



UNIVERSIDAD NACIONAL  
AUTÓNOMA DE  
MÉXICO

**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**



Facultad de Ingeniería

Facultad de Filosofía y Letras

## **DISEÑO DE UN DICCIONARIO ONOMASIOLÓGICO PARA LA WEB**

**QUE PARA OBTENER EL TÍTULO DE:**

**INGENIERA EN COMPUTACIÓN**

**O**

**LICENCIADA EN LENGUA Y LITERATURAS HISPÁNICAS**

**P R E S E N T A N :**

**NADIA BUENDÍA SILVA  
MARTHA GRIZEL DELGADO RODRÍGUEZ  
CARLA ROSANA MÁRQUEZ GONZÁLEZ  
JESSICA ALEJANDRA MÉNDEZ LICEAGA**

**DIRECTOR: DR. GERARDO SIERRA MARTÍNEZ  
CODIRECTOR: M. EN C. GABRIEL CASTILLO HERNÁNDEZ**

**MÉXICO, D. F. SEPTIEMBRE 2004**



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

**A mis padres:**

*Les dedico esta tesis como un agradecimiento a su apoyo incondicional durante mi vida, por todo su amor y porque son parte fundamental de este logro.*

*— Nadia —*

A mis abuelos,  
Por su voluntad inquebrantable

A Marcela Rodríguez,  
Por su amor a la vida

A mis padres y hermanos  
Por su amor, tolerancia y comprensión

Con cariño,  
Grizel.

A mi madre  
Por sus demostraciones de amor y gran apoyo

A Joel y a nuestro pequeñuelo, Dany  
Porque son un impulso para seguir  
adelante

A mi padre y hermano  
Porque son parte de este logro

A mi querido *grupo de amigos* que me  
orientaron y apoyaron  
incondicionalmente...

Al Dr. Gerardo Sierra  
Por su tiempo, paciencia y guía

A mis compañeras de tesis  
Por su comprensión

*A la máxima casa de estudios*

Y principalmente a Dios que me  
permitió ser parte de esta tesis y me  
demostró su gran amor poniéndome en  
el camino a los mejores seres humanos  
Muchas gracias

*Carla Márquez*

*Este trabajo es dedicado a...*

*Mi mamá por ser la persona más especial en mi vida, por brindarme siempre su apoyo incondicional, amor y confianza, te amo mucho*

*Mi papá, por su apoyo y porque gracias a él aprendí muchas cosas que nunca olvidaré, te amo*

*Mis hermanas que siempre han estado conmigo, en especial a Laura que ha sido mi primer ejemplo a seguir y espero corresponder un poco a todo lo que ha hecho por mí, te quiero mucho*

*Javier por su apoyo y porque siempre tiene las palabras correctas en cada momento*

*Mi codirector de tesis por ser ejemplo a seguir como persona y profesionalista*

*Mis amigos porque juntos forman parte de una de las etapas más importantes de mi vida*

*A la UNAM y Facultad de Ingeniería por formarme como profesionalista y darme la oportunidad de vivir esta etapa*

*Gracias...*

*Jessica Méndez*

## ÍNDICE

ÍNDICE	I
1 INTRODUCCIÓN	1
1.1 Descripción del problema	3
1.2 Objetivos	4
1.3 Metodología	4
1.4 Alcances	6
1.5 Dependencias involucradas	6
2 FUNDAMENTOS TEÓRICOS	7
2.1 El lenguaje natural	8
2.2 Reconocimiento de palabras	8
2.2.1 Formas de solución	9
2.3 Lexicografía	10
2.3.1 Búsqueda en los diccionarios	10
2.3.1.1 La búsqueda semasiológica	10
2.3.1.2 La búsqueda onomasiológica	11
2.3.2 Tipología de diccionarios onomasiológicos	11
2.3.2.1 Diccionarios onomasiológicos impresos	13
2.3.2.1.1 Ideológicos	13
2.3.2.1.2 Tesoros	13
2.3.2.1.3 Inversos	14
2.3.2.1.4 Pictóricos	14
2.3.2.2 Limitaciones de los diccionarios onomasiológicos impresos	14
2.3.2.3 Diccionarios onomasiológicos electrónicos	15

2.3.2.3.1	Lenguaje aceptado	16
2.3.2.3.2	Tipo de formato	17
2.3.2.3.3	Recursos exteriores	18
2.3.2.4	Descripción de búsquedas onomasiológicas en diccionarios electrónicos	20
2.3.2.4.1	Descripción de la consulta onomasiológica en el diccionario <i>Vox</i>	20
2.3.2.4.2	Descripción de la consulta onomasiológica en el diccionario <i>Longman</i>	22
2.3.2.4.3	Descripción de la consulta onomasiológica en el diccionario <i>Wordsmyth</i>	24
2.3.2.4.4	Descripción de la consulta onomasiológica en el diccionario <i>OneLook</i>	25
2.4	Descripción del Corpus	27
2.4.1	Naturaleza de los términos	27
2.4.2	Descripción del ejercicio de conceptualización	27
2.4.3	Resultados del ejercicio de conceptualización	28
2.4.4	Metodología general del corpus	29
2.5	Conclusiones	30
3	CONCEPTUALIZACIÓN DE BÚSQUEDAS EN LENGUAJE NATURAL	31
3.1	Escuelas teóricas de la conceptualización	32
3.1.1	Mentalismo	33
3.1.2	Argumento Estándar y Atomismo Informativo	33
3.1.3	Semántica Informativa y Procedural	34
3.2	El concepto	34
3.2.1	El concepto y la percepción	35
3.2.2	El concepto y la referencia	36
3.2.3	El concepto y las creencias	36



3.2.4 El concepto nodo	37
3.2.5 El desarrollo de la red conceptual	37
3.2.6 El concepto y el significado	38
3.2.7 Problemas del concepto	38
3.3 Factores en el hablante	39
3.3.1 Los factores psicológicos	39
3.3.2 Los factores motivacionales	40
3.3.3 Los factores de habilidad	40
3.4 La definición	41
3.4.1 La definición tradicional	41
3.4.2 La definición en la onomasiología	42
3.4.3 La clasificación de las definiciones del corpus por tipo	42
3.4.3.1 El comportamiento del corpus	43
3.4.3.2 La tendencia del corpus	44
3.4.4 La definición libre	45
3.5 La presentación de la información en el corpus	45
3.5.1 La información por género próximo	46
3.5.2 La información por diferencia específica	47
3.5.3 La información por datos sensoriales	48
3.5.4 La información por características científicas	48
3.5.5 La información por verbos y complementos relacionados	50
3.5.6 La información por marcadores situacionales	51
3.5.7 Conclusiones de la presentación de la información	52
3.6 La aplicación de modelos teóricos al corpus	52
3.6.1 El contexto	53

3.6.2	El género próximo y los rasgos semánticos	54
3.6.3	El estereotipo	57
3.6.4	Los elementos diversos	58
3.6.5	La exhaustividad	60
3.6.6	Conclusiones de la aplicación de modelos teóricos al corpus	61
3.7	Conclusiones	62
4	EVALUACIÓN DE LOS DICCIONARIOS ELECTRÓNICOS	63
4.1	La evaluación apreciativa	64
4.1.1	Problemas de lematización	64
4.1.2	Problemas de formas sinónimas	65
4.1.3	Problemas de formas relacionadas semánticamente	66
4.1.4	Problemas del contenido de los diccionarios	67
4.1.5	Problema de errores de correspondencia y signos de puntuación	67
4.1.6	Problemas de presentación de los resultados	68
4.1.7	Problema del diseño	69
4.1.8	Problema de consulta	69
4.2	Evaluación formal	70
4.2.1	Introducción de la información en el diccionario	70
4.2.2	Descripción de la evaluación	71
4.2.2.1	Precisión	72
4.2.2.2	Fallout	72
4.2.2.3	Generallity	72
4.2.3	Construcción de las tablas para la evaluación	73
4.2.4	Evaluación formal entre los diccionarios	73
4.2.5	Conclusiones de la comparación entre diccionarios	77

4.2.6	Conclusiones	77
5	DISEÑO DE BÚSQUEDAS ONOMASIOLÓGICAS EN LENGUAJE NATURAL	79
5.1	Descripción del concepto en lenguaje natural	80
5.2	Identificación de palabras claves	81
5.3	Eliminación de palabras cerradas	81
5.3.1	Definición	82
5.3.2	Antecedentes	82
5.3.3	Procedimiento para la obtención de las palabras cerradas	85
5.3.3.1	Preposiciones	87
5.3.3.2	Artículos	87
5.3.3.3	Conjunciones	87
5.3.3.4	Verbos	88
5.3.3.5	Adjetivos	88
5.3.3.6	Adverbios	89
5.3.3.7	Pronombres	89
5.3.3.8	Sustantivos	89
5.4	Expansión de las palabras claves	89
5.4.1	Expansión morfológica	90
5.4.2	Expansión semántica	92
5.4.2.1	Diccionario de sinónimos	93
5.4.2.2	WordNet	94
5.4.2.3	Algoritmo flexibilizado de agrupamiento semántico	95
5.5	Búsqueda en el banco terminológico	96
5.5	Conclusiones	97
6	HERRAMIENTAS COMPUTACIONALES	98

6.1	Definición y concepto de la base de datos	99
6.1.1	Sistema manejador de bases de datos (DBMS)	100
6.1.1.1	Funciones de un DBMS	101
6.1.1.2	Arquitectura de un DBMS	101
6.1.1.3	Manejador de bases de datos SQL Server	102
6.1.2	Modelo de datos	104
6.1.2.1	Modelo relacional	104
6.1.2.1.1	Modelo Entidad-Relación	106
6.1.2.1.2	Normalización	107
6.1.2.1.3	Diccionario de datos	109
6.2	Servidores WEB	110
6.2.1	WEB	111
6.2.1.1	Páginas Estáticas	112
6.2.1.2	Páginas Dinámicas	113
6.2.2	Arquitectura Cliente-Servidor	114
6.2.3	Servidores de páginas WEB	114
6.2.4	Lenguajes de programación para la WEB	115
6.3	JAVA	116
6.3.1	¿Qué es JAVA?	116
6.3.1.1	Características	116
6.3.1.1.1	Simple	116
6.3.1.1.2	Dinámico	117
6.3.1.1.3	Distribuido	118
6.3.1.1.4	Robusto	118
6.3.1.1.5	De arquitectura neutral	119

6.3.1.1.6 Seguro	119
6.3.1.1.7 Portable	120
6.3.1.1.8 Interpretado	121
6.3.1.1.9 Máquina virtual	122
6.3.1.1.10 Multithreaded	122
6.3.2 Programación Orientada a Objetos	123
6.3.2.1 Mecanismos de la POO	123
6.3.2.1.1 Objetos	123
6.3.2.1.2 Métodos	124
6.3.2.1.3 Mensajes	124
6.3.2.1.4 Clases	125
6.3.2.2 Características de la POO	125
6.3.2.2.1 Abstracción	125
6.3.2.2.2 Encapsulamiento	126
6.3.2.2.3 Herencia	126
6.3.2.2.4 Polimorfismo	126
6.3.2.3 Constructores y destructores	127
6.3.3 Paquetes	127
6.3.4 Interfaces	128
6.3.5 Excepciones	129
6.3.6 Applets	129
6.4 Conclusiones	130
7 ANÁLISIS, DISEÑO Y DESARROLLO DEL SISTEMA	131
7.1 Metodología para el desarrollo de sistemas	132
7.1.1 Metodología del ciclo de la vida	132

7.1.1.1	Análisis del Sistema	133
7.1.1.2	Determinación de los Requerimientos	133
7.1.1.3	Diseño del sistema	133
7.1.1.4	Desarrollo de software	134
7.1.1.5	Pruebas del sistema	134
7.1.1.6	Implantación y evaluación	134
7.2	Análisis del Sistema	135
7.3	Determinación de los Requerimientos	137
7.3.1	Hardware	138
7.3.1	Software	138
7.4	Diseño del Sistema	139
7.4.1	Diagrama Entidad-Relación	139
7.4.2	Tablas del Sistema	140
7.4.3	Diccionario de Datos	141
7.4.4	Diseño Conceptual y Físico de la Base de Datos	144
7.4.5	Diagrama jerárquico funcional	146
7.4.6	Diseño de la pantalla principal	147
7.5	Desarrollo del Sistema	151
7.6	Migración de datos	153
7.7	Interfaz WEB para el DO	157
7.8	Conclusiones	158
8	PRUEBAS, IMPLANTACIÓN Y EVALUACIÓN DEL SISTEMA	159
8.1	Pruebas del Sistema	160
8.1.1	Pruebas unitarias	160
8.1.2	Pruebas de integración	160

8.1.3 Pruebas de validación	161
8.1.4 Pruebas volumétricas	161
8.2 Implantación	162
8.2.1 Implantación del prototipo en computadora independiente	162
8.2.2 Implantación final en Servidor WEB	163
8.3 Pruebas de Evaluación	166
8.3.1 Bateria de pruebas individuales	167
8.3.2 Aplicación de la Bateria de pruebas	169
8.3.3 Análisis de los resultados	171
8.4 Índices de Precisión y Recuperación (Precision and Recall)	171
8.5 Comparación de resultados	172
8.6 Conclusiones	174
9 CONCLUSIONES	176
9.1 Resultados Obtenidos	178
9.2 Trabajos Futuros	178
APÉNDICES	
Apéndice Capítulo 2	180
Apéndice Capítulo 3	198
Apéndice Capítulo 4	217
Apéndice Capítulo 5	256
Apéndice Capítulo 8	263
BIBLIOGRAFÍA	267

# CAPÍTULO 1

## INTRODUCCIÓN



Hoy en día se consolidan centros de investigación interdisciplinarios desde la raíz, ya que las tendencias actuales obligan a los investigadores a combinar perspectivas para la solución de problemas en común. Bajo este panorama surge la Ingeniería Lingüística (IL) como un área interdisciplinaria en reciente y creciente desarrollo. La IL involucra principalmente a ingenieros en computación y lingüistas que conjuntamente desarrollan modelos de lenguaje y herramientas para manipular el lenguaje natural.

En el presente capítulo se presenta un proyecto interdisciplinario, el desarrollo de un diccionario onomasiológico, en el que participan las facultades de Ingeniería y Filosofía y Letras de la UNAM. Para ello, se da a conocer la descripción del problema, los objetivos del proyecto, la metodología y los alcances a lograr.



El lenguaje permite al hombre transmitir ideas, sentimientos, inquietudes, crear o resolver malentendidos. Comparado con otras formas de comunicación como el lenguaje de señas, la mímica y los símbolos, el lenguaje escrito o hablado supera muchas limitaciones, permite entre otras funciones que la comunicación sea fluida y con mejores posibilidades de entendimiento y expresión.

El lenguaje natural ha sido estudiado desde varias perspectivas ya que su importancia no restringe su función al papel de una herramienta de comunicación, sino que también actúa como vía de transmisión del conocimiento a lo largo del tiempo<sup>1</sup>. Una de las disciplinas que lo han estudiado desde varios ángulos es la Filosofía, la cual ha hecho hincapié en el carácter de verdad y falsedad del lenguaje natural, así como otros aspectos entre los cuales se puede mencionar a las estructuras gramaticales y sintácticas que caracterizan cada lengua. La Lingüística, por su parte, lo considera como un sistema de signos usado para la comunicación. Sea cual fuere la perspectiva, es importante reconocer que el lenguaje natural implica ambigüedad pues no se trata de un sistema perfecto, unívoco, finito y cuantificable. Asimismo, una disciplina interesada en el lenguaje natural es la ingeniería donde a través de la implementación de sistemas se busca la interacción entre sistemas computacionales y el usuario.

A pesar de ser un sistema de uso constante, el lenguaje no está exento de sufrir problemas como el olvido o desconocimiento de las palabras y sus significados, debido a que en el nivel léxico el número de palabras es enorme y sólo pocas son de uso frecuente. A cualquier ser humano le ha ocurrido el problema de tener una palabra “en la punta de la lengua”, es decir, saber que existe la palabra, pero no poder decirla ya que sólo se recuerdan segmentos de ésta, o bien, algunas de sus características particulares, como el número de sílabas.

La solución de tal problema, si se aplica la Ingeniería Lingüística, implica un soporte lingüístico pero con un desarrollo computacional que lo haga más eficiente y rápido.

Debido al paulatino desarrollo de la Ingeniería Lingüística, se han visto beneficiadas áreas de la lingüística, una de ellas es la lexicografía –disciplina encargada de hacer diccionarios-, pues gracias a las distintas aplicaciones computacionales se mejoran los productos de esta disciplina.

---

<sup>1</sup> L. Moreno et al. (1999), *Introducción al lenguaje natural.*, Alicante, Universidad de Alicante, p. 13.

## 1.1 Descripción del problema

Con frecuencia sucede que al hablar o al escribir nos detenemos porque hemos olvidado una palabra, ya sea alguna usada con mucha frecuencia o una que rara vez se emplea, es decir, se sabe el concepto pero no se recuerda el término de ese concepto. A este problema se le denomina comúnmente “tener la palabra en la punta de la lengua”. Desde un punto de vista semántico, la onomasiología es la encargada de estudiar cómo los conceptos se asocian con las palabras.

No obstante, la lexicografía tradicional ha dado poco énfasis a este tipo de estudios por diversas razones:

- No había conceptualizaciones fiables o completas.
- La consulta de obras impresas (como el Diccionario Casares en lengua española) resultaba muy compleja.
- Debido a su formato, la información se perdía o resultaba obsoleta. Por tales motivos, los lingüistas dejaron de investigar esta área, reduciendo su importancia sólo a los diccionarios sinonímicos y antonímicos.

Actualmente, comienza a haber una serie de artículos por parte de diferentes lexicógrafos, lingüistas computacionales e investigadores relacionados que evidencian la falta y necesidad de los diccionarios onomasiológicos no sólo como herramientas utilizables por expertos sino por la mayoría de gente.

Las nuevas ramas de la lingüística, como la lexicografía computacional, vuelcan nuevamente la atención sobre la onomasiología (partir de una definición para llegar a la palabra). Sin embargo, se topan con el problema de la conceptualización dada, para solucionar esto:

- restringen las definiciones sólo a aquellas que cuentan con información científica, y
- emplean conceptualizaciones orientadas taxonómicamente. Tales condiciones serían idóneas si la gente común definiera bajo parámetros lexicográficos, pero como no es así, los diccionarios onomasiológicos electrónicos aun no pueden ser eficientes ni estar al alcance para cualquier usuario.

## 1.2 Objetivos

Debido a la escasez de diccionarios orientados al usuario común, el objetivo de este trabajo es desarrollar un Diccionario Onomasiológico (DO) electrónico con el fin de realizar búsquedas de términos a partir de la descripción del concepto en lenguaje natural, por medio de una interfaz en la WEB.

Es decir, Se busca realizar un Diccionario Onomasiológico que cumpla con las siguientes características:

- A partir de una idea o concepto dada por el usuario, la finalidad será encontrar el término más adecuado, debido a que éste no se recuerda.
- Que esté en formato electrónico.
- La forma de búsqueda aceptará lenguaje natural, es decir, una vez teniendo el concepto claro en mente, al momento de escribirlo, esta descripción no tendrá limitantes. Por ello, se aceptarán palabras sueltas, definiciones bien redactadas o breves. Asimismo, se aceptarán diversos signos de puntuación, etc.
- Se trabajará con la serie de palabras que se le ocurran al usuario.
- Se contará con una Interfaz WEB. La interfaz es el sector de la imagen en la pantalla del monitor que muestra distintas posibilidades de operar, o bien las opciones, el desarrollo de ciertos procesos, etcétera. Se emplea el término WEB para definir a un servidor World Wide Web (WWW), así como para definir el universo de Internet en su totalidad. Por tanto la interfaz WEB es visible desde Internet.

## 1.3 Metodología

El equipo de trabajo para la realización de la presente tesis estará compuesta por 4 estudiantes: dos por parte de la Facultad de Filosofía y Letras que se dedicarán a la parte lingüística y dos de la Facultad de Ingeniería dedicadas al desarrollo computacional; además, trabajarán en conjunto para el diseño conceptual y la evaluación del diccionario.

A fin de lograr el objetivo propuesto, se propone la siguiente metodología correspondiente a los capítulos de la tesis:

- **Capítulo 2: Fundamentos teóricos.** En esta unidad se hablará de la onomasiología, los diccionarios tanto impresos como electrónicos. Asimismo se describirá el corpus trabajado en las unidades tres y cuatro.
- **Capítulo 3: Conceptualización de Búsquedas en Lenguaje Natural:** Como el objetivo es lograr que una definición en lenguaje natural sea tratada por nuestro sistema, entonces es necesario hacer una revisión de los aspectos cognitivos referentes a la conceptualización. En otras palabras, se intenta saber qué es importante para el usuario, qué tipo de información se puede esperar de él. Además de ello, se hablará de la definición con la finalidad de enriquecer el concepto de “significado”. A partir de la teoría se harán comparaciones con el corpus descrito anteriormente en la unidad 2.
- **Capítulo 4: Evaluación de los Diccionarios Electrónicos.** Permitirá conocer el estado del arte del problema. Se evaluarán diccionarios descritos en la unidad 2 para conocer sus limitaciones y, la diversidad de estos., así como su formato enriquecerá el panorama para saber cuáles son las ventajas y limitaciones de los diccionarios electrónicos.
- **Capítulo 5: Diseño de Búsquedas Onomasiológicas en Lenguaje Natural.** Se hará un análisis de DEBO, se tratará cuál es la condición actual y limitación del mismo, se hará un diseño de procedimiento de búsquedas el cual se dará una introducción del concepto en lenguaje natural, se tratará acerca del Stoplist y una vez identificadas las palabras que pertenecen al mismo se identificarán las palabras claves y se hará una expansión de las mismas. Éste preámbulo permitirá dar inicio a la parte computacional, el cual se trata en los siguientes capítulos.
- **Capítulo 6: Herramientas computacionales.** El objetivo de este capítulo es dar una explicación teórica de las herramientas a utilizar para el desarrollo del sistema, se hablará acerca de la definición y concepto de las bases de datos y que todo ello implica, de los servidores Web, así también se manejará el tema del lenguaje de programación Java y con sus características.
- **Capítulo 7: Análisis, Diseño y Desarrollo del Sistema.** En éste capítulo se deja la parte teórica para dar inicio a la parte del análisis, diseño y desarrollo del sistema, cada etapa que conforma esta unidad esta constituida por elementos que se complementan para lograr el objetivo de esta unidad. Este capítulo es muy importante ya que dará un preámbulo de cómo es que el sistema funcionará para que el usuario pueda interactuar con el mismo.

- **Capítulo 8: Pruebas, Implantación y Evaluación del Sistema.** El objetivo de este capítulo es crear una conexión física más directa usuario-sistema, esto mediante las pruebas, implantación y evaluación del mismo, gracias a este capítulo se podrá confirmar el buen funcionamiento del mismo con resultados óptimos esperados por el usuario.
- **Capítulo 9: Conclusiones.** En éste capítulo se dan a conocer los resultados obtenidos a lo largo del planteamiento y desarrollo de la presente tesis, se dan a conocer los resultados por parte del análisis que se realizó para los diccionarios onomasiológicos y de los resultados esperados y obtenidos de la implantación del sistema y se hablará de trabajos posteriores al desarrollado en esta tesis para tener la posibilidad a futuras versiones.

#### 1.4 Alcances

Para el progreso de esta tesis, dada la interdisciplinariedad y lo extenso del tema –además de que implica cuestiones muy específicas de cada área– se puede dividir en dos tipos de estudios:

- **Estudios Lingüísticos:** Para los estudios cognitivos de la conceptualización, investigación sobre Diccionarios Onomasiológicos y su evaluación, así como para los aspectos lingüísticos del desarrollo del diccionario como depuración de paradigmas semánticos, metodología de Palabras Clave, etcétera.
- **Estudios Computacionales:** Para todo lo que implica la parte del análisis, desarrollo e implementación del diccionario, es decir, lo relacionado con la programación, la construcción de la interfaz y la administración de las bases de datos.

#### 1.5 Dependencias involucradas

Como tema de tesis, el trabajo se desarrolla en dos facultades: La Facultad de Filosofía y Letras y la Facultad de Ingeniería, las cuales pueden trabajar conjuntamente en el Grupo de Ingeniería Lingüística, parte del Instituto de Ingeniería de la UNAM. Estas dependencias trabajan en conjunto para lograr el objetivo planteado y para que posteriormente se realicen trabajos futuros en la investigación y desarrollo de Diccionarios Onomasiológicos.

## CAPÍTULO 2

# FUNDAMENTOS TEÓRICOS



En el primer capítulo se introdujeron conceptos básicos de ingeniería lingüística y se delimitaron los alcances del presente trabajo.

Para desarrollar un diccionario onomasiológico en formato electrónico, se debe estar familiarizado con ciertos conceptos lingüísticos. Por ello, en esta unidad se introducirán nociones de lenguaje (2.1), reconocimiento de palabras (2.2), lexicografía (2.3), búsqueda onomasiológica (2.3.1.2), además de los diversos tipos de diccionarios onomasiológicos (2.3.2) tanto impresos como electrónicos. Asimismo, se describirá la metodología del corpus tratado (2.4), que permitirá estudiar los fenómenos de conceptualización (capítulo tres) y evaluar los diccionarios electrónicos existentes (capítulo cuatro).

Con la finalidad de delimitar la metodología del diccionario onomasiológico, se explicará en esta unidad una serie de términos que intervienen en la confección de diccionarios, tales como lenguaje (2.1), reconocimiento de palabras (2.2), lexicografía (2.3), búsqueda onomasiológica (2.3.1.1), además de los diversos tipos de diccionarios onomasiológicos (2.3.2) tanto impresos como electrónicos. Asimismo, se describirá la metodología del corpus tratado (2.4), que permitirá estudiar los fenómenos de conceptualización (capítulo tres) y evaluar los diccionarios electrónicos existentes (capítulo cuatro).

## 2.1 El lenguaje natural

El lenguaje natural permite al hombre transmitir ideas, sentimientos, inquietudes; crear o resolver malentendidos. Comparado con otras formas de comunicación como la mímica y los símbolos, el lenguaje escrito o hablado supera muchas limitaciones y permite, entre otras funciones, que la comunicación sea fluida y con mejores posibilidades de entendimiento y expresión. Asimismo, permite al ser humano –bajo ciertos parámetros sintácticos, morfológicos y fonéticos– expresar una misma idea de distintas maneras dependiendo del contexto, su intención comunicativa y su capacidad como hablante.

A pesar de ser un sistema de uso constante, el lenguaje no está exento de sufrir problemas como el olvido o desconocimiento de las palabras y de sus significados, debido a que en el nivel léxico el número de palabras es enorme y sólo pocas palabras son de uso frecuente.

## 2.2 Reconocimiento de palabras

A cualquier ser humano le ha ocurrido el problema de tener una palabra “en la punta de la lengua”; es decir, saber que existe la palabra, pero no poder decirla ya que sólo recuerda segmentos de ésta, o bien algunas de sus características particulares, como el número de sílabas.

Existen varias formas de exponer la forma usada por el cerebro humano para evocar una palabra o término. Tales propuestas teóricas tratan de explicar cómo se da el proceso de reconocimiento de palabras, con la finalidad de conocer también cuáles son los factores que motivan la confusión u olvido de éstas.

Una de las posturas concibe tal proceso como una **búsqueda jerárquica**, en la cual la palabra no se reconoce inmediatamente, sino que se presupone la existencia de nodos de distintos niveles, los

cuales contienen representaciones de palabras. Entre los modelos más difundidos, Engelkamp menciona el de Johnston y MacClelland, el cual consta de tres niveles (acústico, morfológico y semántico) y en donde las búsquedas se realizan siguiendo estrictamente estos tres niveles. De esta manera, no sólo se reconoce una palabra sino que se eliminan a la vez aquellas que, pudiendo parecerse, no son las buscadas<sup>2</sup>.

Por la rigurosidad de consulta en el proceso anterior y además a sabiendas que el reconocimiento de palabras llevado a cabo por un ser humano se realiza a una velocidad sorprendente, Engelkamp señala el **modelo de cohorte** (*cohort model*) de Marslen-Wilson<sup>3</sup>. Este modelo es menos estricto que el anterior y explica el reconocimiento a partir de la interacción entre los diferentes niveles. Sin embargo, como lo señala Aitchison, el problema del modelo es que funciona cuando se recuerda el comienzo de la palabra, siendo que existen casos en los que el hablante recuerda fragmentos de la palabra que no siempre corresponden al principio<sup>4</sup>.

### 2.2.1 Formas de solución

El olvido o desconocimiento de un vocablo puede ser solucionado de diversas maneras. Cuando un hablante sostiene una conversación, puede reparar el olvido empleando un sinónimo o una perífrasis que le permitirá al interlocutor interpretar la palabra que ha olvidado. Otra forma consiste en solicitar explícitamente al interlocutor que a partir de una descripción de la palabra, ya sea por significado o por residuos de su forma, elucubre tanto la referencia como su equivalente lingüístico en una lengua determinada.

No siempre es posible encontrar un sinónimo o una perífrasis adecuada, ni siempre es posible contar con una persona que conozca la lengua en cuestión y ayude a recordar el vocablo. Por ello, otra forma de compensar estos problemas es acudir al **diccionario**, que tradicionalmente se puede definir como un libro de consulta o un compendio de palabras asociadas con sus significados.

<sup>2</sup> J. Engelkamp (1983), "Word meaning and Word recognition", Seiler & Wannenmacher (ed.), *Concept development and the development of word meaning*, Berlin: Springer-Verlag, p. 19-22.

<sup>3</sup> *Ibidem*, p. 20

<sup>4</sup> J. Aitchison (2001), *Words in the mind*, Oxford: Blackwell, pp. 206, 212.



## 2.3 Lexicografía

La **lexicografía** es la parte de la lingüística encargada de hacer diccionarios, para ello perfecciona métodos que estandarizan el significado así como los vocablos. Tiene dos vertientes principales: la **semasiología** (de la palabra al significado) y la **onomasiología** (del significado a la palabra). La primera cuenta con numerosos trabajos en la mayoría de las lenguas existentes. Sin embargo, la segunda tiene un número reducido tanto de investigaciones como de diccionarios terminados.

### 2.3.1 Búsquedas en los diccionarios

El proceso ordinario de consulta del diccionario consiste en partir de una entrada léxica para obtener como resultado un significado que en ocasiones viene acompañado con la información gramatical, etimológica o semántica de ésta. A tal proceso se le conoce como **búsqueda semasiológica**.

#### 2.3.1.1 La búsqueda semasiológica

Cuando un hablante desea conocer o precisar el significado de una forma lingüística, debe saber algunas cosas antes de iniciar la consulta:

- conocer al menos su forma escrita.
- identificar su forma canónica.

Actualmente existen sistemas computacionales –como diccionarios electrónicos o buscadores– que pueden corregir la ortografía, buscar la forma canónica e incluso ser capaces de desplegar una lista de palabras parecidas gráficamente por las letras tecleadas, entre otras características.

Cuando se localiza la palabra buscada en el diccionario impreso o electrónico, puede ocurrir alguno de los siguientes casos con respecto a la definición:

- que sea tautológica,
- que sea fácil de entender,
- que implique nuevas palabras desconocidas y, por tanto, resulte necesario hacer más búsquedas.

La mayoría de estas palabras ‘desconocidas’ a menudo remiten al conocimiento taxonómico o especializado. No obstante, en términos funcionales, sucede que tal información es innecesaria para el usuario del diccionario. Así lo hace notar Ayto al revisar la definición tradicional de ‘tigre’ para el inglés:

*“Such a definition as this, even if correct, is of no use to the reader; it tells him many things he either already knows, albeit unconsciously”* (AYTO, 90).

### 2.3.1.2 La búsqueda onomasiológica

Así como es posible consultar un significado, como se vio en el apartado anterior, puede también buscarse una palabra; a tal proceso se le conoce como **búsqueda onomasiológica**. En este tipo de consulta el hablante puede estar en cualquiera de las siguientes situaciones:

- recordar algunos sonidos de la palabra o el número de sílabas que tiene la palabra perdida;
- saber que puede existir el objeto pero desconocer el nombre;
- conocer el significado pero no recordar el nombre.

Al tomar en cuenta sólo el tercer caso, aquello que el hablante sabe, pero no puede nombrar, deja a su vez varios caminos posibles para encontrar el término:

- buscar a través de categorías,
- buscar un sinónimo o una imagen, o
- usar lenguaje natural.

Las posibilidades son varias. Para conocer la mejor forma de consulta, en el siguiente apartado se mostrarán diferentes diccionarios onomasiológicos, mostrando sus ventajas y desventajas.

### 2.3.2 Tipología de diccionarios onomasiológicos

La oposición presentada en los diccionarios, semasiológico/onomasiológico, no es única. A lo largo de la historia se han utilizado otras, como: diccionarios ideológico/ordinarios,

ortográfico/ideográficos, alfabético/temáticos, basados en la palabra/basados en el objeto, definicional/conceptual, sintagmático/paradigmático, general/terminológico, entre otras.<sup>5</sup>

En este trabajo se utiliza la oposición dada por Sierra<sup>6</sup>, que alude a una perspectiva semántica siguiendo la propuesta de Baldinger, en la que se clasifica al diccionario a partir de criterios de decodificación y codificación. El primero refiere al diccionario semasiológico que sirve para procesos pasivos como la lectura o comprensión auditiva. El segundo toma en cuenta al diccionario onomasiológico que sirve para procesos activos de escritura o producción oral.

Los especialistas consideran que los diccionarios onomasiológicos se pueden clasificar a partir de diversos criterios. Hartmann<sup>7</sup> los agrupa dependiendo de la forma de consulta que presentan. Al partir de los diccionarios existentes para el inglés, el autor menciona que en esta lengua hay 6 tipos diferentes de diccionarios onomasiológicos: a) los tesauros puros, como el *New Oxford Thesaurus of English*; b) los tesauros híbridos, como el *Chambers Concise Dictionary & Thesaurus*; c) tesauros con una finalidad determinada para hablantes nativos o de segunda lengua, como el *Oxford Study Thesaurus*; d) los diccionarios bilingües de tipo híbrido, como el *Collins/Robert Comprehensive Dictionary & Thesaurus French-English English-French*; e) los diccionarios de sinónimos, expresiones, colocaciones, parónimos, o pictóricos, como el *Oxford Visual Dictionary* y f) los productos electrónicos como el *Cambridge Advanced Learner's Dictionary*.

Por su parte, Sierra<sup>8</sup> agrupa a los diccionarios en: ideológicos (2.3.2.1.1), sinonímicos, bilingües, tesauros (2.3.2.1.2), inversos (2.3.2.1.3) y pictóricos (2.3.2.1.4) –clasificación que sirve para los propósitos de la presente investigación. Cada grupo de diccionarios se describirá se describirá a continuación sucintamente. Sin embargo, los diccionarios sinonímicos y los bilingües no se tomarán en cuenta por separarse del interés de esta investigación, ya que se trabajará en la misma lengua y a través de varias palabras o de conceptos y no de equivalentes léxicos.

Por los objetivos del presente trabajo, se hará además una presentación de los diccionarios de acuerdo al formato que posean, esto es, diccionarios onomasiológicos impresos (2.3.2.1) o electrónicos (2.3.2.3).

<sup>5</sup>R. Hartmann , “Onomasiological Dictionaries in 20<sup>th</sup>-Century Europe”, en *Lexicographica. International Annual of Lexicography*, n.20/2004 Werner Hüllen (ed), pp. 1-17. En este artículo se evidencia la carencia de metodología y diccionarios en las lenguas indoeuropeas masivas como el inglés, alemán, español, ruso, francés, etc. p. 2.

<sup>6</sup>G. Sierra (1999), *Design of a concept-oriented tool for terminology* (Tesis doctoral), Manchester: Centre for Computational Linguistics, p. 15.

<sup>7</sup> Ibidem.

<sup>8</sup> Véase mayor información sobre la problemática en Sierra (1999), capítulo 2.

### 2.3.2.1 Diccionarios onomasiológicos impresos

Siguiendo la clasificación de Sierra<sup>9</sup>, los diccionarios impresos se pueden clasificar según el usuario, tema, ordenamiento de la información o estructura del diccionario. Este último criterio permite contrastar diferentes tipos de diccionarios a partir de su información y opciones de consulta.

#### 2.3.2.1.1 Ideológicos

Los diccionarios ideológicos son el resultado de clasificaciones conceptuales y temáticas de formas lingüísticas. Se caracterizan porque evidencian relaciones entre diferentes términos que apuntan hacia una misma idea. Su importancia, según Santana<sup>10</sup>, es que sirven para la clasificación de documentos y su recuperación. Para este autor, las ideas tienen correspondencias con palabras u oraciones que se comunican con estructuras más o menos complejas. Éstas se pueden “traducir” en relaciones físicas a través de relaciones semánticas de hiponimia e hiperonimia, como ocurre en estos diccionarios. Lo mismo sucede con los diccionarios sinónimos y los tesauros. Bajo esta categoría se puede mencionar al *diccionario ideológico de la lengua española*.

#### 2.3.2.1.2 Tesauros

Los tesauros son los diccionarios de referencia más comunes y antiguos en la onomasiología. Su macroestructura puede ser alfabética o temática. El inicio de la búsqueda precisa de una palabra entrante. Sin embargo, su ordenamiento frecuentemente obliga al usuario a repetir la búsqueda varias veces hasta que su palabra clave lo lleve al término buscado.

En la microestructura el orden es generalmente alfabético, criterio que según varios lexicógrafos resulta problemático porque no existen relaciones semánticas entre los términos de un grupo. R. Hartmann<sup>11</sup> hace notar que algunos diccionarios sinónimos se publicaron con el nombre de tesauros y que, además, se publican actualmente unos que poseen una estructura híbrida la cual obedece a intereses comerciales. Como ejemplo de estos está el *Roget's Thesaurus of English Words and Phrases*.

<sup>9</sup> G. Sierra (2000), “The onomasiological dictionary a gap in lexicography” en *Proceedings of Euralex*, Oxford, p. 224.

<sup>10</sup> O. Santana, et al. “Herramienta para el manejo de diccionarios ideológicos”, Departamento de Informática y sistemas, Universidad de Las Palmas de Gran Canaria, <http://protos.dis.ulpgc.es>, p.1.

<sup>11</sup> Hartmann, p. 7.

### 2.3.2.1.3 Inversos

Los diccionarios inversos poseen características específicas que permiten a un usuario buscar una palabra a partir de una palabra clave situada en un índice o árbol conceptual. Estos diccionarios teóricamente podrían partir del lenguaje natural (una descripción), para llegar a una palabra (entrada). Su mayor complejidad se debe a las infinitas formas de clasificar la información. Es decir, la clasificación de un diccionario inverso difiere de un autor a otro, como ocurre entre un experto y un usuario común. Lo anterior repercute en la eficacia del diccionario. Un ejemplo para esta categoría es el *Reader's Digest Reverse Dictionary*.

### 2.3.2.1.4 Pictóricos

Los diccionarios pictóricos se pueden definir como un compendio de imágenes, donde cada una pretende ilustrar el significado de una palabra. Estos diccionarios son frecuentemente usados como material de apoyo para el aprendizaje de lenguas extranjeras. Sin embargo, su principal defecto consiste en que sólo una parte del vocabulario de una lengua se puede representar a través de imágenes, como objetos físicos (*mesa, silla*) y cierto tipo de acciones (*bailar, correr*). No así con conceptos meramente abstractos (*amor, amistad*). Un ejemplo en lengua española es el *Diccionario ortográfico ideovisual*.

### 2.3.2.2 Limitaciones de los diccionarios onomasiológicos impresos

A partir de las características de los diccionarios que se han visto hasta ahora, a continuación se mencionará brevemente una serie de limitantes que poseen los diccionarios onomasiológicos impresos<sup>12</sup>.

Los tesauros no son diccionarios adecuados para el usuario común porque la clasificación realizada por los expertos puede no ser de ayuda para el hablante que carece de un conocimiento preciso o lexicográfico. Aunado a ello, la microestructura puede no ser la adecuada para guiar al usuario, quien se puede confundir o desorientar al buscar la palabra. Es posible que para el usuario común este tipo de consulta sea tediosa si su conceptualización de la realidad difiere de la del diccionario, pues esto lo llevará a repetir numerosas veces su consulta. En el caso de los

<sup>12</sup> Si se desea profundizar en el tema, véase Sierra (2000), en el que, a través de un análisis comparativo, el autor revisa y evalúa varios diccionarios impresos.

diccionarios pictóricos, las imágenes, por otro lado, son capaces de ilustrar objetos pero, como se vio, no pueden ejemplificar con certeza referentes abstractos como los conceptos creados por el hombre. En los diccionarios inversos se cuenta con las mismas limitantes del tesoro: si el usuario común no posee la misma conceptualización del mundo que la ideada en dichos diccionarios, entonces la búsqueda se volverá complicada.

Por otro lado, en el apartado 2.2.1 se mencionaron las formas comunes que emplea el ser humano para reparar el olvido de las palabras, en donde se observó que la búsqueda de una palabra puede realizarse sencillamente empleando lenguaje natural. Tal opción es cómoda porque así el hablante expresará con cierta facilidad, a través de sus propias palabras y bajo sus propios estándares, cómo concibe al referente. Contrastando esta opción de búsqueda con los diccionarios enumerados, se considera que el diccionario inverso puede ser una posibilidad válida. En otras palabras, el funcionamiento del diccionario inverso ideal consiste en dar una definición del término basada en los recursos del hablante, y a partir de ella buscar en el diccionario onomasiológico una entrada. Sierra señala que la búsqueda onomasiológica ideal sería:

*The ideal onomasiological search must allow writers to input the concept to be searched through the ideas they may have, using any words in any order. The system must be so through the ideas they may have, using any words in any order. The system must be so constructed that it accepts a wide range of words which it then analyses in order to point the user to the word that most closely approaches the concept he had in mind when he started the search. (Sierra (1999), 34).*

Es decir, la metodología ideal debe tomar en cuenta como punto de partida al usuario. De lo anterior se derivan nuevos retos: el tratamiento de lenguaje natural, el formato adecuado, los tipos y opciones de búsqueda, el diseño y las herramientas que servirán para el desarrollo, entre otros.

A continuación se describirán varios tipos de diccionarios electrónicos que serán de utilidad para la metodología que se desarrolla en este trabajo.

### **2.3.2.3 Diccionarios onomasiológicos electrónicos**

Los diccionarios electrónicos, como su nombre lo indica, poseen un formato electrónico. La información que contienen no se presenta impresa en papel sino que se trata de información que se recupera a través de un sistema computacional.

Al igual que los diccionarios impresos, los diccionarios electrónicos pueden ser ideológicos, tesauros, etc. Sin embargo, su naturaleza los provee de varias ventajas sobre los impresos, debido a que pueden actualizarse con mayor facilidad y rapidez, y son capaces de contener una mayor cantidad de información. Estos diccionarios facilitan, a su vez, la interacción con recursos multimedia (imagen, sonido o video). Algunos de ellos permiten diferentes tipos de búsqueda, como la sinonímica, la inversa, la búsqueda por operadores lógicos o palabras relacionadas, entre otros. Los diccionarios electrónicos cuentan también con recursos como la lematización automática de su información o herramientas que sobreextienden la búsqueda para extraer antónimos, colocaciones o palabras relacionadas.

Uno de los primeros diccionarios onomasiológicos electrónicos fue el *Casey's Snow Day Reverse Dictionary* desarrollado por Los Alamos National Laboratory<sup>13</sup>, el cual intenta resolver el problema del olvido de las palabras a partir de una descripción de la cosa olvidada. A través de una técnica conocida como *análisis de n-gramas*, el sistema busca correspondencias entre las letras introducidas y su propio acervo, para posteriormente extraer una lista de 48 términos sencillos ordenados por similitud con la medida de los *n-gramas*.

Actualmente se han desarrollado nuevos productos basados en diferentes técnicas; Hartmann<sup>14</sup> documenta los siguientes tipos: los que muestran relaciones sinonímicas (como *WordNet*), constelaciones semánticas (como *Visual Thesaurus*) y contextos de uso (como *FrameNet*), entre otros.

Los diccionarios electrónicos que ofrecen consulta onomasiológica pueden clasificarse principalmente: a) por el tipo de recursos que emplean exteriormente, es decir, por aquellas opciones que el usuario puede manipular; b) por el tipo de lenguaje aceptado; c) por la forma de presentar la información (texto, palabra dentro de un diagrama o imagen); d) por el tipo de usuario, e) por la forma de consulta o f) por la temática de su vocabulario.

Con el fin de delimitar la metodología, en este trabajo se contemplan tres clasificaciones: el lenguaje aceptado, su formato de consulta y el tipo de recursos que emplean.

### 2.3.2.3.1 Lenguaje aceptado

El tipo de información escrita que se introduce en un diccionario electrónico puede ser de dos tipos: descripción libre y palabras claves.

<sup>13</sup> Para profundizar el tema consúltense Sierra (1999), p. 30.

<sup>14</sup> Hartmann, p. 4.

Los diccionarios que aceptan **descripción libre** permiten al usuario formular su búsqueda sin restricciones de expresión. Por ello, la consulta no se ve entorpecida por palabras funcionales o irrelevantes, ya que el sistema tiene la capacidad de eliminar esas palabras internamente, como sucede en el diccionario *Onelook* y parcialmente en el *Wordsmyth*.

Los diccionarios que funcionan a través de **palabras claves** consideran que cada palabra introducida por el usuario es relevante para la búsqueda; por eso no las eliminan, como ocurre con los diccionarios *Vox* y *Longman*. Debido a ello, es recomendable no usar palabras funcionales (preposiciones, conjunciones, etc.) en estos diccionarios y además es necesario adecuarse lo más posible a las categorías gramaticales que el diccionario contemple.

Los diccionarios que admiten palabras clave generalmente permiten el uso de palabras cortadas, esto es, permiten extender la búsqueda a partir de la similitud gráfica de un conjunto de caracteres introducidos. La extensión puede realizarse ya sea hacia la izquierda o hacia la derecha de la palabra, dando así una gama mayor de posibilidades y evitando así que la categoría gramatical empleada por un usuario afecte la búsqueda. La mayoría de estos aconsejan no introducir palabras funcionales.

A continuación se compararán los dos tipos de lenguaje aceptado usando como base la descripción “eso que es verde y ácido”.

La descripción libre aceptará las seis palabras introducidas y probablemente busque en su acervo aquellas palabras plenas como “verde” y “ácido”, o bien extienda su búsqueda incluyendo las palabras funcionales “que”, “es”, “y”.

El diccionario de palabras claves conservará cada palabra y la buscará. Sin embargo, al no encontrar una palabra, puede ocurrir que no traiga resultados ligados a la descripción introducida.

Mediante el empleo palabras cortadas, este diccionario buscaría todo aquello que comenzara con “verd-“ y “acid-”, si es que el usuario considera que ése es el segmento que le sirve. Tal búsqueda cubre una mayor cantidad de posibilidades, pues de esta forma pide al sistema que le regrese resultados tales como “verdes ácidos” o “verdosos ácidamente”.

Las posibilidades anteriores varían dependiendo del funcionamiento de cada diccionario, pero a grandes rasgos funcionan de la forma indicada.

### 2.3.2.3.2 Tipo de formato

Bajo este criterio se toma en cuenta el formato electrónico final del diccionario. Existen tres divisiones: diccionarios en Internet, estáticos o diccionarios en cd-rom.



Los diccionarios en **Internet** implican el uso de la red para consultarlos. Normalmente tienen ligas con otros diccionarios en red, lo cual los dota con un repertorio mucho mayor que los diccionarios en disco o los estáticos. Una de las mayores ventajas que estos presentan es la posibilidad de actualizarlos frecuentemente. En esta categoría se pueden clasificar los diccionarios *Onelook* y *Wordsmyth*.

Los diccionarios **estáticos** son aquellos que se pueden descargar de la red o de un disco y no necesitan de éstos durante su consulta, ya que las bases de información están cargadas en la computadora. En esta categoría se puede mencionar al diccionario *Vox*, al *RAE*, o al *e-Großwörterbuch Deutsch als Fremdsprache*.

Los diccionarios en **cd-rom** necesitan que el disco esté en la computadora cuando se consulta el diccionario, ya que éste no puede ser cargado totalmente en la computadora; como ejemplos está el diccionario *Longman* para el inglés y *lo Zingarelli 2003* para el italiano.

### 2.3.2.3.3 Recursos exteriores

La tercera clasificación, los recursos exteriores, toma en cuenta las herramientas que proporciona el diccionario y que pueden ser manipuladas por el usuario. Engloba tres tipos: operadores lógicos, filtros, y combinación de operadores lógicos y filtros.

En los diccionarios que utilizan palabras clave, estos recursos son visibles para el usuario y éste tiene que saber utilizarlos. Por el contrario, los diccionarios electrónicos de lenguaje natural aceptan cadenas de palabras no importando el estilo y recursos de la persona, y generalmente no ofrecen filtros ni operadores al usuario. Una vez iniciada la búsqueda, devuelven una interpretación de ésta. En otras palabras, el usuario debe entonces introducir la información y es el sistema el que internamente se encarga de aplicar los filtros y operadores lógicos. Por ello, este tipo de consulta evita que el usuario aplique erróneamente estos recursos exteriores. Si cuenta con la suficiente información y un motor de búsqueda eficaz, es posible que el diccionario evoque aquello que el usuario busca.

Los diccionarios que proporcionan consulta a través de **operadores lógicos** o boléanos emplean tres decisiones básicas: conjunción (Y/AND), disyunción (O/OR), y negación (No/Not). La búsqueda toma en cuenta estas relaciones, ya se presenten solas o combinadas, y extrae aquello que satisfaga una condición dada.

Cabe señalar que la mayoría de los diccionarios que utilizan palabras clave, de la clasificación de lenguaje aceptado, también utilizan operadores lógicos para relacionar estas palabras clave.

Siguiendo el ejemplo anterior, la descripción “eso que es verde y ácido”, luego de depurarse, puede buscarse de distintas maneras:

- Verde Y. ácido
- Verde O. ácido

En el primer caso, el diccionario buscará un término que comprenda en la definición *verde* pero también *ácido*, pues es forzoso que ambas condiciones se cumplan. En el segundo caso, cualquiera de ellas puede cumplirse.

Es importante señalar que, si las características proporcionadas están relacionadas de forma incorrecta, difícilmente se extraerá lo buscado. Por otro lado, si son muy generales, el número de palabras extraídas será muy grande.

El siguiente tipo de recurso exterior es el que ofrece filtros. Por **filtro** se entiende aquella condición que debe cumplir un conjunto de datos y cuya función es la delimitación y depuración de información buscada. Debido a que en este caso se trata de una consulta lexicográfica, los diccionarios pueden ofrecer algunos de los siguientes filtros:

- gramaticales: especifican la categoría gramatical que será devuelta al usuario (verbo, sustantivo, adjetivo, etc.);
- por registro o fuentes: busca una palabra dependiendo su origen de uso (hablado/ escrito, libros, revistas, etc.);
- temáticos: busca en un léxico definido (ciencias, artes, humanidades, medicina, botánica, etc.);
- número de palabras encontradas: limita el número de resultados para mostrar.

El número de filtros varía de un diccionario a otro, así como su uso. Normalmente, si las opciones de un filtro no son excluyentes (registro, fuentes, etc.), se manejan a través de cajas de opciones (*checkboxes*), pero cuando lo son (categoría gramatical) su control se da por medio de botones de opción (*options*).

La **combinación de filtros y operadores lógicos** intenta aprovechar las ventajas de cada una de éstas, a fin de que el usuario encuentre con mayor probabilidad lo que busca.

### 2.3.2.4 Descripción de búsquedas onomasiológicas en diccionarios electrónicos

Con el fin de explicar las divisiones anteriormente vistas, se presentarán ejemplos de consulta a partir de diccionarios prototípicos, los cuales fueron seleccionados por poseer sistemas de consulta interesantes de evaluar. Además, individualmente proporcionan características extras, las cuales se mencionarán ampliamente en el capítulo cuatro. Tres de ellos permiten la búsqueda en inglés (*Wordsmyth*, *OneLook* y *Longman*), y sólo uno en español (*Vox*). Algunos diccionarios cuentan con varias de las opciones mencionadas; cabe mencionar que para la evaluación del capítulo cuatro se trató de usar las opciones que permitieran mejores resultados en las búsquedas.

Se describirá la consulta por operadores lógicos en el diccionario *Vox*; la consulta por filtros en el diccionario *Longman*; la consulta por combinación de filtros y operadores lógicos en *Wordsmyth* y la consulta mediante lenguaje natural en el diccionario *Onelook*. En los tres primeros se hace uso de palabras claves.

#### 2.3.2.4.1 Descripción de la consulta onomasiológica en el diccionario *Vox*

El diccionario elegido para describir el uso de operadores lógicos fue el *Vox*, que es un diccionario estático, trabaja con la lengua española y funciona a través de palabras claves. Se eligió este diccionario porque posee una sencilla interfaz (figura 2.1) que permite aplicar los operadores de forma sencilla, evitando que el usuario se confunda.

Este diccionario cuenta con la tradicional búsqueda semasiológica, búsqueda inversa (de la última letra hacia la primera), búsqueda de sinónimos, búsqueda múltiple y búsqueda asistida. Para fines de este trabajo se eligió la búsqueda múltiple porque es la que simula el proceso onomasiológico. Tal opción consiste en introducir una serie de palabras que pueden ser relacionadas con operadores lógicos. El diccionario extrae entonces los términos que contengan las condiciones introducidas. Además posee diversos menús: para modificar la forma en que se visualizará la definición; para copiar la información del diccionario; para visualizar las ventanas de la definición; y un menú de ayuda.

La búsqueda múltiple es la que permite consultas onomasiológicas y en este diccionario se realiza a partir de cuatro pasos. Primero se selecciona el tipo de búsqueda desde el menú *Ver*, o bien, directamente en el botón que muestra una letra *a* apuntada por cuatro flechas rojas. Después se introduce la expresión que fungirá como información clave (siempre aparece en letras mayúsculas),

la cual puede ser modificada con operadores lógicos. Finalmente, se especifica dónde se buscará la condición introducida (en la entrada, en los ejemplos, en los localismos o en los tecnicismos).

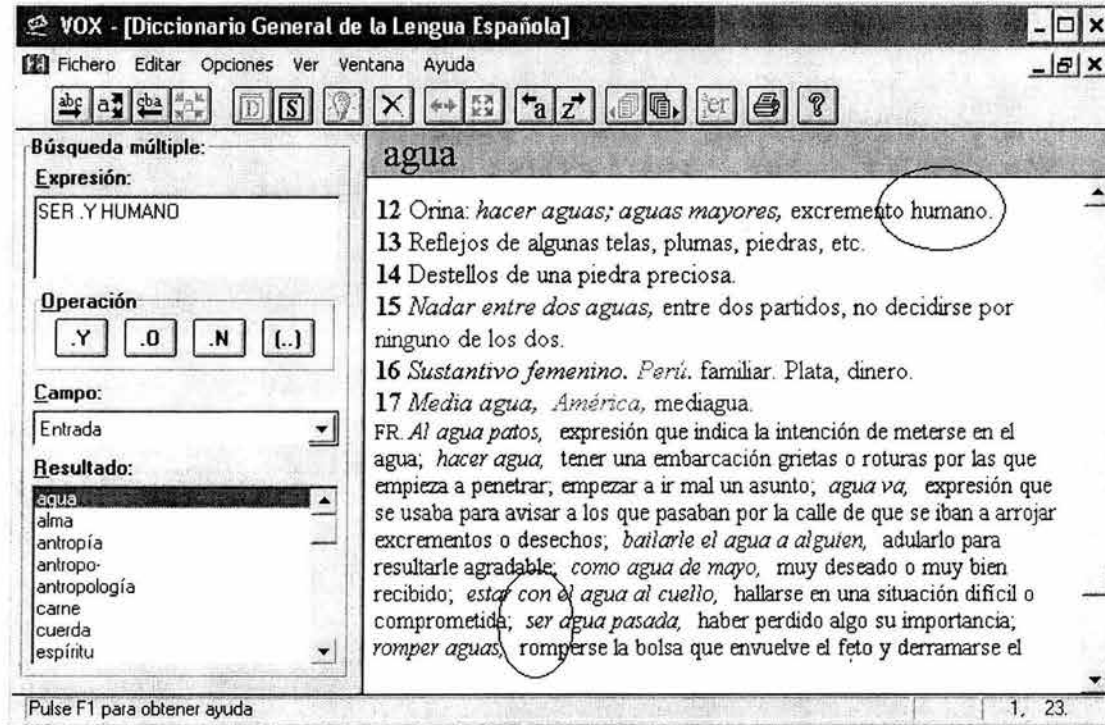


Figura 2.1 Interfaz de consulta del diccionario Vox

A continuación se enumeran los pasos necesarios para realizar una consulta onomasiológica en el diccionario Vox cuando se tiene como expresión: SER .Y HUMANO.

1. previamente se ha cargado el software del diccionario Vox;
2. se cierra la primera ventana, la cual indica con qué cuenta el diccionario;
3. se selecciona la *búsqueda múltiple*;
4. se introduce SER .Y HUMANO con operador lógico de conjunción;
5. se selecciona la opción *entrada* para que la expresión sea buscada en toda la entrada;
6. se pulsa *enter*;
7. se busca el término en la caja de términos relacionados que se encuentra al final de la interfaz del lado izquierdo, y
8. si no se encuentra el término, se comienza desde el tercer punto.

#### 2.3.2.4.2 Descripción de la consulta onomasiológica en el diccionario *Longman*

El diccionario *Longman* permite la búsqueda a través de filtros y está en formato cd-rom. Está diseñado fundamentalmente para los hablantes no nativos del inglés. El diccionario proporciona una gran cantidad de herramientas estilísticas, ejemplos de uso, relaciones semánticas, material didáctico para el estudiante y para el profesor de lengua inglesa. Posee también una gran cantidad de recursos multimedia. Entre sus características especiales está la posibilidad de interactuar con procesadores de texto como *Word*. Además de la consulta semasiológica, ofrece también consulta onomasiológica.

En efecto, *Longman* acepta lenguaje natural para búsquedas onomasiológicas, siempre y cuando todas las palabras introducidas por el usuario se encuentren en la definición del término que se quiere buscar en el diccionario. Por ello, características como *small* y *little* son diferentes a pesar de su semejanza semántica. Con todo, tiene un programa interno de lematización que permite que el usuario introduzca, por ejemplo, *measuring* y busque definiciones con *measure*.

Según la descripción del mismo diccionario, la consulta onomasiológica consiste en introducir una breve descripción en lenguaje natural (Figura 2.2). La consulta puede mejorarse con el uso de filtros y operadores lógicos. Además de ello, la consulta cuenta con una ventaja única en relación con los otros tres diccionarios: el usuario sabe si su información cuenta con referentes al momento de introducir las palabras claves, ya que existe un indicador de términos evocados en la interfaz.

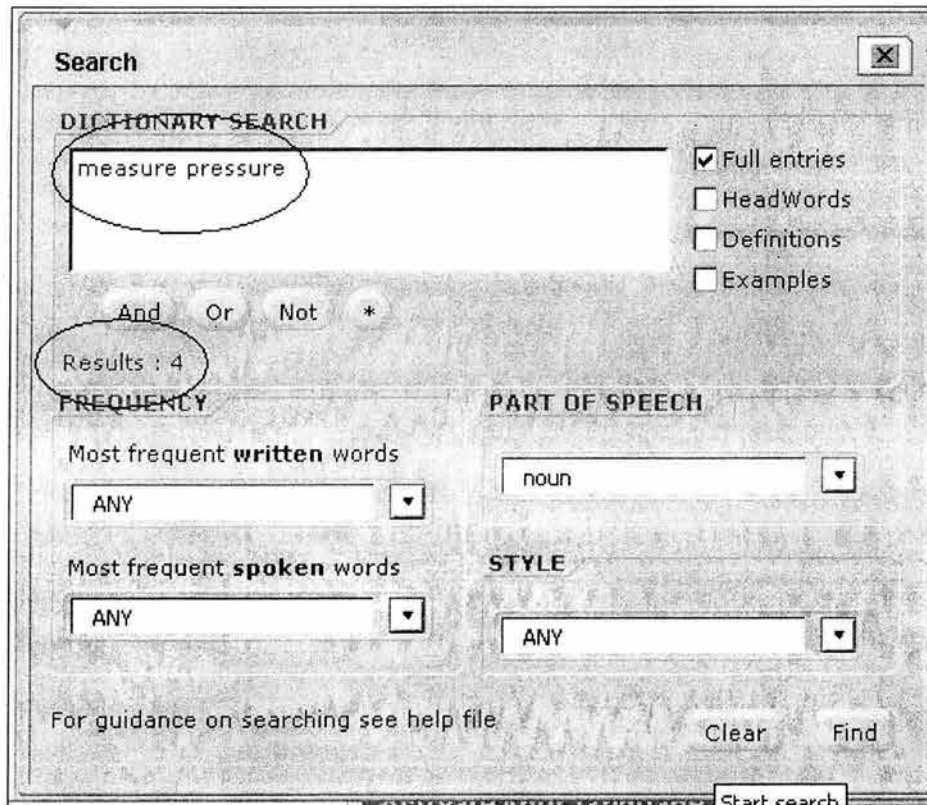


Figura 2.2 Interfaz de consulta del diccionario *Longman*

A continuación se describen los pasos necesarios para realizar una consulta onomasiológica en el diccionario *Longman*, cuando se tiene la descripción *measures pressure*.

1. se carga el diccionario en la computadora,
2. se abre el diccionario,
3. se pulsa el botón *search*,
4. se elige *dictionary search*, acción que abre la ventana *search*,
5. se introduce *measures pressure*,
6. se deja el filtro *full entries*
7. se elige *noun* en *part of speech*
8. se inicia la búsqueda con el botón *find*, que abre una nueva ventana
9. se busca el término,
10. si no se encuentra, se pasa a las siguientes ventanas o se modifica la búsqueda empezando desde el paso tres.

### 2.3.2.4.3 Descripción de la consulta onomasiológica en el diccionario *Wordsmyth*

Con el diccionario *Wordsmyth*, cuya dirección en Internet es <http://www.wordsmyth.net>, se ejemplificará la combinación de filtros y operadores lógicos porque la interacción de ambos recursos tiene buenos resultados y no resulta complicada su formulación. Cuenta con dos posibilidades de consulta: una realizable en el diccionario para niños (*Wordsmyth Children's dictionary*) y otra que se lleva a cabo en el diccionario-tesauro (*Wordsmyth Dictionary-thesaurus*), donde se ofrecen búsquedas semasiológicas y onomasiológicas.

Este último ofrece una consulta onomasiológica bajo el nombre de *advanced search*. En la página electrónica hay un vínculo que permite ingresar a las opciones de esta búsqueda (figura 2.3), las cuales son varias en cuanto a la delimitación de la información buscada y la información devuelta por el diccionario.

fd3003df [details] [logout]

**Wordsmyth**  
Dictionary - Thesaurus

Look up the word...

exact  
 broad  
*"dog" matches "guide dog"*  
 spelled-like  
*"dawg" matches "dog"*

**advanced search**  
[How to Search](#)  
 Helpful hints for using Wordsmyth.

**Wordsmyth**  
Children's Dictionary  
 Look up the word...

**Wordsmyth Advanced Search**

Please separate search terms by spaces or commas. Enclose any phrases within double quotation marks; i.e. "infectious disease".

1. Find all words whose  include(s)

2. Optionally, you may restrict matches to words whose  
 Part of speech is...

3. Choose the result data you'd like to display:

Pronunciations    Definitions    Examples  
 Syllables    Synonyms    Similar Words  
 Inflected Forms    Derived Forms    Cross References

Figura 2.3 Interfaz del diccionario *Wordsmyth*

La consulta onomasiológica en este diccionario se realiza en tres pasos. En el primero se introducen las palabras o frases claves del usuario. A la vez, se selecciona en dónde se buscará tal

información. Es posible seleccionar como campo de búsqueda: la definición, los ejemplos de la entrada, los sinónimos o todas las relaciones de palabra posibles.

El segundo paso de la búsqueda es opcional, en esta parte se puede acotar la búsqueda por categoría gramatical (se especifica o no alguna categoría).

El tercer paso de la búsqueda se refiere al tipo de información que mostrará el diccionario. Por defecto, todos los cuadros de verificación (*checkboxes*) están activados, son nueve opciones que refieren a pronunciación, definiciones, formas derivadas, ejemplos, etc. Una vez que se han seleccionado las opciones requeridas, se inicia la búsqueda a partir del botón *perform search*.

A continuación se enumeran los pasos para realizar una búsqueda onomasiológica en el diccionario *Wordsmyth*, tomando en cuenta la descripción *transparent liquid rivers*.

1. se visita la página [www.wordsmyth.net](http://www.wordsmyth.net);
2. se selecciona *advanced search*;
3. se introduce la definición: *transparent liquid rivers*;
4. se dejan los campos de búsqueda *definition* y *all of the words/phrases* (la búsqueda se prefiere conjuntiva);
5. se elige la categoría gramatical de sustantivo (*noun*);
6. se deshabilitan todas las opciones de información a mostrar menos la opción *definitions*;
7. se inicia la búsqueda con el botón *perform search*, que despliega otra pantalla;
8. se busca la palabra en el listado alfabético;
9. si no se encontró la palabra, se vuelve al punto tres modificando la información.

#### 2.3.2.4.4 Descripción de la consulta onomasiológica en el diccionario *OneLook*

Para ilustrar la consulta a través de lenguaje natural se eligió el diccionario *Onelook*, cuya página es [www.onelook.com](http://www.onelook.com). Además de ofrecer consultas semasiológicas, traducciones automáticas, búsquedas en otros diccionarios, permite también la búsqueda onomasiológica. Se eligió este diccionario porque no existen limitaciones al momento de escribir, acepta sin problemas signos de puntuación e incluso palabras mal escritas.

La consulta onomasiológica (figura 2.4) consiste en introducir la descripción del concepto ya sea a través de definiciones en lenguaje natural o de palabras sueltas, ejemplos, enumeraciones, etc. El sistema analiza los datos introducidos y devuelve un listado de 1000 palabras divididas en páginas de cien términos. El listado está ordenado según la afinidad que existe entre el acervo del diccionario y la información introducida.



La información de cada página es presentada en cuatro columnas, donde los términos son numerados y resaltados en azul. Para consultar el significado de alguno de ellos, se hace clic en el término correspondiente.

