

*LINGÜÍSTICA APLICADA A LA ENSEÑANZA DEL
ÁRABE COMO LENGUA EXTRANJERA:*

*COMPILACIÓN Y ANÁLISIS DE UNA LISTA DE FRE-
CUENCIA LÉXICA EN LA NOVELA ÁRABE CONTEM-
PORÁNEA PARA EL DESARROLLO DE LECTURAS
EXTENSIVAS GRADUADAS*

Autora:

Laila Hasan Familiar

Directores:

Dra. Nieves Paradela (Universidad Autónoma de Madrid)

Dr. Manuel Feria (Universidad de Granada)

TESIS DOCTORAL

Universidad Autónoma de Madrid

Facultad de Filosofía y Letras

Departamento de Estudios Árabes e Islámicos y Estudios Orientales

MADRID, 2017

*“Laughter, joy, excitement, sorrow and anger can promote learning. Neutrality and nullity cannot”,
Brian Tomlinson.*

AGRADECIMIENTOS

Esta investigación no habría sido posible sin el apoyo incondicional de mis dos directores de Tesis, Nieves Paradela y Manuel Feria. Gracias por creer en mí, por desafiar mi intelecto y por animarme, siempre.

El corpus que conforma la base de mi estudio existe gracias a la ayuda de muchos amigos y colegas que me pusieron en contacto con decenas de novelistas. Durante el proceso, Ibrahim Farghali, Doria ElKerdani y Khaled Haroub jugaron un papel esencial. Otros muchos amigos y mi colega Dil Parkinson también me ayudaron. A ellos y a todos los editores y novelistas que me facilitaron sus obras les doy las gracias de todo corazón. Por razones de confidencialidad no me es posible nombrarlos.

La ayuda de Nizar Habash y Hind Saddiki, en New York University Abu Dhabi, ha sido fundamental para el análisis de mi corpus y la extracción de datos a través del programa informático MADAMIRA.

También quisiera recordar las estimulantes conversaciones mantenidas con Mustafa Mughazy, Marilyn Booth y Nizar Habash en el ámbito de la enseñanza del árabe, la literatura y la morfología, que tanto me inspiraron.

A todos ellos les agradezco su ayuda, pero a ninguno hago responsable de cualquier opinión o error cometido durante el transcurso del trabajo o en la tesis. Cualquier falta o incongruencia, es responsabilidad mía.

No quisiera concluir sin agradecer a mis padres, mis hermanos, mi marido y mis amigos su paciencia y espera. Perdonadme si no os he atendido como debiera durante el tiempo que me llevó realizar la tesis. Lo mejor está aún por llegar.

Finalmente, este trabajo se lo dedico a mi profesora, mentora y amiga Inmaculada García Barrientos, por inspirarme en el amor a la literatura, y al inigualable El-Said Badawi, que en paz descansa, por introducirme al irresistible mundo de la lexicografía árabe.

RESUMEN

La enseñanza de árabe como lengua extranjera (EALE) no dispone de lecturas graduadas basadas en datos de frecuencia léxica provenientes de corpora diseñados para tal propósito. Esta tesis pretende cubrir algunas lagunas en el área del desarrollo de material pedagógico destinado a la EALE, a través de instrumentos y métodos propios de la lingüística de corpus.

La presente investigación comienza reconociendo el papel central de la lectura y repasa la extendida opinión académica de que la grafía árabe es una carga desde el punto de vista de la decodificación visual. Mi tesis se desvía de tal afirmación y argumenta que los enfoques pedagógicos y la calidad del material didáctico son los factores más determinantes a la hora de construir y desarrollar una robusta destreza lectora en aprendientes de árabe como lengua extranjera (ALE).

La tesis pone de relieve, asimismo, el papel de la literatura graduada como vehículo ideal capaz de desarrollar destrezas relativas al reconocimiento visual automático y a la adquisición del léxico, especialmente en el contexto de la lectura extensiva.

La principal contribución de la presente investigación es el diseño y la construcción de un corpus representativo de novelas árabes contemporáneas, a partir del cual se extrajo una lista de frecuencia léxica de 2.000 lemas que, además, fue comparada con el *Frequency Dictionary of Arabic* desarrollado por Buckwalter and Parkinson (2011), considerada obra de referencia en el campo de la frecuencia léxica árabe.

Los resultados del análisis muestran que, desde el punto de vista de la frecuencia léxica, la lista de Familiar es el instrumento pedagógico más adecuado, hasta el momento, para el diseño de lecturas graduadas destinadas a estudiantes de ALE, como mínimo, en los niveles A1 y A2 del Marco Común Europeo de Referencia para las lenguas.

Palabras Clave: Lingüística de Corpus, Enseñanza de Árabe como Lengua Extranjera,
Lecturas Graduadas, Lectura Extensiva, Diseño Curricular

ABSTRACT

Teaching Arabic as a Foreign Language (TAFL) lacks Graded Readers grounded on lexical frequency data drawn from adequately designed corpora. The present PhD dissertation is an attempt to fill some of the gap in the area of teaching material development, by means of using corpus linguistics tools and methods.

The present study starts by recognizing the central role of reading and it reviews the widespread idea in the academia that Arabic script is a burden from a visual decoding perspective. My dissertation deviates from this claim and it argues that pedagogical approaches and quality teaching material are the main factors when building robust reading skills in learners of Arabic as a Foreign Language (AFL).

The dissertation also highlights graded narrative literature as an ideal vehicle that can help develop learners automatic visual decoding skills and promote vocabulary acquisition, especially in the context of Extensive Reading.

The main contribution of the present study is the designing and compiling a representative corpus of contemporary Arabic novels, from which a reliable frequency list of 2.000 lemmas was finally extracted and, later on, compared to the *Frequency Dictionary of Arabic* by Buckwalter and Parkinson (2011).

The results of the analysis show that, from a lexical frequency perspective, Familiar's list is the most adequate tool available to date for the design of Graded Readers that target AFL learners that are, at least, at the A1 and A2 proficiency levels according to the Common European Framework of Reference for Languages.

Keywords: Corpus Linguistics, Teaching Arabic as a Foreign Language, Graded Readers, Extensive Reading, Curricular Design

ÍNDICE

AGRADECIMIENTOS	5
RESUMEN.....	7
ABSTRACT.....	11
ÍNDICE DE TABLAS	19
ÍNDICE DE GRÁFICOS	21
SISTEMA DE TRANSLITERACIÓN.....	23
CAPÍTULO I: LA LECTURA COMO DESTREZA ESENCIAL.....	27
A. LA LECTURA COMO PROCESO COGNITIVO	29
1. Modelos de lectura en L1	30
2. La lectura en L2.....	32
B. LA LECTURA EN ÁRABE	34
1. Leer árabe como L1.....	34
2. Leer árabe como L2.....	37
C. LA COMPLEJIDAD DE LEER EN ÁRABE	40
1. La grafía	41
2. Opacidad ortográfica	42
3. Homografía.....	44
4. Morfología.....	45
5. Diglosia	47
6. Sistema educativo.....	47
7. Conclusión.....	49
D. MODALIDADES DE LECTURA	51
1. Lectura intensiva	51
2. Lectura extensiva.....	52

CAPÍTULO II: LOS TEXTOS LITERARIOS EN LA ENSEÑANZA DE LENGUAS EXTRANJERAS.....	55
A. EL TEXTO LITERARIO Y LOS ENFOQUES METODOLÓGICOS.....	55
B. LOS BENEFICIOS DE LEER TEXTOS LITERARIOS.....	57
1. Beneficios lingüísticos	58
2. Beneficios culturales	59
3. Beneficios motivacionales.....	61
4. Beneficios psicológicos.....	63
5. Beneficios humanistas.....	63
6. Implicaciones para la EALE.....	64
C. CÓMO SE PERCIBE LA LITERATURA EN EL CONTEXTO DE LA EALE.....	65
1. La situación en España	66
2. La situación en EEUU	75
3. Conclusión.....	83
CAPÍTULO III: VOCABULARIO Y LECTURA EXTENSIVA.....	85
A. LÉXICO MENTAL Y LECTURA	86
B. CONOCER EL VOCABULARIO	90
C. AUTOMATICIDAD Y VELOCIDAD LECTORA.....	91
D. ADQUIRIR VOCABULARIO DE FORMA INCIDENTAL	95
E. CONCLUSIÓN	98
CAPÍTULO IV: LAS LECTURAS GRADUADAS COMO VÍA PEDAGÓGICA	101
A. ¿QUÉ SON LAS LECTURAS GRADUADAS?	101
B. TEXTOS GRADUADOS VERSUS TEXTOS AUTÉNTICOS.....	103
C. EVIDENCIAS EMPÍRICAS A FAVOR DE LAS LECTURAS GRADUADAS.....	105
D. LA IDONEIDAD DE LA LITERATURA COMO TEXTO GRADUADO	106
E. UNA PROPUESTA PEDAGÓGICA PARA LA EALE	107
CAPÍTULO V: LINGÜÍSTICA DE CORPUS PARA EL DISEÑO DE MATERIALES DIDÁCTICOS.....	111
A. LINGÜÍSTICA DE CORPUS Y LECTURAS GRADUADAS.....	112
1. Listas de frecuencia léxica	112
2. La elaboración de lecturas graduadas.....	115
B. LINGÜÍSTICA DE CORPUS Y DIDÁCTICA DE ALE.....	117
1. Listas de frecuencia léxica.....	117
2. La elaboración de lecturas graduadas en árabe.....	119
CAPÍTULO VI: LA PRESENTE INVESTIGACIÓN	122

A. OBJETIVOS	122
B. PREGUNTAS DE INVESTIGACIÓN	122
C. JUSTIFICACIÓN.....	123
D. VIABILIDAD	124
E. DISEÑO DE LA INVESTIGACIÓN	124
CAPÍTULO VII: CRITERIOS DE REPRESENTATIVIDAD DEL CORPUS	
LITERARIO.....	127
A. REPRESENTAR LA NOVELA ÁRABE CONTEMPORÁNEA.....	127
1. Género	127
2. Diacronía.....	128
3. Registro	128
B. CORPORA DISPONIBLES.....	128
C. ADECUACIÓN DE LA POBLACIÓN	130
D. DESCRIPCIÓN DE LA POBLACIÓN	133
E. POBLACIÓN y VARIABLES.....	136
1. Autoría.....	136
2. Edad.....	136
3. Variación diatópica.....	137
4. Variación temática.....	137
5. Sexo.....	138
E. CONCLUSIÓN	139
CAPÍTULO VIII: COMPOSICIÓN Y CONSTRUCCIÓN DEL CORPUS	
A. DISPONIBILIDAD DE OBRAS	141
B. TAMAÑO DE POBLACIÓN	142
C. TAMAÑO DE LA MUESTRA.....	143
D. MÉTODO DE MUESTREO	144
E. SELECCIÓN DE LA MUESTRA.....	144
CAPÍTULO IX: ANÁLISIS.....	
A. HERRAMIENTAS DE ANÁLISIS	149
1. Segmentación, tokenización y lematización.....	149
2. Elaboración de una lista de frecuencia	150
B. FRECUENCIA Y DISTRIBUCIÓN LÉXICA	151
C. FRECUENCIAS ABSOLUTA Y RELATIVA.....	155
D. ELECCIÓN Y AFINACIÓN DE UNA FRANJA LÉXICA	156

E. AJUSTES INTRODUCIDOS EN LA LISTA DE BUCKWALTER Y PARKINSON Y AJUSTES FINALES EN MI LISTA.....	157
F. NORMALIZACIÓN DE LAS LISTAS DE FRECUENCIA.....	160
CAPÍTULO X: RESULTADOS	163
A. SEMEJANZAS Y DISCREPANCIAS ENTRE MI LISTA DE FRECUENCIA Y LA DE BUCKWALTER Y PARKINSON	163
1. Los 10 lemas más frecuentes.....	165
2. Los 30 sustantivos más frecuentes.....	166
3. Los 30 verbos más frecuentes.....	169
4. Los 30 adjetivos más frecuentes.....	171
5. Lemas exclusivos del corpus Familiar.....	173
6. Lemas propios del lenguaje periodístico en Buckwalter y Parkinson.....	175
7. Discrepancias léxicas recurrentes	176
9. Conclusión.....	178
B. LOS LEMAS MÁS FRECUENTES	179
CAPÍTULO XI: CONCLUSIONES, APLICACIONES DIDÁCTICAS Y FUTURAS LÍNEAS DE INVESTIGACIÓN	183
A. CONCLUSIONES	183
B. APLICACIONES DOCENTES	186
1. Elaborar un diccionario de frecuencia	186
2. Desarrollar y adaptar lecturas graduadas.....	186
C. FUTURAS LÍNEAS DE INVESTIGACIÓN EN LA EALE.....	187
1. Elaborar listas de frecuencia que cubran los niveles A1-C1.....	187
2. Elaborar listas de frecuencia especializadas	188
3. Aplicar fórmulas estadísticas para cuantificar el comportamiento léxico	188
4. Estudiar características gramaticales	188
5. Estudiar aspectos léxicos propios del árabe.....	189
6. Unidades multiléxicas	190
D. FUTURAS LÍNEAS DE INVESTIGACIÓN EN OTRAS ÁREAS	190
1. Sociolingüística y discurso de género	191
2. Variación diatópica y diglosia	191
3. Estilometría	192
4. Variación entre géneros.....	193
5. Listas de frecuencia en otras lenguas y lectura extensiva.....	193
CONCLUSIONS.....	195

APÉNDICE A	199
APÉNDICE B	201
APÉNDICE C	203
APÉNDICE D	217
APÉNDICE E.....	219
APÉNDICE F.....	223
APÉNDICE G	225
APÉNDICE H	289
APÉNDICE I.....	299
BIBLIOGRAFÍA.....	369

ÍNDICE DE TABLAS

Tabla	Título	Página
1	Promedio de lemas adquiridos por estudiantes de inglés L2 en Grecia	84
2	Composición del corpus de Buckwalter y Parkinson (2011)	115
3	Descripción de la población	131
4	Coincidencia del léxico literario y el léxico general en lengua árabe (2.000 lemas desglosados en franjas de 500)	159
5	Los 10 lemas más frecuentes (Buckwalter y Parkinson vs. Familiar)	161
6	Los 30 sustantivos más frecuentes (Buckwalter y Parkinson vs. Familiar) con resaltado de los lemas coincidentes	163
7	Los 30 verbos más frecuentes (Buckwalter y Parkinson vs. Familiar), con resaltado de los lemas coincidentes	165
8	Los 30 adjetivos más frecuentes (Buckwalter y Parkinson vs. Familiar) con resaltado de los lemas coincidentes	167
9	Partes del cuerpo en Familiar vs. Buckwalter y Parkinson	169
10	Lemas de alta frecuencia en Buckwalter y Parkinson que no figuran en Familiar	170
11	Lemas de uso cotidiano en Buckwalter y Parkinson vs. Familiar	172
12	Uso del lema <i>zāra</i> en el corpus Familiar	172
13	Uso del lema <i>gādara</i> en el corpus Familiar	173
14	Uso del lema <i>ŷiwār</i> en el corpus Familiar	173
15	Frecuencia de lemas de contenido semántico cotidiano en Familiar	175
16	Ranquin de lemas típicos del lenguaje figurativo en Familiar vs. Buckwalter y Parkinson	175

ÍNDICE DE GRÁFICOS

Gráfico	Título	Página
1	Procedencia nacional de los autores	131
2	Distribución Léxica Zipf en mi corpus (10.241 lemas)	148
3	Perfil ranquin/frecuencia de lemas en mi corpus (10.241 lemas)	150
4	Espectro de frecuencia de lemas en mi corpus (10.241 lemas)	151

SISTEMA DE TRANSLITERACIÓN

Sistema de la Escuela de Arabistas Españoles

ء	ʾ	
ب	b	
ت	t	
ث	ṯ	Ṭ
ج	ǰ	ǰ
ح	ḥ	Ḥ
خ	j	
د	d	
ذ	ḏ	Ḑ
ر	r	
ز	z	
س	s	
ش	š	ŠŠ
ص	ṣ	Ṣ
ض	ḏ	Ḑ
ط	ṭ	Ṭ
ظ	ẓ	Ẓ
ع	ʿ	
غ	g	
ف	f	
ق	q	
ك	k	
ل	l	
م	m	
ن	n	

ه	h
و	w
ي	y

Otros símbolos:

- Vocales cortas = *a, i, u*
- Vocales largas = *ā, ī, ū*
- diptongos = *ay, aw*
- ʾ = *a* (en estado absoluto) y *at* (en estado constructo)
- ʿ = *à*
- ء al inicio no se transcribe

El artículo = *al-* (aun ante solares; ej.: *Awrāq al-nar̄yis*), *l-* (precedido de palabra terminada en vocal larga; ej.: *fī l-hunā*), *-l-* (precedido de partícula terminada en vocal corta; ej.: *wa-l-kitāb*)

- Sufijo de *nisba* (ّ) = *ī* (para masculino) *-iyya* (para femenino)

CAPÍTULO I: LA LECTURA COMO DESTREZA ESENCIAL

El objetivo principal de la presente investigación es la extracción de una lista de vocabulario frecuente en la literatura escrita a partir de un corpus de novelas árabes contemporáneas, con el fin último de facilitar el desarrollo de materiales de lectura extensiva apropiados para el aprendizaje del árabe como lengua extranjera (ALE) a través del uso del vocabulario que con más frecuencia aparece en la novela árabe contemporánea. Por tanto, comenzaré definiendo el concepto de *lectura extensiva* y repasando algunos factores implicados en la destreza de la lectura, considerada uno de los procesos cognitivos más importantes y complejos que ocurren en la mente humana.

El término *lectura extensiva* ha sido definido por Day y Bamford (1998: xiii) como la lectura de grandes cantidades de libros y de otros materiales de nivel lingüístico accesible, es decir, escritos al nivel de los aprendientes y en consonancia con sus competencias lingüísticas. El conocido investigador neozelandés Paul Nation añade que los textos empleados en la práctica de la lectura extensiva deben estar compuestos por vocabulario conocido por el aprendiente al menos en el 95% de su totalidad, para que así pueda convertirse en una lectura placentera al tiempo que beneficiosa (Laufer 1992, Nation 2013:219). Para los objetivos de la presente investigación, la lectura extensiva significará “la práctica de leer fuera del aula narrativa escrita y adaptada al nivel lingüístico de los aprendientes, con el objetivo de promover el placer de la lectura y desarrollar diversas destrezas lingüísticas y culturales”.

El rol central que desempeña la lectura extensiva, al menos en la cultura occidental, tanto en el aprendizaje de la lengua materna como en la adquisición de lenguas extranjeras, ha sido destacado por muchos investigadores. Se ha dicho, por ejemplo, que la lectura es la única forma posible de que el hablante pueda desarrollar la ortografía, un vocabulario amplio, un buen estilo de escritura y una competencia gramatical avanzada (Krashen 2004:37). En el campo de la enseñanza de lenguas extranjeras se considera que la lectura es, en el proceso de aprendizaje, la destreza lingüística más

estable y duradera (Rivers 1981; Badawi 2002:156-157) y la vía fundamental para adquirir independencia (Day y Bamford 1998: xiii; Grabe y Stoller 2001:187).

En árabe, la destreza de la lectura ocupa un lugar especial (Badawi 2002) debido a la naturaleza diglósica de esa lengua (Ferguson 1959) y al hecho de que el hablante nativo de árabe utiliza en la lectura un registro lingüístico diferente al utilizado en la expresión oral. Según el académico egipcio Elsaid Badawi (1973), en el contexto del aprendizaje del árabe formal o estándar (MSA)¹, la lectura constituye probablemente la destreza activa más importante y uno de los objetivos primordiales que cualquier aprendiente de ALE debe alcanzar si quiere identificarse con el hablante nativo educado y emular sus competencias lingüísticas² (Badawi 2002:159).

Según la experta estadounidense Karin Ryding, la lectura continúa ocupando un lugar central en la mayoría de los programas académicos y no académicos dedicados a la enseñanza del árabe como lengua extranjera (EALE) (Ryding 2013:147), en contraposición con las destrezas habladas, tradicionalmente clasificadas como comunicativas. Sin embargo, como explicaré más adelante, la lectura que se suele practicar en el aula del ALE está enfocada hacia el análisis de textos mediáticos breves de tipo informativo, mientras que la lectura extensiva independiente, por lo general, es una práctica que no se lleva al aula hasta que los aprendientes alcanzan el nivel C1 (Badawi 2002).

La presente investigación pretende contribuir al diseño de materiales de lectura extensiva fuera del aula del ALE, en los niveles A1-B2, a partir de la elaboración de una lista de vocabulario frecuente que facilite el desarrollo de materiales pedagógicos nivelados ajustados a las necesidades y capacidades lingüísticas del alumno. Seguidamente expondré algunos aspectos cognitivos relacionados con la lectura en general y con la lectura en árabe en particular, y destacaré el papel que desempeña el vocabulario en el desarrollo de la competencia lectora.

¹ A partir de aquí, utilizaremos la aceptada abreviatura MSA para referirnos a lo que en inglés viene denominándose Modern Standard Arabic, es decir, el registro formal o estándar moderno de la lengua árabe, utilizado por norma general en los textos escritos.

² Con el término “educado” (*educated* en inglés) nos referimos al hablante que ha sido escolarizado hasta la etapa universitaria y que posee cierto nivel cultural derivado de la lectura.

A. LA LECTURA COMO PROCESO COGNITIVO

Según el Diccionario de la Real Academia Española de la Lengua, leer es “pasar la vista por lo escrito o impreso comprendiendo la significación de los caracteres empleados”³. Todas las definiciones de lectura existentes, muchas y variadas, coinciden en que las destrezas psicolingüísticas y cognitivas involucradas en el proceso de la lectura son heterogéneas y complejas (Grabe 2009:13-14).

La *comprensión lectora* es, desde los años 70 del siglo XX, la destreza que mayor atención ha recibido en la literatura referente al aprendizaje de primeras lenguas (L1) (Urquhart y Weir 1998) y de lenguas extranjeras (L2) (Grabe 2009), posiblemente por ser una de las competencias más fáciles de observar y de manipular. Como proceso mental, la lectura se entiende como “*the reader’s active participation in the creation of meaning, as a manipulation of strategies, as a receptive rather than a passive skill*” (Barnett 1989:2). Definiciones más recientes se refieren a la lectura en los siguientes términos: “*To construct text meaning based on visually encoded information. Essentially, it entails converting print into language and then to the message intended by the author*” (Koda 2007:1). La lectura implica, por tanto, la interdependencia de procesos cognitivos y psicolingüísticos varios que el lector, no necesariamente con consciencia de ello, debe activar de forma eficaz.

