

PERCEPCIÓN DE LA CONCRECIÓN Y LA FAMILIARIDAD DEL LÉXICO EN ESTUDIANTES UNIVERSITARIOS. UN ESTUDIO NORMATIVO

MIGUEL A. BARGETTO*
Universidad de Concepción – Chile

BERNARDO RIFFO OCARES*
Universidad de Concepción – Chile

RESUMEN: Los estudios de alfabetización académica han demostrado que las principales falencias que presentan los estudiantes que se incorporan a una comunidad académica están centradas en la dificultad de comprender textos propios de su disciplina. El léxico juega un papel relevante en este punto, puesto la cantidad y la calidad del lexicón personal incidirá en dicho proceso. En el caso de este trabajo, expondremos los resultados del estudio normativo aplicado a estudiantes de la Facultad de Educación de una universidad regional chilena para medir las variables de concreción y familiaridad. Las encuestas que componen dicho estudio fueron creadas a partir de una base de 536 palabras, obtenidas de cinco referencias bibliográficas obligatorias de una asignatura del área de sociología de la educación mediante análisis computacional y depuradas según criterios de extensión, sílaba inicial y frecuencia. Las conclusiones del trabajo arrojan que las relaciones entre ambos índices son lineales, pero no fuertes, lo que valida la hipótesis de que entre más familiar es una palabra, más concreta se percibe por el hablante. Los datos también establecen que las palabras que son percibidas como poco familiares y poco concretas son las menos, aunque su distribución no responde a una curva normal.

PALABRAS CLAVE: Psicolingüística, estudios normativos, frecuencia léxica, familiaridad, concreción.

*Para correspondencia, dirigirse a: Miguel A. Bargetto (mbargetto@udec.cl), Bernardo Riffo Ocares (bernardo@udec.cl)

PERCEPTION OF CONCRETENESS AND FAMILIARITY OF LEXICON AMONG COLLEGE STUDENTS. A NORMATIVE STUDY

ABSTRACT: Academic literacy studies have shown that the main weakness manifested by students who join an academic community are focused on the difficulty of texts comprehension from their discipline. The lexicon plays an important role on this point, since the quantity and quality of the personal lexicon will influence this process. In the case of this paper, we will present the results of the normative study applied to students of the Faculty of Education of a Chilean regional university to measure the variables of concreteness and familiarity. The surveys that make up this study were created from a base of 536 words, obtained from five compulsory bibliographical references of an area of sociology of education through computational analysis and refined according to criteria of extension, initial syllable and frequency. The conclusions of the paper shows that the relationships between both indices are linear, but not strong, which validates the hypothesis that the more familiar a word is, the more concrete it is perceived by the speaker. The data also establish that the words that are perceived as unfamiliar and not very specific are the least, although their distribution does not respond to a normal curve.

KEYWORDS: Psycholinguistics, normative studies, lexical frequency, lexical familiarity, concreteness.

1. INTRODUCCIÓN

El presente artículo tiene como finalidad dar cuenta de un estudio normativo efectuado para recopilar un conjunto de palabras que serán utilizadas posteriormente en la realización de pruebas experimentales cronométricas de reconocimiento y procesamiento léxico. Este trabajo forma parte de la tesis doctoral del primer autor y se enmarca en una línea de investigación psicolingüística¹ sobre la relación entre habilidades léxicas, por un lado, y el desempeño lector, por otro, en el contexto de los procesos de alfabetización académica que universidades chilenas desarrollan con sus estudiantes novatos.

Una de las mayores dificultades que los participantes de esta investigación -estudiantes novatos de carreras de pedagogía- revelan es la dificultad para entender las lecturas contenidas en la bibliografía de asignaturas regulares que forman parte del plan de estudios de dichos programas de formación profesional. Este problema alcanza su mayor nivel en la asignatura de *Fundamentos socioculturales de la educación*, que ha tenido resultados críticos durante los últimos años, con una tasa de aprobación inferior al 60%². Los propios académicos que dictan dicha materia

¹ El Grupo de Psicolingüística de la Universidad de Concepción incluye, dentro de sus líneas de trabajo, el estudio de la relación entre el dominio del vocabulario y el desempeño en pruebas de comprensión lectora.

² Los estudiantes pertenecen a una universidad privada no selectiva -que cumple con las exigencias legales para la admisión de estudiantes en carreras de pedagogía-. Los índices son preparados por unidades especializadas de la Vicerrectoría Académica para que a partir de ellos las Direcciones de Escuela generen planes de acción para resolver las debilidades manifestadas en el rendimiento.

señalan que los estudiantes tienen graves falencias para comprender los textos y que la principal debilidad para alcanzar la meta de lectura es la pobreza de su vocabulario. Este contexto motiva a revisar los índices de *lecturabilidad* de los textos –entendida como la dificultad/facilidad intrínseca de un texto para su comprensión– y a enfocarnos en la complejidad léxica como factor relevante –pero no excluyente– en la explicación de las dificultades de comprensión en este contexto específico.

2. PRINCIPIOS GENERALES

2.1. La alfabetización académica y la lectura

La alfabetización académica ha sido ampliamente estudiada en las últimas décadas. Los resultados de las investigaciones realizadas en universidades de distintos países han señalado que los nudos críticos más relevantes se ubican en la lectura y la escritura académicas (Carlino, 2003), al mismo tiempo que señalan la relevancia de su papel en la formación de los estudiantes y la conciencia que se debe tener respecto del modo de incorporarlos a la comunidad académico-discursiva. Parodi y cols. (2010) han insistido en que las características del género discursivo académico no pueden desatenderse, por cuanto implican una graduación de la complejidad misma del texto. En este sentido, el contexto de circulación de los textos proporciona valiosa información para comprender si lo leído está al alcance del lector, en este caso, el universitario novato. Sin embargo, las peculiaridades de los géneros académicos no siempre son tan claras de reconocer, especialmente en lo que a léxico se refiere. En efecto, Schröder (1991) señalaba que la lengua con propósitos específicos es un subsistema lingüístico que utiliza recursivamente palabras y estructuras del lenguaje corriente. Por lo tanto, no podemos argumentar que el léxico de los géneros académicos sea exclusivamente especializado –como lo sería un *tecnolecto*–, sino que se vale de las mismas palabras, aunque con configuraciones proposicionales y sintácticas diferentes. Consecuentemente, los géneros académicos presentan especificidades lingüísticas a las que deben sumarse factores sociocognitivos (Parodi *et al.*, 2010) y, sobre estas características, diseñar programas de alfabetización académica *ad hoc* para estudiantes de diferentes programas académicos y áreas del conocimiento.