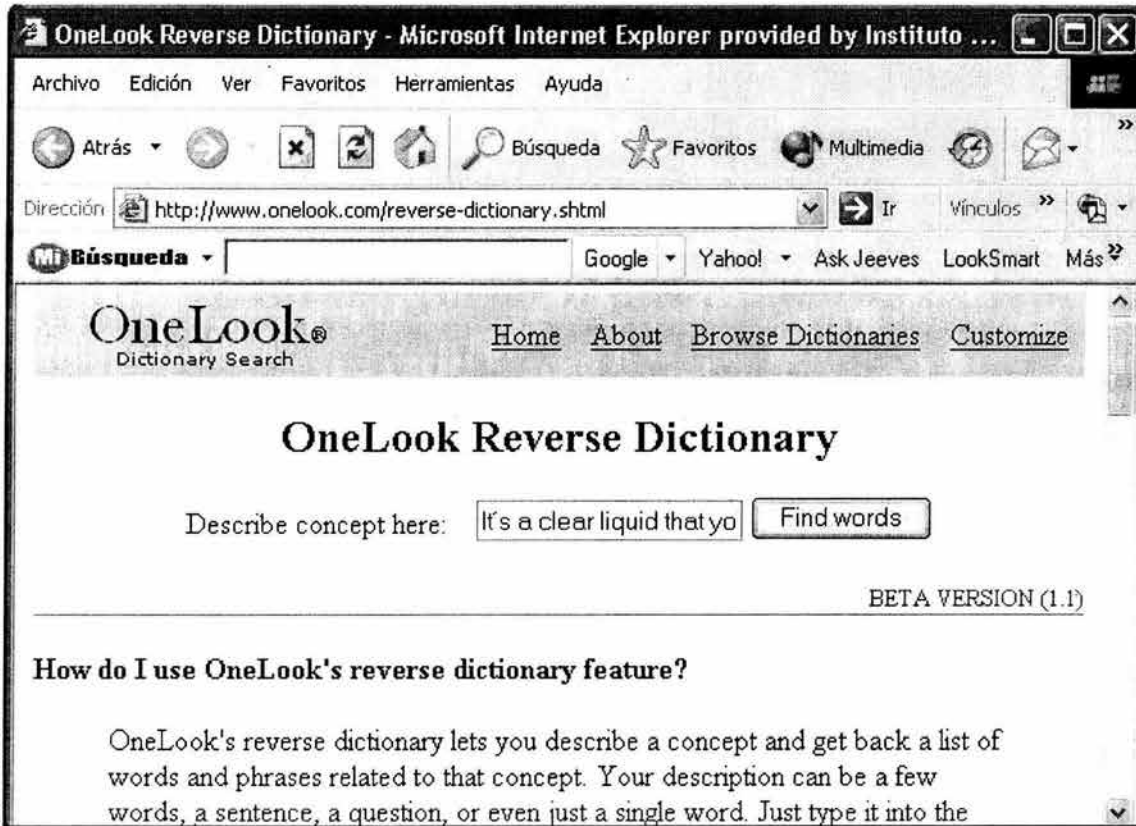


Figura 2.4 Interfaz del diccionario OneLook

A continuación se enumeran los pasos para la búsqueda inversa del diccionario *onelook* cuando se tiene como definición *It's a clear liquid that you get from a tap*:

1. se visita la página: [www.onelook.com](http://www.onelook.com);
2. se selecciona el vínculo *reverse dictionary*;
3. en la caja de texto se escribe el concepto: *It's a clear liquid that you get from a tap*;
4. se comienza la búsqueda presionando *Find words*;
5. se busca el término en el primer listado de 100 palabras;
6. si no se encontró, se puede buscar en los siguientes listados o se puede repetir la búsqueda, modificando la expresión introducida del paso tres.

## 2.4 Descripción del Corpus

Ya que el objetivo fundamental del trabajo es garantizar que el prototipo final sepa interpretar la información proporcionada por un hablante común, es necesario tanto conocer acerca del proceso de conceptualización, como evaluar los diccionarios anteriormente descritos. Para ello, se recurrirá a un corpus bilingüe que permita ahondar en ambas cuestiones. A continuación se describirán la metodología y características del corpus empleado.

Se trabajó con un corpus bilingüe que comprende diez conceptos, los cuales varían de complejidad y campo semántico<sup>15</sup>. Para cada término hay diez definiciones en inglés y diez definiciones en español. Los términos elegidos fueron: *limón/lemon*, *ardilla/squirrel*, *agua/water*, *banca/bench*, *cubeta/bucket*, *ropa/clothes*, *monopolio/monopoly*, *eutanasia/euthanasia*, *huracán/hurricane* y *barómetro/barometer*.

### 2.4.1 Naturaleza de los términos

La selección de términos permite observar además las oposiciones: vocabulario general/terminológico, concreto/abstracto y singular/colectivo. Cabe señalar que se procuró escoger aquellos conceptos que pertenecieran al vocabulario de la gente aunque sea pasivamente, lo cual aumentaría la posibilidad de que todos los términos preguntados fueran definidos.

Los términos elegidos tanto en español como en inglés corresponden a las formas dialectalmente apropiadas para cada país. Por ello, en México se buscó *cubeta* y no *balde* o *cubo*.

### 2.4.2 Descripción del ejercicio de conceptualización

El ejercicio de conceptualización sirvió para recopilar definiciones de estudiantes universitarios del Reino Unido y de México<sup>16</sup>.

El primer paso consistió en dividir al grupo de alumnos en dos mitades. A la mitad A se le dio cinco términos del listado, a la mitad B los cinco restantes. El profesor pedía entonces que los alumnos leyeran el listado y que definieran el concepto usando sus propios recursos pero no debían bajo ningún motivo mencionar la palabra a definir. Así, los estudiantes definieron el concepto en

<sup>15</sup> Para observar las características particulares de los términos, consúltese la tabla 2.1 en el apéndice, construida a partir de la red semántica Eurowordnet.

<sup>16</sup> Diez definiciones en inglés fueron recolectadas y utilizadas por Sierra (1999) para su tesis doctoral. Las definiciones en español fueron recopiladas también por él, como parte de un experimento en una clase universitaria; cuenta con 15 definiciones por cada término, se han tomado sólo 10. En los apéndices se pueden consultar las definiciones a partir de la tabla 2.2.

otra hoja sin mencionar el término. No se permitió ayuda entre los estudiantes, cada uno definió dependiendo de sus capacidades. Asimismo, no se permitió que supieran cuáles eran los términos definidos por la otra mitad de estudiantes.

La segunda parte del ejercicio consistió en identificar la referencia de las definiciones. Para ello, se recogieron las definiciones de cada mitad, de manera que a la mitad A se le entregaba las definiciones de la mitad B y viceversa. En esta ocasión, los estudiantes recibieron las definiciones sin el listado de palabras, por lo que debían adivinar y escribir a qué hacía alusión la definición.

De esta manera, el ejercicio cumplió dos propósitos. Uno, observar el tipo de descripciones conceptuales dadas por los alumnos, lo que se utilizará para evaluar los diccionarios onomasiológicos existentes. Dos, observar el fenómeno de reconocimiento de palabras, esto es, la parte última del proceso onomasiológico.

### 2.4.3 Resultados del ejercicio de conceptualización

Como se observa en las tablas del apéndice del capítulo dos, puede suceder que el primer alumno (el que define):

- sea preciso y construya una definición inteligible; o
- no conozca con certeza las palabras adecuadas pero mencione suficiente información del término; o
- mencione información irrelevante de la palabra, insuficiente, escasa o errada.

En la segunda parte del ejercicio de conceptualización, puede ocurrir con el segundo alumno (el que identifica) que:

- cuente con el conocimiento suficiente como para interpretar la información y acertar al término; o
- entienda la definición pero no conozca el referente; o
- conozca el término definido pero al no haber suficiente información mencione otro término del mismo campo semántico; o
- identifique un término totalmente apartado del definido porque la definición no fue precisa; o
- no proponga ningún término.

Tomando en cuenta lo anterior, se obtuvo por parte de los estudiantes casos en los que 1) la definición y la identificación fueron buenas, por lo tanto hubo éxito; 2) la definición o la identificación fueron inseguras, por lo que se proponen varios términos; 3) se propuso otro término debido a la vaguedad de la definición o al desconocimiento del informante; y 4) ni la definición ni la identificación condujeron al éxito. Lo anterior se puede ejemplificar por medio de la tabla 2.1:

Caso	Clave	Definición	Identificación
1	5ardE	Roedor que trepa a los árboles, tiene cola larga y come bellotas.	Ardilla
2	2ropI	The collective name for the items which we wear; i.e. trousers, shirts, jumpers, etc.	clothes, attire, gear, rags
3	7hurE	Precipitación natural de niveles mayores (nivel 5) que ocasionan catástrofes, principalmente en zonas costeras. Esta precipitación se caracteriza porque a su centro se le llama ojo.	Maremoto
4	10barI	Used to measure pressure, e.g. weather reports	--

Tabla 2.1 Funcionamiento de las definiciones

A grandes rasgos, los informantes ingleses en varias ocasiones proporcionaron no sólo un término sino un listado posible de palabras relacionadas como referentes. Tales listados evidencian que la definición disparó características generales de un grupo semántico y por ello fue necesario mencionar más de un término; lo anterior debe tomarse en cuenta porque de esta manera trabajarán los diccionarios electrónicos como se verá posteriormente.

#### 2.4.4 Metodología general del corpus

Para fines de esta investigación se ha asignado una clave a cada definición, que consiste en numerar la definición del corpus (del 1 al 10), seguido de las primeras tres letras del término en español y al final se indica el idioma con una letra mayúscula: I para el inglés y E para el español. Es decir, la clave "7ardI", se refiere a la séptima definición de ardilla para el inglés.

Las tablas, que se encuentran en el apéndice, seguirán generalmente el siguiente orden en las columnas: clave, definición, palabra obtenida del usuario (fase dos del ejercicio anteriormente descrito). Las columnas anexas serán explicadas en su momento pues dependen del diccionario consultado.

## 2.5 Conclusiones

El panorama teórico visto aquí delimita la opción de búsqueda que se quiere proporcionar al usuario: el lenguaje natural. La búsqueda en lenguaje natural resulta sencilla y cómoda para el hablante, por eso se debe privilegiar esta opción. Eso implica tener una metodología que ayude a clasificar la información introducida, y además relacionarla con una conceptualización en el diccionario.

La construcción de diccionarios onomasiológicos posiblemente en versión electrónica sea más fructífera que en versión impresa, pues permitirán la actualización de la información de una forma continua. Según las prestaciones del sistema, se puede agilizar la consulta y permitir otras formas de búsqueda.

# CAPÍTULO 3

## CONCEPTUALIZACIÓN DE BÚSQUEDAS EN LENGUAJE NATURAL



En el capítulo dos se presentaron los aspectos teóricos relacionados con la onomasiología y los diccionarios onomasiológicos. Además, se ilustraron formas de consulta en cuatro diccionarios y se describió el corpus que se usará en este y en el cuarto capítulo.

Para el diseño del diccionario onomasiológico es necesario conocer más acerca del proceso conceptual que realiza el hablante para formular las búsquedas. Por ello, en este capítulo se presentan los aspectos teóricos de conceptualización, con los cuales se pretende mostrar un panorama general tanto de la conceptualización (3.1), como de la unidad misma, el concepto (3.2). Posteriormente, se hablará de los factores en el hablante (3.3) que modifican el proceso conceptual. Posteriormente, se hablará de la definición (3.4), dado que en este trabajo se concibe como un vínculo entre el concepto y la referencia. Finalmente, se retomará el corpus para contrastar el marco teórico con la praxis (3.5) y para obtener así aquellos elementos teóricos que permitan desarrollar una metodología útil en la onomasiología.

Para el diseño del diccionario onomasiológico es necesario conocer más acerca del proceso conceptual que realiza un hablante para realizar las búsquedas en un diccionario onomasiológico. Por ello, en este capítulo se presentan los aspectos teóricos de conceptualización, con los cuales se pretende mostrar un panorama general tanto de la conceptualización (3.1), como de la unidad misma, el concepto (3.2). Así también, se hablará de los factores en el hablante (3.3) que modifican el proceso conceptual.

Posteriormente, se hablará de la definición (3.4), dado que en este trabajo se concibe como un vínculo entre el concepto y la referencia. Finalmente, se retomará el corpus para contrastar el marco teórico con la praxis (3.5). Lo anterior se realiza para obtener aquellos elementos teóricos que permitan desarrollar una metodología útil en la onomasiología.

Debido a la estructura del capítulo, los primeros apartados resultan en su mayoría teóricos, aunque cabe señalar que se ha buscado formar un panorama general y no cuestionar profundamente los términos introducidos.

### 3.1 Escuelas teóricas de la conceptualización

Hablar de conceptos, palabras y significados remite a los mecanismos de abstracción que el ser humano debe hacer para entender, ordenar y comunicar la realidad en la que vive. La organización o conceptualización resultante es en su mayoría socialmente aprendida.

Si se acepta que uno abstrae su realidad y la ordena, entonces es posible suponer que existen conceptos ordenados quizás en categorías que constantemente se emplean al comunicarse.

Los procesos mentales llevan a cabo la tarea de ordenamiento y clasificación de la información recibida, estos pueden ser denominados **actos mentales**. Sin embargo, la naturaleza y funcionamiento de éstos no es clara para los especialistas. Por ejemplo, Grossmann<sup>17</sup> supone que hay dos clases de actos; unos generados a partir de las percepciones –experiencias sensoriales del sujeto–, y otros formados intelectualmente como resultado de la información que ya tiene el sujeto.

Existen varias disciplinas que desde distintas perspectivas y objetivos han estudiado el proceso de la conceptualización, la existencia de los conceptos y los problemas del significado a continuación se hará un breve recorrido por algunos de estos enfoques. Como la interacción entre áreas ha generado aproximaciones similares al proceso de conceptualización, posturas parecidas pueden verse entre investigadores de materias diferentes.

---

<sup>17</sup> R. Grossmann (1975), "Definite descriptions" en *Philosophical studies* 27, Reidel Publishing Company: Dordrecht-Holland, p. 138.

### 3.1.1 Mentalismo

El mentalismo, una de las escuelas más difundidas y atacadas, ha tratado de explicar la naturaleza de los conceptos dando un papel preponderante a la mente, sus fundadores son Katz y Fodor. La escuela afirma que los conceptos son entidades de la mente que se componen de representaciones semánticas innatas. Es decir, se concibe un lenguaje propio de la mente que se manifiesta en una lengua natural determinada. Si tal supuesto es cierto, se desprende que la información de las representaciones es la misma entre hablantes de distintos entornos. En el mentalismo, afirma Putnam, el significado de una palabra es la suma de una cadena de marcadores semánticos.<sup>18</sup>

Katz usa el término **marcador semántico** para referirse a aquella información innata que no corresponde a una palabra del lenguaje natural, sino del mental. Además de ello, el modelo de Katz presupone una serie de reglas discursivas, que permiten las posibles lecturas de oraciones completas.

Fodor afirma que algunas relaciones son necesarias. Aunque el hablante desconozca ciertas categorías, es posible que emplee conceptos ligados a aquello que no conoce, “es necesario que 2 sea un número primo; pero seguramente, se puede tener el concepto 2 y no tener el concepto de un primo”<sup>19</sup>. Sin embargo, como Aitchison señala, en el mentalismo existe el problema del significado de la representación mental y del aprendizaje del entorno<sup>20</sup>, ya que el segundo no tiene cabida si difiere entre los seres humanos.

### 3.1.2 Argumento Estándar y Atomismo Informacional

El Argumento Estándar (AE) considera que aprender un concepto implica un desarrollo inductivo en el que el hablante crea y prueba hipótesis de las características posibles de un concepto, de esta forma los organiza. Su contraparte es el Atomismo Informacional (AI), el cual consiste en el hablante está situado en una relación mente-mundo y más que saber algo referente a su realidad, posee conceptos cuyo contenido depende de tal relación<sup>21</sup>.

---

<sup>18</sup> H. Putnam (2000) *Representación y realidad*, Barcelona: Gedisa, p 15.

<sup>19</sup> J. Fodor (1998), *Conceptos*, Barcelona: Gedisa, p. 73.

<sup>20</sup> Aitchison, p. 46.

<sup>21</sup> *Ibidem*, pp. 175-176.



### 3.1.3 Semántica Informativa y Procedural

Según Fodor, la Semántica Informativa afirma que los conceptos se diferencian por las propiedades que denotan, éstas a su vez se diferencian por las relaciones que necesariamente se dan entre sí. Es posible, según esta teoría, “tener el concepto DOS pero no el concepto PRIMO”<sup>22</sup>.

### 3.2 El concepto

Como cualquier teoría, los modelos vistos tienen ventajas y deficiencias, algunas de ellas ofrecen respuestas completas a problemas de abstracción y significado. A pesar de sus diferencias, todas coinciden en que debe existir una unidad básica que es el concepto, la cual se piensa como un almacén mínimo de información, que en interacción con otros permite la conceptualización.

Como varios especialistas hacen ver, la definición de un concepto no es clara. En los siguientes apartados se abordarán diversos aspectos del mismo con la finalidad de enumerar los distintos factores que intervienen en la formación y desarrollo del concepto.

El punto de partida para la conceptualización es el concepto o la representación mental, según sea la perspectiva tomada en cuenta. A grandes rasgos, el concepto puede definirse como el conjunto de información acerca de un objeto que el sujeto podría comunicar lingüísticamente.

Según Putnam<sup>23</sup>, el concepto está asociado con su referencia, la cual es contingente y varía dependiendo del mundo o la cultura. Además de ello, no puede ser equiparable al significado de una palabra, porque ello implicaría que el significado es mental y está ceñido por el mundo.

Según Fodor<sup>24</sup>, un concepto es un constituyente del pensamiento y a su vez de otros conceptos más. Asimismo, considera que es de carácter público, lo cual garantiza la comunicación.

Para Engelkamp<sup>25</sup>, el significado de concepto se desprende del de representación. Considera que existen dos tipos de representaciones: unas refieren a palabras como eventos físicos (significante siguiendo a Saussure); las otras contienen información asociada al significado de la representación de la palabra (significado). Basándose en lo anterior, el concepto es la representación del significado.

---

<sup>22</sup> Ibidem, p. 112.

<sup>23</sup> Putnam (2000), pp. 49, 52.

<sup>24</sup> Fodor, p. 51.

<sup>25</sup> Engelkamp, p. 18, 23-26.

Seiler<sup>26</sup>, haciendo un balance de estudios teóricos y aplicados en lingüística, considera que los conceptos pueden concebirse como categorías que sirven para ordenar el mundo y para actividades cuya finalidad sea resolver un problema. Es decir, conforman la memoria en sí. Para establecer esas categorías, el ser humano se basa, en primera instancia, en su percepción. Posteriormente, crea nuevas formas de ordenamiento, basadas en sus capacidades intelectuales.

### 3.2.1 El concepto y la percepción

La percepción es un proceso fundamental en la formación de los conceptos, ya que es por medio de los sentidos que el ser humano interactúa con el mundo exterior.

La percepción consiste en abstraer datos sensoriales, los cuales provienen de diferentes estímulos (visuales, auditivos, táctiles, gustativos u olfativos). Tales datos permiten al sujeto formarse una imagen del objeto aunque no lo tenga presente, y le permiten referirse en futuras ocasiones a éste.

Las percepciones son representaciones internas que mantienen una independencia mínima de las situaciones concretas, y que permiten al sujeto ordenar y explicar la realidad exterior e interior de formas nuevas.<sup>27</sup>

Según Aitchison<sup>28</sup>, la relación entre el concepto y la percepción es básica para la asociación de un concepto a una referencia, ya que percibir a través de los sentidos completa distintos ángulos de una representación.

En resumen, la percepción, además de abstraer información del mundo, le permite al ser humano referirse a objetos que pueden o no estar presentes, o que al ser nuevos para él, pero poseyendo características de alguno ya experimentado, puede identificar y categorizar.

---

<sup>26</sup> B. Seiler, W. Wannemacher (1983) "How can We Assess Meaning and Investigate Meaning development: Theoretical and Methodological Considerations from an Epistemological Point of View" en *Concept development and the development of the word meaning*, Berlin: Springer series in Language and Communication (vol. 12), pp. 320-321.

<sup>27</sup> Seiler, p.322

<sup>28</sup> Aitchison, p. 42.

### 3.2.2 El concepto y la referencia

A pesar de que existen modelos que describen casi completamente el proceso de percepción, no ocurre lo mismo con el proceso inverso, la unión de la información conceptual al referente. La conexión para algunos lingüistas no es clara y ha sido motivo de numerosas discusiones.

Engelkamp<sup>29</sup> considera que el hablante une la referencia al concepto de diferentes formas. Según él, el sujeto reconoce características apreciables a través de los sentidos (como tamaño, color, textura) y las va clasificando en categorías más abstractas o intelectuales no relacionadas directamente con el conocimiento sensible (vertebrado, canino, etc.); el contenido del concepto se compone entonces de la información perceptible que posee el hablante y la información que ha ido aprendiendo sobre el mundo, de manera que el sujeto puede reconocer los distintos referentes.

Sin embargo, como lo señala Putnam, ocurre con frecuencia que un mismo referente tenga diferentes representaciones mentales. Pone como ejemplo la palabra *oro*; para el hablante promedio contiene información general (es un metal con tales características), pero para un geólogo, un químico o un joyero, la representación de *oro* consistirá en las características ya mencionadas más el conocimiento específico de cada área<sup>30</sup>.

Además de la anterior problemática mencionada por Putnam, también existen una serie de creencias, que sin ser información percibida o intelectual, interfiere en la formación de los conceptos.

### 3.2.3 El concepto y las creencias

Es un hecho que la experiencia sensorial influye en la construcción de conceptos; sin embargo, la influencia de las creencias sociales o culturales interfiere de igual manera en el contenido de los mismos. Tal influencia puede verse en palabras cuyo referente es inexistente, como *bruja*, ya que el conocimiento social ha definido la apariencia física, el comportamiento y demás características de ésta. Las creencias se comparten a lo largo del tiempo dependiendo de la cultura y sociedad en las que se desarrolle el ser humano. Por ejemplo, en lengua española han ido quedando estas expresiones: *es una bruja*, *parece bruja*, *estoy bien bruja*.

---

<sup>29</sup> Engelkamp, p. 25.

<sup>30</sup> Putnam (2000), p. 53.

Fodor considera que existen conceptos cuya existencia se difunde independientemente del tipo de cultura, el momento histórico, la sofisticación intelectual, y las experiencias de aprendizaje, como *perro*, *triángulo*, *casa*, aunque su contenido no sea del todo similar<sup>31</sup>.

### 3.2.4 El concepto nodo

Al haber mencionado las principales características del concepto, conviene ahora describirlo como una entidad susceptible de interacción.

Si se piensa que los conceptos son una red de nodos interconectados, es posible suponer que debido a la diferencia de información que poseen, sus relaciones sean también de diversos tipos. Engelkamp<sup>32</sup>, al seguir su clasificación de experiencia sensorial, contempla tres tipos de relaciones entre nodos: dos especifican el tipo de percepción (atributo estático o cambio de estado) y otra se refiere a interrelaciones con información no sensorial. El nodo conceptual puede verse dividido: por un lado está la marca de la palabra o imagen acústica; por el otro, la información referente al objeto. La suma de nodos relacionados y el mismo nodo conceptual conforman el concepto.

El desarrollo y la complejidad de la red conceptual se han descrito de varias maneras, a continuación se explican dos tendencias generalizadas.

### 3.2.5 El desarrollo de la red conceptual

La construcción de la red neuronal se ha concebido de dos maneras a partir de estudios lingüísticos y neurológicos. Para algunos especialistas, la red va adquiriendo complejidad paulatinamente, el número de conexiones va aumentando dependiendo de la experiencia del sujeto. Para otros, implica un proceso de error y corrección, lo cual significa que algunas interrelaciones desaparecen y otras se consolidan.

Según Lara<sup>33</sup>, el desarrollo de la red neuronal puede verse como un proceso individual o colectivo. En el primero, se considera que es el sujeto el que se encarga del desarrollo. En el segundo, se otorga un papel primordial a la sociedad, ya que, según él, el desarrollo ocurre gracias a la continua aculturación social, en la que el sujeto –inmerso en un contexto social y cultural definidos– va adecuando sus estructuras a las de su entorno social.

---

<sup>31</sup> Fodor, p. 52.

<sup>32</sup> Engelkamp, pp. 24-25.

<sup>33</sup> L. Lara (1996), *Teoría del diccionario monolingüe*, México: COLMEX, p. 172.

Ambas formas incluso pueden interactuar, lo que importa es que estos modelos ponen en tela de juicio la relación entre el contenido del concepto y el significado. El problema es si el significado es parte del contenido del concepto, si se adquiere en la primera experiencia con el objeto o incluso si el significado puede ser asido en los conceptos.

### 3.2.6 El concepto y el significado

Debido a los distintos componentes de un concepto (la percepción, las creencias, las categorizaciones intelectuales), resulta difícil decir a partir de dónde se compone el significado, o bien, es problemático decir qué características no son certeramente componentes del significado.

Para Fodor, el contenido o significado de un concepto es todo aquello que hereda de otros nodos<sup>34</sup>. Sin embargo, desde una perspectiva más amplia, el significado no puede equipararse con el concepto. Si se considera al concepto como una estructura cambiante, el sujeto al reactivarlo cancela, rehace o establece nuevas relaciones entre nodos.

Desde la perspectiva de Engelkamp, el significado de un concepto tiene diferentes naturalezas. Para este autor, parte del significado se produce al activar el nodo a través de la imagen acústica de éste o al percibir nuevamente el referente. El significado del concepto consistiría en la interacción entre la información contenida en un nodo con una imagen acústica definida y la información semántica vinculada (sensible e intelectualmente aprendida)<sup>35</sup>.

A pesar de los estudios, de la información fiable y de la existencia de varias teorías y modelos, el funcionamiento del cerebro humano es en gran parte desconocido. Con el desarrollo tecnológico se han ido desarrollando nuevas herramientas que permiten monitorear la actividad cerebral; a partir de ellas se espera que con el tiempo se cuente con mayor certeza en este campo.

### 3.2.7 Problema del concepto

Actualmente los modelos conceptuales tienen limitantes metodológicas, pues el estudio empírico no es posible del todo. No obstante, para los estudiosos del tema, puede existir un acercamiento, hasta cierto punto, al proceso de conceptualización.

---

<sup>34</sup> Ibidem, p. 143.

<sup>35</sup> Engelkamp, p. 25.

Seiler<sup>36</sup> hace énfasis en que la planeación de los estudios empíricos se realice cuidadosamente debido a que las estructuras mentales del investigador pueden limitar y afectar su metodología. Asimismo, señala que cuando el trabajo teórico no fuerza al empírico, entonces será más valiosa la investigación.

Por otro lado, existen también posturas que se oponen a la existencia de los conceptos. Una de ellas es la de Larson<sup>37</sup>, ya que, según la neurología, no existen datos suficientes que permitan elucidar el tema y, más aún, que afirmen o nieguen fases en el proceso de adquisición del lenguaje.

### **3.3 Factores en el hablante**

Debido a que el ser humano está inmerso en un grupo social, aprende a desarrollarse en un contexto definido. Ésta es la razón por la que crea, desarrolla y describe su conceptualización del mundo a partir de una lengua materna, una posición social determinada, una cultura específica, y también una situación psicológica individual.

Una vez descrito el proceso de conceptualización, se procede a mencionar los factores que lo modifican; éstos refieren al hablante y se pueden clasificar en tres tipos: psicológicos, motivacionales y de habilidad.

#### **3.3.1 Los factores psicológicos**

Los factores psicológicos se refieren no sólo a la relación de afectividad entre el sujeto y el objeto, sino también al proceso de aprendizaje (pasos necesarios que tuvo que seguir para individuar y comunicar su mundo según su propia sociedad).

El hablante, a través del error y la corrección, va paulatinamente aprendiendo aquellas combinaciones sintácticas correctas, además de reacciones pertinentes y poco pertinentes durante las conversaciones. Asimismo, va individuando los objetos que se presentan como un continuum en su mundo; lentamente va separando unidades y partes constitutivas de éstas. Sin embargo, tal proceso depende de los requerimientos sociales de un grupo determinado.

Lara<sup>38</sup> expresa que las primeras acciones que realiza el ser humano se reducen a designar. Posteriormente el sujeto aprende a utilizar connotaciones, formas complejas y significados extensivos.

---

<sup>36</sup> Seiler, p. 329.

<sup>37</sup> R. Larson & G. Segal (1996), *Knowledge of meaning: an introduction to semantic theory*, Massachussets: MIT, p. 551.

Aitchison<sup>39</sup> por su parte, sintetiza las etapas del aprendizaje. Menciona tres problemas fundamentales a los que el niño se enfrenta; a pesar de que están interconectados, pueden separarse en: problemas de etiquetado (cómo y bajo qué palabra designar un objeto *X*); problemas de empaquetamiento (qué información semántica, cultural y social va dentro de cada concepto-palabra), y problemas para crear una red (relacionar la información, esto es, establecer vínculos entre referentes y desarrollar conexiones entre estos que no se presentan explícitamente o que se deben abstraer de la realidad observada).

### 3.3.2 Los factores motivacionales

Los factores motivacionales están vinculados con la necesidad de la información para un hablante y con su funcionalidad. En otras palabras, si se considera la información necesaria para el hablante, y a la vez el hablante depende o se relaciona con el objeto, entonces existe un interés por conocer tal entidad y la disposición del hablante permitirá que la búsqueda sea más próspera.

### 3.3.3 Los factores de habilidad

Los factores de habilidad incluyen el conocimiento del tema, así como la facilidad para expresarlo estilística como precisamente. Las actividades cotidianas, la instrucción y el contexto social son factores que delimitan el léxico, el estilo y los recursos expresivos de una persona. Dos hablantes pueden haber crecido en entornos semejantes, ambos pueden poseer similares léxicos pasivos, no obstante uno de ellos a su vez puede tener un amplio léxico activo y claridad de pensamiento, lo que se percibe por las construcciones verbales que realiza. Sin embargo, el usuario no siempre tendrá un mismo perfil, por ello es importante contemplar que la persona que introduce el concepto en un diccionario onomasiológico pueda ser tanto un experto como un hablante común. Es posible suponer que el primero hará uso de mayores características especializadas, y el segundo empleará conocimiento de la entidad descrita, que un lexicógrafo tradicional podría desdeñar por carecer de rigurosidad y no reflejar las características “esenciales” de la cosa.

---

<sup>38</sup> Lara, pp. 173-174.

<sup>39</sup> Aitchison, p. 170.

### 3.4 La definición

Los tres factores del hablante descritos se manifiestan en el corpus cuando el informante define. Las definiciones proporcionadas pueden verse como vínculos entre parte de la información de un concepto y una referencia. En este apartado se hará una revisión de la definición desde un punto de vista lingüístico.

Desde la antigüedad se ha cuestionado la naturaleza del significado y la definición. Para algunos lingüistas, es imposible hacer equivalente el significado con la definición, y a juicio de los teóricos, la definición carece de un papel primordial en las teorías del significado<sup>40</sup>. Fodor señala que no hay modelos defendibles de definiciones, las cuales 1) no expresan el significado de la cosa, 2) resaltan rasgos que a juicio de otros lexicógrafos no son sobresalientes de ésta, 3) presentan definiciones tautológicas.

Por el contrario, Putnam considera que la definición suele ser factible en condiciones ideales, no así en el mundo actual<sup>41</sup>. Lara considera que la definición no es equivalente a una llana descripción, sino que los lexicógrafos reconstruyen aquello del significado que sirve a los intereses de la sociedad<sup>42</sup>.

Para algunos lingüistas es claro que el significado va cambiando a través del tiempo; por ello, es insostenible decir que cada objeto tiene sólo una definición y que ésta permanecerá inmutable. Aun así, la lexicografía ha buscado métodos para describir aquellos componentes semánticos perdurables de la cosa.

#### 3.4.1 La definición tradicional

De manera muy general, la definición es la manera en la que se expresa el significado. Tal forma ha sido depurada con el tiempo a través de la lexicografía, ciencia que se ha encargado de desarrollar una metodología en este campo: presentar la forma canónica, la información gramatical pertinente y la definición misma, además de posibles acepciones o expresiones nuevas, entre otras cosas.

La definición tradicional sigue el modelo aristotélico de componer el significado por un *genus* y una *differentia*, lo que en la lexicografía moderna se conoce como género próximo y diferencia

---

<sup>40</sup> Fodor, p. 72.

<sup>41</sup> Putnam 1975, p. 251.

<sup>42</sup> Lara, p. 205.



específica. De “hombre”, el género próximo podría ser *mamífero* y la diferencia específica podría ser *bípido*.

La definición tradicional, a pesar de sus limitantes, puede emplearse en la semasiología. Sin embargo, en la onomasiología esta disposición del significado no parece favorable.

### 3.4.2 La definición en la onomasiología

Los diccionarios onomasiológicos existentes permiten desglosar el significado a través de categorías, sinónimos o dibujos, de manera que la definición es una secuencia de categorías, un equivalente léxico o bien una imagen. En otras palabras, no se basan en la definición tradicional, pues no se le pide al usuario que proporcione un género próximo y una diferencia específica.

Debe considerarse que el hablante, al no poseer una metodología para expresar una definición, proporciona información de la representación mental que en esos momentos le pareció importante o pertinente acerca del objeto. Debido a ello, interesa en este apartado conocer cuáles son los tipos de definición a los que se puede expresar en un sistema computacional. Por ello, se hará una revisión a distintos tipos de definiciones. Cuando se haga referencia a las definiciones del corpus, se empleará el término “definiciones libres”.

En otras palabras, la **definición libre** es la forma empleada por el usuario que describe a su parecer un objeto y puede contener cualquier tipo de información: un prototipo, características esenciales o secundarias, género próximo, etcétera.

Al tomar en cuenta que tanto la definición tradicional como la libre están realizando una abstracción de la realidad, entonces en algún momento deberían coincidir o mencionar la misma información. Sin embargo, no es posible decir que una definición libre posee sólo información verdadera acerca de un objeto; tampoco es imposible llegar al significado usando definiciones libres porque, de lo contrario, se pondría en tela de juicio la comunicación.

### 3.4.3 La clasificación de las definiciones del corpus por tipo

El hablante define partiendo de su propia reserva de conocimiento del mundo. Debido a los diferentes factores involucrados –psicológicos, motivacionales y de habilidad–, cabe esperar que la forma empleada por una persona no sea la misma en todos los casos ni en todos los hablantes. Por

ello, es importante concebir la idea de que no puede existir sólo un tipo de definición o bien, no se debe suponer que sólo un tipo de definición es válido.

Sierra<sup>43</sup>, retomando el trabajo de Ndi Kimbi<sup>44</sup>, señala que, además de la definición tradicional, hay ocho métodos o tipos de definición según la forma de describir o presentar al objeto descrito. A continuación se presentan en la tabla 3.1 los tipos de definición<sup>45</sup>, en qué consisten y finalmente se muestra un ejemplo.

Tipo	Características	Ejemplo
Sinónimo	Uso de equivalente léxico	Ropa: vestido, ropaje, vestidura, traje
Estipulación	Nuevo significado de "X"	Basic selection process: "We understand by "basic selection process" the rule which uses the request to select the class of documents whose tags are logically compatible with the logic and tags of the request"
Síntesis	Por descripción de un sistema de conceptos	Water: "Liquid, clear, boils at 100°C, freezes at 0°C" (8agul)
Extensión	Enumeración o ejemplificación	Clothes: "...i.e. trousers, shirts, jumpers, etc..." (2rop1)
Ostensiva o Ilustrativa	Imagen o descripción definida de un objeto no presente	Water: " It's a clear liquid that you get from a tap (1agul)
Paráfrasis	Ampliación de la palabra descrita	Respuesta: "acción de responder"
Implicación	Contexto explicativo que extiende el significado.	Monopoly: "When a company or individual has a major share of the market they have a ..." (7mon1)
Analogía	Se explica X a partir de Y porque son similares	Barometer: "Like thermometer but measures air pressure" (3bar1)

Tabla 3.1 Tipos de definición<sup>46</sup>

La anterior clasificación permite observar algunas formas de presentar la información, lo cual aplicado al corpus permitirá observar las diferencias o igualdades entre las definiciones libres propuestas.

### 3.4.3.1 El comportamiento del corpus

Partiendo de la clasificación presentada por Sierra y ya explicada en el apartado anterior, se clasificaron las definiciones del corpus con la intención de conocer cuáles eran los métodos de definición más comunes empleados por los informantes. Es de suponer que en una definición exista

<sup>43</sup> Sierra (1999), pp. 44-45.

<sup>44</sup> N. Kimbi, (1994) *The conceptual structure of terminological definitions and their linguistic realisation: with special reference to data and information processing terms*. Tesis doctoral, Manchester, Universidad de Manchester.

<sup>45</sup> Para ver los ejemplos del corpus, véase el apéndice del capítulo tres, de la tabla 3.30 a la 3.49.

<sup>46</sup> Definiciones tomadas del diccionario electrónico *Vox*, el diccionario terminológico del GIL y el corpus trabajado.

más de una forma empleada, por ello al identificar más de un tipo en una sola definición, se han tomado en cuenta los diferentes métodos usados.

De las ocho formas posibles se descarta la definición por **paráfrasis** ya que se pidió expresamente a los informantes no mencionar la palabra ni formas parecidas. La definición por **síntesis** se ha considerado como aquella que introduce categorías del término, está estrechamente relacionada con la definición tradicional. La definición por **extensión** se toma en cuenta cuando enumera parte de los integrantes de la palabra descrita, así como ejemplos de éste. Para la definición **ostensiva**, se tomaron como elementos deícticos a *that, it, this, esto* cuando no eran anáfora; y en el caso del inglés, cuando no se trató de nexos introductor de oración subordinada. Finalmente, en las definiciones por **implicación** se agrupó aquella información que vinculaba la cosa descrita con su entorno, o bien, donde se manifestaba la relación entre el sujeto que define y la cosa, es decir, las implicaciones de la cosa en la vida del hombre. Por definición **anafórica** se toma en cuenta la definición tanto que iguala a dos elementos *A* y *B* como aquella que lo hace y posteriormente menciona su diferencia.

### 3.4.3.2 Las tendencias del corpus

De las siete definiciones esperadas sólo dos fueron muy frecuentes, mientras que las otras aparecieron esporádicamente. Los métodos de definición con frecuencias prácticamente nulas tanto en inglés como español fueron el sinonímico, el de estipulación, el de ostensión, el anafórico y obviamente el método de paráfrasis. Mientras que el método por síntesis como el de implicaciones fueron muy recurrentes en las definiciones de los informantes.

Aunado a lo anterior, en ningún grupo de definiciones el informante ha optado por elegir sólo uno de los métodos de definición lo cual es importante porque la naturaleza del objeto no restringe las posibilidades del informante.

Los métodos más empleados, síntesis e implicación, normalmente concurren. Lo anterior se debe posiblemente a que el informante trata de imitar la tendencia lexicográfica a través de fórmulas como *se dice de, relativo a*. Pero se ve forzado también a proporcionar información de mundo que él considera importante comunicar, y quizás sin total consciencia introduce la relación hombre-objeto a través de formas como *llevamos, nosotros*, y también a través de formas cuyas presuposiciones no implican explícitamente la existencia del ser humano pero que de alguna manera la manifiestan haciendo mención de su mundo como en *se lleva casas por el aire, se encuentran en parques o a las afueras de algún edificio, se usa como condimento para las ensaladas*, etc.

El hablante común carece de la rigurosidad del lexicógrafo al definir entidades, por ello cabe esperar que introduzca información cultural, social o científica del objeto que pretende definir. Un diccionario onomasiológico debe tomar en cuenta lo anterior, no se debe considerar que hay definiciones erradas, sino que la información mencionada está de alguna forma vinculada con el objeto descrito, ya sea mencionando características de éste, su origen o su uso.

#### **3.4.4 La definición libre**

Ahora bien, una vez que se han establecido los tipos de definiciones, conviene retomar a la definición libre. La descripción en ésta puede o no ser estricta, e incluso no necesariamente contiene la información relevante de la cosa –en términos lexicográficos. Con todo, establece vínculos suficientes de significado que permiten entender al receptor cuál es la referencia de la descripción dada.

En otras palabras, la definición libre puede ser lexicográfica, o bien puede tender hacia métodos diversos de definición como los ya vistos. Generalmente no sigue orden al introducir las características mencionadas, y su información puede llegar a ser subjetiva y poco confiable.

Al observar las definiciones libres del corpus, se percibe uno de los principales problemas a los que se enfrentará el diseño del diccionario del presente trabajo: no siempre se contará con información válida, o en ocasiones ésta será extremadamente vaga, lo cual impedirá al receptor o diccionario proponer con certeza un término.

#### **3.5 La presentación de la información en el corpus**

Cuando se comparte información, se puede mencionar una característica de un objeto a través de diversas variaciones, tanto morfológica como semánticamente. Por ejemplo, para referir a la característica *medir* del objeto *barómetro*, se puede decir *medida*, *medición*, *mide*, *estimar*, *lo que determina*, etc. Así, un adjetivo puede estar proporcionando la misma información semántica que un sustantivo, un verbo, un adverbio o incluso que una oración completa. Para un sistema onomasiológico esto es de vital importancia, porque no se sabe con total certeza cuál es la forma que empleará el usuario para presentar su información.

Así como el usuario puede emplear formas reducidas expresables en adjetivos, sustantivos o verbos, también puede introducir elementos estilísticos o discursivos que, si bien mejoran la

presentación de su definición, añaden información irrelevante que pueden afectar los resultados. Por ello, es aconsejable que los usuarios introduzcan menos elementos discursivos en su descripción, a fin de que el sistema no tenga que limpiar la definición separando entre palabras funcionales y palabras plenas.

A continuación se revisará la información contenida en las definiciones del corpus que ejemplifican la diversidad de posibilidades para describir conceptos. La presentación se realiza bajo los siguientes criterios: género próximo, diferencia, datos sensoriales, características científicas, verbos relacionados y marcadores situacionales. Para ejemplificar y clarificar lo expuesto, se mostraran tres tablas por cada criterio<sup>47</sup>.

### 3.5.1 La información por género próximo

En algunos casos resulta sencillo para el hablante establecer el género próximo de la cosa que pretende definir; por ello, las variantes posibles se reducen a un número reducido. Tal caso ocurre en las definiciones de *limón* tanto en español como en inglés, la tendencia fue identificarlo como *fruta* o *cítrico* y generalmente al principio de la definición. Sin embargo, en otros casos se observa que el *genus* no es claro para el hablante común, como en el caso de *monopolio*, donde tanto los informantes ingleses como los mexicanos lo omiten.

Clave	g. próximo
l1imI	fruit
l1imE	cítrico

Tabla 3.2 Gen Limón

Clave	g. próximo
l1ardI	Rodent
8ardE	Animal mamífero, roedor

Tabla 3.3 Gen Ardilla

Clave	g. próximo
l1monI	Boardgame
7monE	concentración

Tabla3.4 Gen. Monopolio

En el caso de ardilla, la mayoría identifica cualquiera de tres géneros próximos. Es interesante observar que en la descripción 8ardE se describen los tres géneros próximos. En un caso, l1ardE, el *genus* aparece hasta el final de la descripción, como una oración unimembre.

<sup>47</sup> Las tablas completas pueden consultarse en el apéndice del capítulo 3, de la tabla 3.50 a la 3.67.

### 3.5.2 La información por diferencia específica

La diferencia específica puede ser variable dependiendo de la categorización y del *genus* propuesto por el informante. En algunas ocasiones resulta imposible para el hablante expresarla por medio de una sola palabra, por ello emplea formas complejas. Las palabras elegidas permiten ilustrar varios casos que pueden ocurrir al momento de definir.

Clave	d. específica
1aguI	clear
1aguE	indispensable

Tabla 3.5 Diferencia Agua

Clave	d. específica
3ropI	fashionable or unfashionable
1ropE	tejida

Tabla 3.6 Diferencia Ropa

Clave	d. específica
2cubI	a handle
3cubE	asa

Tabla 3.7 Diferencia Cubeta

La diferencia específica en *agua*, *indispensable*, ocurre en diferentes formas: desde *necesario* hasta la forma compleja *imposible vivir sin ella*. La anterior característica pudo haberse simplificado a través de un adjetivo; sin embargo, el informante prefirió presentar de esta manera la información.

En la tabla de *cubeta* la diferencia específica no es muy variable, aunque llama la atención que se encuentra relacionada de diversas maneras con el género próximo. Simplemente, en relación con el asa para prender la cubeta, como diferencia específica, en la definición 3cubE aparece anidada al género próximo propuesto por el estudiante; sin embargo esta unión no se evidencia, porque existe de por medio una oración subordinada. Es decir, la relación se da:

1cubE (a) Lugar + con asa para sujetarse

Pero discursivamente la lejanía entre ambos es evidente:

1cubE (b) Lugar + donde se puede almacenar líquidos + o cualquier otra cosa + con asa para sujetarse

En la oración 6cubE la relación entre *oreja/ agarradera* y *cubeta* se establece a través de un verbo que se encuentra en la oración subsecuente:

6cubE (a) [el] recipiente + se sujeta por una oreja o agarradera

Otro caso lo proporciona la oración 10cubE, que introduce la diferencia específica a través de una oración subordinada:

10cubE (a) Recipiente + que cuenta con un asa + y que se utiliza generalmente para desplazar agua

### 3.5.3 La información por datos sensoriales

Los datos sensoriales refieren a información producto de la percepción, indican propiedades físicas de la cosa como tamaño, color, peso, tipo de superficie.

Clave	d. sensorial
1limI	Yellow
1limE	Verde

Tabla 3.8 Dato sensorial Limón

Clave	d. sensorial (la cola)
1ardI	big
10ardE	Grande

Tabla 3.9 Dato Sensorial Ardilla

Clave	d. sensorial
1cubI	Plastic
9cubE	Grande

Tabla 3.10 Dato sensorial Cubeta

En el primero, *limón*, se observa la diferencia de estereotipos entre ambos informantes: para el hablante inglés el *limón* es comúnmente amarillo, para el mexicano es típicamente verde. Respecto del sabor, en inglés existe una relación mucho más estrecha entre el verbo *to taste* y el tipo de sabor que entre el sabor y un estado expresado por el verbo *to be*. Por ejemplo, se encuentra: *sour tasting*, *bitter taste*, *tastes bitter*, *yellow sour fruit*. En español es más común referirse al sabor como una cualidad de una cosa (fruto o jugo): *fruta ácida*, *jugo agrio*.

### 3.5.4 La información por características científicas

La información científica es aquella que, según alguna disciplina como la física, química, matemáticas o biología, se conoce gracias a investigaciones y diversas pruebas. Se trata de información que intenta no ser subjetiva. Para ejemplificar la forma de incluir las características científicas se han tomado ejemplos de las palabras *agua*, *limón* y *ardilla*.

Clave	C. Científica
8 aguI	Boils at 100°, freezes at 0°C
4aguE	Su fórmula química es H <sub>2</sub> O

Tabla 3.11 C. Científica Agua

Clave	C. Científica
10 limI	citrus fruit
4limE	fruta, cítrico

Tabla 3.12 C. Científica Limón

Clave	C. Científica
10 ardI	mammal, rodent
10ardE	Animal, roedor

Tabla 3.13 C. Científica Ardilla

Posiblemente, uno de los conocimientos científicos más difundidos en el mundo consiste en saber la fórmula química del agua. A pesar de que la fórmula es sólo una, las maneras que tiene un hablante para expresarla pueden ser diversas, como lo ejemplifican 4aguI (*constituents are hydrogen and oxygen*), y 4aguE (*su fórmula química es H<sub>2</sub>O*). No obstante, no es el único conocimiento científico que un hablante común pueda transmitir, existen otras particularidades de este líquido que también se mencionan, como el punto de ebullición y el de congelación (8aguI). En algunos casos, las definiciones de *agua* se sostienen en el conocimiento científico, lo cual fuerza al diccionario onomasiológico a considerar características especializadas si es que pretende adecuarse al usuario y no solamente considerar aquellos datos sensoriales o socialmente compartidos.

El conocimiento científico también interfiere en palabras que impliquen clasificaciones dadas por la botánica y la zoología. Sin embargo, un diccionario onomasiológico no debe depender exclusivamente de estas características, debido a que el usuario conocerá sólo en algunos casos a qué taxonomía pertenece tal animal o planta; el segundo problema es que la categoría tomada en cuenta puede ser muy amplia como *animal*, específica como *roedor*, o errónea. Tómese como ejemplo la definición libre 9limE, el usuario ha introducido para *limón* la característica *verdura*, la cual es una característica errónea. Es decir, la definición cuenta con información que desvía la búsqueda. Con todo, es posible llegar a su referente, ya que de tres características, dos son aceptables y una errada:

9limE (a)      *Verdura (No) + de color verde (Si) + Sabor agrio (Si).*

La pregunta es ¿debe un diccionario onomasiológico ofrecer alternativas de búsqueda a este tipo de definiciones? Como se verá en las siguientes unidades y en la misma confección del diccionario, es importante tomar en cuenta que si el usuario no proporciona información confiable, entonces el diccionario deberá ofrecer términos alternativos.



### 3.5.5 La información por verbos y complementos relacionados

El ser humano al hacer uso de su idioma establece sin notarlo ciertas combinaciones entre verbos y complementos. A continuación se dará el ejemplo del comportamiento de algunos verbos en las definiciones, para los conceptos *barómetro*, *ardilla* y *monopolio*.

Clave	Verbo	Complemento
2barI	for measuring	pressure
5barE	medir	la presión atmosférica

Tabla 3.14 Verbo Barómetro

Clave	Ver.	Compl.
2ardI	living	in trees
5ardE	tropa	los árboles

Tabla 3.15 Verbo Ardilla

Clave	Verbo	compl.
2monI	buying and selling	of the property
5monE	decidir	Algo

Tabla 3.16 Verbo Monopolio

Uno de los verbos más ligados a *barómetro* es *medir*. En español como en inglés existen varias formas de presentarlo, ya sea por uso del verbo directamente, o bien, explicitando la funcionalidad a través de sintagmas como “para medir” o en inglés la forma “used to measure” o “for measuring”. Además de las formas perifrásticas y el verbo pleno, también se puede expresar la acción realizada por el barómetro a través de verbos sinónimos o relacionados, como ocurre en 1barI (*detects*), 7barE (*indica*) y 8barE (*recibe*). El complemento expresa la cosa medida, y como se observa, tiene a su vez varias posibilidades, desde modificadores de la cosa como en 1barI (*changes in atmospheric pressure*) o 7barE (*la presión atmosférica del lugar en que se encuentra*), o bien, el usuario tiene conciencia de que el barómetro es un *aparato* que mide algo que no sabe con certeza qué y por eso agrega *datos de medición y así poder interpretarlos* (8barE).

En el caso de *ardilla* se han tomado los verbos *vivir* o *habitar*, se observa que las formas pueden variar tanto en español como en inglés (*to live, to build, to like, to find, habitar, vivir, trepar, encontrarse*). El complemento regido por estos verbos puede constar de uno o varios elementos locativos, como ocurre en 10ardI (*in both countryside and town*) y 7ardE (*en zonas selváticas o bosque o arboleda*).

El último concepto es *monopolio*, el cual fue un concepto muy complicado de explicar para los estudiantes. En inglés sólo hubo dos verbos relacionados, *to buy* y *to sell*, con complementos diferentes, *property* y *land/railways*, aunque, si se observa, el primero complemento puede contener

al otro en una relación hiponímica<sup>48</sup>. En español, los verbos son más variados, por esto, los complementos no se pueden igualar con facilidad. A pesar de que verbos como *tener* y complementos como *algo, cualquier tipo* pueden no acotar contundentemente el significado de la cosa descrita, existen otras palabras claves que sí especifican el significado como *acaparar, control total, dominio o mercado*. En ocasiones, el usuario no empleará la palabra adecuada, ya que se trata de un léxico no familiar para él; por ello debe suplirlo por verbos comodines como *tener, hacer, ser*. En estos casos, el diccionario debe apoyarse en posibilidades sinónimicas tanto de los complementos como del mismo verbo para proponer una serie de términos posibles.

### 3.5.6 La información por marcadores situacionales

Los marcadores situacionales son propios de las definiciones en las que la persona que define no encuentra formas simples para expresar las características de la entidad, o que la misma definición precisa un contexto explicativo con el cual la persona refuerza, explica o da sentido a los elementos introducidos en su definición. Los marcadores situacionales se reconocen porque se presentan a través de formas explicativas como *cuando, como, en la que*. En el corpus, las palabras como *monopolio, eutanasia* se caracterizan porque el informante emplea recurrentemente estos marcadores.

Clave	M. Situacional
7cubI	When
3cubE	Donde se puede

Tabla 3.17 M. Sit. Cubeta

Clave	M. Situacional
7 monI	when
10monE	en la que

Tabla 3.18 M Sit. Monopolio

Clave	M. Situacional
10eutI	If
10eutE	Cuando

Tabla 3.19 M Sit. Eutanasia

Ningún tipo de objeto descrito está exento de ser descrito usando marcadores situacionales. En *cubeta* la situación referida por el informante sirvió para especificar el tipo de recipiente o contenedor del que se hablaba. En casos como *monopolio*, la presencia de situaciones explicativas son más frecuentes, aunque no todas van introducidas a través de marcas situacionales. En *eutanasia*, el informante debe por un lado, definir el tipo de muerte que es *eutanasia* y por el otro, debe justificar el porqué de tal muerte, ya que el concepto conlleva implicaciones éticas.

<sup>48</sup> Los conceptos como *hiponimia, hiperonimia, meronimia y holonimia* serán introducidos en el capítulo cinco.

La importancia de los marcadores situacionales es que establecen una relación de tipo distinto con el objeto pero aún así, según el usuario, pueden ser piezas claves del significado, dejarían entonces de ser palabras irrelevantes para volverse claves en la búsqueda de ciertos términos.

### **3.5.7 Conclusiones de la presentación de la información**

Todas las anteriores categorías no agotan los tipos de información que puede contener una definición libre; pueden delimitarse aún más pero para los fines del presente estudio, el nivel de revisión sólo pretendió esquematizar la complejidad de la definición libre y los posibles problemas a los que se enfrentará un sistema. De alguna forma, un sistema computacional debe contemplar más de una manera de buscar lo que el usuario quiere. Aitchison, al hacer una comparación entre los seres humanos y los libros señala:

“But the biggest difference between a book dictionary and the mental lexicon is that the latter contains far, far more information about each entry. All book dictionaries are inevitably limited in the amount they contain, just because it would be quite impracticable to include all possible data about each word. In any case it is unlikely that anyone has ever assembled the total range of knowledge which could be brought together about one dictionary entry”. (Aitchison, 13)

### **3.6 La aplicación de modelos teóricos al corpus**

Para finalizar este capítulo, se realizará un contraste entre el corpus y varios enfoques que intentan explicar la naturaleza del significado y la conceptualización, los cuales se derivan del marco teórico presentado al inicio del capítulo. A partir de estas comparaciones, se expondrán las ventajas y desventajas que cada enfoque propone para una aplicación onomasiológica.

Las primeras propuestas juzgan que la definición aristotélica (género próximo y diferencia específica) no es suficiente para hablar de significado y además consideran que un componente es el conocimiento sociocultural que la gente tiene de los objetos. Las últimas, principalmente la de Puntam y la Wierzbicka consideran que sólo un amplio panorama puede expresar el significado.

### 3.6.1 El contexto

Para muchos antropólogos y lingüistas el contexto o los factores situacionales –relacionados con los marcadores situacionales– finalmente son los que definen la acepción de una palabra y permiten acercarse al significado. Robinson<sup>49</sup> considera que la objetividad del significado puede alcanzarse mediante el contexto, aunque señala asimismo que existe un número infinito de contextos posibles, por lo que resulta imposible clasificar y estudiar las razones y funcionamiento de cada caso.

El significado, según los lexicógrafos tradicionales, no puede depender del contexto, así que en sus definiciones intentan no mencionar marcadores situacionales en la definición, ya que se supone existe en las palabras un núcleo semántico, que aunque sea mínimo, hace a la entidad ser lo que es.

En el corpus, el contexto puede expresarse a través de complementos adverbiales, locativos o explicativos y de marcadores situacionales. La presencia del contexto se debe a la relación entre sujeto, objeto y entorno, la cual se expresa en las definiciones de los informantes a través de los verbos en primera persona del plural (*usamos, bebemos, hacemos*) o en impersonal y a través de la forma “uno” cuando se refiere a un sujeto humano (*se usa comúnmente, cuando uno tiene*).

Para el hablante es importante transmitir el contexto, porque supone que al hacerlo se garantiza la transmisión del significado. En la tabla 3.20 se muestran algunas definiciones en las que se remarcan aquellos complementos que implican transmisión de contexto o marcadores situacionales.

Clave	Definición
10limE	Fruta ácida, color verde, <b>con que se puede hacer agua; y se usa como condimento para las ensaladas.</b>
7ardE	Animal que vive en zonas selváticas o en bosque o arboleda, que come regularmente nueces o semillas, vive en los árboles y <b>en CU hay muchas</b>
8banE	Artículo donde nos sentamos. <b>Se encuentran en los parques o a las afueras de algún edificio, iglesia o centro donde se reúnen varias personas</b>
4ropI	Things used to cover up and keep warm, <b>used by humans</b>
8eutI	Ending life before natural end, to relieve from suffering <b>due to illness</b>
7monI	<b>When a company or individual has a major share of the market they have a ...</b>

Tabla 3.20 Contexto y marcadores situacionales

<sup>49</sup>E. Robinson (1999), “11 The cognitive foundations of pragmatic principles: implications for theories of linguistic and cognitive representation” en *Language and conceptualization*. Jan Nuyts and Eric Pederson (ed), Cambridge: Cambridge University Press, p. 262.

### 3.6.2 El género próximo y los rasgos semánticos

La definición lexicográfica se basa en la concepción aristotélica, la suma del género próximo y una diferencia o característica particular del objeto. Sin embargo, muchos autores como Ayto<sup>50</sup>, han notado que encontrar el género próximo y la diferencia específica del objeto puede variar notablemente dependiendo de la cultura y sociedad.

Además, al momento de consultar un diccionario, muchas veces el género próximo no es entendido por el usuario ni es de relevancia para él. El objetivo del autor es hacer funcional el significado para el hablante. Por ello, continúa con la propuesta de Pottier, la cual consiste en tomar campos semánticos y clasificar los rasgos semánticos de los integrantes para diferenciarlos. Él señala dos cosas que deben incluirse en una definición funcional (en la que se da la función y no qué es el elemento en sí): la primera, un género próximo de dominio público, que ayudará mucho más a entender el significado y por consiguiente será de mayor utilidad para la gente que consulta un diccionario; la segunda, el listado de características que perceptualmente diferencian un objeto de otro dentro del campo semántico.

La propuesta de Ayto resulta enriquecedora si se tienen grupos con elementos distinguibles y si se tienen características que puedan ser reconocidas inmediatamente para construir la distinción entre objetos. No obstante, aquellas palabras, cuyo grupo no sea grande, ni cuente con varias características, serán difíciles de definir bajo estos supuestos. En la tabla 3.21 se muestra un posible análisis semántico para definir *banca*<sup>51</sup>.

---

<sup>50</sup> J. Ayto (1983), "On specifying meaning" in *Lexicography: principles and practice*, R.R.K Hartmann (ed.), London: Academic Press, pp. 89-98.

<sup>51</sup> Resumido de Ayto, 1983.

	Para sentarse	Varias personas	Una persona	Fijo	Movible	Tapicería	Sin tapicería	Ambos	Sin respaldo	Con	Interiores	Exteriores	Ambos	Recostar/sentar	Funcional	Para confort	Ambos
Asiento	+																
Silla	+		+		+			+		+			+				+
Banca	+	+		+			+			+		+			+		
Taburete	+		+		+			+	+				+		+		
Sofá	+	+			+	+				+	+					+	
Diván	+	+			+	+				+	+			+		+	

Tabla 3.21 Análisis componencial de “asiento”

Lo primero que se observa es que la única característica constante es *para sentarse*, por lo cual tal constante pueda ser el género próximo del grupo. Ahora bien, retomando el corpus, se realizará una confrontación entre el esquema propuesto por Ayto y las definiciones de *banca*. Al revisar las definiciones, se contempló que una característica puede expresarse de diversas formas sintáctica o semánticamente. Así, *de madera o longer than a chair* funcionan como *sin tapicería* y *para varias personas* respectivamente.

En la siguiente tabla, se observan cinco características mencionadas por Ayto para definir *banca*, elegidas por sintetizar las posibilidades del grupo.

Clave	Asiento	Para algunas personas	Sin tapicería	Exterior	Funcional
1banI	+		+	+	
3banI	+	+			
5banI	+	+		+	+
7banI		+			+
10banI	+	+			+
2banE					+
3banE	+				+
7banE	+			+	
9banE	+			+	+
10banE	+	+		+	

Tabla 3.22 Características semánticas de ‘banca’

Como lo permiten ver las definiciones, los rasgos semánticos no siempre aparecen mencionados en las definiciones del hablante; esto es porque el sujeto considera que con pocas características se da a entender, o porque desconoce las características relevantes, o incluso, porque no sabe cómo presentar la información que conoce.

El análisis propuesto por Ayto es interesante y útil para el lexicógrafo. En varios casos la mera comparación entre objetos del mismo grupo puede mostrar una diferencia “real” o lexicográfica. Sin embargo, el hablante puede preferir otras formas de definición. Además de ello, en el grupo de definiciones relativas a *banca*, lo primero que se observa es que el género próximo no es constante (tabla 3.23), y de hecho en algunos casos incluso se puede eludir.

Clave	Genus
2banI	Seat
3banI	Object
5banI	Object
6banI	Something
7banI	Platform
8banI	Apparatus
2banE	Artículo
3banE	Mueble
5banE	Objeto
6banE	Mueble
7banE	Objeto
8banE	Artículo
9banE	Mueble
10banE	Objeto

Tabla 3.23 Género próximo en definiciones de ‘banca’

### 3.6.3 El estereotipo

El siguiente enfoque considera que el estereotipo puede ser un candidato como difusor del significado. El estereotipo, conjunto de características típicas de un objeto, representa para Hank<sup>52</sup> un aspecto muy importante en la lexicografía y en el análisis de corpus. Considera que los avances en lingüística cognitiva contribuyen a la construcción de una teoría sólida del estereotipo basada en el parecido de familia.

Para Aitchison, los estereotipos –aproximaciones ingenuas al significado y muy probablemente inconscientes– se construyen cultural e individualmente, aunque su contacto con la realidad es parcial. Además de ello, el estereotipo difiere entre niños y adultos, ya que los primeros prestan mayor atención a la apariencia o forma<sup>53</sup>. Agrega que los estereotipos no son adecuados para construir una definición debido a la variabilidad y dificultad de encontrar un estereotipo para diferentes objetos<sup>54</sup>.

Tanto para Putnam<sup>55</sup> como para Lara<sup>56</sup>, el estereotipo no tiene lugar en la teoría de los conceptos y que es imposible estudiar su formación. Piensan que tampoco se puede explicar con total certeza el paso de la singularidad a la generalidad. Putnam afirma que además de imágenes, los conceptos contienen creencias que se formalizan en palabras.

Según Rosch, la creación de estereotipos es independiente de la lengua, refiere a capacidades inherentes de la especie y su función es resaltar los objetos porque así adquieren pertinencia en su comunidad lingüística<sup>57</sup>.

Por su parte, Wiegand<sup>58</sup> intenta encontrar las propiedades que deben ser incluidas en una definición, y considera que el estereotipo es una opción viable para esto. Para encontrar los estereotipos, se basa en una escala de uso, la cual se obtiene estadísticamente. En un experimento a partir de 100 cuestionarios, Wiegand propuso 21 propiedades de *limón* (ácido, oblongo, cítrico, etc) y los informantes juzgaron si la propiedad en cuestión era *buena*, *no muy buena* o *no buena* para describir tal entidad.

---

<sup>52</sup> P. Hank (2000), "Contributions of Lexicography and Corpus Linguistics to a Theory of Language Performance" in *Proceedings of Euralex*, Oxford, p. 9.

<sup>53</sup> Aitchison, pp. 70, 175.

<sup>54</sup> *Ibidem*, p. 67.

<sup>55</sup> H. Putnam (1975), *Mind language and reality* (Philosophical papers, vol.2), Cambridge University Press, p. 250.

<sup>56</sup> Lara, p.180.

<sup>57</sup> Lara, p.182 y Aitchison, p. 55.

<sup>58</sup> H. Wiegand (1984), "On the structure and contents of a general theory of lexicography", en *LEXeter '83 Proceedings*. R.R.K. Hartman (ed), pp. 13-30.



Sin embargo, al comparar el corpus de Sierra<sup>59</sup> con las clasificaciones de Wiegand, se muestra la ineficacia del método, ya que entre ambos datos no se encuentran suficientes paralelismos. Por ello, Sierra sugiere que el estereotipo definido con los criterios de Wiegand no puede fungir como base para una metodología onomasiológica. La tabla 3.24 ilustra algunos ejemplos de la comparación:

Propiedad	Sierra	Wiegand
Yellow/citrus	Buena	No muy buena
Oval, sour pulp	No buena	Buena
Many uses in cooking and drinks, Variable protuberant tip	No buena	No buena
Similar to limes	No buena	---

Tabla 3.24 Análisis comparativo de las propiedades de 'limón'.

La problemática principal de la teoría basada en estereotipos se debe a que no todas las características son compartidas socialmente y el número de características relevantes en las definiciones de los hablantes puede ser reducido e incapaz de distinguir un referente<sup>60</sup>.

No obstante, la información que proporciona el estereotipo debe ser tomada en cuenta, ya que como señala Putnam<sup>61</sup>, representa el primer conocimiento que adquiere un ser humano de un objeto. En una búsqueda onomasiológica es posible que el usuario común recurra al estereotipo para comunicarse.

### 3.6.4 Los elementos diversos

Una postura que intenta englobar varios elementos del significado es la de Putnam<sup>62</sup>, quien propone: marcadores sintácticos (categoría de palabra), semánticos (rasgos semánticos), estereotipo (conjunto de características compartido por una sociedad) y extensión (estructura profunda que hace a la cosa ser lo que es, vinculada con el conocimiento científico).

La propuesta de Putnam contempla la categoría gramatical y el estereotipo; elementos que a pesar de no ser contundentes para determinar el significado, siguen transmitiendo información socialmente compartida. Por otro lado, considera igualmente importantes los rasgos semánticos, características mínimas de información. Sin embargo, la parte fundamental del significado es la

<sup>59</sup> Sierra (1999), capítulo 2.

<sup>60</sup> Putnam (2000), p. 88.

<sup>61</sup> Putnam (1975), pp. 239-240.

<sup>62</sup> Ibidem, 215-271.

extensión, la que se relaciona frecuentemente con el conocimiento científico. La extensión refiere a aquellas características primarias que componen las causas de las propiedades que se manifiestan al usuario común<sup>63</sup>. En la tabla 4.25 se presenta el significado de *agua* según Putnam:

Marcadores Sintácticos	Marcadores Semánticos	Estereotipo	Extensión
mass noun	clase natural	Incolora	H <sub>2</sub> O
Concreto	Líquido	Transparente	con impurezas
		Insípida	
		Refrescante	
		Etc.	

Tabla 3.25 Propiedades para “agua”

Con la finalidad de comparar las definiciones dadas con el significado de *agua* propuesto por Putnam se muestra en la tabla 3.26 las características mencionadas por los estudiantes. Las definiciones del corpus incluyen la propiedad *fluid* que se puede entender como un marcador semántico, así también los puntos de ebullición y congelamiento pueden ser parte de la extensión. De igual manera, se contempló que calificar al objeto *agua* como *vital* para el ser humano integra parte del estereotipo. Los marcadores sintácticos por otro lado, no ocurren comúnmente; sin embargo, en conceptos como *ropa* (sustantivo colectivo) pueden ser de gran utilidad, ya que destacan que el tipo de sustantivo al que el hablante se refiere es de diferente naturaleza<sup>64</sup>.

Marcadores Sintácticos	Marcadores Semánticos	Estereotipo	Extensión
	líquido/ fluido (clase natural)	Incolora	H <sub>2</sub> O
		transparente/ clara	Hierve a 100°C
		Inodora	Congela a 0°C
		Insípida	
		se bebe	
		se encuentra en ríos, lagos...	
		Vital	

Tabla 3.26 Clasificación de Putnam

La propuesta de Putnam resulta muy flexible en tanto que contempla características lexicográficas, semánticas y estereotípicas. No obstante, existe un sinnúmero de conceptos que el hablante no podrá expresar usando más de una categoría. Aunado a eso, concentra su principal

<sup>63</sup> Putnam (1975), p. 230.

<sup>64</sup> véanse las definiciones 2rop1 y 3rop1 en el apéndice.

atención al significado contenido en la extensión, lo cual a pesar de su importancia no puede guiar la búsqueda onomasiológica, debido a que el hablante no posee el conocimiento formal de todas aquellas entidades que puede nombrar.

Lo que se puede retomar de esta propuesta es que el significado se compone de diversos elementos, y un hablante puede hacer uso de cualquiera de ellos no importando su naturaleza.