Algunas definiciones más antiguas se centraban en las destrezas operacionales involucradas en la lectura. Bajo este prisma, la lectura era concebida como un proceso compuesto esencialmente de dos partes: 1) la identificación del símbolo y 2) la obtención de significado a partir del símbolo identificado o, lo que es lo mismo, la decodificación del símbolo con el fin de comprender su significado (Russel y Fea 1963, citados en Calfee y Venezky 1969:92). Aunque esa definición pueda parecer anticuada, me parece importante recuperarla, pues, como veremos más adelante, la identificación y la decodificación son procesos esenciales para la comprensión del texto impreso; cabe añadir, incluso, que se trata de pasos fundamentales para automatizar el reconocimiento de vocabulario en su forma escrita, lo cual es requisito primordial para lograr la fluidez en la lectura.

³ Versión electrónica del diccionario, consultada el 26 de septiembre de 2015.

1. Modelos de lectura en L1

Resumir los distintos modelos psicolingüísticos y cognitivos desarrollados a lo largo de los últimos cincuenta años para explicar la naturaleza de la lectura es tarea imposible, por la complejidad misma de los procesos involucrados y las múltiples teorías propuestas (Alderson, 2000:1)⁴. Sin embargo, creemos que antes de situar la lectura en lenguas extranjeras en un marco teórico particular, es necesario hacer referencia a algunos modelos de lectura muy generales que se han barajado en el campo de la adquisición de L1.

En los *modelos de procesamiento descendente* (desde arriba hacia abajo) -fundamentalmente representados por Goodman (1967:126-135, 1968: 15-26, 1969:9-30)- el proceso de lectura comienza en la mente del lector. Goodman es conocido por su definición de la lectura como un juego de predicción psicolingüístico en el que pensamiento y lenguaje interactúan de forma constante (Goodman 1978: vii). Según esta visión, la comprensión de un texto escrito ocurre a partir de la interacción con la letra impresa; dicha interacción implica la creación de continuas hipótesis que se testan y contrastan durante el proceso de la lectura, con el objetivo de aceptar o rechazar las expectativas mentales sobre lo que se encontrará a continuación en el texto. Estas hipótesis predictivas son posibles porque el lector utiliza su conocimiento del mundo y un conocimiento lingüístico previo (en inglés, *schema*), que ayuda al lector a reconstruir el mensaje (Goodman 1968, Smith 1971). En estos modelos, lo primordial es el papel que desempeña el lector; la letra impresa, por su parte, queda relegada a un segundo plano⁵.

Por el contrario, los *modelos de procesamiento ascendente* (de abajo a arriba) de la lectura consideran que el conocimiento del mundo y el conocimiento lingüístico no tienen tanta influencia a la hora de extraer información a partir del texto escrito. Según Gough (1972) y LaBerge y Samuels (1974), el punto de partida del proceso de lectura es el texto impreso, donde las mismas letras, palabras y oraciones conducen a la comprensión. Uno de los aspectos más importantes de este modelo es el papel central

⁴ Una de las obras recientes que mejor describe los principales modelos de lectura es Grabe (2009:83-106).

⁵ Grabe (2009:102-103) afirma que la evidencia empírica de los últimos 20 años demuestra que estos modelos son absolutamente erróneos, pues el movimiento ocular que se produce durante la lectura es automático y el lector no está en posición de decidir hacia dónde dirige la mirada para adivinar el significado de palabras desconocidas. Únicamente los lectores menos eficientes pondrían en marcha este tipo de estrategias.

otorgado al reconocimiento visual o desciframiento de los signos gráficos y de las palabras, y el acceso mental relacionado con su correspondiente representación fonética, como mecanismo esencial en la comprensión (Barnett 1989, Rayner y Pollatsek 1989, Hudson 2007:34-36). Para estos modelos, los diferentes procesos que tienen lugar en la mente del lector se producen de forma secuencial, de modo que cada uno de los componentes involucrados en la lectura conduce al siguiente en el orden siguiente: grafía-realización fonética-sintaxis-semántica (Alderson 2000:16-17, Hudson 2007:36). En cualquier caso, ni la deducción del significado de las palabras impresas a partir de conocimientos previos ni a partir de signos gráficos individuales cuya representación fonética se realiza mentalmente son mecanismos suficientes para entender el complejo proceso cognitivo de la lectura, por lo que los dos modelos tradicionales⁶ anteriores fueron sustituidos por *modelos interactivos o mixtos*, considerados más integrales porque combinan mecanismos derivados de los modelos descendentes y, también, de los ascendentes (Hudson 2007:33-34,39).

Según uno de los modelos interactivos más conocidos, el de Just y Carpenter (1987), el acceso al significado de palabras individuales es la base fundamental para la comprensión de textos escritos. Las palabras deberán ser, en primer lugar, codificadas y después relacionadas con conceptos semánticos que se activan. A continuación, el lector relaciona cada palabra individual con las que conforman su contexto en un intento de solventar cualquier inconsistencia en la oración y el texto y, así, obtener un significado coherente (Hudson 2007:44-46).

Otros modelos interactivos menos secuenciales conciben la lectura como un conjunto de operaciones cognitivas que ocurren de forma paralela en la mente del lector⁷. Tanto la decodificación del texto impreso, como el conocimiento que tiene el lector del mundo y, en particular, del tema sobre el que está leyendo, desempeñan papeles esenciales para alcanzar una comprensión satisfactoria. Los lectores más hábiles, pues, recurren a todo un abanico de recursos para extraer un significado coherente a partir de la letra impresa. El principio subyacente a estos modelos es, por tanto, la coocurrencia, y no la secuencialidad, de una multiplicidad de procesos cognitivos en la mente del lector (Alderson 2000:18).

⁶ Grabe (2009) los denomina *modelos metafóricos*, por su carácter simplista.

⁷ Para un recuento exhaustivo de los diferentes modelos interactivos ver Grabe (1988) y Hudson (2007:39-57).

Aunque los *modelos interactivos* han dado paso a una multiplicidad de modelos mucho más concretos y coherentes (Grabe 2009:89-106) y, aún no están claros muchos de los procesos cognitivos involucrados en la lectura, recientemente se vuelve a priorizar el papel que desempeña el reconocimiento automático del *input* visual, es decir, el léxico, lo cual incluye información ortográfica, fonológica, morfológica y semántica (Alderson 2000:18-20, Hudson 2007:90, Grabe 2009:90).

2. *La lectura en L2*

Los investigadores subrayan la existencia de similitudes entre la lectura en L1 y en L2, puesto que en ambos casos la lectura, en definitiva, es un proceso de reconstrucción del significado en el que varias destrezas han de coordinarse (*problem solving*) (Grabe 2009, Koda 2005, 2007, Hudson 2007). Con todo, es importante recordar que, aunque la evidencia científica referente a la lectura en L1 es vital para la investigación en L2, los procesos cognitivos involucrados en ambas tareas no son del todo equiparables ni transferibles (Hudson 2007:31-32, 59-74, 77, Koda 2007: 1-3, Grabe 2009: xiii-xiv, Grabe y Stoller 2013:34-35).

Otro factor que los investigadores destacan es el papel que juega la competencia lingüística que posee el aprendiente de L2 en su capacidad de comprender textos escritos (Alderson 1984, Hudson 2007). Aunque los resultados no son concluyentes, parece ser que el nivel de competencia sería un factor determinante en la capacidad lectora del aprendiente. Por ejemplo, se cree que los estudiantes de nivel inicial centran toda su atención lectora en la decodificación del vocabulario, puesto que no habrían automatizado aún el reconocimiento visual de una gran cantidad de léxico. Las destrezas lingüísticas acumuladas no le permitirían, por tanto, procesar correcta y eficazmente las estructuras sintácticas y morfológicas presentes en el texto y que dan paso a una comprensión adecuada. Son dichas destrezas, adquiridas en diferentes fases y etapas, las que determinan el tipo de tareas que realmente podría realizar el aprendiente al enfrentarse a un texto escrito (Hudson 2007:73-90).

La identificación de diferentes destrezas lingüísticas asociadas a la lectura es la razón por la cual, en el ámbito de la enseñanza de lenguas extranjeras, a menudo se habla del *enfoque de destrezas múltiples* (*Component Skills Approach*) (Hudson 2007:77-103). Según este enfoque, nacido a partir de estudios en L1, la lectura es el

producto de un complejo sistema de procesamiento de la información, en el que diversas destrezas y subdestrezas cognitivas, diferenciables entre sí, ocurren en la mente del lector (Carr y Levy 1990 citado en Koda 2005 or Carr et al. 1990 citado en Hudson 2007). Además, para algunos investigadores, cada una de esas destrezas y sub-destrezas podría aislarse y medirse de forma independiente (Sabatini et al. 2014:37).

Este enfoque ha sido adoptado por Keiko Koda (2005), quien destaca los tres componentes fundamentales de este proceso: a) la decodificación de información lingüística extraída a partir del texto impreso, b) la composición y reconstrucción de información textual a partir de la reconstrucción de frases, oraciones y párrafos, y c) la incorporación de *schema* o conocimiento previo para otorgar significado al texto impreso (Koda 2007:4-10).

De estos tres componentes, el que más me interesa en el contexto de la presente investigación es la decodificación de los elementos lingüísticos extraídos a partir del texto impreso, a saber: la ortografía, el sistema fonológico, el vocabulario y el sistema morfológico de la L2 (Koda 2007:4-6), aunque me centraré en el vocabulario. A continuación, destacaré la importancia de este tipo de información lingüística en el proceso de capacitación lectora del estudiante.

Según Koda (2007), una lectura fluida y sin esfuerzo requiere que el lector sea capaz de reconocer una gran cantidad de vocabulario impreso de forma instantánea y automática para, de ese modo, poder acceder a su significado. El aprendiente de L2 deberá haber interiorizado el nuevo sistema de escritura y el sistema fonológico asociado, y deberá conocer la relación que existe entre símbolos y sonidos. Otro factor que contribuye a una lectura eficaz es el conocimiento de suficiente vocabulario, algo que se adquiere, en parte, a través del proceso mismo de la lectura y del consiguiente aprendizaje incidental de vocabulario adicional (también en Grabe 2009:266). Por último, Koda señala que el conocimiento del sistema morfológico y la capacidad del lector de analizar y descomponer las palabras en sus correspondientes componentes morfológicos lo ayudan a comprender el significado del léxico.

En el contexto de la EALE, el sistema de escritura árabe provoca cierta extrañeza en aquellos aprendientes cuya L1 utiliza el alfabeto latino, particularmente, en los niveles iniciales. El proceso de lectura, en tal caso, suele ser lento y laborioso. Sin embargo, al igual que Koda (2007), creo que si el aprendiente logra reconocer una gran cantidad de léxico de forma automática el desarrollo de la destreza lectora será más fluido y la comprensión del texto impreso, mayor.

En el siguiente apartado repasaré la literatura sobre la lectura en árabe como L1 y L2 y analizaré algunos aspectos directamente relacionados con la escritura árabe y sus implicaciones para la lectura.

B. LA LECTURA EN ÁRABE

Por “árabe” entiendo aquí MSA escrito utilizando el alfabeto árabe, pues parto de la base de que la mayor parte de los textos escritos están en esa modalidad y son, además, los textos que generalmente manejan los aprendientes de ALE.

La literatura sobre la lectura en árabe es, hasta el día de hoy, muy escasa. Los investigadores en la materia han seguido principalmente los modelos interactivos, aunque en ocasiones sus definiciones de lectura resultan demasiado vagas o incluso ambiguas. A continuación, hago un repaso de los estudios más relevantes sobre la lectura en árabe como L1 y como L2.

1. Leer árabe como L1

Uno de los académicos más activos en la materia es el palestino-israelí Abu-Rabia, quien realiza sus experimentos lingüísticos utilizando como participantes a estudiantes de árabe escolarizados en escuelas israelíes. En uno de sus primeros estudios, el investigador deja claro que se adhiere a los modelos de lectura interactivos y afirma la importancia de las estrategias descendentes y ascendentes y, en especial, el papel que desempeñan el contexto y el reconocimiento del vocabulario. La premisa de la que parte Abu-Rabia es que el reconocimiento del léxico en árabe pasa por la presencia de los signos diacríticos que marcan las vocales breves (Abu-Rabia 1997:480).

En el estudio citado (Abu-Rabia 1997) los participantes leen en voz alta palabras independientes, oraciones y párrafos. Un año después, Abu-Rabia replica (1998) la investigación y vuelve a concluir que la presencia de los signos gráficos correspondientes a las vocales breves facilita el reconocimiento del vocabulario a lectores hábiles y no tan hábiles. Sin embargo, el investigador no ofrece, en ningún momento, una definición concreta de qué entiende por “reconocimiento de vocabulario” y tampoco cuestiona si la comprensión del vocabulario desempeña algún papel determinante en la facilitación de la lectura. Creo que esto es problemático, puesto que la lectura en voz alta y la lectura para la comprensión implican procesos psicológicos y cognitivos distintos,

y se realizan con fines diferentes⁸. Es más, aunque Abu-Rabia reconoce el vocabulario como elemento central en la comprensión lectora, el investigador llega a posicionarse clara y explícitamente en contra de los modelos que dan primacía al reconocimiento léxico⁹, pues argumenta que sin la presencia de vocales breves, la comprensión no es posible. Para el investigador “Arabic is perhaps the only language in the world in which readers must first understand the sentence in order to recognize the word” (1997:480)¹⁰, y sin la presencia de marcadores de vocales breves la comprensión lectora se vería seriamente mermada. Esta es la misma conclusión a la que llega Abu-Rabia (1999, 2001) tras realizar otros estudios similares en los que los participantes leen en silencio diferentes textos (vocalizados y sin vocalizar) y, después, son testados acerca del grado de comprensión lectora. Aunque la conclusión principal de dicho estudio (Abu-Rabia 1999, 2001) es que los marcadores de vocales breves facilitan la comprensión lectora, los resultados empíricos no indican, a mi juicio, una diferencia tan significativa en el grado de comprensión entre aquellos estudiantes que leyeron textos vocalizados y aquellos que leyeron textos sin vocalizar.

En estudios posteriores, el investigador vuelve a hacer referencia a los modelos de lectura interactivos al decir que “*reading comprehension with unvowelized Arabic texts appears to be an interactive process composed of root identification¹¹ and a top-down process for bridging semantic gaps*” (Abu-Rabia 2001:53). Aunque Abu-Rabia no lo explicita, parece que el modelo de lectura con el que se identifica o, al menos al que más se acerca, es el Modelo Compensatorio Interactivo¹²(Abu-Rabia 2002:305), pues tanto el papel facilitador de los marcadores de vocales breves, como la identificación de la raíz de las palabras en árabe constituyen recursos esenciales, en el proceso de reconocimiento del léxico escrito. Sin embargo, la predilección de Abu-Rabia por

⁸ Esta distinción no aparece en sus escritos hasta 2001:53-54 y 2002:308.

⁹ Se trata de modelos basados en teorías conectivistas según las cuales el reconocimiento automático del léxico es el elemento cognitivo central en la lectura (Grabe 2009:101-102).

¹⁰ Esta idea se vuelve a repetir en Abu-Rabia 1998:116, aunque en la última versión de esta afirmación el investigador señala: “must first understand the gist of the text in order to read correctly the words of the text” y admite que esta es una tarea difícil para los lectores menos hábiles, pero más manejable en el caso de lectores más hábiles (Abu-Rabia 2002:305).

¹¹ La importancia de extraer mentalmente la raíz del léxico árabe no es resultado de sus propias investigaciones, sino que son tomadas de otros autores, aunque las incorpora claramente en su teoría de la lectura en árabe (Abu-Rabia 2001:53, 2002). Más tarde, el investigador las rechaza (Abu-Rabia 2004).

¹² Según este modelo, propuesto por Stanovich, cuando alguno de los procesos automáticos involucrados en la lectura falla (como el reconocimiento visual de las palabras), el lector recurre a otro tipo de recursos para construir el significado del mensaje escrito (como puede ser el contexto).

los modelos de lectura descendentes aplicados al árabe queda patente en varias ocasiones, puesto que –argumenta- el contexto contribuye a adivinar el significado del léxico y por tanto facilita la comprensión de textos no vocalizados (Abu-Rabia 2001:53,57). Lo que queda claro es que Abu-Rabia no está de acuerdo con los modelos de lectura que dan primacía al reconocimiento automático del léxico, un aspecto que el investigador relaciona con los modelos de lectura ascendentes (Abu-Rabia 2002:305).

El investigador estadounidense de origen egipcio Mustafa Mughazy (2005), por su parte, se adhiere explícitamente a los modelos de lectura interactivos y señala que tanto el procesamiento lingüístico como el conocimiento previo (*schema*) son necesarios para reconstruir el significado del texto árabe en la mente del lector L1. Sin embargo, a diferencia de Abu-Rabia, Mughazy demuestra que los hablantes nativos de árabe no se apoyan en los signos diacríticos de vocales breves para comprender el texto, ni siquiera cuando aparecen explícitos, inclusive cuando se trata de lectura en voz alta¹³. Los lectores hábiles de árabe en L1 se apoyan en estrategias derivadas del *conocimiento intuitivo* que poseen de la morfología, la sintaxis y la frecuencia léxica en árabe. En otras palabras, Mughazy (2005:66) demuestra que los signos diacríticos desempeñan un papel marginal en la lectura eficaz de textos árabes.

Por otra parte, Mughazy aporta datos de investigaciones previas que muestran cómo, para textos en inglés, los hablantes nativos de árabe utilizan estrategias descendentes en mayor medida que los hablantes nativos de inglés. De igual modo, cuando leen en árabe, los hablantes nativos de árabe ponen en marcha procesos de lectura descendentes al tiempo que se inician procesos ascendentes que producen diversas hipótesis interpretativas que los procesos descendentes acaban por aceptar o rechazar para, así, llegar a una correcta lectura del texto (Mughazy 2005:68). Por último, en consonancia con las conclusiones de Mughazy (2015), Hermena et al. (2015) han demostrado con un experimento de seguimiento ocular que los lectores nativos de árabe no procesan eficazmente los signos diacríticos. Cuando el texto aparece completamente desambiguado, esos signos se perciben como elementos superfluos y, en su mayor parte, innecesarios para una comprensión correcta, incluso cuando se trata de desambiguar el significado de homógrafos heterofónicos, como es el caso de la voz pasiva. Por el contrario, los marcadores diacríticos son útiles solo cuando aparecen sobre una palabra que

¹³ Los participantes de Mughazy leen, como parte del experimento, algunas frases donde el léxico aparece intencionalmente vocalizado de forma errónea, y aun así las leen correctamente, incluidos algunos homógrafos.

realmente puede crear confusión en el contexto (por ejemplo, la voz pasiva vs. la voz activa).

2. Leer árabe como L2

Salim Khaldieh uno de los académicos que más ha investigado los procesos de lectura en la EALE, afirma en uno de sus primeros artículos sobre el reconocimiento del vocabulario en el texto escrito su predilección por los modelos de lectura ascendentes y la Teoría de Codificación Dual¹⁴, aunque admite que otras destrezas contempladas en los modelos de lectura descendentes también son importantes (Khaldieh 1996). En consecuencia, aconseja a los profesores de ALE que se centren en la fonología y en la ortografía, especialmente en los niveles iniciales, para que el aprendiente pueda llegar a reconocer visualmente el vocabulario de forma más eficaz (Khaldieh 1996:146).

En un artículo posterior, la posición de Khaldieh (1999) se desplaza hacia una defensa de los modelos de lectura interactivos, y prioriza las destrezas propias de los modelos de lectura descendentes. El autor define la lectura como un “*developmental skill which involves the integration of both textual and extratextual information to construct a satisfying meaning in the mind of the L2 learner*” (Khaldieh 1999:117), definición que reconoce tanto la importancia del texto impreso y las destrezas básicas de decodificación del sistema de escritura como el papel que desempeña el *schema*, o el conocimiento previo que el lector tiene del mundo, además de diferentes procesos involucrados en la negociación del significado (Khaldieh 1999:118).

Por otro lado, Khaldieh (1999:143-144) defiende con vehemencia el uso de textos auténticos¹⁵ para fomentar en los lectores de ALE la activación de mecanismos de negociación del significado propios de los modelos de lectura descendentes, a pesar de que demuestra que los lectores menos eficientes, en comparación con los más avanzados y los nativos, decodifican el vocabulario de forma más lenta (Khaldieh 1999:146). Para concluir, Khaldieh (2011) aborda en su último artículo el papel del vocabulario y la flexión sustantiva (*i' rāb*) en la comprensión lectora y subraya de nuevo la naturaleza

¹⁴ Según la *Dual-Coding Theory* (Grabe 2009:100-101) existen dos vías o rutas fundamentales que conducen al reconocimiento del vocabulario escrito, a saber, una ruta visual y otra fonológica. Esto requiere que el lector tenga un conocimiento completo de los grafemas y su correspondiente realización fonológica, para así poder acceder al significado del léxico de forma correcta.

¹⁵ Con el término “textos auténticos” me refiero a textos producidos para una audiencia árabe, y no textos creados explícitamente para el aprendiente de ALE. Tampoco me refiero a textos modificados o simplificados con el objetivo de ser utilizados en el contexto educativo, ya sea de aprendientes nativos o de aprendientes de ALE.

interactiva y dinámica de la lectura, aunque añade que el conocimiento del vocabulario es el elemento más importante para la comprensión lectora, en contraste con el conocimiento de la gramática o el *i'rāb* (Khaldieh 2011:416-417).

Por su parte, la estadounidense Kristen Brustad (2006) –coautora del famoso libro de texto *al-Kitaab*¹⁶ - acepta, en principio, que el proceso de lectura sea en esencia interactivo, puesto que combina procesos cognitivos ascendentes y descendentes. Sin embargo, define la lectura fluida como “*the ability to construct meaning from a text using all available semantic, syntactic, and morphological clues, without using a dictionary and without reading word by word in a linear fashion*” (2006:343). A mi juicio, Brustad mezcla en esta definición aspectos teóricos y didácticos; es decir, incorpora conceptos cognitivos referentes al proceso de la lectura y técnicas de estudio y aprendizaje a utilizar por el aprendiente. Por otro lado, su definición afirma que el lector no tiene por qué leer palabra por palabra para llegar a comprender un texto escrito, pese a que la literatura especializada proporciona suficientes pruebas empíricas de que el lector eficaz ha de leer cada una de las palabras del texto, de forma lineal en su mayor parte, puesto que el reconocimiento de vocabulario escrito es un proceso que se produce de forma automática (Carver 1992:85,93, Alderson 2000:15-16, Grabe 2009:102-103).