Las acciones de alfabetización académica requieren, por lo tanto, conciliar las particularidades específicas de los textos y, a partir de ellas, generar planes de mejora. La comprensión lectora es uno de los nudos críticos de este proceso y, a la vez, el que ha generado más evidencia, incluyendo la comprensión de textos especializados (Neira, 2015; Neira, Riffo y Reyes, 2015). Asimismo, se ha medido el efecto de la complejidad textual como factor gravitante en la comprensión (Campos, Contreras, Riffo, Véliz y Reyes, 2014). En esta línea investigativa, pensamos que considerar el factor léxico contribuirá a profundizar nuestro conocimiento sobre el papel del vocabulario en la comprensión de textos académicos en el contexto de la educación terciaria. Perfetti (2010) propone el *triángulo dorado de la comprensión*, donde el vocabulario sirve de pivote entre la decodificación y la comprensión, puesto que,

sin acceso léxico, la decodificación se torna vacía al no existir una representación conceptual que consecuentemente permita generar el proceso mental de integración de la información nueva. Los efectos del léxico en la comprensión han sido ampliamente estudiados, especialmente en sus características semánticas. En ese sentido, Verhoeven y Perfetti (2003) señalan que la palabra es el enlace entre la comprensión del texto y la construcción del conocimiento, lo que trae como consecuencia el aumento del léxico³, debido a que se establece un vínculo estrecho entre el significado y la forma, generándose en aquel una restricción que conlleva a una mayor precisión. En palabras de los autores, se pasa desde acepciones gruesas a acepciones finas, donde estas últimas constituyen un indicio de la *calidad léxica*, la que representa un factor fundamental para los procesos de alfabetización académica: entre más palabras conocidas tiene disponibles el lector, mayor facilidad debería existir al momento de enfrentarse a una tarea de lectura académica. En efecto, el aprendizaje de nuevas palabras es considerado una habilidad de lectura que tiene dos consecuencias: un aumento en la cantidad de representaciones (más palabras) y un aumento en la calidad de representación del significado. Entonces, más lecturas aportan nuevas acepciones (no siempre más palabras), debido a la exposición a dicho léxico, el que va afinándose a medida que más se lee. Este aporte es coherente con la evidencia psicolingüística respecto de la posibilidad de almacenar en la memoria todos los significados de una palabra (Domínguez & Estévez, 1999).

2.2. Las palabras: su reconocimiento y características

El reconocimiento de palabras es un proceso mental complejo, que ha motivado el desarrollo de varios modelos teóricos para explicarlo (véase para una revisión, Belinchón, Rivière & Igoa, 1992; De Vega y Cuetos, 1999). Entre los más conocidos se encuentran el *Modelo del Logogén* (Morton, 1969), que propone el acceso léxico como proceso de activación de representaciones a partir de información estimular (escrita o auditiva); el *Modelo de Búsqueda* (Forster, 1976), explica el acceso léxico a partir de la metáfora de la biblioteca, es decir, de un ordenamiento de las palabras en la memoria de los individuos; en tercer lugar, el *Modelo de Cohorte*, propuesto por Marslen-Wilson en 1984, se ubica entre el de Morton (interactivo) y el de Forster (serial) y postula que, a partir de un estímulo léxico, se activan serialmente todos los elementos que comparten los rasgos iniciales (fónicos u ortográficos) y luego se desechan interactivamente aquellos que contengan rasgos incompatibles con la palabra. De estos modelos teóricos, se derivan importantes consecuencias psicolingüísticas: el contacto léxico inicial rara vez activa una sola palabra, por lo que se recupera de la memoria un conjunto de *vecinos léxicos*⁴, encabezados por las palabras que tienen mayor frecuencia y son más adecuadas al contexto de uso. En el proceso de selección

³ En Psicolingüística, se denomina *lexicón* al vocabulario que forma parte del conocimiento lingüístico de un individuo y se sitúa en la memoria de largo plazo.

⁴ Los vecinos léxicos corresponden al conjunto de palabras que presentan alta similitud ortográfica entre sí.

de la palabra, los candidatos menos adecuados son abandonados y se configura una “lista” ordenada por frecuencia de uso. La activación de las características post-léxicas entran en juego en esta etapa. Finalmente, cuando se accede a la entrada léxica exacta, se activan las demás características que estaban latentes. Esto implica que el reconocimiento de una palabra exige no solo el proceso de decodificación, acoplando las estructuras fonéticas y morfológicas, sino también su correcta combinación.

Si pensamos, entonces, que el léxico se incrementa a medida que el lector entra en contacto con palabras y significados nuevos, cabe preguntarse cómo las características de las palabras permiten que esta tarea se facilite o dificulte. En este punto, índices psicolingüísticos como *concreción* e *imaginabilidad* cobran especial importancia, ya que inciden en los tiempos de procesamiento, incluso llegando a producirse interferencias entre ellas (Kroll & Merves, 1986; Bleasdale, 1987). Por lo mismo, se puede argumentar que hay rutas de acceso léxico diferido según sea la naturaleza del contenido de la palabra (Pavio, 1991). Asimismo, y junto con la concreción y la imaginabilidad, la *frecuencia* y la *familiaridad* juegan un rol importante en la comprensión. De hecho, la frecuencia se considera la variable más influyente en el reconocimiento de palabras. En su trabajo, Monsalve y Cuetos (2001) consideran una palabra de alta frecuencia a aquella con un índice superior a 170 ocurrencias por millón, mientras que una de baja frecuencia registra frecuencias inferiores a 4 veces por millón. Kroll y Merves (1986), para su trabajo en inglés, establecen como umbral una ocurrencia de 70 veces por millón. Tal es la importancia de esta propiedad, que se han elaborado diccionarios de frecuencias léxicas, como el de Alameda y Cuetos (1995), el LIFCACH (Sadovsky y Martínez-Gamboa, 2012) o bien bases de datos *online* como el CORPES XXI, disponible en el sitio de la RAE, que entregan índices a partir de distintos corpus de textos.