### 3.6.5 La exhaustividad

Wierzbicka<sup>65</sup> propone que la mayor cantidad posible de características conforma el significado. Acepta características que refieran a la cercanía entre objeto y hablante, las implicaciones culturales y la información científica. Por ejemplo, para animales proporciona cinco tipos de características: Hábitat, tamaño, apariencia, comportamiento y relación con la gente. Así, la descripción para *ardilla* según Wierzbicka se ejemplifica en la tabla 3.27.

Propiedad general	Descripción
Habitat	Viven en lugares donde hay muchos árboles.
Tamaño	No son muy grandes, una persona puede tener una fácilmente en ambas manos.
Apariencia	Tienen una cola peluda grande. Su piel es rojiza o grisácea.
Comportamiento	Recolectan y comen pequeñas cosas que crecen en los árboles de ciertos tipos.
Relación con la gente	La gente las considera como criaturas simpáticas y divertidas.

Tabla 3.27 Definición exhaustiva de 'ardilla'

En primer lugar, la concepción de significado, para Wierzbicka, no puede simplificarse a los dos elementos tradicionalmente aristotélicos (género próximo y diferencia específica), ya que la minuciosidad de factores relacionados con el objeto permite distinguirla e identificarla mejor. Al establecer paralelismos entre las definiciones de *ardilla* de los informantes y la clasificación de exhaustividad, sí se encontraron correspondencias, aunque sintácticamente sus formas sean distintas. En la tabla 3.28 se ilustran dos ejemplos por categoría:

<sup>65</sup> A. Wierzbicka (1985), *Lexicography and conceptual analysis*. Karoma Publishers.

Propiedad general	Clave	Descripción
<b>Habitat</b>		Viven en lugares donde hay muchos árboles...
	2ardI	...living in trees...
	1ardE	...habita en zonas boscosas...
<b>Tamaño</b>		No son muy grandes, una persona puede tener una fácilmente en ambas manos.
	9ardI	A small-sized animal...
	3ardE	Animal pequeño...
<b>Apariencia</b>		Tienen una cola peluda grande. Su piel es rojiza o grisácea.
	4ardI	...grey/red, bushy tail...
	10ardE	cuenta con una cola grande y ancha
<b>Comportamiento</b>		Recolectan y comen pequeñas cosas que crecen en los árboles de ciertos tipos.
	3ardI	...collects nuts...
	6ardE	...se alimenta de nueces y diversas semillas
<b>Relación con la gente</b>		La gente las considera como criaturas simpáticas y divertidas.
	7ardI	A small funny animal...

Tabla 3.28 Exhaustividad en las definiciones libres de “ardilla”

La categoría *Relación con la gente* considera información del corpus que el hablante no puede evitar mencionar, los elementos contextuales.

La propuesta de Wierzbicka incluye una multiplicidad de componentes semánticos según la entidad que se describa; por ello la mayoría de características que propusieron los estudiantes al describir *ardilla* pudo clasificarse dentro de alguna categoría.

Debe agregarse que la exhaustividad implica contemplar siempre características o categorías nuevas, situación que se ilustra en la tabla 3.29:

2ardE	...tiene sus dientes de adelante un poco más grandes que los demás
4ardE	...hay dos muy famosas en Disney
7ardE	...en CU hay muchas
4ardI	...buries nuts
6ardI	... builds nests in trees called dreys

Tabla 3.29 Características nuevas de ‘ardilla’

### 3.6.6 Conclusiones de la aplicación de modelos teóricos al corpus

A través de la revisión de varios enfoques teóricos, se ha evidenciado la complejidad de la consulta onomasiológica y la imposibilidad de adoptar un panorama lexicográfico tradicional. Como lo muestran las diferentes vertientes y el corpus mismo, el género próximo no es una constante necesaria del significado, ya que puede omitirse a través de la mención de diferentes características.

El criterio de Wiegand, destacar propiedades por uso, se enfrenta a la problemática del estereotipo, no puede ser el modelo a seguir, porque cambia a través de las culturas, sociedades, regiones y tiempo. La propuesta del contexto no puede extenderse por igual a todo el vocabulario, sobre todo, cuando el hablante prefiere describir el objeto por su apariencia. La propuesta de Ayto, al momento de ser comparada con las definiciones libres, también resulta incompleta, pues no existe una tendencia sobresaliente en el hablante que contemple características distintivas del objeto. Por otro lado, al contemplar un panorama de elementos diversos, los hablantes no tendieron a mencionar todas las características que Putnam menciona cuando escribieron sus definiciones libres, lo anterior no impidió que en la segunda fase el informante no reconociera la referencia. La propuesta de Wierzbicka es atractiva, sin embargo cabe aún la posibilidad de que, a pesar de poseer un listado detallado de características, el sistema se confunda por la excesividad de datos o bien, ninguno de ellos sea suficiente para distinguir un objeto.

### **3.7 Conclusiones**

Luego de hacer una revisión tanto teórica como comparativa, los resultados de este trabajo muestran que el diseño de un diccionario onomasiológico debe contemplar la multiplicidad de propiedades y todos los tipos de definición posibles.

Los modelos teóricos tienen una limitante en la onomasiología, no toman en cuenta la intención del hablante, por ello resulta inconveniente seguir alguno de ellos.

Se deben tomar en cuenta las siguientes observaciones:

- La conceptualización tomada como base debe ser construida de una forma flexible que no limite las posibilidades del usuario.
- El diccionario debe contemplar que el usuario introduzca definiciones simples pero también complejas con información irrelevante que deberá remover.
- El diccionario deberá contemplar que una misma característica pueda presentarse de distintas formas, debe interpretarla y compararla con su información.

# CAPÍTULO 4

## EVALUACIÓN DE LOS DICCIONARIOS ELECTRÓNICOS



En el capítulo tres se habló de los factores involucrados con la conceptualización, de las formas de definir y del tipo de información contenida en las definiciones libres.

En este capítulo se evaluarán cuatro diccionarios electrónicos con la finalidad de conocer su eficacia. Los cuatro fueron descritos en el capítulo dos señalando sus características principales. En este apartado se concentrará la atención en la funcionalidad de éstos.

Se señalarán además los inconvenientes de cada diccionario (4.1). Posteriormente se compararán los resultados obtenidos a través de tres índices: *precision*, *fallout* y *generality* (4.2).

En este capítulo se evaluarán cuatro diccionarios con la finalidad de conocer su eficacia: *Onelook*, *Wordsmyth*, *Longman* y *Vox*. Los cuatro fueron descritos en el capítulo dos señalando sus características principales. En este apartado se concentrará la atención en su funcionalidad. Asimismo, se señalarán sus deficiencias y ventajas. Posteriormente se compararán los resultados a través de tres criterios: *precision*, *fallout* y *generality*, los cuales determinaron la construcción de las tablas de este capítulo. En el apéndice se podrán consultar las consultas completas por cada diccionario, en donde se muestra la información arrojada y los resultados de los índices aplicados. Se utilizarán estas tablas para conocer qué diccionario obtuvo los mejores resultados.

La evaluación consiste en dos pasos, en el primero de ellos se hizo una evaluación apreciativa (4.1), basada en criterios intuitivos. En el segundo se realizó un análisis formal (4.2), tomando en cuenta los resultados de cada diccionario.

## 4.1 La evaluación apreciativa

Dentro de esta evaluación se explican algunas complicaciones de la consulta, así como la forma en la que el diccionario funciona. Los distintos diccionarios poseyeron tanto ventajas como inconvenientes. En los siguientes apartados se explicarán las limitaciones a partir de problemas de: lematización (4.1.1), formas sinónimas (4.1.2), formas relacionadas semánticamente (4.1.3), información en los contenidos del diccionario (4.1.4), errores de correspondencia y signos de puntuación (4.1.5), presentación de resultados (4.1.6) y diseño (4.1.7).

### 4.1.1 Problemas de lematización

La **lematización** puede brevemente definirse como el proceso en el que se elimina la información flexiva o derivativa de una palabra para conservar de esta manera la raíz. Este proceso puede realizarse de dos formas: automática o manualmente. En la primera, es el diccionario quien hace el proceso, según un algoritmo interno. En la segunda, es el usuario el que decide dónde hacer el corte a través de comodines al final de la raíz introducida.

La lematización sirve para expandir morfológicamente las búsquedas del usuario, lo cual puede mejorar los resultados. Por ejemplo, si se lematiza la palabra *medir*, la búsqueda incluirá los resultados que encuentren no sólo esta palabra, sino que además incluirán aquellos que contemplen *mediciones*, *medida*, *medible*, etc.

En los diccionarios consultados el proceso se realizó automáticamente por *Onelook*, y de forma parcial por *Longman* (unía los plurales con los singulares y las formas verbales con infinitivo). Sin embargo, en el *Vox* y el *Wordsmyth* sólo se contempla lematización manual. Por ejemplo, si se deseaba lematizar *instrumento*, un corte posible podría haberse efectuado en *instrument\**, lo cual uniría a *instrumento* palabras como *instrumental*, *instrumentar*, *instrumentación*, *instrumentista*.

Este proceso es básico para una buena consulta onomasiológica, ya que en los diccionarios que no contemplan el proceso automáticamente, se presentan casos en los que no se encuentra una palabra, sólo porque la palabra clave difiere de la palabra propuesta en la definición del diccionario hasta por un cambio mínimo en el género o número; por ejemplo, al buscar *ardilla* en el *Vox*, se utilizó *peludo*, pero la palabra en la definición era *peluda*. En el diccionario *Longman*, al buscar *water*, se utilizó *drinkable*, pero la palabra en la definición era *drink*. Al no corresponder con las formas del diccionario, en ambos casos las características proporcionadas arrojaron información irrelevante.

#### 4.1.2 Problemas de formas sinónimas

Ocurre con frecuencia que el usuario emplea diversos sinónimos o formas complejas en lugar de una sola forma. A pesar de que semánticamente tales formas refieren a lo mismo, el diccionario no las puede comparar con sus propias definiciones, porque gráficamente son distintas.

Para que pueda existir una posible correspondencia entre la información del diccionario y la información del usuario, se debería contemplar una comparación de formas a nivel semántico, con lo cual se expandiría la búsqueda. Es decir, se debería permitir que, si la palabra clave es *pequeño*, el usuario no sólo pudiera acceder al término buscado a través de ésta, sino también a través de formas como: *reducido*, *corto*, *chico*, *que es chiquito*, etc.

En las consultas, se documentan varios casos en los que la información se ve frenada por esta limitación: la diferencia de formas. Las palabras empleadas de una definición libre se vuelven irrelevantes, si no se parecen a las contenidas en la definición del diccionario. Por ejemplo, para buscar la palabra *euthanasia* en el diccionario *Wordsmyth*, la palabra que se toma como clave es *killling*; aunque hubo estudiantes que contemplaron tal palabra, también hubo casos en los que se emplearon otras sinónimas o eufemísticamente semejantes, que les parecieron a los estudiantes



igualmente adecuadas como *to die* (3eutI), *to turn off* (4eutI), *ending* (8eutI). Un ejemplo extremo ocurrió en la consulta de *ardilla* en el diccionario *Vox*, donde siete definiciones poseían palabras claves sinónimas a la palabra contenida en la definición (*arborícola*), éstas fuer: *habita en zonas boscosas* (1ardE), *vive en los árboles* (3ardE), *vive en los bosques* (4ardE), *trepa a los árboles* (5ardE), *en bosque o arboleda [...] vive en los árboles* (7ardE), *habita en los árboles* (8ardE), *habita en los bosques* (9ardE); sin embargo, al ser formas sinónimas, no se tomaron en cuenta como palabras claves. El diccionario *Onelook* fue el único que permitió utilizar formas sinónimas para obtener buenos resultados, aunque nunca se mostró su funcionamiento.

#### 4.1.3 Problemas de formas relacionadas semánticamente

Cuando el usuario no sabe definir con certeza una palabra, a veces prefiere mencionar relaciones de la palabra con otras que de alguna forma también la explican, o que dejan ver su funcionalidad, o que sostienen algún parecido por ser del mismo campo semántico<sup>64</sup>. Sin embargo, al hacerlo, el sistema que lee la información puede no entender la intención del usuario. Tales relaciones podrían expresar no sólo sinonimia sino hiponimia, hiperonimia, holonimia, etc.

Para rescatar e interpretar este tipo de información, sería conveniente que el diccionario onomasiológico contara con alguna red semántica que permitiera establecer relaciones en estos niveles también. De esta forma, si el usuario para definir la palabra *sangre* introduce palabras como *plaquetas*, *corazón*, *oxigenación*, u otras, muy posiblemente el diccionario le traiga *sistema circulatorio*, y no la esperada.

En el corpus este tipo de casos sucedió al definir *limón*. Los estudiantes mencionan *para hacer agua dulce* (1limE, 4limE y 10limE), el diccionario *Vox* no entiende esta relación y por tanto no evoca *limón*, pero, caso curioso, extrae *limonada*. En el diccionario *onelook*, al buscar *barometer*, el estudiante emplea una definición por analogía, *like thermometer* (3limI), y a pesar de que se encuentra la palabra buscada, extrae en primer sitio *weather situation*. Un caso similar ocurre también en el *Wordsmyth* cuando se busca *bench* a través de *a type of chair*, el diccionario encuentra sólo una palabra, *recliner*.

---

<sup>64</sup> Véase el capítulo tres, cuando se describen los diferentes tipos de definición.

#### 4.1.4 Problema del contenido de los diccionarios

Los diccionarios consultados para la evaluación comparan la información entrante con su propia definición o alguna parte de ésta, por ello la vastedad de la definición en el diccionario será una clave fundamental para que el usuario encuentre con más probabilidad las palabras que busca. Los recursos de los diccionarios varían dependiendo del formato, la mayoría no sólo cuenta con información de la definición lexicográfica, sino también incluyen en algunos casos con sinónimos, ejemplos, contextos de uso o relaciones semánticas con otras palabras del diccionario.

Durante la consulta, el diccionario que más problemas tuvo en este sentido fue el *Vox*, porque muchas veces sus definiciones fueron muy breves. Obsérvese el caso de *limón*, que se presenta como *fruto del limonero* más dos acepciones figurativas. Por el contrario, en el *Longman* aparece como *a fruit with a hard yellow skin and sour juice*, e incluye más definiciones, contextos de uso, ejemplos y palabras relacionadas. En el *Wordsmyth lemon* aparece como *a small oval citrus fruit with pungent yellow skin and sour juice*, además cuenta con otra entrada, varias definiciones en la primera entrada, información lexicográfica y palabras relacionadas. El *onelook* lo define como *yellow oval fruit with juicy acidic flesh*, presenta varias definiciones breves y 40 hipervínculos con definiciones de otros diccionarios.

Mientras más grande sea la definición del diccionario, es más probable que la información de un usuario quede comprendida en ésta, y así éste pueda obtener éxito en su búsqueda.

#### 4.1.5 Problema de errores de correspondencia y signos de puntuación

Cuando a un sistema se le introduce un símbolo o dato ajenos a los contemplados, la intromisión puede producir errores en el sistema. Los datos desconocidos, en este caso, pueden ser signos de puntuación como la coma, el punto, punto y coma, paréntesis, guión largo, diagonales, puntos suspensivos, entre otros. Otro tipo de dato ajeno lo conforman los errores de correspondencia, los cuales se pueden deber a tres causas: porque no se escribe una palabra ortográficamente bien, porque se cometió un error al escribir la información, o porque la palabra introducida es desconocida por el sistema.

Dado que es una posibilidad que los usuarios introduzcan enumeraciones, varias ideas o presenten palabras compuestas, entonces podría suceder que los mismos separaran su información a

través de comas, puntos, etcétera. También sería posible que cometieran faltas de ortografía o tuvieran errores de dedo al escribir.

Véase por ejemplo la definición 2aguE, *líquido incoloro, inoloro e insípido*, en la que se ejemplifica el uso de la coma y un error ortográfico, *inoloro*. Otro ejemplo sería la definición 5hurI, *a very strong wind, very destructive, often named after people's first names, i.e. the weather researchers who first discover each one*, que cuenta con comas, apóstrofes y puntos. Si alguno de estos casos llegara a pasar, el diccionario debe procurar no frenar la búsqueda:

En el diccionario *Onelook*, la intromisión de diferentes signos de puntuación (punto, apóstrofe, coma, diagonal, paréntesis, guión corto) no causó problemas, asimismo no tuvo problemas con los errores ortográficos, porque cuando éstos venían acompañados de otra palabra, el diccionario realizaba la búsqueda con los términos que estuvieran bien escritos. *Wordsmyth* pide al usuario separar la información por comas o comillas, signos que acepta al igual que a las diagonales, pero tiene problemas con el punto y coma, el punto, los dos puntos y el apóstrofo; además, no acepta formas ortográficamente erróneas. El diccionario *Vox* tiene problemas con los dos puntos, el punto, pero sí acepta la diagonal, la coma y los paréntesis; respecto a la ortografía, no respeta subíndices y no contempla los errores ortográficos como la falta de acentos. El diccionario *Longman* acepta comas, punto, punto y coma, dos puntos, paréntesis, apóstrofo y diagonales pero no acepta comillas dobles; no acepta formas escritas erróneamente.

#### 4.1.6 Problema de presentación de los resultados

Cuando ha terminado el proceso de consulta, el usuario espera encontrar de manera rápida la información que busca, por ello le conviene que ésta aparezca ordenada con base en sus necesidades y no le sirve tanto el hecho de que se distribuya la información alfabéticamente. Para que aparezca por similitud a la información introducida, es necesario que el sistema ordene la información según la semejanza entre las palabras extraídas y las palabras claves.

Si durante el proceso se ha dado una buena definición y la información está ordenada por relevancia, el usuario sólo tendrá que buscar en los primeros lugares la información. No obstante, si está en orden alfabético, este orden sólo será benéfico para el usuario si su palabra comienza con las primeras letras del alfabeto, o si éste sabe cómo empieza la palabra, porque así podrá reconocerla.

En el diccionario *Onelook* se ordena la información por relevancia. En los tres restantes es por orden alfabético; sin embargo, la localización del término en el *Vox* se realiza de forma rápida al

revisar el listado a través de una barra de desplazamiento, pero en el *Longman* y el *Wordsmyth* la localización es lenta debido a que la información está dividida en ventanas.

#### **4.1.7 Problema del diseño**

Uno de los factores que determinan el uso de los diccionarios electrónicos es el diseño de la interfaz. Si ésta es amistosa y cómoda para el usuario, éste se sentirá motivado para seguirlo utilizando. Sin embargo, si el diseño resulta confuso, el usuario sentirá poco interés por volver a consultarlo.

Un buen diseño evita que el usuario se pierda durante la consulta, ofrece instrucciones claras para la búsqueda y permite que, a través de pocos pasos, el usuario obtenga un resultado satisfactorio.

En la consulta se observó que la interfaz de cada diccionario tenía diferentes características y contaba con diversos recursos de consulta y multimedia. Las frecuentes consultas en cada uno de ellos, permitió concluir que un ejemplo de funcionalidad lo proporciona el diccionario *Vox*, pues la consulta se realiza en una sola ventana, lo cual permite agilizar el proceso y reintentar rápidamente nuevas consultas si la primera no fue exitosa. El *Longman* tiene un diseño lento y confuso, cuando se desea repetir la búsqueda se obliga al usuario a partir casi desde el principio. El *Onelook* tiene un diseño sencillo, incluye sólo la información necesaria, pero al localizar el resultado, la interfaz no resulta eficiente. El diccionario *Wordsmyth* tiene un diseño sencillo, pero también presenta la información en ventanas, y si uno seleccionó ver varios campos, entonces la localización será lenta porque se mostrarán menos términos por cada ventana.

#### **4.1.8 Problema de consulta**

El objetivo de la evaluación apreciativa es conocer el estado en el que se encuentra la consulta onomasiológica y a su vez delimitar aquellas características que debería tener un buen sistema.

Diccionario	Onelook	Longman	Wordsmyth	Vox
Lematización	Automática	Automática (parcial), manual	Manual	Manual
Formas Sinónimas	Sí (nivel desconocido)	No	No	No
Formas relacionadas	Sí (nivel desconocido)	No	No	No
Contenido del diccionario	definiciones y varios hipervínculos a otros diccionarios en Internet	Varias definiciones, acepciones, relaciones entre palabras, contexto de uso	Varias definiciones, información, lexicográfica, ejemplos de uso	Definiciones (con pocas palabras clave) y acepciones
Errores ortográficos	Contemplados	No contemplados	No contemplados	No contemplados
Signos aceptados	Contemplados	Punto, apóstrofo, coma, diagonal, paréntesis, guión corto	Comas, comillas, diagonales	Diagonal, paréntesis, comas
Presentación de la información	Por relevancia	Por orden alfabético	Por orden alfabético	Por orden alfabético
Diseño	Sencillo	Lento, confuso	sencillo	Funcional

Tabla 4.1 Diferencias cualitativas de los diccionarios cualitativamente

Las diferentes ventajas de cada diccionario pueden apreciarse en la tabla 4.1, en la que se observa qué ofrece cada uno de ellos.

## 4.2 Evaluación formal

En este apartado se realizará una evaluación basada en tres índices: *precision*, *generality* y *fallout*. Esta evaluación permite por un lado, homogeneizar los resultados de cada diccionario para posteriormente establecer comparaciones entre todos. En seguida se explica la metodología.

### 4.2.1 Introducción de la información en el diccionario

Para todos los diccionarios fue necesario hacer varias pruebas previas con la finalidad de aprovechar mejor los recursos particulares de cada uno. A grandes rasgos, la primera división debe hacerse entre el diccionario *Onelook*, que aceptó sin problemas el lenguaje natural así como símbolos ortográficos de distinta índole; y los diccionarios *Longman*, *Wordsmyth* y *Vox*, los cuales forzaron a hacer una preselección de información, porque no siempre aceptaban todas las palabras de las definiciones libres, ni los signos de puntuación que venían en ellas. Con el fin de introducir sólo aquellas palabras que produjeran resultados apropiados, se buscaron las palabras claves de las definiciones libres. Así, se introdujo en el primer diccionario la definición completa del corpus, pero en los segundos se introdujeron sólo las palabras claves en combinación con filtros u operadores.

Cuando se realizaron las consultas, se anotaba: el número de palabras extraídas, el lugar donde fue hallado el término buscado y los dos primeros términos extraídos<sup>65</sup>. Asimismo, en las tablas se remarcó aquella información relevante que fue introducida en la consulta. En la tabla 4.2 se muestran un ejemplo por cada diccionario.

Diccionario	Clave	Definición libre del concepto 'agua'	Forma introducida	Fase 2	1º Lugar	2º Lugar	Total	L. Ref
OneLook	lagul	It's a clear liquid that you get from a tap		Water	escape	water	1000	2
Longman	lagul	<b>It's a clear liquid</b> that you get from a tap	It's a clear liquid	Water	acetone	ammonia	15	14
Wordsmyth	lagul	It's a clear <b>liquid</b> that you get from a tap	a liquid	Water	acetal	Acetic acid	180	175
Vox	2aguE	<b>liquido incoloro, inoloro e insípido</b>	LIQUIDO .Y INCOLORO .Y INSÍPIDO		agua	-	1	1

Tabla 4.2 Introducción de definiciones en los diccionarios evaluados

Para todos los casos se buscó alcanzar los mejores resultados en la obtención de términos. Para ello, se utilizaron distintas combinaciones de las palabras de la definición libre para identificar aquella que permitiera el mejor resultado. De esta manera, como se observa de la tabla 4.2, de la definición de lagul el mejor resultado en el Wordsmyth para llevar al término fue *a liquid*, incluso mejor que solo *liquid*. Esto es entonces lo que se reporta en la cuarta columna, forma introducida. En el diccionario *Onelook* no se depuró información y por eso la cuarta columna de la tabla está vacía. En los diccionarios *Longman* y *Wordsmyth* se depuró la información, por ello aparece tanto la definición libre como la forma que finalmente se introdujo. En el diccionario *Vox* también existió una depuración, pero en este caso la forma introducida necesitaba a la vez aplicar los operadores lógicos, y como lo muestra el ejemplo automáticamente recibe la forma en mayúsculas. Hay que observar la relación entre la columna que muestra la definición libre y la que ilustra la forma introducida porque permite ver que en algunos casos la información relevante de las definiciones fue mínima.

#### 4.2.2 Descripción de la evaluación

Una vez completados los resultados, se evaluaron a los diccionarios bajo criterios rigurosos dentro de la lingüística computacional. La finalidad de esta evaluación es comparar los diferentes

<sup>65</sup> Las tablas completas pueden consultarse en el apéndice del capítulo cuarto.

diccionarios no importando su formato o formas de consulta, para ello se ha hecho uso de tres índices empleados dentro de esta área: *precision*, *fallout* y *generality*.

#### 4.2.2.1 Precision

Esta medida señala si lo que se buscó fue encontrado o no; por ello, el resultado ideal es 1 cuando se ha encontrado el término y el peor valor es 0 cuando no se encontró.

Por otro lado, debido a que *Fallout* emplea el concepto “subconjunto” o “subgrupo” conviene definirlo a continuación. Por subconjunto se entiende el listado de términos desde el primer término que aparece hasta el término buscado. Si el listado es de 1000 palabras y el término buscado aparece en el lugar 25 entonces, el subconjunto comprende las palabras desde el primer puesto al 25.

#### 4.2.2.2 Fallout

*Fallout* indica el número de términos no relevantes antes de topar con lo buscado: el mejor valor es 0 y el peor es 1. Es decir, si el primer término leído es el buscado, entonces se tiene un valor de cero, pero si es el último del total de términos obtenidos, entonces el valor es uno. Cabe señalar que aquellas definiciones que no evocaron ningún término poseen un valor de uno. Su fórmula es:

$$f = \frac{R-1}{T-1}$$

En donde  $f$  es *fallout*,  $R$  el *ratio* o bien subconjunto,  $T$  es el total de términos evocados. Cabe señalar que sólo en el diccionario *Onelook* este límite es constante, siempre extrae 1000 términos.

#### 4.2.2.3 Generality

Finalmente *Generality* señala cuál es la relevancia del vocablo buscado en el grupo de términos implicados, el valor ideal es 1 cuando el total de términos sólo posee un miembro y éste es el término buscado y el peor valor es cero, el cual ocurre cuando no aparece el término en el grupo de elementos evocados. Su fórmula es:

$$G = \frac{1}{T}$$

Donde  $T$  es el total de términos extraídos.

### 4.2.3 Construcción de las tablas para la evaluación

Una vez terminada la consulta de cada diccionario, se completaron las tablas que se muestran en el apéndice de este capítulo. En las tablas de los diccionarios se agregaron tres columnas: *precision*, *fallout* y *generality*, los cuales son índices que permiten establecer juicios formales y comparaciones entre los diccionarios..

En la tabla 4.3 se muestran dos ejemplos de los resultados para *Onelook* y *Longman*,:

Di c.	Clave	Definición Libre	Forma Introducida	Fase 2	1° Lugar	2° Lugar	Total	L. Ref	F	P	G
O L	9barl	Measures pressure		-	Anaeroid barometer	anaeroid	1000	3	0.002	0.001	0.001
L M	10ropl	Things you wear	Things wear	Clothes...	Best	box	38	8	0.18919	1	0.026

Tabla 4.3 Descripción de las tablas

### 4.2.4 Evaluación formal entre los diccionarios

El presente apartado tiene como finalidad mostrar, según los índices, los resultados en los cuatro diccionarios evaluados para saber cuáles de ellos pueden servir como base del sistema final en este trabajo.

En cuanto a *precision*, el diccionario *Vox* obtuvo malos resultados: ocho resultados fueron totalmente negativos, de los cuales *monopolio* (0.1) y *huracán* (0.1) fueron los más bajos. El diccionario *Longman* obtuvo buenos resultados, con excepción de *bench* (0.3) y *monopolio* (0.4). Por otro lado, los resultados de *Wordsmyth* fueron muy dispersos, mientras *water* (1) *lemon* (1) y *barómetro* (1) consiguieron evocar en todas sus definiciones la palabra respectiva, hubo otros como  *euthanasia* (0.3), *bucket* (0.22) y *hurricane* (0.4) que poseen resultados negativos. Finalmente, el diccionario *Onelook* extrajo resultados positivos en la mayoría de los términos; aunque para *euthanasia* (0.4) y *monopoly* (0.6) no fueron satisfactorios. A grandes rasgos, el diccionario está cumpliendo con su función, la cual es asociar definiciones a términos concretos, sin embargo la basura que genera también es considerable, lo cual se puede contrarrestar con el ordenamiento por similitud. Los resultados pueden observarse en la figura 4.4.



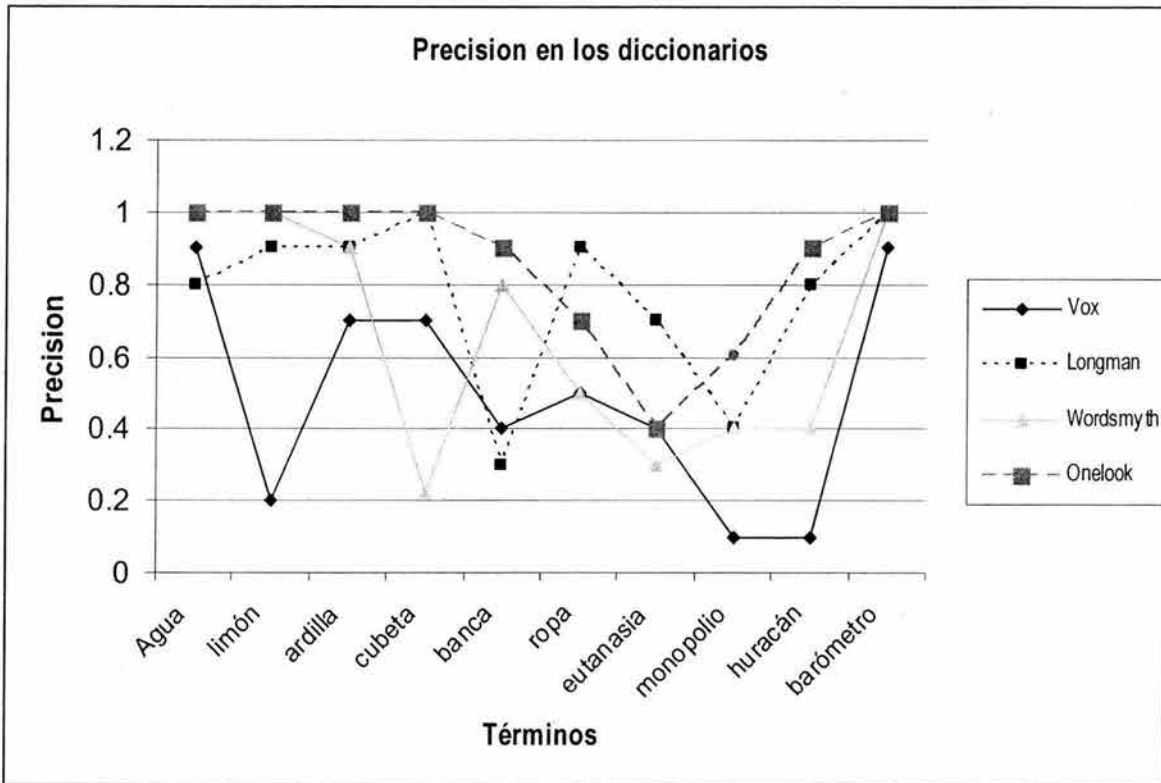


Figura 4.4 Precision en los diccionarios

El siguiente índice tomado en cuenta es *fallout*. La mayor eficiencia la tuvo *Onelook* y la peor la obtuvo el diccionario *Vox*. Este último obtuvo su mejor resultado con *barómetro* (0.17786), de ahí en adelante los resultados se incrementan notablemente. En la mayoría de los términos buscados más de la mitad de lo leído no es de relevancia para el diccionario. La tendencia es mostrar un 69% de términos no relacionados antes de leer lo buscado.

El diccionario *Longman* obtuvo un resultado perfecto en *barometer* (0), le sigue *bucket* (0.06541) y de ahí en adelante los resultados se alejaron. Destacó el resultado de *water* (0.96947) porque evidenció que el diccionario y el usuario no coincidieron en mencionar la misma información, pues el sistema menciona primero otros términos antes de evocar el concepto *agua*. Al hacer el promedio de todos los términos, se puede ver que la tendencia es mostrar un 47% de términos no relacionados antes de leer lo buscado.

El diccionario *Wordsmyth* obtuvo su mejor resultado en *barometer* (0.04644), de ahí en adelante los resultados se incrementaron. En este diccionario, la tendencia fue mostrar un 58% de términos no relacionados antes de leer lo buscado.

El diccionario *Onelook* consiguió buenos resultados en cuanto a las palabras *lemon* (0.001401) y *water* (0.005205). Los peores resultados se encontraron en *euthanasia* (0.617717) y *monopoly* (0.513012). Al hacer el promedio de todos los términos, se obtuvo que la tendencia fue mostrar un 21.99% de basura o términos no relacionados antes de leer lo buscado. La dificultad mayor del diccionario se encuentra en términos que carecen de datos sensoriales y que tienden a la abstracción. La figura 4.5 ilustra el contraste del índice *fallout* entre los diferentes diccionarios.

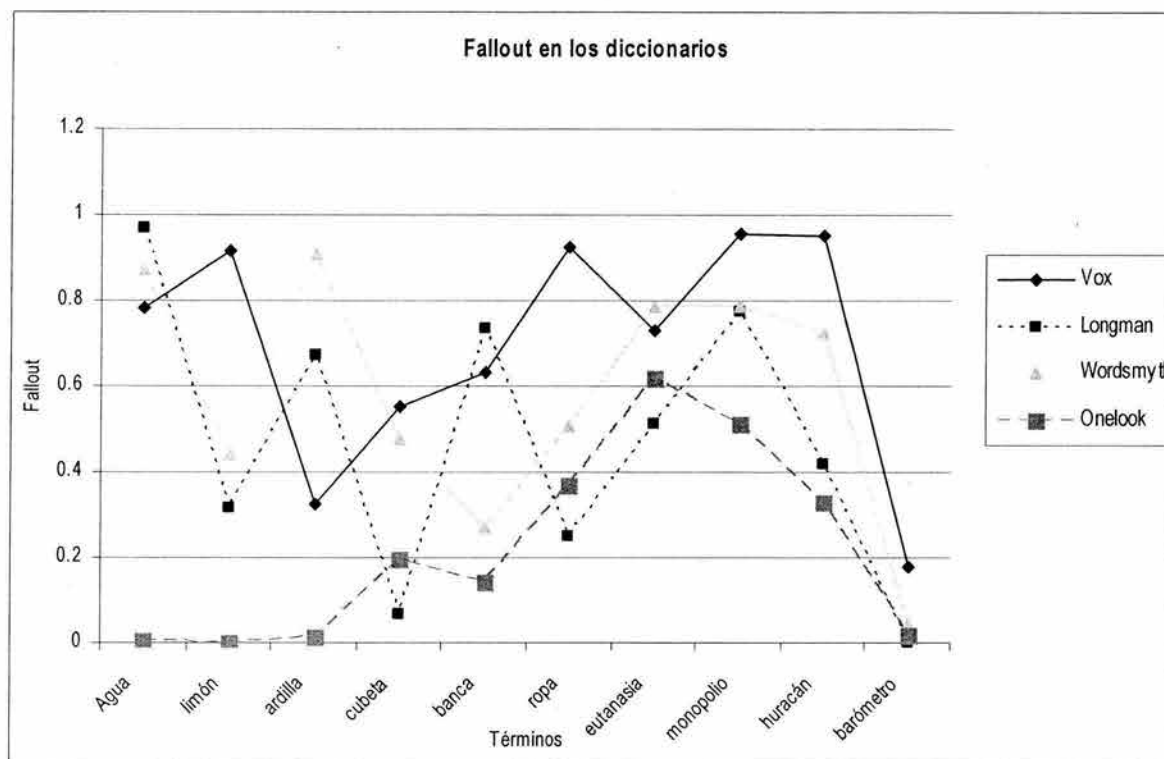


Figura 4.5 Fallout en los diccionarios

En cuanto a *generality*, el diccionario *Longman* consiguió el primer sitio. Cabe señalar que en este punto los cuatro diccionarios tuvieron resultados bajos debido a que el número de términos extraídos fue siempre un problema. Todos los diccionarios no se extraen uno o dos términos por consulta, sino que extraen todo un conjunto amplio de posibilidades, lo cual por un lado aumenta la probabilidad de que se encuentre lo buscado, pero también aumenta el volumen de información donde se debe buscar el término. Nuevamente el diccionario *Vox* no arrojó buenos resultados.

En el *Vox* el mejor resultado lo proporcionaron los conceptos *barómetro* (0.20459) y *agua* (0.30959) el peor lo obtuvo *huracán* (0.000526). En esta última consulta sólo existió una palabra clave.

El diccionario *Longman* obtuvo un resultado satisfactorio. Por ejemplo, un buen resultado fue *barometer* (0.8), no así *euthanasia* (0.00999). Lo cual quiere decir que en este diccionario fue posible tener resultados buenos siempre y cuando la información del usuario coincidiera con la del diccionario.

En el diccionario *Wordsmyth* el mejor resultado lo proporcionó el concepto *bucket* (0.22030), los resultados siguientes no son tan positivos, el peor de ellos lo obtiene *hurricane* (0.03929).

Finalmente, en el diccionario *Onelook* se muestran resultados muy similares dado que el sistema siempre extrajo mil términos. Los peores resultados estuvieron en *bench* (0.0009) y *hurricane* (0.0009). En la figura 4.6 se muestran los resultados de los cuatro diccionarios tomando en cuenta el índice *generality*.

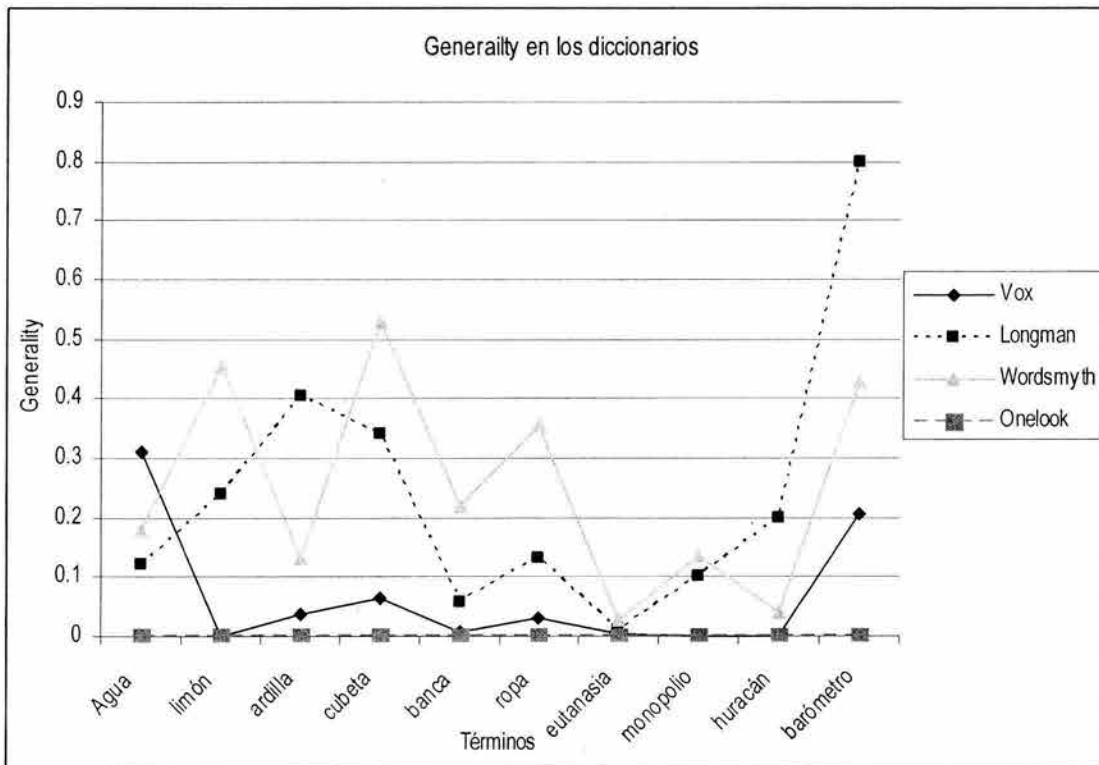


Figura 4.6 Comparación por Generality

#### 4.2.5 Conclusiones de la comparación entre diccionarios

En términos generales, los resultados en los diccionarios electrónicos que funcionan a través de palabras claves no fueron satisfactorios, a pesar de que se realizó una etapa de selección de información relevante. Por otro lado, los resultados del diccionario que funciona a través de lenguaje natural, *Onelook*, corroboran la hipótesis de que la consulta onomasiológica con definiciones libres es posible, y más aún, que puede ser efectiva. Sin embargo, existen aún deficiencias sobre las que se debe trabajar.

El problema de la creación de un diccionario onomasiológico para lengua española no se debe al sistema lingüístico mismo, ya que como se vio, existen actualmente herramientas que ayudan a filtrar la información. El problema es la escasez de trabajos que satisfaga este tipo de búsqueda.

#### 4.2.6 Conclusiones

En algunos casos las personas que quieran consultar un diccionario onomasiológico, poseerán suficientes conocimientos y palabras para definir un objeto; sin embargo, hay factores que complican el proceso de conceptualización, tales como la cercanía entre sujeto y objeto, las implicaciones sociales y culturales de éste, la complejidad del objeto en sí y la propia capacidad del hablante para definir la palabra que han olvidado. Las consultas hechas en los diccionarios electrónicos han evidenciado que el proceso onomasiológico no está del todo perfeccionado.

A pesar de que los diccionarios cuentan con diferentes herramientas y filtros de ayuda, no hay garantía de que se llegue al vocablo deseado y mucho menos de que se llegue a éste de forma rápida.

Por otro lado, las definiciones libres no siempre remiten al mismo tipo de definición e información. En algunos casos el usuario emplea en sus definiciones creencias, estereotipos o contexto. Esta información resulta confusa para un sistema, y habría que contemplarla en un futuro.

Las diferentes definiciones del corpus tratado son sólo algunas de las posibilidades que pueden servir como definición libre, ésta es la razón fundamental para que el sistema no concentre su búsqueda en una sola forma escrita.

Por otro lado, el listado de problemas en los diccionarios evaluados permite establecer bases para la construcción del prototipo para este trabajo: lematización de información, expansión

semántica y expansión morfológica. Lo anterior mejoraría notablemente los resultados de un diccionario onomasiológico.

## CAPÍTULO 5

# DISEÑO DE BÚSQUEDAS ONOMASIOLOGICAS EN LENGUAJE NATURAL



En el capítulo anterior se evaluaron algunos de los diccionarios onomasiológicos en formato electrónico. A partir de este punto se comienza a construir el modelo aquí propuesto, describiendo el procedimiento de búsquedas onomasiológicas en lenguaje natural.

Como punto de partida, se verá la descripción del concepto en lenguaje natural dada por el usuario (5.1). Con esta información, el sistema de cómputo identifica las palabras claves (5.2), discriminando las palabras cerradas (5.3), y expande las palabras claves morfológica y semánticamente (5.4), para realizar la búsqueda de los términos en el banco de datos (5.5). En el capítulo siete se analizaron los requerimientos básicos para llevar a cabo la implantación del sistema, en este capítulo se verá cómo se lleva a cabo dicha implantación.

Como ya se señaló en el capítulo dos, un diccionario onomasiológico servirá para que un usuario encuentre una palabra o un término; para ello es necesario introducir la descripción del concepto, o bien la definición libre de éste. Desde un punto de vista computacional, tal información se denomina **query**.

El diccionario que acepta lenguaje natural permite que el usuario pueda introducir una serie de palabras y será el sistema que evalúe, a través de utilizar filtros, la información. Con esto será posible que el sistema encuentre lo que el usuario desea.

El diccionario onomasiológico permitirá que las personas con conocimientos especializados o no, lo consulten sin problemas, es decir, cualquier persona que haga uso del diccionario podrá buscar un término partiendo de lo que él entienda por su concepto.

Como no es posible trabajar con todas las palabras del idioma español, para la consulta del diccionario únicamente se utilizara una base de datos o banco terminológico que incluya la terminología de dos áreas determinadas, Física y Desastres. Entiéndase por **terminología** al conjunto de palabras relacionadas con un área específica del conocimiento humano. Por ejemplo, Física es el área terminológica, en tanto *aceleración angular*, *cuerpo*, *energía cinética*, *fuerza* y *movimiento lineal* son algunos de los términos especializados de esa área.

Al englobar sólo áreas específicas del conocimiento es posible tener un mejor dominio sobre la efectividad del diccionario. Por ello, no se abarcaron todas las palabras del idioma español.

A continuación se describe, a mayor detalle, el procedimiento de búsquedas onomasiológicas en lenguaje natural. Como punto de partida, se verá la descripción del concepto en lenguaje natural dada por el usuario (5.1). Con esta información, el sistema de cómputo identifica las palabras claves (5.2), discriminando las palabras cerradas (5.3), y expande las palabras claves morfológica y semánticamente (5.4), para realizar la búsqueda de los términos en el banco de datos (5.5).

### **5.1 Descripción del concepto en lenguaje natural**

Para que el usuario pueda introducir una frase, debe tener el concepto en mente, tener claro qué es lo que quiere expresar y qué es a lo que se quiere referir al momento de introducir la definición.

El usuario necesita describir el concepto. Teniéndolo en su mente al momento de escribir la definición libre, no tendrá límites: pueden existir palabras sueltas, puede ser una definición bien

redactada o una corta, puede haber signos de puntuación, etc<sup>66</sup>. A todas estas palabras que introduzca el usuario se le llama **input** desde un punto de vista computacional, y desde el punto de vista lingüístico se llama **definición libre**.

## 5.2 Identificación de palabras claves

Debido a que el usuario puede introducir el concepto de un término como él lo entienda y de manera libre, sin restricciones, la búsqueda estará compuesta por dos tipos de palabras: **palabras claves** y **palabras cerradas**. Las primeras, siendo necesarias, son las que aportan significado y serán las que servirán para los propósitos de recuperación de información. Por ejemplo, si el usuario busca el término *agua* introducirá:

*Líquido vital transparente que sirve para vivir*

Las palabras marcadas en negro son las palabras claves, las que aportan significado; las demás son las cerradas, las que únicamente sirven para conectar una oración con otra.

Para la identificación de las palabras claves, el sistema toma la definición introducida por el usuario, es decir, el input. De aquí, el sistema analiza las palabras, identificando aquellas que son significativas de las que no lo son.

Aquellas palabras que no sean significativas son las palabras cerradas y se desechan; las restantes son llamadas palabras claves. Por tanto, las palabras claves serán identificadas como aquellas palabras que no sean cerradas, es decir, que no estén en la lista de palabras cerradas.

## 5.3 Eliminación de palabras cerradas

El sistema evalúa las palabras que el usuario introduce como input y discrimina las palabras cerradas. El sistema trabaja más rápido al desechar las palabras cerradas, ya que estas ocurren un 40% de las palabras que aparecen en un texto<sup>67</sup>.

---

<sup>66</sup> Para ahondar en el tema consúltense los capítulos dos y tres de este trabajo.

<sup>67</sup> N. Ziviani 1999, Capítulo 7 "Text operations" en *Modern Information Retrieval*, Edimburgo, ACM Press, Addison-Wesley, p. 167.



### 5.3.1 Definición

Las palabras cerradas son aquellas palabras que aparecen frecuentemente en un texto sin agregar demasiado significado y provocan que la búsqueda de información sea lenta e ineficaz. Por esta razón, pueden ser omitidas las preposiciones, artículos y conjunciones. A éstas se les llama en español palabras cerradas y en inglés son conocidas como **stoplist**.

### 5.3.2 Antecedentes

Desde que el área de recuperación de información surgió, en la década de los 50's, se observó que muchas de las palabras que aparecen con frecuencia en un texto, como *de, a, con, la, y*, entre otras, no son útiles para recuperar información, pues sólo las palabras claves sirven para este propósito. Si un usuario del diccionario onomasiológico hace su búsqueda con palabras como *el, con, a, y o*, es muy probable que recupere casi todos los términos que existen en la banco terminológico, por lo que su búsqueda resultaría ineficaz. Por esta razón se deben separar las palabras cerradas de las palabras claves.

Christopher Fox<sup>68</sup> reconoce que no se sabe a ciencia cierta qué palabras deben ser incluidas en una lista de palabras cerradas. Se ha mostrado que las palabras con mayor aparición en un texto formarían la lista de palabras cerradas. Sin embargo, si se sigue este criterio estrictamente se corre el riesgo de desechar palabras importantes solamente porque son frecuentes. Por ejemplo, en el banco terminológico del área de Física aparece *movimiento* en 322 definiciones, es decir, es una palabra que aparece en varias definiciones y es sumamente importante porque ayuda a definir a los términos en los que se encuentra. Por ejemplo, se define *movimiento uniformemente acelerado*:

*Es un movimiento en línea recta en el que la aceleración es constante*

Aquí *movimiento* es un sustantivo del que se habla, por eso es una palabra clave. Debido a esto las palabras consideradas como cerradas fueron analizadas una por una para evitar descartar palabras que aportan significado a la definición de un término.

---

<sup>68</sup> C. Fox 1992, Capítulo 7 "Lexical analysis and stoplists" en *Information Retrieval: data structures and algorithms*, New Jersey, USA, Prentice Hall, p. 113.

Por otro lado, una lista grande de palabras será más útil que una pequeña. De esta manera, se hace hincapié en la necesidad de añadir otro tipo de palabras aparte de los artículos, conjunciones y preposiciones, tales como algunos verbos, adverbios y adjetivos.<sup>69</sup>

A continuación se agrega una lista de palabras propuesta por Fox en la que se incluyen preposiciones, artículos, conjunciones, adverbios, verbos, sustantivos, pronombres y adjetivos.

a about above across after again against all almost alone along already also although always among an and another any anybody anyone anything anywhere are area areas around as ask asked asking asks at away
b back backed backing backs be because become becomes became been before began behind being beings best better between big both but by
c came can cannot case cases certain certainly clear clearly come could
d did differ different differently do does done down downed downs during
e each early either end ended ending ends enough even evenly ever every everybody everyone everything everywhere
f face faces fact facts far felt few find finds first for four from full fully further furthered furthering furthers
g gave general generally get gets give given gives go going good goods got great greater greatest group grouped grouping groups
h had has have having he her herself here high higher highest him himself his how however
i if important in interest interested interesting interests into is it its itself
j just
k keep keeps kind knew know known knows
l large largely last later latest least less let lets like likely long longer longest
m made make making man many may me member members men might more most mostly mr mrs much must my myself
n necessary need needed needing needs never new newer newest next no non not nobody noone nothing now nowhere number numbered numbering numbers
o of off often old older oldest on once one only open opened opening opens or order ordered ordering orders other others our out over
p part parted parting parts per perhaps place places point pointed pointing points possible present presented presenting presents problem problems put puts
q quite
r rather really right room rooms
s said same saw say says second seconds see sees seem seemed seeming seems several shall she should show showed showing shows side sides since small smaller smallest so some somebody someone something somewhere state states still such sure
t take taken than that the their them then there therefore these they thing things think thinks this those though thought thoughts three through thus to today together too took toward turn turned turning turns two
u under until up upon us use uses used
v very
w want wanted wanting wants was way ways we well wells went were what when where whether which while who whole whose why will with within without work worked working works would
x
y year years yet you young younger youngest your yours
Z

Tabla 5.1 Lista de palabras de Fox

El autor Jacques Savoy<sup>70</sup> siguió las normas descritas de Fox (1990) para establecer una lista de palabras cerradas en otros idiomas diferentes al inglés. Primero, clasificó todas las palabras que

<sup>69</sup> Ziviani, p. 167.

encontró en un corpus de acuerdo con su frecuencia de aparición y tomó las 200 palabras que aparecen con mayor frecuencia. Segundo, examinó esta lista para no incluir todos los números a la lista de palabras cerradas, como “1994”, “1”, “19”, etcétera. También todos los sustantivos y adjetivos que más o menos tenían relación con el tema principal del corpus. Tercero, incluyó en la lista de palabras cerradas varios pronombres personales o posesivos, tales como *meine* (mía), en alemán; preposiciones *nello* (en el), en italiano; y conjunciones *ou* (donde), en francés; incluso, si no eran de las primeras 200 palabras más frecuentes.

El resultado de la lista de palabras cerradas contiene un gran número de pronombres, artículos, preposiciones y conjunciones. También fueron consideradas como palabras cerradas algunas formas verbales como, *sein* en alemán, y *essere* en italiano, que son equivalentes a *to be* en inglés.

La lista de palabras cerradas es la siguiente:

A acuerdo adelante ademas además adrede ahí ahí ahora al allí allí alrededor antano antaño ante antes apenas aproximadamente aquel aquél aquella aquélla aquellas aquéllas aquello aquellos aquéllos aquí aquí arriba abajo así así aun aún aunque
B bajo bastante bien breve
C casi cerca claro como cómo con conmigo contigo contra cual cuál cuales cuáles cuando cuándo cuanta cuánta cuantas cuántas cuanto cuánto cuantos cuántos
D de debajo del delante demasiado dentro deprisa desde despacio despues después detras detrás día día días días donde dónde dos durante
E el él ella ellas ellos en encima enfrente enseguida entre es esa ésa esas ésas ese ése eso esos ésos esta está ésta estado estados estan están estar estas éstas este éste esto estos éstos ex excepto
F final fue fuera fueron
G general gran
H ha había había habla hablan hace hacia han hasta hay horas hoy
I incluso informo informé
J junto
K
L la lado las le lejos lo los luego
m mal mas más mayor me medio mejor menos menudo mi mí mia mía mias mías mientras mio mío mios míos mis mismo mucho muy
N nada nadie ninguna no nos nosotras nosotros nuestra nuestras nuestro nuestros nueva nuevo nunca
O os otra otros
P país país para parte pasado peor pero poco por porque pronto próximo puede
Q que qué quien quién quienes quiza quizá quizás quizás
R raras repente
S salvo se sé segun según ser sera será si si sido siempre sin sobre solamente solo sólo son suyos su supuesto sus suya suyas suyo
T tal también también tampoco tarde te temprano ti tiene todavía todavía todo todos tras tu tú tus tuya tuyas tuyo tuyos
U un una unas uno unos usted ustedes
V veces vez vosotras vosotros vuestra vuestras vuestro vuestros
W
X
Y ya yo
Z

Tabla 5.2 Lista de palabras cerradas de Savoy

<sup>70</sup> <http://www.unine.ch/info/clef/>

### 5.3.3 Procedimiento para la obtención de las palabras cerradas

Lo que arriba se menciona son las listas de palabras cerradas que algunos autores proponen y los criterios que siguieron para formar las listas. A continuación se explica el procedimiento para determinar la lista de palabras cerradas para los propósitos del diccionario onomasiológico en las áreas de Física y Desastres.

- En primer lugar, se analizaron las palabras que por la mayoría de los autores son consideradas cerradas, a fin de tener criterios acordes con los trabajos realizados y utilizar su experiencia.
- Se obtuvieron como archivos electrónicos en formato texto al conjunto de definiciones capturadas en el banco terminológico de Física y Desastres del Grupo de Ingeniería Lingüística.<sup>71</sup> Este banco terminológico sirve de base para el diccionario onomasiológico que se desarrolla en el presente trabajo y está integrado por los términos de cada una de las áreas más varias definiciones para cada término según diferentes fuentes. De esta manera, se asegura que se tenga una multiplicidad de características semánticas y léxicas para que se facilite la búsqueda onomasiológica.
- Después, se comprobó en el archivo del banco terminológico si efectivamente las palabras son cerradas en todos los casos. Para ello, se debía cumplir el requisito que al eliminarlas del texto no repercutiera en el significado de la definición. Para facilitar este proceso, fue necesario utilizar el programa *Wordsmith*. Es una de las herramientas de la Ingeniería Lingüística que permite, entre otras posibilidades, que la palabra que se desea analizar sea desplegada en todas las definiciones en las que aparezca, es decir, en su contexto. A esto se le llama **concordancias**.
- Sin embargo, en algunos casos, una palabra que se consideraba cerrada en la mayoría de las definiciones tenía sus excepciones, por lo que fue necesario señalar en qué término o términos era importante dicha palabra. De esta manera se creó una lista de palabras ambiguas. Por ejemplo, la preposición *bajo* podía ser sustituida por *con*, que no aporta

---

<sup>71</sup> <http://iling.torreingenieria.unam.mx/diccionarios/>

significado en todos los casos, excepto en el término *marea* del banco terminológico del área de Física, cuya definición dice:

*Horas después estará en su nivel más bajo*

En este caso es relevante porque está adjetivando a *nivel*. Pero como es el único caso, se considera una palabra ambigua. Por ello se hizo una división de palabras en dos tipos:

- **Palabras cerradas**, aquellas palabras que serán irrelevantes para todos los casos.
- **Palabras ambiguas**, aquellas palabras que son cerradas en la mayoría de los casos, excepto para algunos términos.

Al ser ordenadas las palabras cerradas en su contexto, esto es, su concordancia, se siguieron los siguientes criterios:

- Si en todos los casos eran irrelevantes, se incluyeron como palabras cerradas.
- Si la palabra cerrada aparecía en términos en los que algunas veces era importante, se indicó en qué casos sucedía esto. Así, se formó la lista de palabras ambiguas.
- Si en otros casos la palabra cerrada aparecía en términos en los que era importante y términos en los que no lo era, el criterio a seguir fue: si eran más las veces en que era importante, se consideró clave; si eran menos las veces en que era importante, se consideró ambigua.

En los siguientes apartados se muestra la lista de palabras cerradas ordenada por categoría gramatical en la que desfilan artículos, preposiciones, conjunciones, adverbios, adjetivos y verbos.

### 5.3.3.1 Preposiciones

Hay que aclarar que estas palabras por sí solas, fuera de un contexto, no dicen mucho para el que las ve escritas en un texto, sobre todo las que se usan más frecuentemente como: *a, de, en, con, por*. En cambio, existen otras que se usan con menor frecuencia, pero que insinúan por sí solas la idea que encierran, como *sin* y *hasta* que indican, respectivamente, la idea de excepción y término espacial o temporal.<sup>72</sup>

Las preposiciones forman una lista pequeña de palabras, de las cuales:

- No se encontraron en nuestro corpus: *ante, cabe, pro, so y tras*.
- Cerradas: *a, al<sup>73</sup>, con, de, del, en, entre, para, por, según, sobre*.
- Ambiguas: *bajo*.

### 5.3.3.2 Artículos

Los artículos se clasifican en determinados o definidos y en indeterminados o indefinidos. En todas las definiciones en que aparecen en el corpus fueron palabras cerradas.

- Cerradas: *el, los, la, las, lo, un, una, unas, unos*.

### 5.3.3.3 Conjunciones

Las conjunciones sirven para unir frases u oraciones, por lo que normalmente no aportan ningún significado. De las que se encontraron en el corpus, se consideran las siguientes:

- Cerradas: *aún, aunque, bien, cuando, e, entonces, ni, o, pero, porque, pues, que, si, sino, u, y, ya*.

---

<sup>72</sup> Real Academia Española 1985, *Esbozo de una gramática de la Lengua Española*, Madrid, Espasa-Calpe, p. 435.

#### 5.3.3.4 Verbos

Con base en el procedimiento descrito arriba, se analizaron las concordancias para cada uno de los verbos del corpus, en cada una de sus flexiones. De esta manera, se identificaron a los verbos auxiliares (ser, haber, poder, tener) como palabras cerradas en todas sus flexiones, excepto en el infinitivo del verbo ser. Si bien hay varios verbos entre los definidos como palabras cerradas que resultan relevantes, hay que considerar que para el caso específico de Desastres y Física, estas palabras son irrelevantes, ya que no proporcionan ningún elemento significativo para la descripción del término. En la lista inmediata sólo aparecen los verbos en la forma flexionada tal y como aparece en el corpus.

- Cerradas: cuidado, debe, deben, debido, debida, decir, deja, dice, dícese, dicha, dichas, dicho, dichos, era, eran, es, está, estaba, estamos, estan, están, estar, fue, fuera, fueron, ha, haber, había, han, hay, haya, hubiera, incluso, ir, podemos, poder, podía, podría, posee, poseen, poseída, pudo, pueda, puede, pueden, queda, quedó, quiere, quiero, saber, sea, será, sería, sido, siendo, son, tener, tenemos, tenga, tenía, tenido, tiene, tienen, tomar, ve, vista, visto, viene, vienen.
- Ambiguas: conocido, conocida, hacer, hecho, pesar, va.

#### 5.3.3.5 Adjetivos

Se analizaron uno por uno para verificar que efectivamente agregaran una cualidad o característica a un sustantivo que contribuyera a definir al término; de no ser así, aparecen en la siguiente lista.

- Cerradas: algún, alguna, algunas, alguno, algunos, ambos, ambas, cada, ciertas, cierto, ciertos, consiguiente, demás, esa, esas, ese, esos, esta, estas, este, esto, estos, fácil, mejor, muchas, mucho, muchos, ningún, ninguna, nueva, poco, poca, posible, primer, primeras, puras, siguientes, sola, solo, tal, tales, tanto, toda, todas, todos, otra, otras, otro, otros, último, única, uno, varias, varios.
- Ambiguas: cierta, general, primero, primera, todo.

---

<sup>73</sup> *Al y del* forman una contracción compuesta de una preposición y un artículo. Se considerarán como preposiciones.

### 5.3.3.6 Adverbios

Los adverbios que a continuación se enumeran son de lugar, de tiempo, de modo, relativo, de negación, de afirmación y de cantidad.

- Cerradas: acerca, actualmente, además, ahí, ahora, así, bastante, casi, cuanto, cuánto, donde, exactamente, mas, más, menos, mientras, muy, nunca, no, realmente, relativamente, sí, siempre, solamente, sólo, también, tan, todavía, totalmente, únicamente, ya.
- Ambiguas: como.

### 5.3.3.7 Pronombres

Los pronombres son categorizados normalmente como palabras cerradas. En el corpus se encontraron las siguientes:

- Cerradas: algo, aquél, aquel, aquella, aquellas, aquello, aquellos, cual, cuál, cuales, cualesquiera, cualquiera, cualquier, cuya, cuyas, cuyo, cuyos, él, ella, ellas, ello, ellos, ésa, eso, ésta, éstas, éste, éstos, le, les, nos, nosotros, nuestra, nuestras, nuestro, nuestros, quien, se, su, sus, suya, ti.

### 5.3.3.8 Sustantivos

Los siguientes son clasificados como sustantivos y considerados como palabras cerradas.

- Cerradas: ejemplo, etc., vez.

## 5.4 Expansión de las palabras claves

El usuario, al requerir el término de una definición, introducirá palabras que, como ya se demostró en la unidad cuatro ampliamente, no serán las que el lexicógrafo empleó para definir ese término. Por ejemplo, un usuario pretende encontrar el término *barómetro* que el lexicógrafo define así:



### *Instrumento para medir la presión atmosférica*

Sin embargo, el usuario introduce lo siguiente:

### *Aparato para calcular la presión atmosférica*

Es decir, se introdujeron dos palabras que el lexicógrafo no utilizó para definir al término *barómetro*: *aparato* y *calcular*. Estas palabras son sinónimos de *instrumento* y *medir*, respectivamente. Esto implica que el sistema debe contar con los sinónimos de cada una de las palabras claves. La expansión significa agrupar todas las posibles palabras que un usuario pueda emplear para hacer su búsqueda. Existen dos formas de expandir las palabras claves, la expansión morfológica y la expansión semántica, las cuales se tratarán a continuación.

#### **5.4.1 Expansión morfológica**

La expansión morfológica se refiere a los cambios morfológicos que puede tener una palabra, a estas variaciones se les conoce como flexión, composición y derivación.

- **Flexión**, que define las variaciones posibles de la forma radical o raíz, variaciones que se agrupan bajo el término de paradigmas. Por ejemplo, *pasearíamos*. Sin modificar lo que dice la raíz, el sufijo proporciona información complementaria.
- **Composición**, que consiste en la combinación de palabras completas para dar origen a nuevas formas. Por ejemplo, *pelirrojo*, *pelegrino*, *boquirrubio*, *hombre-rana*.
- **Derivación**, que se ocupa de la formación de nuevas raíces (flexionables) a partir de otras raíces que, a menudo, pertenecen a categorías gramaticales diferentes. Por ejemplo, *nación* (N), *nacional* (Adj), *nacionalizar* (V), *nacionalismo* (N), *nacionalista* (N), *internacionalizar* (V). En tanto la raíz dice una cosa, los afijos la convierten en otra cosa.

Como no se restringe la búsqueda que un usuario hace, puede introducir su input utilizando las flexiones verbales o derivaciones que puede tener una palabra. Por ejemplo:

Aparato	en el que se	<i>determina</i>	la presión del aire.
Aparato	por el cual	<i>determinarías</i>	la presión del aire
Aparato	para	<i>determinar</i>	la presión del aire.

Tabla 5.3 Ejemplo de expansión morfológica

Estos cambios no aparecen en el banco terminológico debido a que se ocuparía mucho espacio en la base de datos. Para resolverlo y mejorar la efectividad de la recuperación de información, se utiliza la lematización, un proceso automático que se utiliza para separar la raíz o lexema del morfema de una palabra, dejando únicamente la raíz. Por ejemplo:

*determin a*  
*determin arias*

La raíz es la parte del verbo que contiene el significado. El morfema proporciona la información gramatical del verbo: persona, número, tiempo, modo y voz.

Para que no se capturen cada uno de los cambios que una palabra puede tener, se lematizan. Con ello, en vez de tener tres palabras con una misma raíz y de ocupar demasiado espacio en la base de datos, se tiene solamente la raíz que permite abarcar todas las variaciones de una palabra. De esta manera, se mejora la efectividad de la búsqueda del usuario.

Uno de los métodos que existen para lematizar es el Algoritmo de Porter<sup>74</sup>. Originalmente se creó para el idioma inglés, cuyas flexiones y derivaciones son más sencillas que las del español. Actualmente, existe una adaptación para el español. El algoritmo lematiza las palabras claves que el usuario introduce para realizar su búsqueda. Sin embargo, hay que tomar en cuenta que este algoritmo presenta algunos errores, como **sobrelematización** y **sublematización**. En la primera, el algoritmo lematiza de más una palabra, provocando que se confunda con otras, como sería el caso lematizar *manera* por *man*. En la segunda, el algoritmo no realiza una lematización completa, por lo que dos palabras que debieran tener el mismo lema no lo tienen, como es el caso de *orgullosísimo* y

<sup>74</sup> M. F. Porter 1980, *An algorithm for suffix stripping*. Program, 14(3).

*orgullosamente*, que los lematiza como *orgullosisim* y *orgull*, respectivamente. Aunque presenta estas fallas, el algoritmo se ha comprobado que es el mejor de los algoritmos existentes.<sup>75</sup>

#### 5.4.2 Expansión semántica

Para resolver el problema de la diferencia de terminología empleada por el usuario y el lexicógrafo, se ha demostrado, en el campo de recuperación de información, la ventaja de expandir automáticamente la formulación original del usuario con agrupamientos semánticos de palabras claves relacionadas. La expansión semántica es una herramienta que consiste en proporcionar palabras que tienen el mismo o similar significado, pero que su forma es totalmente diferente, esto es, sinónimos. Significa que se necesita encontrar todos los sinónimos de las palabras claves introducidas por el usuario, para que sea compatible con la base del lexicógrafo. Por ejemplo, si se busca el término *segunda Ley de Kepler* que el lexicógrafo define así:

*La línea imaginaria que une a un planeta con el sol barre áreas iguales en tiempos iguales.*

Pero el usuario introduce:

*La línea entre un planeta y el sol describe áreas iguales en el mismo tiempo.*<sup>76</sup>

Con la expansión semántica es posible encontrar el término que el usuario busca aunque no utilice las mismas palabras que el lexicógrafo. En este contexto, el significado de las dos palabras, *barre* y *describe*, es similar.