Es interesante que, para Brustad, la clave para la comprensión lectora sea el conocimiento gramatical y sintáctico, y su puesta al servicio de la construcción del significado (Brustad 2006:343). En este sentido, la autora realiza otra afirmación que contradice la literatura especializada: “*Before processing the meaning of individual lexical items, the reader needs to deconstruct the syntax of the sentence*” (Brustad 2006:344)¹⁷ e, incluso, afirma que los lectores nativos y de nivel Superior de árabe¹⁸ procesan sintácticamente los textos escritos de forma no lineal (Brustad 2006:344). Sin embargo, la autora no aporta datos empíricos ni estudios previos válidos que apoyen esas afirmaciones¹⁹. Su intuición acerca de que las estrategias de lectura descendentes desempeñan

¹⁶ A partir de aquí utilizaremos la abreviatura *al-Kitaab* para referirnos a la serie de libros publicada bajo el nombre *al-Kitaab fii Ta'allum al-'Arabiyya* (Brustad et al.), en sus diferentes ediciones.

¹⁷ La afirmación parece decantarse más por los modelos de lectura descendentes (Grabe 2009:89, 102-103) que por los modelos interactivos, pues Brustad apela a la capacidad del lector de dirigir la mirada a ciertos aspectos gramaticales y sintácticos importantes presentes en el texto a partir de supuestos mecanismos de inferencia que, en el caso de Brustad no aparecen especificados, y que son propios de los modelos de lectura descendentes.

¹⁸ El nivel Superior en ACTFL equivale a un nivel C1 del *Marco*.

¹⁹ Aunque las afirmaciones de Brustad no encajan con los datos empíricos disponibles en el área de investigación generalmente aceptadas (Grabe 2009:103), sí tienen sentido en el marco de la pedagogía propuesta por *al-Kitaab*, pues se trata de un libro de texto que ofrece al aprendiente inicial textos de lectura que contienen un 90% de léxico que le es desconocido (Hansen 2010:568). Esto es también,

un papel central en la lectura en árabe (por ejemplo, el análisis de los pronombres clíticos y la relectura parcial) fue bien explicada por Mughazy (2005)²⁰.

La escasa literatura sobre la lectura en ALE se enriqueció significativamente gracias al análisis exhaustivo de la investigadora danesa Gunna Hansen. En un artículo publicado en 2010 (*Word Recognition in Arabic as a Foreign Language*), la autora señala que los modelos de lectura universales no son suficientes para explicar el proceso de la lectura en árabe, pues están basados en investigaciones realizadas en el ámbito de lenguas que utilizan el alfabeto latino y las evidencias indican que cada sistema de escritura se procesa de forma diferente (Hansen 2010:567). Aunque Hansen no es pionera en afirmar la importancia de este asunto²¹, creo que es la primera que lo hace en el contexto de la EALE y subrayando sus posibles implicaciones didácticas.

La línea de investigación de Hansen se enmarca en el *enfoque de destrezas múltiples* y, por tanto, la investigadora parte de la necesidad de reconocer el léxico de forma automática para poder desarrollar una buena competencia lectora (Hansen 2010:577-579, 2014:55-62). En definitiva, Hansen se suma a la teoría conectivista (Grabe 2009:86-88)²², según la cual el reconocimiento automático del léxico es esencial para la comprensión fluida del texto impreso.

En un segundo artículo titulado *Word Recognition in Arabic: Approaching a Language-Specific Model* (2014), la investigadora danesa propone un modelo de lectura que incorpora diversos aspectos pertenecientes a los modelos ascendentes y descendentes. Aunque Hansen se apresura a afirmar que no se adscribe a los modelos interactivos más tradicionales, sí acepta, en cambio, la existencia de diferentes destrezas y subdestrezas cognitivas que están muy interconectadas y que se retroalimentan para

quizás, lo que lleva a Brustad a afirmar que los aprendientes de árabe deben analizar las oraciones sintácticamente antes de recurrir al diccionario (Brustad 2006:344). Sin embargo, esta estrategia compensatoria es propia de lectores no eficaces, es decir, propia de aprendientes que no han llegado a automatizar el reconocimiento visual del léxico (Grabe 2009:102-103). Abu Rabia (1997:466), sin embargo, rechaza la afirmación de Grabe y señala que esto no se aplica al árabe debido a la naturaleza misma de la ortografía árabe y a la presencia de un alto número de homógrafos. Hablaré de este punto en el siguiente apartado.

²⁰ La literatura especializada sobre la lectura en lenguas que utilizan el alfabeto latino también apunta a que el lector da saltos y, en ocasiones, ha de releer partes del texto para una comprensión eficaz (Nation 2009:31), pese a lo cual no necesariamente realiza un análisis sintáctico de las oraciones para poder comprenderlas. De hecho, el lector nativo eficaz se fija ocularmente en 95 de cada 100 palabras, en su mayor parte de forma lineal (Nation 2009:31).

²¹ Abu Rabia (1997, 1998, 2001, 2002), Urquhart y Weir (1998:34) y Koda (2005) ya advertían de los peligros de aplicar modelos de lectura basados en el inglés a lenguas que utilizan otros sistemas de escritura.

²² Grabe señala que la teoría conectivista no es un modelo de lectura *per se*, sino una teoría cognitiva que está ganando fuerza en el campo del procesamiento de la información y que se ajusta a varios modelos de lectura.

la decodificación exitosa de la escritura (Hansen 2014:60-62). Todo ello, en definitiva, recuerda la posición de Mughazy (2005) en relación con la lectura en L1.

En esta investigación adopto, con Hansen (2010), la idea de que el reconocimiento visual del vocabulario es el elemento fundamental en el desarrollo de la destreza lectora en aprendientes de ALE. Por tanto, me posiciono a favor del *enfoque de destrezas múltiples* y de los *modelos de reconocimiento léxico* de corte conectivista basados en datos empíricos obtenidos mediante experimentos de seguimiento ocular (Grabe 2009:101-102).

C. LA COMPLEJIDAD DE LEER EN ÁRABE

Diversos autores han puesto de relieve la complejidad que presenta la lectura en árabe (Justice 1987, Khaldieh 1996, Abu-Rabia 1997, Abu-Rabia 2001, Al-Batal 2006, Mughazi 2008, Habash 2010, Hansen 2010, Ryding 2013, Eviatar y Ibrahim 2014, Hansen 2014). Según la literatura especializada, el aprendiente de ALE se enfrenta, entre otras, a las siguientes dificultades durante la lectura: la ausencia de marcadores de vocales breves, de la *shadda* y de las mayúsculas, la similitud que presentan muchas letras entre sí, y los limitados signos de puntuación utilizados en algunos textos, además del complejo sistema morfológico y derivativo del léxico, y la enorme cantidad de vocabulario, en comparación con otros idiomas.

Como afirmaba con anterioridad, la lectura implica varias y muy complejas operaciones cognitivas. Mientras que aprender a leer en la L1 supone aprender una grafía para poder codificar, decodificar y relacionar entre sí un vocabulario ya conocido por el hablante (Hudson 2007:60, Grabe 2009), aprender a leer en una L2 requiere adquirir tanto el vocabulario nuevo como su correspondiente forma escrita y realización fonética.

Cuando la L2 utiliza un sistema de escritura diferente al de la L1, decodificar la nueva grafía se convierte en un reto y, en ocasiones, en un problema que persiste durante largo tiempo. De hecho, para algunos aprendientes, esa puede convertirse en la tarea más difícil; y, si no es tratada de forma rápida y adecuada, puede llegar a transformarse en un serio obstáculo en el proceso de aprendizaje de la L2. Para aquellos aprendientes cuya L1 utiliza el alfabeto latino, la lengua árabe representa un sistema de escritura nuevo que conlleva, sin duda, algunos retos y dificultades. A continuación

hago un repaso a algunos factores que afectan directamente a la comprensión lectora en árabe (L1 y L2).

1. La grafía

El alfabeto árabe está compuesto por 28 letras²³. Se escribe de derecha a izquierda y la mayoría de las letras requieren conectarse entre sí. Esto conlleva que, dependiendo de su posición, muchas letras cambian de forma, en mayor o menor medida. Además, existen algunos grupos de letras que comparten un esqueleto gráfico básico, diferenciándose únicamente por el número y posición de los puntos diacríticos (por ejemplo, *ب ت ث*). El alfabeto árabe es, fundamentalmente, consonántico y normalmente las vocales breves no se representan, aunque sí las largas. Estas se representan a través de las letras (*ا - و - ي*). Las vocales breves se representan a través de tres signos diacríticos raramente utilizados:

- فتحة (/fatha/) para la /a/, representada a través de una (*َ*) encima de la consonante.
- ضمة (/Damma/) para la /u/, representada a través de una (*ُ*) encima de la consonante.
- كسرة (/kasra/) para la /i/, representada a través de una (*ِ*) debajo de la consonante.

Una de las complejidades que habitualmente subraya la literatura especializada (Khalidieh 1996, Eviatar y Ibrahim 2014:82, Hansen 2014:63) es la carga cognitiva que teóricamente supone procesar letras que son gráficamente muy similares (por ejemplo, *ح خ*) y que solo se distinguen por el número y/o posición de los puntos, o que varían de forma dependiendo de su posición en la palabra (*ه هـ هـ*). Como consecuencia, argumentan algunos investigadores, leer en árabe supone más tiempo que hacerlo en otras lenguas semíticas, como el hebreo (Abu-Rabia 2001, Eviatar y Ibrahim 2014:83-84). Se ha afirmado, incluso, que leer lenguas escritas en alfabetos semíticos requiere una mayor fijación visual en las palabras que hacerlo utilizando el alfabeto latino (Shimron y Sivan 1994) en L1, aunque esto no significa que leer en árabe o en hebreo sea un proceso más lento (Roman y Pavard 1987:439), pues el léxico árabe impreso se

²³ A veces se habla de 29 letras, por contabilizar la *ا* y la *أ* como dos letras separadas. Para una discusión de esta cuestión, ver Habash (2010:9).

caracteriza por ser altamente clítico, a diferencia de otros idiomas. De hecho, algunos investigadores opinan que la densidad morfológica del léxico árabe podría ser la causa principal de que la fijación visual que se produce durante el procesamiento del árabe sea mayor que la de lenguas escritas en el alfabeto latino (Hermena et al. 2015).

En definitiva, la supuesta complejidad propia del sistema de escritura utilizado por la lengua árabe sería responsable de que la adquisición automática de las relaciones grafema-fonema no se produzca de forma adecuada, lo cual entorpecería la automatización del reconocimiento visual del léxico, tan necesaria para que pueda producirse una lectura fluida (Eviatar y Ibrahim 2014:82).

Lo anterior parece abonar la opinión de la investigadora de origen iraní Niloofar Haeri, quien niega que el sistema de escritura árabe en sí sea, por ejemplo, culpable de la baja alfabetización y práctica lectora de los árabes, pues la lengua persa utiliza el mismo alfabeto, con alguna variación, y los resultados son bien distintos (Haeri 2009:424). De hecho, Wagner (1993) ya había probado en un trabajo etnográfico-lingüístico longitudinal, utilizando a niños marroquíes como sujetos de su investigación, que aprender a leer árabe es un proceso similar a aprender a leer en otros idiomas que no utilizan el alfabeto árabe, y que el predictor de lectura exitosa más importante es aprender a decodificar las letras (forma, variaciones, y pronunciación) (Wagner 1993:79-106).

En definitiva, se diría que las dificultades asociadas a la decodificación de la grafía árabe, en comparación con la grafía latina u otras grafías distintas, no han sido empíricamente probadas en su totalidad.

2. *Opacidad ortográfica*

La ortografía árabe suele considerarse muy *opaca* (Abu-Rabia 1998, 1999, 2004, 2013, Mughazy 2005, Eviatar y Ibrahim 2014), en comparación con ortografías más *transparentes* características de otras lenguas, como el turco y el español o, incluso, en comparación con otras ortografías *opacas*, como es la del inglés o el francés. Por *transparencia ortográfica* se entiende la correspondencia directa entre grafema y fonema, mientras que la *opacidad ortográfica* supone una correspondencia no clara entre ambos (Feldman y Turvey 1983, Frost et al. 1987, Katz y Frost 1992). Sin embargo, el investigador norteamericano de origen palestino Nizar Habash señala que la ortografía

árabe o, lo que es lo mismo, la representación fonética de la grafía, se encuentra en realidad a medio camino entre la ortografía del español o el finés, que poseen una ortografía transparente, y la ortografía de lenguas como el inglés o el francés, considerada opaca (Habash 2010:31).

La razón principal por la que la ortografía árabe se considera opaca es que los marcadores diacríticos (vocales breves, *šadda*, y *tanwīn*) no se utilizan en la mayoría de los textos escritos. De hecho, únicamente el 1.6% de las palabras empleadas en la prensa árabe aparecen diacritizadas (Habash 2010:11). Sin embargo, cabe la posibilidad de escribir textos árabes completamente desambiguados y así ocurre, de hecho, en algunas publicaciones religiosas, literarias y escolares (Khaldieh 2001:419, Maamouri et al. 2004, Habash 2010:11, Hermena et al. 2015). Mughazy indica, acertadamente, que hablar de escritura árabe *transparente* (como en el caso de El Corán, por estar el texto completamente desambiguado) y de escritura árabe *opaca* (como en el caso de la prensa o la literatura) es simplificar la realidad lingüística, pues existen diversos grados de opacidad en la escritura árabe, en lo que el autor denomina el *continuum de opacidad* (Mughazy 2005:60). En cualquier caso, por norma general los textos árabes no utilizan los signos diacríticos que marcan las vocales breves, a menos que el contexto requiera alguna aclaración, como suele ocurrir en el caso de algunos homógrafos y, en ocasiones, la voz pasiva. Sin embargo, la mayoría de las veces, aunque se necesiten, los marcadores diacríticos no aparecen.

Conviene señalar que, aunque la grafía árabe pueda ser percibida como opaca desde el punto de vista del aprendiente de árabe como L2, desde el punto de vista del hablante nativo de árabe el sistema de escritura árabe no necesariamente ha de serlo. Datos empíricos parecen indicar que, de hecho, los hablantes árabes con formación académica no experimentan mayores dificultades para leer y reconstruir el significado de textos árabes escritos (Khaldieh 1999:143, Mughazy 2005). Aunque también es cierto que quizás esos datos sean solo aplicables a lectores árabes avezados (o, según la terminología de la enseñanza de lenguas extranjeras, “lectores hábiles”) que poseen una larga experiencia lectora y un vocabulario mental amplio.

Otra opinión a tener en cuenta es la de Hermena et al. (2015), quienes afirman a partir de datos empíricos que, si los textos árabes aparecen completamente desambiguados, los lectores hábiles utilizan una estrategia lectora eficaz, a saber, ignorar los

marcadores diacríticos. Esto indica, a mi juicio, la idiosincrasia del proceso cognitivo y psicolingüístico involucrado en la decodificación del árabe como L1, y no la ineficacia del sistema de escritura que utiliza. Por tanto, el reto de leer en árabe como L2 pasa por identificar métodos didácticos apropiados que capaciten al aprendiz a desarrollar estrategias de lectura similares a las que desarrolla el hablante nativo, y no depender de la presencia de marcadores diacríticos para llevar a cabo la lectura de forma eficaz, al menos a largo plazo.

Un marcador diacrítico que ha recibido poca atención en el estudio de la lectura es la *šadda*. Algunos investigadores que se ocupan del procesamiento de lenguajes naturales señalan que la ausencia de la *šadda* dificulta el procesamiento electrónico del vocabulario árabe (Habash 2010:60-61), y probablemente este fenómeno también dificulta el proceso de la lectura en árabe, tanto en L1 como en L2.

De igual modo, la ausencia de mayúsculas que indiquen –como en el caso de lenguas que utilizan el alfabeto latino- la presencia de nombres propios, supone un reto en el procesamiento del léxico árabe (Habash 2010:15). Este fenómeno adquiere especial relevancia si tenemos en cuenta que en árabe casi todos los nombres propios (a excepción de los nombres propios adoptados de lenguas extranjeras) van asociados a un significado; lo cual contribuye a aumentar el número de homógrafos.

3. Homografía

Esta omisión de los marcadores diacríticos ha conducido a algunos investigadores a afirmar que la lengua árabe abunda en homógrafos, lo que impide la decodificación eficaz de gran parte del léxico (Abu-Rabia 1997:466, 469, Abu-Rabia 1998:105, Abu-Rabia 2001:40, Abu-Rabia 2002:305, Mughazy 2005:60-63, Hansen 2010:569, Hansen 2014:65). Abu-Rabia llega a afirmar que leer en árabe es una tarea difícil incluso para los lectores avezados por el alto índice de homógrafos presentes en esa lengua (Abu-Rabia 1997:466; 1998:106-107)²⁴.

²⁴ A pesar de que Abu-Rabia realizó sus estudios en colegios israelíes, sorprende que cite como ejemplo la palabra الله (“Allah”, sin signos diacríticos). Abu-Rabia afirma textualmente que esta palabra “can be very confusing, even with vowels, for advanced readers” (1998:107). Quizás el entorno pueda haber influido en sus opiniones y en los resultados de sus investigaciones.

Cuantificar de forma exacta el nivel de homografía en árabe no es tarea fácil. Según Abu-Rabia (1998:109), 1 de cada 3 palabras en árabe es homógrafa con 2, 3 o incluso más significados posibles, aunque su afirmación no está avalada por datos empíricos y se diría una mera estimación personal. Farghaly y Shaalan (2009:10), basándose en experimentos realizados por SYSTRAN (compañía coreana dedicada a la traducción automática), afirman que cada token árabe escrito se presta, de media, a 19.2 interpretaciones, frente a 2.3 interpretaciones en inglés. Salloum y Habash (2014), por su parte, reducen el promedio de posibles interpretaciones a 12 por token y 2.7 por lema basándose en los resultados de un experimento realizado con el analizador morfológico BAMA (Diab y Habash 2014:54).

Conviene advertir que la ambigüedad léxica a la que se refieren dichos investigadores tiene únicamente en cuenta vocablos aislados²⁵. Es decir, la teórica ambigüedad que puede presentar un vocablo, podría resolverse, en muchísimas ocasiones, si dicha palabra apareciera en contexto. Por otro lado, otros investigadores señalan que la ambigüedad léxica no se debe únicamente a la homografía, sino que en ella intervienen otros factores como la estructura interna del léxico o el contexto sintáctico (Farghaly y Shaalan 2009:10-12).

En definitiva, no se sabe a ciencia cierta hasta qué punto afecta la homografía existente en árabe en la comprensión de textos escritos. La única forma de hacerlo sería desarrollando un instrumento que midiera empíricamente la capacidad de comprensión de textos escritos, especialmente para hablantes nativos, pero también para aprendientes de ALE. Sin duda alguna, el hablante nativo de árabe desarrolla mecanismos lingüísticos y cognitivos que le ayudan a desambiguar los homógrafos, pero hace falta llevar a cabo estudios experimentales para conocerlos con exactitud.

4. Morfología

Existen otras razones que me llevan a la conclusión de que la grafía árabe no es necesariamente más complicada de procesar, a nivel cognitivo, que la grafía latina. De hecho, los modelos de lectura más recientes parecen indicar que los textos árabes escritos y, en particular, el vocabulario, son procesados de forma diferente –que no más

²⁵ Para ejemplos concretos, puede visitarse la siguiente página web, que indica las diferentes lecturas que puede tener cualquier palabra árabe escrita: <http://quest.ms.mff.cuni.cz/cgi-bin/elixir/index.fcgi?mode=resolve>

difícil- a cómo se procesan otras lenguas. Sami Boudelaa (2014), por ejemplo, aporta datos provenientes de la psicología cognitiva, la neuropsicología y la neurociencia cognitiva que apuntan a la existencia de dos entidades cognitivas separadas en la mente del hablante nativo adulto que facilitarían la decodificación del léxico escrito y oral: la raíz y la forma (*al-ʔadr* y *al-wazn*) de las palabras. El investigador tunecino señala que, a la hora de enfrentarse a un texto escrito, el hablante nativo de árabe extrae de forma automática y por separado la raíz y la forma de todas las palabras contenido, en lo que viene a denominar *Decomposición Morfológica Obligatoria* (Boudelaa 2014:47-48). Datos de esta naturaleza podrían significar que los hablantes nativos de árabe poseen, antes de comenzar a leer y a escribir, un conocimiento morfológico del registro hablado de la lengua que, por ser compartido en cierta medida con el registro escrito, facilitaría el reconocimiento léxico y de su significado durante la lectura²⁶. Este tipo de investigaciones vendría a confirmar datos obtenidos por la investigadora marroquí Fatima Badry (1983) quien aporta información que confirma que los niños árabes adquieren durante la infancia estrategias que les permiten crear vocabulario novedoso utilizando su conocimiento de las formas nominales y verbales (citado en Wagner 1993:102, Badry 2009) y su capacidad extraer la raíz de palabras conocidas para combinarlas con formas adquiridas (Badry 2005:250). Por tanto, parece razonable pensar que el hablante nativo interioriza el sistema morfológico del árabe cuando aprende a hablar (registro oral); lo cual le ayudaría más tarde a adquirir las destrezas lectoras (registro escrito) necesarias para decodificar y comprender con éxito el léxico que aparece en los textos escritos²⁷, en consonancia con la teoría de la *Decomposición Morfológica Obligatoria* de Boudelaa.

En el caso de los aprendientes de ALE, dicho conocimiento morfológico no existe a priori, sino que ha de ser construido poco a poco hasta llegar a interiorizarlo. Por tanto, el estudio explícito del sistema morfológico, acompañado de una metodología adecuada, podría facilitar la tarea de la lectura en árabe como L2 y del reconocimiento automático del léxico impreso.

²⁶ En realidad, esto está en consonancia con el principio general de que la lengua materna hablada adquirida durante la infancia influye en la posterior adquisición de la destreza lectora.

²⁷ Mughazy (2005:61) alude a esta idea cuando se refiere al *conocimiento intuitivo*.