Concreción y *abstracción*, a su vez, constituyen índices gravitantes para medir los tiempos de procesamiento cognitivo del lenguaje, puesto que la presencia o ausencia de referentes extra lingüísticos incidirá en dicho procesamiento. Para Iliev y Axelrod (2016), estos valores pueden ser considerados puntos en el espectro de un campo mayor y que no es oponible, sino graduado pues las palabras pueden ser fácilmente accesibles dependiendo de la información disponible en el lexicón. Esto quiere decir que un concepto abstracto puede ser más fácil de procesar precisamente porque la información es menor y más general, lo que se opone a los hallazgos precedentes (Crutch & Warrington, 2005; Vega & Fernández, 2011; Connell & Lynott, 2012) que plantean que un término concreto puede ser fácilmente accesible debido a la ventaja de contar con referentes perceptibles por la experiencia. Esto deriva en el concepto propuesto por los autores: la *precisión*. Lo anterior implica que la cantidad de información permite disminuir la incertidumbre, porque el sujeto efectúa inferencias más precisas sobre la naturaleza conceptual de una palabra. El trabajo de Iliev y Axelrod (2016) ha permitido organizar la semántica de las palabras en nodos que distinguen entre *inclusión* y *profundidad*. La consideración de redes jerárquicas de procesamiento léxico ofrece la posibilidad de clasificar las palabras según sus posibilidades de organización.

En el caso del vocabulario que constituye la bibliografía de la asignatura de *Fundamentos Socioculturales de la Educación*, antes referida, consideramos que esta dimensión nos puede aportar un interesante enfoque en la robustez del lexicón mental de los participantes en el estudio. Desde el modelamiento propuesto más arriba, pensamos que la familiaridad de la palabra estará en directa relación el grado de concreción que los individuos le asignen. Por lo mismo, el objetivo de este estudio es normar los valores de *concreción* y *familiaridad* a fin de obtener un indicador específico para los estudiantes universitarios de ciencias de la educación que permita construir luego pruebas cronométricas, controlando los indicadores mencionados como variables que inciden en la respuesta experimental.

2.3. Los estudios normativos

Los estudios normativos se emplean en la investigación psicolingüística debido a la necesidad de conocer las características de los estímulos que se utilizarán en experimentos. Su validez no es absoluta, toda vez que la información puede variar según los participantes, el tiempo y el lugar donde se recoja la muestra. Por lo mismo, cada uno entrega valores actualizados y precisos para la comunidad en la cual se realizó el estudio. En su recuento de diversos estudios en español, Pérez, Campoy y Navalón (2001) encontraron números variables de sujetos y estímulos para normar; en el caso de los estudios sobre *concreción e imaginabilidad*, se da cuenta de trabajos que consideran desde 120 hasta 2300 palabras y de 100 a 2000 sujetos. En el cuadro se puede apreciar una muestra de los realizados específicamente sobre concreción:

Autores	Palabras	N.º participantes	País
Blanch y Bañados (1996)	125 palabras	212 adultos	España
Campos y Astorga (1989)	300 palabras	100 adultos	España
Moltó y cols. (1999)	480 fotografías	1102 adultos	España
Muñoz Yago (1983) y López Burlo (1985)	2332 palabras	2000 adultos	España

Tabla 1: *Estudios normativos sobre concreción (Tomado de Pérez, Campoy y Navalón)*

3. EL ESTUDIO

3.1. Materiales

Se consideraron seis lecturas obligatorias contempladas en la bibliografía de la asignatura *Fundamentos Socioculturales de la Educación*, correspondientes a cuatro géneros discursivos diferentes: artículo, conferencia, capítulo de libro y blog. Los

textos fueron analizados computacionalmente mediante el etiquetador Connexor[®] y con el programa Trunajod⁵. De los resultados obtenidos, solo fueron considerados los relativos al léxico nocional. Las palabras fueron subclasificadas por su frecuencia, extensión silábica, y configuración de la sílaba inicial.

Del conjunto analizado, fueron excluidos los monosílabos, adverbios terminados en *_mente* y las palabras que a lo largo del texto se repetían. Después de esta reducción, quedaron palabras de tres y cuatro sílabas, de las cuales se seleccionaron, finalmente, 242, puesto que se descartaron aquellas palabras con frecuencias menores a 10 ocurrencias por millón en el límite inferior y 300 ocurrencias por millón en el superior.

Con esta cantidad, se construyó la encuesta para el estudio normativo con tres formas: las formas 1 y 2 con 81 palabras y la 3 con 80, las que fueron presentadas aleatoriamente. Las encuestas estaban construidas como escala tipo Likert, que evaluaban la percepción de los estudiantes respecto de ambos índices. La encuesta presentaba dos preguntas por cada una de las palabras: la primera pregunta señalaba “escucho y empleo la palabra xxxx” seguida de “me es fácil imaginar el concepto de la palabra xxxx”. La encuesta se aplicó en formato papel a estudiantes de segundo semestre, durante tres jornadas de clases. Previamente, el instrumento fue piloteado con alumnos de curso superior para calibrar las preguntas y el tiempo de aplicación.

3.2. Participantes

Tomaron parte en la investigación 56 individuos—40 mujeres, 11 hombres—, estudiantes de la Facultad de Educación de una universidad privada de la ciudad de Concepción (Chile), cuyas edades fluctuaban entre los 18 y los 30 años (promedio: 22,11; SD: 3,59). Seis de ellos no se individualizaron en la muestra, es decir, omitieron género, edad y carrera. Los estudiantes cursaban programas de Educación Diferencial (22), Educación Parvularia (16), Pedagogía en Educación Básica (3) y Pedagogía en Educación Física (10). Del total, el 95% cursaba primer año y el 5% restante, segundo año. A los estudiantes se les informó sobre la naturaleza de la encuesta y no recibieron retribución alguna por su participación voluntaria.