Sin embargo, en la práctica es diferente. Resulta que al buscar en un diccionario de sinónimos las dos palabras antes mencionadas se encontrará que una no es sinónima de la otra. Al consultar el *Diccionario Anaya* se encontró que los sinónimos de *barrer* son: carcomer, demoler, derribar, derrumbar, derruir, descomponer, deshacer, limpiar, cepillar, escobillar, barriquear, arrastrar, desembrozar, arrollar.

---

<sup>75</sup> G. Castillo 2002, Tesis de Maestría *Algoritmo revisado para la extracción automática de agrupamientos semánticos*, México, p. 16

<sup>76</sup> Las dos definiciones de la Ley de Kepler son obtenidas del banco terminológico del GIL.

En tanto los sinónimos de *describir* para el mismo diccionario son: dilucidar, divulgar, elucidar, esclarecer, especificar, explanar, glosar, disponer, tener, considerar, explicar, definir, determinar, exponer, detallar, referir.

De igual manera puede buscarse en otros diccionarios de sinónimos y se podrá notar que *barrer* y *describir* no son sinónimos para estos diccionarios, puesto que en ninguno de ellos la búsqueda coincide en que lo sean. El hecho de que no sean sinónimos para estos diccionarios no quiere decir que de ninguna manera lo sean, pues para el hablante sí lo son. Además, como se verá más adelante, dos palabras son sinónimas si pueden ser sustituidas una por otra dentro de cierto contexto. Los diccionarios de sinónimos están pensados para la lengua y no pueden abarcar todos los contextos. Por ello, si bien es necesario utilizar un diccionario de sinónimos, también es conveniente utilizar otros medios alternativos que permitan una mayor exhaustividad en la expansión semántica y que permitan cubrir una mayor cobertura de usuarios.

Para expandir semánticamente las palabras se utilizan de manera complementaria tres herramientas de las que se dispone: un diccionario de sinónimos, *WordNet* y el *Algoritmo flexibilizado de agrupamiento semántico*. A continuación se explicarán los beneficios y desventajas de cada una de ellos.

#### 5.4.2.1 Diccionario de sinónimos

En sí, los diccionarios de sinónimos tienen dos desventajas: la primera es que al buscar los sinónimos de un par de palabras, como en el ejemplo de *barre* y *describe*, se encuentra que una y otra no son sinónimas; la segunda es que cuando un par de palabras son sinónimas, el diccionario aporta demasiadas palabras que no sirven.

Un diccionario de sinónimos que se obtuvo en formato electrónico es el *Anaya*. Con esto se facilita su manipulación para el proceso de búsquedas y la expansión semántica. Para cada entrada se enlistan los sinónimos sin ninguna distinción o clasificación más detallada. Cuando el usuario introduce una palabra clave para realizar la búsqueda en el diccionario onomasiológico, el sistema busca con todas las palabras ofrecidas por el *Anaya*.

### 5.4.2.2 WordNet

Una herramienta muy utilizada para distintos desarrollos de la ingeniería lingüística es *WordNet*. Es un sistema de representación léxica que intenta desarrollar un lexicón multipropósito en forma de aplicación informática. Es un sistema electrónico de referencia léxica, desarrollado en forma de bases de datos léxica, cuyo diseño está en consonancia con teorías psicolingüísticas relativas a la organización de la información léxica en la mente del hablante. Está disponible gratuitamente en línea a través de Internet.

Las relaciones léxicas de *WordNet* son la sinonimia, antonimia, superordinación (hiperonimia), subordinación (hiponimia) y meronimia (todo-parte). Si bien todas las relaciones son importantes, para el diccionario onomasiológico sólo se utilizan las relaciones de sinonimia. Los hiperónimos e hipónimos quedan pendientes para trabajos futuros.

Al igual que un diccionario de sinónimos, *WordNet* presenta los mismos problemas. Aunque es más completo que el *Anaya*, presenta saturación de información. Sin embargo, la organización de *WordNet* permite separar los sinónimos de una palabra por acepciones. Gracias a esto se puede resolver el problema de separar la información que resulta inútil. Así, se utilizan solo las acepciones correspondientes al mismo contexto, para lo cual es necesario conocer el área de conocimiento a la que se va a referir para saber si una acepción será de utilidad o no. Por ejemplo, tenemos la palabra *cuerpo*, cuyas acepciones son: *volumen* y *cadáver*.

La primera acepción servirá para el área de Física porque se encuentra en el banco terminológico, dicha palabra aparece en el término *masa total de un sistema*. Pero el segundo, servirá para un área relacionada con medicina pues *cadáver* no se encuentra en ninguno de los bancos terminológicos. Sin embargo, si en *cadáver* existiera una palabra que se encuentra en dicho banco, por esa sola palabra, se utilizará todo el contenido de esa acepción y será utilizada como sinónimo de *cuerpo*. Se tomará el ejemplo, que arriba se menciona, de *medir*:

*competir, valorar, calcular, evaluar, determinar, enfrentarse, arquear, contenerse, refrenarse, establecer, dominarse, cronometrar, contar, mesurar, comprobar, moderarse, calibrar*

Se busca en la primera serie de sinónimos si alguna de esas palabras se encuentra en el banco terminológico. La palabra que se encontró fue *determinar* en la definición del término *pesa*. Se

continúa buscando hasta encontrar todas las palabras que existan en el banco. Si solamente se encontró determinar se usará todo el contenido de esa acepción como sinónimo de *medir*. En la tercera serie de sinónimos ninguno de ellos se encontró en la base terminológica de Física por lo que se deshecha esa serie. Lo mismo se hace con cada una de los sinónimos que se encuentren.

Con este proceso se elimina un 50% aproximadamente de palabras que se encuentren en el diccionario de sinónimos y que no sirvan.

### 5.4.2.3 Algoritmo flexibilizado de agrupamiento semántico

Para solucionar el problema de la sinonimia usamos el *Algoritmo Flexibilizado de Agrupamiento Semántico*<sup>77</sup>, llamado en inglés *clustering*. Este algoritmo, mejorado por Castillo, se basa en uno desarrollado por Sierra y McNaught<sup>78</sup>. Utiliza como entrada un conjunto de términos y sus definiciones (provenientes de diferentes fuentes), compara estas definiciones e identifica qué palabra con que otra palabra coincide semánticamente hablando. Reúne palabras cuyo significado, en el contexto en el que aparecen, se pueden considerar como sinónimos. Los pares de palabras que son detectados se pueden sustituir por la pareja con la que fue comparada o alineada de tal manera que no afecta al significado de las definiciones en las que aparecen. Por ejemplo:

*Movimiento de un cuerpo en un campo gravitatorio...*

*Descenso de un cuerpo sometido únicamente a la...*

Este es un par de palabras que el algoritmo alineó y encontró como sinónimos, pues si se emplea una en lugar de la otra, se nota que el significado de la definición no cambia considerablemente:

*Descenso de un cuerpo en un campo gravitatorio...*

*Movimiento de un cuerpo sometido únicamente a la...*

---

<sup>77</sup> G. Castillo y G. Sierra 2003, "Algoritmo flexibilizado de agrupamiento semántico". *Estudios de Lingüística Aplicada*, Núm. 38.

<sup>78</sup> G. Sierra y J. McNaught 2000 "Extracting semantic clusters from MRDs for an onomasiological search dictionary". *International Journal of Lexicography*. Vol. 13 (4).

De esta manera el algoritmo encuentra que *movimiento* y *descenso* son sinónimos. Es importante señalar que este algoritmo fue creado originalmente con el fin de obtener pares semánticos a partir de la comparación de definiciones y resulta adecuado para los propósitos de diccionarios onomasiológicos. Además, los pares obtenidos son bastante prometedores y, lo más interesante, es que si buscamos estos pares en los diccionarios de sinónimos, observaremos que no se consideran como sinónimos. Por ejemplo, para el caso del Banco Terminológico de Física, se encontraron, entre otros, los siguientes pares semánticos: *parte-rama*, *considerar-importar*, *atenerse-atender*, *relaciona-estudia-provoca-trata*, *empleado-invertido*, *revoluciones-veces*. Por ello es de utilidad este método que utiliza las definiciones en su contexto.

### 5.5 Búsqueda en el banco terminológico

Una breve descripción de la manera en que se realiza la búsqueda en el banco terminológico es la siguiente:

1. Una vez que el usuario tiene el concepto en mente y lo introduce, este input es separado en palabras.
2. La lista de palabras obtenidas es comparada con la lista de palabras cerradas, si la palabra comparada está dentro de la lista de palabras cerradas, entonces la palabra se elimina. Las palabras restantes se consideran palabras claves de la búsqueda.
3. Las palabras claves deben someterse a un proceso de lematización, es decir, se obtiene la raíz de una palabra, removiendo su información derivativa o flexiva como se mostró anteriormente.
4. Una vez lematizadas las palabras, cada una de ellas se busca en una lista que contiene todas las palabras dentro del contexto de búsqueda: a éste proceso se le denomina expansión de términos:
  - a. Se incorpora cada palabra clave a una nueva lista de palabras, denominada *lista de palabras expandida*.

- b. Se buscarán los agrupamientos semánticos de sinónimos correspondientes a las palabras clave, en caso de existir uno o más agrupamientos se incorporan todas las palabras del agrupamiento a la lista de palabras expandida.
5. Una vez obtenida la lista expandida de palabras y para cada palabra en la lista, se procede a la búsqueda en una lista que contiene agrupamientos de palabras claves y términos. Identificando para cada palabra de la lista expandida qué términos la contienen. Se debe llevar una lista de cuántos términos se han encontrado y cuantas palabras clave los han seleccionado. A este conjunto se le denomina lista de términos y ocurrencias.
6. Con base en la lista de términos y ocurrencias, se presentan los términos con mayor ocurrencia y se ofrece la alternativa de mostrar todos los términos identificados. Además como segunda opción se deberá permitir que el usuario consulte la definición del término.

## 5.6 Conclusiones

Este capítulo tuvo como objetivo dar a conocer el proceso del diseño de búsquedas onomasiológicas en lenguaje natural. Se describió la forma en que el sistema (a partir de la descripción del concepto en lenguaje natural proporcionada por el usuario) realiza el proceso de identificación de palabras claves, las expande (morfológica y semánticamente) y finalmente identifica los términos más probables.

La lista de palabras cerradas, en este trabajo presentado, no fue hecha para ser utilizada en cualquier contexto, es decir, no es universal. Se creó para ser utilizada en el banco terminológico de dos áreas, Física y Desastres. Sin embargo, aunque se pretendiera hacer una lista de palabras cerradas universal no es posible debido a que si se desea definir *artículo* por enumeración, entonces *la* sería una palabra clave. Por tanto, se debe tomar en cuenta la base de datos que se utilizará, es decir, si serán términos o palabras de uso común y el usuario al que se enfocará el diccionario onomasiológico.



# CAPÍTULO 6

## HERRAMIENTAS COMPUTACIONALES



En el capítulo anterior se manejó el diseño de búsquedas onomasiológicas en lenguaje natural, es decir, la forma en que opera el sistema onomasiológico y la manera en como se lleva a cabo.

Sin embargo, para realizar este proceso de manera automática contamos con diferentes herramientas computacionales que nos van a permitir manipular cantidades muy grandes de información cómo lo son las Bases de Datos, para que los usuarios puedan interactuar con dicha información necesitamos un sistema que lo permita y para realizarlo son necesarias otras herramientas como los lenguajes de programación y los Servidores WEB.

La tecnología de las bases de datos se ha descrito como “una de las áreas de la ciencia de la computación y la información de más rápido desarrollo”. El procesamiento de la información es esencial para la administración de los gobiernos, los negocios y la educación, etc. En nuestra sociedad es vital para una organización o empresa proporcionar información correcta y puntual para apoyar la toma de decisiones y otras actividades gerenciales.

En el presente capítulo se analizarán las herramientas computacionales que se utilizarán para el desarrollo del sistema y se manejarán definiciones y conceptos para su mejor entendimiento.

Inicialmente, las aplicaciones individuales se desarrollaban independientemente, y cada programa de aplicación procesaba sus propios archivos privados. Como resultado, algunas actividades se duplicaban y la información redundante se almacenaba para usarla en distintas operaciones.

Al final de los años 60 surgió el sistema de bases de datos para superar los problemas asociados con los sistemas de información tradicionales. En vista de la centralización de los datos por medio de un sistema de base de datos, los requerimientos de todos los usuarios se podían coordinar de una manera efectiva para alcanzar la mejor utilidad general para la organización.

A continuación se dará la definición de bases de datos.

### 6.1 Definición y concepto de las bases de datos

Las **bases de datos** son una colección de archivos interrelacionados creados con un DBMS (Sistema Manejador de Bases de Datos). El contenido de una base se obtiene combinando datos de todas las diferentes fuentes en una organización, de tal manera que los datos estén disponibles para todos los usuarios, y los datos redundantes puedan eliminarse, o al menos minimizarse.

La figura 6.1 muestra la base de datos como un recipiente de datos a ser compartidos por varios programas. El usuario podrá recobrar datos de varias partes de la base ya que los archivos ahí almacenados, están conectados directa o indirectamente.

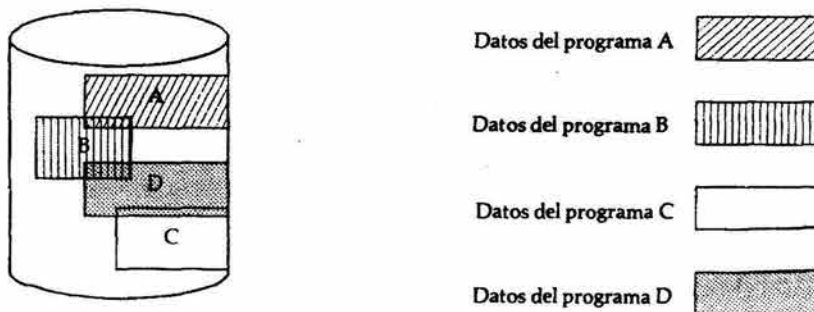


Fig. 6.1. Vista esquemática de las bases de datos

### 6.1.1 Sistema manejador de bases de datos (DBMS)

El sistema de manejo de bases de datos es la porción más importante del software de un sistema de bases de datos. Es una colección de numerosas rutinas de software interrelacionadas, cada una de las cuales es responsable de alguna tarea específica.

El DBMS interpreta y procesa las peticiones del usuario para recobrar información de la base. La figura 6.2 muestra como un DBMS sirve de interfase entre las peticiones del usuario y la base de datos. Las preguntas a la base pueden tener distintas formas, pueden teclearse directamente desde la terminal, o codificarse como programas en lenguajes de alto nivel y presentarse para procesamiento interactivo o por lotes. En la mayoría de los casos, una petición de consulta tendrá que atravesar varias capas de software en el DBMS y en el sistema operativo antes de que se pueda acceder la base de datos física.

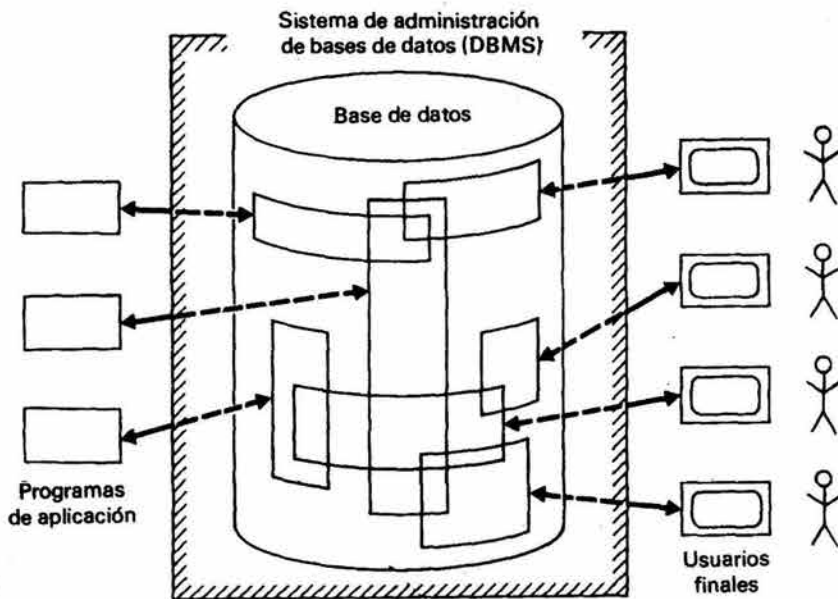


Fig. 6.2. Interfase entre las peticiones del usuario y la base de datos.

El DBMS responde a una pregunta llamando a los subprogramas apropiados, cada uno de los cuales realizará su función especial para interpretar la petición o localizar los datos deseados en la base y presentarlos en el orden solicitado. Así, el DBMS protege a los usuarios de la tediosa programación que tendrían que hacer para organizar el almacenamiento de los datos, o accederlos una vez almacenados.

Un sistema de base de datos es un sistema de mantenimiento de registros basado en computadoras, es decir, un sistema cuyo propósito general es registrar y mantener información. Tal información puede estar relacionada con cualquier cosa que sea significativa para la organización donde el sistema opera, es decir, cualquier dato necesario para los procesos de toma de decisiones inherentes a la administración de esa organización

#### **6.1.1.1 Funciones de un DBMS**

Las funciones principales de un DBMS son:

- i. Crear y organizar la base de datos
- ii. Establecer y mantener las trayectorias de acceso a la base de datos, de tal manera que los datos en cualquier parte de la base se puedan acceder rápidamente
- iii. Manejar los datos de acuerdo con las peticiones de los usuarios
- iv. Mantener la integridad y seguridad de los datos
- v. Registrar el uso de las bases de datos

#### **6.1.1.2 Arquitectura de un DBMS**

La arquitectura se divide en tres en tres niveles generales: **interno, conceptual y externo**. El nivel interno es más cercano al almacenamiento físico, es decir, el que concierne a la manera como los datos se almacenan en realidad; el nivel externo es el más cercano a los usuarios, es decir, el que atañe a la manera cómo cada usuario ve los datos; y el nivel conceptual es un nivel de mediación entre los otros dos. Si el nivel externo se desarrolla con las vistas de los usuarios individuales, el nivel conceptual puede considerarse como el que define una vista de la comunidad de usuarios. Las figuras 6.3 y 6.4 esquematizan la arquitectura general de un DBMS.

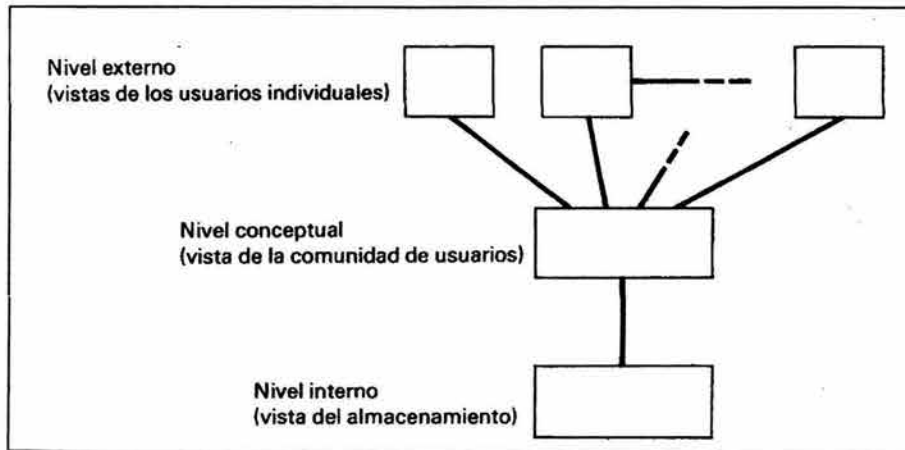


Fig. 6.3. Los tres niveles de la arquitectura

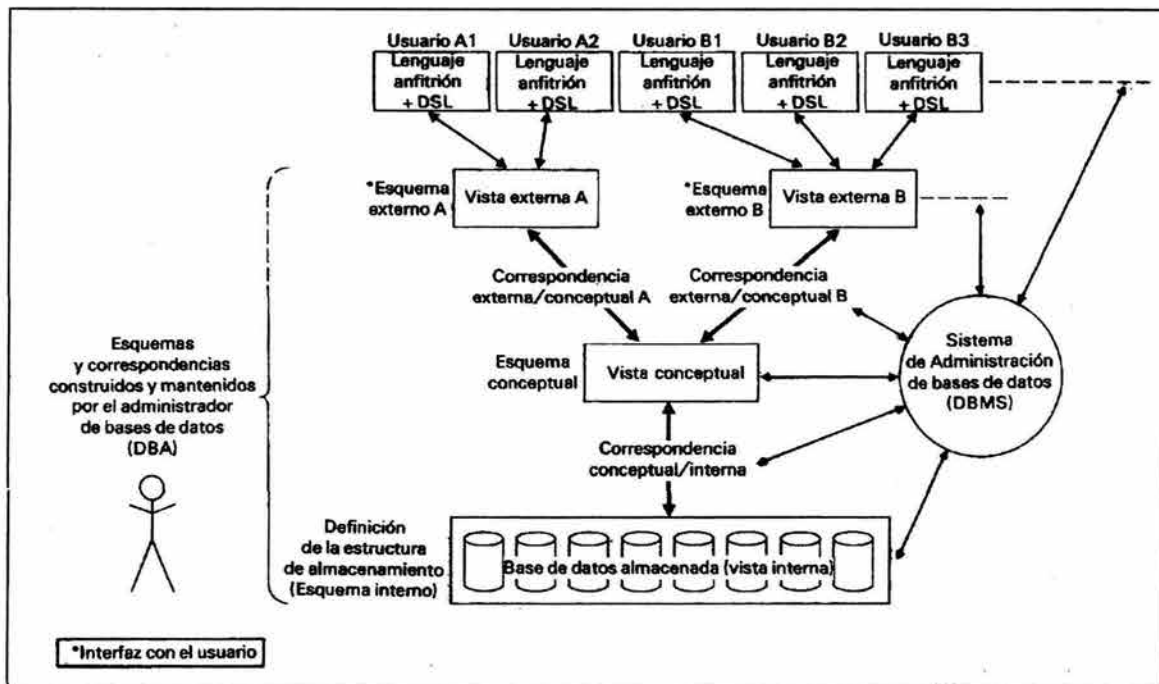


Fig. 6.4 Diagrama detallado de la arquitectura de un DBMS

### 6.1.1.3 Manejador de bases de datos SQL Server

SQL (*Standar Query Language*) es un lenguaje estandarizado de base de datos, el cual nos permite realizar tablas y obtener datos de ella de manera muy sencilla.

SQL es un lenguaje que consta de varias partes:

- Lenguaje de definición de datos ( DDL): Proporciona ordenes para definir esquemas de relación, eliminar relaciones, crear índices y modificar esquemas de relación.
- Lenguaje de manipulación de datos interactivos (DML): Incluye un lenguaje de consultas que permite rescatar datos de las relaciones. También incluye órdenes para insertar, suprimir y modificar tuplas.
- Lenguaje de manipulación de datos inmerso (EDML): La forma inmersa de SQL esta diseñada para usar dentro de los lenguajes de programación de lenguaje general.
- Definición de vistas (DDL): incluye órdenes para definir vistas.

### Estructura básica

La estructura básica de una expresión para consulta SQL consta de tres cláusulas:

- *SELECT*
- *FROM*
- *WHERE*

La cláusula *SELECT* se usa para listar los atributos que se desean en el resultado de una consulta.

La cláusula *FROM* lista las relaciones que se van a examinar en la evaluación de la expresión

La cláusula *WHERE* costa de un predicado que implica atributos de las relaciones que aparecen en la cláusula *FROM*.

Una consulta básica en SQL tiene la forma:

*SELECT A1,A2,...,An*

*FROM r1,r2,...,rn*

*WHERE P*

Donde  $A_i$  = atributo (Campo de la tabla)

$r_i$  = relación ( Tabla )

$P$  = predicado (condición)

Por razones institucionales, para el desarrollo de la tesis se utilizará SQL Server de Microsoft.

## 6.1.2 Modelo de datos

Un **modelo** es una representación de la realidad que sólo conserva los detalles relevantes. Una base de datos incorpora un modelo de la realidad.

Un modelo de datos es una técnica para estructurar la información a base de definir las entidades informativas necesarias y las relaciones existentes entre ellas.

El manejo de información implica el saber como organizar los datos. Un apoyo lo encontramos en las herramientas de bases de datos que a su vez se apoyan en el modelo de datos. Para que éste se encargue de organizar y manipular los datos, en los sistemas de bases de datos, encontramos a los manejadores de bases de datos, que son los que nos ayudan a almacenar la información y recuperarla. Entre los modelos que se pueden encontrar y los que se utilizaron para el desarrollo del sistema son:

- **Modelo relacional**
- **Modelo entidad-relación**
- **Modelo jerárquico:** Ver capítulo 7

### 6.1.2.1 Modelo relacional

La relación es el elemento básico en el modelo relacional y se puede representar como una tabla. Las bases de datos relacionales son el tipo de bases de datos actualmente más difundido. Los motivos de este éxito son fundamentalmente dos:

1. Ofrecen sistemas simples y eficaces para representar y manejar los datos.
2. Se basan en un modelo, el relacional, con sólidas bases teóricas.

El modelo relacional fue propuesto originariamente por E.F. Codd. Gracias a su facilidad de uso, el modelo se ha convertido en el más usado para la producción de DBMS.

La estructura fundamental del modelo relacional es precisamente esa "relación", es decir una tabla bidimensional constituida por líneas (tuplas<sup>79</sup>) y columnas (atributos<sup>80</sup>). Las relaciones representan

---

<sup>79</sup> *Tupla:* Es cada uno de los elementos que contiene una instancia de la relación (filas). Una tupla es una lista inmutable. Una tupla no puede modificarse de ningún modo después de su creación. Los elementos de una tupla tienen un orden definido, como los de una lista. Las tuplas tienen primer índice 0, como las listas, de modo que el primer elemento de una tupla no vacía es siempre  $t[0]$ .

<sup>80</sup> *Atributo:* Las columnas de las tablas se denominan atributos. Es una característica de interés o un hecho sobre una entidad o sobre una relación. Los atributos representan las propiedades básicas de las entidades y de las relaciones. Toda la información extensiva es portada por los atributos. Gráficamente, se representan mediante bolitas que cuelgan de las entidades o relaciones a las que pertenecen. Cada atributo tiene un conjunto de valores asociados denominado dominio.

las entidades que se consideran interesantes en la base de datos. Cada instancia de la entidad encontrará sitio en una tupla de la relación, mientras que los atributos de la relación representarán las propiedades de la entidad.

El modelo se compone de tres partes principales:

- Estructura de datos: Colección de objetos abstractos formados por datos (ejemplo: Dominios<sup>81</sup>, relaciones, atributos, tuplas). Las tablas tienen un número concreto de columnas y un número arbitrario de filas. Cada atributo debe tener un dominio. En la intersección de cada fila y columna se almacena un valor simple.
- Integridad de datos: Conformada por dos reglas generales de integridad:
  - Integridades de las entidades: Impone que toda tabla tenga exactamente una clave primaria. Una clave primaria es una combinación de uno o más atributos cuyo valor ubica unívocamente a cualquier fila de la tabla (una clave primaria no puede tener valor nulo).
  - Integridad referencial: Exige que el RDBMS mantenga la congruencia de todas las claves externas o ajenas con su correspondiente clave primaria. Una clave externa o ajena es una clave primaria de una tabla, que está incluida dentro de otra tabla.
- Manejo de datos: Operaciones del álgebra relacional (o equivalentes del cálculo): Unión, intersección, diferencia, producto, restricción, proyección, reunión, división asignación relacional.

Las ventajas del modelo relacional son:

- Compatibilidad y estandarización.
- Fiabilidad.
- Garantía de independencia de los datos.
- Existencia de numerosos sistemas comerciales entre los que escoger y consiguiente apoyo técnico.
- Conectividad garantizada con los lenguajes de programación estándar.

Las desventajas de modelo relacional son:

- Imposibilidad de representar conocimiento en forma de reglas.
- Inexistencia de mecanismos de herencia de propiedades (y por supuesto de métodos).
- Falta de poder expresivo (por ejemplo, para representar jerarquías).

---

<sup>81</sup> *Dominio*: Un dominio es un conjunto de valores validos. El dominio define todos los valores posibles que puede tomar un atributo. Puede haber varios atributos definidos sobre un mismo dominio.



- Dificultad para gestionar datos no atómicos (por ejemplo, los valores estructurados de una estructura de rasgos).
- Incompatibilidad entre los tipos de estructuras de datos que se transfieren o desadaptación de impedancia (*impedance mismatch*).

#### 6.1.2.1.1 Modelo Entidad-Relación

El **modelo entidad-relación** es el modelo conceptual más utilizado para el diseño ideal de bases de datos. Fue introducido por Peter Chen en 1976. El modelo entidad-relación está formado por un conjunto de conceptos que permiten describir la realidad mediante representaciones gráficas y lingüísticas. Para un mejor entendimiento a continuación se definen los conceptos de entidad y relación:

- **Entidad.**- Objeto del mundo real sobre el que queremos almacenar información. Las entidades están compuestas de atributos que son los datos que definen el objeto. De entre los atributos habrá uno o un conjunto de ellos que no se repite; a este atributo o conjunto de atributos se le llama clave de la entidad. En toda entidad siempre hay al menos una clave que en el peor de los casos estará formada por todos los atributos de la tabla. Ya que pueden haber varias claves y necesitamos elegir una, lo haremos atendiendo a estas normas:
  - Que sea única.
  - Que se tenga pleno conocimiento de ella.
  - Que sea mínima, ya que será muy utilizada por el gestor de base de datos.
- **Relación.**- Asociación entre entidades, sin existencia propia en el mundo real que estamos modelando, pero necesaria para reflejar las interacciones existentes entre entidades. Las relaciones pueden ser de tres tipos:
  - Relaciones 1-1.- Las entidades que intervienen en la relación se asocian una a una.
  - Relaciones 1-n.- Una ocurrencia de una entidad está asociada con muchas (n) de otra.
  - Relaciones n-n.-Cada ocurrencia, en cualquiera de las dos entidades de la relación, puede estar asociada con muchas (n) de la otra y viceversa.

Sus características fundamentales son:

- Reflejan tan sólo la existencia de los datos sin expresar lo que se hace con ellos.

- Es independiente de las bases de datos y de los sistemas operativos.
- Incluye todos los datos que se estudian sin tener en cuenta las aplicaciones que se van a tratar.
- Las entidades se representan como rectángulos, los atributos como elipses y las relaciones como rombos.

### 6.1.2.1.2 Normalización

**Normalización** es un conjunto de reglas que sirven para ayudar a los diseñadores a desarrollar un esquema que minimice los problemas de lógica. Cada regla está basada en la que le antecede. La normalización se adoptó porque el viejo estilo de poner todos los datos en un solo lugar, como un archivo o una tabla de la base de datos, era ineficiente y conducía a errores de lógica cuando se trataba de manipular los datos.

Básicamente, las reglas de Normalización están encaminadas a eliminar redundancias e inconsistencias de dependencia en el diseño de las tablas. Son las técnicas para prevenir las anomalías en las tablas. Dependiendo de su estructura, una tabla puede estar en primera forma normal, segunda forma normal o en cualquier otra, existen cinco formas normales y podemos observarlas en la figura 6.5.

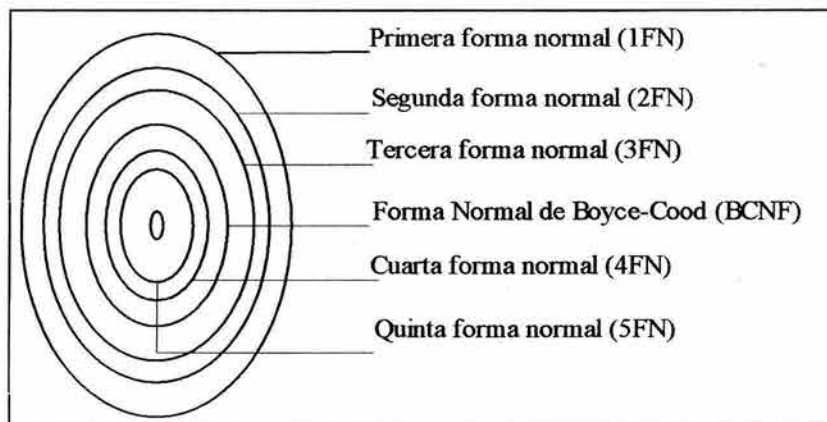


Figura 6.5. Formas Normales

La figura anterior muestra la secuencia de las formas normales, en principio debemos empezar la normalización por la parte externa y continuar hacia las internas.

A continuación se presenta una breve explicación de las formas normales que se muestran en la figura 6.5. Para la realización de la presente tesis se utilizaron las tres primeras formas normales.

- **FN1.** Se pretende garantizar la no existencia de grupos repetitivos. Un grupo repetitivo es un conjunto de atributos de igual semántica en el problema y dominio, que toman valores distintos para la misma clave. Cualquier esquema que tenga claves correctas está seguro en FN1. Con la primera forma normal se pretende:
  - Eliminar los grupos repetitivos de las tablas individuales.
  - Crear una tabla separada por cada grupo de datos relacionados.
  - Identificar cada grupo de datos relacionados con una clave primaria.
- **FN2.** Si FN1 y cada atributo de la tabla que no forma parte de la clave depende funcionalmente de forma completa de la clave primaria. Es decir, depende de toda la clave y no de ningún subconjunto de ella. Se pretende garantizar una correcta elección de claves y eliminar redundancias. Si la clave está formada por un único atributo entonces ese esquema estará seguro en segunda forma normal. Con la segunda forma normal se pretende:
  - Crear tablas separadas para aquellos grupos de datos que se aplican a varios registros.
  - Relacionar estas tablas mediante una clave externa.
- **FN3.** Si FN2 y cada atributo no primo de la tabla no depende funcionalmente de forma transitiva de la clave primaria. Con la tercer forma normal se pretende:
  - Eliminar aquellos campos que no dependan de la clave.
- **FNBC (Forma Normal de Boyce-Codd).** Se basa en el concepto de determinante funcional: uno o varios atributos de una tabla de los cuales dependen funcionalmente de forma completa algún otro atributo de la misma tabla. Una relación está en FNBC si FN1 y cada determinante funcional es una clave candidata de la tabla. Así se garantiza que se han elegido bien las claves al no existir dependencias funcionales entre atributos que no son clave. Cada vez que se verifica una dependencia funcional  $a \rightarrow b$  entonces  $a$  es clave primaria o alterna con seguridad. Todas las dependencias funcionales cumplen que en su parte izquierda solo aparecen atributos que son parte de una clave candidata. Esta forma normal es más restrictiva que la tercera y tiene la interesante propiedad de que su cumplimiento implica la satisfacción de FN3 o sea que  $FNBC \rightarrow FN3$ .

- **FN4.** En las relaciones varios-con-varios, entidades independientes no pueden ser almacenadas en la misma tabla. Ya que sólo se aplica a las relaciones varios-con-varios, la mayoría de los desarrolladores pueden ignorar esta regla de forma correcta.
- **FN5.** Existe otro nivel de normalización que se aplica a veces, pero es de hecho algo en el que en la mayoría de los casos no es necesario para obtener la mejor funcionalidad de la estructura de datos o aplicación. Su principio sugiere:
  - La tabla original debe ser reconstruida desde las tablas resultantes en las cuales a sido particionada.

Los beneficios de aplicar esta regla aseguran que no se ha creado ninguna columna extraña en las tablas y que la estructura de las tablas que se ha creado sea del tamaño justo que tiene que ser. Es una buena práctica aplicar esta regla, pero a no ser que se esté tratando con una extensa estructura de datos probablemente no se necesitará.

#### **6.1.2.1.3 Diccionario de datos**

El diccionario de datos se ha propuesto como una gramática casi formal para describir el contenido de los elementos de información, un listado organizado de todos los datos pertinentes al sistema con definiciones precisas y rigurosas para que tanto el usuario como el analista tengan un entendimiento común de todas las entradas, salidas, componentes de almacenes y cálculos intermedios.

El diccionario de datos define los datos haciendo lo siguiente:

- Describe el significado de los flujos y almacenes que se muestran en el DFD (Diagrama de Flujo de Datos)
- Describe la composición de agregados de paquetes de datos que se mueven a lo largo de los flujos, es decir, paquetes complejos que pueden descomponerse en unidades más elementales.
- Describe la composición de los paquetes de datos de los almacenes

Especifica los valores y unidades relevantes de piezas elementales de información en los flujos de datos y los almacenes de datos

Razones para su utilización:

Para manejar los detalles en sistemas muy grandes, ya que tienen enormes cantidades de datos, aun en los sistemas más chicos hay gran cantidad de datos.

1. Para asignarle un solo significado a cada uno de los elementos y actividades del sistema.

2. Para documentar las características del sistema, incluyendo partes o componentes así como los aspectos que los distinguen. También es necesario saber bajo que circunstancias se lleva a cabo cada proceso y con que frecuencia ocurren. Produciendo una comprensión más completa. Una vez que las características están articuladas y registradas, todos los participantes en el proyecto tendrán una fuente común de información con respecto al sistema.
3. Para facilitar el análisis de los detalles con la finalidad de evaluar las características y determinar donde efectuar cambios en el sistema.

## 6.2 Servidores WEB

Un **servidor Web** es un programa que implementa el protocolo HTTP (*hypertext transfer protocol*). Este protocolo está diseñado para transferir lo que llamamos hipertextos, páginas web o páginas HTML (*hypertext markup language*): textos complejos con enlaces, figuras, formularios, botones y objetos incrustados como animaciones o reproductores de sonidos.

Sin embargo, el hecho de que HTTP y HTML estén íntimamente ligados no debe dar lugar a confundir ambos términos. HTML es un formato de archivo y HTTP es un protocolo.

Un servidor web se encarga de mantenerse a la espera de peticiones HTTP llevadas a cabo por un cliente HTTP que solemos conocer como navegador. El navegador realiza una petición al servidor y éste le responde con el contenido que el cliente solicita. El servidor responde al cliente enviando el código HTML de la página; el cliente, una vez recibido el código, lo interpreta y lo muestra en pantalla. Como vemos con este ejemplo, el cliente es el encargado de interpretar el código HTML, es decir, de mostrar las fuentes, los colores y la disposición de los textos y objetos de la página; el servidor tan sólo se limita a transferir el código de la página sin llevar a cabo ninguna interpretación de la misma.

Sobre el servicio web *clásico* podemos disponer de aplicaciones web. Éstas son fragmentos de código que se ejecutan cuando se realizan ciertas peticiones o respuestas HTTP. Hay que distinguir entre:

- **Aplicaciones de cliente:** el cliente web es el encargado de ejecutarlas en la máquina del usuario. Son las aplicaciones tipo Java o Javascript: el servidor proporciona el código de las aplicaciones al cliente y éste lo ejecuta.

- **Aplicaciones de servidor:** el servidor web ejecuta la aplicación; ésta, una vez ejecutada, genera cierto código HTML; el servidor toma este código recién creado y lo envía al cliente por medio del protocolo HTTP. Esta estructura suele ser mejor opción a la hora de crear una aplicación web.

### 6.2.1 WEB

Las **páginas web** son los documentos con los que se construye la WWW, y su número crece exponencialmente. Una página web es un documento multimedia, ya que puede contener información de distintos tipos de medios: textos, imágenes, sonidos, animaciones, etc. Y a ello se une la posibilidad de crear enlaces con otras páginas que pueden estar ubicadas en lugares remotos, éste tipo de páginas se convierten entonces en documentos hipermedia.

Las páginas web son diseñadas para ser accedidas a través de Internet, por lo cual serán visualizadas por distintos navegadores en máquinas con características muy diversas. Por ello resulta conveniente comprobar su correcta visualización al menos en las configuraciones más habituales. Cuando esto no sea posible se deberá indicar en la página la configuración más adecuada para visualizarla correctamente.

Por sus características, las páginas web representan un medio excelente para la publicación de información multimedia en Internet. Sin embargo, el uso de las páginas web no está restringido a la red. Las páginas se pueden visualizar localmente sin necesidad de conexión a la red. Para ello basta con disponer de un navegador, y como éste es ya un componente fijo en cualquier máquina se pueden utilizar las páginas web para distribuir cualquier tipo de información multimedia.

En una página web pueden colocarse distintos tipos de elementos. Los más habituales son los siguientes:

- **Texto**
- **Imágenes**
- **Objetos multimedia**
- **Enlaces**

El texto puede aparecer en forma de párrafo o de lista. Las imágenes complementan al texto proporcionando información visual. También pueden utilizarse para hacer la página más atractiva. En cuanto a los objetos multimedia, los más utilizados son sonidos, vídeos y animaciones. Estos objetos permiten crear efectos muy atractivos.

Los enlaces pueden estar asociados a cualquiera de los elementos anteriores y son los responsables de la auténtica potencia de las páginas web. Los enlaces permiten saltar de una página a otra sin importar el lugar donde ésta se encuentre. De esta forma podemos consultar un gran número de documentos relacionados con un mismo tema y que pueden estar dispersos por todo el planeta.

Todos los elementos mencionados deben estar colocados en la página de forma que la información que transmiten sea lo más clara posible. Para ello nos podemos servir de otros elementos como son las tablas y los *frames* o marcos. Estos elementos los podemos considerar como contenedores, pues contienen otros elementos, y nos ayudan a conseguir el formato de visualización que deseamos para ellos.

Las páginas Web se dividen en dos tipos, las páginas estáticas y las páginas dinámicas de las que hablaremos a continuación.

### 6.2.1.1 Páginas Estáticas

Inicialmente, las páginas Web eran estáticas, en el sentido de que, a efectos de usuario, el único proceso realizado era el de visualización de su contenido (escrito en lenguaje HTML) por parte del explorador del cliente.

Las páginas estáticas se siguen utilizando ampliamente debido a que forman la base necesaria para la presentación de datos en muchos tipos de situaciones. También influye decisivamente la sencillez con que se pueden crear, instalar y mantener.

La figura 6.6 representa esquemáticamente la forma en que interactúan las páginas estáticas con el Cliente y el Servidor.

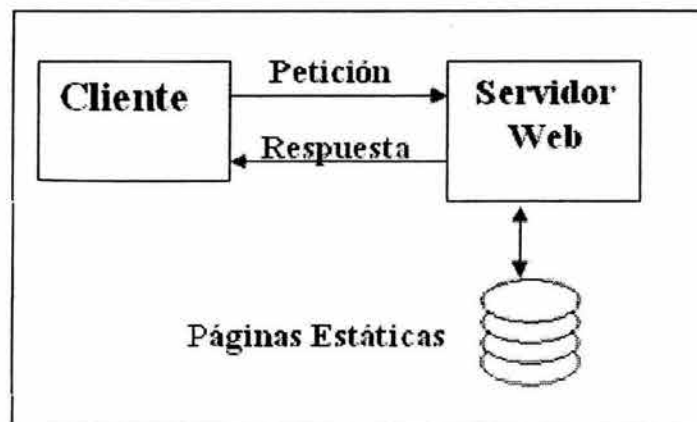


Figura 6.6. Páginas Estáticas

### 6.2.1.2 Páginas Dinámicas

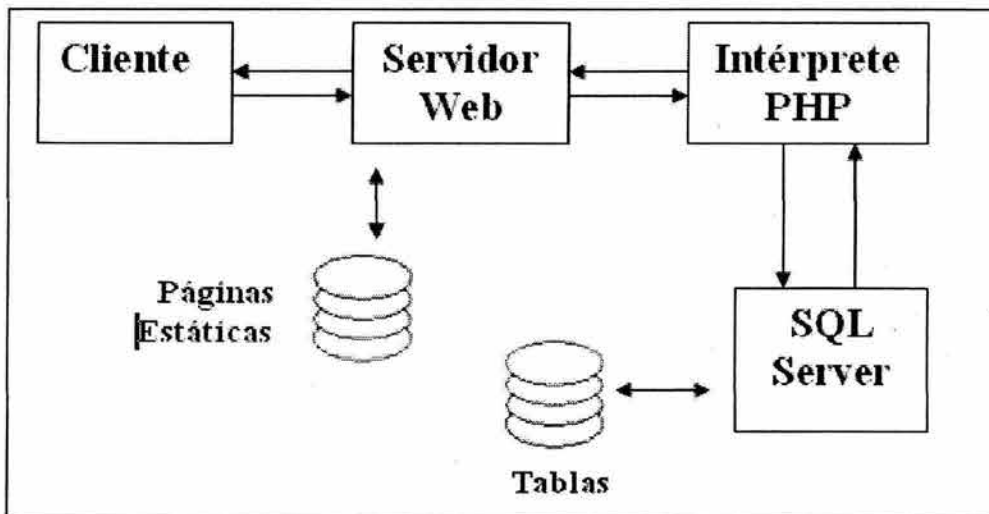
En el momento en que se requiere una interacción mayor entre los usuarios y el sistema que soporta las páginas Web, surge la necesidad de reunir y procesar las peticiones del cliente con el fin de ofrecerle informaciones mejor dirigidas, escogidas y elaboradas. El procesamiento de intercambio de información entre los usuarios y el servidor de páginas Web, introdujo el concepto de páginas Web activas o dinámicas.

La capacidad de procesamiento que sustenta las páginas dinámicas se puede llevar a cabo siguiendo alguno de estos modelos:

- Procesamiento en el equipo del usuario: páginas activas en el cliente.
- Procesamiento en el equipo donde reside el servidor de Web: páginas activas en el servidor.
- Procesamiento mixto: páginas activas en el cliente + páginas activas en el servidor.

La figura 6.7 muestra el intercambio de información entre los usuarios y el servidor de páginas Web.

Una ventaja de las páginas activas es el ahorro de comunicaciones (ancho de banda<sup>82</sup>) que se puede experimentar en muchas aplicaciones que ejecutan procesos en el cliente, y evitan de esta manera realizar continuos trasposos de información con el servidor de Web.



Pág. 6.7 Páginas Dinámicas

<sup>82</sup> Ancho de Banda: Velocidad a la que puede transmitir información un canal de comunicación, con independencia del soporte físico que se utiliza



### 6.2.2 Arquitectura Cliente-Servidor

La arquitectura Cliente-Servidor es utilizada por todas las aplicaciones de Internet y consta de los siguientes elementos:

- Un **cliente**, el cual funciona en la computadora del usuario, se comunica con el servidor remoto (computadora remota), y solicita a éste información.
- El **servidor**, que es una computadora remota tiene como tarea proporcionar la información solicitada por el cliente.
- Un **formato de comunicación**, que establece la manera en que el cliente y el servidor se comunican.

Los sistemas Cliente-Servidor pueden ser de muchos tipos, dependiendo de las aplicaciones que el servidor pone a disposición de los clientes. Entre ellos, existen:

- Servidores de Impresión, mediante el cual los usuarios comparten impresoras.
- Servidores de Archivos, con el cual los clientes comparten discos duros
- Servidores de Bases de Datos, donde existe una única base de datos
- Servidores de páginas Web, mediante los cuales podemos publicar información en Internet y realizar consultas de la misma.

A continuación profundizaremos en las características de los servidores de páginas Web.

### 6.2.3 Servidores de páginas WEB

Los **Servidores Web** son aquellos que permiten a los clientes compartir datos, documentos y multimedia en formato HTML. Aunque se basa en la tecnología Cliente-Servidor, el servidor Web aporta ventajas adicionales en aspectos muy importantes como son:

- El Web se crea normalmente como un sistema abierto al que cualquiera puede contribuir y acceder desde cualquier punto de la red de la empresa; normalmente no requieren claves de acceso.
- Los Servidores Web regresan información con un simple clic del ratón a través de un proceso de hipervínculo.

- La información entregada puede ser de cualquier tipo (datos, documentos, multimedia, etc...), gracias a la utilización de los estándares de Internet. Esa información es únicamente de lectura pues el usuario no puede hacer cambios en el documento original de los datos.
- Dado que el servidor de Web es de fácil acceso, ello hace posible publicar información de forma instantánea en toda la organización mediante su almacenamiento en el servidor.
- Un servidor Web en la Intranet<sup>83</sup> puede servir la misma copia de un archivo o documento a toda la organización, de la misma forma que un único servidor web en Internet puede servir al mundo entero
- La amplitud de la red suele ser mayor que otros sistemas Cliente- Servidor, pues la mayor parte de Intranet se construye sobre redes de Área Extensa (Wide Area Networks, WAN por sus siglas en inglés), más que sobre redes locales (Local Area Networks, LAN por sus siglas en inglés).

#### 6.2.4 Lenguajes de programación para la WEB

Existen diferentes lenguajes de programación que nos permiten crear de forma sencilla páginas estáticas y dinámicas en la Web, así como desarrollar aplicaciones para interactuar fácilmente con el usuario.

Entre los principales lenguajes de programación que se utilizan para el desarrollo del sistema se tienen:

##### HTML

HTML es un lenguaje de formateo de textos que se utiliza para la creación de páginas en la WWW. Por página entenderemos el documento que aparece en el visualizador. HTML se compone de una serie de comandos, que son interpretados por el visualizador, o programa que utilizamos para navegar por el WWW. En última instancia es el visualizador el que

---

<sup>82</sup> *Intranet*: Una Intranet es una red privada empresarial o educativa que utiliza los protocolos TCP/IP de Internet para su transporte básico. Los protocolos pueden ejecutar una variedad de Hardware de red, y también, pueden coexistir con otros protocolos de red, como IPX. Aquellos empleados que están dentro de una Intranet pueden acceder a los amplios recursos de Internet, pero aquellos en Internet no pueden entrar en la Intranet, que tiene acceso restringido.

ejecuta todas las órdenes contenidas en el código HTML, de forma que un visualizador puede estar capacitado para unas prestaciones, pero no para otras. Así, podremos especificar que una página tenga una imagen de fondo, o un texto parpadeando, pero si nuestro visualizador no está capacitado para esas funciones, no podremos comprobarlas.

## **6.3 JAVA**

Java es el lenguaje de programación que seleccionamos para desarrollar la aplicación WEB debido a que ofrece algunas ventajas con respecto de otros lenguajes, a continuación se verán algunas de esas características.

### **6.3.1 ¿Qué es JAVA?**

Java es un lenguaje de programación de alto nivel; es decir, un conjunto de símbolos, palabras y palabras reservadas que se reúnen con el fin de escribir una serie de líneas de código que puedan ser entendidas por una computadora.

Es un lenguaje de programación orientado a objetos, desarrollado por Sun en 1991 como parte de un proyecto de investigación para crear software para los dispositivos electrónicos. El lenguaje Java se diseñó para ser pequeño, sencillo y portátil a través de plataformas y sistemas operativos, tanto a nivel código fuente como en binario.

#### **6.3.1.1 Características**

Las características principales que nos ofrece Java respecto a cualquier otro lenguaje de programación, son:

##### **6.3.1.1.1 Simple**

Java es simple en su funcionalidad pero muy potente. Ha adoptado de lenguajes como C y C++ las características que hacen grandes a estos lenguajes, a nivel de seguridad y diseño.

Una de las características que elimina la necesidad de liberar manualmente el espacio de memoria es suplida por una gestión automática de memoria donde un recolector de basura (*garbage collector*) es el encargado de liberarla cuando lo cree conveniente.

### 6.3.1.1.2 Dinámico

Esta característica hace referencia a la habilidad con la que Java puede incorporar, por ejemplo, librerías nuevas o actualizadas sin parar las aplicaciones actuales. También se simplifica el uso de protocolos de red nuevos o actualizados; de esta manera si se ejecuta una aplicación Java sobre la red, y se encuentra una pieza de aplicación que no sabe manejar, cargará de forma automática a través de la red, la pieza que el sistema necesita para que pueda funcionar.

Si Java descubre que ciertos objetos o clases son demandados con frecuencia, evita que sean eliminados por el reciclador de memoria y si una aplicación quiere hacer uso de una pieza que no está cargada, ésta puede ser añadida dinámicamente a través de la red.

En la figura 6.8 se describe de forma esquemática esa petición, y cómo, automáticamente, se tramita ésta para buscar información sobre el objeto desconocido.

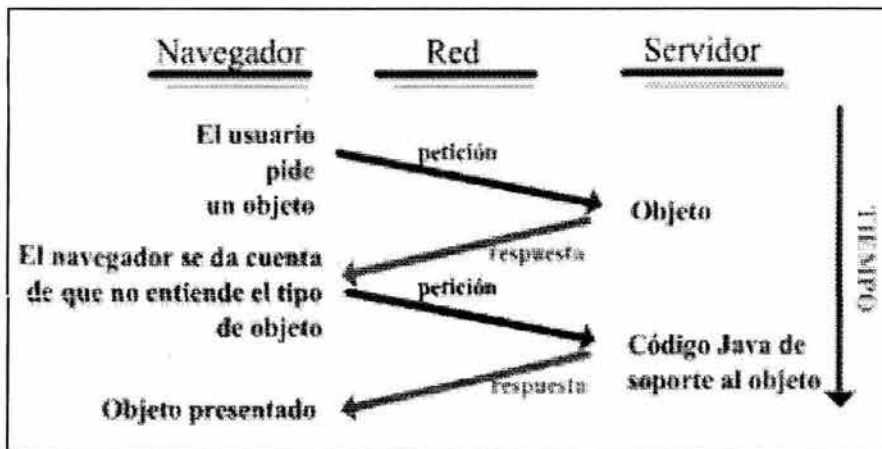


Figura 6.8 Peticiones del usuario a través de la red.

### 6.3.1.1.3 Distribuido

Java se ha construido con extensas capacidades de interconexión TCP/IP<sup>84</sup>. Existen librerías de rutinas para acceder e interactuar con protocolos como HTTP y FTP<sup>85</sup>. Esto permite a los programadores acceder a la información a través de la red con tanta facilidad como a los directorios y archivos locales.

Java proporciona una colección de clases para su uso en aplicaciones de red, que permiten abrir *sockets*<sup>86</sup> y establecer y aceptar conexiones con servidores o clientes remotos, facilitando así la creación de aplicaciones distribuidas.

### 6.3.1.1.4 Robusto

Java realiza verificaciones en busca de problemas tanto en tiempo de compilación como en tiempo de ejecución. La comprobación de tipos en Java ayuda a detectar errores, lo antes posible, en el ciclo de desarrollo. Java obliga a la declaración explícita de métodos, reduciendo así las posibilidades de error. Maneja la memoria para eliminar las preocupaciones por parte del programador de la liberación o corrupción de memoria. También implementa los *arrays* auténticos, en vez de listas enlazadas de punteros, con comprobación de límites, para evitar la posibilidad de sobrescribir o corromper memoria resultado de punteros que señalan a zonas equivocadas. Estas características reducen drásticamente el tiempo de desarrollo de aplicaciones en Java.

Además, para asegurar el funcionamiento de la aplicación, realiza una verificación de los *bytecodes*, que son el resultado de la compilación de un programa Java. Es un código de máquina virtual que es interpretado por el intérprete de Java. No es el código máquina directamente entendible por el hardware, pero ya ha pasado todas las fases del compilador: análisis de instrucciones, orden de operadores, etc., y ya tiene generada la pila de ejecución de órdenes.

Java proporciona, pues: Comprobación de apuntadores, comprobación de límites de arrays, excepciones, verificación de byte-codes

---

<sup>84</sup> *TCP/IP* Transport Control Protocol/Internet Protocol/ Internet Protocol protocolo que define cómo debe ser la comunicación entre computadoras sobre Internet.

<sup>85</sup> *FTP* (File Transfer Protocol) protocolo para la transferencia de Hipertexto y Archivos.

<sup>86</sup> Un *socket* es simplemente una conexión a un número específico de puerto del servidor.

### 6.3.1.1.5 De Arquitectura neutral

Para que pueda ser portable, Java se ha diseñado con una arquitectura que no dependiera de ninguna plataforma, y a la vez, que pudiera ejecutar su código en el mayor número de ellas. A su vez enlaza otra característica, Java es interpretado y por tanto se deberán diseñar sistemas interpretados para cada una de ellas.

La arquitectura de Java es independiente y esto provoca que el código también lo sea. Cuando se compila un programa Java, se genera un bytecode, que es un código de alto nivel independiente de la máquina, pero que para ejecutarse debe ser interpretado por un sistema *runtime* que sí es dependiente de la plataforma.

En la figura 6.9 se observan los elementos que forman parte integrante de la arquitectura de Java.

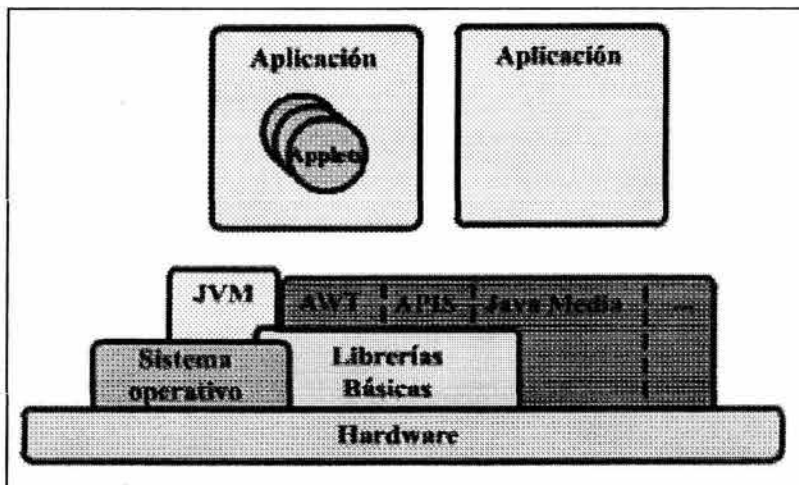


Figura 6.9 Partes principales de la plataforma Java.

Las partes principales de la plataforma Java son: la máquina virtual de Java (MVJ) y la interfaz de programación para aplicaciones Java (API)

### 6.3.1.1.6 Seguro

El sistema de Java tiene ciertas políticas que evitan que se puedan codificar virus con este lenguaje, además existen muchas restricciones, especialmente para los *Applets* (programas en Java que se cargan en un explorador), que limitan lo que se puede o no se puede hacer con los recursos críticos de una computadora.

La seguridad de Java incorpora 4 niveles que sirven para controlar, entre otras, la posibilidad de ejecutar software de manera distribuida entre diferentes máquinas, y de cargar Applets procedentes de Internet. Los niveles de seguridad son los siguientes:

- Nivel de lenguaje
- Nivel de verificación de los bytecode
- Nivel de cargador de clases
- Nivel de API de Java

### **Seguridad para internet**

La relación más conocida de Java con Internet es la que mantienen a través de los Applets. La posibilidad de cargar un Applet, desde cualquier punto de la red, es una puerta abierta para virus y programas de resultados dudosos.

Para evitar sorpresas, todos los exploradores actuales, a partir de las versiones de 1998, incorporan una serie de restricciones:

- Los Applets no pueden ejecutar programas sobre el equipo local
- Tampoco pueden leer o escribir sobre este equipo
- Sólo se podrán conectar al servidor que los contenía antes, es decir, desde donde fueron bajados o descargados.

#### **6.3.1.1.7 Portable**

Esta característica es distintiva de los grandes lenguajes de programación. La posibilidad como tal se puede definir como la posibilidad que tiene el software desarrollado por un lenguaje de programación sobre una plataforma concreta, de poder ser utilizado sobre otra plataforma completamente distinta sin tener que realizar ninguna modificación sobre el código fuente.

Gracias a que se trata de un lenguaje compilado-interpretado, a su arquitectura independiente, y a su definición de tipos de datos, se trata de un lenguaje portable.

En la figura 6.10 se observa un esquema sobre el contenido de este punto.

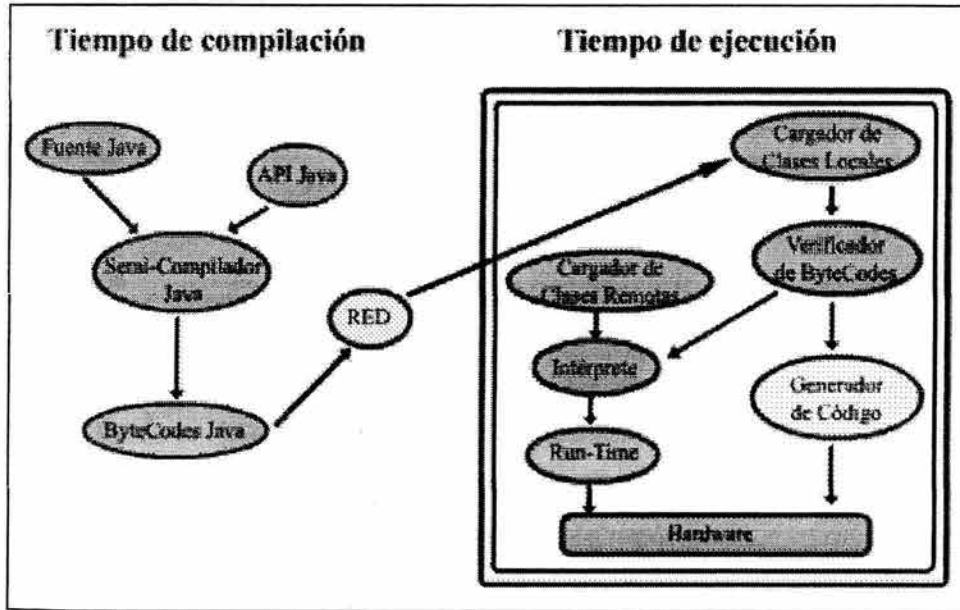


Figura 6.10 Portabilidad de Java

### 6.3.1.1.8 Interpretado

La idea de que Java fuese interpretado, es decir, que no cree un ejecutable del código fuente una vez compilado, si no un código objeto (código intermedio) que el intérprete de Java pueda ejecutar directamente, se pensó para consumir menos recursos y, de esta manera, dedicar mas tiempo al desarrollo que al proceso de creación de un ejecutable.

Pero además, para que sea realmente un lenguaje totalmente independiente de la maquina y del sistema operativo, debe ser compilado e interpretado.

El código intermedio que se genera al ser compilado, los *bytecodes* que vimos en puntos anteriores, es código maquina de muy bajo nivel, y corresponde al 80% de las instrucciones del programa.

Sin embargo, ese código es genérico y, para adaptarlo a una plataforma en concreto, debemos interpretarlo para añadir el 20% que falta para su ejecución.

Evidentemente, esto conlleva dos problemas a tener en cuenta. El primero hace que Java sea entre un 10% y un 20% más lento respecto de otros lenguajes de programación orientados a objetos.

Y el segundo de ellos, que deberá existir un *runtime* para cada una de las plataformas mas habituales a fin de conseguir diseñar programas multi-plataformas (portables).



### 6.3.1.1.9 Máquina virtual

Dentro de la arquitectura de Java, la máquina virtual (MV) es un elemento dependiente de la plataforma sobre la cuál se desarrollan las aplicaciones. Ella es la encargada de interpretar los bytecode al lenguaje de máquina, para adaptar la ejecución de la aplicación al hardware de esa plataforma.

En la figura 6.11 se observa dónde queda situada esta máquina dentro del proceso de compilación e interpretación del código.

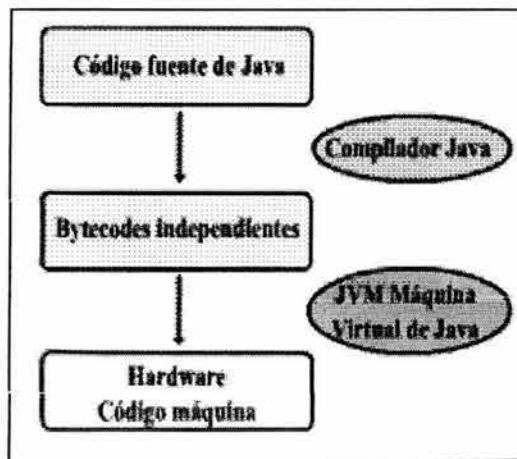


Figura 6.11 Máquina virtual en el proceso de compilación

La principal ventaja es que dota a Java de la característica de la portabilidad (a nivel de código fuente y bytecode, ya que en el proceso de compilado es independiente de la plataforma).

También proporciona seguridad a la hora de ejecutar Applets, donde la máquina virtual se presenta en forma de *plug-in* dentro del explorador.

Y además genera resultados con el tamaño adecuado para la implementación de sistemas distribuidos.

### 6.3.1.1.10 Multithreaded

La característica de multithread permite a Java ejecutar actividades simultáneas dentro de la aplicación. El principal beneficio que aporta esto es un mejor rendimiento interactivo y comportamiento en tiempo real.

Mientras un hilo se encarga de la comunicación, otro puede interactuar con el usuario mientras otro presenta una animación en pantalla y otro realiza cálculos.

### 6.3.2 Programación Orientada a Objetos

Java incorpora en su arquitectura las tres características propias de la filosofía para la programación orientada a objetos: **herencia, encapsulación y polimorfismo**.

La programación orientada a objetos se basa en la idea natural de la existencia de un mundo lleno de objetos, con características que los diferencian, y con un conjunto de acciones propias que pueden realizarse sobre ellos.

Un lenguaje está basado en objetos si soporta a éstos como característica fundamental del mismo.

Los objetos o instancias podrán tener características comunes si son de un determinado tipo o clase (la clase define las características o atributos para sus objetos), y a su vez, una clase podrá heredar características de otra si existe una relación de familia.

Java está orientado a objetos, tanto a nivel de arquitectura por incorporar las características propias de esta filosofía, como a nivel funcional, porque trabaja con sus datos dándoles un tratamiento de objeto en todo momento.

#### 6.3.2.1 Mecanismos de la POO

Los mecanismos básicos de la programación orientada a objetos son: **objetos, mensajes, métodos y clases**.

##### 6.3.2.1.1 Objetos

Un programa orientado a objetos se compone solamente de objetos, entendiéndose por objeto una encapsulación genérica de datos y de los métodos para manipularlos. Dicho de otra forma, un objeto es una entidad que tiene unos atributos particulares, las propiedades, y unas formas de operar sobre ellos, los métodos.

### 6.3.2.1.2 Métodos

Los métodos son funciones que pueden ser llamadas dentro de la clase o por otras clases.

Un método se implementa en una clase de objetos y determina cómo tiene que actuar el objeto cuando recibe el mensaje vinculado con ese método. A su vez, un método puede también enviar mensajes a otros objetos solicitando una acción o información.

En adición, las propiedades (atributos) definidas en la clase permitirán almacenar información para dicho objeto.

Cuando se diseña una clase de objetos, la estructura más interna del objeto se oculta a los usuarios que lo vayan a utilizar, manteniendo como única conexión con el exterior, los mensajes. Esto es, los datos que están dentro de un objeto solamente podrán ser manipulados por los métodos asociados al propio objeto.

Según lo expuesto, podemos decir que la ejecución de un programa orientado a objetos realiza fundamentalmente tres cosas:

1. Crea los objetos necesarios.
2. Los mensajes enviados a unos y a otros objetos puedan viajar, dando lugar a que se procese internamente la información.
3. Finalmente, cuando los objetos no son necesarios, son borrados, liberándose la memoria ocupada por los mismos.

### 6.3.2.1.3 Mensajes

Cuando se ejecuta un programa orientado a objetos, los objetos están recibiendo, interpretando y respondiendo a mensajes de otros objetos. Esto marca una clara diferencia con respecto a los elementos de datos pasivos de los sistemas tradicionales.

En la POO un mensaje está asociado con un método, de tal forma que cuando un objeto recibe un mensaje la respuesta a ese mensaje es ejecutar el método asociado.

#### 6.3.2.1.4 Clases

Una clase es un tipo de objetos definido por el usuario. Una clase equivale a la generalización de un tipo específico de objetos. Por ejemplo, piense en un molde para hacer flanes; el molde es la clase y los flanes los objetos.