5. *Diglosia*

Los factores que pueden afectar a la comprensión del vocabulario escrito son muy heterogéneos y difícilmente testables. Algunos autores (Wagner 1993, Haeri 2009, Hansen 2014:71-72, Saiegh-Haddad y Spolsky 2014) subrayan el impacto de la diglosia en el proceso de aprendizaje de la lectura y de la escritura en árabe como L1, es decir, aprender a leer y a escribir en una variedad lingüística diferente a la hablada. Existen diversos estudios realizados por investigadores israelíes –muchos de ellos publicados en Saiegh-Haddad y Joshi (2014)- que atribuyen el bajo índice de alfabetización y de lectura en los países árabes principalmente al fenómeno de la diglosia en árabe y al hecho de que los árabes no sean alfabetizados en su lengua materna. Estos investigadores proponen que la alfabetización comience en la variante dialectal del árabe (Myhill 2014), y que tenga en cuenta el vocabulario propio de los registros hablados (Saiegh-Haddad y Spolsky 2014).

Si bien es cierto que la diglosia puede influir en la adquisición de ciertas destrezas lingüísticas, este tipo de generalizaciones no ayuda a esclarecer la supuesta dificultad que presenta la decodificación de la escritura árabe en comparación con otras grafías, ni los niveles de alfabetización de los árabes. A mi juicio, es preciso observar y analizar más detenidamente los métodos pedagógicos implementados en la enseñanza de la escritura y la lectura (como L1) y evaluar la adecuación de los materiales didácticos utilizados, algo que ya se ha empezado a estudiar (Haeri 2009, Papadopoulus 2013, 2014, 2015).

En el contexto de la EALE, la diglosia podría tener cierto impacto en el caso de estudiantes que estén aprendiendo simultáneamente el MSA y algún registro oral dialectal, o en el caso de estudiantes que empiecen aprendiendo alguna variante dialectal y pospongan el MSA y la lectura en dicho registro para etapas posteriores. Sin embargo, no existen estudios al respecto.

6. *Sistema educativo*

Entre las investigaciones en L1 conviene destacar una serie de artículos publicada por la sociolingüista egipcia Zeinab Ibrahim en colaboración con varios investigadores (Papadopoulos et al. 2013, 2014, 2015). Estos estudios apuntan a que la enseñanza del alfabeto árabe a niños de 5 años a través de un currículum que combine el

método comunicativo, el aprendizaje comunitario y el uso de tecnología educativa contribuye a un mejor reconocimiento de las letras, ya que genera mayor confianza en el niño. Los autores citados demuestran que ese tipo de currículum también contribuye al desarrollo de una actitud positiva del aprendiente hacia la lengua árabe escrita, a diferencia de lo que experimenta el aprendiente árabe en un contexto educativo tradicional. Teniendo en cuenta que aún no existen estudios acerca del bagaje léxico que los niños árabes poseen en torno a los 5 años (Papadopoulos et al. 2013, 2014), sería interesante hacer un seguimiento longitudinal a los niños participantes en los estudios piloto citados, que tuvieron lugar en Catar, para cuantificar la adquisición de vocabulario en comparación con niños educados con pedagogías tradicionales, y para medir la velocidad a la que leen los niños (palabras por minuto). El estudio podría, incluso, ofrecer algunos datos interesantes acerca del posible impacto de una adquisición exitosa del alfabeto árabe en la práctica de la lectura extensiva en años posteriores. Dicho seguimiento podría también aportar datos acerca de la influencia que pueden tener las políticas educativas implementadas en el contexto escolar árabe²⁸ y, así esclarecer las razones de la supuesta dificultad que experimentan muchos hablantes nativos de árabe a la hora de leer en árabe y del bajo índice de lectura que se observa entre los árabes en general²⁹ en comparación con otras poblaciones cuya L1 no es el árabe.

Esta misma postura la sostiene la conocida lingüista y antropóloga estadounidense de origen iraní Niloofar Haeri, quien, en un artículo pionero sobre la situación de la lengua árabe y el alfabetismo en los países árabes, sostiene que la educación pública en Egipto –y por extensión en casi todos los países árabes– *“falla a todos los niveles: desde los libros de texto, pasando por aulas aglomeradas, profesores mal entrenados, etc. De hecho, incluso enseñando un registro lingüístico que no es la lengua materna*

²⁸ No solo en la edad más temprana, sino en la enseñanza de las humanidades hasta la edad universitaria, en tanto que recurre a textos escritos para transmitir el conocimiento. Así, por ejemplo, no existen, hasta donde conozco, estudios que hayan medido la velocidad a la que leen los estudiantes árabes a diferentes edades, tal y como se ha hecho en inglés.

²⁹ Según un informe publicado por la UNESCO en 2007-2008, los países árabes están a la cola de países alfabetizados (Myhill 2014:198-199). Además, según fuentes periodísticas supuestamente basadas en estadísticas científicas, el ciudadano árabe lee de promedio menos de un libro al año (<http://raseef22.com/culture/2015/04/23/reading-habits-in-the-arab-world/>). Como ya he dicho, sería interesante examinar hasta qué punto el sistema educativo es responsable de tal desinterés por la lectura o si existen otros factores socioculturales involucrados, como podrían ser: el estatus socioeconómico, los valores que se tienen acerca de la lectura, la percepción de la L1, o los hábitos parentales en el entorno familiar, especialmente en la infancia. Estos factores son muy influyentes en el desarrollo de la lectura (Grabe 2009:152-174, Grabe y Stoller 2013:52-53), y algunos de ellos ya han sido considerados y estudiados por Wagner (1993) y Haeri (2009) en relación con el árabe.

de nadie (MSA) si hubiera mejores libros de texto y escuelas, el índice de alfabetismo sería más alto” (Haeri 2009:420).

En el contexto de la EALE, la didáctica implementada dentro y fuera del aula en la enseñanza/aprendizaje de la lectura tiene, sin duda, un gran impacto en la competencia lectora de los aprendientes; pero, hasta donde conozco, no ha sido investigada, a excepción de los artículos de Hansen en relación a la adquisición del vocabulario escrito, y que ya repasé. En esta misma línea, sería importante realizar estudios experimentales para determinar qué prácticas didácticas son las más efectivas en la enseñanza del reconocimiento visual del léxico.

7. Conclusión

Teniendo en cuenta todos los factores anteriores, parto de la suposición de que, en principio, el alfabeto árabe no presenta un serio problema para los hablantes nativos que han recibido una educación escolar sólida (al igual que ocurre con otros alfabetos e idiomas), aunque sí parece suponerlo para aquellos que no han sido escolarizados durante suficiente tiempo y para los que han aprendido a leer siguiendo metodologías no adecuadas. Incluso si se aceptara la premisa de que el alfabeto árabe es problemático, parece razonable aceptar la postura de Mughazy (2005:71-72) de que el lector árabe hábil desarrolla ciertas *estrategias adaptativas* para poder procesar con éxito un sistema de escritura que no es del todo transparente, como puede ser recurrir a su intuición lingüística, adquirida durante la infancia. En cualquier caso, opino que, por normal general, el contexto socio-económico árabe y los métodos didácticos implementados en los diferentes sistemas educativos son inadecuados para desarrollar destrezas lectoras y hábitos de lectura comparables a los utilizados en algunos países desarrollados.

Además, parto de la suposición de que la riqueza morfológica del árabe no es una desventaja a la hora de procesar textos escritos. De hecho, la mayor fijación visual que experimentan los hablantes nativos sobre el léxico impreso se debe a que éste es altamente clítico. Si acaso, podría decirse que el sistema de escritura árabe es altamente económico, pues requiere menos signos gráficos que, por ejemplo, las lenguas indoeuropeas que utilizan el alfabeto latino. En este sentido, no estoy de acuerdo con la visión extendida entre algunos investigadores de que el alfabeto árabe es, de por sí, inadecuado o problemático.

Conviene aceptar, sin embargo, que para el aprendiente de ALE, la ausencia de marcadores diacríticos que señalen las vocales breves y la *shadda* en el léxico escrito supone, probablemente, el mayor reto al que se enfrenta, puesto que la ausencia de un léxico mental inicial y de un sistema morfológico interiorizado, como el que adquiere el hablante nativo durante su infancia, dificulta la tarea de relacionar el léxico escrito con su correspondiente realización fonética y su correspondiente significado. Puedo, incluso, afirmar que el aprendiente de ALE experimenta, en muchas ocasiones, dificultades a la hora de reconocer la representación escrita del léxico, puesto que no dispone de referencias léxicas mentales de naturaleza semántica y fonética a las que recurrir, tal y como ocurre en el caso del hablante nativo, incluso en el caso de palabras que han sido visualizadas y estudiadas con anterioridad³⁰. Esto dificulta y ralentiza el proceso de aprendizaje de vocabulario árabe y entorpece la comprensión de textos escritos en general, a veces durante muchos años, especialmente si los métodos didácticos implementados no son los adecuados.

Para los aprendientes iniciales de ALE, cuya L1 está basada en el alfabeto latino, leer en árabe es una tarea que supone un desafío por varias razones: tienen que decodificar un nuevo alfabeto que se escribe de derecha a izquierda, deben prestar mucha atención al número y posición de los puntos diacríticos y han de reemplazar todas las vocales breves, *šaddas*, y *tanwīn* omitidos por su correspondiente realización fonética; todo lo cual ralentiza el proceso de lectura. En conclusión, para los aprendientes de ALE, el alfabeto árabe es incompleto, desde el punto de vista de la realización fonética; aunque es difícil saber hasta qué punto este factor socava la eficacia del mismo, especialmente si no se tienen en cuenta los métodos didácticos implementados en la enseñanza de la grafía, la ortografía, la lectura y la fonología.

Teniendo en cuenta, pues, la literatura existente y la naturaleza misma del sistema de escritura empleado por la lengua árabe, puede decirse que, para poder decodificar el léxico de forma eficaz, los aprendientes de ALE necesitan disponer previamente de bastante información ortográfica, fonológica y morfológica. Una forma de reducir el hándicap que supone la decodificación del léxico impreso es centrarse en aprender

³⁰ Como se verá más adelante, estudiar el vocabulario no es garantía de adquisición. De hecho, conocer una palabra implica conocer diversos aspectos de la misma en fases diferentes (Schmidt 2009).

el vocabulario que es utilizado con más frecuencia en la lengua hasta llegar a automatizar el reconocimiento visual del mismo. Este concepto de frecuencia no es únicamente aplicable al léxico del árabe, sino que puede extenderse a las raíces mismas del léxico, cuya productividad y consiguiente frecuencia deberían ser un factor esencial, a la hora de diseñar un currículum de ALE sólido que ayude al aprendiente a desarrollar estrategias cognitivas que faciliten la comprensión léxica a nivel semántico y contribuyan a su consiguiente aprendizaje.

A continuación, expondré brevemente las diferencias existentes entre los dos tipos de lectura más utilizados en la enseñanza de lenguas extranjeras (lectura intensiva vs. lectura extensiva) y después realizaré un repaso a la literatura científica que ha investigado el papel que desempeña el vocabulario en la lectura extensiva y la importancia de automatizar el reconocimiento del léxico escrito más frecuente para desarrollar una competencia lectora fluida y exitosa.

D. MODALIDADES DE LECTURA

En la vida cotidiana leemos numerosos tipos de material impreso en múltiples circunstancias por varios motivos y de diferentes modos (Grabe 2009:4-13, Grabe y Stoller 2013:5-10). Casi siempre, la finalidad de la lectura condiciona en gran medida la modalidad en que lo hacemos y determina las diferentes destrezas cognitivas y psicolingüísticas que se ponen en marcha. Esto ocurre tanto en la L1 como en la L2 y, de hecho, existen varias clasificaciones de los tipos de lectura (Grabe 2009:8). Para los objetivos de la presente investigación, y teniendo en cuenta la realidad del aula de lenguas extranjeras, paso a distinguir entre los dos tipos de lectura más comunes: la lectura intensiva y la lectura extensiva.

1. Lectura intensiva

La primera de las dos grandes modalidades de lectura es la lectura intensiva, la cual se caracteriza por el uso de textos más bien breves, generalmente 300-500 palabras, y el análisis directo de los mismos (Nation, 2013:219). Además de tener como objetivo la comprensión del texto mismo, el profesor suele guiar la atención de los aprendientes

hacia el vocabulario, la gramática, las características discursivas y el desarrollo de estrategias de lectura (Renandya y Jacobs 2002:296, Nation 2009:25-27, Nation, 2013:219-220).

Este tipo de lectura puede llevarse a cabo de forma eficaz o no, dependiendo del tipo de ejercicios que acompaña al texto (Nation 2009:28-37); y no debe utilizarse como método principal para desarrollar el léxico de los aprendientes (Nation 2009:37-38). En todo caso, la lectura intensiva debe prestar atención solo al vocabulario frecuente que aparece en los textos (Nation 2009:38).

En general, la lectura intensiva es el tipo de lectura que, de forma mayoritaria, se practica en el aula de lenguas extranjeras, en detrimento de la lectura extensiva.

2. *Lectura extensiva*

En contraposición a la lectura intensiva, la lectura extensiva implica leer grandes cantidades de textos durante largos periodos de tiempo, sin aparente esfuerzo y sin necesidad de interrumpir constantemente dicha actividad (Grabe 2009:311). En el contexto de la enseñanza de lenguas extranjeras, la lectura extensiva también se denomina *lectura placentera* o *lectura voluntaria* (Day y Bamford 1998, Krashen 2004).

El principal objetivo de la lectura extensiva es promover el placer por la lectura silenciosa, sostenida e independiente, a través de la cual los aprendientes pueden consolidar sus destrezas lingüísticas (Day y Bamford 1998:xiii-xiv). Según el conocido investigador estadounidense William Grabe, los beneficios de esta práctica son varios: adquisición de léxico y de estructuras sintácticas, automaticidad en el reconocimiento del vocabulario, desarrollo del conocimiento que se tiene del mundo y mejora de las destrezas de comprensión lectora en general. Otros investigadores destacan que la lectura extensiva fomenta el interés y la motivación del aprendiente durante el aprendizaje de lenguas extranjeras y promueve el sentimiento de confianza en sus capacidades lingüísticas (Day y Bamford 1998:6). Es más, se afirma que la lectura extensiva debe ser un fin en sí mismo (Nation 2009:49), a través del cual se pueden obtener todos los beneficios anteriormente citados.

Tradicionalmente, la lectura extensiva ha sido confundida con la lectura practicada bajo el prisma del método Gramática-Traducción; es decir, leer con la ayuda del

diccionario cualquier tipo de texto y por muy difícil que sea. Es por ello que la lectura extensiva ha sido y sigue siendo, en gran medida, rechazada por muchos docentes de lenguas extranjeras.

En este contexto, es necesario tener en cuenta que, para que puedan lograrse los beneficios anteriormente citados, una de las características fundamentales de la lectura extensiva es que los textos elegidos estén al nivel del aprendiente (Day y Bamford 1998: xiii, Grabe 2009:311). Sin embargo, es difícil encontrar textos auténticos adecuados y variados, a diferentes niveles de competencia lingüística. Como explicaré más adelante, esto ha derivado en la creación de materiales de lectura nivelados o graduados para ser utilizados en la enseñanza de lenguas extranjeras. De esto hablaré más adelante.

La lectura extensiva, a veces denominada *lectura para la comprensión general*, es un tipo de lectura que evoluciona y se desarrolla a través de una larga práctica, que conduce a la automatización del reconocimiento del léxico impreso, de las relaciones sintácticas y semánticas que existen entre las palabras, y de otro tipo de operaciones más complejas involucradas en la comprensión del texto (Grabe 2009:10, 311). Sobre el vocabulario y la automatización del mismo hablaré más adelante.

Para los objetivos de la presente investigación me interesa la práctica de la lectura extensiva y el proceso de automatización que desarrolla el lector (en este caso, el aprendiente de ALE) en el reconocimiento del vocabulario escrito. Dentro de este marco, me parece oportuno reiterar la definición de lectura extensiva que ofrecí en la primera página de esta investigación: “*La práctica de leer fuera del aula, narrativa escrita y adaptada al nivel lingüístico de los aprendientes, con el objetivo de promover el placer de la lectura y desarrollar diversas destrezas lingüísticas y culturales*”³¹. A continuación, repasaré el uso que se ha dado al texto literario en el aula de lenguas extranjeras y destacaré la importancia de recuperarlo en la enseñanza del ALE.

³¹ La atención que merece la narrativa como género empleado en la práctica de la lectura extensiva será argumentada en el capítulo titulado *Uso del Texto Literario en la Enseñanza de Lenguas Extranjeras*.

CAPÍTULO II: LOS TEXTOS LITERARIOS EN LA ENSEÑANZA DE LENGUAS EXTRANJERAS

En este capítulo repasaré el uso del texto literario en la enseñanza de lenguas extranjeras bajo diferentes enfoques metodológicos; expondré los beneficios de utilizar textos literarios como material de lectura y, finalmente, analizaré el papel que desempeña la literatura en el currículo del ALE en los contextos académicos español y estadounidense.

En el marco de la presente investigación utilizaré el término “literatura” para referirme, en esencia, al texto narrativo en prosa, en particular la novela y el cuento corto. En caso de referirme a otros géneros literarios (como la poesía o el teatro), lo haré notar de forma expresa.

A. EL TEXTO LITERARIO Y LOS ENFOQUES METODOLÓGICOS

Numerosos autores han resumido de forma magistral el tratamiento que a lo largo del último siglo se le ha otorgado a la literatura en el aula de lenguas extranjeras (Duff y Maley 1990, Kramsch y Kramsch 2000, Martín Peris 2000, Acquaroni 2007, Carter 2007, entre otros). El objetivo de esta investigación no es ahondar en la cuestión. Baste con señalar, a grandes rasgos, que cuando a principios del siglo XX primaba el Enfoque Gramática-Traducción, de tradición filológica³², los textos literarios constituían el eje de la enseñanza de lenguas extranjeras, la “esencia” de la misma (Maley 1990:3) y, a menudo, la razón principal por la que se estudiaban lenguas extranjeras (Richards y Rodgers 1986:11). Los textos se elegían atendiendo a su valor canónico y

³² Aunque en el siglo XIX este enfoque estaba ya institucionalizado en toda Europa, sus raíces se remontan al siglo XVI cuando comienza a extenderse el estudio riguroso de la gramática del latín y el griego (Acquaroni 2007, Richards y Rodgers 1986). Teóricamente, esta corriente pedagógica quedó obsoleta a finales de los años 30 del siglo XX, aunque ha continuado aplicándose en numerosos contextos educativos hasta hoy.

el profesor y el aprendiente los sometían a una minuciosa disección lingüística que contribuía a sacralizar esos textos literarios (Acquaroni 2007:49). En definitiva, la literatura representaba, ante todo, el canon superior de la buena escritura y un modelo lingüístico y estilístico que había de ser traducido, analizado e imitado.

Cuando en los años 20 se implanta el Método Directo, de base estructural y centrado en el análisis de estructuras gramaticales y léxicas como vía para el dominio de la lengua extranjera, se abandona la traducción (Richards y Rodgers 1986:18). Entonces, la literatura comienza a ser marginada y se convierte en simple pretexto para llevar a la práctica ejercicios mecánicos enmarcados dentro de la teoría conductista del aprendizaje, que se basaba en el concepto de estímulo-respuesta (Martín Peris 2000:101, Acquaroni 2007:49-50). El texto literario pasa entonces a considerarse parte del ‘método tradicional’.

Cuando a mediados de los años 40 aparece el Enfoque Audio-lingüístico (en EEUU) o Enfoque Oral (en Gran Bretaña), ya se entiende que la literatura solo es accesible al estudiante que domina las estructuras lingüísticas de la lengua extranjera (Kramersch y Kramersch 2000:563). En esta nueva corriente pedagógica lo que prima es perfeccionar la expresión oral, que pasa a ocupar a consecuencia de las necesidades surgidas tras el fin de la II Guerra Mundial un lugar central en la práctica didáctica, tanto en el ejército estadounidense (Richards y Rodgers 1986:49-50) como en ciertos programas académicos europeos conscientes de los importantes cambios sociopolíticos que sucedían en Europa (Martí Tormo 2012:40). En este contexto, y a modo de reacción, algunos especialistas comienzan a criticar la deshumanización de la enseñanza de lenguas extranjeras (Kramersch y Kramersch 2000:565) y llaman al uso de la literatura como vehículo de entendimiento (Johnson 1956:234); sin embargo, los profesionales tardarían aún en hacerse eco de este llamamiento.

Con el advenimiento del Enfoque Comunicativo y los métodos nocional-funcionales en los años 80, el texto literario queda relegado a una posición marginal. De hecho, se intercala en los métodos de lenguas extranjeras de forma anecdótica e inconsistente y, en la mayoría de los casos, sin pautas didácticas para su explotación pedagógica (Duff y Maley 1990:3, Acquaroni 2007:50). La literatura parece ser irrelevante en el proceso de adquisición de habilidades comunicativas, y solo aparece en los libros de texto de forma esporádica y para añadir un toque de variedad. Con todo, y en contra

de lo que pudiera parecer, es en el seno de este enfoque donde el uso del texto literario comienza a experimentar una lenta recuperación (Acquaroni 2007:50-51).

La aparición en el Enfoque Comunicativo del concepto de competencias lingüísticas supuso la inclusión del texto literario entre los textos auténticos que pueden servir como pretexto para la comunicación (Collie y Slater 1987:3, Carter 2007:6). Al mismo tiempo se redescubre la literatura como instrumento útil para acceder a la ‘cultura auténtica’ de los hablantes de las lenguas extranjeras (Collie y Slater 1987:4). Se recupera, pues, el texto literario partiendo de que puede ser utilizado como ventana a la cultura meta y como vía para la adquisición de vocabulario, además de como instrumento para el desarrollo de estrategias y habilidades analíticas relacionadas con la lectura. La literatura, pues, “*inicia un camino sin retorno de normalización, revalorización y paulatina incorporación a los materiales didácticos, gracias al reconocimiento de su particular fuerza procedimental*” (Acquaroni 2007).

La reincorporación completa de la literatura al currículo de lenguas extranjeras viene de la mano de la Instrucción por Contenidos, corriente que surge en los años 80 en el seno del Enfoque Comunicativo y recupera el texto literario como texto auténtico que puede contribuir a formar a alumnos competentes en los planos lingüístico y literario (Brinton et al. 1989 citado en Elyasi 2013). Esta filosofía de la enseñanza de lenguas extranjeras da finalmente paso, a principios del siglo XXI, al nacimiento de una corriente educativa humanista de Ciudadanía Global, cuyo objetivo es desarrollar en los aprendientes (ciudadanos) competencias de índole global que les permita resolver situaciones reales complejas, sirviéndose de un abanico de conocimientos y habilidades cognitivas, sociales, morales y emocionales aprendidas a través de diferentes disciplinas, entre otras, la literatura utilizada en el aula de lenguas extranjeras (Díaz-Pérez et al. 2013).