4. RESULTADOS

En primer lugar, detallaremos las características que arrojó el análisis computacional de las palabras nocionales: respecto de la distribución por la cantidad de sílabas y la configuración de sílaba inicial, podemos apreciar que el grueso de las palabras se constituye con aquella más frecuente en español, que es CV, equivalente al 36,9%

⁵ Connexor es una plataforma que desarrolla software para la investigación lingüística. En este caso, se empleó el *demo* “Machine Syntax” que es un etiquetador sintáctico que analiza fragmentos de texto y entrega las categorías gramaticales, dependencias sintácticas entre otras. Por su parte, Trunajod es un programa computacional desarrollado en la Universidad de Concepción que analiza dichas etiquetas para entregar información sobre la lecturabilidad del texto.

de la base, seguida de la configuración CVC, con un 19,8% y en tercer lugar, la configuración VC, con un 15,7%. Otras combinaciones menos frecuentes aportan cantidades menores de palabras, las que no superan el 2% del total. Lo anterior se detalla en la siguiente tabla:

	Sílabas por palabra	1	2	3	4	5	6	Total
Configuración sílaba inicial	CCV		13	20	5	3		41
	CCVC					1		1
	CCVCC			4	1			5
	CV	4	61	66	44	22	1	198
	CVC	9	33	45	14	5		106
	CVCC		1	4				5
	V		14	36	21	6		77
	VC	1	10	29	31	12	1	84
	VCC				1	1		2
	CVVC		7		1			8
	CVV		1					1
	VVC			1				1
	CVV	1	4		1			6
	CCVV			1				1
Total		15	144	206	119	50	2	536

Tabla 2: *Distribución de palabras por cantidad y configuración de sílaba inicial*

Las palabras de tres sílabas son las que se presentan en mayor cantidad en las lecturas, con un 38,4%, seguidas de las de dos sílabas, con un 26,9% y las de cuatro sílabas con 22,2%. Estos resultados son coherentes con la literatura, en cuanto a que la mayor cantidad de palabras está concentrada en unidades de dos, tres y cuatro sílabas. Por lo tanto, las lecturas escogidas no escapan a los patrones normales de la constitución léxica.

En cuanto a la frecuencia léxica, se clasificó según lo que recomienda la evidencia científica. Por lo tanto, se consideraron *infrecuentes* aquellas que presentaban un valor menor a 70 veces por millón de ocurrencias. En general, en estos textos académicos el uso de palabras infrecuentes tiene una alta incidencia. Tal como se aprecia en el gráfico, la proporción es de 60-40:

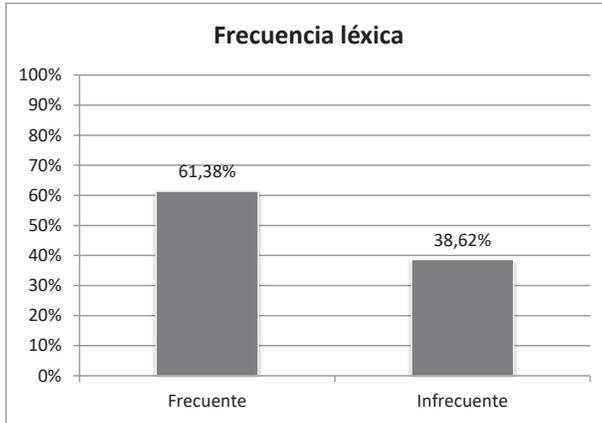


Gráfico 1: *Porcentaje de palabras clasificadas por frecuencia léxica*

Desglosando la frecuencia por la cantidad de sílabas de las palabras consideradas en el estudio, se observan las siguientes curvas:

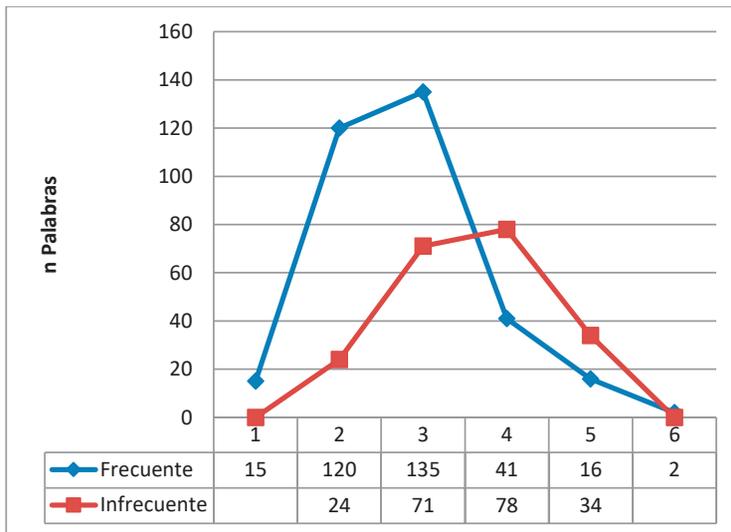


Gráfico 2: *Volumen de palabras frecuentes e infrecuentes según sílabas*

Como se puede apreciar, las palabras frecuentes e infrecuentes aumentan en cantidad a medida que aumenta su número de sílabas, es decir, su incremento es constante. Los valores más altos están ubicados en tres y cuatro sílabas para palabras frecuentes e infrecuentes, respectivamente. Por último, palabras de cuatro y cinco sílabas aparecen en cantidades muy reducidas. En consecuencia, esta distribución ha permitido optar por las de tres y cuatro sílabas, dada su mayor cantidad. Este criterio de selección no corresponde a palabras especializadas de las ciencias de la educación, debido a que el concepto de *léxico especializado* está centrado en acepciones

restringidas solamente al ámbito de la disciplina, situación que en este caso no se da. Dicho de otro modo, el léxico incluido en la bibliografía disciplinar no está emparejado necesariamente con sus acepciones exclusivas del ámbito educativo, por lo tanto, las palabras seleccionadas para el estudio normativo corresponden solamente a aquellas que cumplen las condiciones de frecuencia, extensión silábica y configuración de la sílaba inicial.

Respecto de la familiaridad y la concreción, llama la atención que los resultados de ambos índices tengan valores fluctuantes entre el 3 y el 5. Son pocas las palabras que los estudiantes consideran “no familiares” y “abstractas”, lo que se aprecia en la tabla 2:

Estadísticos familiaridad	
Media	4,23
Error típico	0,02
Mediana	4,28
Moda	4,5
Desviación estándar	0,35
Varianza de la muestra	0,12
Mínimo	2,8
Máximo	4,9
Nivel de confianza (95,0%)	0,045