Cuando escribe un programa utilizando un lenguaje orientado a objetos, no se definen objetos verdaderos, se definen clases de objetos, donde una clase se ve como una plantilla para múltiples objetos con características similares.

Un programa consta de una o más clases, éstas pueden organizarse en paquetes. Las clases de Java definen objetos de software encerrando los miembros de los datos (campos) y a los miembros de función (métodos). Los miembros pueden designarse como privados, protegidos de paquete o públicos, lo que proporciona una manera conveniente de definir la interfaz pública y el dominio privado de un objeto.

#### 6.3.2.2 Características de la POO

Las características fundamentales de la POO son: **abstracción, encapsulamiento, herencia y polimorfismo.**

##### 6.3.2.2.1 Abstracción

Por medio de la abstracción conseguimos no detenernos en los detalles concretos de las cosas que no interesen en cada momento, sino generalizar y centrarse en los aspectos que permitan tener una visión global del problema. Por ejemplo, el estudio de una computadora podemos realizarlo a nivel de funcionamiento de sus circuitos electrónicos, en términos de corriente, tensión, etc., o a nivel de transferencia entre registros, centrándose así el estudio en el flujo de información entre las unidades que lo componen (memoria, unidad aritmética, unidad de control, registros, etc.), sin importarnos el comportamiento de los circuitos electrónicos que componen estas unidades.

### 6.3.2.2 Encapsulamiento

Esta característica permite ver un objeto como una caja negra en la que se ha introducido de alguna manera toda la información relacionada con dicho objeto. Esto nos permitirá manipular los objetos como unidades básicas, permaneciendo oculta su estructura interna.

La abstracción y la encapsulación están representadas por la clase. La clase es una abstracción, porque en ella se definen las propiedades o atributos de un determinado conjunto de objetos con características comunes, y es una encapsulación porque constituye una caja negra que encierra tanto los datos que almacena cada objeto como los métodos que permiten manipularlos.

### 6.3.2.3 Herencia

La herencia es el mecanismo mediante el cual un objeto adquiere (o hereda) las propiedades de otro. De esta forma se consigue la clasificación jerárquica. Si no se hiciera una clasificación jerárquica de los objetos, cada objeto debería definir todas sus características explícitamente, y esto no sería viable.

Sin embargo, utilizando la herencia, un objeto puede heredar sus atributos generales de otro objeto (su padre), y definir explícitamente sólo aquellas cualidades que lo hacen único dentro de su clase. Por tanto, la herencia es el mecanismo que permite a un objeto ser una instancia específica de un caso más general.

Una clase que hereda de otra, se denomina *subclase*, y aquella clase de la que se hereda, se denomina *superclase*.

Una subclase hereda todos los atributos de cada uno de sus antecesores en la jerarquía de clases. Esta característica permite a los programas orientados a objetos crecer en complejidad de manera lineal en vez de geométrica.

### 6.3.2.4 Polimorfismo

Esta característica permite implementar múltiples formas de un mismo método, dependiendo cada una de ellas de la clase sobre la que se realice la implementación.

Esto hace que se pueda acceder a una variedad de métodos distintos (todos con el mismo nombre) utilizando exactamente el mismo medio de acceso.

### 6.3.2.3 Constructores y destructores

Un **constructor** es un procedimiento especial de una clase que es llamado automáticamente siempre que se crea un objeto de esa clase. Su función es iniciar el objeto.

Un **destructor** es un procedimiento especial de una clase que es llamado automáticamente siempre que se destruye un objeto de esa clase. Su función es realizar cualquier tarea final en el momento de destruir el objeto.

### 6.3.3 Paquetes

El lenguaje Java proporciona una serie de paquetes que incluyen ventanas, utilidades, un sistema de entrada/salida general, herramientas y comunicaciones. En la versión 1.1.8 el JDK, los paquetes Java que se incluyen son:

#### java.applet

Este paquete contiene clases diseñadas para usar con Applets. Hay una clase Applet y tres interfaces: AppletContext, AppletStub y AudioClip.

#### java.awt

El paquete Abstract Windowing Toolkit (awt) contiene clases para generar widgets y componentes GUI (Interfaz Gráfico de Usuario). Incluye las clases Button, Checkbox, Choice, Component, Graphics, Menu, Panel, TextArea y TextField.

#### java.io

El paquete de entrada/salida contiene las clases de acceso a ficheros: FileInputStream y FileOutputStream.

#### java.lang

Este paquete incluye las clases del lenguaje Java propiamente dicho: Object, Thread, Exception, System, Integer, Float, Math, String, etc.

### java.net

Este paquete da soporte a las conexiones del protocolo TCP/IP y, además, incluye las clases Socket, URL y URLConnection.

### java.util

Este paquete es una miscelánea de clases útiles para muchas cosas en programación. Se incluyen, entre otras, Date (fecha), Dictionary (diccionario), Random (números aleatorios) y Stack (pila FIFO).

## 6.3.4 Interfaces

Para diseñar interfaces gráficas (ventanas con componentes, llamados también controles, como etiquetas, cajas de texto, botones, barras de desplazamiento, etc.) Java proporciona una biblioteca de clases denominada JFC (*Java Foundation Classes* - clases base de Java). Actualmente bajo esta denominación se agrupan las siguientes APIs:

- *Swing*. Conjunto de componentes escritos en Java para diseñar interfaces gráficas de usuario que se ejecutan uniformemente en cualquier plataforma nativa que soporta la máquina virtual de Java.
- *AWT* (*Abstract Window Toolkit* - kit de herramientas de ventanas abstractas). Grupo de componentes para diseñar interfaces gráficas de usuario común a todas las plataformas nativas. Este grupo ha sido sustituido en gran medida por el conjunto de componentes *Swing*; muchos de estos heredan de sus correspondientes componentes *AWT*.
- *Accesibilidad*. Ofrece soporte para usuarios discapacitados.
- *Java 2D*. Permite incorporar en los programas, gráficos 2D de alta calidad, texto e imágenes.
- Soporte para *arrastrar y colocar* (*drag and drop*). Permite la transferencia de datos entre aplicaciones mediante la simple operación de arrastrarlos hasta el lugar de destino.

### 6.3.5 Excepciones

A diferencia de otros lenguajes de programación orientados a objetos, Java incorpora en el propio lenguaje el manejo de errores. El mejor momento para detectar los errores es durante la compilación. Sin embargo prácticamente sólo los errores de sintaxis son detectados durante este periodo.

El resto de problemas surgen durante la ejecución de los programas. En el lenguaje Java, una **Exception** es un cierto tipo de error o una condición anormal que se ha producido durante la ejecución de un programa. Algunas excepciones son fatales y provocan que se deba finalizar la ejecución del programa. En este caso conviene terminar ordenadamente y dar un mensaje explicando el tipo de error que se ha producido. Otras, como por ejemplo no encontrar un archivo en el que hay que leer o escribir algo, pueden ser recuperables. En este caso el programa debe dar al usuario la oportunidad de corregir el error (indicando una nueva localización del archivo no encontrado).

### 6.3.6 Applets

Todo lo estudiado hasta este momento relativo a aplicaciones Java y a paginas HTML es aplicable al desarrollo de Applets, ya que un Applet, desde el punto de vista de código Java, lo podemos definir como un pequeño programa diseñado para ser ejecutado en el contexto de una pagina Web y desde el punto de vista de una pagina HTML, representa una pequeña ventana grafica dentro de la misma.

Por otra parte, la diferencia entre una aplicación Java y un Applet radica en cómo se ejecutan. Sabemos que para cargar y ejecutar una aplicación Java utilizamos directamente el interprete Java. En cambio, un Applet se carga y se ejecuta desde cualquier explorador que soporte Java.



## 6.4 Conclusiones

En este capítulo se vieron todos los conceptos teóricos que se aplicarán en el siguiente capítulo tales como base de datos, manejador de bases de datos SQL Server, así como también servidores Web, definiciones fundamentales de Java el cual será el programa de desarrollo para el sistema.

Todo lo visto en este capítulo servirá como base para desarrollar el sistema Diseño de un Diccionario Onomasiológico para la Web como se presenta en el siguiente capítulo.

# CAPÍTULO 7

## ANÁLISIS, DISEÑO Y DESARROLLO DEL SISTEMA



En el capítulo anterior se mostraron conceptos fundamentales de las herramientas que utilizaremos para el desarrollo del sistema, conceptos como lenguaje de programación Java, bases de datos y servidores Web.

Cualquier sistema implementado por el hombre, pasa por diferentes fases. El ciclo de vida de un sistema empieza en un grupo de planeación, al principio como una idea conceptual muy general, que posteriormente toma una forma en un grupo de diseño.

Si se considera factible un proyecto desde diversos puntos de vista, se desarrolla e implementa para ponerse en marcha y culminar con su operación, es decir, el sistema satisface una necesidad.

En el presente capítulo tiene como finalidad explorar el enfoque de la metodología del ciclo de la vida, abordando sus características, así como también se analizan y desarrollan los pasos que se siguieron en esta faceta de análisis y diseño del sistema del diccionario onomasiológico para una mejor concepción de su desarrollo.

La complejidad tecnológica de los sistemas que actualmente implementa el hombre, es tan grande, que es necesario planear con todo cuidado estos proyectos, analizar diversas alternativas, tomando en cuenta todos los aspectos de un proyecto. En resumen, se requiere una secuencia lógica de pasos detallados de análisis, para resolver estos problemas.

## **7.1 Metodología para el desarrollo de sistemas**

El método más conveniente a utilizar para la implementación del Diccionario Onomasiológico para la Web es el del ciclo de vida para el desarrollo de sistemas como se presenta a continuación.

### **7.1.1 Método del ciclo de la vida**

El método de ciclo de vida para el desarrollo de sistemas es el conjunto de actividades que los analistas, diseñadores y usuarios realizan para desarrollar e implantar un sistema de información.

Una idea básica del ciclo de vida del desarrollo de sistemas es que en un proceso bien definido en el cual se percibe una aplicación, se desarrolla y se realiza. El ciclo de vida le da estructura a un proceso creativo. Con el fin de manejar y controlar el esfuerzo de desarrollo, se hace necesario conocer qué se debería haber hecho, que se ha hecho y aún qué falta por cumplir. Las fases en el ciclo de vida proveen una base para la administración y el control, en razón de que:

1. definen los segmentos del flujo de trabajo que se pueden identificar para propósitos administrativos y
2. especificar los documentos u otros resultados intermedios que van a ser producidos en cada fase.

Las fases en el ciclo de vida para el desarrollo de sistemas de información son descritos por varios autores, pero las diferencias están principalmente en la cantidad de detalle y en la forma de categorización. Hay un acuerdo general en el flujo de los pasos de desarrollo y la necesidad de procedimientos de control en cada etapa.

El método de ciclo de vida para el desarrollo de sistemas consta de las siguientes actividades:

1. Análisis del sistema
2. Determinación de los requerimientos
3. Diseño del sistema
4. Desarrollo de software
5. Prueba del sistema
6. Implantación y evaluación

### 7.1.1.1 Análisis del Sistema

Para esta etapa se analiza la información, basándonos en un proceso sistemático para razonar un problema dentro de una organización y encontrar una solución detallada. Además, el análisis permite establecer si la implantación de un sistema de información automatizado es posible.

El proceso de análisis consta de tres etapas fundamentales, que son:

- **Análisis de Objetivos:** En esta etapa se define el problema a solucionar, se entiende y razona para plantear una solución y se establecen los lineamientos para el desarrollo del sistema.
- **Análisis de Requerimientos:** Aquí se estudia con detalle la conducta que debe tener el sistema para ser útil a la organización, también se definen las restricciones del medio en el cual operará normalmente el sistema.
- **Análisis de Implantación:** Por último se determinan con precisión los costos y tiempos de desarrollo y se define el plan de acciones necesarias para diseñar, construir, probar e implantar el sistema.

En cada una de las etapas se define un conjunto de productos a obtener, y que están orientados a reflejar fielmente la operación esperada del sistema.

### 7.1.1.2 Determinación de los Requerimientos

Con el fin de obtener de manera efectiva un conjunto completo y correcto de los requerimientos, es necesario utilizar un método o métodos que tengan en cuenta hasta que punto los requerimientos son completamente conocidos o por el contrario, si se está necesitando ser explorados o descubiertos.

El resultado del análisis de los requerimientos de información del ciclo de vida de desarrollo es un reporte que detalla los requerimientos de la aplicación.

### 7.1.1.3 Diseño del sistema

El diseño de un sistema produce los detalles que establecen la forma en la que el sistema cumplirá con los requerimientos identificados durante la fase de análisis.

Se comienza el proceso de diseño identificando los reportes y demás salidas que debe producir el sistema. Hecho lo anterior, se determinan con toda precisión los datos específicos para cada reporte y salida, es común que los diseñadores hagan un bosquejo del formato o pantalla que esperan que aparezcan cuando el sistema esté terminado.

El diseño de un sistema también indica los datos de entrada, aquellos que serán calculados y los que deben ser almacenados. Los procedimientos que se escriben indican cómo procesar los datos y producir las salidas.

#### **7.1.1.4 Desarrollo de software**

Los encargados de desarrollar software pueden instalar (o modificar y después instalar) software comprado a terceros o escribir programas diseñados a la medida del solicitante. La elección depende del costo de cada alternativa, del tiempo disponible para escribir el software y de la disponibilidad de los programadores. Por regla general, los programadores (o analistas programadores) que trabajan en las grandes instituciones pertenecen a un grupo permanente de profesionales.

Los programadores también son responsables de la documentación de los programas y de proporcionar una explicación de cómo y porqué de ciertos procedimientos se codifican en determinada forma. La documentación es esencial para probar el programa y llevar a cabo el mantenimiento una vez que la aplicación se encuentre instalada.

#### **7.1.1.5 Pruebas del sistema**

Durante la fase de pruebas del sistema, el sistema se emplea de manera experimental para asegurarse de que el sistema no tenga fallas, es decir, que funciona de acuerdo con las especificaciones y en la forma en que los usuarios esperan que lo haga. Se alimentan con un conjunto de datos para su procesamiento y después se examinan los resultados. En ocasiones se permite que varios usuarios utilicen el sistema para que los analistas observen si tratan de emplearlo en formas no previstas. Es preferible descubrir cualquier sorpresa antes de que el sistema sea implantado y que los usuarios dependan de él.

#### **7.1.1.6 Implantación y evaluación**

La implantación es el proceso de verificar e instalar la aplicación. Una vez instaladas, las aplicaciones se emplean durante muchos años. Sin embargo, los usuarios y el ambiente cambian con el paso del tiempo. Por consiguiente, es indudable que debe darse mantenimiento a las aplicaciones; realizar cambios y modificaciones en el sistema. Dado que la Web y los usuarios experimentan cambios de manera continua, los sistemas deben mantenerse siempre al día. En este sentido, la implantación es un proceso en constante evolución. La evaluación de un sistema se lleva a cabo para identificar puntos débiles y fuertes.

Una vez definidas las etapas del método del ciclo de vida de los sistemas a continuación se presenta de forma desglosada las etapas antes mencionadas.

## 7.2 Análisis del Sistema

El proyecto “Diseño de un diccionario onomasiológico para la Web”, pretende cubrir el problema de tener un sistema que permita encontrar el término de la idea o concepto que el usuario tiene en mente, pero que no recuerda, es decir, lo que se conoce como el problema de “tener la palabra en la punta de la lengua”.

Esta tesis permitirá a los usuarios introducir una idea o concepto en lenguaje natural dentro del área de contexto de desastres y física, con toda la información que el usuario desee poner.

Para resolver la parte del análisis del problema se presenta el Algoritmo de Búsqueda Onomasiológica<sup>86</sup> que describe la forma en que el sistema operará:

1) El usuario introduce una frase que a su juicio describa el término buscado. La frase introducida por el usuario es separada en palabras y cada una de ellas es comparada con el arreglo de la tabla de STOPLIST, si la palabra comparada esta dentro del arreglo de STOPLIST, entonces esa palabra se elimina. Las palabras restantes se consideran palabras clave de la búsqueda, almacenándose en un arreglo.

2) Las palabras clave deben someterse a un proceso de **lematización**. El proceso de lematización consiste en obtener automáticamente la raíz de una palabra, removiendo su información derivativa o flexiva (visto en el capítulo 5), a partir de las palabras clave lematizadas para cada una de ellas se efectúa un proceso de búsqueda en la tabla PALABRA, en caso de no existir se incluye dicha palabra en la tabla PALABRA2, en caso contrario se busca en la tabla AGRUPAMIENTO\_PALABRA los sinónimos asociados a las palabras, a este proceso se le denomina expansión de términos. En caso de no encontrar en la tabla AGRUPAMIENTO\_PALABRA la palabra clave buscada, ésta es desechada. Al conjunto de todos los sinónimos de todas las palabras clave se denomina lista expandida de palabras.

3) Con la lista expandida de palabras y para cada palabra en la lista, se busca en la tabla PCLAVE\_TERMINO qué términos contienen dicha palabra. Se debe llevar una lista de cuántos términos se han encontrado y cuantas palabras clave los han seleccionado. A esta lista se le denomina lista de términos y ocurrencias.

4) Con base en la lista de términos y ocurrencias, se presentan los tres términos con mayor ocurrencia y se ofrece la alternativa de mostrar todos los términos identificados. Además deberá permitirse que el usuario consulte la definición del término.

---

<sup>86</sup> G. Sierra (1996), *Bases para la búsqueda onomasiológica de términos* (Tesis maestría), Universidad Nacional Autónoma de México. Facultad de filosofía y letras. Capítulo 2

Para ejemplificar lo anteriormente descrito, tenemos el siguiente ejemplo de operación:

**Ejemplo de operación.**

Considere que la base de datos contiene la siguiente información:

PALABRA			PCLAVE TERMINO	
IdPalabra	Palabra	LemaPalabra	IdTermino	IdPalabra
1	caer	cae	12	2
2	agua	ag	15	2
3	fuerte	fuerte	15	56
4	cielo	ciel	15	178

**AGRUPAMIENTO PALABRA**

IdAgrupamiento	IdPalabra
1	7
1	3
1	5
1	27
1	35
1	47
1	13
1	21
3	45
3	101
3	44
3	83
3	11
3	23
3	38
3	189
3	245
3	237
3	88
3	150
3	123
3	56
3	78
3	17
4	178
4	222

**DEFINICION**

IdTérmino	Término	Definición
12	Lluvia	Precipitación de gotas de agua líquida o sobreenfriada que produce consecuencias desastrosas cuando son de gran intensidad.
15	Tormenta de granizo	Es agua congelada que desciende con violencia de las nubes en granos mas o menos duros y gruesos.

Paso 1.

El usuario introduce la siguiente oración: “es cuando cae agua fuertemente del cielo y con hielo”

Paso 2.

Las palabras que pertenecen a la tabla de STOPLIST son: **es**, **cuando**, **del** e **y**, una vez identificadas se eliminan.

En un arreglo se almacenan las palabras clave de búsqueda: **cae, agua, fuertemente, cielo.**

Paso 3.

Las palabras contenidas en el arreglo de palabras clave, pasan por un proceso de lematización, una vez lematizadas se buscan en la tabla PALABRA (en caso de no existir la palabra se agrega a la tabla PALABRA2), como todas las palabras anteriores existen, nos vamos a la tabla de AGRUPAMIENTO\_PALABRA en donde podemos encontrar las relaciones de sinónimos y sus agrupamientos semánticos. La lista expandida de palabras sería:

Palabras lematizadas	Sinónimos
cae	Decaer, bajar, descender, perder el equilibrio, dar de bruces, irse de cabeza, dar en tierra, faltarle a uno el suelo.
ag	
fuerte	Duro, sólido, resistente, robusto, vigoroso, recio, forzado, sano, entero, firme, enérgico, impetuoso, violento, esforzado, animoso, varonil.
ciel	Atmósfera, firmamento.

Paso 4.

Con la lista de palabras expandidas se busca en la tabla PCLAVE\_TERMINO las palabras clave que están relacionadas con algún término y la tabla de término y ocurrencias quedaría de la siguiente manera:

Término	Ocurrencias
12	1
15	3

Paso 5.

En la tabla anterior, vemos que el término que tiene mayor número de ocurrencias es el 15 y es él que será presentado al usuario en primer lugar y que correspondería a “Tormenta de granizo” y en segundo lugar se le presentaría el término “Lluvia”.

### 7.3 Determinación de los Requerimientos

Cabe mencionar que las personas que conforman el grupo de ingeniería lingüística no pertenecen a la misma institución educativa, y tampoco cuentan con el mismo nivel de conocimientos en cuanto a Tecnologías e Internet, por lo que es necesario considerar que el sistema sea sencillo y tenga una interfaz amigable y simple, ya que a través de él se planea no sólo consultar la información contenida en la base de datos, sino que también se requiere que los participantes del proyecto puedan realizar las manipulaciones de la información necesarias.



Así también implantar el algoritmo que permita realizar búsquedas onomasiológicas de forma que los usuarios puedan introducir la frase o concepto en lenguaje natural y a través de una búsqueda interna del sistema se obtenga como resultado los términos.

Al usuario se le presentan dos opciones de búsqueda: la búsqueda onomasiológica y la búsqueda semasiológica, la primera definida anteriormente y la segunda la cual consiste en que una vez obtenidos los términos, se podrá consultar su definición, con esto, el usuario tendrá dos formas de utilizar el diccionario.

El desarrollo de un diccionario electrónico permitirá tener acceso al sistema a través del Web. La plataforma de trabajo que se eligió para el desarrollo del sistema por requisitos institucionales será Java.

### 7.3.1 Hardware

El hardware es el equipo que se necesita para que el sistema funcione óptimamente, para evaluar esto, es necesario tener la clara concepción de qué es lo que el usuario espera del sistema.

De acuerdo a lo anterior el servidor es una computadora de gran potencia, que se encarga de "prestar un servicio" a otras computadoras que se conectan a él. En general, se ocupan de gestionar todo el tráfico de información que produce en la red.

El servidor que se utilizará para el sistema se llama Tabasco y tiene las siguientes especificaciones:

EQUIPO	CARACTERÍSTICAS
Marca	Acer
Procesador	Intel Pentium II 350 MHz
Memoria RAM	384 MB
Disco duro	74.5 GB

Las características del equipo antes mencionado garantizan la efectividad y buen desempeño del sistema.

### 7.3.2 Software

El software se refiere a los programas de aplicación que se pueden utilizar para desarrollar un sistema, éstos se clasifican de acuerdo a: efectividad en el desempeño, eficiencia, facilidad de uso, portabilidad, calidad de la documentación, ayuda en línea, costo, disponibilidad, recursos de hardware, ámbito de trabajo y soporte.

La propuesta de software considera: la plataforma a utilizar, el lenguaje de programación, el manejador de bases de datos y el servidor Web necesario para visualizar el sistema desde Internet.

Considerando las necesidades del cliente, explicadas en la etapa de análisis, se decidió utilizar:

- El sistema operativo Windows como plataforma de desarrollo. De esta forma se contará con un sistema con capacidad de escalabilidad al servidor con las características antes mencionadas.
- Por otro lado, se buscó un lenguaje de programación que pudiera cumplir con las especificaciones requeridas. Por requisitos institucionales se eligió a Java como lenguaje de programación para el desarrollo del sistema
- En cuanto al manejador de bases de datos se utilizará SQL Server,
- Y como servidor Web se utilizará *Internet Information Server* (IIS). IIS engloba un conjunto de herramientas destinadas al control de servicios de Internet como el Web, FTP, correo y servidores de noticias. Además incluye el soporte necesario para la creación de páginas dinámicas en el servidor mediante el lenguaje ASP<sup>87</sup>.

Con base en estas herramientas se diseñó el sistema que se presenta en esta tesis.

## 7.4 Diseño del Sistema

A este momento se han definido de manera general los objetivos y características generales del sistema.

El diseño de la base de datos que se realizó se muestra a continuación en el diagrama entidad-relación. Con el diseño propuesto el usuario podrá hacer consultas de forma efectiva.

### 7.4.1 Diagrama Entidad-Relación

El modelo de la base de datos es el siguiente:

---

<sup>87</sup> ASP (Active Server Pages): Mediante ASP se pueden crear y ejecutar aplicaciones sobre la Web.

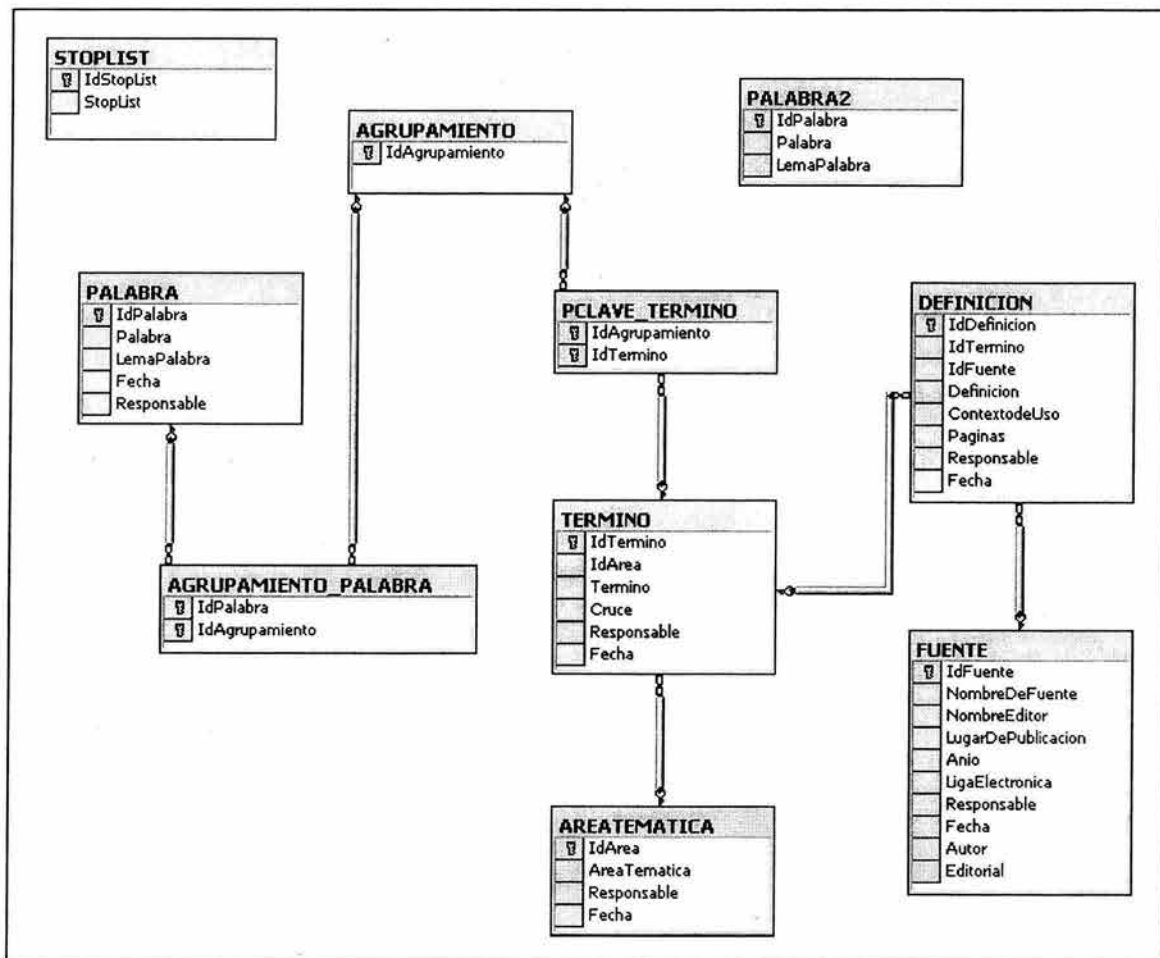


Fig. 7.1 Modelo Entidad-Relación

## 7.4.2 Tablas del Sistema

Las tablas que se definieron durante el proceso de diseño son las siguientes:

TABLA	DESCRIPCIÓN
STOPLIST	Palabras que resultan redundantes para la búsqueda.
PALABRA	Palabras que se encuentran dentro del contexto de acuerdo a la base de datos asignada.
AGRUPAMIENTO_PALABRA	Asociaciones de la palabra que tecleó el usuario con las agrupaciones respectivas a dicha palabra.
AGRUPAMIENTO	Agrupamientos de palabras.
PCLAVE_TERMINO	Agrupamientos asociado al término
TÉRMINO	Términos relacionados dentro del contexto de la base de datos seleccionada.
DEFINICIÓN	Definiciones que corresponden al contexto de la base de datos elegida.
FUENTE	Información bibliográfica de las definiciones correspondientes.
AREATEMATICA	Información del contexto en que se encuentra la búsqueda.
PALABRA2	Información nueva que no existe en la tabla PALABRA y que se va incluyendo conforme se solicita una búsqueda.

### 7.4.3 Diccionario de Datos

El diccionario de datos correspondiente a la base antes diseñada, se muestra a continuación, se incluyen las validaciones pertinentes para los campos que lo requieran.

El diccionario muestra:

- Nombre del campo
- Descripción del campo
- Tipo de dato
- Longitud

#### TABLA STOPLIST

CAMPO	DESCRIPCIÓN	TIPO	LONGITUD
IdStopList	Identificador del StopList	INT, not null	4
StopList	Palabras redundantes para la búsqueda	TEXT, null	16

#### INDICES

CAMPO	TIPO
IdStopList	Llave primaria

#### TABLA PALABRA

CAMPO	DESCRIPCIÓN	TIPO	LONGITUD
IdPalabra	Identificador de la Palabra	INT, not null	4
Palabra	Palabras que se comparan con las introducidas por el usuario	VARCHAR, not null	50
LemaPalabra	Palabras lematizadas	TEXT, null	16
Fecha	Fechas de actualizaciones	CHAR, null	10
Responsable	Responsable de proyecto	TEXT, null	16

#### INDICES

CAMPO	TIPO
IdPalabra	Llave primaria

#### TABLA AGRUPAMIENTO\_PALABRA

CAMPO	DESCRIPCIÓN	TIPO	LONGITUD
IdPalabra	Identificador de la Palabra	INT, not null	4
IdAgrupamiento	Identificador del Agrupamiento	INT, not null	4

**INDICES**

CAMPO	TIPO
IdPalabra + IdAgrupamiento	Llave primaria

**TABLA AGRUPAMIENTO**

CAMPO	DESCRIPCIÓN	TIPO	LONGITUD
IdAgrupamiento	Identificador del Agrupamiento	INT, not null	4

**INDICES**

CAMPO	TIPO
IdAgrupamiento	Llave primaria

**TABLA PCLAVE\_TERMINO**

CAMPO	DESCRIPCIÓN	TIPO	LONGITUD
IdAgrupamiento	Identificador del Agrupamiento	INT, not null	4
IdTermino	Identificador del Término	INT, not null	4

**INDICES**

CAMPO	TIPO
IdAgrupamiento + IdTermino	Llave primaria

**TABLA TÉRMINO**

CAMPO	DESCRIPCIÓN	TIPO	LONGITUD
IdTermino	Identificador del Termino	INT, not null	4
Diarrea	Identificador del Área Temática	INT, not null	4
Termino	Termino que corresponde a la agrupación seleccionada	VARCHAR, not null	50
Cruce		TEXT, null	16
Responsable	Fechas de actualizaciones	TEXT, null	16
Fecha	Responsable de proyecto	DATETIME, null	8

**INDICES**

CAMPO	TIPO
IdTermino	Llave primaria

### TABLA DEFINICIÓN

CAMPO	DESCRIPCIÓN	TIPO	LONGITUD
IdDefinición	Identificador de la Definición	INT, not null	4
IdTermino	Identificador del Término	INT, not null	4
IdFuente	Identificador de la Fuente	INT, not null	4
Definición	Definición que corresponde al término seleccionado	TEXT, not null	16
ContextodeUso		TEXT, null	16
Paginas	Número de páginas de la información consultada	INT, null	4
Responsable	Responsable de proyecto	TEXT, null	16
Fecha	Fechas de actualizaciones	INT, null	4

### INDICES

CAMPO	TIPO
Indefinición	Llave primaria

### TABLA FUENTE

CAMPO	DESCRIPCIÓN	TIPO	LONGITUD
IdFuente	Identificador de la Fuente	INT, not null	4
NombreDeFuente	Nombre de fuente de la información consultada	TEXT, not null	16
NombreEditor	Nombre de editor de la información consultada	TEXT, null	16
LugarDePublicacion	Lugar de publicación de la información consultada	TEXT, null	16
Anio	Año de publicación de la información consultada	INT, null	4
LigaElectronica	Liga electrónica de la información consultada	TEXT, null	16
Responsable	Responsable de proyecto	TEXT, null	16
Fecha	Fechas de actualizaciones	INT, null	4
Autor	Nombre del autor de la edición	TEXT, null	16
Editorial	Nombre de la editorial de la edición	TEXT, null	16

### INDICES

CAMPO	TIPO
IdFuente	Llave primaria

**TABLA AREATEMATICA**

CAMPO	DESCRIPCIÓN	TIPO	LONGITUD
IdArea	Identificador del Area Temática	INT, not null	4
AreaTematica		TEXT, not null	16
Responsable	Responsable de proyecto	TEXT, null	16
Fecha	Fechas de actualizaciones	INT, null	4

**INDICES**

CAMPO	TIPO
Diarrea	Llave primaria

**TABLA PALABRA2**

CAMPO	DESCRIPCIÓN	TIPO	LONGITUD
IdPalabra	Identificador de la Palabra 2	INT, not null	4
Palabra	Palabras que no se encuentran en la tabla PALABRA	VARCHAR, null	50
Lema Palabra	Palabras lematizadas	TEXT, null	16

**INDICES**

CAMPO	TIPO
IdPalabra	Llave primaria

**7.4.4 Diseño Conceptual y Físico de la Base de Datos**

En la figura 7.2 se puede observar el diagrama conceptual de la base de datos, en este diseño están relacionadas las tablas originales, las propuestas y las adicionales, así como aquellos atributos que representan las llaves primarias de cada tabla.

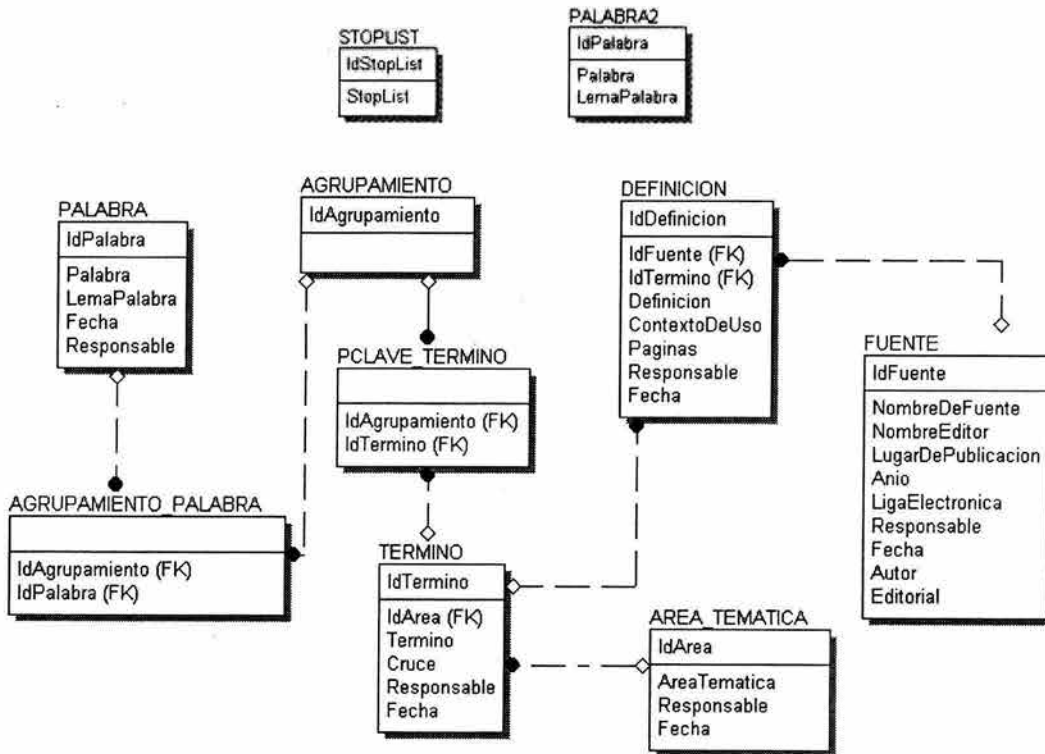


Fig. 7.2. Diagrama Conceptual de la base de datos

En la figura 7.3 se presenta el diseño físico de la base de datos, éste representa el diseño final a partir de cual se generará el *script* para la creación de la base.



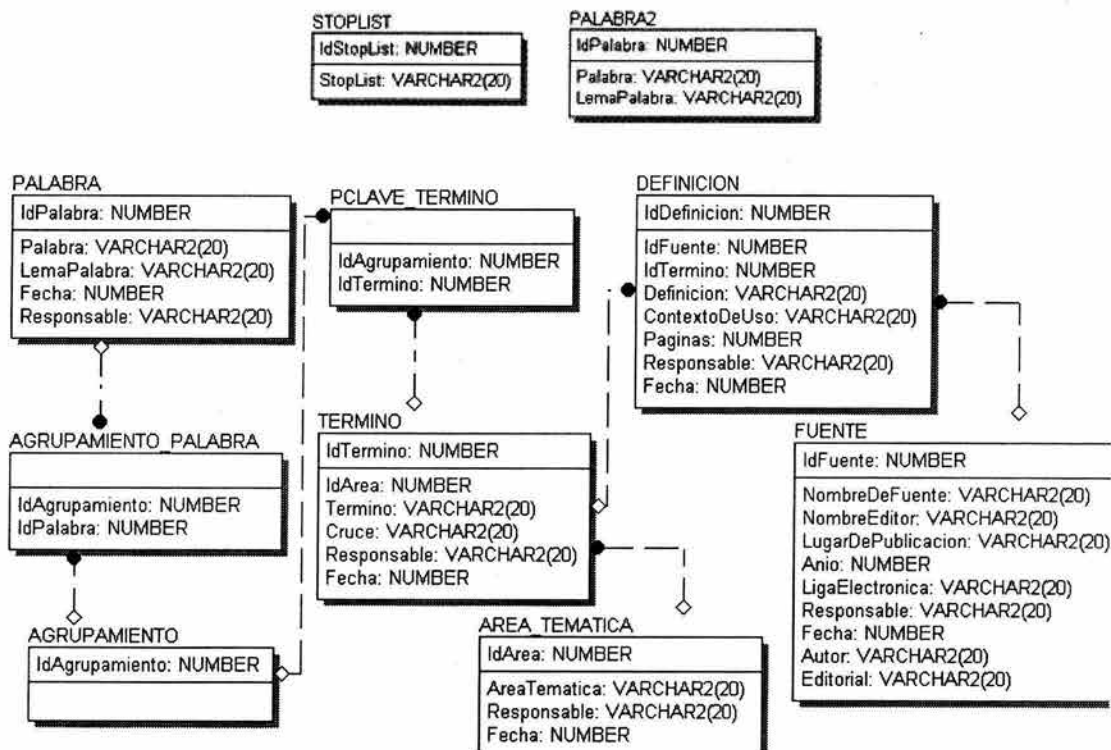


Fig. 7.3. Diagrama Físico de la base de datos

### 7.4.5 Diagrama jerárquico funcional

El sistema cuenta con una única pantalla principal en la que se realizan todas las búsquedas, en el segundo nivel esta elegir la base de datos, es decir, elegir el contexto al que se va a orientar la búsqueda, como tercer nivel esta el conectar la base de datos seleccionada, realizar la búsqueda interna y los resultados que arroja dicha búsqueda.

Dentro del registro Búsqueda interna se encuentra una serie de actividades que lleva a cabo el sistema internamente tales como el separa las palabras de la frase del usuario, hacer la búsquedas en la base de datos, lematizar las palabra. Al final de la búsqueda interna, los resultados que arroja el sistema son los términos y sus definiciones correspondientes a cada término.

En la figura 7.4 se presenta el diagrama jerárquico funcional correspondiente al sistema:

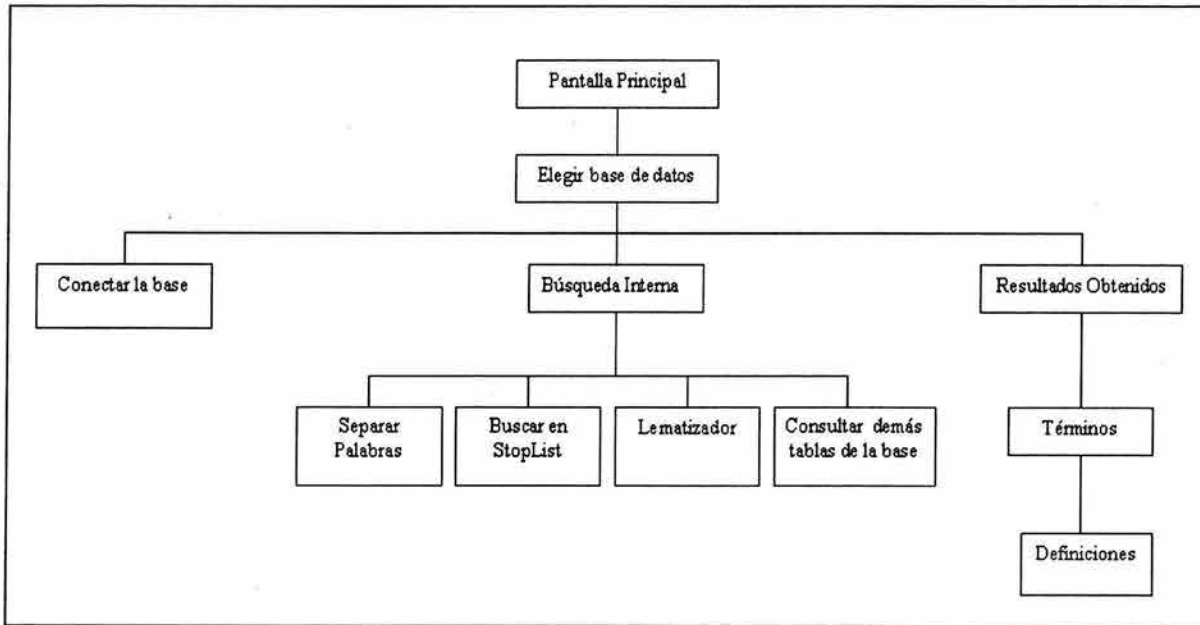


Fig. 7.4 Diagrama Jerárquico funcional de la base de datos

#### 7.4.6 Diseño de la pantalla principal

A continuación se presentan cuatro lineamientos para el diseño de pantallas, que si bien son importantes, no son exhaustivos.

1. Mantener la pantalla sencilla: El primer lineamiento para un buen diseño de pantallas es mantener la sencillez de la pantalla. Las secciones que son útiles para simplificar la pantalla son: la parte superior de la pantalla contiene la sección del encabezado, parte de la cual se encuentra programada para indicar al usuario en donde se encuentra dentro de la aplicación, la siguiente sección es el cuerpo, el cual utiliza las convenciones de la lectura: de izquierda a derecha y de arriba para abajo, con el fin de que se facilite al usuario la captura de datos.

2. Mantener una presentación consistente: El segundo lineamiento para un buen diseño de pantalla es el mantenimiento de una imagen consistente. La consistencia de la pantalla también se mantiene, si la información se localiza en la misma área cada vez que se accesa una nueva pantalla. La información que tenga alguna relación lógica entre si, debe presentarse en forma agrupada.

3. Facilitar los movimientos del usuario entre pantallas: El tercer lineamiento para un buen diseño de pantalla es la factibilidad de desplazarse con facilidad en una pantalla. Un método comúnmente utilizado, es hacer sentir al usuario que se mueve físicamente hacia una nueva pantalla.

4. Crear pantallas atractivas: El cuarto lineamiento para el buen diseño de pantallas es el desarrollo de una pantalla atractiva para el usuario. Si el usuario se siente atraído por la pantalla, es muy probable que sea más productivo, requiera de menos supervisión y cometa menos errores. Algunos de los principios de diseño utilizados para las formas también tienen una aplicación y se observan ciertos principios estéticos, si bien en un contexto ligeramente diferente. Las pantallas deben atraer al usuario y mantener su atención. Esto se logra con el uso de espacios abiertos que rodeen los campos de captura de datos, de tal forma que la pantalla no se vea sobrecargada. Nunca se debe saturar una forma, así como nunca se debe saturar una pantalla. Siempre será mejor utilizar pantallas múltiples, que amontonar todo en una pantalla. Para la presente tesis se requiere de una sola pantalla que contenga todos los elementos necesarios para realizar las búsquedas. Al crear pantallas que de primera instancia son fáciles de recordar, atraerá tanto a los usuarios inexpertos como a los expertos.

Al plantear pantallas es necesario utilizar un flujo lógico. Organizar el material de la pantalla para tomar ventaja de la forma en que la gente trabaja y así facilitarles su ubicación en la pantalla. También es necesario distribuir de manera consistente la información en las subsecciones indicadas con anterioridad.

Se trata de una sola pantalla que será la principal en la cual se harán todas las consultas correspondientes. La definición básica de la pantalla es la siguiente:

- Etiqueta de título: Esta etiqueta servirá para darle título a la pantalla la cual lleva por título: “Diccionario Electrónico de Búsqueda Onomasiológica”, como se presenta en el esquema. El formato de esta etiqueta es el siguiente:

Fuente: *Sans Serif*

Tamaño de fuente: 18

Resaltado en negritas

Label

**“ Diccionario Electrónico de Búsqueda Onomasiológica “**

- Etiqueta de selección de base de datos: Esta etiqueta indicará el sitio donde se podrá seleccionar la base de datos deseada por el usuario. La etiqueta quedará como se muestra a continuación. El formato de esta etiqueta es el siguiente:

Fuente: *Dialog*

Tamaño de fuente: 14

Resaltado en negritas

*Label*

**Elija una base de Datos:**

- Área de selección de la base de datos: Se eligió el elemento *choice* para representar el nombre de las bases de datos, mediante una indicación en forma de flecha que despliega el menú del nombre de las bases contenidas. Esta representación se muestra a continuación:



- Etiqueta de definición: Esta etiqueta indica el área donde el usuario podrá introducir su definición en lenguaje natural, la forma de la etiqueta es como se presenta en el esquema, esta etiqueta tiene el siguiente formato:

Fuente: *SansSerif*

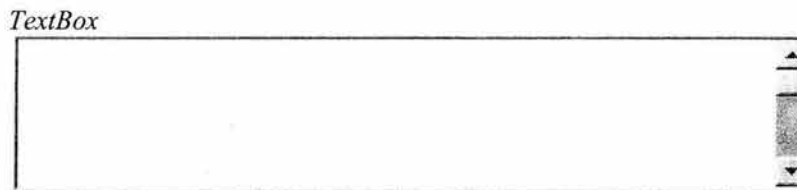
Tamaño de fuente: 13

Resaltado en negritas

*Label*

**Introduzca su definición:**

- Área de definición: El elemento que se utilizará es un *TextBox* como se presenta en el esquema siguiente, el usuario podrá insertar la definición en lenguaje natural.



- Botón Buscar: Una vez que el usuario inserte su definición, al dar clic en el botón buscar como lo muestra el esquema, comienza la búsqueda interna. El formato del botón es el siguiente:

Fuente: *Dialog*

Tamaño de fuente: 12

*Button*

**BUSCAR**

- Etiqueta Términos: Esta etiqueta indica el área donde el usuario podrá ver los resultados de la búsqueda. El formato de la etiqueta es el siguiente:

Fuente: *SansSerif*

Tamaño de fuente: 13

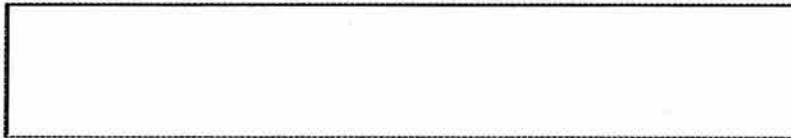
Resaltado en negritas

*Label*

**Términos:**

- Área de términos: El elemento que se utilizará para mostrar los términos al usuario es un *TextField*. En ésta área el usuario podrá consultar la lista de términos que el sistema encontró después de haber realizado una búsqueda.

*TextField*



- Etiqueta definición: Esta etiqueta indica el área donde el usuario podrá consultar las definiciones de los términos encontrados por el sistema. El formato de la etiqueta es el siguiente:

Fuente: *SansSerif*

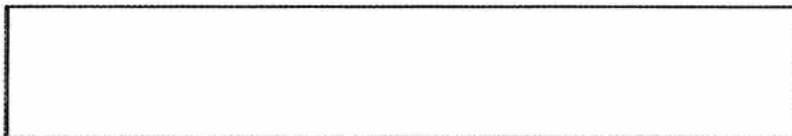
Tamaño de fuente: 13

*Label*

**Definición:**

- Área de definición: El elemento que se utilizará para mostrar las definiciones al usuario es un *TextField*. En ésta área el usuario podrá consultar las definiciones de los términos que él mismo seleccione.

*TextField*



- Botón Nueva Búsqueda: El elemento que se utilizará para hacer nuevas búsquedas es un botón como lo muestra el esquema. Este elemento permite al usuario realizar una nueva búsqueda, limpiando lo que se encuentre escrito en los demás elementos. El formato del botón es el siguiente:

Fuente: *Dialog*

Tamaño de fuente: 12

Button



## 7.5 Desarrollo del sistema

Para esta parte del sistema se desarrolló código en Java. Presentando una parte de código del desarrollo del sistema se tienen los siguientes ejemplos. Los dos métodos presentados a continuación forman parte de la clase BaseDeDatos y hacen la conexión y desconexión al servidor de Tabasco para poder consultar la pantalla, el código se ejecuta una vez que el usuario ingresa a la página.

```
/**
 * * Se realiza la conexión de la base
 *
 * @param base Nombre de la bases a desconectar
 * @throws ClassNotFoundException Muestra en pantalla el error que se genere al realizar la
desconexión de la base
 * @throws SQLException Si se genera algún error, se realiza la desconexión de la base
 */
```

```
public void conectarBase(String base) throws ClassNotFoundException, SQLException {
    System.out.println("Se inicia la conexion Proveedor");
    mvarProveedor = "com.microsoft.jdbc.sqlserver.SQLServerDriver";
    Class.forName(mvarProveedor);
    System.out.println("Se inicia la conexion Data Source");
    mvarDataSource =
"jdbc:microsoft:sqlserver://tabasco.torreingenieria.unam.mx:7598;DataBaseName="+base;
    conn=DriverManager.getConnection(mvarDataSource,"UsBD2003","ILing#20032004");
    System.out.println("conexion establecida con: "+base);
}
```

```
/**
 * * Se realiza la desconexión de la base
 *
 * @param base Nombre de la bases a desconectar
 * @throws ClassNotFoundException Muestra en pantalla el error que se genere al realizar la
desconexión de la base
 * @throws SQLException Si se genera algún error, se realiza la desconexión de la base
 */
```

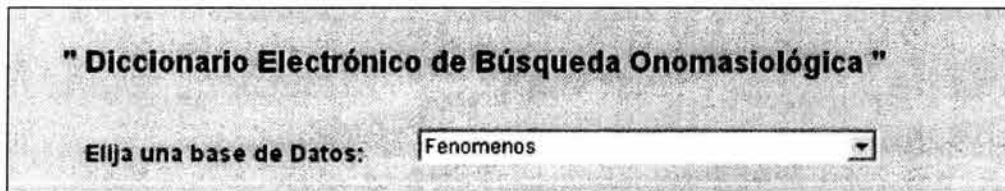
```
public void desconectarBase(String base) throws ClassNotFoundException, SQLException {
```

```

if (conn.isClosed()){
    System.out.println("Ya esta desconectada");
}
else{
    conn.close();
    System.out.println("Desconectado de: "+base);
}
}

```

Como siguiente ejemplo se presenta el siguiente código que realiza la conexión y desconexión de la base de datos que el usuario selecciona, ya sean estas las opciones de Fenómenos y Física, el código se ejecuta cuando el usuario desea seleccionar alguna de las bases de datos disponibles.



```

/**
 * * El método motor se conecta a la base de datos
 *
 * @param nombre Es el nombre de la base de datos a conectarse, desconectarse y al cargar
 palabras claves se define la base de datos de donde se obtendrán los datos *
 * @param lstb Combo que contiene la lista de nombres de las bases de datos
 */

public motor(String nbase) {
}

/**
 * * Se realiza la conexión a la base deseada y manda llamar al método cargarPalabrasClave
 *
 * @param nombre Nombre de la base de datos a consultar
 */

public void conectaBase(String nombre) {
    try {
        base = new BaseDeDatos();
        base.conectarBase(nombre);
// System.out.println("Llama a conectar Base");
    }
}

```

```
catch(SQLException ex){
    // mostraría una pantalla con el error generado...
    System.out.println(ex.toString());
}
catch(ClassNotFoundException ex){
    // mostraría una pantalla con el error generado...
    System.out.println(ex.toString());
}
}

/**
 * * Se realiza la desconexión de la base seleccionada
 *
 * @param nombre Nombre de la base de datos a desconectar
 */

public void desconectaBase(String nombre) {
    try {
        base = new BaseDeDatos();
        base.desconectarBase(nombre);
    }
    catch(SQLException ex){
        // mostraría una pantalla con el error generado...
        System.out.println(ex.toString());
    }
    catch(ClassNotFoundException ex){
        // mostraría una pantalla con el error generado...
        System.out.println(ex.toString());
    }
}
```

## 7.6 Migración de datos

La calidad de los datos es una medida de cómo la consistencia y corrección de los datos se encuentran dentro de ciertos límites preestablecidos. La codificación eficiente de los datos facilita la captura precisa al reducir el tiempo y al aumentar la cantidad de datos, y en consecuencia el tiempo requerido para la entrada.



Cuando los datos se introducen en forma eficiente, la captura satisface los parámetros establecidos de desempeño que se apegan a una relación entre el tiempo invertido en la captura y el número de datos capturados.

Es de fundamental importancia asegurarse de que los datos se capturen con precisión. El analista de sistemas puede apoyar la entrada precisa de los datos mediante el logro de tres objetivos generales, que son:

1. La codificación eficaz: Una de las maneras mediante las cuales pueden introducirse los datos, con más precisión y eficiencia es utilizando diversos códigos previamente establecidos. Al proceso de reexpresar datos ambiguos o complicados en dígitos o letras que se capturan con facilidad se le denomina codificación (no confundirlo con la codificación del programa). La codificación ayuda al analista de sistemas a alcanzar el objetivo de eficiencia, ya que los datos que se encuentran codificados, requieren de menor tiempo para capturarse y reducen el número de elementos procesados. Los datos codificados pueden reducir espacio valioso de almacenamiento y de memoria. En resumen, la codificación es una manera de ser elocuente, pero a la vez concisa, en la captura de los datos. Además de proporcionar precisión y eficiencia, los códigos deben contar con un propósito.
2. La captura y la entrada de datos eficaz y eficientes: Es importante capturar los datos con efectividad con el fin de asegurar la calidad de los datos que entran al sistema. La captura de datos ha recibido mayor atención por ser la etapa del proceso informático, de la cual se pueden obtener grandes logros en productividad.
3. El aseguramiento de la calidad a través de la validación: Hasta el momento se ha expuesto el aseguramiento de la captura de los datos sobre los documentos fuente y la entrada eficiente al sistema, a través de varios dispositivos de entrada. Aunque éstas son condiciones necesarias para asegurar la calidad de los datos, por sí solos no son suficientes. Los errores no pueden evitarse por completo y la importancia fundamental de detectar los errores durante la entrada, previa al procesamiento y almacenamiento, no puede dejar de destacarse. El aseguramiento de la calidad a través de la validación se verá en el capítulo 8.

Previamente se contaba con una base de datos y un banco terminológico ya definidos, en donde los datos fueron proporcionados por el Grupo de Ingeniería Lingüística, estos datos sirvieron para poblar la base de datos actual del sistema.

La figura 7.6 muestra la estructura de la base de datos original:

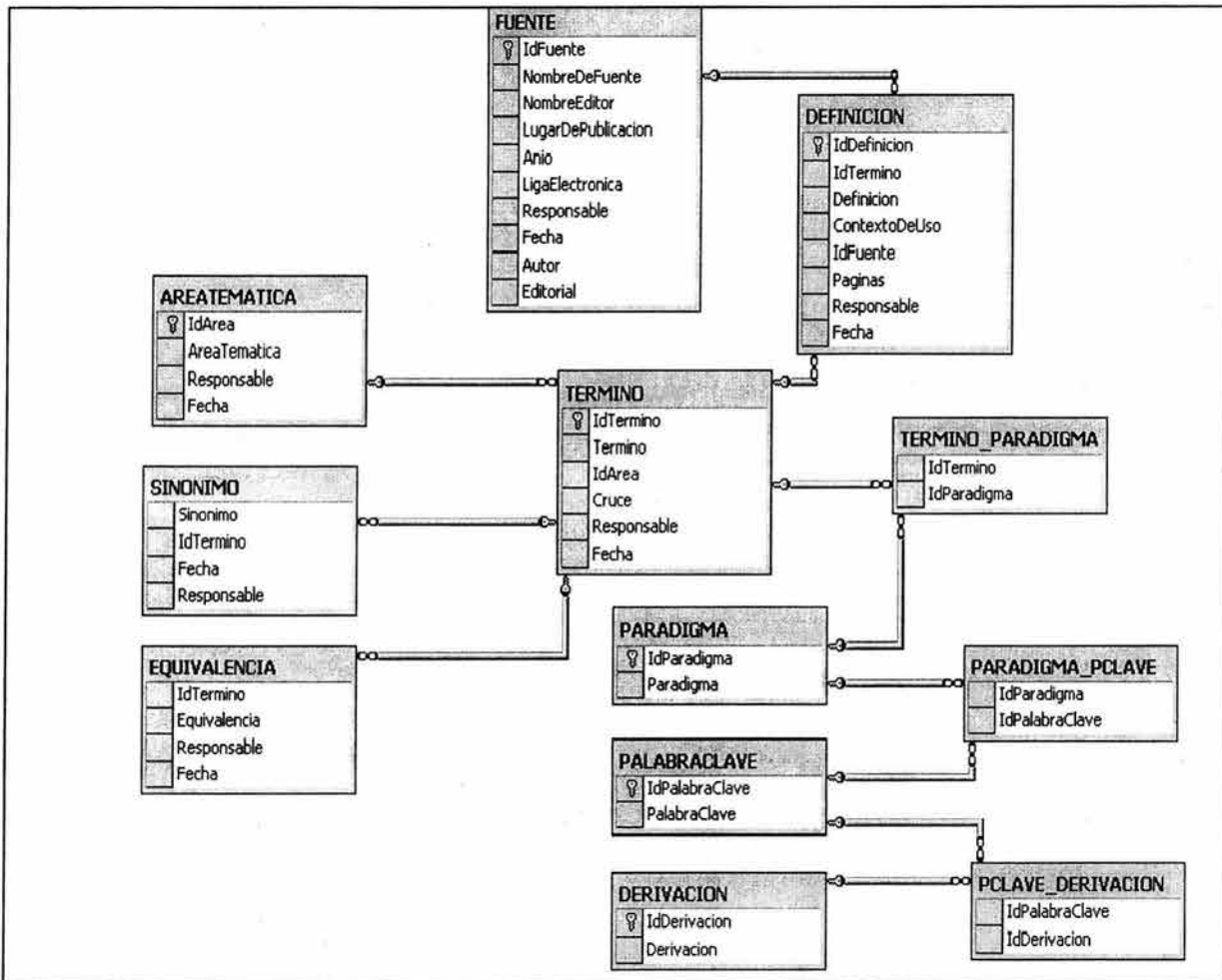


Fig. 7.6 Diagrama de base de datos original.

La estructura de la base de datos actual para el desarrollo del diccionario onomasiológico es el que se muestra en la figura 7.7.

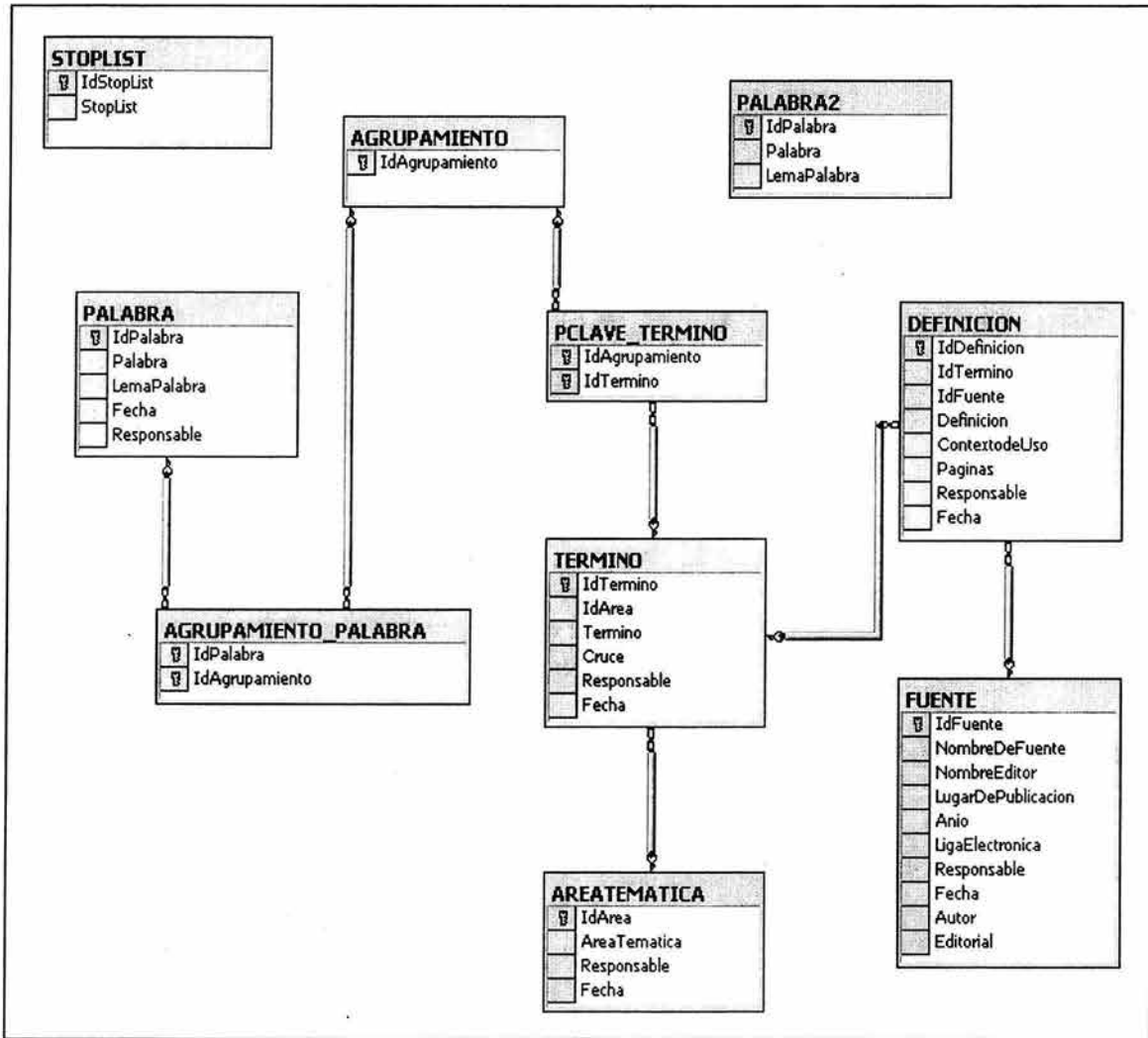


Fig. 7.7 Diagrama actual de la base de datos

El número de registros que se migraron para cada tabla es como se muestra a continuación:

Nombre de la tabla	Número de registros migrados
StopList	412
Palabra	50785
Palabra 2	Incrementa conforme a lo que inserta el usuario
Agrupamiento_Palabra	331275
Agrupamiento	331275
Pclave_Término	51832
Termino	32
Area Temática	1
Definición	32
Fuente	75

Esta migración de datos se hizo con la ayuda del manejador de bases de datos SQL Server. Se presentaron problemas al momento de ingresar los datos a las tablas correspondientes por la cantidad de datos que se manejaron. De los problemas más comunes fueron que se bloqueaba la máquina por la cantidad de datos que deseaban ingresar, al hacer la copia de la base de datos de Física se presentaba el mismo problema, es decir la cantidad de información era muy grande y esto ocasionaba el bloqueo de la máquina.

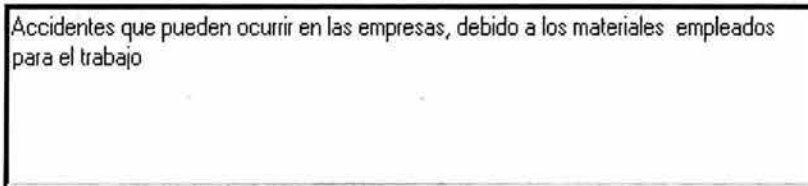
## 7.7 Interfaz WEB para el DO

La forma en que debe funcionar el Diccionario Onomasiológico para la Web conforme al algoritmo de búsqueda onomasiológica es como se presenta a continuación:

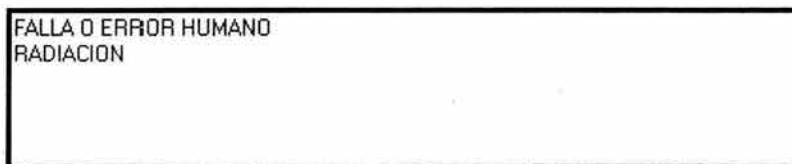
El usuario selecciona la base de datos dentro del contexto que hará la búsqueda, para ello, en el área de selección de base de datos, elige la base deseada, para este ejemplo se seleccionará la base de Fenómenos como se muestra en el esquema:



El usuario introduce la definición en lenguaje natural en el área de definición como se muestra en el siguiente esquema:



Una vez introducida la definición, el usuario da clic en el botón buscar para que se inicie la búsqueda interna, los resultados de la búsqueda son los términos que se muestran a continuación:



Como ya se había mencionado, como segunda opción se le presenta al usuario la oportunidad de consultar la definición de los términos que fueron encontrados durante la búsqueda, el usuario podrá seleccionar una opción para ver su definición, para este ejemplo, se seleccionará el primer elemento, es decir, el que tienen mayor número de ocurrencias como se muestra a continuación:

Es la falla ocasionada por el ser humano en forma involuntaria o por descuido, de la cual resultan malas operaciones en los servicios, accidentes que pueden causar heridas graves o pérdidas de la vida, errores de diseño, etc.

## 7.8 Conclusiones

Este capítulo tuvo como objetivo la realización del análisis, diseño y desarrollo del sistema, cada etapa que conformaron a este capítulo esta constituido por elementos que se complementaron para lograr dicho objetivo. Mediante la aplicación del método del ciclo de vida se logró mostrar los pasos a seguir para el desarrollo del sistema.

Gracias al presente capítulo se permite tener un panorama general de la forma en que el usuario participará con el sistema una vez implantado como se manejará en el siguiente capítulo.

# CAPÍTULO 8

## PRUEBAS, IMPLANTACIÓN Y EVALUACIÓN DEL SISTEMA



En el capítulo siete se analizaron los requerimientos básicos para llevar a cabo la implantación del sistema, en este capítulo se verá cómo se lleva a cabo dicha implantación. Además se hablará acerca de cuáles son las pruebas realizadas para comprobar que el sistema cumple con las especificaciones en cuanto a los requerimientos iniciales, dichas pruebas permitirán también hacer una evaluación de los resultados obtenidos.

En esta unidad se realizarán las pruebas a cada uno de los componentes del sistema de información. Una vez codificados, se busca comprobar que su estructura es correcta y que se ajustan a la funcionalidad establecida.