B. LOS BENEFICIOS DE LEER TEXTOS LITERARIOS

Los argumentos a favor de la incorporación del texto literario en el aula de lenguas extranjeras existen desde principios del siglo XX. Las publicaciones que se ocupan de la necesidad de incorporar el texto literario al currículo, aún en épocas en que las

prioridades pedagógicas se centraban exclusivamente en los planos lingüístico y comunicativo, abundan en las revistas académicas especializadas desde, por lo menos, los años 20 del siglo pasado (Kramersch y Kramersch 2000). Es complejo, por tanto, realizar un seguimiento completo de ellas.

Los manuales teóricos dedicados a la enseñanza de lenguas extranjeras, particularmente los monográficos que se ocupan de la destreza de la lectura y la adquisición de vocabulario también recogen el uso de la literatura como material didáctico esencial en el desarrollo de ciertas competencias lingüísticas (Nuttall 1982, Alderson y Urquhart 1984, Brumfit y Carter 1986, Nation 2013).

Las constantes referencias a la integración del texto literario como elemento central en el currículo dan paso, en los años 80 y 90, a la publicación de obras de impacto dedicadas, en exclusiva, a destacar el papel de la literatura en la enseñanza de lenguas extranjeras, y a proporcionar al profesor multitud de ideas prácticas para explotar pedagógicamente los textos literarios. Entre esas obras destacan *Literature and Language Teaching* (Brumfit y Carter 1986), *Literature in the Language Classroom: A resource book of ideas and activities* (Collie y Slater 1987), *Literature* (Duff y Maley 1990), y *Literature and Language Teaching: A guide for teachers and trainers* (Lazar 1993).

A continuación, resumiré los beneficios más relevantes del uso del texto literario, según han sido expuestos por los expertos en enseñanza de lenguas extranjeras, en los últimos 50 años.

1. Beneficios lingüísticos

Los beneficios lingüísticos que obtiene el estudiante de lenguas extranjeras al exponerse a textos literarios son de índole diversa. A nivel léxico, la literatura prueba el amplio y sutil uso del vocabulario (Povey 1967:41-42). Por tanto, la práctica de la lectura extensiva aumenta el vocabulario receptivo del lector, lo cual facilita su reutilización y producción en la modalidad oral y en la escrita (Widdowson en Rossner 1983, Collie y Slater 1987).

Por otra parte, la lectura extensiva, si se sostiene a lo largo de toda una novela, por ejemplo, desarrolla en el aprendiente la capacidad de hacer inferencias y de deducir el significado de palabras por contexto, destreza que puede extrapolarse a la lectura de cualquier otro tipo de textos (Collie y Slater 1987:5). Este beneficio se extiende a la competencia auditiva, pues a partir de la escucha de material literario grabado en audio puede adquirirse asimismo una gran cantidad de vocabulario (Lazar 1993). Como explicaré más adelante, este beneficio lingüístico está estrechamente relacionado con la adquisición incidental de vocabulario a partir de la lectura extensiva.

A nivel pragmático y sociolingüístico, la literatura es ideal para que el estudiante tome conciencia del uso del lenguaje en diferentes contextos (McKay 1982:530). Al exponerse a una gran variedad de discursos literarios, el aprendiente se familiariza con diferentes usos, formas y convenciones lingüísticas (Collie y Slater 1987:4). Además, los textos literarios ofrecen una gran variedad de estilos, registros y tipologías (Duff y Maley 1990) que permiten al aprendiente acercarse a un discurso desmarcado de las reglas estándares del lenguaje (Widdowson en Rossner 1983) y que le ofrece otro tipo de sintaxis, colocaciones y connotaciones que contribuyen a sensibilizarlo ante la diversidad sociolingüística (Lazar 1993, Martín Peris 2000, Acquaroni 2007).

Además de todo esto, algunos investigadores han recalcado que el texto literario expone a los estudiantes a diferentes recursos estilísticos (Brumfit y Carter 1986, Lazar 1993, Carter 2007), en concreto a los metafóricos (Macdonald y Macdonald 1983).

Todos esos beneficios lingüísticos se reflejarían, además, en una producción lingüística más creativa por parte del aprendiente (Collie y Slater 1987), puesto que, como dice Povey “*literature may guide [...] students towards their own creativity by example derived from their reading of successful writers*” (Povey 1967:42).

2. Beneficios culturales

Más allá de los beneficios lingüísticos, tradicionalmente centrados en aspectos formales de la lengua, muchos autores se han referido a la literatura como vehículo

transmisor de la cultura. Uno de los primeros en afirmar que el aprendiente debe conocer la cultura de aquellos que hablan la lengua objeto de estudio fue el académico estadounidense Scott (1964:489), quien defendía la práctica de una ‘orientación cultural’ de los alumnos. Para Scott, el conocimiento de la cultura a través de la literatura promueve, como fin último, el entendimiento entre personas y culturas, pues la literatura, como producto cultural, constituye, en palabras de este autor, una vía “obvia” y “valiosa” para hacernos comprender la cultura meta, ya que el texto literario “*frequently isolates for close scrutiny the aesthetic, moral, spiritual, and social values that a given community cherishes –values and patterns of behavior which too often are not clear to the casual observer*” (Scott 1964:490). Con el tiempo, esta idea evoluciona. La académica estadounidense McKay, por ejemplo, afirma que la literatura “*may work to promote a greater tolerance for cultural differences for both the teacher and the student*” (McKay 1982:531).

El argumento, basado en la idea de que los textos literarios presentan una gran cantidad de información cultural beneficiosa para el aprendiente de lenguas extranjeras, fue particularmente desarrollado por la canadiense Joanne Collie y el australiano Steven Slater, reputados expertos en la enseñanza de lenguas extranjeras y en el diseño curricular de materiales docentes. Los autores del popular manual *Literature in the Language Classroom* (1987) señalan que el lector puede aprender acerca de los diferentes modos de vida existentes en otros países, al exponerse a mundos literarios vívidos que reflejan aspectos tan diversos como “*their thoughts, feelings, customs, possessions; what they buy, believe in, fear, enjoy; how they speak and behave behind closed doors*” (Collie y Slater 1987:4). Esto sería especialmente beneficioso para aquellos estudiantes que no pueden viajar o disfrutar de programas de inmersión total en los países en los que se habla la lengua extranjera.

Esta forma indirecta de exponerse a la cultura meta también fue ensalzada por la reconocida especialista británica en diseño curricular Gillian Lazar. Aun admitiendo que la literatura no refleja necesariamente de forma cabal o directa la cultura del pueblo en cuya lengua está escrita la obra y considerando el riesgo de que el aprendiente ge-

neralice los aspectos culturales presentes en una determinada obra a una o más sociedades en su conjunto³³, Lazar afirma que la literatura, en efecto, “puede constituir un acceso a la cultura” (Lazar 1993:16). Estas constataciones tienen lugar en una época en la que se acepta el uso de la literatura como elementos clave para desarrollar la competencia y el entendimiento culturales que, como docentes, procuramos promover en nuestros estudiantes (Bernhardt 1995:6).

Poco después aparecen dos obras básicas de Claire Kramersch, académica europea afincada en EEUU, que incorporan a la ecuación pedagógica de lenguas extranjeras el concepto de “tercer espacio” que, en realidad, es una nueva conciencia que nace en el aprendiente como consecuencia de haberse expuesto a un nuevo sistema de valores y prácticas culturales (Kramersch 1993) y que promueve la empatía cultural.

Por último, según algunos investigadores, uno de los mayores beneficios culturales que se obtiene a partir de la lectura de textos literarios es que el aprendiente logra poner en perspectiva la representación de las manifestaciones culturales (Harklau 1999, Kubota 1999 y 2003). Al presentar discursos culturales complejos y varios, la literatura contribuye a evitar una imagen monolítica de la cultura meta. En otras palabras, la literatura presenta perspectivas culturales matizadas y ayuda a evitar estereotipos y generalizaciones. Esta visión poliédrica de la cultura, que a su vez permite al lector desarrollar una empatía para con la cultura meta, permite adquirir la competencia intercultural (Kramersch 2005).

3. *Beneficios motivacionales*

La literatura académica anglosajona ha afirmado de los textos y materiales de lectura tradicionalmente utilizados en la enseñanza de lenguas extranjeras que “recrean mundos [...] demasiado familiares” y que, por eso mismo, los aprendientes “no se interesan por ellos” (Widdowson en Rossner 1983). Se ha llegado incluso a manifestar

³³ La advertencia de Lazar se funda en que el inglés es vehículo de comunicación en muchas sociedades y países y, por tanto, no puede decirse que las obras literarias escritas en inglés representen la cultura de un país concreto. Lo mismo puede decirse de la literatura árabe, escrita, en su mayoría, en una variante lingüística (MSA), que no refleja necesariamente la cultura de un único país o sociedad.

que se trata de textos que “no tienen ningún valor” (Breadsley 1936:4), que son “insípidos” (Breadsley 1936:3), “triviales” (Gamel 1984:8, Duff y Maley 1990:6) y que “hacen al estudiante perder el tiempo, y son peligrosos” (Gamel 1984:8). En definitiva, textos irrelevantes para los intereses y preocupaciones reales de nuestros estudiantes (Lazar 1993:3).

Por el contrario, la literatura es fuente de motivación para el estudiante de lenguas extranjeras. La autenticidad del texto literario, entendida como producto que surge de las preocupaciones reales del autor, provoca en el lector una motivación genuina, pues las temáticas presentes en el mismo están directamente relacionadas con las vivencias del receptor-aprendiente. Al tratar “temas complejos” y “dilemas propios de adultos” (Lazar 1993:15), el texto literario es capaz de interesar y absorber a los estudiantes más que las “pseudo-narrativas” que normalmente se incluyen en los libros de texto. Para el investigador británico y experto en lingüística aplicada y enseñanza de lenguas extranjeras Henry Widdowson, las habilidades interpretativas que el estudiante ha de activar a la hora de decodificar el discurso literario interesante y atractivo es lo que le mantiene “enganchado” (en palabras de Rossner 1983) a ese mundo de ficción, pues, en definitiva, lo que hace el lector del texto literario es resolver un rompecabezas lleno de suspense y emoción.

En este sentido, existen otros factores de naturaleza psicológica que pueden proporcionar a los aprendientes motivos para leer, como son los factores afectivos, de actitud y vivenciales (McKay 1982). El texto literario, si está bien escogido, provoca respuestas emocionales intensas en el lector-aprendiente, lo que indicaría que estamos involucrando a la persona que es el aprendiz (Lazar 1993:3). Como conclusión de las tesis expuestas, traigo a colación las palabras de Brian Tomlinson, experto en desarrollo de materiales pedagógicos para la enseñanza de lenguas extranjeras: “*Laughter, joy, excitement, sorrow and anger can promote learning. Neutrality, numbness and nullity cannot*” (Tomlinson 2003:18).

4. *Beneficios psicológicos*

Así pues, el texto literario permite involucrar emocionalmente a los aprendientes como personas con capacidad afectiva; y, por este motivo, los estudiantes sienten que “lo que hacen en el aula es relevante y significativo para sus propias vidas” (Lazar 1993:15). Este proceso de aprendizaje, facilitado por el texto literario, especialmente si va acompañado de ejercicios para explotarlo pedagógicamente (Collie y Slater 1987), desarrolla en el aprendiente una mayor confianza en sí mismo, al sentirse capacitado lingüística y culturalmente, el estudiante experimenta una sensación de logro que refuerza aún más su autoestima.

5. *Beneficios humanistas*

El impacto de la literatura en la formación educativa del individuo, entendida en su sentido más amplio y global, ha sido destacado desde principios del siglo XX; aunque es, con la aparición de la psicología humanista, en los años 60, cuando toma una mayor relevancia en el contexto de la enseñanza de lenguas extranjeras. Así, los académicos británicos Bright y McGregor venían a decir en su manual *Teaching English as a Second Language* que aprender un idioma extranjero a través de la lectura de textos literarios contribuye a recibir una educación global y a desarrollar todo el potencial de la persona, pues “*literature sharpens sight and insight, widens sympathies and experience and provides occasions for the exercise of judgment about man and his condition. It helps with the main business of education, the production of men and women capable of an appropriate response to life*” (Bright y McGregor 1970: 53).

También se ha puesto en evidencia que la diversidad de cuestiones humanas fundamentales que aparece en la literatura, como el amor, la muerte, la amistad, la venganza o el odio, tiene un gran impacto en el lector-aprendiente, pues desarrollan en él el pensamiento crítico y la conciencia emocional (Collie y Slater 1987). Es más, al remitirnos constantemente a “experiencias y vivencias personales, a sentimientos, a opiniones”, el texto literario enriquece a la Persona y le otorga relevancia (Lazar 1993).

Como he analizado, interactuar con el texto literario expone necesariamente a los aprendientes a dilemas morales, éticos y sociales, pero también a dilemas de naturaleza cognitiva (como la decodificación o la interpretación del texto). Por tanto, se puede afirmar que la literatura desarrolla en los aprendientes habilidades críticas y un pensamiento más agudo (Lazar 1993, Díaz-Pérez et al. 2007). Para las corrientes pedagógicas humanistas y comprometidas con la educación de ciudadanos globales, el uso de textos literarios en el aula de lenguas extranjeras contribuye a educar a las personas en la empatía y, por tanto, a formar a personas en la responsabilidad ciudadana, moral, social y política (Phillip 2013).

6. *Implicaciones para la EALE*

A mi juicio, todos los beneficios arriba expuestos son sumamente valiosos y deben tenerse en cuenta a la hora de diseñar un currículo de ALE sólido, especialmente si va enfocado a la preparación de estudiantes universitarios. En el marco del presente estudio me interesa, sobre todo, recalcar el potencial del texto literario en prosa como fuente de lectura extensiva capaz de exponer al aprendiente a un importante volumen de vocabulario de forma recurrente y, por tanto, capaz de desarrollar automaticidad en el reconocimiento del léxico impreso. El texto narrativo ofrece múltiples beneficios lingüísticos y, si es diseñado y ajustado al nivel del aprendiente, resultará un instrumento muy potente en el desarrollo de la destreza lectora de los estudiantes.

A continuación, repasaré el uso que se ha dado al texto literario en prosa en el aula de ALE y analizaré la percepción que tienen los docentes acerca de su incorporación en el currículo, en los contextos español y estadounidense.

C. *CÓMO SE PERCIBE LA LITERATURA EN EL CONTEXTO DE LA EALE*

Los manuales dedicados a la didáctica del árabe como lengua extranjera en los contextos académicos español y estadounidense han incorporado el texto literario al menos desde mediados del siglo XX, aunque haya sido de forma asistemática y, en la mayor parte de los casos, desvinculada de un marco pedagógico explícito. Esto no es exclusivo de la enseñanza del árabe; se trata, más bien, de un fenómeno generalizado en la didáctica de las lenguas extranjeras (Kramersch y Kramersch 2000, Martín Peris 2000). Sin embargo, hay que reconocer que en el contexto de la EALE existen muy pocos materiales desarrollados con el objetivo claro y específico de promover la lectura literaria y extensiva.

En este sentido, la disciplina de la EALE ha tardado mucho en elaborar un discurso teórico favorable a la incorporación del texto literario en el aula si se compara, por ejemplo, con la enseñanza del inglés o del español como lenguas extranjeras. De hecho, en los años 30 del siglo pasado había ya, en el contexto de la enseñanza del inglés como L2, artículos muy críticos hacia los materiales de lectura utilizados en los manuales y la actitud del profesorado respecto a su incorporación al currículo (Beardsley 1936 y 1937). Por contraposición, en el contexto de la EALE, los pedagogos y teóricos que llaman la atención sobre el papel que debe desempeñar la literatura en el proceso de aprendizaje de la lengua, no aparecen hasta mediados de los años 80, década que destaca por algunos momentos históricos de gran relevancia en cuanto al uso del texto literario.

En el contexto académico español, Federico Corriente (1984:11) subrayó la necesidad de conectar a los alumnos con la literatura árabe medieval y moderna, como manifestación cultural, en una época en la que el material literario se utilizaba en las aulas como pretexto para enseñar a traducir textos escritos en una lengua que se consideraba muerta (Corriente 2002:14, Paradela 2010:33). En cuanto al contexto estadounidense, es importante recordar la propuesta del arabista y crítico literario Roger Allen (1985) en las *Arabic Proficiency Guidelines* encargadas por el American Council on the Teaching of Foreign Languages (ACTFL) (Allen 1984). Las guías, redactadas por el propio Allen y su equipo de la Universidad de Pennsylvania, son relevantes porque

mencionan la comprensión de textos en prosa sencillos en el nivel Intermediate High (equivalente al B1 del *Marco*) y niveles posteriores, en un periodo en que el campo de la EALE estadounidense estaba inmerso en el desarrollo de las competencias lingüísticas orales de los estudiantes.

Como explicaré a continuación, la escasez de publicaciones científicas que se han ocupado de la didáctica del ALE (Martí Tormo 2012:59) en el contexto académico español es destacable y son mínimas las que abordan la incorporación del texto literario al currículo. En cuanto al contexto estadounidense, las referencias a la utilización de la literatura en el aula son algo más numerosas, aunque en ningún momento han centrado el debate pedagógico. A continuación, voy a repasar la literatura científica que se ha ocupado de la explotación del texto literario en la EALE y las opiniones prevalentes respecto a su utilidad, como vía para el desarrollo de las destrezas lingüísticas y culturales de los alumnos, tanto en España como en EEUU.

1. La situación en España

Muchas gramáticas elaboradas en el siglo XX para la enseñanza del árabe en España recurrían a textos literarios clásicos para ofrecer al aprendiente un modelo lingüístico que ejemplificara las reglas gramaticales explicadas en dichos manuales. Las introducciones redactadas por los autores de dichas obras gramaticales son un buen reflejo del discurso imperante en la época y, por tanto, paso a analizarlas.

La *Crestomatía* de Asín Palacios³⁴, por ejemplo, incluía 37 textos medievales andalusíes y orientales, la mayoría literarios, destinados a “*facilitar los ejercicios de análisis y traducción*” (Asín Palacios 1959:5). La *Crestomatía* iba dirigida a estudiante que quería iniciarse en el aprendizaje del árabe y la traducción, y fue diseñada para un curso de árabe elemental. Por esa razón, argumenta Asín Palacios, todos los textos aparecen completamente vocalizados. Sin embargo, el análisis detenido de los textos seleccionados parece indicar que requieren, al menos, un nivel B2 según la clasificación actual del *Marco*. Esta clasificación que se hace de los textos como “elementales”,

³⁴La edición que he consultado es la edición corregida de 1959.

cuando en realidad son mucho más complejos de lo que en principio se argumenta, será durante décadas una de las características más comunes en la descripción y presentación editorial de materiales pedagógicos dedicados a la EALE, tanto en España como en EEUU, como ejemplificaré más adelante.

Otra obra particularmente importante es la *Gramática* que el renombrado investigador Federico Corriente ha elaborado y ampliado a lo largo de varias décadas³⁵, con el fin de formar a arabistas y traductores. Quiero destacar, en concreto, una idea que no debe pasar desapercibida, y es la que nos ofrece Corriente en la tercera edición de su *Gramática árabe*, de que la literatura es reflejo de la cultura de los árabes y, por tanto, pedagógicamente, ha de estar presente en el currículo del ALE. De lo contrario, argumenta Corriente, estaríamos formando a arabistas que serían “*analfabetos funcionales*” y objeto de “*conmiseración*” por parte de los árabes (Corriente 1984:11). Aunque esta afirmación pueda resultar algo exagerada, lo que Corriente viene a decir es que los arabistas españoles que aspiren a una formación académica sólida necesitan alcanzar una competencia cultural que les permita comunicarse con los árabes de forma efectiva, al menos en la esfera intelectual y académica. Se trata, en el fondo, de lo mismo que diría Elsaid Badawi, años más tarde, al destacar el papel que debe desempeñar la lectura extensiva de literatura, en la formación de aprendientes capaces de identificarse lingüística y culturalmente con el hablante nativo educado (Badawi 2002:159).

Esa necesidad de ofrecer al alumno textos literarios la plasma Corriente, en cierta medida, en su *Gramática y textos árabes elementales* (1990), al incluir 28 textos literarios en prosa, la mayoría medievales y algunos modernos. Sin embargo, en el prefacio a dicha obra, Corriente justifica la variedad de textos utilizados por su deseo de mostrar a los estudiantes de árabe “*la notoria estabilidad gramatical del idioma árabe a través de sus al menos quince siglos de historia literaria*” (Corriente 1990:8). Aunque el argumento es válido pedagógicamente, parece que el texto literario posea únicamente una función lingüística; o, al menos, así se expresa.

No cabe duda de que, en el contexto universitario español, el texto literario ha constituido tradicionalmente uno de los pilares del proceso de enseñanza y aprendizaje

³⁵ Federico Corriente publicó la primera edición de su *Gramática Árabe* en Madrid con el Instituto Hispano-Árabe de Cultura (1980), que luego fue reimpressa a partir de 1986 bajo el título de *Introducción a la gramática y textos árabes*.

del árabe. El texto literario, casi siempre medieval, era el modelo lingüístico por antonomasia que se ofrecía al estudiante de árabe para aprender la gramática y el vocabulario. La literatura moderna y contemporánea, por su parte, se utilizó a partir de los 70 y 80 traducida (Gómez Camarero 1994, Paradela 2000), para iniciar una nueva corriente de investigación literaria apartada del ‘andalusismo’ imperante en la época y, también, para acercar a los estudiantes a la producción literaria árabe y, así, prepararles para la crítica literaria.

En el campo de la didáctica moderna del ALE, sin embargo, las publicaciones que se ocupan de la integración de la literatura en el currículo solo aparecen a partir de los 90. En este sentido cabe mencionar, en primer lugar, una serie de tres artículos publicados por Arvide Cambra entre los años 1997 y 2000, con el título *Apuntes sobre métodos de enseñanza y técnicas de aprendizaje de la lengua árabe*. En ellos la investigadora hace unas observaciones personales referentes a las técnicas pedagógicas utilizadas en algunas universidades estadounidenses y europeas que visitó durante los 90. Arvide Cambra cuenta que, a principios de aquella década, la literatura era, al igual que en España, asignatura obligatoria en todos los programas académicos que impartían cursos de árabe. La materia se estudiaba en el último año de carrera, y el texto literario solía ser moderno y medieval. En cuanto a la metodología aplicada en el aula, los docentes utilizaban la traducción y el comentario de texto como técnicas didácticas para acercar el texto literario al aprendiente. El objetivo de incorporar la literatura, dice Arvide Cambra, era, en esencia, que el estudiante perfeccionara su conocimiento de la gramática árabe. En algunos programas académicos, incluso, los textos literarios eran comentados por los alumnos en lengua árabe, un fenómeno que Arvide Cambra percibe como una novedad extraordinaria (Arvide Cambra 1997:45, 2000:113), pues en aquellos años esto era ajeno por completo a la práctica extendida en el sistema universitario español.