Tabla 2: *Estadísticos familiaridad*

En cuanto al índice de familiaridad, lo que más llama la atención es la fluctuación de los términos en los tramos de 3 y 5 puntos. A partir de esto, hemos organizado los resultados en tres grandes grupos: palabras de alta, media y baja familiaridad, debido a la clase de respuestas que los sujetos entregaron. Del total de las palabras normadas, 31 obtuvieron un promedio entre 4,5 y 5 (muy de acuerdo), por lo que se consideran de “alta familiaridad”; 192 palabras obtuvieron puntajes entre 4,0 y 4,59, lo que las ubicó en el rango de “familiares” mientras que solo 49 palabras obtuvieron una puntuación promedio inferior a 4, por lo que fueron clasificadas de “baja familiaridad”. Al graficar los datos, apreciamos una curva muy asimétrica, inclinada hacia la derecha, lo que indica claramente que los datos están agrupados y muy concentrados en el parámetro de “familiar-alta familiaridad” por lo que se podría establecer que todos los sujetos son conocedores de las palabras que componen la bibliografía de las lecturas de *Fundamentos Socioculturales de la Educación*, puesto que las medias en ambos índices están sobre 4 y las medianas y modas se alejan de ella, configurando una distribución sesgada, tal como lo muestra el gráfico 3:

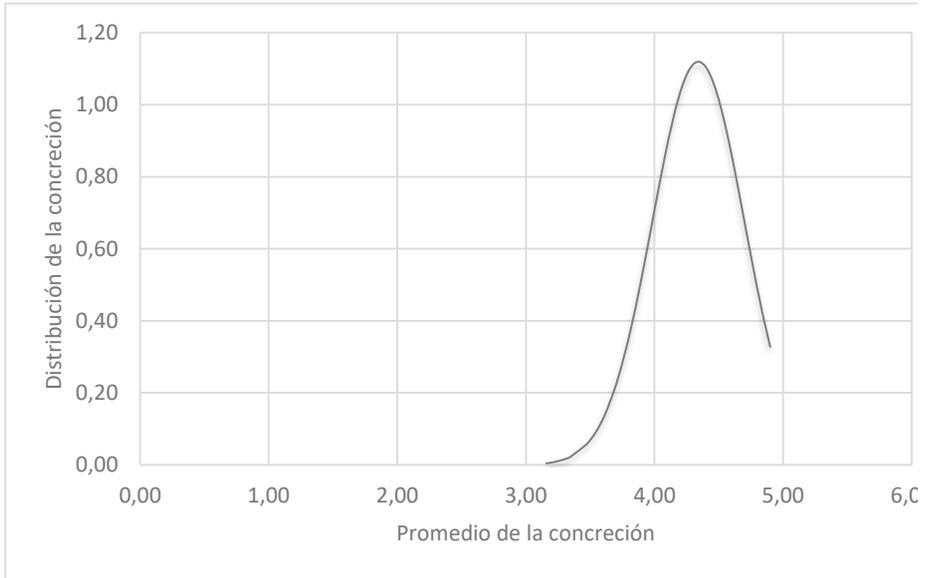


Gráfico 3: *Distribución de la familiaridad.*

En cuanto a la concreción, los estadísticos descriptivos demuestran que las puntuaciones mínimas y máximas son similares a los estadísticos de familiaridad, por ejemplo, la desviación estándar es casi idéntica, lo que señala el comportamiento homogéneo entre ambos índices. Sin embargo, el valor mínimo es ligeramente superior en la concreción, lo que explicaría una menor carga cognitiva que la familiaridad.

Estadísticos concreción	
Media	4,34
Error típico	0,02
Mediana	4,41
Moda	4,61
Desviación estándar	0,36
Varianza de la muestra	0,13
Mínimo	3,1
Máximo	4,9
Nivel de confianza (95,0%)	0,05

Tabla 3: *Estadísticos concreción*

Los resultados de la encuesta para la concreción son similares a los de familiaridad en cuanto a que las palabras se agrupan preferentemente en la percepción de “concreta”, con 142. Las palabras con mayor puntuación, es decir, entre 4,6 y 5 suman 65 y las 36 restantes marcan puntuaciones inferiores a 4, por lo que se consideran “poco concretas”.

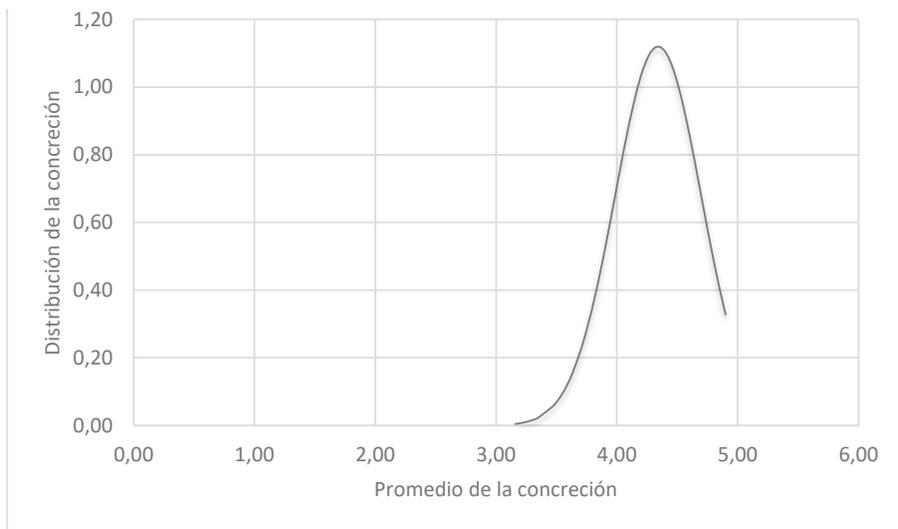


Gráfico 4: *Distribución de la concreción.*

Este aparente “conocimiento” del léxico por parte de los estudiantes podría implicar que la espontaneidad de la respuesta no guarda relación con el significado mismo de la palabra interrogada ni con la cantidad de acepciones que pudiera tener dicha palabra. Por otra parte, tampoco garantiza que la palabra situada en contexto sea cabalmente procesada en una de tarea de comprensión.

Otra característica destacable del estudio fue que tanto familiaridad como concreción están estrechamente relacionadas (prueba r de Pearson = 0.594), por lo que podemos concluir que las palabras recibían la misma puntuación en ambas categorías. No obstante, se puede postular que la familiaridad es una característica de mayor procesamiento, por su valor más bajo, lo que se verifica al comparar los valores mínimos. La valoración que los sujetos realizan de las palabras normadas, asimismo, tiende a ser muy compacta, por cuanto los valores de desviación estándar son muy cercanos a cero y similares (0.1 de diferencia).

En las tablas siguientes se presentan las palabras que obtuvieron menor puntuación en familiaridad y concreción. Las diez palabras con menor puntuación en la familiaridad son todas infrecuentes, sin embargo, se debe considerar que la desviación estándar en cada caso es ligeramente elevada, lo que implica que estas palabras no son concebidas por todos los sujetos como de “baja familiaridad”. En el caso de las palabras de baja concreción, la desviación estándar es más alta, por lo que la posibilidad de que los sujetos tengan percepciones diferentes es mayor.