De igual manera, se comprobará la disponibilidad de los recursos humanos y técnicos necesarios para realizar las pruebas de implantación.

De ser necesario, se creará algún caso de prueba adicional que se considere importante y que no se haya tenido en cuenta hasta entonces. Se prepararán las condiciones que permitan simular las situaciones límite previstas para las pruebas.

## **8.1 Pruebas del sistema**

Las pruebas del sistema fueron realizadas por los desarrolladores y por los usuarios, a continuación se describe cómo se llevaron a cabo.

### **8.1.1 Pruebas unitarias**

Esta prueba es el nivel básico en donde se prueban las funciones que componen un módulo para garantizar que operan de manera correcta. Además se verifica que los datos en cada uno de los procedimientos fluyan de manera correcta y que conserven su valor al dirigirse hacia otros procedimientos.

En este sistema se utilizó el tipo de prueba unitaria de integración incremental que consiste en agregar cada módulo componente individual a un conjunto de componentes existentes y el conjunto resultante de prueba.

Generalmente la prueba de unidad se considera como algo adyacente a la codificación, por ello, estos casos de prueba nos permitieron descubrir errores como comparaciones entre tipos de datos distintos, variables o comparadores incorrectos, terminación de bucles inapropiada o inexistente y variables de bucles modificadas de forma inapropiada.

### **8.1.2 Pruebas de integración**

En estas pruebas se observa la interacción entre los procedimientos los cuales se unen formando la estructura dictada por el diseño. La integración de los módulos se realizó de manera incremental

para facilitar la detección y corrección de errores, es decir los módulos se fueron uniendo sistemáticamente de forma descendente.

Debido a que en las pruebas unitarias la integración fue de manera incremental, resulto más sencillo hacer toda la integración, y además comprobamos que no existían errores y que los datos mantienen los valores necesarios.

### 8.1.3 Pruebas de validación

Durante esta prueba se valida el funcionamiento específico del sistema de acuerdo con los requerimientos determinados. En este caso el procedimiento de prueba para demostrar la conformidad con los requisitos, consistió en la captura de diferentes ideas o conceptos bajo supervisión, la segunda parte de la prueba la realizó el usuario sin supervisión la cual notificó los problemas registrados y sus observaciones. Una vez resueltos los problemas se libera una versión validada del sistema, aunque es necesario hacer las pruebas de evaluación que se detallarán posteriormente.

### 8.1.4 Pruebas volumétricas

Para la presente investigación, el banco terminológico del área de Fenómenos contiene 32 términos, en el área de Física<sup>1</sup> se tienen 320 términos y para Sexualidad 287, esto representa realmente un número muy bajo de términos si consideramos que el Diccionario de la Real Academia Española, contienen alrededor de 91,000 términos. Respecto al número de definiciones que manejamos, el banco terminológico de Fenómenos contiene sólo una definición por cada uno de los términos, en Física y Sexualidad, sí existen para algunos términos, más de una definición.

Adicionalmente, a la tabla *palabra*, se agregaron todas aquellas que pertenecen a cada una de las definiciones existentes en la tabla *definicion*, se eliminaron todas las que estaban duplicadas y además aquellas que pertenecían a la tabla de *stoplist*, para que sólo existieran palabras funcionales, el resultado fue que cada una de las áreas: Fenómenos, Física y Sexualidad, existe un número diferente de palabras clave: 513, 3399, 1921 respectivamente.

---

<sup>1</sup> Consultar el diagrama de la Base de Datos en el capítulo 6.



## 8.2 Implantación

En esta actividad se revisa la estrategia de implantación para el sistema, en el capítulo 6 vimos las herramientas necesarias para su desarrollo.

Una vez estudiado el alcance y los condicionantes de la implantación, se decide si ésta se puede llevar a cabo. Será preciso establecer, en su caso, la estrategia que se concretará de forma definitiva en el plan de implantación.

### 8.2.1 Implantación del prototipo en computadora independiente

El sistema fue desarrollado usando como lenguaje de programación Java, y como herramienta JBuilder (Java Builder), esta herramienta permite al programador ejecutar el sistema en cualquier computadora, aunque para ello es necesario tenerla instalada, además se debe instalar el Microsoft SQL Server 2000 Driver for JDBC para realizar la conexión con el servidor de bases de datos, una vez instalado y usando JBuilder se crea un paquete que contiene las librerías del JDBC como se muestra en la figura 8.1:

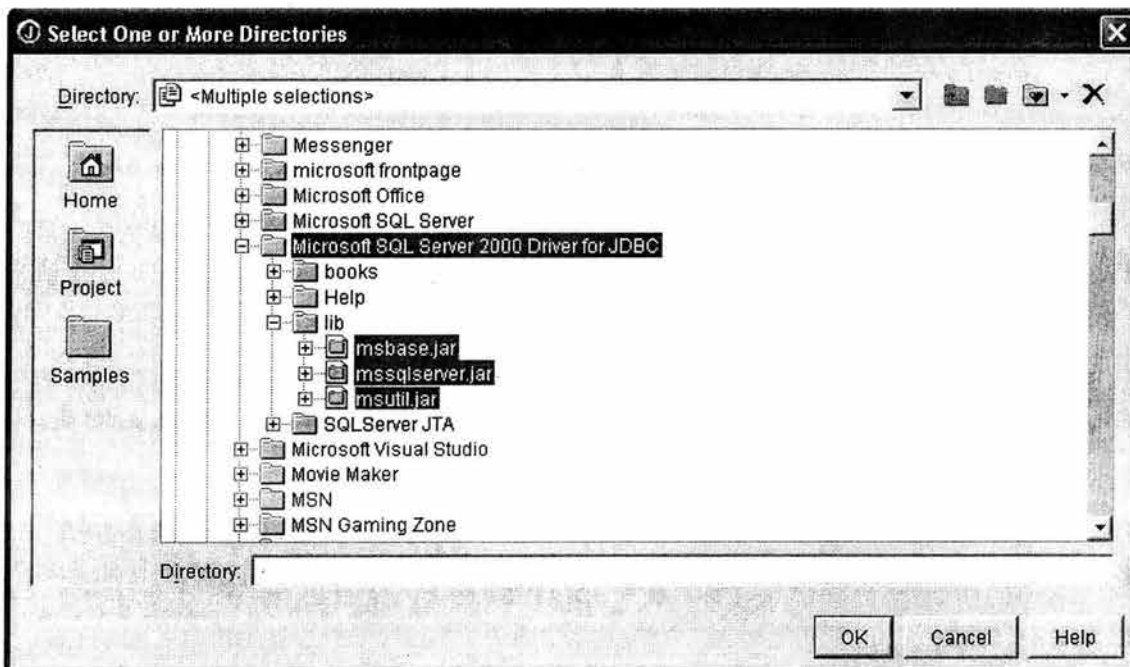


Figura 8.1 Selección de las librerías necesarias para conectar la Base de Datos.

Las librerías mostradas quedan dentro de un paquete que en este caso llamamos JDBC una vez que se tiene este paquete se agrega a las propiedades del proyecto y de esta manera se puede ejecutar el programa desde cualquier computadora.

### 8.2.2 Implantación final en Servidor Web

Para hacer la implantación en el Servidor Web, fue necesario crear una especie de paquete que contiene el Applet, todas las clases, las librerías que se ocuparon del Microsoft SQL Server y todo lo necesario para que el programa funcione, este paquete se crea desde JBuilder y fue colocado en una carpeta del servidor, en la cual además se crea una página en formato HTML que hace referencia a dicho paquete cuyo nombre es Diccionario.jar.

A continuación, en la figura 8.2, se muestra la forma en que se presenta el código HTML:

```

<html>
<head>
<title>Diccionario de Búsqueda Onomasiológica</title>
</head>
<body>
<p align="center"><a href="http://java.com/en/index.jsp">Para poder ver este
  programa debe tener instalado el JVM </a></p>
<center>

<applet
  codebase = "./"
  code    = "diccionario.DO.class"
  archive = "DICCIONARIO.jar, msutil.jar, msbase.jar, mssqlserver.jar"
  name    = "TestApplet."
  width   = "650"
  height  = "550"
  hspace  = "0"
  vspace  = "0"
  align   = "middle"
  >
</applet><p>
</body>
</html>

```

Figura 8.2 Código HTML de la página Diccionario.html

El archivo DICCIONARIO.jar y la página Diccionario.html, se encuentran en el Servidor Web en una carpeta llamada DEBO, por lo tanto la liga para poder ver la aplicación desde internet es:

<http://tabasco.torreingenieria.unam.mx/DEBO/Diccionario.html>

Para que el usuario vea la aplicación no es necesario instalar todas las herramientas que se mencionaron para la implantación en una computadora independiente, en este caso sólo se instala la máquina virtual de java que se puede descargar de la siguiente dirección: <http://java.com/en/index.jsp>.

En la figura 8.3 se observa cómo se visualiza la aplicación del diccionario onomasiológico desde Internet:

Para poder ver este programa debe tener instalado el JVM

**Lingüística**

**" Diccionario Electrónico de Búsqueda Onomasiológica "**

Elija una base de Datos:

Introduzca su definición:

Terminos:

Definición:

Figura 8.3 Aplicación del diccionario onomasiológico vista desde Internet.

En el diccionario se puede elegir entre tres áreas de conocimiento: Fenómenos, Física y Sexualidad, posteriormente el usuario introduce la idea de un concepto que no logra recordar con claridad y posteriormente en la lista de términos aparece él o los términos más probables de acuerdo a su idea.

Como ejemplo para conocer el funcionamiento del sistema, se tomará alguna de las ideas que se usarán para hacer las pruebas de evaluación, en la figura 8.4 se observa que el usuario escribe una frase:

**" Diccionario Electrónico de Búsqueda Onomasiológica "**

Elija una base de Datos:

Introduzca su definición:

Terminos:

Definición:

Figura 8.4 Solicitud del usuario al sistema.

cuando presiona el botón "BUSCAR" debe esperar unos segundos a que aparezca un resultado y a continuación verá cómo la lista de términos se llena con los resultados obtenidos, como en la figura 8.5:

**" Diccionario Electrónico de Búsqueda Onomasiológica "**

Elija una base de Datos:

**Introduzca su definición:**

**Terminos:**

EROSION  
**TERRORISMO**

**Definición:**

Es una sucesión de actos de violencia ejecutados para infundir terror. Generalmente son acciones que atentan contra la vida, la salud y los bienes materiales los demás.

Figura 8.5 Respuesta del sistema al usuario.

se observa que se obtienen dos términos como respuesta a la búsqueda del usuario y que además tiene la posibilidad de conocer la definición exacta de cada uno de los términos encontrados.

### 8.3 Pruebas de Evaluación

Este tipo de pruebas permiten obtener elementos cualitativos y cuantitativos de los resultados obtenidos, con el fin de determinar la eficiencia del sistema y en particular del algoritmo implantado.

El proceso de evaluación consiste en establecer manualmente, con base en el conocimiento del propio idioma, los términos obtenidos en el diccionario onomasiológico, en esta sección las pruebas realizadas son únicamente para el área de Fenómenos. El apéndice del capítulo ocho, muestra los

términos y definiciones que constituyen el banco terminológico sobre el que se aplicaron las pruebas.

### **8.3.1 Batería de pruebas individuales**

Para realizar esta prueba, se proporcionó a un grupo de personas la lista de términos disponibles y les pedimos que eligieran uno o varios conceptos y escribieran una idea que lo o los describieran.

En la tabla 8.1 se puede ver cuáles fueron las consultas que les interesaban y los resultados esperados.

<b>Consultas</b>	<b>Términos esperados</b>
Todos los actos o fenómenos destructivos que son provocados por los humanos	Disturbios sociales Sabotaje Terrorismo Acción bélica Falla o error humano Accidente Crecimiento explosivo de la población
Accidentes que pueden ocurrir en las empresas, debido a los materiales que empleados para el trabajo	Radiación Fuga y derrame de sustancias Incendio Explosión Falla o error humano Accidente
Cuando en el aire existe una gran cantidad de partículas que provocan daño a la salud	Contaminación
Movimiento brusco del mar en el que se producen inundaciones	Maremoto
Superposición y movimiento de placas tectónicas en el planeta tierra	Sismo Terremoto
Caída excesiva de agua	Lluvia Tormenta
Amontonamiento de bichos	Plaga
Falta de información que lleva a un crecimiento acelerado de los seres humanos en el planeta tierra que nos llevara a la autodestrucción	Crecimiento explosivo de la población
El manejo inadecuado de sustancias químicas puede traer consigo la propagación de enfermedades	Radiación Fuga y derrame de sustancias Incendio Explosión Falla o error humano Accidente
El debilitamiento de la capa de ozono puede traer consigo, el deterioro genético de la población mundial	Radiación Contaminación Sequía Erosión
Ejercer medidas violentas con el fin de atentar contra la integridad física de personas inocentes promoviendo la incertidumbre y el terror de una sociedad determinada	Terrorismo
Falta de el vital liquido que da la vida en el planeta	Sequía
Movimiento de los edificios, del suelo y cuando cae mucho agua	Colapso de suelos Sismo Maremoto Lluvia Huracán Inundación
Una persona ingiere algún tóxico que le hace mal y cuando tiran mucha basura y la temperatura se eleva	Envenenamiento Calor
Cuando un país es atacado con otro con armas	Terrorismo
Las personas tienen muchos hijos en un país donde habitan ya muchas personas	Crecimiento explosivo de la población
Fenómenos destructivos que tienen que ver con agua	Maremoto Lluvia Inundación Huracán Tormenta de granizo Tormenta eléctrica Sequía.
Suceso causado por el aumento en la temperatura o la liberación de sustancias fáciles de ocasionar fuego o lesiones en animales, plantas y materia orgánica.	Calor Radiación Incendio Explosión
Población afectada por una misma enfermedad, la cual puede provocar la muerte de esta.	Epidemia
Causa de muerte por falta de agua	Sequía

Tabla 8.1 Consultas del usuario al sistema y términos esperados.

### **8.3.2 Aplicación de la Batería de pruebas**

La aplicación de la batería de pruebas consiste en poner a prueba el sistema de acuerdo al conjunto de ideas que proporcionaron algunos usuarios, estas ideas se capturan de una en una y se anotan los resultados obtenidos.

En la tabla 8.2 se observa cuáles son los términos que se esperaban y cuáles son los términos que el sistema devolvió.



Definición	Términos esperados	Términos devueltos
Todos los actos o fenómenos destructivos que son provocados por los humanos	Disturbios sociales, Sabotaje, <b>Terrorismo</b> , Acción bélica, Falla o error humano, Accidente, Crecimiento explosivo de la Población	Erosión, <b>Terrorismo</b>
Accidentes que pueden ocurrir en las empresas, debido a los materiales que empleados para el trabajo	<b>Radiación</b> , Fuga y derrame de sustancias, Incendio, Explosión, <b>Falla o error humano</b> , Accidente	<b>Falla o error humano, Radiación</b>
Cuando en el ambiente existe una gran cantidad de partículas que provocan daño a la salud	<b>Contaminación</b>	<b>Contaminación</b> Maremoto
Movimiento brusco del mar en el que se producen inundaciones	<b>Maremoto</b>	<b>Maremoto</b> Sismo
Superposición y movimiento de placas tectónicas en el planeta tierra	<b>Sismo</b> , Terremoto	<b>Sismo</b> , Erosión
Caída excesiva de agua	<b>Lluvia</b> , Tormenta	<b>Lluvia</b> , Helada
Amontonamiento de bichos	Plaga	<i>Ninguno</i>
Falta de información que lleva a un crecimiento acelerado de los seres humanos en el planeta tierra que nos llevara a la autodestrucción	Crecimiento explosivo de la población	Sismo, Deslizamiento de talud, Erosión, Inundación
El manejo inadecuado de sustancias químicas puede traer consigo la propagación de enfermedades	Radiación, Fuga y derrame de sustancias, Incendio, Explosión, Falla o error humano Accidente	Desertificación, Sabotaje
El debilitamiento de la capa de ozono puede traer consigo, el deterioro genético de la población mundial	Radiación, Contaminación, Sequía, Erosión	Acción bélica, Nevada, Sismo
Ejercer medidas violentas con el fin de atentar contra la integridad física de personas inocentes promoviendo la incertidumbre y el terror de una sociedad determinada	Terrorismo	<i>Ninguno</i>
Falta de el vital liquido que da la vida en el planeta	Sequía	Explosión, Falla o error humano, Inundación, Deslizamiento de talud
Movimiento de los edificios, del suelo y cuando cae mucho agua	Colapso de suelos, <b>Sismo</b> , Maremoto, <b>Lluvia</b> , Huracán, Inundación	<b>Lluvia, Sismo</b>
Una persona ingiere algún tóxico que le hace mal y cuando tiran mucha basura y la temperatura se eleva	Envenenamiento, <b>Calor</b>	Inundación, <b>Calor</b>
Cuando un país es atacado con otro con armas	Terrorismo	<i>Ninguno</i>
Las personas tienen muchos hijos en un país donde habitan ya muchas personas	Crecimiento explosivo de la población	
Fenómenos destructivos que tienen que ver con agua	Maremoto, <b>Lluvia</b> , Inundación, Huracán, Tormenta de granizo, Tormenta eléctrica, Sequía.	<b>Lluvia</b> , Erosión
Suceso causado por el aumento en la temperatura o la liberación de sustancias fáciles de ocasionar fuego o lesiones en animales, plantas y materia orgánica.	Calor, Radiación, <b>Incendio</b> , Explosión.	<b>Incendio</b> , Maremoto, Terrorismo, Accidente, Calor
Población afectada por una misma enfermedad, la cual puede provocar la muerte de esta.	Epidemia	Acción bélica, Calor, Helada
Causa de muerte por falta de agua	Sequía	Lluvia, Sismo, Deslizamiento de talud, Helada, Inundación

Tabla 8.2 Comparación de resultados esperados contra resultados obtenidos

### 8.3.3 Análisis de los resultados

El proceso de análisis de los resultados debe buscar obtener elementos cuantitativos más que cualitativos para establecer las bondades del sistema.

Cuando se considera el problema de evaluar la eficiencia en los algoritmos de recuperación de información es necesario considerar primero la naturaleza del proceso de recuperación que se debe llevar a cabo. En este sentido, se pueden identificar dos grandes grupos:

- Procesos autónomos: Es un proceso, conocido como “*batch*”, donde el usuario hace una solicitud y la computadora simplemente responde.
- Procesos interactivos: Implican una sesión de trabajo donde el usuario y la computadora interactúan para llegar a un resultado final.

La evaluación de procesos como el que se presenta en esta tesis, sigue la línea planteada por la evaluación de los procesos autónomos.

## 8.4 Índices de Precisión y Recuperación (Precision and Recall)

El objetivo de esta sección es delinear los conceptos y técnicas de sistemas de recuperación de información considerados en esta investigación.

El área de recuperación de información, se encarga principalmente del estudio de sistemas y técnicas para asignar índices, buscar y devolver datos valiosos a los usuarios. Este conjunto de procedimientos se denomina también filtrado de información. Por lo general se procesa texto.

Una vez que se obtiene uno o varios términos que responden a la consulta que hace el usuario, se asocia un cierto valor que indique qué tanto se han satisfecho las necesidades de información del mismo. El valor se conoce comúnmente como relevancia, aunque puede ser una medida subjetiva.

Utilizando la relevancia como punto de partida, existen medidas estándares para evaluar la ejecución en los sistemas de recuperación de información. Una de ellas es (como se vio en el capítulo 4) *precision*, definida como el número de términos relevantes entre el total de términos recuperados.

Por ejemplo, suponiendo que la colección X contiene 32 términos, y que el sistema de recuperación de información Y devuelve 4, de los cuales 1 es relevante. Entonces, la precisión del sistema Y es del 25% (1/4).

Otra medida utilizada ampliamente es **recall**, la cual se calcula cuando el número de términos relevantes que satisfacen una consulta, se divide entre el total de términos relevantes de la colección.

Como ejemplo, suponiendo que en la colección X existen 32 términos relevantes, y que el sistema de recuperación Y regresa 8 de los cuales sólo 4 son relevantes, por tanto, el *recall* es del 12.5% (4/32).

## 8.5 Comparación de resultados

Utilizando las técnicas de *precision* y *recall*, se hace el análisis de los resultados esperados con los resultados obtenidos a través del sistema.

Para *precision*, un valor  $P=100\%$  indica que todos los términos recuperados son correctos y para *recall* un valor de  $R=100\%$  indica que se han recuperado todos los términos que se han identificado manualmente.

En la tabla 8.3 se observa cuáles son los valores de *precision* y *recall* obtenidos:

<b>Definición</b>	<b>Precision</b>	<b>Recall</b>
Todos los actos o fenómenos destructivos que son provocados por los humanos	50% (1/2)	14.29% (1/7)
Accidentes que pueden ocurrir en las empresas, debido a los materiales que empleados para el trabajo	100% (2/2)	33.33% (2/6)
Cuando en el ambiente existe una gran cantidad de partículas que provocan daño a la salud	50% (1/2)	100% (1/1)
Movimiento brusco del mar en el que se producen inundaciones	50% (1/2)	100% (1/1)
Superposición y movimiento de placas tectónicas en el planeta tierra	50% (1/2)	50% (1/2)
Caída excesiva de agua	50% (1/2)	50% (1/2)
Amontonamiento de bichos	0%	0% (0/1)
Falta de información que lleva a un crecimiento acelerado de los seres humanos en el planeta tierra que nos llevara a la autodestrucción	0% (0/4)	0% (0/1)
El manejo inadecuado de sustancias químicas puede traer consigo la propagación de enfermedades	0% (0/2)	0% (0/6)
El debilitamiento de la capa de ozono puede traer consigo, el deterioro genético de la población mundial	0% (0/3)	0% (0/4)
Ejercer medidas violentas con el fin de atentar contra la integridad física de personas inocentes promoviendo la incertidumbre y el terror de una sociedad determinada	0%	0% (0/1)
Falta de el vital liquido que da la vida en el planeta	0% (0/3)	0% (0/1)
Movimiento de los edificios, del suelo y cuando cae mucha agua	50% (1/2)	33.33% (2/6)
Una persona ingiere algún tóxico que le hace mal y cuando tiran mucha basura y la temperatura se eleva	50% (1/2)	50% (1/2)
Cuando un país es atacado con otro con armas	0%	0% (0/1)
Las personas tienen muchos hijos en un país donde habitan ya muchas personas	0%	0% (0/1)
Fenómenos destructivos que tienen que ver con agua	50% (1/2)	14.29% (1/7)
Suceso causado por el aumento en la temperatura o la liberación de sustancias fáciles de ocasionar fuego o lesiones en animales, plantas y materia orgánica.	40% (2/5)	50% (2/4)
Población afectada por una misma enfermedad, la cual puede provocar la muerte de esta.	0% (0/3)	0% (0/1)
Causa de muerte por falta de agua	0% (0/5)	0% (0/1)

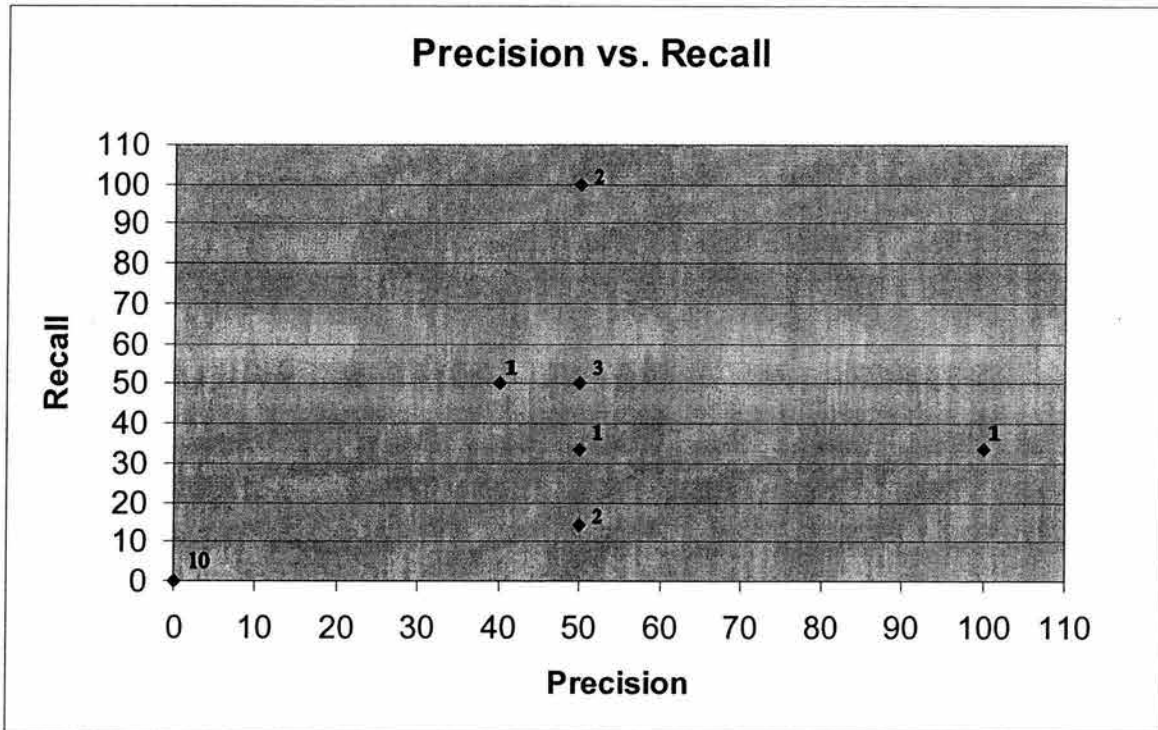
Tabla 8.3 Comparación de resultados utilizando las técnicas *precision* y *recall*.

Resumiendo la tabla anterior:

<b>Precision</b>	<b>Recall</b>	<b>Ocurrencias</b>
0%	0%	10
40%	50%	1
50%	14.24%	2
50%	33.33%	1
50%	50%	3
50%	100%	2
100%	33.33%	1

Tabla 8.4 Resumen de resultados: *precision*, *recall*.

De la tabla anterior se grafican los resultados y se obtiene la siguiente gráfica:



Gráfica 8.1 *Precision vs. Recall*

Como se observa en la gráfica, de los resultados obtenidos 9 de los 20 se encuentran en un porcentaje mayor o igual al 50% de *precision*, aunque se recuperaron pocos términos respecto de los identificados manualmente, y eso se sabe gracias a la evaluación y a los resultados que arroja la medida *recall*, porque 14 de los resultados están por debajo del 50% también, aunque es muy importante resaltar que hasta esta etapa, el proceso de lematización no se ha llevado a cabo, una vez que el lematizador sea implantado en el sistema, se espera que los resultados mejoren.

## 8.6 Conclusiones

En este capítulo se analizaron los requerimientos para llevar a cabo la implantación del sistema, y cómo llevarla a cabo. Las pruebas realizadas indican que el sistema cumple con las especificaciones iniciales y para evaluar los resultados obtenidos usamos las medidas de evaluación *precision* y *recall* las cuáles indican que la recuperación de términos no es muy satisfactoria aunque la *precision* es buena. Al incorporar el proceso de lematización se espera que los resultados mejoren.

Además se comprobó que cuando un usuario introduce una frase que no está relacionada con el área, el sistema no devuelve ningún resultado, por lo que en este caso podemos decir que las medidas *precision* y *recall* son buenas, aunque no es conveniente hacer la evaluación cómo en los otros casos.

# CAPÍTULO 9

## CONCLUSIONES



En el capítulo ocho se llevaron a cabo las pruebas, la implantación y evaluación del sistema, se realizaron pruebas para verificar que el sistema cumple con los requerimientos iniciales y pruebas que permitieron hacer una evaluación de los resultados obtenidos.

En el presente capítulo se darán las conclusiones generales de la tesis, dando un panorama general de los resultados obtenidos a lo largo del trabajo y proponiendo futuras áreas de trabajo.

A continuación se presenta: (1) un panorama final del trabajo, (2) los resultados obtenidos en esta tesis y (3) un conjunto de trabajos que deberán realizarse en el futuro para enriquecer el algoritmo y el sistema desarrollado.

El panorama teórico visto en los capítulos anteriores delimita la opción de búsqueda que se quiere proporcionar al usuario: el lenguaje natural. En las primeras unidades se justificó la necesidad de concebir un diccionario onomasiológico que funcionara a través de lenguaje natural, por eso en las unidades tres y cuatro se estudió teórica y prácticamente esta posibilidad, para que a partir del capítulo seis la metodología a seguir fuera la adecuada.

Según las actuales tendencias, la construcción de diccionarios onomasiológicos en versión electrónica será más fructífera en tanto los diseños permitan la actualización de la información de una forma continua. Dependiendo de las prestaciones del sistema, se podrá agilizar la consulta y permitir otras formas de búsqueda.

Sin embargo, la construcción de un diccionario onomasiológico con las anteriores características debe ser construido de una forma flexible a fin de que no limite al usuario. Es decir, no debe importar si el usuario introduce definiciones simples o complejas, las cuales contendrán con cierta frecuencia información irrelevante que deberá ser removida.

Por otro lado, las definiciones libres o input –como se vio en el capítulo tres- no siempre remiten al mismo tipo de definición e información. En algunos casos el usuario emplea en sus definiciones creencias, estereotipos o contextos, con esto habrá que contemplar en algún momento alguna forma de interpretar esta información.

En el capítulo cuatro se realizó un listado de problemas en los diccionarios evaluados, el cual permitió establecer bases para la construcción del prototipo para este trabajo: lematización de información, expansión semántica y expansión morfológica. Lo anterior mejoraría notablemente los resultados de un diccionario onomasiológico.

En el capítulo cinco se retoma este listado para construir finalmente el proceso del diseño de búsquedas onomasiológicas en lenguaje natural. Se describió la forma en que el sistema (a partir de la descripción del concepto en lenguaje natural proporcionada por el usuario) realiza el proceso de identificación de palabras claves, las expande (morfológica y semánticamente) y finalmente identifica los términos más probables.

En el capítulo seis se explicaron los conceptos teóricos de las herramientas necesarias para iniciar la construcción del diccionario onomasiológico. Todo lo visto en este capítulo sirvió como base para desarrollar el Diccionario Onomasiológico para la Web.

El capítulo siete es importante ya que se profundizó en el análisis, desarrollo y diseño del diccionario, es decir, fue la parte práctica del mismo. Mediante la aplicación del método del ciclo de vida se logró mostrar los pasos a seguir para el desarrollo del sistema. Asimismo, se permite tener un panorama general de la forma en que el usuario participará con el sistema una vez implantado.



El capítulo ocho presenta un informe sobre cómo se realizaron las pruebas, implantación y evaluación del sistema. Las pruebas cualitativas y cuantitativas realizadas verifican que el sistema tiene un funcionamiento adecuado y que proporciona salidas correctas conforme a la entrada del usuario. Los resultados aún no son del todo satisfactorios, pero se espera que mejoren una vez que sea implementado el lematizador.

Con la realización de este trabajo se puede afirmar que se cumplieron los objetivos y alcances que se plantearon al inicio del mismo, logrando satisfactoriamente el trabajo de las dos Facultades de la UNAM para lograr el resultado que se presenta.

## 9.1 Resultados obtenidos

Los índices de *Recall* y *Precision* no pudieron ser adecuadamente interpretados, en parte porque el banco terminológico empleado es pequeño (30 definiciones en el área de desastres) y porque es necesario establecer un mayor tamaño tanto en el número de pruebas como en la diversidad del nivel académico de los sujetos de prueba. Sin embargo, hasta ahora, y con base en los resultados obtenidos, se puede afirmar que el algoritmo propuesta arroja resultados alentadores en la identificación adecuada de un término a partir de una definición libre.

El trabajo interdisciplinario efectuado entre los alumnos de la Facultad de Filosofía y Letras y la Facultad de Ingeniería, bajo la dirección de los investigadores del Grupo de Ingeniería Lingüística, ha entregado logros altamente satisfactorios, y permiten mostrar que el trabajo grupal aún en disciplinas tan distintas no sólo es posible sino también enriquecedor.

## 9.2 Trabajos futuros

Con el desarrollo del sistema, queda abierta la posibilidad de realizar cambios en el funcionamiento del mismo, sin perder la estructura inicial que se estipula en los requerimientos, esto con el fin de que futuros diccionarios que se desarrollen permitan a los usuarios tener la posibilidad de realizar consultas eficientes por medio de una interfaz que sea sencilla y manejable para el usuario.

Por un lado, la incorporación de un proceso de lematización debe enriquecer los resultados. Los algoritmos actuales tienen un buen desempeño en lenguas como el inglés, sin embargo presentan desempeños bajos cuando son aplicados al español. Por ello, también es necesario ahondar en investigaciones lingüísticas para desarrollar un algoritmo de lematización para el español. Por el otro, se debe continuar la investigación en la expansión semántica en relaciones no sólo sinonímicas entre las palabras, sino también holonímicas y meronímicas.

Asimismo, se debe establecer un mejor conjunto de pruebas, aplicarlo utilizando sujetos con diferentes niveles de conocimiento y provenientes de diferentes puntos geográficos permitirá establecer un mejor criterio respecto al algoritmo propuesto, permitiendo identificar además líneas de trabajo para mejorar el desempeño del algoritmo y la calidad y claridad de la interfaz del sistema.



## **Apéndice del Capítulo 2**

### **Diccionario Onomasiológico**

En este apéndice se pueden consultar:

- Una tabla que muestra los diccionarios onomasiológicos desde el siglo XVII al XX.
- Tablas que enumeran algunas características de los términos del corpus.
- Tablas que presentan el corpus empleado en este trabajo.

**Diccionarios Onomasiológicos (Español y Catalán)**

			T(emático)
	Año de publicación (ediciones subsecuentes)	Título Nombre del compilador/ed. Lugar de publicación	A(lfabético)
			C(umulativo) D(istintivo) P(ictórico)
1626		<i>Thesauri linguarum in quo nomenclatura et proverbia Hispanica, Gallica et Italica continentur</i> Juan Angel Sumaran. Ingolstadt	
1789		<i>Examen de la posibilidad de fijar la significación de los sinónimos de la lengua castellana</i> J. López de la Huerta. Vienna 1799, [Valencia] 1807, [Madrid] 1835	A D
1806		<i>Ensayo sobre la distinción de los sinónimos de la lengua castellana</i> Santiago Jonama. Madrid (2) 1936	A D
1834		<i>Sinónimos de la lengua castellana</i> José March. Barcelona (2) 1838	A D
1843		<i>Diccionario de sinónimos de la lengua castellana</i> P.M. de Olive. Madrid 1908	A C
1845		<i>Diccionario de sinónimos</i> . Conde de la Cortina. Mexico (2) 1853	A D
1853		<i>Diccionario de sinónimos de la lengua castellana</i> Societé de Littérature. Paris	A C
1855		<i>Colección de sinónimos de la lengua castellana</i> Juan José de Mora. Madrid, Visor libros (Biblioteca Filológica hispánica/7).	A D
1864		<i>Filosofía de la lengua española. Sinónimos castellanos</i> Roque Barcia. Madrid 1939	A D
1873		<i>Diccionario de sinónimos</i> J. Carrión. Madrid	
1879		(incompleto) <i>Inventario de la lengua castellana. Índice ideológico del Diccionario de la Academia</i> [in <i>Verbos</i> ] José Ruiz León. Madrid	T? C?
1890		<i>Diccionario de sinónimos castellanos</i> R. Barcía. Madrid (5) 1948, rev. edn [Buenos Aires, Sopena Argentina] 1960, 1967	A D
1898		<i>Diccionario de ideas afines</i> Eduardo Benot y Rodríguez. Madrid [Buenos Aires] 1940	
1910		<i>Diccionario de sinónimos</i> . Roque Barcia, México, Oasis, 1983.	A D
1911		<i>Sinónimos castellanos y voces de sentido análogo</i> Benjamin Monroy Ocampo. Madrid	

1916	<i>Sinónimos. Repertorio de palabras usuales ...</i> Pedro de Irizar y Avilés. Barcelona. (2) 1928, 1933, (8) 1961 <i>Pequeño diccionario de sinónimos</i> , 1967, (6) 1976 ed. A. Viñoly et al.	
1925	<i>Diccionario ideológico para facilitar el trabajo literario y enriquecer el estilo</i> Enrique Gómez Carrillo & Alfonso de Sola. Madrid	
1930	<i>Diccionario hispanoamericano de voces sinónimos y análogos</i> Gabriel Maria Vergara y Martín. Madrid	A C
1939	<i>Sinónimos castellanos</i> . Roque Barcia, Ed Joaquín Gil, Buenos Aires, Talleres Gráficos argentinos.	
1941	<i>Diccionario de símiles</i> José Raúl, México, Aguilar Ediciones Lux.	A D A C
1941	<i>Diccionario de sinónimos e ideas afines</i> A. López Bejarano & Casto Peña. Barcelona 1949	A C
1942	<i>Diccionario ideológico de la lengua española (Desde la idea a la palabra; desde la palabra a la idea)</i> Julio Casares. Barcelona, Editorial Gustavo Gili (1959). 1951, (2) 1959 [cf. 1995 <i>Diccionario ideológico</i> ]	T D/C
1944	<i>Grates diccionario de sinónimos castellanos</i> Barcelona 1961 [Buenos Aires]	
1946	<i>Diccionario español de sinónimos y antónimos</i> Federico C. Sainz de Robles. Madrid, Aguilar 1968, 1971, 1976, 1979, 1981 ISBN: 84 03 27029-1	A C
1947	<i>Diccionario ideo-constructivo</i> [in <i>Ciencia del lenguaje y arte del estilo</i> ] Martín Alonso. Madrid (12) 1982, (13) 1984 <i>Diccionario de sinónimos explicados</i>	
1954	<i>Diccionario gramatical</i> . Emilio M. Martínez Amador, Barcelona, Ramón Sopena,	A D
1954	<i>Diccionario de sinónimos españoles</i> Antonio Zamora. Buenos Aires 1961, (5) 1969	A C
1956	<i>Diccionario de parónimos y dudas del castellano</i> S. Lazzati. Buenos Aires. Fondo editorial IPASME (1995) ISBN: 980-6122 44-5	A D
1958	<i>VOX Diccionario de sinónimos</i> Samuel Gili Gaya. Barcelona, Bibliograf 1968, 1972, (5) 1975, (7) 1980, [abreviada edn] 1972, 1981, 1982	A D
1959	<i>Diccionario español de sinónimos, equivalencias e ideas afines</i> M.F. Andrés. Barcelona (8) 1979	A C
1959	<i>Nuevo diccionario de sinónimos y palabras afines</i> Alberto Ruiz Cárdenas. Lima	A? C
1961	<i>Diccionario de sinónimos e ideas afines</i> Julio de la Canal. México 1972.	A C

- 1962 *Diccionario de sinónimos* M. Santos. México
- 1964 *Diccionario ortográfico de sinónimos* G. Sanín Villa. Bogotá.
- 1965 *Antónimos castellanos* José Cruz Aufrere. La Paz
- 1967 *Diccionario de sinónimos* Buenos Aires.
- 1969 *Casals diccionario de sinónimos* Barcelona.
- 1970 *Diccionario de sinónimos e ideas afines y de la rima* ed. J. Horta Massanes. Madrid Paraninfo, 1981. ISBN: 84 283 0317-7 A C
- 1970 *Diccionario de sinónimos, antónimos e ideas afines* Buenos Aires
- 1971 *Gran diccionario de sinónimos, voces afines e incorrecciones*, Fernando Corripio Pérez, Barcelona A C
- 1972 *Diccionario esp... de sinónimos y equivalencias* M.F. Andrés. Barcelona (6?) 1973.
- 1972 *Diccionario manual de sinónimos, antónimos e ideas afines* Andrés Santamaría. Barcelona A C
- 1972 *Novísimo diccionario práctico de sinónimos y antónimos* Jorge Fernandez Mara. Buenos Aires A C
- 1973 *Diccionario razonado de sinónimos y contrarios.* José Maria Zainqui. Barcelona, De Vecchi. ISBN: 84- 315 0931 7 A D
- 1975 *Diccionario de incorrecciones: Dudas y normas gramaticales*, Fernando Corripio Barcelona, Bruguera. ISBN: 84 02 04591-X A T D
- 1975 *Diccionario de incorrecciones, particularidades y curiosidades del lenguaje.* Andrés Santamaría et al. Madrid, Paraninfo. ISBN: 84 283 0112 3 A T D
- 1973 [CATALAN] *Diccionari manual de sinónims VOX.* Barcelona
- 1975 VOX: *Diccionario temático de la lengua española / Prólogo de Juan Alcina Franch.* Barcelona. T D
- 1975 *Voces homófonas, homógrafas y homónimas castellanas / Alvaro J. Moreno.* México : A. J. Moreno, (1977, aumentada). Porrúa. A D
- 1975 [CATALAN] *Diccionari usual de sinónims catalans. Mots i frases* Juana Rospel de Cahué. Barcelona
- 1976 [CATALAN] *Diccionari de sinónims i antónims* Santiago Pey Estrany. Barcelona (3) *económica* edn 1976
- 1976 *Diccionario actualizado de sinónimos y contrarios de la lengua española [in Diccionario ideológico manual de la lengua española]* Distein. ... A? C

1978	<i>VOX: diccionario de sinónimos.</i> Samuel Gili Gaya. Barcelona, Bibliograf.	A C
1979	<i>VOX diccionario escolar de sinónimos y antónimos</i> Barcelona. Bibliograf.	A C
1980	Diccionario de sinónimos. Mariano Orta Manzano. Ed Juventud, Barcelona. ISBN: 84 261 1704-X	A D
1980	<i>Diccionario temático. Americanismos</i> Miguel A. Arias de la Cruz. León, Ed Everest. ISBN: 84 241 1504-X (1987)	A T D
1980	<i>Diccionario de dudas y dificultades de la lengua española</i> Manuel Seco. Madrid, Aguilar. (8) 1980, (9) 1986 ISBN: 84-03-27069-0	A T D
1981	<i>Diccionario ideológico feminista</i> [2 volúmenes] Victoria Sau. Barcelona, Ed. Icaria 1990, 2000, 2001 ISBN: 84 7426 0728	A T D
1984	<i>Diccionario de sinónimos, ideas afines y contrarios</i> Santiago Pey Estrany & Juan Ruiz Calonja. México ISBN: 968-7128-13-5	A C
1984	<i>Thesaurus Gran Sopena de sinónimos y asociación de ideas. Diccionario analógico de la lengua española</i> David Ortega Cavero. Barcelona	A C
1985	<i>Diccionario de ideas afines.</i> Fernando Corripio. Barcelona, Herder, (1996, 2000).	A D
1987	<i>Gran diccionario de sinónimos, antónimos e ideas afines.</i> Autor Tirso R. M. Morrison	A C
1980	<i>Diccionario de Argot</i> Víctor León. Madrid, Alianza editorial. (1992, 1996) ISBN: 84 206 1766 0	A T D
1986	<i>Diccionario de sinónimos y antónimos de la lengua española</i> Ángel López García-Molins. Valencia	A C
1987	<i>Diccionario inverso de la lengua española</i> I. Bosque & M. Perez Hernandez. Madrid ISBN: 84-249 1079-6	A Inverso
1987	<i>Diccionario de sinónimos e ideas afines con antónimos</i> (2 vol), México, Editores Mexicanos Unidos, 1987 ISBN: 968 15 0717-7	A C
1990	<i>Diccionario de sinónimos</i> (Apéndice diccionario ortográfico por José Manuel Parroquín) Roque Barcia. México, Colofón. ISBN: 968 867 032 4	A D
1990	<i>Diccionario de sinónimos y antónimos.</i> (2 vol) Barcelona, Mediterráneo. ISBN: 84 7764 462 6. 1990	A D
1993	<i>Diccionario de palabras olvidadas o de uso poco frecuente.</i> Elvira Muñoz, Madrid, Paraninfo.	A T D
1993	<i>Pequeño diccionario de sinónimos, ideas afines y contrarios,</i> Barcelona, Teide.	A C
1994	<i>Mega-thesaurus sinónimos, antónimos, parónimos e ideas afines,</i> Barcelona,	

	Ramón Sopena.	A D
1994	<i>Larousse sinónimos y antónimos</i> . México, ISBN 970 607 073-7	A C
1994	<i>MASTER 4: Diccionario antónimos y sinónimos</i> (Tomo 3).Barcelona, OLIMPO ediciones. ISBN: 84-7937-136-6	A C
1995	VOX: <i>Diccionario ideológico de la lengua española</i> Barcelona , Bibliograf ISBN: 84 7153 812 1	T C
1996	<i>Diccionario de sinónimos</i> , Madrid : Cultural,	A C
1997	<i>Del concepto a la palabra. Diccionario temático</i> . Rafael Méndez. Madrid, Temas de Hoy. ISBN: 84 7880 735 7	T D
1997	<i>Diccionario temático del español : --de la idea a las palabras, de la palabra a las Ideas</i> . Rafael del Moral, Madrid, Editorial Verbum. ISBN: 84-7962 099 4	T C
1997	<i>Diccionario de equívocos (Definiciones, expresiones, frases y locuciones)</i> . León Deneb, Madrid, Biblioteca Nueva. ISBN: 84- 7030-414-3.	A, T, C (FRASES)
1998	<i>Diccionario de usos y dudas del español actual</i> . José Martínez de Sousa. Barcelona, Bibliograf. (2º ed) ISBN: 84 7153 985 3 D/C	A T
1998	<i>Diccionario avanzado de la lengua española</i> . Barcelona ISBN: 84 7153 967 5	A D
1998	<i>Diccionario de sinónimos, ideas afines, antónimos y parónimos</i> . Tomás del Rey. Madrid, Cultural. ISBN 84-8055131-3	A C
1998	<i>Dudas e incorrecciones del idioma</i> , Fernando Corripio, México, Larousse.	A D
1998	<i>Diccionario de sinónimos y antónimos de la lengua española</i> , Fernando Corripio, Barcelona. Larousse.	A C
1998	<i>Diccionario de blasfemias, irreverencias y reniegos</i> . Ángel López, Madrid, Alderaban	A T D
1999	<i>Dudas e incorrecciones muy frecuentes del idioma español</i> , Pineda Alejandro, México, Comunicación óptima.	A T D
2000	<i>Diccionario de sinónimos y antónimos</i> . Ed Juan Pérez y María Villalón. Barcelona, Océano. ISBN: 84 494 1534 9	A C
2000	<i>Diccionario paremiológico e ideológico de la edad media (Castilla, siglo XIII)</i> , Hugo O. Bizzarri, Buenos Aires, SECRIT, Publicaciones 5. ISBN: 987 99735 5 0	T D
Sin fecha:		
	<i>Diccionario de ideas afines y elementos de tecnología</i> . Mariano Núñez Samper. Madrid (4 vol).	T C



Diccionarios pictóricos:

- |      |  |        |
|------|--|--------|
| 1945 | VOX: <i>Diccionario ilustrado de la lengua española</i> . Barcelona, Bibliograf (1978, 1953, 1973, 1976, 1977, 1978) ISBN: 84 7153 109 7 | A P    |
| 1990 | <i>Diccionario temático visual</i> . Ed José López Andrade et al. México, Reader's digest. ISBN : 968 28 0128 1                          | T D P  |
| 1997 | <i>Diccionario ortográfico ideovisual</i> . Manuel Sanjuán Nájera. Zaragoza, Yalde. ISBN: 84 87705 35 9                                  | A D, P |

Agua	Artefacto	Forma	Función	objeto	origen						
Limón	Comestible	Forma	Función	Grupo	Viviente	Natural	objeto	origen	parte	planta	sustancia
ardilla	Forma	Natural	objeto	origen							
cubeta	Artefacto	contenedor	Forma	Función	instrumento	objeto					
banca	Artefacto	Forma	Función	Mueble	Grupo	instrumento	objeto	origen			
ropa	Artefacto	ubre	Forma	Función	Vestimenta	objeto	origen	sólido	sustancia		
eutanasia	Agentivo	Evento delimitado	Causa	dinámico	existencia	físico	Situación				
monopolio	Condición	Propiedad	Situación	estático							
huracán	causa	dinámico	existencia	fenómeno	Físico	Situación					
barómetro	Artefacto	forma	Función	Instrumento	objeto	origen					

Tabla 2.1 Características de los términos según Wordnet

### Definiciones del corpus

Clave	Definición libre del concepto 'agua'	Fase 2
1agul	It's a clear liquid that you get from a tap	Water
2agul	The colourless transparent liquid occurring on rivers	Water
3agul	A clear, neutral liquid that surrounds us everywhere	Water
4agul	Liquid, clear, drinkable – constituents are hydrogen and oxygen	Water
5agul	Liquid, clear, H <sub>2</sub> O	Water
6agul	Liquid form, scientific term H <sub>2</sub> O	Water
7agul	Liquid, freezes at 0°C	Water
8agul	Liquid, clear, boils at 100°C, freezes at 0°C	Water
9agul	Fluid, clear, tasteless, colourless	Water
10agul	Wash with it; drink it; used for dilution; H <sub>2</sub> O; found in springs, rivers, lakes, seas, oceans	Water

Tabla 2.2 Definiciones de 'agua' en inglés

Clave	Definición libre del concepto 'limón'	Fase 2
1liml	It's a yellow fruit, like limes. Citrus. Used in cooking for sharpness	Lemon
2liml	A yellow citrus fruit. Sour tasting. Often used as an accompaniment to drinks	Lemon
3liml	a yellow citrus fruit with a bitter taste often sliced and put in drinks	Lemon
4liml	It's a citrus fruit, yellow, used with sugar on pancakes	Lemon
5liml	It's a yellow citrus fruit. Tastes bitter. Oval shaped	Lemon
6liml	A yellow sour fruit	lemon, grapefruit
7liml	A yellow citrus fruit	Lemon
8liml	Yellow, citrus, fruit	Lemon
9liml	Citrus fruit which is yellow	Lemon
10liml	Yellow citrus fruit	Lemon

Tabla 2.3 Definiciones de 'limón' en inglés

Clave	Definición libre del concepto 'ardilla'	Fase 2
1ardl	It's a little rodent and can be red or grey, it has a big bushy tail	squirrel
2ardl	A small rodent living in trees with a long bushy tail	squirrel
3ardl	A small rodent which lives in trees, collects nuts and has a bushy tail	squirrel
4ardl	Animal, grey/red, bushy tail, lives in trees, buries nuts	squirrel
5ardl	Small animal, lives in trees, eats acorns, has a bushy tail	squirrel
6ardl	Animal, bushy tail, eats nuts, builds nests in trees called dreys	squirrel
7ardl	Small funny animal with big, bushy tail, likes nuts, likes trees	squirrel
8ardl	Animal that lives in trees and collects acorns, has a long tail	squirrel
9ardl	A small- sized animal, habitat in trees	bird
10ardl	Small grey mammal, relative to the rodents, found in both countryside and town	rat?

Tabla 2.4 Definiciones de 'ardilla' en inglés

Clave	Definición libre del concepto 'cubeta'	Fase 2
1cubl	It's used for carrying water or other liquids. Also things like sand. A beach it is often with a spade and is used to make sandcastles	Bucket
2cubl	A vessel for carrying liquid or solids. Has a handle. Often seen in wells and accompanies a spade when see on the beach	Bucket
3cubl	item used to carry certain objects - mainly sand, water or soil. Children use them to make sandcastles at the beach	Bucket
4cubl	Used to carry water in/ make sandcastles; has a handle	Bucket
5cubl	Used to collect/ hold water, also on seaside to measure sand	Bucket
6cubl	A device made out of metal or plastic used mainly to carry water or other fluids	Bucket
7cubl	When cleaning the floor you put water in this and dip the mop into it	Bucket
8cubl	A device usually used to contain water	Bucket
9cubl	Item with a handle used for carrying things in, usually made of plastic	Bag
10cubl	A sort of container used to carry something	Bag

Tabla 2.5 Definiciones de 'cubeta' en inglés

Clave	Definición libre del concepto 'banca'	Fase 2
1banl	You can sit on it in the street or a park and they are made of wood	Bench
2 banl	A long hard seat for several persons on which the players on a sport team sit	Bench
3 banl	An object for sitting on, usually long which can seat many people	Bench
4 banl	Sit on it (a few people can) in parks, made of wood or iron	Bench
5 banl	Object used for sitting on. Often found in public places such as parks and gardens. Used to seat 1 or more people at a time	Bench
6 banl	Something you seat on, is longer than a chair, usually made of wood	Bench
7 banl	Long platform for sitting on (fit many people on one)	Bench
8 banl	Apparatus for sitting on, designed for more than one person, often found in parks	Bench
9 banl	A kind of seat found in parks, made of wood	bench
10 banl	A type of chair	Sofa

Tabla 2.6 Definiciones de 'banca' en inglés

Clave	Definición libre del concepto 'ropa'	Fase 2
1ropl	It describes the set of things you wear like trousers, jumpers, shirts, dresses, etc., anything in fact that you wear	clothes
2ropl	The collective name for the items which we wear, i.e. trousers, shirts, jumpers, etc.	clothes, attire, gear, rags
3ropl	Buy it in shops, keeps us warm, can be fashionable or unfashionable. Consists of things such as jumpers, trousers - the collective name	clothes, woolies
4ropl	Things used to cover up and keep warm, used by humans	clothes
5ropl	Items people can wear	clothes
6ropl	Items that we wear, to keep us warm	coat
7ropl	Garments worn on body	clothes
8ropl	What you wear on your body	clothes
9ropl	Items that we all wear	clothes
10ropl	Things you wear	clothes (or more specific? trousers, shoes, hats, etc.)

Tabla 2.7 Definiciones de 'ropa' en inglés

Clave	Definición libre del concepto 'eutanasia'	Fase 2
1eutl	It's when old or disabled people are killed legally and happens in the Netherlands	euthanasia
2 eutl	Act or practice of causing death painlessly	euthanasia
3 eutl	The decision taken by someone as to whether they die or not	suicide, euthanasia
4 eutl	Idea of turning off life -support machine; coma	euthanasia
5 eutl	Illegal act of helping somebody die who is terminally ill, who is in lot of pain, who needs help – cannot die alone	euthanasia
6 eutl	The right to decide to end one's own life	suicide
7 eutl	Right to die under special conditions	euthanasia
8 eutl	Ending life before natural end, to relieve from suffering due to illness	euthanasia
9eutl	Killing someone when they are terminally ill with their permission	Euthanasia
10eutl	Allowing people to die if they wish to, giving them the drugs or means to terminate their lives	euthanasia

Tabla 2.8 Definiciones de 'eutanasia' en inglés

Clave	Definición libre del concepto 'monopolio'	Fase 2
1monl	Boardgame with houses and hotels, properties, jail, "go"	monopoly
2monl	A board game involving the buying and selling of property	monopoly
3monl	Board game - buy/sell properties - contains Mayfair (Eng. Version) and park line; when you have complete control over something - normally business	monopoly
4 monl	A game where you buy land/railways and build houses/hotels or to hold a majority amount of something	monopoly
5 monl	Something to do with business, literally translated from Latin – one of many. Is also a board game	monopoly
6 monl	A company has control in the economy	monopoly
7 monl	When a company or individual has a major share of the market they have a ...	monopoly
8 monl	To have majority control, e.g. of a company's shares, or of a divisional field	monopoly
9 monl	An authoritative power	monopoly
10 monl	The interaction of economical communities	

Tabla 2.9 Definiciones de eutanasia en inglés

Clave	Definición libre del concepto 'huracán'	Fase 2
1hurl	A kind of severe storm which involves a lot of strong winds	hurricane, monsoon, gale, typhoon, tomado
2hurl	A violent tropical cyclone	typhoon, hurricane
3 hurl	Very strong winds of over 100 mph which cause great destruction	gales, tornados, hurricanes
4 hurl	Strong winds and rain, gale, destruction over large areas	hurricane, tomado
5 hurl	A very strong wind, very destructive, often named after people's first names, i.e. the weather researchers who first discover each one	hurricane?
6 hurl	Big amount of wind going from country to country, ruining everything as it goes!	cyclone, hurricane
7 hurl	Violent winds which can cause large scale destruction	Gale
8 hurl	Natural disaster, whipping wind, destroys anything in its path	Tomado
9 hurl	A type of typhoon	Hurricane
10 hurl	A very very strong wind	Hurricane

Tabla 2.10 Definiciones de 'huracán' en inglés

Clave	Definición libre del concepto 'barómetro'	Fase 2
1barl	Instrument which detects changes in atmospheric pressure	barometer
2barl	An instrument for measuring pressure	barometer
3 barl	Like thermometer but measures air pressure	barometer
4 barl	Used to measure air pressure	barometer
5 barl	A device to measure pressure	barometer
6 barl	Instrument used in physics to measure pressure	barometer
7 barl	An instrument used to measure pressure, used in physics	barometer
8 barl	A device used to measure air pressure	air gauge
9 barl	Measures pressure	
10 barl	Used to measure pressure, e.g. weather reports	

Tabla 2.11 Definiciones de 'barómetro' en inglés

Clave	Definición libre del concepto 'agua'	Fase 2
1aguE	es algo que es indispensable para el ser humano, el ser humano está hecho del 75% de este	
2aguE	liquido incoloro, inoloro e insípido	
3aguE	liquido vital para la supervivencia	
4aguE	Líquido vital para la vida, su fórmula química es H <sub>2</sub> O.	agua
5aguE	Se encuentra en estado líquido, se puede beber, es indispensable para los seres vivos	
6aguE	Líquido incoloro e insaboro que bebemos usualmente	agua
7aguE	Líquido inodoro, incoloro e insípido, necesario para la vida animal y vegetal	
8aguE	Tiene alta capacidad de solubilidad.	agua
9aguE	Elemento vital para la vida que abunda en todo el Planeta Tierra.	agua
10aguE	Contamos con una extensión muy grande en el mundo. Es imposible vivir sin ella. Se puede evaporar.	agua

Tabla 2.12 Definiciones de 'agua' en español

Clave	Definición libre del concepto 'limón'	Fase 2
1limE	cítrico verde, usado comúnmente con agua y azúcar	
2limE	fruta ácida	
3limE	fruta agria, redonda, pequeña	
4limE	fruta, cítrico verde utilizado para hacer agua	
5limE	Es un fruto color verde, cuyo sabor es muy ácido.	limón
6limE	Es una fruta ácida y de color verde.	limón
7limE	Fruta color verde, redonda, del tamaño de una pelota de golf. Se puede obtener de ella un jugo agrio al exprimirla.	limón
8limE	Fruto de forma redonda, color verdoso de sabor agrio.	limón
9limE	Verdura de color verde redonda, sabor agrio.	limón
10limE	Fruta ácida, color verde, con que se puede hacer agua; y se usa como condimento para las ensaladas.	limón

Tabla 2.13 Definiciones de 'limón' en español



Clave	Definición libre del concepto 'ardilla'	Fase 2
1ardE	peludo, habita en zonas boscosas. Mamífero	
2ardE	es algo que come nueces y cacahuates, pequeño y se caracteriza porque tiene sus dientes de adelante un poco más grandes que los demás	
3ardE	animal pequeño que come nueces y vive en los árboles, existen de color café, gris, negro, tiene una cola larga y esponjosa	
4ardE	vive en los bosques, come nueces y hay dos muy famosas en Disney	
5ardE	Roedor que trepa a los árboles, tiene cola larga y come bellotas.	ardilla
6ardE	Mamífero que suele encontrarse en los bosques, se alimenta de nueces y diversas semillas	ardilla
7ardE	Animal que vive en zonas selváticas o en bosque o arboleda, que come regularmente nueces o semillas, vive en los árboles y en CU hay muchas	ardilla
8ardE	Animal mamífero, roedor, habita en los árboles, se alimenta de nueces.	ardilla
9ardE	Animal pequeño y con una cola larga que habita en los bosques y se alimenta de semillas	ardilla
10ardE	Animal pequeño, es un roedor, gusta comer de nueces. Cuenta con una cola grande y ancha	ardilla

Tabla 2.14 Definiciones de 'ardilla' en español

Clave	Definición libre del concepto 'cubeta'	Fase 2
1cubE	recipiente grande para líquidos, generalmente de plástico, lo usamos como auxiliar para lavar el coche	
2cubE	recipiente de varios materiales para contener líquidos generalmente	
3cubE	lugar donde se puede almacenar líquidos o cualquier otra cosa con asa para sujetarse	
4cubE	artículo utilizado como contenedor de agua	
5cubE	Un objeto que nos sirve para acarrear agua.	cubeta
6cubE	Recipiente para poder transportar agua; se sujeta por una oreja o agarradera.	olla
7cubE	Recipiente de forma redonda.	olla
8cubE	Recipiente de forma variable utilizado para contener y transportar agua.	garrafón
9cubE	Es un objeto que sirve para acarrear agua.	cubeta
10cubE	Recipiente que cuenta con un asa y que se utiliza generalmente para desplazar agua	cubeta

Tabla 2.15 Definiciones de 'cubeta' en español

Clave	Definición libre del concepto 'banca'	Fase 2
1banE	la utilizamos para poner objetos, es de madera	
2banE	artículo utilizado para descansar	
3banE	Mueble que sirve para sentarse y puede o no tener respaldo	
4banE	sirve para sentarse	
5banE	Objeto que sirve para sentarse pero no tiene respaldo, generalmente se utiliza en la escuela	
6banE	mueble para sentarse se encuentra generalmente en sitios públicos	
7banE	Objeto que sirve para sentarse, generalmente la ponen en los parques y jardines	banca
8banE	Artículo donde nos sentamos. Se encuentran en los parques o a las afueras de algún edificio, iglesia o centro donde se reúnen varias personas	bancas
9banE	Mueble para sentarse de diversos materiales, se encuentra en parques y jardines	banca
10banE	Objeto en el que se puede sentar más de una persona, por lo regular existe en jardines	bancas

Tabla 2.16 Definiciones de 'banca' en español

Clave	Definición libre del concepto 'ropa'	Fase 2
1ropE	prenda tejida para uso personal	
2ropE	artículo con el cual se puede cubrir a un ser humano	
3ropE	artículos para cubrir el cuerpo humano del clima	
4ropE	Es todo aquello que usamos para salir a la calle. Lo llevamos puesto.	ropa
5ropE	Se dice de prendas hechas generalmente con tela o piel de animales para cubrir el cuerpo	abrigo
6ropE	Prenda de vestir utilizada para cubrir la desnudez de nuestro cuerpo	
7ropE	Actualmente fabricadas con fibras sintéticas predominantemente.	ropa
8ropE	Tela cortada y confeccionada de diferentes formas que sirve para cubrir el cuerpo humano	ropa
9ropE	Indumentaria que se usa para vestirse diariamente	
10ropE	Algo que nos cubre la piel del medio ambiente, con la que podemos cubrirnos del frío	chamarra

Tabla 2.17 Definiciones de 'ropa' en español

Clave	Definición libre del concepto 'eutanasia'	Fase 2
1eutE	decisión para morir por uno mismo, pero que se pide a un médico que te aplique	
2eutE	la practican los médicos para saber	
3eutE	inyección letal	
4eutE	facultad que se le da a una persona de quitarse la vida por padecer enfermedad terminal	
5eutE	Se practica cuando alguien presenta muerte cerebral para quitarle la vida.	eutanasia
6eutE	Dejar que muera o provocar la muerte de alguien, que por enfermedad tenga una muy mala calidad de vida, sin posibilidad de mejorarla.	eutanasia
7eutE	Acto voluntario en el cual una persona podría ser auxiliada para morir.	
8eutE	Proceso médico por el cual se provoca la muerte de pacientes terminales bajo consentimiento	eutanasia
9eutE	Método aceptado en pocos países que consiste en quitarle la vida a alguien durante una enfermedad terminal	eutanasia
10eutE	La llevan a cabo los médicos cuando llega a ser necesario, está relacionada con la muerte	eutanasia

Tabla 2.18 Definiciones de 'eutanasia' en español

Clave	Definición libre del concepto 'monopolio'	Fase 2
1monE	tener el control total de producción de cierto producto o servicio. Es el nombre de un juego de mesa	
2monE	es el dominio de un mercado o área en general por una sola persona, grupo o consorcio	
3monE	la forma en la que varias sociedades se unen por un solo fin principalmente	
4monE	mercado	
5monE	Decidir algo por sí mismo, sin cuenta de otras opiniones.	injusticia
6monE	Acaparar cualquier tipo de elementos y teniendo solamente el control de ellos, sin que nadie más tenga acceso a estos.	
7monE	Concentración de comercios	mercado
8monE	Predominio que ejerce cierta compañía comercial que acapara el mercado, y no da lugar a sus competidores	monopolio
9monE	Es la dominación del mercado por una empresa.	imperialismo
1monE	Es una situación en la que una empresa tiene total dominio del ramo al que pertenece	monopolio

Tabla 2.19 Definiciones de 'monopolio' en español

Clave	Definición libre del concepto 'huracán'	Fase 2
1hurE	remolino de agua y viento en forma circular, es en tierra	
2hurE	es un acontecimiento natural o fenómeno natural que provoca grandes cantidades de lluvia y aire	
3hurE	fenómeno climatológico que se presenta en los lugares que están cerca del mar, característico por sus fuertes lluvias y vientos	
4hurE	fenómeno meteorológico provocado por aire caliente que se combina con el aire frío. Hay una película con su nombre. Se lleva casas por el aire	
5hurE	Vientos a gran velocidad con lluvia en el mar.	huracán
6hurE	Fenómeno atmosférico que ocasiona movimiento de corrientes de aire con gran fuerza y velocidad alrededor de un centro u origen	
7hurE	Precipitación natural de niveles mayores (nivel 5) que ocasionan catástrofes, principalmente en zonas costeras. Esta precipitación se caracteriza porque a su centro se le llama ojo.	maremoto
8hurE	Fenómeno meteorológico provocado por el choque de dos frentes de aire con lluvia y vientos muy fuertes	tormenta
9hurE	Fenómeno destructivo que consiste en lluvias acompañadas de fuertes vientos	
10hurE	Fenómeno físico que puede llegar a ser peligroso, tiene la fuerza de arrancar casas frágiles u otro tipo de objetos	

Tabla 2.20 Definiciones de 'huracán' en español

Clave	Definición libre del concepto 'barómetro'	Fase 2
1barE	es un instrumento que mide la velocidad del aire	
2barE	instrumento de medición de la presión atmosférica	
3barE	instrumento de medición de presiones	
4barE	instrumento para medir presiones	
5barE	Es un instrumento que sirve para medir la presión atmosférica	barómetro
6barE	Aparato que se utiliza para medir la presión	
7barE	Instrumento de medición, cuya lectura nos indica la presión atmosférica del lugar en que se encuentra	barómetro
8barE	Aparato que recibe datos de medición y así poder interpretarlos.	computadora
9barE	Instrumento para medir la presión atmosférica	barómetro
10barE	Instrumento de medición	regla

Tabla 2.21 Definiciones de 'barómetro' en español



## **Apéndice del Capítulo 3**

### **Conceptualización de búsquedas en lenguaje natural**

En este apéndice se podrán consultar:

- Las tablas que ilustran las clasificaciones de las definiciones del corpus.
- Las tablas en las que se marca el tipo de información por tipo en las definiciones del corpus.

