máxima aparición en el sistema de base de datos de ADESSE, y por lo tanto, su clasificación semántica. Lo anterior no significa que sucedió así en nuestro corpus. Por lo tanto, este estudio fue un primer acercamiento y también una descripción general.

La siguiente tabla muestra los diez verbos que coincidieron como los más frecuentes tanto en nuestro corpus (C-TEN) con respecto a la base de ADESSE.

**Tabla 18**

*Los Diez Verbos Más Frecuentes tanto del C-TEN como de ADESSE*

N	VERBO	N. C-TEN	N. ADESSE	FRECUENCIA ADESSE	N. ENTRADAS / CLASES SEMÁNTICAS	CLASE SEMÁNTICA MAYOR FREC. APARICIÓN ADESSE
1	Decir	1032	3	2642	3	Comunicación
2	Ver	661	1	3461	2	Percepción
3	Llegar	343	13	953	1	Desplazamiento
4	Saber	324	2	3044	2	Conocimiento
5	Salir	246	11	1019	3	Desplazamiento
6	Vivir	256	20	659	2	Vida
7	Llevar	246	17	835	9	Desplazamiento
8	Venir	231	9	1168	2	Desplazamiento
9	pensar	214	6	1462	1	Cognición
10	preguntar	205	12	900	1	Comunicación

Los datos de la Tabla 18 dan cuenta de la gama de posibles clases semánticas que pudieron registrarse en los verbos del vocabulario fundamental. La gama de posibles clases semánticas se pueden considerar como criterio en la toma de decisiones para la búsqueda exhaustiva en los contextos del C-TEN. El verbo *llevar*, por ejemplo, presenta el mayor número de entradas (9), y por lo tanto, de clases semánticas: desplazamiento, adquisición, posesión actividad, propiedad, uso, tiempo, relación, relaciones sociales en la base de ADESSE. Aunado a lo anterior, al interior de las clases y subclases presentaron sub-acepciones. Estos datos pueden ser indicador de que el verbo *llevar* puede resultar interesante por la diversidad semántica provocada por el significado que se completa con el contexto (SV).

## CAPÍTULO 4

### CONCLUSIONES Y CONSIDERACIONES FINALES

La investigación que presentamos consistió en la descripción formal y sintáctico-semántica de los verbos más frecuentes del Corpus de Textos Escritos para Niños (C-TEN). Este trabajo nos dio la oportunidad de reflexionar sobre la construcción del concepto de verbo desde una perspectiva histórica para encontrar la forma de delimitar los criterios que orientaron el análisis del corpus de verbos. Los principales lingüistas y los académicos de la lengua española han demostrado las dificultades que presenta la categoría gramatical del verbo por su riqueza formal, sintáctica y semántica. De acuerdo con este debate sobre la conceptualización del verbo, delimitamos los criterios que guiaron esta investigación. Los criterios formales y funcionales (sintáctico-semánticos) del verbo fueron los que preferimos.

Recordamos al lector que este estudio se enfocó en una pregunta central: ¿cuáles y cómo son los verbos más frecuentes en los textos escritos para niños mexicanos en momentos iniciales de alfabetización inicial? y, más específicamente, ¿cómo son las características sintáctico-semánticas (léxicas) más frecuentes de los verbos que componen el vocabulario fundamental contenido en el C-TEN?

Respecto a la primera pregunta, encontramos que entre los verbos más frecuentes (100 primeros verbos del C-TEN) hubo una presencia importante de verbos no personales o no conjugados. Aunado a lo anterior, hubo una presencia considerable de perífrasis verbales donde participan estos verbos no personales. Las perífrasis verbales son un fenómeno interesante tanto

para la lexicografía como para la psicolingüística porque son una sola unidad léxica compuesta por dos verbos (a veces una preposición) y el verbo no conjugado desempeña una función diferente a la verbal (como sustantivo, adjetivo o adverbio).

La longitud (por número de sílabas) de los 100 primeros verbos fueron los verbos monosílabos y bisílabos que juntos sumaron el 75% de la muestra. Estos resultados nos dieron la oportunidad de reflexionar sobre las dificultades que plantea, en el caso de los verbos, delimitar la longitud. Por un lado, por la cualidad flexiva del verbo es difícil establecer la unidad de análisis para el conteo: el lema o qué verbo del paradigma al que pertenece. Por otro lado, el fenómeno de la perífrasis verbal que es una unidad léxica compuesta por dos verbos (a veces con una preposición), cuál sería el criterio para el conteo.

Podemos caracterizar a los verbos más frecuentes en los textos escritos para niños en momentos iniciales de alfabetización como verbos bisilábicos, que responden a la vocal temática *-er*, irregulares, conjugados en 3ª persona del singular, en tiempo presente del modo indicativo y predicativos. Esta fue la caracterización formal de los verbos conjugados más frecuentes.