Palabra	Frecuencia	Promedio familiaridad	Desviación estándar
conferir	Infrecuente	2,8	1,38
occidente	Infrecuente	3,2	1,25
mundialista	Infrecuente	3,3	1,33
obtención	Infrecuente	3,3	1,38
acotar	Infrecuente	3,4	1,39
dimensión	Infrecuente	3,4	1,25
arribar	Infrecuente	3,5	1,47
constar	Infrecuente	3,5	1,00
científico	Infrecuente	3,5	1,54
horizonte	Infrecuente	3,5	1,31

Tabla 4: *Palabras de baja familiaridad*

Palabra	Frecuencia	Promedio concreción	Desviación estándar
aptitud	Infrecuente	3,1	1,42
énfasis	Infrecuente	3,1	1,42
conferir	Infrecuente	3,2	1,59
obtención	Infrecuente	3,3	1,45
constar	Infrecuente	3,3	1,34
arribar	Infrecuente	3,3	1,53
clásico	Infrecuente	3,4	1,50
reforma	Frecuente	3,4	1,26
proveer	Infrecuente	3,5	1,40
contrastar	Infrecuente	3,5	1,43

Tabla 5: *Palabras de baja concreción*

Las palabras que aparecen tanto en la lista de baja familiaridad como de baja concreción son *conferir*, *obtención*, *arribar* y *constar*. Su repetición indica que si frecuencia, familiaridad y concreción confluyen en los niveles inferiores, la dificultad para reconocerlas y asociarlas a un concepto semántico aumenta, lo que se comprueba estadísticamente y que es coherente con investigaciones previas: la familiaridad de las palabras facilita la concreción porque los referentes (o la cantidad de información en los términos que Iliev y Axelrod (2016) proponen) las vuelven más accesibles.

Por otra parte, las diez palabras con mayor puntuación en familiaridad y concreción son:

Palabra	Promedio familiaridad	Desviación estándar
Alumno	4,9	0,31
Futuro	4,9	0,31
Aplicar	4,9	0,37
Entender	4,8	0,51
Atención	4,8	0,41
Ejemplo	4,8	0,52
Éxito	4,8	0,41
Examen	4,8	0,70
Escuela	4,8	0,42

Tabla 6: *Palabras de alta familiaridad*

Palabra	Promedio concreción	Desviación estándar
Alumno	4,9	0,31
Futuro	4,9	0,31
Examen	4,9	0,31
Oferta	4,9	0,31
Éxito	4,8	0,37
Aportar	4,8	0,37
Ayudar	4,8	0,37
Invitación	4,8	0,37
Escuela	4,8	0,37

Tabla 7: *Palabras de alta concreción*

En este comparativo, cinco palabras se repiten *alumno*, *futuro*, *éxito*, *examen* y *escuela*. La alta puntuación promedio obtenida y su baja desviación estándar permite deducir que los sujetos no difieren en el juicio hacia dichas palabras. Asimismo, todas las palabras están asociadas al léxico propio de la formación docente, por lo que su procesamiento se vuelve aún más fácil.

5. DISCUSIÓN

El vocabulario utilizado en las lecturas de *Fundamentos Socioculturales de la Educación* no presenta características fono-ortográficas que otorguen, por sí solas, dificultad a los textos. Esto quiere decir que el proceso de decodificación, al momento de enfrentarse a una lectura de estudio, no debería resultar complejo más allá del desconocimiento de aquellas palabras de baja frecuencia léxica y silábica. Esto es coherente, efectivamente con lo propuesto por Schröder (1991) en cuanto las palabras del léxico para propósitos específicos no constituyen un sector aparte de la lengua. Además, en su conjunto, la mayoría de las palabras es frecuente, lo que *a priori* no representa un obstáculo para la comprensión, ya que su ocurrencia en otros tipos de texto (de naturaleza distinta a la asignatura estudiada) debería facilitar el reconocimiento y, consecuentemente, el acceso léxico, tal como lo postulan los principales modelos teóricos de reconocimiento de palabras. Entonces, analizados los factores de configuración de sílaba inicial y frecuencia léxica, es esperable que las palabras sean fácilmente reconocidas en un primer contacto visual, debido a su familiaridad, la que se refleja en una campana ubicada hacia la izquierda de los valores. Esto nos permite afirmar que la frecuencia léxica, que es el dato objetivo, está relacionada estrechamente con la familiaridad, que es el índice subjetivo.

Si consideramos que la encuesta fue contestada por estudiantes de segundo año y que las palabras fueron obtenidas de lecturas que fueron realizadas en su momento, podemos suponer que no basta con la mera exposición a ellas para que se adquieran en propiedad por los participantes. En este sentido, lo propuesto por Adlof y Perfetti (2012) respecto del incremento del vocabulario está en aparente contradicción, no obstante, este fenómeno se puede explicar porque las lecturas no son objeto de revisión permanente.

6. CONCLUSIONES

La primera conclusión que se puede obtener después del análisis computacional de los textos seleccionados para este trabajo es que, a pesar de pertenecer a géneros discursivos diferentes, no hay grandes diferencias en las características del léxico que los componen. Del mismo modo, el vocabulario no cuenta peculiaridades que permitan clasificarlo “especializado” por cuanto:

- a. La distribución de frecuencia/infrecuencia se divide en proporciones 60-40 por ciento, por lo que la mayoría de las palabras que conforman las lecturas es potencialmente conocida por los lectores previstos (estudiantes universitarios).
- b. Según la conformación silábica de las palabras, 325 cuentan 3 o 4 sílabas, lo que las hace, teóricamente, de fácil procesamiento, por cuanto no exigen un gasto excesivo de recursos cognitivos en el proceso de decodificación, lo que debería redundar en mejores índices de comprensión lectora.