Clave	Definición libre del concepto 'agua'	Tipo de Definición
1aguE	es algo que es indispensable para el ser humano, el ser humano está hecho del 75% de este	Implicación
2aguE	liquido incoloro, inoloro e insípido	Síntesis
3aguE	liquido vital para la supervivencia	Síntesis
4aguE	Líquido vital para la vida, su fórmula química es H <sub>2</sub> O.	Síntesis/ Implicación
5aguE	Se encuentra en estado líquido, se puede beber, es indispensable para los seres vivos	Síntesis/ Implicación
6aguE	Líquido incoloro e insaboro que bebemos usualmente	Síntesis/ Implicación
7aguE	Líquido inodoro, incoloro e insípido, necesario para la vida animal y vegetal	Síntesis/ Implicación
8aguE	Tiene alta capacidad de solubilidad.	Implicación
9aguE	Elemento vital para la vida que abunda en todo el Planeta Tierra.	Síntesis/ Implicación
10aguE	Contamos con una extensión muy grande en el mundo. Es imposible vivir sin ella. Se puede evaporar.	Implicación

Tabla 3.30 Tipos de definición de 'agua' en español

Clave	Definición libre del concepto 'limón'	Tipo de definición
1limE	cítrico verde, usado comúnmente con agua y azúcar	Síntesis/ Implicación
2limE	fruta ácida	Síntesis
3limE	fruta agria, redonda, pequeña	Síntesis
4limE	fruta, cítrico verde utilizado para hacer agua	Síntesis/ Implicación
5limE	Es un fruto color verde, cuyo sabor es muy ácido.	Síntesis
6limE	Es una fruta ácida y de color verde.	Síntesis
7limE	Fruta color verde, redonda, del tamaño de una pelota de golf. Se puede obtener de ella un jugo agrio al exprimirla.	Síntesis/ Implicación
8limE	Fruto de forma redonda, color verdoso de sabor agrio.	Síntesis
9limE	Verdura de color verde redonda, sabor agrio.	Síntesis
10limE	Fruta ácida, color verde, con que se puede hacer agua; y se usa como condimento para las ensaladas.	Síntesis/ Implicación

Tabla 3.31 Tipos de definición de 'limón' en español

Clave	Definición libre del concepto 'ardilla'	Tipos de definición
1ardE	peludo, habita en zonas boscosas. Mamífero	Síntesis
2ardE	es algo que come nueces y cacahuates, pequeño y se caracteriza porque tiene sus dientes de adelante un poco más grandes que los demás	Síntesis
3ardE	animal pequeño que come nueces y vive en los árboles, existen de color café, gris, negro, tiene una cola larga y esponjosa	Síntesis
4ardE	vive en los bosques, come nueces y hay dos muy famosas en Disney	Síntesis/Extensiva
5ardE	Roedor que trepa a los árboles, tiene cola larga y come bellotas.	Síntesis
6ardE	Mamífero que suele encontrarse en los bosques, se alimenta de nueces y diversas semillas	Síntesis
7ardE	Animal que vive en zonas selváticas o en bosque o arboleda, que come regularmente nueces o semillas, vive en los árboles y en CU hay muchas	Síntesis/ Implicación
8ardE	Animal mamífero, roedor, habita en los árboles, se alimenta de nueces.	Síntesis
9ardE	Animal pequeño y con una cola larga que habita en los bosques y se alimenta de semillas	Síntesis
10ardE	Animal pequeño, es un roedor, gusta comer de nueces. Cuenta con una cola grande y ancha	Síntesis

Tabla 3.32 Tipos de definición de 'ardilla' en español

Clave	Definición libre del concepto 'cubeta'	Tipos de definición
1cubE	recipiente grande para líquidos, generalmente de plástico, lo usamos como auxiliar para lavar el coche	Síntesis/ Implicación
2cubE	recipiente de varios materiales para contener líquidos generalmente	Síntesis/ Implicación
3cubE	lugar donde se puede almacenar líquidos o cualquier otra cosa con asa para sujetarse	Síntesis/ Implicación
4cubE	artículo utilizado como contenedor de agua	Síntesis
	Un objeto que nos sirve para acarrear agua.	Síntesis/ Implicación
6cubE	Recipiente para poder transportar agua; se sujeta por una oreja o agarradera.	Síntesis
7cubE	Recipiente de forma redonda.	Síntesis
8cubE	Recipiente de forma variable utilizado para contener y transportar agua.	Síntesis
9cubE	Es un objeto que sirve para acarrear agua.	Síntesis
10cubE	Recipiente que cuenta con un asa y que se utiliza generalmente para desplazar agua	Síntesis/ Implicación

Tabla 3.33 Tipos de definición de 'cubeta' en español

Clave	Definición libre del concepto 'banca'	Tipos de definición
1banE	la utilizamos para poner objetos, es de madera	Síntesis/ Implicación
2banE	artículo utilizado para descansar	Síntesis
3banE	Mueble que sirve para sentarse y puede o no tener respaldo.	Síntesis
4banE	sirve para sentarse	Implicación
5banE	objeto que sirve para sentarse pero no tiene respaldo, generalmente se utiliza en la escuela	Síntesis/ Implicación
6banE	mueble para sentarse se encuentra generalmente en sitios públicos	Síntesis
7banE	Objeto que sirve para sentarse, generalmente la ponen en los parques y jardines	Síntesis/ Implicación
8banE	Artículo donde nos sentamos. Se encuentran en los parques o a las afueras de algún edificio, iglesia o centro donde se reúnen varias personas	Síntesis/ Implicación
9banE	Mueble para sentarse de diversos materiales, se encuentra en parques y jardines	Síntesis/ Implicación
10banE	Objeto en el que se puede sentar más de una persona, por lo regular existe en jardines	Síntesis/ Implicación

Tabla 3.34 Tipos de definición de 'banca' en español

Clave	Definición libre del concepto 'ropa'	Tipos de definición
1ropE	prenda tejida para uso personal	Síntesis
2ropE	artículo con el cual se puede cubrir a un ser humano	Síntesis/ Implicación
3ropE	artículos para cubrir el cuerpo humano del clima	Síntesis
4ropE	Es todo aquello que usamos para salir a la calle. Lo llevamos puesto.	Síntesis/ Implicación
5ropE	Se dice de prendas hechas generalmente con tela o piel de animales para cubrir el cuerpo	Síntesis
6ropE	Prenda de vestir utilizada para cubrir la desnudez de nuestro cuerpo.	Síntesis/ Implicación
7ropE	Actualmente fabricadas con fibras sintéticas predominantemente.	Implicación
8ropE	Tela cortada y confeccionada de diferentes formas que sirve para cubrir el cuerpo humano	Síntesis
9ropE	Indumentaria que se usa para vestirse diariamente	Síntesis/ Implicación
10ropE	Algo que nos cubre la piel del medio ambiente, con la que podemos cubrimos del frío	Implicación

Tabla 3.35 Tipos de definición de 'ropa' en español



Clave	Definición libre del concepto 'eutanasia'	Tipos de definición
1eutE	decisión para morir por uno mismo, pero que se pide a un médico que te aplique	Síntesis/ Implicación
2eutE	la practican los médicos para saber	implicación
3eutE	inyección letal	Síntesis
4eutE	facultad que se le da a una persona de quitarse la vida por padecer enfermedad terminal	Síntesis/ Implicación
5eutE	Se practica cuando alguien presenta muerte cerebral para quitarle la vida.	implicación
6eutE	Dejar que muera o provocar la muerte de alguien, que por enfermedad tenga una muy mala calidad de vida, sin posibilidad de mejorarla.	implicación
7eutE	Acto voluntario en el cual una persona podría ser auxiliada para morir.	Síntesis/ Implicación
8eutE	Proceso médico por el cual se provoca la muerte de pacientes terminales bajo consentimiento	Síntesis
9eutE	Método aceptado en pocos países que consiste en quitarle la vida a alguien durante una enfermedad terminal	Síntesis/ Implicación
10eutE	La llevan a cabo los médicos cuando llega a ser necesario, está relacionada con la muerte	implicación

Tabla 3.36 Tipos de definición de 'eutanasia' en español

Clave	Definición libre del concepto 'monopolio'	Tipos de definición
1monE	tener el control total de producción de cierto producto o servicio. Es el nombre de un juego de mesa	Implicación/ Estipulación
2monE	es el dominio de un mercado o área en general por una sola persona, grupo o consorcio	Síntesis
3monE	la forma en la que varias sociedades se unen por un solo fin principalmente	Síntesis
4monE	mercado	Sinonímica
5monE	Decidir algo por sí mismo, sin cuenta de otras opiniones.	Implicación
6monE	Acaparar cualquier tipo de elementos y teniendo solamente el control de ellos, sin que nadie más tenga acceso a estos.	Implicación
7monE	Concentración de comercios	Síntesis
8monE	Predominio que ejerce cierta compañía comercial que acapara el mercado, y no da lugar a sus competidores	Síntesis
9monE	Es la dominación del mercado por una empresa.	Síntesis
10monE	Es una situación en la que una empresa tiene total dominio del ramo al que pertenece	Síntesis/ Implicación

Tabla 3.37 Tipos de definición de 'monopolio' en español

Clave	Definición libre del concepto 'huracán'	Tipos de definición
1hurE	remolino de agua y viento en forma circular, es en tierra	Síntesis
2hurE	es un acontecimiento natural o fenómeno natural que provoca grandes cantidades de lluvia y aire	Síntesis/ Implicación
3hurE	fenómeno climatológico que se presenta en los lugares que están cerca del mar, característico por sus fuertes lluvias y vientos	Síntesis
4hurE	fenómeno meteorológico provocado por aire caliente que se combina con el aire frío. Hay una película con su nombre. Se lleva casas por el aire	Síntesis/ Implicación
5hurE	Vientos a gran velocidad con lluvia en el mar.	Síntesis
6hurE	Fenómeno atmosférico que ocasiona movimiento de corrientes de aire con gran fuerza y velocidad alrededor de un centro u origen	Síntesis
7hurE	Precipitación natural de niveles mayores (nivel 5) que ocasionan catástrofes, principalmente en zonas costeras. Esta precipitación se caracteriza porque a su centro se le llama ojo.	Síntesis/ Implicación
8hurE	Fenómeno meteorológico provocado por el choque de dos frentes de aire con lluvia y vientos muy fuertes	Síntesis/ Implicación
9hurE	Fenómeno destructivo que consiste en lluvias acompañadas de fuertes vientos	Síntesis/ Implicación
10hurE	Fenómeno físico que puede llegar a ser peligroso, tiene la fuerza de arrancar casas frágiles u otro tipo de objetos	Síntesis/ Implicación

Tabla 3.38 Tipos de definición de 'huracán' en español

Clave	Definición libre del concepto 'barómetro'	Tipos de definición
1barE	es un instrumento que mide la velocidad del aire	Síntesis
2barE	instrumento de medición de la presión atmosférica	Síntesis
3barE	instrumento de medición de presiones	Síntesis
4barE	instrumento para medir presiones	Síntesis
5barE	Es un instrumento que sirve para medir la presión atmosférica.	Síntesis
6barE	Aparato que se utiliza para medir la presión	Síntesis
7barE	Instrumento de medición, cuya lectura nos indica la presión atmosférica del lugar en que se encuentra	Síntesis/ Implicación
8barE	Aparato que recibe datos de medición y así poder interpretarlos.	Síntesis
9barE	Instrumento para medir la presión atmosférica	Síntesis
10barE	Instrumento de medición	Síntesis

Tabla 3.39 Tipos de definición 'barómetro' en español

Clave	Definición libre del concepto 'agua'	Tipos de definición
1agul	It's a clear liquid that you get from a tap	Ostensión/Implicación
2agul	The colourless transparent liquid occurring on rivers	Síntesis/ Implicación
3 agul	A clear, neutral liquid that surrounds us everywhere	Síntesis/ Implicación
4 agul	Liquid, clear, drinkable – constituents are hydrogen and oxygen	Síntesis
5 agul	Liquid, clear, H <sub>2</sub> O	Síntesis
6 agul	Liquid form, scientific term H <sub>2</sub> O	Síntesis
7 agul	Liquid, freezes at 0°C	Síntesis
8 agul	Liquid, clear, boils at 100°C, freezes at 0°C	Síntesis
9 agul	Fluid, clear, tasteless, colourless	Síntesis
10 agul	Wash with it; drink it; used for dilution; H <sub>2</sub> O; found in springs, rivers, lakes, seas, oceans	Síntesis/ Implicación

Tabla 3.40 Tipos de definición 'agua' en inglés

Clave	Definición libre del concepto 'limón'	Clas. Definición
1liml	It's a yellow fruit, like limes. Citrus. Used in cooking for sharpness	Ostensión/ Analogía/ Síntesis/ Implicación
2 liml	A yellow citrus fruit. Sour tasting. Often used as an accompaniment to drinks	Síntesis/ Implicación
3 liml	a yellow citrus fruit with a bitter taste often sliced and put in drinks	Síntesis/ Implicación
4 liml	It's a citrus fruit, yellow, used with sugar on pancakes	Ostensión/ Síntesis/ Implicación
5 liml	It's a yellow citrus fruit. Tastes bitter. Oval shaped	Ostensión/ Implicación
6 liml	A yellow sour fruit	Síntesis
7 liml	A yellow citrus fruit	Síntesis
8 liml	Yellow, citrus, fruit	Síntesis
9 liml	Citrus fruit which is yellow	Síntesis
10 liml	Yellow citrus fruit	Síntesis

Tabla 3.41 Tipos de definición de 'limón' en inglés

Clave	Definición libre del concepto 'ardilla'	Tipos de definición
1ardl	It's a little rodent and can be red or grey, it has a big bushy tail	Ostensión/ Síntesis
2ardl	A small rodent living in trees with a long bushy tail	Síntesis
3ardl	A small rodent which lives in trees, collects nuts and has a bushy tail	Síntesis
4ardl	Animal, grey/red, bushy tail, lives in trees, buries nuts	Síntesis
5ardl	Small animal, lives in trees, eats acorns, has a bushy tail	Síntesis
6ardl	Animal, bushy tail, eats nuts, builds nests in trees called dreys	Síntesis
7ardl	Small funny animal with big, bushy tail, likes nuts, likes trees	Síntesis
8ardl	Animal that lives in trees and collects acorns, has a long tail	Síntesis
9ardl	A small-sized animal, habitat in trees	Síntesis
10ardl	Small grey mammal, relative to the rodent, found in both countryside and town	Síntesis/ Implicación

Tabla 3.42 Tipos de definición de 'ardilla' en inglés

Clave	Definición libre del concepto 'cubeta'	Tipos de definición
1cubl	It's used for carrying water or other liquids. Also things like sand. A beach it is often with a spade and is used to make sandcastles	Ostensiva/ Síntesis/ Implicación
2cubl	A vessel for carrying liquid or solids. Has a handle. Often seen in wells and accompanies a spade when seen on the beach	Síntesis/ Implicación
3cubl	Item used to carry certain objects - mainly sand, water or soil. Children use them to make sandcastles at the beach	Síntesis/ Implicación
4cubl	Used to carry water in/make sandcastles; has a handle	Implicación
5cubl	Used to collect/hold water, also on seaside to measure sand	Implicación
6cubl	A device made out of metal or plastic used mainly to carry water or other fluids	Síntesis
7cubl	When cleaning the floor you put water in this and dip the mop into it	Implicación
8cubl	A device usually used to contain water	Síntesis
9cubl	Item with a handle used for carrying things in, usually made of plastic	Síntesis
10cubl	A sort of container used to carry something	Síntesis

Tabla 3.43 Tipos de definición de 'cubeta' en inglés

Clave	Definición libre del concepto 'banca'	Tipos de definición
1banl	You can sit on it in the street or a park and they are made of wood	Implicación
2 banl	A long hard seat for several persons on which the players on a sport team sit	Síntesis/ Implicación
3 banl	An object for sitting on, usually long which can seat many people	Síntesis/ Implicación
4 banl	Sit on it (a few people can) in parks, made of wood or iron	Implicación
5 banl	Object used for sitting on. Often found in public places such as parks and gardens. Used to seat 1 or more people at a time	Síntesis/ Implicación
6 banl	Something you seat on, is longer than a chair, usually made of wood	Síntesis/ Implicación
7 banl	Long platform for sitting on (fit many people on one)	Síntesis/ Implicación
8 banl	Apparatus for sitting on, designed for more than one person, often found in parks	Síntesis/ Implicación
9 banl	A kind of seat found in parks, made of wood	Síntesis/ Implicación
10 banl	A type of chair	Analogía

Tabla 3.44 Tipos de definición de 'banca' en inglés

Clave	Definición libre del concepto 'ropa'	Tipos de definición
1ropl	It describes the set of things you wear like trousers, jumpers, shirts, dresses, etc., anything in fact that you wear	Ostensiva/ Síntesis/ Extensión / Implicación
2ropl	The collective name for the items which we wear; i.e. trousers, shirts, jumpers, etc.	Síntesis/ Extensión
3ropl	Buy it in shops, keeps us warm, can be fashionable or unfashionable. Consists of things such as jumpers, trousers - the collective name	Implicación/ Extensión
4ropl	Things used to cover up and keep warm, used by humans	Implicación
5ropl	Items people can wear	Implicación
6ropl	Items that we wear, to keep us warm	Implicación
7ropl	Garments worn on body	Síntesis
8ropl	What you wear on your body	Implicación
9ropl	Items that we all wear	Implicación
10ropl	Things you wear	Implicación

Tabla 3.45 Tipos de definición de 'ropa' en inglés

Clave	Definición libre del concepto 'eutanasia'	Tipos de definición
1eutl	It's when old or disabled people are killed legally and happens in the Netherlands	Ostensión/ Implicación
2 eutl	Act or practice of causing death painlessly	Síntesis
3 eutl	The decision taken by someone as to whether they die or not	Síntesis/ Implicación
4 eutl	Idea of turning off life-support machine; coma	Síntesis/ Implicación
5 eutl	Illegal act of helping somebody die who is terminally ill, who is in lot of pain, who needs help – cannot die alone	Síntesis/ Implicación
6 eutl	The right to decide to end one's own life	Síntesis/ Implicación
7 eutl	Right to die under special conditions	Síntesis
8 eutl	Ending life before natural end, to relieve from suffering due to illness	Síntesis/ Implicación
9eutl	Killing someone when they are terminally ill with their permission	Síntesis/ Implicación
10eutl	Allowing people to die if they wish to, giving them the drugs or means to terminate their lives	Implicación

Tabla 3.46 Tipos de definición de 'eutanasia' en inglés

Clave	Definición libre del concepto 'monopolio'	Tipos de definición
1monl	Boardgame with houses and hotels, properties, jail, "go"	Síntesis
2monl	A board game involving the buying and selling of property	Síntesis /Extensión
3monl	Board game - buy/sell properties - contains Mayfair (Eng. Version) and park line; when you have complete control over something - normally business	Síntesis/ Extensión/ Implicación
4 monl	A game where you buy land/railways and build houses/hotels or to hold a majority amount of something	Síntesis/ Implicación
5 monl	Something to do with business, literally translated from Latin – one of many. Is also a board game	Síntesis/Estipulación
6 monl	A company has control in the economy	Implicación
7 monl	When a company or individual has a major share of the market they have a ...	Implicación
8 monl	To have majority control, e.g. of a company's shares, or of a divisional field	Implicación
9 monl	An authoritative power	Síntesis
10 monl	The interaction of economical communities	Síntesis

Tabla 3.47 Tipos de definición de 'monopolio' en inglés

Clave	Definición libre del concepto 'huracán'	Tipos de definición
1hurl	A kind of severe storm which involves a lot of strong winds	Síntesis/ Implicación
2hurl	A violent tropical cyclone	Síntesis
3 hurl	Very strong winds of over 100 mph which cause great destruction	Síntesis/ Implicación
4 hurl	Strong winds and rain, gale, destruction over large areas	Síntesis/ Implicación
5 hurl	A very strong wind, very destructive, often named after people's first names, i.e. the weather researchers who first discover each one	Síntesis /Extensión/ Implicación
6 hurl	Big amount of wind going from country to country, ruining everything as it goes!	Síntesis/ Implicación
7 hurl	Violent winds which can cause large scale destruction	Síntesis/ Implicación
8 hurl	Natural disaster, whipping wind, destroys anything in its path	Síntesis/ Implicación
9 hurl	A type of typhoon	Analogía
10 hurl	A very very strong wind	Síntesis/ Implicación

Tabla 3.48 Tipos de definición de 'huracán' en inglés

Clave	Definición libre del concepto 'barómetro'	Tipos de definición
1barl	Instrument which detects changes in atmospheric pressure	Síntesis
2barl	An instrument for measuring pressure	Síntesis
3 barl	Like thermometer but measures air pressure	Analogía
4 barl	Used to measure air pressure	Implicación
5 barl	A device to measure pressure	Síntesis
6 barl	Instrument used in physics to measure pressure	Síntesis
7 barl	An instrument used to measure pressure, used in physics	Síntesis
8 barl	A device used to measure air pressure	Síntesis
9 barl	Measures pressure	Implicación
10 barl	Used to measure pressure, e.g. weather reports	Implicación /Extensión

Tabla 3.49 Tipos de definición de 'barómetro' en inglés

Información Género próximo

Clave	Género Próximo
1liml	fruit
2liml	fruit
3liml	fruit
4liml	fruit
5liml	fruit
6liml	fruit
7liml	fruit
8liml	fruit
9liml	fruit
10liml	fruit
1limE	cítrico
2limE	fruta
3limE	fruta
4limE	fruta
5limE	fruto
6limE	fruta
7limE	fruta
8limE	fruto
9limE	verdura
10limE	fruta

Tabla 3.50 limón'

Clave	Género próximo
1ardl	rodent
2ardl	rodent
3ardl	rodent
4ardl	animal
5ardl	animal
6ardl	animal
7ardl	animal
8ardl	animal
9ardl	animal
10ardl	animal
1ardE	mamífero
2ardE	algo
3ardE	animal
4ardE	
5ardE	roedor
6ardE	mamífero
7ardE	animal
8ardE	animal
9ardE	animal
10ardE	animal

Tabla 3.51 ardilla'

Clave	Género próximo
1monl	boardgame
2monl	board game
3monl	board game
4monl	game
5monl	something
6monl	
7monl	
8monl	
9monl	power
10monl	interaction
1monE	
2monE	dominio
3monE	forma
4monE	mercado
5monE	
6monE	
7monE	concentración
8monE	predominio
9monE	dominación
10monE	situación

Tabla 3.52 monopolio'



Información Diferencia específica

Clave	diferencia específica
1agul	clear
2agul	transparent
3agul	clear
4agul	clear
5agul	clear
6agul	
7agul	
8agul	clear
9agul	clear
10agul	
1aguE	indispensable
2aguE	
3aguE	vital
4aguE	vital
5aguE	indispensable
6aguE	
7aguE	necesario
8aguE	
9aguE	vital
10aguE	imposible vivir sin ella

Tabla 3.53 Agua

Clave	diferencia específica
1ropl	
2ropl	
3ropl	fashionable or unfashionable
4ropl	
5ropl	
6ropl	
7ropl	
8ropl	
9ropl	
10ropl	
1ropE	tejida
2ropE	
3ropE	
4ropE	
5ropE	hechas con tela o piel de animales
6ropE	de vestir
7ropE	fabricadas con fibras sintéticas
8ropE	cortada y confeccionada de diferentes formas
9ropE	
10ropE	

Tabla 3.54 Ropa

Clave	diferencia específica
1cubl	
2cubl	a handle
3cubl	
4cubl	a handle
5cubl	
6cubl	
7cubl	
8cubl	
9cubl	with a handle
10cubl	
1cubE	
2cubE	
3cubE	con asa
4cubE	
5cubE	
6cubE	una oreja, o agarradera
7cubE	
8cubE	
9cubE	
10cubE	un asa

Tabla 3.55 Cubeta

Información Datos sensoriales

Clave	datos sensoriales
1limI	yellow
2limI	yellow
3limI	yellow
4limI	yellow
5limI	yellow
6limI	yellow
7limI	yellow
8limI	yellow
9limI	yellow
10limI	yellow
1limE	verde
2limE	
3limE	
4limE	verde
5limE	color verde
6limE	color verde
7limE	verde
8limE	color verdoso
9limE	color verde
10limE	color verde

Tabla 3.56 Limón

Clave	datos sensoriales
1ardI	big
2ardI	long
3ardI	
4ardI	
5ardI	
6ardI	
7ardI	big
8ardI	long
9ardI	
10ardI	
1ardE	
2ardE	
3ardE	larga
4ardE	
5ardE	larga
6ardE	
7ardE	
8ardE	
9ardE	larga
10ardE	grande

Tabla 3.57 Ardilla (la cola)

Clave	datos sensoriales
1cubI	
2cubI	
3cubI	
4cubI	
5cubI	
6cubI	of metal or plastic
7cubI	
8cubI	
9cubI	plastic
10cubI	
1cubE	grande
2cubE	
3cubE	
4cubE	
5cubE	
6cubE	
7cubE	forma redonda
8cubE	forma variable
9cubE	
10cubE	

Tabla 3.58 Cubeta

Información Características científicas

Clave	Característica científica
1agul	
2agul	
3agul	neutral liquid
4agul	constituents are hydrogen and oxygen
5agul	H <sub>2</sub> O
6agul	scientific term H <sub>2</sub> O
7agul	freezes at 0°C
8agul	boils at 100°C, freezes at 0°C
9agul	
10agul	used for dilution, H <sub>2</sub> O
1aguE	el ser humano está hecho del 75% de éste
2aguE	
3aguE	
4aguE	Su fórmula química es H <sub>2</sub> O
5aguE	
6aguE	
7aguE	
8aguE	Tiene alta capacidad de solubilidad
9aguE	
10aguE	Se puede evaporar

Tabla 3.59 Agua

Clave	Característica científica
1liml	fruit. Citrus
2liml	citrus fruit
3liml	citrus fruit
4liml	citrus fruit
5liml	citrus fruit
6liml	fruit
7liml	citrus fruit
8liml	citrus fruit
9liml	citrus fruit
10liml	citrus fruit
1limE	citrico
2limE	fruta
3limE	fruta
4limE	fruta, citrico
5limE	fruto
6limE	fruta
7limE	fruta
8limE	fruto
9limE	verdura
10limE	fruta

Tabla 3.60 Limón

Clave	Característica científica
1ardl	rodent
2ardl	rodent
3ardl	rodent
4ardl	animal
5ardl	animal
6ardl	animal
7ardl	animal
8ardl	animal
9ardl	animal
10ardl	mammal, rodent
1ardE	Mamifero
2ardE	
3ardE	Animal
4ardE	
5ardE	Roedor
6ardE	Mamifero
7ardE	Animal
8ardE	Animal, mamifero, roedor
9ardE	Animal
10ardE	Animal, roedor

Tabla 3.61 Ardilla

Información verbo relacionado y complemento

Clave	Verbo	Complemento
1barI	detects	changes in atmospheric pressure
2barI	for measuring	pressure
3barI	measures	air pressure
4barI	used to measure	air pressure
5barI	to measure	pressure
6barI	used...to measure	
7barI	used to measure	
8barI	used to measure	air pressure
9barI	measures	pressure
10barI	used to measure	pressure
1barE	mide	la velocidad del aire
2barE	medición	de la presión atmosférica
3barE	medición	de presiones
4barE	medir	presiones
5barE	medir	la presión atmosférica
6barE	medir	la presión
7barE	medición cuya lectura nos indica	la presión atmosférica del lugar en que se encuentra
8barE	que recibe	datos de medición y así poder interpretarlos
9barE	medir	la presión atmosférica
10barE	medición	

Tabla 3.62 Barómetro

Clave	Verbo	Complemento
1ardI		
2ardI	living	in trees
3ardI	lives	in trees
4ardI	lives	in trees
5ardI	lives	in trees
6ardI	builds	nets in trees called dreys
7ardI	likes	trees
8ardI	lives	in trees
9ardI	(habitat)	in trees
10ardI	found	in both countryside and town
1ardE	habita	en zonas boscosas
2ardE		
3ardE	vive	en los árboles
4ardE	vive	en los bosques
5ardE	trepa	los árboles
6ardE	encontrarse	en los bosques
7ardE	vive	en zonas selváticas o en bosque o arboleda
8ardE	habita	en los árboles
9ardE	habita	en los bosques
10ardE		

Tabla 3.63 Ardilla

Clave	Verbo	Complemento
1monI		
2monI	buying and selling	of the property
3monI	buy/ sell	
4monI	buy	land/railways
5monI		
6monI		
7monI		
8monI		
9monI		
10monI		
1monE	tener	el control total
2monE		
3monE	se unen	
4monE		
5monE	decidir	algo
6monE	acaparar	cualquier tipo
7monE		
8monE	que ejerce, que acapara	el mercado
9monE		
10monE	tiene	total dominio

Tabla 3.64 Monopolio

Información Marca situacional

Clave	Marca situacional
1cubl	
2cubl	when
3cubl	
4cubl	
5cubl	
6cubl	
7cubl	when
8cubl	
9cubl	
10cubl	
1cubE	
2cubE	
3cubE	donde se puede
4cubE	
5cubE	
6cubE	
7cubE	
8cubE	
9cubE	
10cubE	

Tabla 3.65 Cubeta

Clave	Marca situacional
1monl	
2monl	
3monl	when
4 monl	where
5 monl	
6 monl	
7 monl	when
8 monl	
9 monl	
10 monl	
1monE	
2monE	
3monE	en la que
4monE	
5monE	
6monE	
7monE	
8monE	
9monE	
10monE	en la que

Tabla 3.66 Monopolio

Clave	Marca situacional
1eutl	when
2 eutl	
3 eutl	
4 eutl	
5 eutl	who
6 eutl	
7 eutl	
8 eutl	due
9eutl	when
10eut10	if
1eutE	pero que (se)
2eutE	
3eutE	
4eutE	por (padecer)
5eutE	cuando
6eutE	que por (enfermedad)
7eutE	en el cual
8eutE	por el cual
9eutE	que consiste en
10eutE	cuando

Tabla 3.67 Eutanasia



## **Apéndice del Capítulo 4**

### **Evaluación de los diccionarios electrónicos**

En este apéndice se podrán consultar:

- Las tablas que muestran los resultados de la consulta en los diccionarios *Vox*, *Longman*, *Onelook* y *Wordsmyth*.



Clave	Definición libre de 'agua'	Forma introducida	Fase 2	1º Lugar	2ºLugar	Total	L.Ref	F	P	G
1aguE	es algo que es indispensable para el <b>ser humano</b> , el <b>ser humano</b> está hecho del 75% de este	SER .Y HUMANO		agua	alma	23	1	0,00000	1,00000	0,04348
2aguE	<b>líquido incoloro</b> , inoloro e <b>insípido</b>	LÍQUIDO .Y INCOLORO .Y INSÍPIDO		agua		1	1	0,00000	1,00000	1,00000
3aguE	<b>líquido vital</b> para la supervivencia	LÍQUIDO		abducción	absorciómetro	514	25	0,04678	1,00000	0,00195
4aguE	<b>Líquido vital</b> para la vida, su fórmula química es H2O.	LÍQUIDO	agua	abducción	absorciómetro	514	25	0,04678	1,00000	0,00195
5aguE	Se encuentra en <b>estado líquido</b> , se <b>puede beber</b> , es indispensable para los seres vivos	ESTADO .Y LÍQUIDO .Y PUEDE .Y BEBER		agua		1	1	0,00000	1,00000	1,00000
6aguE	<b>Líquido incoloro</b> e insaboro que bebemos usualmente	LÍQUIDO .Y INCOLORO	agua	acetona	acrílico	21	4	0,15000	1,00000	0,04762
7aguE	<b>Líquido inodoro</b> , <b>incoloro</b> e <b>insípido</b> , necesario para la vida animal y vegetal	LÍQUIDO .Y INODORO .Y INCOLORO .Y INSÍPIDO		agua		1	1	0,00000	1,00000	1,00000
8aguE	<b>Tiene</b> alta capacidad de solubilidad.	TIENE	agua	ab-	abecero	3679	115	0,03100	1,00000	0,00027
9aguE	Elemento vital para la vida que abunda en todo el Planeta Tierra.	0	agua	0	0	0	0	1,00000	0,00000	0,00000
10aguE	Contamos con una extensión muy grande en el mundo. Es imposible vivir sin ella. Se <b>puede</b> evaporar.	PUEDE	agua	abarrancadero	abatible	1502	45	0,02931	1,00000	0,00067

Tabla 4.7 Consulta de 'agua' en el diccionario Vox

Clave	Definición libre de 'limón'	Forma introducida	Fase 2	1° Lugar	2°Lugar	Total	L.Ref	F	P	G
1limE	citrico verde, usado comúnmente con agua y azúcar	0		0	0	0	0	1,00000	0,00000	0,00000
2limE	fruta ácida	0		0	0	0	0	1,00000	0,00000	0,00000
3limE	fruta agria, redonda, pequeña	0		0	0	0	0	1,00000	0,00000	0,00000
4limE	fruta, citrico verde utilizado para hacer agua	0		0	0	0	0	1,00000	0,00000	0,00000
5limE	Es un <b>fruto</b> color verde, cuyo sabor es muy ácido.	FRUTO	limón	ababábite	abatido	895	519	0,57942	1,00000	0,00112
6limE	Es una fruta ácida y de color verde.	0	limón	0	0	0	0	1,00000	0,00000	0,00000
7limE	Fruta color verde, redonda, del tamaño de una pelota de golf. Se puede obtener de ella un jugo agrio al exprimirla	0	limón	0	0	0	0	1,00000	0,00000	0,00000
8limE	<b>Fruto</b> de forma redonda, color verdoso de sabor agrio.	FRUTO	limón	ababábite	abatido	895	519	0,57942	1,00000	0,00112
9limE	Verdura de color verde redonda, sabor agrio.	0	limón	0	0	0	0	1,00000	0,00000	0,00000
10limE	Fruta ácida, color verde, con que se puede hacer agua; y se usa como condimento para las ensaladas.	0	limón	0	0	0	0	1,00000	0,00000	0,00000

Tabla 4.8 Consulta de 'limón' en el Vox

Clave	Definición libre de 'ardilla'	Forma introducida	Fase 2	1° Lugar	2°Lugar	Total	L.Ref	F	P	G
1ardE	peludo, habita en zonas boscosas. <b>Mamifero</b>	MAMIFERO		acleido	adive	296	16	0,05085	1,00000	0,00338
2ardE	es algo que come nueces y cacahuates, pequeño y se caracteriza porque tiene sus dientes de adelante un poco más grandes que los demás	0		0	0	0	0	1,00000	0,00000	0,00000
3ardE	animal pequeño que come nueces y vive en los árboles, existen de color café, gris, negro, tiene una <b>cola larga</b> y esponjosa	COLA .Y LARGA		águila	alondra	68	5	0,05970	1,00000	0,01471
4ardE	vive en los bosques, come nueces y hay dos muy famosas en Disney	0		0	0	0	0	1,00000	0,00000	0,00000
5ardE	<b>Roedor</b> que trepa a los árboles, tiene <b>cola larga</b> y come bellotas.	ROEDOR .Y COLA .Y LARGA	ardilla	ardilla	coendú	5	1	0,00000	1,00000	0,20000
6ardE	<b>Mamifero</b> que suele encontrarse en los bosques, se alimenta de nueces y diversas semillas	MAMIFERO	ardilla	acleido	adive	296	16	0,05085	1,00000	0,00338
7ardE	Animal que vive en zonas selváticas o en bosque o arboleda, que come regularmente nueces o semillas, vive en los árboles y en CU hay muchas	0	ardilla	0	0	0	0	1,00000	0,00000	0,00000
8ardE	Animal <b>mamifero, roedor</b> , habita en los árboles, se alimenta de nueces.	MAMÍFERO .Y ROEDOR	ardilla	aguti	ardilla	32	2	0,03226	1,00000	0,03125
9ardE	Animal pequeño y con una <b>cola larga</b> que habita en los bosques y se alimenta de semillas	COLA .Y LARGA	ardilla	águila	alondra	68	5	0,05970	1,00000	0,01471
10ardE	Animal pequeño, es un <b>roedor</b> , gusta comer de nueces. Cuenta con una <b>cola</b> grande y ancha	ROEDOR .Y COLA	ardilla	ardilla	castor	11	1	0,00000	1,00000	0,09091

Tabla 4.9 Consulta de 'ardilla' en el Vox

Clave	Definición libre de 'cubeta'	Forma introducida	Fase 2	1° Lugar	2°Lugar	Total	L.Ref	F	P	G
1cubE	<b>recipiente</b> grande para líquidos, generalmente de plástico, lo usamos como auxiliar para lavar el coche	RECIPIENTE		abducción	abocar	100	44	0,43434	1,00000	0,01000
2cubE	<b>recipiente</b> de varios materiales para contener líquidos generalmente	RECIPIENTE		abducción	abocar	100	44	0,43434	1,00000	0,01000
3cubE	lugar donde se puede almacenar líquidos o cualquier otra cosa con <b>asa</b> para sujetarse	ASA		agarradero	aldabón	45	19	0,40909	1,00000	0,02222
4cubE	artículo utilizado como contenedor de agua	0		0	0	0	0	1,00000	0,00000	0,00000
5cubE	Un objeto que nos sirve para acarrear agua.	0	cubeta	0	0	0	0	1,00000	0,00000	0,00000
6cubE	<b>Recipiente</b> para poder transportar agua; se sujeta por una oreja o agarradera.	RECIPIENTE	olla	abducción	abocar	100	44	0,43434	1,00000	0,01000
7cubE	<b>Recipiente de forma</b> redonda.	RECIPIENTE .Y FORMA	olla	ascidio	balde	28	12	0,40741	1,00000	0,03571
8cubE	<b>Recipiente de forma</b> variable utilizado para contener y transportar agua.	RECIPIENTE .Y FORMA	garrafón	ascidio	balde	28	12	0,40741	1,00000	0,03571
9cubE	Es un objeto que sirve para acarrear agua.	0	cubeta	0	0	0	0	1,00000	0,00000	0,00000
10cubE	<b>Recipiente</b> que cuenta con un <b>asa</b> y que se utiliza generalmente para desplazar agua	RECIPIENTE .Y ASA	cubeta	cubeta	porongo	2	1	0,00000	1,00000	0,50000

Tabla 4.10 Consulta de 'cubeta' en el Vox

Clave	Definición libre de 'banca'	Forma introducida	Fase 2	1º Lugar	2ºLugar	Total	L.Ref	F	P	G
1banE	la utilizamos para poner objetos, es de <b>madera</b>	MADERA		ababán	ábaco	920	110	0,11861	1,00000	0,00109
2banE	artículo utilizado para descansar	0		0	0	0	0	1,00000	0,00000	0,00000
3banE	Mueble que sirve para sentarse y puede o no tener	RESPALDO		antimacasar	banca	28	2	0,03704	1,00000	0,03571
4banE	sirve para sentarse	0		0	0	0	0	1,00000	0,00000	0,00000
5banE	objeto que sirve para sentarse pero no tiene <b>respaldo</b> , generalmente se utiliza en la escuela	RESPALDO		antimacasar	banca	28	2	0,03704	1,00000	0,03571
6banE	mueble para sentarse se encuentra generalmente en sitios públicos	0		0	0	0	0	1,00000	0,00000	0,00000
7banE	Objeto que sirve para sentarse, generalmente la ponen en los parques y jardines	0	banca	0	0	0	0	1,00000	0,00000	0,00000
8banE	Artículo donde nos sentamos. Se encuentran en los parques o a las afueras de algún edificio, iglesia o centro donde se reúnen varias personas	0	bancas	0	0	0	0	1,00000	0,00000	0,00000
9banE	Mueble para sentarse de diversos materiales, se encuentra en parques y jardines	0	banca	0	0	0	0	1,00000	0,00000	0,00000
10banE	Objeto en el que se puede sentar más de una <b>persona</b> , por lo regular existe en jardines	PERSONA	bancas	a	ababol	4433	562	0,12658	1,00000	0,00023

Tabla 4.11 Consulta de 'banca' en el Vox

Clave	Definición libre de 'ropa'	Forma introducida	Fase 2	1° Lugar	2°Lugar	Total	L.Ref	F	P	G
1ropE	prenda tejida para <b>uso personal</b>	USO .Y PERSONAL		antes	bolso	4	4	1,00000	1,00000	0,25000
2ropE	artículo con el cual se puede cubrir a un ser humano	0		0	0	0	0	1,00000	0,00000	0,00000
3ropE	artículos para cubrir el <b>cuerpo</b> humano del clima	CUERPO		abadejo	abambolo	1760	1500	0,85219	1,00000	0,00057
4ropE	Es todo aquello que usamos para salir a la calle. Lo llevamos puesto.	0	ropa	0	0	0	0	1,00000	0,00000	0,00000
5ropE	Se dice de prendas hechas generalmente con <b>tela</b> o piel de animales para cubrir el <b>cuerpo</b>	TELA .Y CUERPO	abrigo	angorra	calar	37	29	0,77778	1,00000	0,02703
6ropE	Prenda de vestir utilizada para cubrir la desnudez de nuestro <b>cuerpo</b>	CUERPO		abadejo	abambolo	1760	1500	0,85219	1,00000	0,00057
7ropE	Actualmente fabricadas con fibras sintéticas	0	ropa	0	0	0	0	1,00000	0,00000	0,00000
8ropE	<b>Tela</b> cortada y confeccionada de diferentes formas que sirve para cubrir el <b>cuerpo</b> humano	TELA .Y CUERPO	ropa	angorra	calar	37	29	0,77778	1,00000	0,02703
9ropE	Indumentaria que se usa para vestirse diariamente	0		0	0	0	0	1,00000	0,00000	0,00000
10ropE	Algo que nos cubre la piel del medio ambiente, con la que podemos cubrirnos del frío	0	chamarra	0	0	0	0	1,00000	0,00000	0,00000

Tabla 4.12 Consulta de 'ropa' en el Vox

Clave	Definición libre de 'eutanasia'	Forma introducida	Fase 2	1º Lugar	2ºLugar	Total	L.Ref	F	P	G
1eutE	decisión para morir por uno mismo, pero que se pide a un médico que te aplique	0		0	0	0	0	1,00000	0,00000	0,00000
2eutE	la practican los médicos para saber	0		0	0	0	0	1,00000	0,00000	0,00000
3eutE	inyección letal	0		0	0	0	0	1,00000	0,00000	0,00000
4eutE	facultad que se le da a una persona de quitarse la vida por padecer enfermedad terminal	0		0	0	0	0	1,00000	0,00000	0,00000
5eutE	Se practica cuando alguien presenta <b>muerte</b> cerebral para quitarle la vida.	MUERTE	eutanasia	a	abadía	171	55	0,31765	1,00000	0,00585
6eutE	Dejar que muera o provocar la <b>muerte</b> de alguien, que por enfermedad tenga una muy mala calidad de vida, sin posibilidad de mejorarla.	MUERTE	eutanasia	a	abadía	171	55	0,31765	1,00000	0,00585
7eutE	Acto voluntario en el cual una persona podría ser auxiliada para morir.	0		0	0	0	0	1,00000	0,00000	0,00000
8eutE	Proceso médico por el cual se provoca la <b>muerte</b> de pacientes terminales bajo consentimiento	MUERTE	eutanasia	a	abadía	171	55	0,31765	1,00000	0,00585
9eutE	Método aceptado en pocos países que consiste en quitarle la vida a alguien durante una enfermedad terminal	0	eutanasia	0	0	0	0	1,00000	0,00000	0,00000
10eutE	La llevan a cabo los médicos cuando llega a ser necesario, está relacionada con la <b>muerte</b>	MUERTE	eutanasia	a	abadía	171	55	0,31765	1,00000	0,00585

Tabla 4.13 Consulta de 'eutanasia' en el Vox

Clave	Definición libre de 'monopolio'	Forma introducida	Fase 2	1º Lugar	2º Lugar	Total	L.Ref	F	P	G
1monE	tener el control total de producción de cierto producto o servicio. Es el nombre de un juego de mesa	0		0	0	0	0	1,00000	0,00000	0,00000
2monE	es el <b>dominio</b> de un mercado o área en general por una sola persona, grupo o consorcio	DOMINIO		abandono	absoluto	78	44	0,55844	1,00000	0,01282
3monE	la forma en la que varias sociedades se unen por un solo fin principalmente	0		0	0	0	0	1,00000	0,00000	0,00000
4monE	mercado	0		0	0	0	0	1,00000	0,00000	0,00000
5monE	Decidir algo por sí mismo, sin cuenta de otras opiniones.	0	injusticia	0	0	0	0	1,00000	0,00000	0,00000
6monE	Acaparar cualquier tipo de elementos y teniendo solamente el control de ellos, sin que nadie más tenga acceso a estos.	0		0	0	0	0	1,00000	0,00000	0,00000
7monE	Concentración de comercios	0	mercado	0	0	0	0	1,00000	0,00000	0,00000
8monE	Predominio que ejerce cierta compañía comercial que acapara el mercado, y no da lugar a sus competidores	0	monopolio	0	0	0	0	1,00000	0,00000	0,00000
9monE	Es la dominación del mercado por una empresa.	0	imperialismo	0	0	0	0	1,00000	0,00000	0,00000
10monE	Es una situación en la que una empresa tiene total dominio del ramo al que pertenece	0	monopolio	0	0	0	0	1,00000	0,00000	0,00000

Tabla 4.14 Consulta de 'monopolio' en el Vox



Clave	Definición libre de 'huracán'	Forma introducida	Fase 2	1º Lugar	2º Lugar	Total	L.Ref	F	P	G
1hurE	remolino de agua y viento en forma circular, es en tierra	VIENTO		abandonar	abanico	389	199	0,51031	1,00000	0,00257
2hurE	es un acontecimiento natural o fenómeno natural que provoca grandes cantidades de lluvia y aire	0		0	0	0	0	1,00000	0,00000	0,00000
3hurE	fenómeno climatológico que se presenta en los lugares que están cerca del mar, característico por sus fuertes lluvias y vientos	0		0	0	0	0	1,00000	0,00000	0,00000
4hurE	fenómeno meteorológico provocado por aire caliente que se combina con el aire frío. Hay una película con su nombre. Se lleva casas por el aire	0		0	0	0	0	1,00000	0,00000	0,00000
5hurE	Vientos a gran velocidad con lluvia en el mar	0	huracán	0	0	0	0	1,00000	0,00000	0,00000
6hurE	Fenómeno atmosférico que ocasiona movimiento de corrientes de aire con gran fuerza y velocidad alrededor de un centro u origen	0		0	0	0	0	1,00000	0,00000	0,00000
7hurE	Precipitación natural de niveles mayores (nivel 5) que ocasionan catástrofes, principalmente en zonas costeras. Esta precipitación se caracteriza porque a su centro se le llama ojo.	0	maremoto	0	0	0	0	1,00000	0,00000	0,00000
8hurE	Fenómeno meteorológico provocado por el choque de dos frentes de aire con lluvia y vientos muy fuertes	0	tormenta	0	0	0	0	1,00000	0,00000	0,00000
9hurE	Fenómeno destructivo que consiste en lluvias acompañadas de fuertes vientos	0		0	0	0	0	1,00000	0,00000	0,00000
10hurE	Fenómeno físico que puede llegar a ser peligroso, tiene la fuerza de arrancar casas frágiles u otro tipo de objetos	0		0	0	0	0	1,00000	0,00000	0,00000

Tabla 4.15 Consulta de 'huracán' en el Vox

Clave	Definición libre de 'barómetro'	Forma introducida	Fase 2	1º Lugar	2ºLugar	Total	L.Ref	F	P	G
1barE	es un <b>instrumento</b> que mide la velocidad del <b>aire</b>	INSTRUMENTO .Y AIRE		abanico	abano	35	10	0,26471	1,00000	0,02857
2barE	<b>instrumento</b> de medición de la <b>presión atmosférica</b>	INSTRUMENTO .Y PRESIÓN .Y ATMOSFÉRICA		barómetro	manoscopio	2	1	0,00000	1,00000	0,50000
3barE	<b>instrumento</b> de medición de presiones	INSTRUMENTO		a	ábaco	947	118	0,12368	1,00000	0,00106
4barE	<b>instrumento</b> para medir presiones	INSTRUMENTO		a	ábaco	947	118	0,12368	1,00000	0,00106
5barE	Es un <b>instrumento</b> que sirve para medir la <b>presión atmosférica</b>	INSTRUMENTO .Y PRESIÓN .Y ATMOSFÉRICA	barómetro	barómetro	manoscopio	2	1	0,00000	1,00000	0,50000
6barE	Aparato que se utiliza para medir la <b>presión</b>	PRESIÓN		abombar	aceite	148	22	0,14286	1,00000	0,00676
7barE	<b>Instrumento</b> de medición, cuya lectura nos indica la <b>presión atmosférica</b> del lugar en que se encuentra	INSTRUMENTO .Y INDICA .Y PRESIÓN .Y ATMOSFÉRICA	barómetro	barómetro	manoscopio	2	1	0,00000	1,00000	0,50000
8barE	Aparato que recibe datos de medición y así poder interpretarlos.	0	computadora	0	0	0	0	1,00000	0,00000	0,00000
9barE	<b>Instrumento</b> para medir la <b>presión atmosférica</b>	INSTRUMENTO .Y PRESIÓN .Y ATMOSFÉRICA	barómetro	barómetro	manoscopio	2	1	0,00000	1,00000	0,50000
10barE	<b>Instrumento</b> de medición	INSTRUMENTO	regla	a	ábaco	947	118	0,12368	1,00000	0,00847

Tabla 4.16 Consulta de 'barómetro' en el Vox

Término	F	P	G
Agua	0,78205	0,9	0,30959
limón	0,91588	0,2	0,00022
ardilla	0,32534	0,7	0,03583
cubeta	0,55269	0,7	0,06237
banca	0,63193	0,4	0,00727
ropa	0,92599	0,5	0,03052
eutanasia	0,72706	0,4	0,00234
monopolio	0,95584	0,1	0,00128
huracán	0,95103	0,1	0,00026
barómetro	0,17786	0,9	0,20459

Tabla 4.17 Promedios Vox

Diccionario Longman

Clave	Definición libre del 'agua'	Forma introducida	Fase 2	1° Lugar	2°Lugar	Total	L.Ref	F	P	G
1 agul	<b>It's a clear liquid</b> that you get from a tap	it's a clear liquid	water	acetone	ammonia	15	14	0,92857	1,00000	0,06667
2 agul	<b>The colourless transparent liquid</b> occurring on rivers	The liquid on rivers	Water	draught	water	2	2	1,00000	1,00000	0,50000
3 agul	<b>A clear, neutral liquid</b> that surrounds us everywhere	A clear liquid surrounds	Water	0	0	0	0	1,00000	0,00000	0,00000
4 agul	<b>Liquid, clear</b> , drinkable – constituents are hydrogen and oxygen	clear, liquid, are	Water	shit	water	2	2	1,00000	1,00000	0,50000
5 agul	<b>Liquid, clear</b> , H2O	liquid, clear	Water	acetone	ammonia	15	14	0,92857	1,00000	0,06667
6 agul	<b>Liquid form</b> , scientific term H2O	liquid	Water	liquid	abscess	405	397	0,98020	1,00000	0,00247
7 agul	<b>Liquid</b> , freezes at 0°C	liquid	Water	liquid	abscess	405	397	0,98020	1,00000	0,00247
8 agul	<b>Liquid, clear</b> , boils at 100°C, freezes at 0°C	liquid, clear	Water	acetone	ammonia	15	14	0,92857	1,00000	0,06667
9 agul	Fluid, <b>clear</b> , tasteless, colourless	clear	Water	clear	clearing	176	167	0,94857	1,00000	0,00568
10 agul	<b>Wash with it; drink it; used for dilution; H2O; found in springs, rivers, lakes, seas, oceans</b>	Wash with it drink it used for found in springs, rivers, lakes, seas, oceans	Water	0	0	0	0	1,00000	0,00000	0,00000

Tabla 4.18 Consulta de 'agua' en Longman

Clave	Definición libre 'limón'	Forma introducida	Fase 2	1º Lugar	2ºLugar	Total	L.Ref	F	P	G
1liml	<b>It's a yellow fruit, like limes.</b> Citrus. Used in cooking for sharpness	it's a yellow fruit	Lemon	apple	apricot	18	7	0,35294	1,00000	0,05556
2 liml	<b>A yellow citrus fruit. Sour tasting.</b> Often used as <b>an accompaniment to drinks</b>	A yellow fruit. Sour tasting. To drinks	lemon	lemon		1	1	0,00000	1,00000	1,00000
3 liml	<b>a yellow citrus fruit with a bitter taste</b> often <b>sliced</b> and put in <b>drinks</b>	a yellow fruit with a taste sliced and drinks	lemon	0	0	0	0	1,00000	0,00000	0,00000
4 liml	<b>It's a citrus fruit, yellow,</b> used <b>with</b> sugar on pancakes	It's a fruit, yellow	lemon	apple	apricot	18	7	0,35294	1,00000	0,05556
5 liml	<b>It's a yellow citrus fruit. Tastes</b> bitter. Oval shaped	It's a yellow fruit. Tastes	lemon	lemon	lime	2	1	0,00000	1,00000	0,50000
6 liml	<b>A yellow sour fruit</b>	A yellow sour fruit	lemon, grapefruit	lemon	lime	2	1	0,00000	1,00000	0,50000
7 liml	<b>A yellow citrus fruit</b>	A yellow fruit	lemon	apple	apricot	18	7	0,35294	1,00000	0,05556
8 liml	<b>Yellow, citrus, fruit</b>	Yellow, fruit	lemon	apple	apricot	18	7	0,35294	1,00000	0,05556
9 liml	Citrus <b>fruit</b> which is <b>yellow</b>	fruit is yellow	lemon	apple	apricot	9	4	0,37500	1,00000	0,11111
10 liml	<b>Yellow citrus fruit</b>	Yellow fruit	lemon	apple	apricot	18	7	0,35294	1,00000	0,05556

Tabla 4.19 Consulta de 'limón' en Longman

Clave	Definición libre de 'ardilla'	Forma introducida	Fase 2	1° Lugar	2°Lugar	Total	L.Ref	F	P	G
1ardl	It's a little rodent and can be red or grey, it has a big bushy tail	a tail	squirrel	alligator	ape	67	58	0,86364	1,00000	0,01493
2ardl	A small rodent living in trees with a long bushy tail	A small, trees, with a long tail	squirrel	gibbon	marten	4	4	1,00000	1,00000	0,25000
3ardl	A small rodent which lives in trees, collects nuts and has a bushy tail	A small, trees, nuts, a tail	squirrel	squirrel		1	1	0,00000	1,00000	1,00000
4ardl	Animal, grey/red, bushy tail, lives in trees, buries nuts	Animal, tail, trees, nuts	squirrel	squirrel		1	1	0,00000	1,00000	1,00000
5ardl	Small animal, lives in trees, eats acorns, has a bushy tail	Small animal, trees, eats, a tail	squirrel	koala	marten	3	3	1,00000	1,00000	0,33333
6ardl	Animal, bushy tail, eats nuts, builds nests in trees called dreys	Animal, tail, eats nuts, trees, dreys	squirrel	0	0	0	0	1,00000	0,00000	0,00000
7ardl	Small funny animal with big, bushy tail, likes nuts, likes trees	Small animal with tail, nuts, trees	squirrel	squirrel		1	1	0,00000	1,00000	1,00000
8ardl	Animal that lives in trees and collects acorns, has a long tail	Animal, trees and a long tail	squirrel	gibbon	marten	4	4	1,00000	1,00000	0,25000
9ardl	A small- sized animal, habitat in trees	A small animal, trees	bird	gibbon	koala	5	5	1,00000	1,00000	0,20000
10ardl	Small grey mammal, relative to the rodents, found in both countryside and town	small	rat?	small	a	2116	1788	0,84492	1,00000	0,00047

Tabla 4.20 Consulta de 'ardilla' en Longman

Clave	Definición libre de 'cubeta'	Forma introducida	Fase 2	1º Lugar	2º Lugar	Total	L.Ref	F	P	G
1 cubl	<b>It's used for carrying water</b> or other liquids. Also things like sand. A beach it is often with a spade and is used to make sandcastles	It's used for carrying water or liquids, things	bucket	absorption	bucket	29	2	0,03571	1,00000	0,03448
2 cubl	<b>A vessel for carrying liquid</b> or solids. Has a handle. Often seen in wells and accompanies a spade when see on the beach	for carrying liquid, a handle,	bucket	bucket	pail	2	1	0,00000	1,00000	0,50000
3 cubl	Item used to carry certain objects - mainly sand, water or soil. Children use them to make sandcastles at the beach	used to carry water	bucket	bucket	dyke	5	1	0,00000	1,00000	0,20000
4 cubl	<b>Used to carry water in/ make sandcastles; has a handle</b>	Used to carry water in/ make, a handle	bucket	bucket		1	1	0,00000	1,00000	1,00000
5 cubl	<b>Used to collect/ hold water</b> , also on seaside to measure sand	Used to hold water,	bucket	bucket	course	8	1	0,00000	1,00000	0,12500
6 cubl	<b>A device made out of metal or plastic used mainly to carry water</b> or other fluids	used to carry water	bucket	bucket	dyke	5	1	0,00000	1,00000	0,20000
7 cubl	When cleaning the floor you put water in this and dip the mop into it	water in it	bucket	antifreeze	ballast	99	5	0,04082	1,00000	0,01010
8 cubl	<b>A device usually used to contain water</b>	used to water	bucket	apple	backwater	117	10	0,07759	1,00000	0,00855
9 cubl	Item with a handle used for carrying things in, usually made of plastic	with a handle used for carrying things in,	bag	bucket		1	1	0,00000	1,00000	1,00000
10 cubl	<b>A sort of container used to carry something</b>	A container used to carry something	bag	bag	bucket	3	2	0,50000	1,00000	0,33333

Tabla 4.21 Consulta de 'cubeta' en Longman

1,00000 0,34115

Clave	Definición libre de 'banca'	Forma introducida	Fase 2	1° Lugar	2°Lugar	Total	L.Ref	F	P	G
1banl	You can <b>sit on</b> it in the street or <b>a park</b> and <b>they are</b> made of wood	sit on, a park	bench	bench	front	6	1	0,00000	1,00000	0,16667
2 banl	<b>A long hard seat</b> for several persons on which <b>the players on a sport team</b> sit	A long seat, on the players, a sport team	bench	back		1	0	1,00000	0,00000	0,00000
3 banl	<b>An object for sitting on</b> , usually <b>long</b> which can seat many people	An, for, long	bench	antecedent	antenna	143	0	1,00000	0,00000	0,00000
4 banl	<b>Sit on it (a few people can) in parks</b> , made of wood or iron	Sit on, in parks	bench	bench	front	6	1	0,00000	1,00000	0,16667
5 banl	Object <b>used for sitting on</b> . Often found in public places such as <b>parks</b> and gardens. <b>Used to seat 1 or more</b> people at a time	parks, Used to seat	bench			0	0	1,00000	0,00000	0,00000
6 banl	Something you <b>seat on</b> , is longer than a chair, usually made of wood	seat on, is longer	bench	back	bench	4	2	0,33333	1,00000	0,25000
7 banl	<b>Long platform for sitting on</b> (fit many people on <b>one</b> )	Long for one	bench	antecedent	antenna	100	0	1,00000	0,00000	0,00000
8 banl	Apparatus <b>for sitting on</b> , designed for <b>more</b> than one person, often found in	for more, parks	bench	ground	hour	11	0	1,00000	0,00000	0,00000
9 banl	<b>A kind of seat</b> found in <b>parks</b> , made of wood	A seat, parks	bench	big wheel	Ferris wheel	3	0	1,00000	0,00000	0,00000
10 banl	<b>A type of chair</b>		0 sofa	0	0	0	0	1,00000	0,00000	0,00000

Tabla 4.22 Consulta de 'banca' en Longman

Clave	Definición libre de 'ropa'	Forma introducida	Fase 2	1º Lugar	2º Lugar	Total	L.Ref	F	P	G
1ropl	It describes the set of <b>things</b> you <b>wear</b> like trousers, jumpers, shirts, dresses, etc., anything in fact <b>that you wear</b>	things, that wear	clothes	best	box	33	7	0,18750	1,00000	0,03030
2ropl	The collective name for the items which we <b>wear</b> ; i.e. trousers, shirts, jumpers, etc.	wear	clothes, attire, gear, rags	wear	abaya	485	88	0,17975	1,00000	0,00206
3ropl	Buy it in <b>shops</b> , <b>keeps</b> us <b>warm</b> , can be fashionable or unfashionable. Consists of <b>things</b> such as jumpers, trousers - the collective name	shops, keeps warm, things	clothes, woolies			0	0	1,00000	0,00000	0,00000
4ropl	<b>Things</b> used to <b>cover</b> up and <b>keep warm</b> , used by humans	Things, cover and keep warm,	clothes	clothes		1	1	0,00000	1,00000	1,00000
5ropl	Items <b>people</b> can <b>wear</b>	people wear	clothes	account	bagde	84	15	0,16867	1,00000	0,01190
6ropl	Items <b>that we wear</b> , to <b>keep</b> us <b>warm</b>	that wear, to keep warm	coat	clothes	coat	5	1	0,00000	1,00000	0,20000
7ropl	Garments <b>worn</b> on <b>body</b>	worn body	clothes	abaya	armour	38	12	0,29730	1,00000	0,02632
8ropl	<b>What</b> you <b>wear</b> on your <b>body</b>	wear body	clothes	abaya	armour	38	12	0,29730	1,00000	0,02632
9ropl	Items <b>that we all wear</b>	that wear	clothes	acesory	account	301	57	0,18667	1,00000	0,00332
10ropl	<b>Things</b> you <b>wear</b>	things wear	clothes (or more specific? trousers, shoes, hats, etc.)	best	box	38	8	0,18919	1,00000	0,02632

Tabla 4.23 Consulta de 'ropa' en Longman



Clave	Definición libre de 'eutanasia'	Forma introducida	Fase 2	1º Lugar	2ºLugar	Total	L.Ref	F	P	G
1eutl	It's when old or disabled people are <b>killed</b> legally and happens in the Netherlands	killed	euthanasia	kill	killing	216	74	0,33953	1,00000	0,00463
2eutl	Act or practice of causing death painlessly		0 euthanasia	0	0	0	0	1,00000	0,00000	0,00000
3eutl	<b>The</b> decision taken by someone as to whether they <b>die</b> or not	to die	suicide, euthanasia	air	annuity	118	37	0,30769	1,00000	0,00847
4eutl	Idea of turning off life -support machine; coma		0 euthanasia			0	0	1,00000	0,00000	0,00000
5eutl	Illegal act of helping somebody <b>die who</b> is terminally ill, <b>who</b> is in lot of pain, <b>who</b> needs help – cannot <b>die</b> alone	who, die	euthanasia	beneficiary	breed	72	16	0,21127	1,00000	0,01389
6eutl	<b>The</b> right to decide to end one's own life		0 suicide	0	0	0	0	1,00000	0,00000	0,00000
7eutl	Right to <b>die</b> under special conditions	to die	euthanasia	air	annuity	118	37	0,30769	1,00000	0,00847
8eutl	Ending life before natural end, to relieve from <b>suffering</b> due to illness	to, suffering	euthanasia	abuse	area	82	23	0,27160	1,00000	0,01220
9eutl	<b>Killing</b> someone when they are terminally ill with their permission	killing	Euthanasia	killing	kill	218	74	0,33641	1,00000	0,00459
10eutl	Allowing people to <b>die</b> if they wish to, giving <b>them the</b> drugs or means to terminate their lives	die them	euthanasia	blight	breed	21	8	0,35000	1,00000	0,04762

Tabla 4.24 Consulta de 'eutanasia' Longman

Clave	Definición libre de 'monopolio'	Forma introducida	Fase 2	1º Lugar	2º Lugar	Total	L.Ref	F	P	G
1monl	Boardgame with houses and hotels, properties, jail, "go"		0 monopoly			0	0	1,00000	0,00000	0,00000
2monl	A board game involving the buying and selling of property		0 monopoly			0	0	1,00000	0,00000	0,00000
3monl	Board game - buy/sell properties - contains Mayfair (Eng. Version) and park line; when you have complete control over something - normally business	have complete control, business	monopoly	monopoly		1	1	0,00000	1,00000	1,00000
4 monl	A game where you buy land/railways and build houses/hotels or to hold a majority amount of something	a game, something	monopoly	ace	amusement	129	0	1,00000	0,00000	0,00000
5 monl	Something to do with business, literally translated from Latin – one of many. Is also a board game	something, business	monopoly	agent	appropriat	102	55	0,53465	1,00000	0,00980
6 monl	A company has control in the economy	has control	monopoly	about-face	act	143	91	0,63380	1,00000	0,00699
7 monl	When a company or individual has a major share of the market they have a ...		0 monopoly			0	0	1,00000	0,00000	0,00000
8 monl	To have majority control, e.g. of a company's shares, or of a divisional field	control, of a field	monopoly	part		1	0	1,00000	0,00000	0,00000
9 monl	An authoritative power	power	monopoly	power	absolutism	412	228	0,55231	1,00000	0,00243
10 monl	The interaction of economical communities		0			0	0	1,00000	0,00000	0,00000

Tabla 4.25 Consulta de 'monopolio' Longman

Clave	Definición libre de 'huracán'	Forma introducida	Fase 2	1° Lugar	2°Lugar	Total	L.Ref	F	P	G
1hurl	A kind of severe <b>storm</b> which involves a lot of <b>strong winds</b>	storm, strong winds	hurricane, monsoon, gale, typhoon, tornado	dust storm	hurricane	7	2	0,16667	1,00000	0,14286
2hurl	A violent tropical cyclone		0 typhoon, hurricane			0	0	1,00000	0,00000	0,00000
3 hurl	<b>Very strong winds</b> of over 100 mph which cause great destruction	very strong winds over	gales, tornadoes, hurricanes	hurricane		1	1	0,00000	1,00000	1,00000
4 hurl	<b>Strong winds</b> and rain, gale, destruction over large areas	strong winds over	hurricane, tornado	hurricane	wind	2	1	0,00000	1,00000	0,50000
5 hurl	<b>A very strong wind, very</b> destructive, often named after people's first names, i.e. the weather researchers who first discover each one	a very strong wind, very	hurricane?	blast	cyclone	6	4	0,60000	1,00000	0,16667
6 hurl	Big amount of <b>wind</b> going from country to country, ruining everything as it goes!	wind	cyclone, hurricane	wind	anorak	105	29	0,26923	1,00000	0,00952
7 hurl	Violent <b>winds</b> which can cause large scale destruction	winds	gale	wind	anorak	105	29	0,26923	1,00000	0,00952
8 hurl	Natural disaster, whipping <b>wind</b> , destroys anything in its path	wind	tornado	wind	anorak	105	29	0,26923	1,00000	0,00952
9 hurl	A type of typhoon		0 hurricane			0	0	1,00000	0,00000	0,00000
10 hurl	<b>A very very strong wind</b>	A very very strong wind	hurricane	blast	cyclone	6	4	0,60000	1,00000	0,16667

Tabla 4.26 Consulta de 'huracán' Longman

Clave	Definición libre de 'barómetro'	Forma introducida	Fase 2	1º Lugar	2ºLugar	Total	L.Ref	F	P	G
1barl	Instrument which detects <b>changes</b> in atmospheric <b>pressure</b>	instrument, changes, pressure	barometer	barometer		1	1	0,00000	1,00000	1,00000
2barl	<b>An instrument for measuring pressure</b>	An instrument, measuring pressure	barometer	barometer		1	1	0,00000	1,00000	1,00000
3 barl	Like thermometer but <b>measures air pressure</b>	measures air pressure	barometer	barometer		1	1	0,00000	1,00000	1,00000
4 barl	Used to <b>measure air pressure</b>	measure air pressure	barometer	barometer		1	1	0,00000	1,00000	1,00000
5 barl	A device to <b>measure pressure</b>	measure pressure	barometer	barometer	head	4	1	0,00000	1,00000	0,25000
6 barl	<b>Instrument</b> used in physics to <b>measure pressure</b>	instrument, measure, pressure	barometer	barometer		1	1	0,00000	1,00000	1,00000
7 barl	<b>An instrument</b> used to <b>measure pressure</b> , used in physics	an instrument, measure pressure	barometer	barometer		1	1	0,00000	1,00000	1,00000
8 barl	A device used to <b>measure air pressure</b>	measure air pressure	air gauge	barometer		1	1	0,00000	1,00000	1,00000
9 barl	<b>Measures pressure</b>	Measures pressure		barometer	head	4	1	0,00000	1,00000	0,25000
10 barl	Used to <b>measure pressure</b> , e.g. <b>weather</b> reports	measure pressure, weather		barometer	high	2	1	0,00000	1,00000	0,50000