La vocal temática de estos verbos (100 primeros verbos) presentó un comportamiento no canónico de los verbos escritos del español ya que la mitad de la muestra respondió a la clase *-er*. Esta clase no es la más productiva en el español.

Como se pudo apreciar la información que se deriva de la estructura morfé mica en el caso de los verbos del español da cuenta de un sistema complejo, variado y con esquemas definidos. Lo anterior, nos permitió anticipar que puede resultar interesante la descripción morfológica de los verbos más frecuentes en los textos escritos para niños para abonar en las

reflexiones de futuras investigaciones sobre la adquisición de la lectura (lengua) y la conciencia morfológica.

En cuanto a la segunda pregunta que versaba sobre las características sintáctico-semánticas de los verbos del vocabulario fundamental del C-TEN encontramos que el vocabulario fundamental estuvo conformado por 185 cabezas lexicas donde estuvieron incluidos los verbos con significado pleno. Las clases semánticas con mayor presencia en el vocabulario fundamental fueron los procesos material y mental. El proceso mental superó al resto de forma significativa. El dominio conceptual del proceso mental que presentó mayor frecuencia fue la clase de espacio. El marco conceptual que tuvo mayor presencia, tanto al interior de la macroclase como del vocabulario fundamental, fue la subclase de desplazamiento. El verbo con mayor frecuencia de este marco conceptual fue *llegar*.

La base de datos de ADESSE consultada para la asignación de los procesos, dominios y marcos conceptuales ha encontrado en este verbo (*llegar*) nueve acepciones y entre ellas figuras y locuciones verbales. Lo anterior, nos lleva a anticipar que puede resultar interesante el análisis de los contextos registrados de *llegar* en el C-TEN. Estos contextos podrían convertirse en observables interesantes para estudios psicolingüísticos.

El proceso mental fue el segundo proceso de mayor frecuencia tanto en número de lemas como de frecuencia absoluta. El dominio conceptual, al interior de esta macro-clase, que presentó mayor frecuencia fue el de percepción. El verbo de mayor presencia de este dominio conceptual fue *ver*. Este verbo presentó cinco sub-acepciones, entre ellas figuras, en la base consultada. Incluye la percepción física pero también cognitiva.

Aunado a lo anterior, *ver* es un verbo identificado potencialmente con significados metafóricos. Lo anterior, nos lleva a anticipar que puede resultar interesante realizar una búsqueda en los contextos del C-TEN para determinar si estos textos escritos para niños contienen significados metafóricos (locuciones verbales). *Pasar* y *llevar* fueron los verbos que presentaron mayor número de acepciones, y por lo tanto, mayor variedad de clases verbales en la base de datos de ADESSE. Este dato puede sugerir la revisión de estos verbos en los contextos del C-TEN para determinar la variedad de las acepciones y clases semánticas a las que están en contacto los niños en momentos de alfabetización inicial.

De acuerdo con Halliday y Matthiessen (2004) los procesos de adquisición más temprana en la lengua oral son precisamente los que nosotros encontramos como los más frecuentes en los textos escritos para niños: los procesos material y mental. Esta coincidencia entre la adquisición oral y escrita puede tener múltiples explicaciones, entre ellas está el que representa el mayor índice de uso de conceptos del mundo exterior y del mundo de la conciencia precisamente en los textos escritos para la población infantil.

Los distintos inventarios con criterios específicos elaborados en este trabajo de investigación podrían facilitar recursos para estudios posteriores. Desde la psicolingüística se podría indagar sobre el efecto de la disponibilidad léxica que presentan los niños con relación a los verbos escritos disponibles para niños en estos momentos iniciales de alfabetización. Podría resultar interesante también estudiar la significación de los verbos menos comunes. También se podrían realizar estudios sobre la anticipación visual en la lectura de verbos más y menos frecuentes utilizados en contextos específicamente dirigidos al público infantil. Por otro lado, sería pertinente indagar sobre la conciencia morfológica en los momentos iniciales de lectura de palabras en contexto. En el caso de la evaluación del lenguaje en general y de la lectura en

particular, el verbo es una variable recurrentemente utilizada. Los inventarios eventualmente podrían servir de apoyo para tomar decisiones para realizar pruebas de lectura, entre otros usos.

En el campo de la lexicografía, se podrían buscar sistemas de análisis perifrástico verbal que permitiera describir con mayor profundidad el fenómeno para convertirlo en un observable concreto para estudios psicolingüísticos. También se podría generar un sistema y un estudio contextual de las locuciones verbales como recurso para convertirlas en una variable interesante para estudios psicolingüísticos semánticos.