En cuanto a los índices frecuencia, familiaridad y concreción, se postula que existe una relación entre:

- a. Frecuencia léxica y familiaridad: Las palabras que cuentan con frecuencias bajas tienden a ser menos familiares para los sujetos. Sin embargo, esto no está en una relación directa, dado que los valores de desviación estándar en los niveles inferiores se ubican en 1,3 y 1,4, lo que explica diferencias de percepción entre los sujetos. Sin embargo, en los niveles superiores, la baja desviación estándar permite concluir que los sujetos no tienen duda respecto de la familiaridad de la palabra. Dicho de otro modo, las palabras muy familiares presentan dudas respecto de dicha condición, mientras que las menos familiares dependen más de la percepción del sujeto.
- b. Familiaridad y concreción: Los índices de familiaridad y concreción se comportan homogéneamente. La dispersión entre ambos valores es muy baja, lo que indica que la familiaridad permite una mayor concreción de la palabra. Por el contrario, menores índices de familiaridad, afectan la concreción, lo que conllevaría el empleo de más recursos cognitivos para la comprensión.
- c. Familiaridad y apropiación del significado: Los altos índices de familiaridad no implican conocimiento acabado del significado de las palabras, por cuanto el estudio normativo solo apunta al reconocimiento de palabras y no a su empleo en propiedad.

Conocer las peculiaridades del léxico de la bibliografía de los estudiantes universitarios que se incorporan a la comunidad académica entrega referentes para abordar los problemas de comprensión que tanto los docentes como los aprendientes declaran al momento de enfrentar tareas de lectura disciplinar. Las palabras propias de la formación docente presentan altas frecuencias, por cuanto son de uso habitual por la naturaleza misma de la enseñanza. Sin embargo, son otras las palabras – no relacionadas con el ámbito de la formación docente– las que pueden generar dificultades en la lectura. Esto implica que el resguardo que los académicos deben tener tiene que focalizarse en estas y no en aquellas de modo tal de garantizar una tarea de lectura exitosa.

Siguiendo las ideas del triángulo dorado de la comprensión de Perfetti (2012), conocer las características léxicas de los textos y trabajarlas como estrategias de lectura traerá como consecuencia la adquisición de significados nuevos en las palabras de menor frecuencia y familiaridad y una definición más fina en aquellas que aún no están del todo asentadas, lo que derivará en mejores desempeños en la comprensión de los aprendientes.

7. REFERENCIAS

- ALAMEDA, J. y F. CUETOS. (1995) Diccionario de frecuencias de las unidades lingüísticas del castellano. Oviedo: Servicio de publicaciones de la Universidad de Oviedo.
- BELINCHÓN, M., A. RIVIÉRE y J. IGOA. 1992. Psicología del lenguaje. Investigación y teoría. Madrid: Trotta.
- BLEASDALE, F. 1987. Concreteness-dependent associative priming: Separate lexical organization for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 13(4): 582-594.
- CAMPOS, D., P. CONTRERAS, B. RIFFO, M. VÉLIZ y A. REYES. 2014. Complejidad textual, lecturabilidad y rendimiento lector en una prueba de comprensión en escolares adolescentes. *Universitas Psychologica* 13(3): 15-26.
- CARLINO, P. (2003). Alfabetización académica: un cambio necesario, algunas alternativas posibles. *Educere* 6(20): 409-420.
- CONNELL, L. y D. LYNOTT. 2012. Strength of perceptual experience predicts word processing performance better than concreteness or imageability. *Cognition*, 125(3): 452-465.
- CRUTCH, S. y E. WARRINGTON. 2005. Abstract and concrete concepts have structurally different representational frameworks. *Brain* 128(3): 615-627.
- DOMÍNGUEZ, A. y A. ESTÉVEZ (1999) El reconocimiento de las palabras: procesamiento morfológico y semántico 131-161. En M. de Vega y F. Cuetos (Eds.). *Psicolingüística del español*. Pp. 131-161. Madrid: Trotta.
- FORSTER, K. (1976) Accessing the mental lexicon. En R. J. Wales y E. Walker (eds.). *New approaches to language mechanisms*. Amsterdam: North-Holland. Pp. 257-287.
- ILIEV, R. y R. AXELROD. 2016. The Paradox of Abstraction: Precision Versus Concreteness. *Journal of Psycholinguistic Research* 46(3): 1-15.
- KROLL, J y J. MERVES. 1986. Lexical access for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 12(1): 92.
- MARSLEN-WILSON, W. Function and process in spoken word recognition: A tutorial review. En R.J. Wales y E.C.T. Walker (Eds.) *Attention and performance: Control of language processes*. Pp. 125-150. Amsterdam: North Holland.
- MONSALVE, A. y F. CUETOS. 2001. Asimetría hemisférica en el reconocimiento de palabras: efectos de frecuencia e imaginabilidad. *Psicothema* 13(1): 24-28.
- MORTON, J. (1969) Interaction of information in word recognition. *Psychological review*, 76(2): 165-178.
- NEIRA, A. 2015. Lectura en la educación superior: uso de estrategias en la comprensión de textos especializados y no especializados en estudiantes de primer año. Tesis para optar al grado de Doctor en Lingüística. Universidad de Concepción.
- NEIRA, A., REYES, F. y RIFFO, B. 2015. Experiencia académica y estrategias de comprensión lectora en estudiantes universitarios de primer año. *Literatura y Lingüística*, 31: 221-244.
- PAVIO, A. (1991). Dual coding theory: Retrospect and current status. *Canadian journal of psychology* 45(3): 255-287.

- PARODI, G., R. IBÁÑEZ, R. VENEGAS, y C. GONZÁLEZ. 2010. Identificación de géneros académicos y géneros profesionales: Principios teóricos y propuesta metodológica. En G. Parodi (Ed.), *Alfabetización académica y profesional en el siglo XXI: Leer y escribir desde las disciplinas*. Pp. 249-289. Santiago: Ariel
- PÉREZ, M., G. CAMPOY y C. NAVALÓN. 2001. Índice de estudios normativos en idioma español. *REMA Revista electrónica de metodología aplicada* 6(2): 85-105.
- PERFETTI, C. 2010. Decoding, vocabulary and comprehension. En M. McKeown, y L. Kuncan (Eds.), *Bringing reading research to life*. Pp. 291-303. New York: The Guilford Press.
- PERFETTI, C. y S. ADLOF (2012) Reading comprehension: A conceptual framework from word meaning to text meaning. En J. Sabatini, E. Albro Y T. O'Reilly (Eds.) *Measuring up: Advances in how we assess reading ability*. Pp. 3-20. Maryland: Rowman & Littlefield Education.
- SADOWSKY, S. y R. MARTÍNEZ-GAMBOA (2012). LIFCACH 2.0: Word Frequency List of Chilean Spanish (Lista de Frecuencias de Palabras del Castellano de Chile) [en línea]. Disponible en <http://sadowsky.cl/lifcach.html> [Consulta 05/08/2018].
- SCHRÖDER, H. 1991. Linguistic and text-theoretical research on languages for special purposes. A thematic and bibliographical guide. En H. Schröder (Ed.). *Subject-oriented texts*. Pp. 1-49. Berlín: De Gruyter.
- VEGA, M. DE y F. CUETOS. 1999. *Psicolingüística del español*. Madrid: Trotta.
- VEGA, M., y A. FERNÁNDEZ. 2011. Datos normativos de concreción de 730 palabras utilizadas por sujetos de habla castellana. *Psicológica* 32(2): 171-206
- VERHOEVEN, L. y C. PERFETTI. 2003. Introduction to this special issue: The role of morphology in learning to read. *Scientific Studies of Reading* 7(3): 209-217.