Tabla 4.27 Consulta de 'barómetro' Longman

Término	F	P	G
Water	0,96947	0,8	0,12106
lemon	0,31397	0,9	0,23889
squirrel	0,67086	0,9	0,40487
bucket	0,06541	1	0,34115
bench	0,73333	0,3	0,05833
clothes	0,25064	0,9	0,13265
euthanasia	0,51242	0,7	0,00999
monopoly	0,77208	0,40000	0,10192
hurricane	0,41744	0,8	0,20048
barometer	0	1	0,80000

Tabla 4.28 Promedios

Longman

Diccionario Wordsmyth

Clave	Definición libre de 'agua'	Forma introducida	Fase 2	1° Lugar	2°Lugar	Total	L.Ref	F	P	G
1agul	It's a clear <b>liquid</b> that you get from a tap	a liquid	Water	acetal	acetic acid	180	175	0,97207	1,00	0,00556
2agul	The colourless <b>transparent liquid</b> occurring on <b>rivers</b>	transparent liquid, rivers	Water	water	0	1	1	0,00000	1,00	1,00000
3agul	A clear, neutral <b>liquid</b> that surrounds us everywhere	a liquid	Water	acetal	acetic acid	180	175	0,97207	1,00	0,00556
4agul	<b>Liquid</b> , clear, drinkable – constituents are hydrogen and oxygen	liquid	Water	ablution	acetal	272	267	0,98155	1,00	0,00368
5agul	<b>Liquid</b> , clear, H2O	liquid	Water	ablution	acetal	272	267	0,98155	1,00	0,00368
6agul	<b>Liquid form</b> , scientific term H2O	liquif form	Water	alcohol	cesium	8	7	0,85714	1,00	0,12500
7agul	<b>Liquid</b> , freezes at 0°C	liquid	Water	ablution	acetal	272	267	0,98155	1,00	0,00368
8agul	<b>Liquid</b> , clear, boils at 100°C, freezes at 0°C	liquid	Water	ablution	acetal	272	267	0,98155	1,00	0,00368
9agul	Fluid, clear, <b>tasteless</b> , colourless	tasteless	Water	air	arsenic	8	8	1,00000	1,00	0,12500
10agul	Wash with it; drink it; used for dilution; H2O; found in springs, <b>rivers, lakes</b> , seas, oceans	rivers, lakes,	Water	grass carp	water	2	2	1,00000	1,00	0,50000

Tabla 4.29 Consulta de 'agua' en Wordsmyth

Clave	Definición libre de 'limón'	Forma introducida	Fase 2	1° Lugar	2°Lugar	Total	L.Ref	F	P	G
1liml	It's a <b>yellow fruit</b> , like limes. Citrus. Used in cooking for sharpness	yellow fruit	Lemon	agrimony	apple	26	11	0,40000	1,00	0,03846
2 liml	A <b>yellow citrus fruit</b> . <b>Sour</b> tasting. Often used as an accompaniment to drinks	yellow citrus fruit, sour	lemon	lemon	0	1	1	0,00000	1,00	1,00000
3 liml	a <b>yellow citrus fruit</b> with a bitter taste often sliced and put in drinks	yellow citrus fruit	lemon	citron	grapefruit	4	3	0,66667	1,00	0,25000
4 liml	It's a <b>citrus fruit, yellow</b> , used with sugar on pancakes	citrus fruit yellow	lemon	citron	grapefruit	4	3	0,66667	1,00	0,25000
5 liml	It's a <b>yellow citrus fruit</b> . Tastes bitter. <b>Oval</b> shaped	yellow citrus fruit, oval	lemon	lemon	0	1	1	0,00000	1,00	1,00000
6 liml	A <b>yellow sour fruit</b>	yellow sour fruit	lemon, grapefruit	lemon	0	1	1	0,00000	1,00	1,00000
7 liml	A <b>yellow citrus fruit</b>	yellow citrus fruit	lemon	citron	grapefruit	4	3	0,66667	1,00	0,25000
8 liml	<b>Yellow, citrus, fruit</b>	yellow, citrus, fruit	lemon	citron	grapefruit	4	3	0,66667	1,00	0,25000
9 liml	<b>Citrus fruit</b> which is <b>yellow</b>	citrus fruit yellow	lemon	citron	grapefruit	4	3	0,66667	1,00	0,25000
10 liml	<b>Yellow citrus fruit</b>	yellow citrus fruit	lemon	citron	grapefruit	4	3	0,66667	1,00	0,25000

Tabla 4.30 Consulta de 'limón' en *Wordsmyth*

Clave	Definición libre de 'ardilla'	Forma introducida	Fase 2	1° Lugar	2°Lugar	Total	L.Ref	F	P	G
1ardl	It's a little rodent and can be red or grey, it has a big <b>bushy tail</b>	a bushy tail	squirrel	aye-aye	chipmunk	11	9	0,80000	1,00	0,09091
2ardl	<b>A small rodent living in trees with a long bushy tail</b>	a small, a long bushy tail	squirrel	lesser panda	phalanger	3	3	1,00000	1,00	0,33333
3ardl	<b>A small rodent which lives in trees, collects nuts and has a bushy tail</b>	a small, a bushy tail	squirrel	chipmunk	lesser panda	5	5	1,00000	1,00	0,20000
4ardl	Animal, grey/red, <b>bushy tail</b> , lives in trees, buries nuts	bushy tail	squirrel	aye-aye	chipmunk	11	9	0,80000	1,00	0,09091
5ardl	<b>Small animal, lives in trees, eats acorns, has a bushy tail</b>	small, a bushy tail	squirrel	chipmunk	lesser panda	5	5	1,00000	1,00	0,20000
6ardl	Animal, <b>bushy tail</b> , eats nuts, builds nests in trees called dreys	bushy tail	squirrel	aye-aye	chipmunk	11	9	0,80000	1,00	0,09091
7ardl	<b>Small funny animal with big, bushy tail, likes nuts, likes trees</b>	small, bushy tail	squirrel	chipmunk	lesser panda	5	5	1,00000	1,00	0,20000
8ardl	Animal that lives in trees and collects acorns, has a <b>long tail</b>	a long tail	squirrel	ant bear	aye-aye	61	50	0,81667	1,00	0,01639
9ardl	A small-sized animal, habitat in trees	0	bird	0	0	0	0	1,00000	0,00	0,00000
10ardl	<b>Small grey mammal, relative to the rodents, found in both countryside and town</b>	small, rodents	rat?	dormouse	gerbil	13	11	0,83333	1,00	0,07692

Tabla 4.31 Consulta de 'ardilla' *Wordsmyth*

Clave	Definición libre de 'cubeta'	Forma introducida	Fase 2	1° Lugar	2°Lugar	Total	L.Ref	F	P	G
1 cubl	It's <b>used</b> for carrying water or <b>other</b> liquids. Also things like sand. A beach it is often with a spade and is <b>used to</b> make sandcastles	used, other, used to	bucket	abortifacient	activated carbon	88	15	0,16092	1,00	0,01136
2 cubl	A vessel for carrying liquid or solids. Has a <b>handle</b> . Often seen in wells and accompanies a spade when see on the beach	a handle	bucket	adz or adze	ax or axe	93	11	0,10870	1,00	0,01075
3 cubl	Item <b>used to</b> carry certain objects - mainly sand, water or soil. Children use them to make sandcastles at the beach	used to	bucket	abacus	abortifacient	1491	174	0,11611	1,00	0,00067
4 cubl	<b>Used to</b> carry water in/ make sandcastles; has a <b>handle</b>	used to handle	bucket	adz or adze	cat-o'-nine-tails	28	0	1,00000	0,00	0,00000
5 cubl	<b>Used to</b> collect/ hold water, also on seaside to measure sand	used to,	bucket	abacus	abortifacient	1491	174	0,11611	1,00	0,00067
6 cubl	A device made out of metal or plastic <b>used</b> mainly <b>to</b> carry water or <b>other</b> fluids	used to, other	bucket	abortifacient	activated carbon	90	15	0,15730	1,00	0,01111
7 cubl	When cleaning the floor you put water in this and dip the mop into it	0	bucket	0	0	0	0	1,00000	0,00	0,00000
8 cubl	A device usually <b>used</b> to contain water	used	bucket	abaca	abacus	3772	428	0,11323	1,00	0,00027
9 cubl	Item with a <b>handle used</b> for carrying things in, usually made of plastic	a handle, used	bag	adz or adze	ax or axe	50	0	1,00000	0,00	0,02000
10 cubl	A sort of <b>container used</b> to carry something	container, used	bag	autoclave	bag	37	0	1,00000	0,00	0,00000

Tabla 4.32 Consulta de 'cubeta' Wordsmyth



Clave	Definición libre de 'banca'	Forma introducida	Fase 2	1° Lugar	2°Lugar	Total	L.Ref	F	P	G
1banl	You can sit on it in the street or a park and they are made of wood	0	bench	0	0	0	0	1,00000	0,00	0,00000
2banl	A <b>long</b> hard <b>seat</b> for several persons on which the players on a sport team sit	long, seat	bench	bench	settle	4	1	0,00000	1,00	0,25000
3banl	An object for sitting on, usually long which can <b>seat</b> many <b>people</b>	seat, people	bench	barouche	bench	3	2	0,50000	1,00	0,33333
4banl	Sit on it (a few <b>people</b> can) in parks, made of wood or iron	people	bench	accent	acquaintance	475	36	0,07384	1,00	0,00211
5banl	Object used for sitting on. <b>Often</b> found in public places such as parks and gardens. Used to <b>seat</b> 1 or <b>more people</b> at a time	often seat, more people	bench	bench	0	1	1	0,00000	1,00	1,00000
6banl	Something you <b>seat</b> on, <b>is</b> longer than a chair, usually made of wood	seat, is	bench	bench	saddle seat	3	1	0,00000	1,00	0,33333
7banl	<b>Long</b> platform for sitting on (fit many <b>people</b> on one)	long, people	bench	bench	counter	5	1	0,00000	1,00	0,20000
8banl	Apparatus for sitting on, designed for <b>more</b> than one person, <b>often</b> found in parks	more, often	bench	account executive	bench	16	2	0,06667	1,00	0,06250
9banl	A kind of <b>seat</b> found in parks, made of wood	seat	bench	accomodatio	back seat	46	4	0,06667	1,00	0,02174
10banl	A type of chair	0	sofa	0	0	0	0	1,00000	0,00	0,00000

Tabla 4.33 Consulta de 'banca' Wordsmyth

Clave	Definición libre de 'ropa'	Forma introducida	Fase 2	1° Lugar	2°Lugar	Total	L.Ref	F	P	G
1ropl	It describes the set of things you wear like trousers, jumpers, <b>shirts, dresses</b> , etc., anything in fact that you wear	like shirts, dresses	clothes	clothes	0	1	1	0,00000	1,00	1,00000
2ropl	The collective name for the items which we wear; i.e. trousers, <b>shirts</b> , jumpers, etc.	shirts	clothes, attire, gear, rags	clothes	0	1	1	0,00000	1,00	1,00000
3ropl	Buy it in shops, keeps us warm, can be fashionable or unfashionable. Consists of things <b>such as</b> jumpers, trousers - the collective name	such as	clothes, woolie	actinic rays	bedclothes	36	4	0,08571	1,00	0,02778
4ropl	Things used to cover up and keep warm, used by humans	0	clothes	0	0	0	0	1,00000	0,00	0,00000
5ropl	Items people can wear	0	clothes	0	0	0	0	1,00000	0,00	0,00000
6ropl	Items that we wear, to keep us warm	0	coat	0	0	0	0	1,00000	0,00	0,00000
7ropl	<b>Garments worn on body</b>	worn on body	clothes	clothes	0	1	1	0,00000	1,00	1,00000
8ropl	<b>What you wear on your body</b>	on body	clothes	clothes	stigmata	2	1	0,00000	1,00	0,50000
9ropl	Items that we all wear	0	clothes	0	0	0	0	1,00000	0,00	0,00000
10ropl	Things you wear	0	clothes (or more specific? trousers, shoes, hats, etc.)	0	0	0	0	1,00000	0,00	0,00000

Tabla 4.34 Consulta de 'ropa' Wordsmyth

Clave	Definición libre de 'eutanasia'	Forma introducida	Fase 2	1° Lugar	2°Lugar	Total	L.Ref	F	P	G
1eutl	It's when old or disabled people are killed legally and happens in the Netherlands	0	euthanasia	0	0	0	0	1,00000	0,00	0,00000
2 eutl	<b>Act</b> or practice of causing <b>death</b> painlessly	act death	euthanasia	decease	euthanasia	4	2	0,33333	1,00	0,25000
3 eutl	The decision taken by someone as to whether they die or not	0	suicide, euthanasia	0	0	0	0	1,00000	0,00	0,00000
4 eutl	Idea of turning off life -support machine; coma	0	euthanasia	0	0	0	0	1,00000	0,00	0,00000
5 eutl	Illegal <b>act</b> of helping somebody die who is terminally ill, who is in lot of pain, who needs help – cannot die alone	act	euthanasia	abatement	abbreviation	1875	610	0,32497	1,00	0,00053
6 eutl	The right to decide to end one's own life	0	suicide	0	0	0	0	1,00000	0,00	0,00000
7 eutl	Right to die under special conditions	0	euthanasia	0	0	0	0	1,00000	0,00	0,00000
8 eutl	Ending life before natural end, to relieve from <b>suffering</b> due to illness	suffering	euthanasia	affliction	agony	51	12	0,22000	1,00	0,01961
9eutl	Killing someone when they are terminally ill with their permission	0	Euthanasia	0	0	0	0	1,00000	0,00	0,00000
10eutl	Allowing people to die if they wish to, giving them the drugs or means to terminate their lives	0	euthanasia	0	0	0	0	1,00000	0,00	0,00000

Tabla 3.35 Consulta de 'eutanasia' Wordsmyth

Clave	Definición libre de 'monopolio'	Forma introducida	Fase 2	1º Lugar	2ºLugar	Total	L.Ref	F	P	G
1monl	Boardgame with houses and hotels, properties, jail, "go"	0	monopoly	0	0	0	0	1,00000	0,00	0,00000
2monl	A board game involving the buying and selling of property	0	monopoly	0	0	0	0	1,00000	0,00	0,00000
3monl	Board game - buy/sell properties - contains Mayfair (Eng. Version) and park line; when you have <b>complete control over something</b> - normally <b>business</b>	complete, over something, business	monopoly	0	0	0	0	1,00000	0,00	0,00000
4 monl	A game where you buy land/railways and build houses/hotels or to hold a majority amount of <b>something</b>	something	monopoly	abbreviation	ability	1617	843	0,52104	1,00	0,00062
5 monl	<b>Something</b> to do with <b>business</b> , literally translated from Latin – <b>one</b> of many. <b>Is</b> also a board game	something, business, one, is	monopoly	trade secret	0	1	0	1,00000	0,00	0,00000
6 monl	A <b>company has control</b> in the economy	company has control	monopoly	monopoly	0	1	1	0,00000	1,00	1,00000
7 monl	When a <b>company</b> or individual <b>has</b> a <b>major share</b> of the market they have a ...	company has	monopoly	agency shop	home office	3	3	1,00000	1,00	0,33333
8 monl	To have majority control, e.g. of a <b>company's</b> shares, or of a divisional field	company	monopoly	account execu	agency shop	78	43	0,54545	1,00	0,01282
9 monl	An authoritative power	0	monopoly	0	0	0	0	1,00000	0,00	0,00000
10monl	The interaction of economical communities	0		0	0	0	0	1,00000	0,00	0,00000

Tabla 4.36 Consulta 'monopolio' *Wordsmyth*

Clave	Definición libre de 'huracán'	Forma introducida	Fase 2	1° Lugar	2° Lugar	Total	L.Ref	F	P	G
1hurl	A kind of severe <b>storm</b> which involves a lot of strong <b>winds</b>	storm, winds	hurricane, monsoon, gale, typhoon, tornado	dust storm	hurricane	7	2	0,16667	1,00	0,14286
2hurl	A violent tropical cyclone	0	typhoon, hurricane	0	0	0	0	1,00000	0,00	0,00000
3hurl	Very strong <b>winds</b> of over 100 mph which cause great destruction	winds	gales, tornadoes, hurricanes	Aeolus	anticyclone	12	5	0,36364	1,00	0,08333
4hurl	Strong <b>winds</b> and rain, gale, destruction over large areas	winds	hurricane, tornado	Aeolus	anticyclone	12	5	0,36364	1,00	0,08333
5hurl	A very strong wind, very destructive, often named after people's first names, i.e. the weather researchers who first discover each one	0	hurricane?	0	0	0	0	1,00000	0,00	0,00000
6hurl	Big amount of wind going from country to country, ruining everything as it goes!	0	cyclone, hurricane	0	0	0	0	1,00000	0,00	0,00000
7hurl	Violent <b>winds</b> which can cause large scale destruction	winds	gale	Aeolus	anticyclone	12	5	0,36364	1,00	0,08333
8hurl	Natural disaster, whipping wind, destroys anything in its path	0	tornado	0	0	0	0	1,00000	0,00	0,00000
9hurl	A type of typhoon	0	hurricane	0	0	0	0	1,00000	0,00	0,00000
10hurl	A very very strong wind	0	hurricane	0	0	0	0	1,00000	0,00	0,00000

Tabla 4.37 Consulta 'huracán' *Wordsmyth*

Clave	Definición libre de 'barómetro'	Forma introducida	Fase 2	1° Lugar	2° Lugar	Total	L.Ref	F	P	G
1barl	<b>Instrument</b> which detects <b>changes</b> in <b>atmospheric pressure</b>	instrument, changes, atmospheric pressure	barometer	barometer	baroscope	2	1	0,00000	1,00	0,50000
2barl	An <b>instrument</b> for measuring <b>pressure</b>	instrument, pressure	barometer	barometer	baroscope	8	1	0,00000	1,00	0,12500
3 barl	Like thermometer but <b>measures</b> air <b>pressure</b>	measures, pressure	barometer	barometer	kymograph	4	1	0,00000	1,00	0,25000
4 barl	<b>Used to</b> measure air <b>pressure</b>	used to, pressure	barometer	acetylcholine	autoclave	13	3	0,16667	1,00	0,07692
5 barl	A device to measure <b>pressure</b>	pressure	barometer	accumulator	acetylcholine	146	20	0,13103	1,00	0,00685
6 barl	<b>Instrument</b> used in physics to measure <b>pressure</b>	instrument used, pressure	barometer	barometer	0	1	1	0,00000	1,00	1,00000
7 barl	An <b>instrument</b> used to measure <b>pressure</b> , used in physics	instrument used to, pressure, used	barometer	barometer	0	1	1	0,00000	1,00	1,00000
8 barl	A device <b>used to</b> measure air <b>pressure</b>	used to, pressure	air gauge	acetylcholine	autoclave	13	3	0,16667	1,00	0,07692
9 barl	<b>Measures</b> <b>pressure</b>	measures pressure		barometer	kymograph	4	1	0,00000	1,00	0,25000
10 barl	<b>Used to</b> measure <b>pressure</b> , e.g. <b>weather</b> reports	used to, pressure, weather		barometer	0	1	1	0,00000	1,00	1,00000

Tabla 4.38 Consulta 'barómetro' Wordsmyth

Término	Fallout	Precision	Generality
<b>Water</b>	0,87275	1,00000	0,17758
<b>lemon</b>	0,44000	1,00000	0,45385
<b>squirrel</b>	0,90500	0,90000	0,12994
<b>bucket</b>	0,47724	0,22000	0,52778
<b>bench</b>	0,27072	0,80000	0,22030
<b>clothes</b>	0,50857	0,50000	0,35278
<b>euthanasia</b>	0,78783	0,30000	0,02701
<b>monopoly</b>	0,78517	0,40000	0,13468
<b>hurricane</b>	0,72576	0,40000	0,03929
<b>barometer</b>	0,04644	1,00000	0,42857

Tabla 4.39 promedios Wordsmyth

Clave	Definición libre del 'agua'	Fase 2	1° Lugar	2°Lugar	Total	L.Ref	F	P	G
1agul	It's a clear liquid that you get from a tap	Water	escape	water	1000	2	0,001	1	0,00100
2agul	The colourless transparent liquid occurring on rivers	Water	Water	pyrrolidine	1000	1	0	1	0,00100
3agul	A clear, neutral liquid that surrounds us everywhere	Water	water	plasma	1000	1	0	1	0,00100
4agul	Liquid, clear, drinkable – constituents are hydrogen and oxygen	Water	water	hydrate	1000	1	0	1	0,00100
5agul	Liquid, clear, H2O	Water	Wash	skim	1000	24	0,02302	1	0,00100
6agul	Liquid form, scientific term H2O	Water	wash	drop	1000	29	0,02803	1	0,00100
7agul	Liquid, freezes at 0°C	Water	Water	Freezing point	1000	1	0	1	0,00100
8agul	Liquid, clear, boils at 100°C, freezes at 0°C	Water	water	H2O	1000	1	0	1	0,00100
9agul	Fluid, clear, tasteless, colourless	Water	Water	propelyne glyco	1000	1	0	1	0,00100
10agul	Wash with it; drink it; used for dilution; H2O; found in springs, rivers, lakes, seas, oceans	Water	water	naiad	1000	1	0	1	0,00100

Tabla 4.40 Consulta de 'agua' en Onelook

Clave	Definición libre de 'limón'	Fase 2	1° Lugar	2°Lugar	Total	L.Ref	F	P	G
1liml	It's a yellow fruit, like limes. Citrus. Used in cooking for sharpness	Lemon	Orange	Lime	1000	7	0,00601	1	0,00100
2liml	A yellow citrus fruit. Sour tasting. Often used as an accompaniment to drinks	Lemon	Lemon	Orange	1000	1	0	1	0,00100
3liml	a yellow citrus fruit with a bitter taste often sliced and put in drinks	Lemon	Lemon	Orange	1000	1	0	1	0,00100
4liml	It's a citrus fruit, yellow, used with sugar on pancakes	Lemon	Lemon	Mandarin	1000	1	0	1	0,00100
5liml	It's a yellow citrus fruit. Tastes bitter. Oval shaped	Lemon	Lemon	Orange	1000	1	0	1	0,00100
6liml	A yellow sour fruit	lemon, grapefruit	Orange	bergamot	1000	3	0,002	1	0,00100
7liml	A yellow citrus fruit	Lemon	Orange	Citron	1000	3	0,002	1	0,00100
8liml	Yellow, citrus, fruit	Lemon	Orange	Citron	1000	3	0,002	1	0,00100
9liml	Citrus fruit which is yellow	Lemon	Lemon	Grapefruit	1000	1	0	1	0,00100
10liml	Yellow citrus fruit	Lemon	Orange	Citron	1000	3	0,002	1	0,00100

Tabla 4.41 Consulta de 'limón' en Onelook

Clave	Definición libre de 'ardilla'	Fase 2	1º Lugar	2ºLugar	Total	L.Ref	F	P	G
1ardl	It's a little rodent and can be red or grey, it has a big bushy tail	squirrel	Squirrel	fox	1000	1	0	1	0,00100
2ardl	A small rodent living in trees with a long bushy tail	squirrel	Squirrel	Phalanger	1000	1	0	1	0,00100
3ardl	A small rodent which lives in trees, collects nuts and has a bushy tail	squirrel	raccoon	Squirrel	1000	2	0,001	1	0,00100
4ardl	Animal, grey/red, bushy tail, lives in trees, buries nuts	squirrel	Raccoon	Squirrel	1000	2	0,001	1	0,00100
5ardl	Small animal, lives in trees, eats acorns, has a bushy tail	squirrel	Raccoon	dormouse	1000	3	0,002	1	0,00100
6ardl	Animal, bushy tail, eats nuts, builds nests in trees called dreys	squirrel	Raccoon	Squirrel	1000	2	0,001	1	0,00100
7ardl	Small funny animal with big, bushy tail, likes nuts, likes trees	squirrel	Raccoon	Squirrel	1000	2	0,001	1	0,00100
8ardl	Animal that lives in trees and collects acorns, has a long tail	squirrel	Kinkajou	raccoon	1000	4	0,003	1	0,00100
9ardl	A small-sized animal, habitat in trees	bird	Beaver	Squirrel	1000	2	0,001	1	0,00100
10ardl	Small grey mammal, relative to the rodent, found in both countryside and town	rat?	Coyote	Lemming	1000	121	0,12012	1	0,00100

Tabla 4.42 Consulta de 'ardilla' en Onelook

Clave	Definición libre de 'cubeta'	Fase 2	1º Lugar	2ºLugar	Total	L.Ref	F	P	G
1cubl	It's used for carrying water or other liquids. Also things like sand. A beach it is often with a spade and is used to make sandcastles	Bucket	Wash	Riot grl	1000	3	0,002	1	0,00100
2cubl	A vessel for carrying liquid or solids. Has a handle. Often seen in wells and accompanies a spade when see on the beach	Bucket	Touch	Block	1000	3	0,002	1	0,00100
3cubl	Item used to carry certain objects - mainly sand, water or soil. Children use them to make sandcastles at the beach	Bucket	Wash	filter	1000	445	0,44444	1	0,00100
4cubl	Used to carry water in/make sandcastles; has a handle	Bucket	Wash	ship	1000	5	0,004	1	0,00100
5cubl	Used to collect/hold water, also on seaside to measure sand	Bucket	Drift	Beach	1000	366	0,36537	1	0,00100
6cubl	A device made out of metal or plastic used mainly to carry water or other fluids	Bucket	Pipe	make	1000	914	0,91391	1	0,00100
7cubl	When cleaning the floor you put water in this and dip the mop into it	Bucket	Wash	set	1000	42	0,04104	1	0,00100
8cubl	A device usually used to contain water	Bucket	Filter	anchor	1000	41	0,04004	1	0,00100
9cubl	Item with a handle used for carrying things in, usually made of plastic	Bag	Pick	glass	1000	115	0,11411	1	0,00100
10cubl	A sort of container used to carry something	Bag	Carrier	bag	1000	37	0,03604	1	0,00100

Tabla 4.43 Consulta de 'cubeta' en Onelook



Clave	Definición libre de 'banca'	Fase 2	1° Lugar	2°Lugar	Total	L.Ref	F	P	G
1banl	You can sit on it in the street or a park and they are made of wood	Bench	Stand	Stick	1000	303	0,3023	1	0,00100
2 banl	A long hard seat for several persons on which the players on a sport team sit	Bench	Bench	Line	1000	1	0	1	0,00100
3 banl	An object for sitting on, usually long which can seat many people	Bench	Counter	trail	1000	13	0,01201	1	0,00100
4 banl	Sit on it (a few people can) in parks, made of wood or iron	Bench	Stand	stick	1000	68	0,06707	1	0,00100
5 banl	Object used for sitting on. Often found in public places such as parks and gardens. Used to seat 1 or more people at a time	Bench	Storied	lounge	1000	3	0,002	1	0,00100
6 banl	Something you seat on, is longer than a chair, usually made of wood	Bench	Settle	baton	1000	5	0,004	1	0,00100
7 banl	Long platform for sitting on (fit many people on one)	Bench	Gallery	Bench	1000	2	0,001	1	0,00100
8 banl	Apparatus for sitting on, designed for more than one person, often found in parks	Bench	Bench	Play	1000	1	0	1	0,00100
9 banl	A kind of seat found in parks, made of wood	bench	Settle	stick	1000	21	0,02002	1	0,00100
10 banl	A type of chair	Sofa	Jampan	coxwell	1000	0	1	0	0,00000

Tabla 4.44 Consulta de 'banca' en Onelook

Clave	Definición libre de 'ropa'	Fase 2	1º Lugar	2ºLugar	Total	L.Ref	F	P	G
1ropl	It describes the set of things you wear like trousers, jumpers, shirts, dresses, etc., anything in fact that you wear	clothes	clothes	Pack	1000	1	0	1	0,00100
2ropl	The collective name for the items which we wear; i.e. trousers, shirts, jumpers, etc.	clothes, attire, gear, rags	Tie	Register	1000	27	0,02603	1	0,00100
3ropl	Buy it in shops, keeps us warm, can be fashionable or unfashionable. Consists of things such as jumpers, trousers - the collective name	clothes, woolies	Cobham	Book	1000	0	1	0	0,00000
4ropl	Things used to cover up and keep warm, used by humans	clothes	Front	Guard	1000	194	0,19319	1	0,00100
5ropl	Items people can wear	clothes	Sports memorabilia	Stand	1000	377	0,37638	1	0,00100
6ropl	Items that we wear, to keep us warm	coat	Carry	Stock	1000	0	1	0	0,00000
7ropl	Garments worn on body	clothes	Gown	Wash	1000	4	0,003	1	0,00100
8ropl	What you wear on your body	clothes	Apon	Touch	1000	67	0,06607	1	0,00100
9ropl	Items that we all wear	clothes	Detail	Take	1000	0	1	0	0,00000
10ropl	Things you wear	clothes (or more specific? trousers, shoes, hats, etc.)	Pack	Battery	1000	20	0,01902	1	0,00100

Tabla 4.45 Consulta de 'ropa' en Onelook

Clave	Definición libre de 'eutanasia'	Fase 2	1º Lugar	2ºLugar	Total	L.Ref	F	P	G
1eutl	It's when old or disabled people are killed legally and happens in the Netherlands	euthanasia	Old	Safsaf massacre	1000	0	1	0	0,00000
2 eutl	Act or practice of causing death painlessly	euthanasia	Kill	Exercise	1000	5	0,004	1	0,00100
3 eutl	The decision taken by someone as to whether they die or not	suicide, euthanasia	Pass	blind	1000	0	1	0	0,00000
4 eutl	Idea of turning off life-support machine; coma	euthanasia	Run	turn	1000	0	1	0	0,00000
5 eutl	Illegal act of helping somebody die who is terminally ill, who is in lot of pain, who needs help – cannot die alone	euthanasia	Relief	service	1000	122	0,12112	1	0,00100
6 eutl	The right to decide to end one's own life	suicide	Take	conscience	1000	0	1	0	0,00000
7 eutl	Right to die under special conditions	euthanasia	Privilege	fall	1000	0	1	0	0,00000
8 eutl	Ending life before natural end, to relieve from suffering due to illness	euthanasia	Short	late	1000	29	0,02803	1	0,00100
9eutl	Killing someone when they are terminally ill with their permission	Euthanasia	Will	Facial hair	1000	25	0,02402	1	0,00100
10eut10	Allowing people to die if they wish to, giving them the drugs or means to terminate their lives	euthanasia	Red cap	Totalitarian democracy	1000	0	1	0	0,00000

Tabla 4.46 Consulta de 'eutanasia' Onelook

Clave	Definición libre de 'monopolio'	Fase 2	1º Lugar	2ºLugar	Total	L.Ref	F	P	G
1monl	Boardgame with houses and hotels, properties, jail, "go"	monopoly	Telecom hotel	Port island	1000	0	1	0	0,00000
2monl	A board game involving the buying and selling of property	monopoly	Commerce	trade	1000	8	0,00701	1	0,00100
3monl	Board game - buy/sell properties - contains Mayfair (Eng. Version) and park line; when you have complete control over something - normally business	monopoly	Nothingness	point	1000	114	0,11311	1	0,00100
4 monl	A game where you buy land/railways and build houses/hotels or to hold a majority amount of something	monopoly	Get	catch	1000	804	0,8038	1	0,00100
5 monl	Something to do with business, literally translated from Latin – one of many. Is also a board game	monopoly	Turn	act	1000	0	1	0	0,00000
6 monl	A company has control in the economy	monopoly	Nauru phosphate corporation	Production equipment control	1000	51	0,05005	1	0,00100
7 monl	When a company or individual has a major share of the market they have a ...	monopoly	Milton Hershey school	public	1000	125	0,12412	1	0,00100
8 monl	To have majority control, e.g. of a company's shares, or of a divisional field	monopoly	Run	point	1000	33	0,03203	1	0,00100
9 monl	An authoritative power	monopoly	Control	crown	1000	0	1	0	0,00000
10 monl	The interaction of economical communities		pluralism	commerce	1000	0	1	0	0,00000

Tabla 4.47 Consulta de 'monopolio' Onelook

Clave	Definición libre de 'huracán'	Fase 2	1º Lugar	2º Lugar	Total	L.Ref	F	P	G
1hurl	A kind of severe storm which involves a lot of strong winds	hurricane, monsoon, gale, typhoon, tomado	Blizzard	gale	1000	4	0,003	1	0,00100
2hurl	A violent tropical cyclone	typhoon, hurricane	Typhoon	Tomado	1000	3	0,002	1	0,00100
3 hurl	Very strong winds of over 100 mph which cause great destruction	gales, tornadoes, hurricanes	Aim	Pass	1000	662	0,66166	1	0,00100
4 hurl	Strong winds and rain, gale, destruction over large areas	hurricane, tomado	Storm	tomado	1000	77	0,07608	1	0,00100
5 hurl	A very strong wind, very destructive, often named after people's first names, i.e. the weather researchers who first discover each one	hurricane?	Beat	point	1000	519	0,51852	1	0,00100
6 hurl	Big amount of wind going from country to country, ruining everything as it goes!	cyclone, hurricane	High	return	1000	670	0,66967	1	0,00100
7 hurl	Violent winds which can cause large scale destruction	Gale	Wind	turn	1000	105	0,1041	1	0,00100
8 hurl	Natural disaster, whipping wind, destroys anything in its path	Tomado	Reverse	whip	1000	51	0,05005	1	0,00100
9 hurl	A type of typhoon	Hurricane	Tropical cyclone	Severe weather	1000	0	1	0	0,00000
10 hurl	A very very strong wind	Hurricane	Blast	High wind	1000	192	0,19119	1	0,00100

Tabla 4.48 Consulta de 'huracán' en Onelook

Clave	Definición libre de 'barómetro'	Fase 2	1º Lugar	2º Lugar	Total	L.Ref	F	P	G
1barl	Instrument which detects changes in atmospheric pressure	barometer	Statoscope	baroscope	1000	5	0,004	1	0,00100
2barl	An instrument for measuring pressure	barometer	Barometer	Manometer	1000	1	0	1	0,00100
3barl	Like thermometer but measures air pressure	barometer	Weather situation	collateral	1000	67	0,06607	1	0,00100
4barl	Used to measure air pressure	barometer	Manometer	Barometer	1000	2	0,001	1	0,00100
5barl	A device to measure pressure	barometer	Sphygmomanometer	tonometer	1000	46	0,04505	1	0,00100
6barl	Instrument used in physics to measure pressure	barometer	Gauge	tonometer	1000	5	0,004	1	0,00100
7barl	An instrument used to measure pressure, used in physics	barometer	Gauge	Tonometer	1000	5	0,004	1	0,00100
8barl	A device used to measure air pressure	air gauge	Sound	sphygmomanometer	1000	37	0,03604	1	0,00100
9barl	Measures pressure		Anaeroid barometer	Anaeroid	1000	3	0,002	1	0,00100
10barl	Used to measure pressure, e.g. weather reports		High	Barometer	1000	2	0,001	1	0,00100

Tabla 4.49 Consulta de 'barómetro' Onelook

Término	F	P	G
water	0,005205	1	0,001
lemon	0,001401	1	0,001
squirrel	0,013012	1	0,001
bucket	0,196295	1	0,001
Bench	0,14084	0,9	0,0009
Clothes	0,368369	0,7	0,0007
Euthanasia	0,617717	0,4	0,0004
Monopoly	0,513012	0,6	0,0006
Hurricane	0,327627	0,9	0,00090
Barometer	0,016316	1	0,00100

Tabla 4.50 Promedio de Onelook



## **Apéndice del Capítulo 5**

### **Diseño de búsquedas onomasiológicas en lenguaje natural**

En este apéndice se podrán consultar:

- El análisis que se realizó para determinar del corpus de definiciones de desastres y física las palabras cerradas y ambiguas. Para estas últimas se indica además para qué términos resultan ambiguas.

## Análisis de Palabras Cerradas y ambiguas

Preposiciones	Ejemplos y análisis
AL	<p>“...se propaga en forma de radiación y <i>al</i> ser reflejada por los objetos...”</p> <p>Es la contracción de la preposición <i>a</i> y el artículo <i>el</i> que son palabras muy frecuentes y que no aportan información importante a la definición. Palabra cerrada.</p>
BAJO	<p>“...horas después estará en su nivel más <i>bajo</i>. A este subir y bajar periódico del...”</p> <p>Especifica la posición de un objeto, está adjetivando a <i>nivel</i>, por lo que es relevante, en este caso. Palabra ambigua para el término <i>marea</i>.</p> <p>“...como se mueven los objetos <i>bajo</i> la acción de una fuerza...”</p> <p>En los demás casos, la preposición <i>bajo</i> puede ser sustituida por <i>con</i> que no aporta significado a la definición. Palabra cerrada.</p>
CON	<p>“...ejercida por A sobre B va aparejada <i>con</i> otra fuerza...”</p> <p>La preposición es frecuente y no es indispensable para definir al término. Por ello, en todos los casos analizados es una palabra cerrada.</p>
DEL	<p>“...de las caídas de los cuerpos y <i>del</i> movimiento...”</p> <p>Contracción de la preposición <i>de</i> y el artículo <i>el</i>. No aportan significado. Palabra cerrada.</p>
EN	<p>“Calor disponible <i>en</i> la electricidad...”</p> <p>La preposición no ayuda a definir al término, por lo que es una palabra cerrada.</p>
PARA	<p>“...el experimento se diseñó <i>para</i> medir la velocidad de la Tierra...”</p> <p>En todos los casos en que aparece la preposición se encontró que al omitirse no altera el significado de la definición, por lo que es palabra cerrada.</p>
POR	<p>“Premio Nobel en 1921 <i>por</i> su ley del efecto...”</p> <p>Preposición muy frecuente. En todos los casos en los que aparece funciona como nexos, por lo que es una palabra cerrada.</p>
SIN	<p>“Movimiento oscilatorio respecto a la posición de equilibrio <i>sin</i> rozamientos...”</p> <p>En todos los casos es importante porque indica ausencia o carencia de algo lo cual añade significado a la definición. En principio no debe considerarse como palabra clave sino como elemento que buscaría el antónimo de la palabra clave a la cual está negando. Sin embargo, por los alcances de este trabajo se consideró como palabra cerrada. Lo mismo sucede con partículas similares, como <i>no</i>, <i>nunca</i>, <i>ningún</i>, <i>ninguna</i>.</p>
SOBRE	<p>“...la tendencia a deslizar y se encuentra <i>sobre</i> las superficies en contacto.”</p> <p>En este caso la preposición significa <i>encima</i> y viene reforzada por el verbo que la precede que puede ser deslizar o desplazar. Palabra ambigua para los términos <i>fuerza de fricción</i>, <i>fricción</i>, <i>fuerza de fricción cinética</i>, <i>coeficiente de fricción</i>, <i>fricción cinética</i>, <i>fuerza de fricción estática</i> y <i>gravedad</i>.</p> <p>“Marco en el que un cuerpo, <i>sobre</i> el cual no actúen fuerzas netas...”</p> <p>En todos los demás casos encontrados, la preposición no aporta mayor significado y se puede sustituir por las preposiciones <i>en</i> o <i>de</i>, por lo que es una palabra cerrada.</p>
<b>Artículos</b>	<b>Ejemplos y análisis</b>
LA, LAS, EL, LOS, UN, UNA, UNOS, UNAS	<p>“Plano inclinado es <i>un</i> dispositivo para multiplicar una...”</p> <p>Artículos que no son relevantes debido a que indican si el sustantivo es conocido por el que habla, pero esto no ayuda a definir el término. Palabras cerradas.</p>



## Conjunciones

AUNQUE

## Ejemplos y análisis

"...la cual se presenta **aunque** no haya deslizamiento..."

Conjunción adversativa que, en todos los casos en que aparece, indica un inconveniente que no impide la realización de una situación. Equivale a *pesar de*, pero para los propósitos de la información recuperada no es útil. Palabra cerrada.

BIEN

"...coordinadas no rota y puede estar fijo o **bien** desplazarse en una dirección..."

En todos los casos encontrados en el corpus, forma parte del nexos disyuntivo *o + bien*, por lo que no altera el significado del término si se omite. En ningún caso del corpus funciona como adverbio. Palabra cerrada.

E, O, U, Y

"Directamente proporcional al producto de sus masas *e* inversamente proporcional al cuadrado..."

Nexos copulativos o disyuntivos que, en todos los casos encontrados, sirven para unir elementos sintácticos, pero esto no es útil para los propósitos de la información recuperada. Palabras cerradas.

NI

"La energía no puede crearse **ni** destruirse, sólo transformarse."

Nexo que une oraciones negativas. A través de la negación, se define al término debido a que es una forma de definir, pero no se consideró palabra clave al igual que *sin* y *no*. Palabra cerrada.

PORQUE

"...esta cuando dos o más ondas se combinan **porque** coinciden en el mismo lugar del..."

Indica causa de algo, pero no sirve para los propósitos de la información recuperada. Palabra cerrada.

PUES

"...la satisfacción de todas sus necesidades, **pues** gracias a la combustión..."

Es una conjunción que indica motivo, pero no añade significado a la definición, por lo que es una palabra cerrada.

SINO

"Los intervalos de tiempo no son absolutos, **sino** que son relativos..."

La conjunción en todos los casos contrapone un juicio positivo con uno negativo que se menciona antes, pero no agrega información que sea importante. Palabra cerrada.

## Adverbios

## Ejemplos y análisis

ADEMÁS

"Cuando un móvil tiene trayectoria circular y **además** si la resistencia al movimiento..."

El adverbio indica que aparte de un hecho sucede otro, pero esto no agrega información relevante a la definición. No hay ningún caso en que dicha palabra indique algo diferente a lo señalado, es una palabra cerrada.

ANTES, DESPUÉS

"...de una pieza sometida a tracción **antes** de romperse..."

Estos adverbios, en todos los casos, indican que algo es anterior o posterior, en tiempo o en espacio, a otra cosa, lo que permite ubicar mejor la definición, por lo que son relevantes. Palabras claves.

ASÍ

"**Así** descubrió que la fuerza ejercida sobre un polo..."

En los casos encontrados funciona como adverbio y equivale a *de este modo* o *de esta manera*. Pero esta información es adicional, por lo que es una palabra cerrada.

CASI

"...las fuerzas que lo originan, abarca **casi** toda la mecánica"

El adverbio indica que el adjetivo al que está modificando prácticamente es igual, pero no del todo. Esa diferencia no lo hace relevante, por lo que es una palabra cerrada.

CERCA

"...cuando el planeta está más **cerca** del sol se mueve más rápido"

El adverbio es importante porque forma parte de las características del término, en todos los casos encontrados. Palabra clave

COMO

"...ser una cosa por la cual es lo que es y como es"

Es importante en este único caso porque es parte de las características de la definición

		<i>cualidad</i> . Palabra ambigua.
		"...los relojes, a esto se le conoce <i>como</i> la dilatación del tiempo." En todos los demás casos encontrados el adverbio está comparando un concepto con otro. Pero para los propósitos de la información recuperada esto no es de utilidad. Palabra cerrada.
DENTRO		"Superficie comprendida <i>dentro</i> de un perímetro" Adverbio que indica la ubicación espacial de algo que aporta información que ayuda a definir al término. Palabra clave.
DIRECTAMENTE		"La fuerza que se le aplica es <i>directamente</i> proporcional..." En el corpus, si se omitiera <i>directamente</i> cambiaría el significado de las definiciones pues es parte de las características del término. Palabra clave.
MIENTRAS		"...cuerpos sigue deslizándose e incrementa su velocidad <i>mientras</i> la fuerza de fricción cinética permanece constante..." El adverbio indica simultaneidad de sucesos, pero para la definición no es importante. Palabra cerrada.
SÍ		"...no inventó el telescopio, como a veces se dice, <i>sí</i> fue el primero en usarlo metódicamente..." Se afirma lo que ya está afirmado, por lo que es una palabra cerrada.
TAN		"Cuando en él existe una fuerza amortiguadora <i>tan</i> grande..." El adverbio sirve para exagerar la idea del adjetivo al que acompaña, pero esto no es relevante para la definición, por lo que es una palabra cerrada.
TOTALMENTE		"...para el segundo sucede otro evento <i>totalmente</i> distinto..." El adverbio no es indispensable, en todos los casos, porque la totalidad que indica es irrelevante frente a la palabra clave a la que acompaña. Palabra cerrada.
<b>Pronombres</b>		<b>Ejemplos y análisis</b>
ALGO		"...personas acometen o se lanzan contra <i>algo</i> ..." Pronombre que se refiere a una cosa, objeto o cuerpo que no está definido por el que habla, siendo así irrelevante. Además, en caso de que se considerara clave, entonces equivaldría a muchísimos sustantivos, por lo que traería demasiados términos en una búsqueda. Por tanto, se queda como palabra cerrada.
AQUEL		" <i>Aquel</i> en que la velocidad..." El pronombre está sustituyendo a un sustantivo, pero no adquiere su significado, en todos los casos encontrados, por lo que es una palabra cerrada.
ELLA		"De figura de elipse o parecido a <i>ella</i> ..." En todos los casos el pronombre sustituye a un sustantivo, pero no adopta su significado. No es una palabra relevante; por tanto, es una palabra cerrada.
ÉSTA		"...y de efectos contrarios a los de <i>ésta</i> ..." Sustituye a un sustantivo, pero no adquiere su significado, por lo que es palabra cerrada.
SU, SUS		"...basada en <i>su</i> principio de equivalencia..." "Enunció en 1609 <i>sus</i> dos primeras leyes..." Los pronombres indican posesión, pero para la definición no es relevante. Palabras cerradas.
<b>Adjetivos</b>		<b>Ejemplos y análisis</b>
ALGÚN, ALGUNAS, ALGUNOS	ALGUNA, ALGUNO,	"...de las propiedades de las partículas y <i>algunas</i> de las propiedades de las ondas..." Estos adjetivos, en todos los casos encontrados, se refieren a algo indefinido, por lo que son palabras cerradas.

CERCANO		<p>“...asteroide, en el que se encuentra mas <i>cercano</i> al Sol...”</p> <p>Está dando una ubicación geográfica que es importante, por lo que ayuda a definir al término. Palabra clave</p>
CIERTA		<p>“...sin ser evidente, se admite como <i>cierta</i> sin demostración”</p> <p>El adjetivo significa <i>verdadera</i>, por lo que está agregando una cualidad a la definición <i>postulado</i>. Palabra ambigua.</p> <p>“...una función seno multiplicada por una cierta constante.”</p> <p>Cuando se tenga: <i>cierta</i> + sustantivo se refiere a algo indeterminado. Palabra cerrada.</p>
CIERTAS, CIERTOS	CIERTO,	<p>“...Fenómeno por el que <i>ciertas</i> células adquieren material...”</p> <p>Cuando se tenga: <i>ciertas, cierto, ciertos</i> + sustantivo se refieren a algo indeterminado. Son palabras cerradas.</p>
DEMÁS		<p>“...que la tierra es el centro y los <i>demás</i> astros, incluido el sol, giraban...”</p> <p>Adjetivo que está separando un sustantivo del resto. Pero esto no es importante porque ese uno junto con el resto son de la misma especie. Palabra cerrada</p>
DIFERENTES		<p>“En otra de frecuencia, tensión o intensidad <i>diferentes</i>...”</p> <p>El adjetivo es importante porque especifica que el sustantivo al que acompaña no es igual, es decir, expresa contrariedad. Palabra clave.</p>
DISTINTAS, DISTINTOS		<p>“Relación de semejanza entre cosas <i>distintas</i>...”</p> <p>Son importantes los adjetivos porque agregan una característica al sustantivo que les antecede o les sigue. Indican la diversidad de algo por lo que son parte de la definición. Palabras claves.</p>
ESTA, ESTAS, ESTE, ESTO, ESTOS		<p>“...en <i>este</i> movimiento las partículas...”</p> <p>Los adjetivos sirven para indicar que algo está cercano o conocido al hablante; sin embargo, esta información no agrega significado a la definición, por lo que son palabras cerradas.</p>
GENERAL		<p>“TEORÍA <i>GENERAL</i> DE LA RELATIVIDAD: Considera...”</p> <p>Parte del nombre de un término que además si se excluye se puede confundir con el término <i>Teoría especial de la relatividad</i>. Palabra ambigua en los términos <i>Postulados de Einstein, teoría de la relatividad y Einstein Alberto</i>.</p> <p>“En <i>general</i>, todo lo que tiene sustancia...”</p> <p>Palabra cerrada en los casos en que le precede la preposición <i>en</i> y <i>lo</i>, porque es parte de la información irrelevante.</p>
GRANDES		<p>“...que se caracteriza por la presencia de <i>grandes</i> fuerzas durante un intervalo...”</p> <p>El adjetivo está especificando el tamaño de las <i>fuerzas</i>, por lo que es palabra clave.</p>
IGUAL		<p>“...toda acción corresponde una reacción de <i>igual</i> magnitud y dirección pero de sentido...”</p> <p>Está agregando una característica al sustantivo, en todos los casos en los que aparece, indicando semejanza, por lo que es una palabra clave.</p>
OTRA, OTRAS, OTRO, OTROS		<p>“...forma ángulo recto con <i>otra</i> línea o con <i>otro</i> plano.”</p> <p>Adjetivos que, para los casos observados, no proporcionan ningún significado relevante a la definición, por lo que son palabras cerradas.</p>
POCA, POCO		<p>“Una partícula posee masa pero de tamaño <i>poco</i> significativo.”</p> <p>Los adjetivos indican cantidad de una palabra clave que no es indispensable para la definición. Palabras cerradas.</p>
PRÓXIMO		<p>“Punto de la órbita de un planeta más <i>próximo</i> al Sol...”</p> <p>Indica cercanía y si el adjetivo se omitiera la definición cambiaría su significado. Palabra clave.</p>
TODO		<p>“Conjunto de cosas que forman un <i>todo</i>”</p> <p>Palabra ambigua en el término <i>transformación de la masa</i>.</p> <p>“Es <i>todo</i> cuerpo que puede ser movido bajo la...”</p> <p>En los demás casos encontrados, indica cantidad de una palabra clave, pero no es</p>

## Sustantivos

CONDICIONES

indispensable para la definición. Palabra cerrada.

### Ejemplos y análisis

"...las fuerzas que lo producen y las *condiciones* de equilibrio..."

En varios de los casos analizados implica una situación en la que se encuentra algo y en otros casos la palabra forma parte del término. Es importante porque aporta significado a la definición. Palabra clave.

EJEMPLO

"...las engloba a todas. La química, por *ejemplo*, se ocupa de la interacción de..."

Antecedido de *por* o *como* es una expresión. En ningún caso es relevante, por lo que es una palabra cerrada.

ETC.

"...se aplica a una palanca, polea, torno, *etc.*, para vencer la resistencia."

Indica omisión de lo que falta por escribir, sin embargo, no agrega significado a la definición, por lo que es una palabra cerrada.

## Verbos

DA

"...cuerpo o a un sistema de cuerpos se les *da* energía cuando se efectúa un trabajo..."

No está auxiliando a otro verbo, por lo que es palabra clave.

DADA, DADO

"...aplicada a él que oscila con frecuencia *dada*..."

Relevantes porque equivalen a *determinada* (*o*), es decir, agregan una cualidad al sustantivo. Palabras claves.

DEBE

"...la dirección es importante y usualmente *debe* ser especificada..."

El verbo funciona como auxiliar que forma parte de una perífrasis, por lo que es una palabra cerrada.

DEBIDO, DEBIDA

"Es la energía *debida* a un campo eléctrico..."

*Debido* (*a*) + *a* forman una locución prepositiva. Indican la causa de algo, pero no sirven para los propósitos de la información recuperada. Palabras cerradas.

DEJA

"Un sistema a menudo no se *deja* oscilar por sí mismo, sino que..."

Forma parte de una perífrasis siendo el auxiliar, por lo que es una palabra cerrada.

DICHA

"...su expresión para medir *dicha* energía es..."

En todos los casos, el verbo puede ser sustituido por el pronombre demostrativo *esa* pues se refiere a algo ya mencionado, por lo que es palabra cerrada.

ERA, ERAN, ES

"Es la teoría según la cual la Tierra *era* el centro del Universo..."

Los verbos copulativos están uniendo al sujeto y al predicado sin aportar información relevante, por lo que son palabras cerradas.

HA, HAN

"Cuando un punto *ha* estado sometido a una fuerza dirigida..."

Son verbos auxiliares que no aportan significado, por lo que son palabras cerradas.

HAYA

"Es la fuerza paralela a dos superficies que se encuentran en contacto, la cual se presenta aunque no *haya* deslizamiento..."

Irrelevante porque es un verbo auxiliar. Palabra cerrada.

HECHA

"Aquello de que una cosa está *hecha*; sustancia extensa, divisible..."

Relevante porque forma parte de una perífrasis siendo el verbo principal. Palabra clave.

HUBIERA

"...permanentes debido a la colisión, o si *hubiera* producido calor durante..."

Irrelevante porque es un verbo auxiliar. Palabra cerrada.

IR

"El aumento o disminución de la energía calorífica debe *ir* acompañado por una..."

Irrelevante porque forma parte de una perífrasis siendo uno de los auxiliares. Palabra cerrada.

PARTIR		<p>“...reposo absoluto no puede determinarse a <i>partir</i> de la posición de los cuerpos en...”</p> <p>En todos los casos encontrados es relevante, debido a que equivale a <i>desde</i>. Palabra clave.</p>
PODEMOS, PUDO	PODER,	<p>“A partir de los cambios de velocidades nosotros <i>podemos</i> determinar...”</p> <p>Forman parte de una perífrasis verbal que agrega un matiz de posibilidad, pero esto no ayuda a definir el término, por lo que son palabras cerradas.</p>
RESULTA		<p>“Cociente que <i>resulta</i> de dividir el límite de adherencia ...”</p> <p>Indica que es el producto de algo, por lo que es una palabra relevante. Palabra clave.</p>
SALE		<p>“En un círculo, línea recta que <i>sale</i> de su centro y llega a un punto...”</p> <p>Es un verbo que indica dirección de adentro hacia afuera, por lo que es una palabra clave.</p>
SIGUE		<p>“Camino o rumbo que <i>sigue</i> algo en su movimiento...”</p> <p>El verbo no es auxiliar de otro, por lo que es un palabra clave.</p>
TIENEN		<p>“Marco cuyos ejes <i>tienen</i> una dirección fija...”</p> <p>No indica movimiento sino posesión. Palabra cerrada.</p>



## **Apéndice del Capítulo 8**

### **Pruebas, implantación y evaluación del sistema**

En este apéndice se podrá consultar:

- La lista de términos y definiciones del Banco Terminológico del área de Fenómenos que se utilizaron en el análisis del capítulo.

## Términos y definiciones del Banco Terminológico del área de Fenómenos

Término	Definición
ACCIDENTE	Es el acto que altera el curso regular de las cosas, puede ser de origen humano o natural, no premeditado, aunque muchas veces previsible, que se presenta en forma súbita o inesperada y que causa la muerte o lesiones de las personas y daños a sus bienes.
ACCION BELICA	Es una acción de lucha armada entre los habitantes de un mismo pueblo, ciudad, de grupos sociales que están en oposición de intereses o entre bandos de una misma nación o de dos o más naciones.
CALOR	Es el alza de temperaturas que provoca deshidrataciones entre niños, mujeres, etc., descomposición de alimentos y otros.
COLAPSO DE SUELOS	Es la falla o hundimiento de una zona, ya sea por carga propia o externa. Su principal característica es que se trata siempre de un desplome vertical.
CONTAMINACION	Es la presencia en el medio ambiente de uno o más elementos que perjudican o molestan la vida, la salud y el bienestar humano, la flora, la fauna o degradan la calidad del aire, del agua, de la tierra, de los bienes.
CRECIMIENTO EXPLOSIVO DE LA POBLACION	Es el crecimiento desproporcionado de una ciudad, ya sea por la inmigración constante de personas, por la mala planificación de la familia, etc. Es el resultado del comportamiento de las variables demográficas (natalidad, mortalidad y migraciones).
DESERTIFICACION	Cambio ecológico que despoja a la tierra de su capacidad para sostener y reproducir vegetación, actividades agropecuarias y condiciones de habitación humana. Dicho cambio es ocasionado por el empobrecimiento del suelo de una región debido a excesos o errores en la explotación de sus recursos o a la evolución natural del clima.
DESLIZAMIENTO DE TALUD	Movimiento hacia debajo de las partículas y porciones del suelo próximas a la superficie libre e inclinada de un talud, ocasionado por la falta de presión normal confinante que ahí existe.
DISTURBIOS SOCIALES	Son originados por el hombre, por desacuerdos en las disposiciones gubernamentales o patronales, esto se refleja en las manifestaciones, huelgas, revueltas, etc.
ENVENENAMIENTO	Es la introducción en el organismo de cualesquiera de los tóxicos de naturaleza química o biológica en cantidades dosificadas, diferentes en cada caso; causa trastornos de carácter grave o mortal.
EPIDEMIA	Es la extensión de una enfermedad infecto-contagiosa, a escala local, regional e incluso mundial y que afecta a la mayoría de individuos de la zona en la que se desarrolla. Las malas condiciones higiénicas favorecen el desarrollo de las epidemias.
EROSION	Deterioración progresiva de la calidad o formas de la tierra como resultado de fenómenos naturales o actividades humanas.

EXPLOSION	Es la liberación violenta, rápida e irreversible de energía. La existencia de productos explosivos en las industrias (materia prima, productos químicos, etc.), da origen a las explosiones
FALLA O ERROR HUMANO	Es la falla ocasionada por el ser humano en forma involuntaria o por descuido, de la cual resultan malas operaciones en los servicios, accidentes que pueden causar heridas graves o pérdidas de la vida, errores de diseño, etc.
FUGA Y DERRAME DE SUSTANCIAS	Liberación accidental ocurrida durante
HELADA	Es una congelación producida por una fuerte baja de temperatura. Las heladas provocan la muerte de personas carentes de vivienda.
HUNDIMIENTO REGIONAL	Es un fenómeno que se genera en los suelos blandos, produciéndoles pérdida de volumen
HURACAN	Es un conjunto de centros de depresión atmosférica con trayectorias parabólicas, generalmente de sureste a noreste, que operan como centros de atracción de masas de aire caliente y húmedo.
INCENDIO	Es la ignición de materiales combustibles y la extensión del fuego resultante en industrias, viviendas, etc. Las zonas más susceptibles a incendios son las concentraciones industriales.
INUNDACION	Es el flujo o encubrimiento de agua que se origina por las tormentas y desbordamientos de ríos y se agrava por la deficiencia de capacidad del drenaje y por la acumulación de basura en las captaciones de aguas pluviales, por las fallas en los bordos
LLUVIA	Precipitación de gotas de agua líquida o sobreenfriada que produce consecuencias desastrosas cuando son de gran intensidad.
MAREMOTO	Movimiento intenso, violento y brusco de las aguas del mar producido por un terremoto o una erupción volcánica en éste, que origina grandes olas que provocan devastaciones e inundaciones en las zonas costeras.
NEVADA	Precipitación atmosférica sólida en pequeños cristales de hielo en forma hexagonal o estrellada que se reúnen en grupos formando copos.
PLAGA	Es el crecimiento desproporcionado de una especie animal, a escala local, regional e incluso mundial; que afecta a personas, bienes materiales, plantaciones, etc.
RADIACION	Es un proceso de diseminación de energía en forma de onda (rayos x, rayos gama), o de partícula atómicas (electrones, protones, neutrones), núcleos de diferentes elementos tales como el helio, por los equipos o materiales radioactivos. Se distinguen tres tipos de radioactividad natural, artificial e inducida.
SABOTAJE	Es el acto material destinado a impedir el funcionamiento normal de un servicio o de una empresa, o a inutilizar una máquina o una instalación.
SEQUIA	Es el fenómeno que se caracteriza por la falta de agua suficiente en el suelo y que impacta la vegetación que pierde el agua por la evaporación y puede ser debida a que se agota la humedad del suelo o a que la precipitación en un periodo es menor a su promedio característico. Cuando esta deficiencia es grande y prolongada daña las actividades humanas; se dividen en diferentes



	tipos: permanente, estacional, contingente o accidental e invisible.
SISMO	Es un conjunto de movimientos de tierra causados por desacomodamientos de placas de la corteza terrestre, que se producen por los esfuerzos generados en la corteza y el manto superior por las corrientes ascendentes y descendentes.
TERRORISMO	Es una sucesión de actos de violencia ejecutados para infundir terror. Generalmente son acciones que atentan contra la vida, la salud y los bienes materiales los demás.
TORMENTA DE GRANIZO	Es agua congelada que desciende con violencia de las nubes en granos mas o menos duros y gruesos.
TORMENTA ELECTRICA	Es una chispa eléctrica gigante entre dos zonas cuyas cargas eléctricas (positivas y negativas)
VULCANISMO	Es la expulsión de magma del interior de la corteza terrestre a la superficie a través de uno a más cráteres y es generalmente acompañado por sismos.



## **Bibliografía**

- Aitchison, J.** 2001. *Words in the mind. An introduction to the mental lexicon*, Oxford and Cambridge.Blackwell.
- Alice Y. H. T. et al.** 1990. *Sistemas de bases de datos, administración y uso*. México, Prentice Hall Hispanoamericana.
- Allen, J.** 1995. *Natural Language Understanding*, Redwood City, The Benjamin/ Cummings Publishing Company, Inc.
- Arnold, et al.** 2001. *El lenguaje de programación Java*, México. Adison Wesley.
- Ayto, J.** 1983. "On specifying meaning" en *Lexicography: principles and practice*, R.R.K Hartmann (ed), London, Academic Press, pp. 89-98.
- Baldinger, K.** 1970. *Teoría semántica. Hacia una semántica moderna*, Alcalá, Madrid, pp. 119-131.
- Bell, D. et al.** 2003. *Java para Estudiantes*. México, Pearson Educación.
- Bizzarri, H.** 2000. "Prefacio" en *Diccionario paremiológico e ideológico de la Edad Media (Castilla, siglo XIII)*, Buenos Aires, SECRET, pp i-iv.
- Ceballos, F. J.** 2000. *Java 2 Curso de Programación*, México, Alfaomega.
- Cuarón, B.** 1999. *Significado*, México, UNAM (Centro de Investigaciones Interdisciplinarias en Ciencias y Humanidades: col. Conceptos).
- Date, C. J.** 2001. *Introducción a los sistemas de bases de datos*. México, Adison Wesley Iberoamericana.
- Engelkamp, J.** 1983. "Word meaning and word recognition" en *Concept development and the development of word meaning* (Springer Series in language and communication vol.12) Editors TH.B Seiler, W. Wannemacher. Springer-Verlag, Berlin, Heidelberg New York, Tokyo, pp.17-33.
- Fodor, J.** 1999. *Conceptos (donde la ciencia cognitiva se equivocó)*, Barcelona, Gedisa.
- Grossmann, R.** 1975. "Definite descriptions" en *Philosophical studies 27*, Reidel Publishing Company, Dordrecht-Holland, pp. 127-144.
- Hanks, P.** 2000. "Contributions of Lexicography and Corpus Linguistics to a Theory of Language Performance" en *Proceedings of Euralex*, Oxford, pp 3-11.
- Hansen, G. W., et al.** 1996. *Database Management and design*, New Jersey, Prentice Hall. Segunda Edición.
- Hartzenbusch, J.** 1855. "Prólogo" de en *Colección de sinónimos de la Lengua castellana*. Visor Libros. Madrid, pp. XVII- XXII.
- Kendall y Kendall** 1998, *Análisis y diseño de sistemas*, México, Hispanoamericana.

- Lara, L.** 1996. *Teoría del diccionario monolingüe*, COLMEX, México.
- Martin, J.** 1977. *Organización de las bases de datos*, México. Prentice-Hall Hispanoamericana, S.A.
- Martín, T. M. A.** 2000. *Manual imprescindible de Java 2*, Madrid, Anaya Multimedia.
- Moreno, L.** et al. 1999. *Introducción al procesamiento del lenguaje natural*, Universidad de Alicante, Alicante.
- Pederson, E.** y **J. Nuyts** 1999. "1 Overview: on the relationship between language and conceptualization" en *Language and conceptualization*, Jan Nuyts and Eric Pederson (ed). Cambridge University Press.
- Pressman, R. S.** 1988. *Ingeniería del Software*. Ed. McGraw-Hill, 3ª. Edición. Madrid.
- Putnam, H.** 1975. "The meaning of meaning" en *Mind language and reality* (Philosophical papers, vol.2), Cambridge University Press, pp. 215-271.
- Putnam, H.** 1983. "¿Es posible la semántica?. Tr. Alejandro Tomasini, en *Cuadernos críticos*, México, UNAM.
- Putnam, H.** 2000. *Representación y realidad*, Barcelona, Gedisa.
- Robinson, E.** 1999. "11 The cognitive foundations of pragmatic principles: implications for theories of linguistic and cognitive representation" en *Language and conceptualization*. Jan Nuyts y Eric Pederson (ed). Cambridge University Press.
- Rosenberg, J.** 1994. *Beyond formalism: Naming and necessity for human beings*. Philadelphia, Temple University Press.
- Santana, O.** et al., "Herramienta para el manejo de diccionarios ideológicos". Departamento de Informática y sistemas, Universidad de Las Palmas de Gran Canaria, <http://protos.dis.ulpgc.es>
- Seiler, B.** y **W. Wannemacher**, 1983. "How can We Asses Meaning and Investigate Meaning development: Theoretical and Methodological Considerations from an Epistemological Point of View" en *Concept development and the development of the word meaning* (springer series in Language and Communication.(vol 12), Ed. Th. Seiler, W. Wannemacher. Berlin, Heidelberg, New York, Tokyo, Springer Verlag, pp. 320- 339.
- Sierra, G.** 1999. *Design of a concept-oriented tool for terminology*, (Tesis doctoral) Centre for Computational Linguistics, Manchester.
- Sierra, G.** 2000. "The onomasiological dictionary: a gap in lexicography" en *Proceedings of Euralex*, Oxford, pp. 223- 235.

**Sinha, Ch.** 1983. "Background Knowledge, Presupposition and Canonality" *Concept development and the development of word meaning*. (Springer series in Language and communication. Vol.2) pp. 269-294.

**System 11 fast track SQL Server.** *Sybase, The architecture for change*. Student guide.

**Ullmann, S.** 1965. *Semántica: introducción a la ciencia del significado*, Madrid, Aguilar.

**Viegas, E.** 1999. *Text, speech and language technology* (vol. 10), Kluwer Academic Publishers.

**Wiegand, H.E.** 1984. "On the structure and contents of a general theory of lexicography" en *LEXeter '83 Proceedings*. R.R.K. Hartman (ed), pp. 13-30.

**Wierzbicka, A.** 1985. *Lexicography and conceptual analysis*. Karoma Publishers.

### **Páginas consultadas:**

[http://mx.geocities.com/buferpa/bd\\_relacional.htm](http://mx.geocities.com/buferpa/bd_relacional.htm)

<http://www.cs.us.es/cursos/bd-2001/temas/normalizacion.html>

<http://rinconprog.metropoliglobal.com/CursosProg/BDatos/IntroBD/index.php?cap=2>

[http://es.wikipedia.org/wiki/Servidor\\_web](http://es.wikipedia.org/wiki/Servidor_web)

<http://www.um.es/~psibm/tutorial/>

<http://www.iec.csic.es/criptonomicon/java/quesjava.html>

<http://fciencias.ens.uabc.mx/~mjava/Java/carac.html>

<http://www.fisica.uson.mx/~carlos/Java/JavaTut/Cap1/concept.html>

<http://www.mflor.mx/materias/temas/java/java.htm>

<http://www.sai.uam.mx/apoyodidactico/po/Unidad8/poo8.html>

<http://www.monografias.com/trabajos4/basesdatos/basesdatos.shtml>

<http://www.sqlmax.com/caracter.asp>

<http://www.verextremadura.com/miguel/introsql.htm>