De hecho, en aquella época, los docentes que se ocupaban de la literatura en los programas de filología árabe acostumbraban a utilizar la lengua española para tratar el contenido de los textos y analizarlos. Esos textos podían ser modernos e, incluso, contemporáneos. Pero, sobre todo, estaba extendido el uso del texto medieval. Por este motivo, no sorprende el comentario de Arvide Cambra tras asistir a un curso de literatura que se impartía en la Universidad de Berkley en 1991 en referencia al tipo de textos manejados, cuando, con total espontaneidad, confiesa: “*Debo señalar que eché en falta*

la utilización de textos medievales” (Arvide Cambra 1997:45). El comentario refleja, en mi opinión, la mentalidad de muchos docentes españoles de aquella época respecto de la tipología de textos árabes que debían ser trabajados en el aula, sobre todo si se seguía, como era el caso de muchos docentes, el Método Gramática-Traducción. Por aquellos años, las técnicas comunicativas de tipo oral no estaban muy extendidas en las universidades españolas y mucho menos en la EALE. Esta realidad didáctica concuerda con una afirmación realizada por la investigadora, al concluir su serie de artículos: “*Se equivocan los que consideran el método de traducción como obsoleto y decimonónico; por el contrario, representa una parcela importantísima [...] en el aprendizaje de una lengua*” (Arvide Cambra 2000:17). Sin lugar a dudas, esta era la opinión prevalente entre los arabistas más tradicionales.

Una década después, en uno de los pocos artículos publicados en España en torno al uso de la literatura árabe en el aula de ALE, el profesor Waleed Saleh (2010a) hace un llamamiento a la incorporación del cuento como elemento indispensable. El autor afirma que los métodos dedicados a la enseñanza del ALE “*han marginado los textos literarios árabes -incluido el cuento corto- por considerar que dichos textos no son relevantes para las necesidades de la vida diaria y por considerar que los aprendientes (solo) desean aprender el lenguaje de la comunicación (verbal)*” (Saleh 2010a:219-220, traducción mía). Tras denunciar la marginación del texto literario en el currículo de lengua árabe, Saleh enumera los muchos beneficios que el alumno puede obtener a partir de la lectura del texto literario (lingüísticos, culturales y psicológicos). Saleh observa, por ejemplo, que leer literatura aproxima a los aprendientes a la cultura de los árabes, elemento indispensable en su formación académica; que el alumno puede aprender nuevos vocablos y estructuras gramaticales que aparecen en el texto literario; que la reutilización del lenguaje presente en el texto otorga al aprendiente una mayor confianza en sí mismo y en sus capacidades lingüísticas; y que el disfrute estético que se obtiene de la literatura “*transporta al lector y al oyente a un mundo imaginario y atractivo lleno de placer*” (Saleh 2010a: 220-222 y Saleh 2012a, traducción mía).

Aunque Saleh menciona diversas utilidades asociadas a la lectura literaria, habría sido útil cierta claridad en la exposición de las ideas planteadas y en la argumentación, y ello por dos razones fundamentales. La primera es que la clasificación que hace el autor de los beneficios que se obtienen a partir del texto literario es un tanto ambigua.

Por ejemplo, Saleh destaca como uno de los *principales*³⁶ beneficios “*el desarrollo de la competencia comunicativa del estudiante*” y “*el estudio de la gramática y la sintaxis*” (Saleh 2010a:221, traducción mía). Sin embargo, pocas líneas después considera *secundarios* otros objetivos que incluyen “*la práctica de la expresión oral en árabe a través de la participación activa*” y “*el estudio de estructuras gramaticales*” (Saleh 2010a:222, traducción mía). Esta ambigüedad en la exposición denota, en mi opinión, una reflexión poco sistemática acerca del papel que puede desempeñar el texto literario en el aula del ALE y de sus posibles objetivos didácticos. Dicho de otro modo, no quedan claras las bases teóricas sobre las que el autor fundamenta su convicción de la necesidad de incorporar la literatura en el aula de ALE

Por otro lado, Saleh no define el término “cuento” y deja al lector la tarea de deducir, por la forma en que discurre el razonamiento y los ejemplos que propone, que el vocablo hace referencia al cuento popular u oral de tradición folklórica y no al cuento corto moderno o contemporáneo. La distinción es importante porque se trata de tipologías textuales bien diferentes en cuanto a estructura, lenguaje y, muchas veces, contenido, y que pueden servir a objetivos pedagógicos también distintos.

Sin lugar a dudas, Saleh ha sido uno de los mayores defensores del uso del texto literario con fines pedagógicos lingüísticos y culturales en el aula de ALE, y así lo han corroborado varias publicaciones suyas, dedicadas al cuento tradicional y folklórico: *Cuentos tradicionales árabes: Antología didáctica y bilingüe* (Saleh 2003) e *Historias jocosas de Yuha: Selección comentada y bilingüe* (Saleh 2010b), sin omitir algunas obritas de teatro introducidas en su manual *Curso Práctico de Lengua Árabe I* (Saleh 2001).

Otro género literario que ha suscitado interés entre los docentes españoles de ALE es el teatro como vehículo pedagógico. Como acabo de mencionar, Saleh (2001) incluye algunas escenas de teatro breves en uno de sus manuales, pero no comunica al profesor ni al estudiante las razones pedagógicas de tal decisión. Aquellas obras breves de teatro habían sido escritas por el propio Saleh en tono humorístico y picaresco, “*inspirándose en historietas y anécdotas tradicionales que aparecen en diferentes fuentes*

³⁶ Las cursivas son mías.

*árabes de forma escueta a veces*³⁷, y algunas de ellas aparecen recogidas en su libro *Cuentos tradicionales árabes*, lo cual indica claramente el interés de Saleh por acercar la cultura popular árabe al alumno.

Herederero de esta tradición es Martínez Castro (2014), quien, por primera vez en una publicación académica, expone argumentos pedagógicos, parcialmente basados en *el Marco*, a favor de la utilización del texto teatral como instrumento que facilita la adquisición del ALE. El autor nos recuerda que el uso del texto literario: a) dota a la lengua de un uso contextualizado, más real que los textos hasta entonces utilizados en el aula, para practicar las estructuras formales de la lengua, b) expone al aprendiente a usos novedosos e imaginativos del lenguaje y, además, c) sirve como medio para acercarse a la realidad sociocultural meta (Martínez Castro 2014:240-241).

Además de repasar los beneficios del teatro en la enseñanza de lenguas extranjeras, Martínez Castro dirige una crítica al *Marco* y afirma que las directrices europeas no mencionan el uso del teatro hasta el nivel B2. El autor critica el hecho de que los niveles iniciales no incluyan el texto dramático, bajo ninguna de las destrezas lingüísticas y que solo aparezca a partir del B2, que es cuando “*figura por primera vez la comprensión auditiva de obras de teatro*” (Martínez Castro 2014:241). En mi opinión, la crítica no es del todo acertada, pues no se puede pretender que *el Marco* incluya la comprensión del texto teatral a partir del A1, dada la naturaleza de dichos textos, casi siempre superior a ese nivel. Lo que el autor pretendía demostrar en su artículo es que el uso del texto dramático es beneficioso desde los niveles A1-A2, a pesar de las posibles dificultades que pueda presentar. Así, por ejemplo, Martínez Castro defiende la memorización de textos teatrales en los niveles iniciales porque, argumenta, conduce a algunos beneficios lingüísticos y tiene un impacto muy positivo en los aprendientes.

Martínez Castro esgrime algunos argumentos a favor de la incorporación del texto dramático desde los niveles A1 y A2, y también en el B1, que pueden resumirse en los siguientes puntos (Martínez Castro 2014:241-248):

- Contribuye a una mejor adquisición de la lengua en general.
- Motiva a los aprendientes.

³⁷Comunicación personal mediante correo electrónico fechado el 24 de julio de 2015. La obra a la que se refiere el autor es *Cuentos tradicionales árabes: Antología didáctica y bilingüe* (2003).

- Les da confianza en sus habilidades lingüísticas orales y, en particular, mejora su pronunciación.
- Contribuye a reducir el índice de abandono entre los estudiantes de primer año³⁸.
- Beneficia, asimismo, a los compañeros de los aprendientes que participan en la obra de teatro, porque les permite observar el progreso lingüístico que puede alcanzarse a partir de dichas representaciones.
- Mejora las relaciones interpersonales en el programa académico, dentro y fuera del aula.
- Finalmente, en el contexto particular de la EALE, el teatro es “*una vía efectiva para suprimir los estereotipos tan patentes en nuestra sociedad en lo relativo a la cultura árabe islámica*”.

A juzgar por los escasos artículos que se han publicado en relación a la EALE, se observa una ligera evolución en el discurso didáctico a favor del texto literario. Sin embargo, cabe preguntarse ¿por qué los arabistas españoles no han tratado más el papel de la literatura como posible instrumento en la enseñanza de la lengua árabe? A continuación, intentaré contestar a la cuestión planteada, sin pretender ser exhaustiva pero aspirando a ser lo más objetiva posible.

Para empezar, en el contexto académico español, los docentes que se ocupan de la enseñanza del ALE suelen ser, al mismo tiempo, profesores de literatura árabe (medieval, moderna o contemporánea) y, a menudo, tienen una formación especializada en la crítica literaria. Por eso quizá no sorprenda y, en cierto modo, sea lógico, que la gran mayoría de las publicaciones que se ocupan de la literatura árabe hagan un tratamiento del texto literario desde el ángulo historicista y de la crítica literaria. Así, después de haber repasado los artículos publicados por docentes e investigadores españoles en *Anaquel de Estudios Árabes* (1990-2014), *Al-Andalus Maghreb* (1993-2014), *Awraq* (1988-2014), *Miscelánea de Estudios Árabes y Hebraicos: Sección Árabe-Islam* (1958-2015)³⁹, *Revista de Estudios Internacionales Mediterráneos* (2007-2014), y también las actas del Congreso ARABELE 2009 y ARABELE 2012, puedo afirmar que la mayoría de los artículos dedicados a la literatura árabe se ocupan de cuestiones englobadas

³⁸ Se trata de uno de los retos más importantes que afectan al campo de la EALE desde hace muchos años, como también afirmaban otros arabistas, en EEUU (Parkinson 1985, Belnap 1987).

³⁹ Los volúmenes 1-6 anteriores a 1958 no se encuentran online y no han podido ser consultados.

bajo la disciplina de la crítica literaria, como el contexto del texto literario, sus interpretaciones históricas y sociológicas a través del prisma literario, apuntes biográficos de autores conocidos e, incluso, estudios de literatura comparada; pero casi nunca se comenta la importancia de explotar el texto literario en el aula del ALE para fines lingüísticos, comunicativos y culturales. Muy de vez en cuando pueden leerse comentarios elogiando el texto literario como recurso que tiene potencial didáctico del tipo “*resultará útil para la enseñanza del árabe marroquí en España, además de para aquellos que quieran conocer de primera mano manifestaciones vivas de la cultura popular marroquí, e incluso se podría explotar en las aulas de las escuelas españolas*” (Fernández Parrilla 2012:156)⁴⁰, pero se trata únicamente de apuntes marginales que no se desarrollan ni ahondan en la cuestión.

A este aparente desinterés académico hacia la didáctica de la literatura como elemento integral del currículo del ALE debido, en parte y, como ya he dicho, a la formación y trayectoria misma de los docentes españoles, quizá habría que añadir otras razones basadas en las creencias mismas del profesorado. A pesar de que esta cuestión no ha sido investigada, algunas publicaciones dejan entrever la opinión extendida de que el texto literario es difícil. Así, por ejemplo, en un magistral artículo sobre la situación del estudio de la literatura árabe moderna y contemporánea en España, la investigadora Nieves Paradela considera que la traducción de obras literarias al español facilita y mejora la docencia del texto literario en el contexto universitario, puesto que de lo contrario el texto “*sería imposible o de muy difícil acceso al estudiante*” (Paradela 2000:237). Esta afirmación no se enmarca en el contexto particular de la didáctica del ALE y, sin embargo, como explicaré más adelante, refleja una percepción del texto literario árabe que ha sido también expresada por académicos estadounidenses. Me atrevería incluso a afirmar que esa es la visión más extendida entre los docentes de árabe hasta nuestros días, a menudo con razón.

Otros investigadores van más allá y añaden que la mayoría de los profesores de ALE en España carecen de una formación especializada y de una acreditación específica en la didáctica y la pedagogía moderna de la lengua (Aguilar 2011, Giménez Reillo 2014a y 2014b). Se ha argüido, incluso, que la situación económica del profesorado

⁴⁰ Comentario de Fernández Parrilla (2012) en relación a la obra de Moscoso *Aproximación al cuento narrado en árabe marroquí* (2012).

español y las políticas educativas de contratación y de financiación puestas en práctica por la universidad española no contribuyen favorablemente al desarrollo profesional continuo de los docentes (Giménez Reillo 2014b). Sin embargo, opino que esta no es razón suficiente ni principal que, por sí sola, pueda justificar la situación en la que se encuentra la didáctica del ALE en España, y mucho menos el número y la calidad de las publicaciones académicas pertinentes.

La cuestión que me ocupa, en realidad, está estrechamente relacionada con la situación del diseño curricular, un área al que sí se ha hecho referencia en algunas publicaciones españolas. Cabe recordar, por ejemplo, la durísima crítica que lanzó Corriente (1999) a la *Crestomatía* de Asín Palacios y sus usuarios, a quienes culpaba de entorpecer el progreso en materia de enseñanza del árabe como lengua extranjera y en materia de investigación lingüística⁴¹. Tales críticas no han sido comunes en el campo de la EALE en España y, posiblemente, hayan sido vistas con recelo. Sin embargo, debe señalarse que en los últimos 10 años, algunos docentes han hecho un llamamiento explícito a la necesidad de reflexionar sobre las metodologías empleadas en la EALE y a la importancia de diseñar nuevos materiales didácticos que ayuden a desarrollar de forma más efectiva las competencias lingüísticas del aprendiente de árabe (Barreda et al. 2007, Aguilar et al. 2010, Martí Tormo 2012:47 y 58-59). También es cierto que ha aumentado el número de publicaciones españolas destinadas a aprender árabe⁴², aunque quizá se ha exagerado, en alguna ocasión, el estado de la cuestión y se ha llegado a decir que estamos viviendo un “*avance vertiginoso*” (Aguilar 2011:84). Es verdad que se ha experimentado un progreso cuantitativo importante; sin embargo, habría que analizar con detenimiento si está siendo acompañado de un avance cualitativo en términos metodológicos y de una recepción positiva de dichos materiales por parte del profesorado y el alumnado español de ALE. De lo que no cabe duda es de que las nuevas publicaciones académicas reflejan una clara conciencia de la necesidad de renovar al-

⁴¹Aunque la *Crestomatía* ya no se utiliza en las universidades españolas, el manual seguía aún en uso hasta bien entrados los años noventa, en instituciones como la Universidad Complutense de Madrid.

⁴²La publicación de los métodos *Paso a paso* (2006), *al-Nafura* (2007), *al-Jadwal* (2008), *al-Shallaal* (2010), *Alatul!* (2011), *Basma* (2011), *B chuiya b chuiya* (2013), y más recientemente *Mabruk* (2014), además de todos los materiales publicados por la editorial Albujaayra, dan fe de la proliferación de materiales dedicados a la EALE en España, muchas veces impulsada desde las Escuelas Oficiales de Idiomas. En los últimos diez años también han sido publicadas algunas gramáticas nuevas, aunque en menor número que los libros de texto.

gunos materiales didácticos todavía en uso y, también, de modificar los métodos pedagógicos en práctica, algo en lo que coinciden incluso los estudiantes de árabe en España (Valera Laborda y Aguilar 2013:191).

Para terminar, quiero referirme brevemente a la situación en la que se encuentra la investigación de la lingüística aplicada y la adquisición del árabe como lengua extranjera en España. La separación existente entre dichas disciplinas y la práctica docente en el aula fue ya denunciada por Federico Corriente a finales de los 90, cuando hablaba del “*desinterés del arabismo español universitario por la lingüística, teórica o aplicada*” (Corriente 1999:69). En realidad, se trata de disciplinas científicas esenciales para el pedagogo que desee entender los mecanismos que intervienen en la adquisición y la enseñanza, de lenguas extranjeras, en particular del árabe en particular⁴³. Sin estudios de investigación en esta área, es muy difícil que pueda mejorar la situación del diseño curricular y las aplicaciones pedagógicas de la EALE en el aula.

2. *La situación en EEUU*

Algo diferente es la situación en el ámbito académico estadounidense, aunque tampoco sea excesivamente llamativo el interés que ha suscitado la incorporación del texto literario en el currículo del ALE. Como ya he señalado, Arvide Cambra (1997, 1998, 2000) reflejó con cierto detalle el uso que los docentes estadounidenses le daban al texto literario y dejó constancia de la aplicación del método Gramática-Traducción, que en aquella época se utilizaba con el objetivo de estudiar la gramática y el vocabulario. En realidad, esta aplicación pedagógica databa de tiempo atrás y así lo atestigua una reseña del filólogo estadounidense de origen ucraniano Jaroslav Stetkevycha a una obra coetánea dedicada a la enseñanza de textos literarios en el contexto académico

⁴³ Sería interesante hacer un estudio exhaustivo del número de tesinas y tesis doctorales producidas en las universidades españolas en el campo de la lingüística aplicada y teórica del árabe en todas sus variantes y, concretamente, aquellas relacionadas con la pedagogía del ALE. Esta cuestión, estaría, también, en estrecha relación con el número de conferencias y congresos dedicados a la didáctica del árabe y que, desgraciadamente, están en declive, sobre todo tras la cancelación de ARABELE por parte de Casa Árabe, y los escasos recursos económicos destinados a la universidad pública para celebrar eventos de tal impacto.

(Stetkevych 1966:352)⁴⁴. Una noticia similar recoge el profesor e investigador estadounidense Dilworth Parkinson cuando alude a la lectura de cuentos cortos originales en el último año de carrera con la ayuda del diccionario, una práctica que él mismo califica de “dolorosa y tediosa” (Parkinson 1985:11); y el profesor estadounidense de origen palestino Peter Abboud, cuando se refiere a los textos literarios medievales como material esencial de lectura en los programas universitarios de grado y posgrado (Abboud 1995:29).

Ahora bien, a pesar de la aproximación pedagógica tradicional al texto literario y el uso de textos medievales, implantado desde hacía décadas y que habría de continuar en los años siguientes, los docentes de ALE en EEUU habían empezado a desarrollar algunos materiales específicamente destinados a promover la lectura de textos literarios más modernos. Así, durante la década de los 60, se publicaron varias obras de entre las cuales destaco *A Reader in Modern Literary Arabic* (Ziadeh 1964), *Contemporary Arabic Readers* (Bellamy et al. 1963) y *Readings in Modern Arab Literature* (Brinner y Khouri 1961 y 1971)⁴⁵. El renacimiento que experimentaba la prosa árabe por entonces concienció a algunos docentes de la necesidad de incorporar textos literarios modernos y contemporáneos al currículo del ALE.

Como ocurría con algunos manuales españoles, las introducciones de esas obras dejan entrever algunas ideas prevalentes sobre el papel de la literatura en la EALE. Bellamy et al. (1963: v) y Brinner y Khouri (1971: vii) defienden la *lectura extensiva* de textos literarios (cuentos cortos y extractos de novelas) para que el alumno expanda su caudal léxico y adquiera expresiones idiomáticas y recursos estilísticos en árabe. Para Brinner y Khouri era esencial que el estudiante de árabe se familiarizase con las particularidades del nuevo lenguaje literario surgido a mediados del siglo XIX, puesto que se apartaba del modo de expresión árabe clásico, y que conociera los diversos géneros literarios que habían nacido en el mundo árabe. Aunque en ningún momento se explicita dónde realiza el estudiante la lectura del texto (fuera o dentro del aula), se insiste en que la lectura debe ser *extensiva*, es decir, amplia y variada⁴⁶.

⁴⁴ Se trata de una reseña a la obra de Ziadeh *A Reader in Modern Literary Arabic* (1964), de la que hablaré en detalle más adelante.

⁴⁵ La primera edición de esta obra fue publicada, según el Prefacio, en el año 1961. Sin embargo, no he logrado tener acceso a ella.

⁴⁶ El concepto de “lectura extensiva” en estas obras difiere del actual.

En cuanto al nivel apropiado para comenzar a leer los textos literarios seleccionados para dichas antologías, Bellamy et al. (1963: v) y Ziadeh (1964: ix) opinan que los alumnos de árabe deben conocer previamente un vocabulario básico y dominar la gramática árabe (morfología y sintaxis). En palabra de dichos autores, los alumnos con un nivel Intermedio estarían ya preparados para leer esos textos. Por su parte, Brinner y Khouri (1971: vii) afirman que la lectura de textos literarios en prosa debe realizarse tras dos años de estudio formal de la lengua. Pero lo más importante era tener un conocimiento sólido de la gramática árabe, algo que, en ocasiones, se lograba tras un año de estudio formal, ya que los glosarios que acompañaban a los textos literarios facilitarían su comprensión (Brinner y Khouri 1971: xi). En cualquier caso, Como solía ocurrir por aquella época, el nivel Intermedio no se corresponde con lo que actualmente entendemos por dicho nivel. Por el contrario, tal y como esclarece la Introducción al libro de texto *Intermediate Standard Arabic: Intermediate Level* (1971: v), el nivel Intermedio era generalmente el segundo semestre o el segundo año de carrera.

En cuanto a la cuestión del diseño curricular, es interesante y, me atrevo a decir, pionero, el comentario de Brinner y Khouri (1971: vii) acerca de la necesidad de diseñar material docente adecuado que aborde el texto literario y el uso moderno de la lengua árabe, especialmente para los niveles intermedio y avanzado. Dicho comentario podría interpretarse, incluso, como una crítica indirecta a los textos medievales que se venían utilizando en el aula.

Las tres obras que acabamos de mencionar no eran un fenómeno aislado en el campo de la EALE estadounidense. De hecho, uno de los libros de texto más populares, el *Modern Standard Arabic: Intermediate Level* de Abboud et al. (1971), incluía textos literarios modernos como parte esencial del material de lectura, textos que se utilizaban para presentar al estudiante notas léxicas y gramaticales esenciales. De igual modo, el método *Let's learn Arabic: A Proficiency-Based Syllabus for Modern Standard Arabic* (1986) parece que incluía textos literarios modernos y contemporáneos (Allen 1990:21)⁴⁷.