La tesis pretendió proveer de un recurso metodológico con descripciones e inventarios que dieran cuenta de los niveles de significación del verbo. El primer nivel de significación que consideramos fue el formal o morfológico flexivo. En él demostramos, con la estructura morfotáctica, que el verbo es un contexto morfológico privilegiado. Podría resultar interesante conocer la lógica con la que los niños explican el funcionamiento de la estructura morfotáctica. También se podría continuar este estudio describiendo la morfología derivativa para desarrollar observables interesantes para comprender cómo los niños explican el funcionamiento de formación de palabras.

El segundo nivel de significación, sintáctico-semántico, en el que se incluyeron en la gramática de la experiencia: los procesos, los dominios y los marcos conceptuales. Podría resultar interesante analizar cómo representan por medio del lenguaje la experiencia los niños. También se podría ahondar en la investigación describiendo los contextos específicos donde aparecen los verbos del vocabulario fundamental. Aunado a lo anterior, continuar con el análisis del sintagma verbal describiendo los participantes y las circunstancias. Con estos observables se

podrían realizar análisis sobre cómo los niños hipotetizan sobre el funcionamiento del sintagma verbal.

## BIBLIOGRAFÍA

- Aguirre, C. y Dressler, A. (2006). On Spanish Verb Inflection. *Folia Linguística*, pp. 40(1-2), 75-96.
- Ahumada, I. (2006). Justificación lingüística de las insuficiencias que presentan los actuales diccionarios de la lengua. In J. y. Moya, *Lexicografía y enseñanza de la lengua española* (pp. 15-32). Granada: Editorial Universidad de Granada.
- Alcoba, R. (1999). La flexión verbal. In *Gramática descriptiva de la lengua española* (pp. 4915-4992). Espasa Calpe.
- Albertuz, F. J.-7.-2. (2007). Sintaxis, semántica y clases de verbos: clasificación verbal en el proyecto ADESSE. *Sintaxis, semántica y clases de verbos: clasificación verbal en el proyecto de ADESSE* (pp. 2015-2030). Santiago de Compostela 3-7 de mayo de 2004: Actas del VI Congreso de Lingüística General.
- Base de Datos ADESSE consulta
- Bosque, I. (2010). *Nueva gramática de la lengua española. Manual*. Madrid: RAE-Espasa Calpe.
- Bosque, I. y Delmonte, I. (eds.) (1999). *Gramática Descriptiva de la Lengua Española*, España: Espasa Calpe.
- Breva, M. (2005 Actas del V Congreso internacional de la Sociedad Española de Historiografía Lingüística: Murcia, 7-11 de noviembre de 2005, vol. 1). El manuscrito "de verbo mentis" del bocense (1523-1601): aclaraciones sobre los fundamentos de su doctrina gramatical. In S. E. Lingüística, *Caminos actuales de la historiografía lingüística* (p. 309). Murcia: Editum.
- Cantos, P. y Sánchez. (2011,). El inglés y el español desde una perspectiva cuantitativa y distributiva: equivalencias y contrastes. *Estudios ingleses de la Universidad Complutense*, pp. vol. 9, 15-44. [http://dx.doi.org/10.5209/rev\\_EIUC.2011.v19.36242](http://dx.doi.org/10.5209/rev_EIUC.2011.v19.36242) consultado el 21 de marzo 2013.
- Chartier, A. y. (2002). *La lectura de un siglo a otro. Discursos sobre la lectura (1980-2000)*. España: Gedisa.
- Chartier, R. (2005). *El mundo como representación*. España: Gedisa.
- De Val, M. C. (1958). *Gramática española*. España: Saeta.
- Demonte, I. B. (1999). *Gramática descriptiva de la lengua española*. Madrid: RAE (colección Nebrija y Bello)-Escasa Calpe.
- Dowty, D. (1979). *Word Meaning and Montague Grammar*. Dordrecht: Reidel.