ANEXO: LISTA DE PALABRAS INCLUIDAS EN EL ESTUDIO NORMATIVO, ORDENADAS ALFABÉTICAMENTE.

PALABRA	FRECUENCIA	RANGO	SÍLABAS	CONFIG PRIM SIL
abandonar	53	Infrecuente	4	V
acceder	91	Frecuente	3	VC
acceso	135	Frecuente	3	VC
afectar	164	Frecuente	3	V
alcanzar	258	Frecuente	3	VC
alimento	67	Infrecuente	4	V
alumno	159	Frecuente	3	V
ámbito	95	Frecuente	3	VC
análisis	119	Frecuente	4	V
animal	78	Frecuente	4	V
antiguo	93	Frecuente	3	VC
aparecer	219	Frecuente	4	V
aparición	30	Infrecuente	4	V
apertura	76	Frecuente	4	V
aplicar	180	Frecuente	3	V
aportar	88	Frecuente	3	V
aprender	89	Frecuente	3	V
aptitud	17	Infrecuente	3	VC
asignar	28	Infrecuente	3	V
asistir	104	Frecuente	3	V
asociación	224	Frecuente	4	V
aspecto	172	Frecuente	3	VC
asumir	199	Frecuente	3	V
atención	203	Frecuente	3	V
avance	107	Frecuente	3	V
ayudar	152	Frecuente	3	V
clásico	55	Infrecuente	3	CCV
complejo	70	Infrecuente	3	CVC
comprensión	19	Infrecuente	3	CVC
comunidad	260	Frecuente	4	CV
concepto	120	Frecuente	3	CVC
conducta	60	Infrecuente	3	CVC
conferir	17	Infrecuente	3	CVC

conjunto	180	Frecuente	3	CVC
conocido	161	Frecuente	4	CV
conquistar	24	Infrecuente	3	CVC
conseguir	183	Frecuente	3	CVC
continuo	31	Infrecuente	3	CVC
convertir	164	Frecuente	3	CVC
cooperar	14	Infrecuente	4	CV
costumbre	36	Infrecuente	3	CVC
creación	118	Frecuente	3	CCV
creciente	35	Infrecuente	3	CCV
creencia	13	Infrecuente	3	CCV
cultura	237	Frecuente	3	CVC
descripción	32	Infrecuente	3	CVC
desempeño	58	Infrecuente	4	CV
detallado	15	Infrecuente	4	CV
deteriorar	15	Infrecuente	4	CV
diferencia	198	Frecuente	4	CV
dinámico	29	Infrecuente	4	CV
distinguir	32	Infrecuente	3	CVC
diversidad	26	Infrecuente	4	CV
enseñar	43	Infrecuente	3	VC
entender	220	Frecuente	3	VC
entorno	44	Infrecuente	3	VC
época	102	Frecuente	3	V
escena	61	Infrecuente	3	VC
escuchar	131	Frecuente	3	VC
escuela	215	Frecuente	3	VC
estudiar	155	Frecuente	3	VC
etapa	170	Frecuente	3	V
evento	155	Frecuente	3	V
evidenciar	16	Infrecuente	4	V
evolución	44	Infrecuente	4	V
examen	82	Frecuente	3	V
existencia	98	Frecuente	4	V
existente	66	Infrecuente	4	V
expresar	202	Frecuente	3	V

expuesto	45	Infrecuente	3	VC
facilitar	65	Infrecuente	4	CV
fenómeno	68	Infrecuente	4	CV
fundación	91	Frecuente	3	CVC
generación	105	Frecuente	4	CV
habitante	83	Frecuente	4	V
hábito	18	Infrecuente	3	V
identidad	77	Frecuente	4	V
iglesia	130	Frecuente	3	V
impacto	118	Frecuente	3	VC
iniciar	289	Frecuente	3	VC
interés	287	Frecuente	3	VC
jornada	221	Frecuente	3	CVC
lectura	45	Infrecuente	3	CVC
limitado	32	Infrecuente	4	CV
objeto	144	Frecuente	3	VC
observar	118	Frecuente	3	VC
obtener	267	Frecuente	3	VC
oferta	125	Frecuente	3	V
oficio	95	Frecuente	3	V
orientar	24	Infrecuente	3	V
origen	115	Frecuente	3	V
originar	29	Infrecuente	4	V
partida	66	Infrecuente	3	CVC
percibir	52	Infrecuente	3	CVC
poesía	39	Infrecuente	4	CV
político	297	Frecuente	4	CV
posterior	61	Infrecuente	3	CVC
preciso	36	Infrecuente	3	CCV
presencia	180	Frecuente	3	CCV
profesor	193	Frecuente	3	CCV
profundo	80	Frecuente	3	CCV
proponer	157	Frecuente	3	CCV
propuesta	199	Frecuente	3	CCV
proveer	27	Infrecuente	3	CCV
realidad	197	Frecuente	4	CV

recolección	14	Infrecuente	4	CV
recorrido	55	Infrecuente	4	CV
referencia	68	Infrecuente	4	CV
religioso	63	Infrecuente	4	CV
repetido	11	Infrecuente	4	CV
representar	183	Frecuente	4	CV
responder	186	Frecuente	3	CVC
sacerdote	30	Infrecuente	4	CV
sentido	284	Frecuente	3	CVC
sostener	262	Frecuente	3	CVC
técnica	168	Frecuente	3	CVC
teoría	37	Infrecuente	4	CV
totalidad	39	Infrecuente	4	CV
tradicción	58	Infrecuente	3	CCV
utilizar	194	Frecuente	4	V
vincular	72	Frecuente	3	CVC