Como puede verse, la incorporación del texto literario moderno al material diáctico destinado al ALE, gozaba de popularidad entre ciertos docentes, aunque no

⁴⁷El método, cuyo coautor es Adel Allouche, no ha sido localizado y, por tanto, no he podido consultarlo.

fuera práctica común. La primera vez que se hace un llamamiento oficial a la incorporación del texto literario en el contexto universitario estadounidense fue en 1985, un momento histórico en el que Roger Allen propone una guía de competencias lingüísticas preliminares específicas al ALE, encargadas por ACTFL y seguidamente consensuadas por los docentes de árabe y adoptadas por la institución. En aquel documento, se explicitaba la necesidad de exponer a los aprendientes de árabe a textos literarios a partir del nivel Intermedio alto (B1 en *el Marco*), aun a sabiendas y admitiendo que el diccionario era todavía un instrumento muy utilizado en la lectura de las obras literarias. Ahora bien, es relevante el hincapié que hace Allen en la necesidad de que dichos pasajes literarios sean “poco complicados” o “especialmente adaptados para tal tarea” (Allen 1985:63), especialmente en nivel Intermedio. Las guías de competencia ofrecían unas pautas más o menos claras acerca del tipo de textos literarios que debían ser utilizados en cada nivel. En los niveles avanzados, por ejemplo, se aconsejaba que los textos fueran elegidos “con cuidado” y que las selecciones pertenecieran a la literatura contemporánea o moderna (Allen 1985:63-64). Únicamente en el nivel superior (C1 en *el Marco*) se aconsejaba introducir textos medievales. Alcanzado ese nivel se suponía que el aprendiente ya era capaz de leer novelas contemporáneas y obras de teatro, aunque aún con dificultad y valiéndose, en ocasiones, del diccionario (Allen 1985:64-65).

Si leemos con detalle la literatura académica estadounidense observamos que Roger Allen es el académico que más insiste, directa e indirectamente, en el uso del texto literario. Quizá esto se deba a su formación académica literaria y a su incesante interés en la pedagogía e incorporación de la literatura en el aula de ALE. Lo cierto es que, pocos años después de la publicación de la guía de competencias, el destacado investigador menciona, aunque solo sea de pasada, uno de los beneficios más importantes que puede ofrecer la literatura como instrumento didáctico: “*Literature from all periods can provide a rich source to illustrate aspects of society, both public and private*” (Allen 1990:13).

Por esa época, ya se sabía que una de las razones principales por la que los estudiantes universitarios se matriculaban en cursos de árabe era para poder acceder a la literatura y a la cultura árabes. Según una encuesta realizada por Belnap (1987), el 56,3% de los estudiantes estadounidenses y canadienses expresaba su deseo de leer literatura árabe y conocer la cultura meta.

En este contexto favorable a la literatura como instrumento pedagógico, Munt-her Younes (1990) publicaba una edición simplificada de cuentos de *Kalila wa Dimna* destinada a estudiantes de nivel intermedio. La reacción no se hizo esperar y Mahmoud al-Batal (1990) escribe una reseña en la que elogia la colección y el desarrollo de este tipo de materiales didácticos, que incorporan el texto literario al currículo del ALE. Parecía, pues, que se continuaba con la línea pedagógica iniciada en los años 60, en consonancia ahora con la opinión de ACTFL y la predisposición de los docentes y aprendientes de ALE.

Sin embargo, a partir de la segunda mitad de la década de los 80, irrumpe el Enfoque Comunicativo y los efectos se hacen sentir de inmediato en la EALE. Así, una encuesta también realizada por Belnap (1995:56) desvela que para un gran número de docentes de ALE la primera prioridad pedagógica es preparar a los estudiantes para la comunicación oral, siendo la segunda prioridad pedagógica capacitarles para la comprensión del lenguaje de los medios de comunicación. En tercer lugar, aparece el interés de los docentes por que los aprendientes puedan leer literatura árabe moderna y, más tarde, para leer literatura medieval. Dicho de otro modo, mientras que el objetivo primordial de los alumnos parecía ser, una década antes, acceder a la literatura y cultura árabes, los docentes están más interesados por un tipo de instrucción que obedece a tendencias pedagógicas oral-comunicativas, centradas en producir hablantes competentes que puedan superar con éxito el, por aquella época, recientemente establecido OPI⁴⁸, una prueba de aptitud lingüística esencial para acceder a ciertos puestos de trabajo en la Administración estadounidense.

Esta nueva tendencia pedagógica explica por qué la mayor parte de textos de lectura utilizados en el aula por aquella época provenían, y siguen proviniendo en la actualidad, de periódicos y revistas (Hammoud 1996:104); es decir, de la prensa. Esta nueva realidad didáctica no impidió, sin embargo, que con la emergencia de la socio-lingüística y el desarrollo de teorías culturales que llaman a la necesidad de incorporar la cultura meta en el currículo de lenguas extranjeras, aparecieran voces como la del reconocido académico Yasir Suleiman (1993), quien, apoyándose en estudios realizados en otras lenguas, defiende el uso del texto literario como medio útil en la transmisión y aprendizaje de los valores culturales de los hablantes meta.

⁴⁸Oral Proficiency Interview.

Al igual que ocurre probablemente en el contexto universitario español, existen experiencias y experimentos didácticos que no han sido trasladados a la literatura académica, o que lo han sido tardíamente. Es el caso, por ejemplo, de unos cursos de lectura literaria extensiva que fueron diseñados e implementados en el programa estadounidense Center for Arabic Study Abroad (CASA) por parte del prestigioso profesor egipcio Elsaid Badawi entre 1970 y 1977 (Badawi 2002)⁴⁹. Como refiere el conocido académico, aquellos cursos estaban dirigidos a estudiantes avanzados de nivel 2-3 en la escala del ILR⁵⁰ (equivalente a un B2-C1 en *el Marco*) cuyo objetivo era entrenarles en la lectura extensiva de textos literarios originales, con el fin de alcanzar el nivel 4 (equivalente a C2 en *el Marco*) y una absoluta independencia en la práctica de la lectura literaria.

En el artículo, Badawi se apoya en teorías y conceptos que favorecen la lectura extensiva para justificar el uso que, en su momento, había dado a textos literarios de diversa índole, como obras de teatro, novelas y cuentos cortos. El autor alude, por ejemplo, al papel central que desempeña la palabra impresa en la educación adulta, no solo durante los años universitarios, sino a lo largo de toda la vida. Partiendo de esta base y de la propia teoría de Badawi referente a los múltiples registros que existen en lengua árabe, el investigador argumenta que el aprendiente de árabe que alcance el nivel 4 (C2 en *el Marco*) debe asemejarse al hablante de árabe nativo educado, puesto dice:

“The educated adult [...] usually identifies with her/his counterpart in the target society, and given that the serious learner aims at gaining competence approximating, as much as possible, that of the educated native speaker, then the question maybe be answered by mapping out the totality of language skills an Arab adult is supposed to possess upon graduation from college” (Badawi 2002:159).

Para alcanzar tal nivel lingüístico, Badawi propone exponer al aprendiente de árabe a enormes cantidades de textos literarios, bajo un formato que denomina *Independent Reading*⁵¹. Badawi no se explaya en describir los beneficios que ofrece el texto

⁴⁹ Dichos cursos continúan hasta el día de hoy, aunque el primero en introducirlos fue Badawi.

⁵⁰ Abreviatura de Interagency Language Roundtable, una organización establecida por el Gobierno federal y el ejército estadounidenses en los años 50 para mejorar la formación lingüística de sus funcionarios.

⁵¹ Badawi (2002:174) señala que en el programa CASA los estudiantes leían unas 2.500 páginas a lo largo de 14 semanas.

literario, aunque señala que este tipo de lectura extensiva mejora el dominio del vocabulario, de las expresiones idiomáticas, y de estructuras lingüísticas que pueden transferirse a otros contextos y géneros textuales (Badawi 2002:172, 175). El autor, sin embargo, es consciente de la riqueza del texto literario y conoce muy bien el impacto que puede tener la lectura extensiva de novelas en el conocimiento lingüístico y cultural del estudiante, pues afirma que *“it is through the literary genres that the reader gains insight into the complexity of language expression and social phenomena in one package”*.

Una de las últimas publicaciones académicas que hace referencia explícita a la incorporación del texto literario en el currículo de ALE es la obra de la experta estadounidense Karin Ryding, *Teaching and Learning Arabic as a Foreign Language: A Guide for Teachers* (2013). En este manual que, en cierto modo, resume el estado de la cuestión actual en EEUU, Ryding reconoce de soslayo la importancia de exponer a los estudiantes de árabe a diferentes tipos de textos literarios. Decimos “de soslayo” porque para Ryding la literatura parece cobrar importancia únicamente en el contexto universitario y, sobre todo, en el caso de aquellos estudiantes que deciden cursar estudios de posgrado en literatura árabe. Para Ryding, la lectura de la literatura implica el “estudio de la literatura” (Ryding 2013:154); dicho de otro modo, la adquisición de las destrezas que habilitan al estudiante para leer textos literarios está muy ligada a la práctica de la crítica literaria. Estudiar literatura, afirma Ryding, significa adquirir una nueva competencia mucho más difícil que las otras competencias lingüísticas que forman parte del proceso de aprendizaje de lenguas extranjeras. Esta creencia se manifiesta en su introducción a la sección dedicada a la enseñanza del texto literario en el contexto del ALE, donde aborda la necesidad de realizar un *“shift in competency that not all students are interested in or likely to make”* (Ryding 2013:154). Se transmite, pues, desde un primer momento, la idea de que aprender a leer literatura es una tarea difícil, diríamos que casi imposible. Por si fuera poco, la propia Ryding asegura, un poco más adelante, que los contenidos, estilos, normas y referencias culturales que forman parte del texto literario *“are not easy to teach”* (Ryding 2013:155), sobre todo si tenemos en cuenta que existe una larguísima y amplísima tradición literaria en árabe. En conclusión, para la autora, la capacidad de analizar y de apreciar el texto literario se convierte en *“benchmarks of ability”*, reservado para una minoría y que distinguiría a los aprendientes de árabe más brillantes del resto (Ryding 2013:155 y 157).

Ryding no está en contra de la incorporación del texto literario al currículo del ALE. Por el contrario, reconoce que la literatura árabe ha sido poco explorada como recurso para estudiar las diversas representaciones culturales, la identidad de los árabes y su autopercepción (Ryding 2013:154). Sin embargo, la autora parece partidaria de utilizar textos que forman parte del canon literario árabe tradicionalmente aceptado por los críticos, pues los textos excluidos del canon, por muy atractivos y relevantes que sean, “*remain to be discussed, studied, and defined in terms of their relation to Arabic cultural literacy*” (Ryding 2013:154). En principio, para Ryding, el hecho de que los expertos en nuestra disciplina no hayan alcanzado aún un consenso en torno a cuáles son los textos literarios canónicos y legítimos que deben utilizarse en el contexto de la EALE supone una desventaja a la hora de planificar el currículo universitario; pero es, al mismo tiempo, una oportunidad para la experimentación, siempre y cuando los departamentos estén dispuestos a negociar y aceptar nuevas tendencias literarias y culturales (Ryding 2013:154). Sin embargo, esta ventana a la experimentación en el aula no parece razón suficiente para dar un giro radical y colocar al texto literario en una posición central o siquiera relevante del currículo. De hecho, Ryding prefiere los textos literarios clásicos y señala que constituyen una parte importante del currículo de ALE y, por tanto, es difícil imaginar la implementación de técnicas pedagógicas comunicativas e interactivas, puesto que la mayor parte de los docentes continúa utilizando el texto literario en la lectura intensiva, la traducción y el análisis detallado del texto (Ryding 2013:155).

Si bien es cierto que la autora cierra esta sección de su manual recordando algunas de las actividades más interactivas propuestas por Kramsch para abordar el texto literario, Ryding no especifica con qué objetivo debemos incorporar el texto literario en el aula del ALE y solo menciona la competencia cultural y la estilística, caracterizada esta última por ser ese “*marcador de aptitud*” que permite al estudiante disfrutar y analizar la literatura. Se ignora, pues, el hecho de que el aprendiente de árabe puede llegar a apreciar el texto literario sin tener que, necesariamente, ser capaz de analizar el estilo y de dominar completamente la lengua árabe, al tiempo que se ignoran los diversos beneficios lingüísticos, motivacionales y psicológicos.

3. *Conclusión*

Hemos podido comprobar que la disciplina de la EALE, tanto en España como en EEUU, no tiene aún una visión común sobre el lugar que debe ocupar el texto literario en el currículo. Aunque existen docentes y académicos que llaman a una mayor utilización de la literatura, la mayoría continúa aceptando el uso de textos periodísticos y expositivos como material de lectura esencial en el currículo de ALE, desde los niveles iniciales a los avanzados. Esporádicamente, se denuncia la escasez y poca variedad de materiales didácticos (Martí Tormo 2012:47 y 58-59, Veerstegh 2006), pero no se llega a elaborar un discurso claro y favorable a la integración del texto literario como fuente de conocimiento lingüístico y cultural. Por el contrario, la mentalidad que subyace al discurso pedagógico actual (la dificultad del discurso literario, la primacía del método comunicativo oral, etc.) contribuye a marginar al texto literario casi de forma automática.

La escasez de publicaciones referentes a la pedagogía del ALE es señal, en mi opinión, de que el profesorado no se ha involucrado aún en una reflexión profunda sobre la adecuación de los métodos didácticos actuales en la formación de estudiantes universitarios, ni sobre el lugar que puede y debe ocupar la literatura en la didáctica del ALE. Dicha afirmación no es óbice para reconocer la existencia, en el pasado y en la actualidad, de docentes que han incorporado el texto literario como medio para promover diferentes habilidades lingüísticas y culturales. De lo que se carece, pues, es de la literatura científica que describa esas experiencias. Para suplir esa carencia es preciso tender puentes entre la lingüística (teórica, aplicada, y de corpus) y la didáctica del ALE.

En el siguiente capítulo vuelvo a retomar la cuestión de la lectura extensiva para analizar la adquisición de vocabulario a partir de dicha práctica y procedo, después, a hacer una propuesta pedagógica que pueda implementarse en el campo de la EALE para promover la lectura extensiva de textos literarios.

CAPÍTULO III: VOCABULARIO Y LECTURA EXTENSIVA

La comprensión lectora, como destreza general, se basa en dos elementos básicos: el léxico y la capacidad de raciocinio (Bernhardt y Kamil 2010:182). No sorprende, pues, que en el aprendizaje de lenguas extranjeras el vocabulario sea la piedra angular del desarrollo de la lectura y de las habilidades lingüísticas en su conjunto (Lewis 1997:19-21, Lee y Hirsh 2012:80, Gass 2013:194). En el contexto particular de la EALE, los especialistas señalan que el conocimiento del léxico es esencial en el aprendizaje y enseñanza del árabe y un elemento imprescindible para la comunicación exitosa en todas las destrezas lingüísticas (Gamal 1998:55, al-Batal 2006:332, Ryding 2013:197); incluso se ha afirmado que el vocabulario constituye el elemento central para la comprensión lectora y no, como tradicionalmente se creía y ponía en práctica, el conocimiento gramatical (Khaldieh 2001:417).

El reto radica en cómo puede el aprendiente adquirir la cantidad de vocabulario necesaria para leer de forma fluida en la L2, teniendo en cuenta el enorme tiempo y esfuerzo que ello requiere. En el caso particular del ALE, me atrevería a afirmar que el reto más importante para los aprendientes cuya L1 no sea semítica y que utilice grafía no árabe es el aprendizaje del léxico, entre otras razones, por la ausencia de signos diacríticos, por la práctica inexistencia de cognados y porque los campos semánticos y las implicaciones culturales asociadas a las palabras no siempre se corresponden con la L1 de los aprendientes.

A continuación, presento algunos conceptos esenciales relacionados con la adquisición del vocabulario en el contexto de la lectura y algunas recomendaciones pedagógicas acordes con la literatura científica.

A. LÉXICO MENTAL Y LECTURA

Los estudios en torno al aprendizaje y el conocimiento de vocabulario en L1 están muy avanzados en inglés. Por ejemplo, se sabe que los hablantes de inglés adquieren el léxico mental a un ritmo aproximado de 1000 familias de palabras⁵² por año (Nation 2013:211, Schmitt 2010:6) y que el adulto educado conoce un promedio de 20.000 familias de palabras (Nation y Waring 1997:7-8, Nation 2014:7)⁵³ o lo que equivale a 70.000 palabras (Matsuoka 2012:153), salvando las obvias diferencias individuales. Para hacerse una idea del alcance de dichos datos, conviene señalar que en inglés existen en la actualidad unas 100.000 voces en uso activo (Grabe 2009:268), lo que equivale a 54.000 familias de palabras (Nation y Waring 1997:7).

La literatura sobre lectura extensiva en L1 y L2 ha demostrado que a mayor conocimiento léxico mayor fluidez en la comprensión del texto (Perfetti y Harts 2002 citado en Nation 2013:350, Grabe 2009:265-266, Nagy 2010:72, Milton 2013:71), y cuanta más práctica de la lectura más efectiva es la adquisición de vocabulario (Nagy 2010, Nation 2014, 2015). Paul Nation afirma con rotundidad que la lectura extensiva requiere un amplio dominio de vocabulario para que la atención del lector pueda desplazarse al significado del texto (Nation 2009:50); de lo contrario, el lector centra su atención en los aspectos lingüísticos (gráficos, semánticos, gramaticales y sintácticos) del texto.

Para leer de manera fluida y para que pueda producirse una comprensión aceptable del texto, este debe contener, como mínimo, un 95% de *input* comprensible (Laufer 1991, 1992, Nation 2009:51, Milton 2013:66); aunque para considerar que un texto

⁵²Según Nation (2013:11, 2014:3-4) las familias de palabras están formadas por palabras *base* que engloban sus posibles realizaciones como los plurales, las conjugaciones verbales y algunas palabras derivadas formadas a partir de prefijos o sufijos que mantienen la transparencia semántica (Nation y Waring 1997:7). Por poner un ejemplo proveniente del inglés, la palabra *base* 'accident' incluiría en su *familia* palabras como 'accidental', 'accidents' y 'accidentally', pues el aprendiente puede comprenderlas por relación gramatical y semántica, y no porque las memorice por separado (Bauer y Nation 1993:253). En lenguas como el inglés, algunos investigadores consideran que la forma más adecuada de contabilizar el léxico es por *familias de palabras* (Nation 2014:4), sobre todo para fines pedagógicos. El concepto es interesante porque tiene muchas implicaciones para el árabe, aunque también es problemático por la dificultad que supone determinar el número de palabras que constituye cada una de las *familias* (Bauer y Nation 1993). Para los objetivos de la presente investigación utilizaré principalmente el concepto de *lema* para referirme al vocabulario, por ser más práctico a efectos cuantitativos en la lingüística de corpus.

⁵³Esta figura excluye nombres propios, palabras compuestas, abreviaciones y extranjerismos (Nation y Waring 1997:8).

está al nivel del lector, dicho texto no debe contener más de un 1% o 2% de vocabulario desconocido (Nation 2006, 2014:7, Grabe 2009:311). Dicho de otro modo, cuando un texto contiene solo ese pequeño porcentaje de palabras nuevas, el lector es capaz de comprenderlo sin mayor dificultad, pues la cantidad de léxico desconocido puede ignorarse o llegar a entenderse por contexto. Sin embargo, cuando el porcentaje de vocabulario desconocido alcanza un 3-5%, el lector necesita recurrir a cierta ayuda externa, como el uso del diccionario, para comprender el texto. A partir de un 6% de vocabulario desconocido la comprensión del texto es inadecuada y, en la enseñanza de lenguas extranjeras, la utilización de este tipo de textos entorpece el proceso de aprendizaje (Nation 2013:207-209).

Teniendo en cuenta estos datos, parece oportuno preguntarse: ¿Cuál es la cantidad de léxico que necesitan los aprendientes de lenguas extranjeras para leer con fluidez? Según Milton (2013:65), el aprendiente de inglés como L2 conoce como promedio las siguientes cantidades de lemas en cada nivel del *Marco*:

Niveles del <i>Marco</i>	Lemas
A1	<1.500
A2	1.500-2.500
B1	2.500-3.250
B2	3.250-3.750
C1	3.750-4.500
C2	4.500-5.000

Tabla 1: Promedio de lemas adquiridos por estudiantes de inglés L2 en Grecia

Estos datos fueron obtenidos tras examinar a 10.000 estudiantes de inglés (L2) en Grecia utilizando instrumentos de evaluación internacionales y deben interpretarse tomando como referencia la lista de 5.000 lemas más frecuentes en inglés. Es decir, esas cifras no representan necesariamente todas las palabras que conoce el aprendiente

en cada nivel, pues puede ocurrir que, por ejemplo, un estudiante de B1 conozca palabras que están en la franja del B2 o del C1. En cualquier caso, dichos datos son un buen indicador de lo que se puede esperar del estudiante en diferentes niveles y, además, miden el vocabulario que posee el aprendiente mejor que el concepto de "familias de palabras" utilizado por Nation.

A diferencia de lo que ocurre en inglés, no he logrado localizar investigaciones cuantitativas acerca del léxico mental que poseen los hablantes nativos de árabe y tampoco parece haber estudios longitudinales sobre la velocidad de adquisición del vocabulario y el caudal léxico.

En árabe, los diccionarios son de las pocas fuentes que ofrecen alguna información respecto al vocabulario en uso activo. Así, por ejemplo, *Mu'jam al-luga al-'arabiyya al-mu'āsira*, cuya primera edición fue publicada por Dār al-Kitāb en 2008, incluye 32.300 entradas (‘Aṭīfī 2015:27), pertenecientes al léxico árabe contemporáneo. Otra fuente que ofrece datos acerca del léxico árabe son los estudios de lingüística de corpus, los cuales señalan que, en el lenguaje periodístico, existen actualmente unos 33.500 lemas en uso activo⁵⁴. Sin embargo, no se sabe con exactitud el número total de palabras en uso activo en otros contextos, como puede ser el lenguaje literario contemporáneo, objeto de la presente investigación.

En el contexto de la EALE, el único estudio que parece haberse ocupado de la adquisición de vocabulario es una tesis doctoral de Robert Ricks (2015). Hasta entonces, la única estimación que se encontraba en la literatura era que los aprendientes de árabe L2 de nivel avanzado (=B2 en el *Marco*) necesitan conocer un promedio de 3.000-3.500 palabras de uso frecuente (al-Batal 2006:333); una cifra prácticamente idéntica a la que ya barajaban los especialistas en enseñanza de inglés L2.