- Faber, P. y Mairal, R. (1999). *Constructing a Lexicon of English Verbs*. Berlin: Mouton de Gruyter.
- Fellbaum, C. (1998). A Semantic Network of English Verbs. In C. F. (ed), *WordNet. An Electronic Lexical Database* (pp. 69-104). Cambridge (Massachusetts): The MIT Press.
- Fillmore, C. J. (16/ 3, 2003). Background to FrameNet. *International Journal of Lexicography*, 235-250.
- García Aldeco, A. (2013). *Corpus de sustantivos más frecuentes en textos escritos para niños mexicanos, en momentos iniciales de alfabetización tesis inédita*. (Tesis inédita de maestría). Universidad Autónoma de Querétaro, Querétaro, Qro.
- García-Miguel, J. M., Costas, L. y Martínez, S. (2003). *Diateisis verbales y esquemas construccionales. Verbos, clases semánticas y esquemas sintáctico-semánticos en el proyecto ADESSE*. Leipzig.
- Goldberg, A. (1995). *Constructions. A Construction Grammar Approach to Argument Structure*. Chicago-London: The University of Chicago Press.
- Gómez-Torrego, L. (2007). *Gramática didáctica del español*. Madrid: Ediciones SM.
- González Calvo, J. M. (1978). El concepto de verbo. *Anuario de estudios filológicos*, (1), 65-90.
- González Calvo, J. M. (Nos. 34-35, 2011-2012). Caracterización del verbo como clase de palabra en español. *Revista de filología y su didáctica*, 181-193.
- gramaticales, L. c. (1989). *Ignacio Bosque*. Madrid: Síntesis.
- Halliday, M.A.K. (1985). *An Introduction to Functional Grammar*. London: Edward Arnold.
- Halliday, M.A.K. (2004b). Lexicology. In M. T. Halliday, *Lexicography and Corpus Linguistics* (pp. 1-22). London-New York: Continuum.
- Halliday, M.A.K. y Matthiesen, C. M. (2004). *An Introduction to Functional Grammar*. London-New York: Hodder Arnold.
- Halliday, M. y Matthiessen, M. (2006). *Contruing experience throug meaning: a language-based approach to cognition*. London-New York: Continuum.
- Ham, R. (1979). Del 1 al 100 en Lexicografía. In L. H. Lara, R. Ham y M. García, *Investigaciones lingüísticas en lexicografía* (pp. 41-83). México: El Colegio de México.
- Hanán, F. (2008). *Leer y mirar el libbro album: ¿un género en construcción?* Colombia: Norma.
- Hernando, L. A. (2011). El verbo en el arte de la lengua española castellana de Gonzalo Correas. *Revista de filología románica*, ISBN 0212-999x, No. 28, pág. 29-47.
- Hualde, J. I., Olarrea, A., Escobar, A. M., & Travis, C. E. (2010). *Introducción a la Lingüística hispánica*. Cambridge-Nueva York-Tokio: Cambridge University Press.

- Hürlimann, B. (1982). *Tres siglos de literatura infantil europea*. Barcelona: Juventud.
- Langacker, R. W. (1990). *Foundations of Cognitive Grammar*, 2 vols. USA: Stanford University Press.
- Muñoz, C. M. (1958). *El español vulgar: descripción de sus fenómenos y métodos de corrección* (Vol. 2). Sección de Publicaciones de la Secretaría General Técnica, Ministerio de Educación Nacional.
- Nebrija, A. D. (1999). Gramática de la lengua castellana. Recuperado de <http://www.antonionebrija.org/libro3.html#10>
- Lara, L. (1979). Del análisis semántico en lexicografía. En L. H. Lara, R. Ham y M. García, *Investigaciones lingüísticas en lexicografía* (pp. 157 - 266). México: El Colegio de México.
- Lara, L. (2006). *Curso de lexicografía*. México: El Colegio de México.
- Lara L.F. (2007). *Resultados numéricos del vocabulario fundamental del español de México*. México: Colmex.
- Levin, B. (1993). *English Verb Classes and Verbal Alternations. A Preliminary Investigation*. Chicago: University of Chicago Press.
- Llorach, E. A. (1994). *Gramática de la lengua española*. Madrid: Espasa.
- Marin, F. M. (2002). *Gramática española*. Madrid: Editorial Síntesis.
- Pottier, B. (1961). Sobre el concepto de verbo auxiliar. *Nueva Revista de Filología Hispánica*, pp. 325-331, Jul.-Dic. Año 15, No. 3/4.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Londres: Oxford University Press.
- Suances-Torres, J. (2000). *Diccionario del verbo español, hispanoamericano y dialectal*. Barcelona: Herder.
- Torrego, L. G. (2007). *La gramática didáctica del español*. España: SM.
- Torruella, J. y Llisteri, J. (1999). Diseño de corpus textuales y orales. In E. G. J.M. Bleca, *Filología e informática. Nuevas tecnologías en los estudios filológicos*. (pp. 45-77). Barcelona: Milenio.
- Vendler, Z. (1967). *Linguistics in Philosophy*. New York: Cornell University Press.

Bases de datos consultadas

Ancora-CAT y AnCora-ESP consultada en <http://clic.ub.edu/ancora/>

SenSem consultada en <http://grial.uab.es/fproj.php?id=1>

ADESSE consultada en <http://webs.uvigo.es/adesse/>

ARTHUS consultada en <http://www.bds.usc.es/>

BDS consultada en <http://www.bds.usc.es/corpus.html>

FrameNet consultada en <http://sfn.uab.es:8080/SFN>

WordNet consultada en <http://wordnet.princeton.edu/>