El estudio de Ricks, sin embargo, muestra una perspectiva algo diferente, pues ofrece una estimación de la cantidad de léxico –en árabe estándar moderno (Ricks 2015:87)- que conoce el aprendiente en las diferentes etapas de aprendizaje. Los datos fueron obtenidos a partir de la administración de 3 pruebas de vocabulario receptivo a 161 participantes, diseñadas en base a las recomendaciones de estudios previos realizados sobre otras lenguas (Ricks 2015:59-102). Las pruebas fueron administradas online, al igual que una encuesta para recoger información referente a los años de estudio de

⁵⁴Comunicación personal con Nizar Habash mediante correo electrónico fechado el 29 de marzo de 2016.

los sujetos participantes y a las instituciones académicas de las que provienen⁵⁵. El siguiente gráfico⁵⁶ ofrece datos acerca del número de lemas que conocen los estudiantes de árabe L2 tras diferentes años de estudio considerando los 5.000 lemas más frecuentes recogidos por Buckwalter y Parkinson en su diccionario de vocabulario frecuente (Ricks 2015:146):

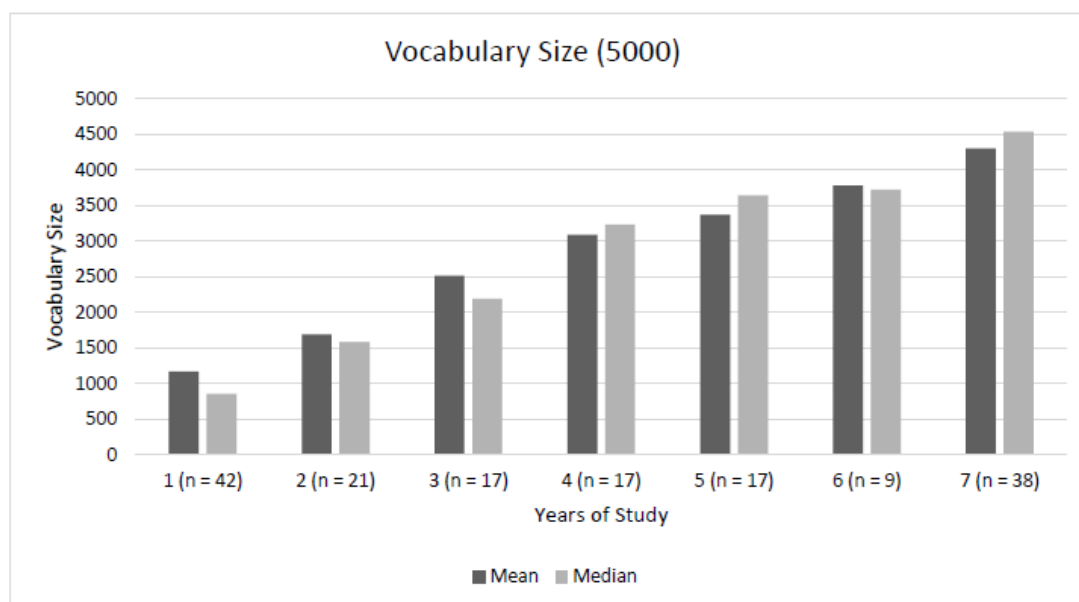


Figure 51: Vocabulary Size (5000 words) by Years of Study

Como conclusión a los datos obtenidos, Ricks estima, que el ritmo de adquisición de vocabulario en MSA es de 800-1000 palabras por año académico, una figura más alta de lo que se encuentra en aprendientes de otros idiomas y que puede deberse a que las personas que deciden estudiar árabe suelen ser personas con experiencia en el aprendizaje de varios idiomas extranjeros⁵⁷ (Ricks 2015:150-151). En cualquier caso, y a pesar de algunas limitaciones que presenta el estudio⁵⁸, si comparamos el gráfico

⁵⁵ Es importante señalar que las 161 poseen diferentes perfiles educativos y no son producto de un único programa académico, aunque 132 utilizaron *al-Kitāb*. Además, el 85% de los participantes ha estudiado algún dialecto y el 50% ha disfrutado de una estancia académica en algún país árabe (Ricks 2015: 114-116, 123). Aunque no se informa a qué instituciones académicas concretas pertenecen los estudiantes, sí conocemos que pertenecen a 20 instituciones diferentes, además de haber aprendientes que ya dejaron de estudiar. En comunicación personal por correo electrónico (9 de junio 2016), Ricks indica que la gran mayoría de dichos estudiantes han estudiado en instituciones académicas estadounidenses, y sólo un puñado en instituciones europeas.

⁵⁶ En el gráfico, *n* equivale al número de participantes.

⁵⁷ Estos datos deben interpretarse teniendo en cuenta los estudios mencionados por Ricks, referentes a estudiantes de lenguas extranjeras en diferentes contextos.

⁵⁸ Algunas de las limitaciones son, por ejemplo, que desconocemos la naturaleza de los programas académicos en los que los participantes de su estudio aprendieron árabe (horas de instrucción y estudio semanales), cuándo dejaron de estudiar los participantes, su edad, o la dificultad de testar el conocimiento

arriba presentado con las estimaciones de Milton (2013:65), mencionadas en las páginas anteriores, parece que hay una gran correspondencia entre la cantidad de léxico que se adquiere en inglés L2 y en árabe L2 en los diferentes niveles de competencia lingüística.

Es importante recordar que el número de participantes en el estudio de Ricks no puede considerarse representativo de toda la población de estudiantes de árabe, tal y como el mismo autor admite (Ricks 2015:123), aunque sí ofrece unas estimaciones que, a mi juicio y por mi experiencia profesional, se aproximan a la realidad. De hecho, por mi experiencia docente puedo afirmar que los estudiantes de árabe que cursan estudios en instituciones académicas estadounidenses suelen alcanzar el B1/B2 de promedio en el cuarto año de carrera al menos en el caso de programas de lengua intensivos (5-6 horas de contacto semanales dentro del aula), lo cual coincide con los datos existentes para el inglés L2 (Milton 2013).

Datos de esta naturaleza podrían guiar el desarrollo de materiales de lectura extensiva en árabe L2, puesto que la cantidad de vocabulario conocida por el aprendiente en los diferentes niveles de competencia lingüística influye de forma directa en su capacidad de leer y comprender textos de forma adecuada.

B. CONOCER EL VOCABULARIO

Aunque la expresión “conocer vocabulario” implica el conocimiento de diversas facetas lingüísticas del léxico (ortografía, realización fonética, significado, morfología, restricciones sintácticas, concordancias, etc.) y grados de conocimiento (Schmitt 2010:15-18, Zhong 2012:26-29, Gass 2013:196-198, 212-213, Milton 2013:74-75), el factor que más me interesa en el contexto de la presente investigación es la capacidad del lector de decodificar el vocabulario escrito y de acceder a su significado de forma rápida y eficaz para poder leer de forma fluida.

léxico de personas que conozcan las 5.000 palabras más frecuentes de ByP. La importancia de esta información es que se trata de factores que pueden afectar a la adquisición del vocabulario. En cualquier caso, los datos obtenidos por Ricks concuerdan con mi experiencia en la enseñanza de árabe en instituciones estadounidenses y representan un paso importante en los estudios léxicos del árabe L2.

Para el especialista británico James Milton, el conocimiento y reconocimiento ortográfico es fundamental para acceder al significado del léxico impreso y, por tanto, para formar a lectores hábiles (Milton 2013:74). De hecho, Perfetti define “vocabulario” como la capacidad de “*conocer el significado de una palabra*” (Perfetti 2010:292), para lo cual es necesario, como primer paso, decodificarlo a nivel visual, al menos en el caso de la lectura.

En árabe, “conocer vocabulario” abarca igualmente diversas facetas del léxico. Sin embargo, es importante resaltar que en el contexto de la lectura existen dos factores esenciales en la construcción de un léxico mental robusto: el reconocimiento automático visual del vocabulario y el conocimiento morfológico.

Este último aspecto ha sido destacado por varios expertos (Gamal 1998:78-79, Khaldieh 1998, 2001:418, Al-Batal 2006:338, Brustad 2006:342, Ryding 2013:152) que abogan por la enseñanza temprana del sistema árabe de raíces y formas para que el aprendiente pueda utilizarlo de forma efectiva en el aprendizaje de vocablos nuevos. De hecho, si se tienen en cuenta los estudios realizados sobre niños marroquíes para investigar la adquisición del léxico durante la infancia (Badry 2004) y la teoría de la *Descomposición Morfológica Obligatoria* desarrollada por Sami Boudelaa⁵⁹, es razonable pensar que recurrir a la raíz y la forma de las palabras es una técnica particularmente útil para los aprendientes de árabe (L2)⁶⁰.

C. AUTOMATICIDAD Y VELOCIDAD LECTORA

Como acabo de apuntar, un concepto muy relacionado con la fluidez lectora y la capacidad óptima de comprensión léxica es la *automaticidad*. En la lectura de L1 se habla de la importancia de poder reconocer rápidamente las palabras impresas para poder activar el significado de las mismas y desarrollar la necesaria automaticidad que

⁵⁹ Véase cap. 1, apart. C.4..

⁶⁰En un estudio experimental no publicado, Familiar (2008) concluyó que, en el contexto de la lectura, el conocimiento de raíces árabes asociadas a vocabulario adquirido facilita el reconocimiento y comprensión de palabras nuevas que comparten la misma raíz (estudio accesible en https://www.academia.edu/7042882/The_Role_of_the_Arabic_Root_in_Guessing_the_Meaning_of_Unfamiliar_Nominals_-_M.A._Thesis_2008).

hace posible una lectura fluida (Samuels 1979, Grabe 1988:59). Según Perfetti, “*decoding a word whose meaning is known strengthens the connection between the word’s orthographic form (its spelling) and its meaning*” (Perfetti 2010:291).

Para los defensores del *component skills approach* -como Koda (1992) y Haynes y Carr (1990)- la automatización de las subdestrezas más básicas involucradas en la decodificación del léxico impreso es esencial para que otras destrezas de orden mayor se ocupen de la comprensión del texto escrito (Hudson 2007:94, Grabe 2009:21). Esto significa que cuanto antes automatice el aprendiente de lenguas extranjeras el reconocimiento visual del vocabulario más frecuente, antes podrá leer de forma fluida. De hecho, cuando, durante el proceso de lectura, el aprendiente se encuentra con léxico que le es desconocido, dedica más tiempo a la decodificación visual de dicho vocabulario (Carver 1992:93) que al significado del texto. Por tanto, si un libro está repleto de vocablos nuevos, el aprendiente se verá forzado a ralentizar el ritmo al que lee.

Un lector eficaz es, pues, aquel que decodifica automáticamente el léxico impreso a nivel visual y ortográfico, lo cual le permite acceder al significado del mismo (Grabe 1988:60, Hudson 2007:96), sin tener que depender del contexto o de su *schema*, o conocimiento del mundo, personal para comprenderlo. En otras palabras, un procesamiento visual correcto es esencial para una comprensión lectora adecuada.

Al mismo tiempo, la velocidad a la que leemos un texto es fundamental para determinar el grado de habilidad que poseemos cuando acometemos la lectura. A mayor velocidad, asumimos que el lector lee con facilidad, es decir, sin aparente complicación. La literatura especializada lo denomina *fluidez* en la lectura.

En el caso de los hablantes de inglés como L1, se sabe que los adultos leen a una velocidad de comprensión óptima de 250-300 palabras por minuto (Grabe 2009:14, Nation 2009:31). Es más, los estudios realizados sobre el inglés como L1 desde hace más de un siglo han alcanzado tal nivel que existen datos empíricos que demuestran que la velocidad lectora es altamente predictiva; es decir, se conoce con exactitud a qué velocidad se lee a determinadas edades de escolarización, en diferentes situaciones, dependiendo incluso del tipo de texto y del motivo por el que se lee (Carver 1992).

En el contexto del inglés como L2, la difusión de estimaciones acerca de la velocidad a la que leen los estudiantes en diferentes niveles es escasa. El reconocido experto estadounidense Neil J. Anderson recomienda alcanzar una velocidad mínima de 180-200 palabras por minuto para poder interactuar con el texto impreso de forma eficaz; aunque, para poseer las destrezas de un hablante nativo educado, el aprendiente necesita poder leer a una velocidad superior (Anderson 1999:3). Para alcanzar tal objetivo, la lectura extensiva ha demostrado ser un instrumento muy útil, según indican varios estudios (Iwahori 2008:75).

En resumen, la velocidad lectora es un factor determinante en la comprensión exitosa, pues para leer de forma fluida el lector debe haber automatizado el reconocimiento visual de una gran cantidad de léxico. A ese tipo de vocabulario se le denomina en la literatura científica *sight vocabulary* y, según los investigadores Renandya y Jacobs (2002:299), el concepto “*refers to words that readers can recognize quickly and effortlessly*”. Dicha idea tiene tal importancia que los investigadores señalan a continuación: “*This rapid and automatic process of word recognition is extremely crucial for reading. If this ability is lacking, subsequent reading processes are likely to be seriously impeded, which in turn makes comprehension difficult, if not impossible*” (Renandya y Jacobs 2002:299).

En el campo de la EALE, la trascendencia de desarrollar automaticidad en el reconocimiento del léxico impreso ha sido mencionada en la literatura científica por muy pocos investigadores y, casi siempre, se ha hecho para hablar del sistema de escritura como causa principal de las dificultades que experimenta el aprendiente a la hora de decodificar y reconocer el vocabulario escrito (Khaldieh 2001, Hansen 2010, Ryding 2013). De hecho, Abu-Rabia (1998:107, 110) rechazó para el árabe la centralidad del reconocimiento automático de vocablos individuales fundamentado solo en el reconocimiento del esqueleto del léxico, e incidió en la importancia de la automaticidad *holística* que, en su opinión, requiere el procesamiento de los signos diacríticos como parte esencial del léxico. Bien diferente era la opinión de Khaldieh (2001), quien se refirió de forma explícita a la importancia de perfeccionar las destrezas lectoras ascendentes para alcanzar una competencia lectora óptima. Siguiendo la opinión de investigadores anglosajones, Khaldieh (2001:418, 426) afirma que desarrollar esas destrezas requiere cierto entrenamiento, aunque no especificó cómo llevarlo a cabo.

En el ámbito pedagógico, la primera referencia al papel que desempeña el material didáctico en el desarrollo de la automaticidad lectora se encuentra en Gunna Hansen (2010). La investigadora danesa toma como ejemplo el libro de texto *al-Kitāb* para criticar la ausencia de ejercicios enfocados a la adquisición de la destreza lectora. De hecho, los textos escritos incluidos en una lección típica de *al-Kitāb* (Parte 1) contiene un 90% de vocabulario desconocido (Hansen 2010: 568), una característica que no permite al aprendiente desarrollar el *sight vocabulary* necesario para desarrollar el reconocimiento automático del léxico y, en consecuencia, una lectura fluida. Mi propia experiencia docente confirma que los estudiantes de árabe experimentan mucha dificultad en la lectura de los textos incluidos en *al-Kitāb* y Hansen es consciente de que el problema no se limita a este libro de texto particular, aun siendo el más utilizado, sino a todos los métodos dedicados a la enseñanza de ALE. De hecho, la investigadora concluye su artículo afirmando lo siguiente:

“The most widely used methods for teaching Arabic communicatively may lack emphasis on automaticity training as a crucial subskill for reading, as most reading practice—beyond the propedeutic course within the first semester of study—often focuses on training of top-down reading strategies like skimming and guessing from context” (Hansen 2010:579).

La convicción de Hansen de que el reconocimiento automático del léxico es una destreza esencial para desarrollar una competencia lectora sólida (Hansen 2014:68) la había llevado ya a publicar una serie de textos sencillos (Hansen y Fettah 2007) para ayudar a los estudiantes de ALE a desarrollar una fluidez lectora a partir de textos que reutilizan de forma reiterada el léxico que aparece en las listas de vocabulario presentes en *al-Kitāb*. En la introducción a dicha obra, la investigadora señala que el estudiante de lenguas extranjeras *“must be exposed to copious amounts of written material to be able to build sufficient decoding competency, so as to acquire a reasonable reading speed”* (Hansen y Fettah 2007:4) y recalca la importancia de reconocer visualmente las palabras de forma automática para poder desarrollar dicha velocidad lectora; en especial, si tenemos en cuenta que los libros de texto de corte comunicativo disponibles en el mercado tienden a utilizar textos auténticos (Hansen y Fettah 2007:5-8) y no textos diseñados para los aprendientes de ALE.

En lo que respecta a la velocidad lectora, Ryding (2013:197) reconoce que el estudiante de árabe solo puede alcanzar una adecuada velocidad si se adquiere previamente una cantidad sólida de vocabulario. En este ámbito, parece haber indicios de que leer textos árabes no desambiguados es más rápido que leer textos desambiguados (Hansen 2010:576), hecho que está en consonancia con la estrategia lectora que desarrollan los hablantes nativos de árabe (Hermena 2015). Se desconoce, sin embargo, la velocidad a la que leen los hablantes nativos de árabe educados de promedio. La única referencia en la literatura científica aparece en un estudio realizado por Hansen (2010:577), donde se concluye que los aprendientes de árabe cuya L1 se basa en letras romanas y que han alcanzado el nivel 3 (ILR) (=C1 en el *Marco*), leen tres veces más lento que los hablantes nativos, aunque no se indica cuántas palabras por minuto pueden leer, quizá porque no se hayan utilizado técnicas propias del seguimiento ocular⁶¹.

D. ADQUIRIR VOCABULARIO DE FORMA INCIDENTAL

Uno de los beneficios que se citan con frecuencia cuando se habla de lectura extensiva es que el aprendiente adquiere léxico nuevo a través de la práctica lectora. Este tipo de aprendizaje se denomina *incidental*, puesto que el objetivo primordial de la lectura extensiva es la comprensión del texto y no el aprendizaje de vocabulario. El investigador británico Norbert Schmitt, especializado en estudios léxicos, ofrece una definición que me parece relevante y apropiada en el contexto de la presente investigación: “*Incidental learning is learning which accrues as a by-product of language usage, without the intended purpose of learning a particular linguistic feature. An example is any vocabulary learned while reading a novel simply for pleasure, with no stated goal of learning new lexical items*” (Schmitt 2010:30).

⁶¹ En su estudio, Hansen da a leer a los participantes varios textos escritos breves (124-192 palabras), tanto vocalizados como no vocalizados, y cronometra la velocidad de lectura en voz alta (grabada en laboratorio de idiomas), un elemento que resulta problemático a la hora de estudiar la lectura pues los factores que intervienen en la lectura silenciosa son diferentes a los de la lectura en voz alta. Sin embargo, los resultados obtenidos por Hansen indican claramente que los hablantes nativos leen mucho más rápido que los estudiantes de árabe de nivel 3 (150 segundos aproximadamente los hablantes nativos vs. 420 segundos aprox. los estudiantes avanzados, Hansen 2010:576).

Aunque la adquisición de vocabulario de forma incidental, a través de la lectura, se ha puesto en duda y se ha minimizado por algunos investigadores (Lee y Hirsh 2012:81), lo cierto es que el estudiante puede llegar a aprender una gran cantidad de vocabulario si practica la lectura extensiva de textos que están a su nivel (Grabe 2009:272-274, Nation 2013:218). Como ya se ha dicho, esto significa que los textos utilizados deben contener un 95%, como mínimo, de vocabulario conocido, para que la experiencia de la lectura extensiva sea placentera al tiempo que beneficiosa⁶² (Nation 2013:219).

Otro de los requisitos indispensables para que pueda adquirirse vocabulario a través de la lectura extensiva es la repetición de léxico. Existe una amplia evidencia científica (Schmitt 2010:30-31, Matsuoka 2012:158-159, Gass 2013:209-211, Nation 2014:2-3) que muestra la correlación positiva entre la visualización reiterada del vocabulario y el desarrollo del conocimiento léxico. Aunque algunos investigadores afirman que el léxico ha de aparecer en el texto un mínimo de 10-12 veces (Webb 2007, Matsuoka 2012:158, Nation 2014:3)⁶³, también hay estudios que aportan datos que demuestran la existencia de otros factores que influyen en la retención del vocabulario durante la lectura, como puede ser el contexto o las similitudes entre la L1 y la L2 (Webb 2007). De hecho, en ocasiones, se pueden adquirir palabras nuevas que aparecen en el texto dos o tres veces (Gass 2013: 209-211, Nation 2014:2-3, Reynolds 2015), e incluso se puede adquirir un conocimiento parcial de dicho vocabulario (Webb 2007:60-61). En cualquier caso, cuanto más se repita el léxico en un mismo texto, más oportunidades y, por tanto, mayor facilidad tiene el estudiante de adquirir dicho vocabulario (Webb 2007).

En el campo de la EALE al concepto de vocabulario incidental ha sido mencionado de forma explícita pocas veces. Al-Batal (2006), por ejemplo, dota al término de ciertas connotaciones negativas en contraposición a la enseñanza explícita de léxico a través de listas de vocabulario y de actividades de corte oral, bajo el pretexto de que

⁶² Esto no significa que el estudiante vaya a aprender el 5% del vocabulario desconocido presente en el texto, sino que la experiencia de la lectura será posible y fluida. La cantidad de léxico que aprende el estudiante a partir de este tipo de lectura depende de múltiples factores, principalmente la repetición del léxico en el texto (como explico a continuación), y el número de horas y textos que lee el aprendiente en un año académico (Nation 2013:211-216).

⁶³ Dichos estudios se refieren a la adquisición de lemas desconocidos, e incluso pseudo-lemas en el curso de la lectura de textos breves, como pueden ser cuentos cortos, ya sean originales o diseñados específicamente para el estudio.

dicha teoría es aplicable al contexto de la lectura en los niveles Avanzado y Superior según ACTFL (=B1/C1 en *el Marco*) y porque asume que el aprendiente de árabe tarda más en desarrollar las destrezas lectoras en comparación con otros idiomas (al-Batal 2006:332). Es más, el aprendizaje incidental de vocabulario no se entiende en el contexto de la lectura extensiva de textos diseñados al nivel del estudiante, sino que se percibe como consecuencia de practicar estrategias para la comprensión del vocabulario por contexto (al-Batal 2006:337). Para al-Batal, el estudiante de árabe no necesita conocer todo el vocabulario que aparece en los textos de lectura, sino que, por el contrario, deben aprender a ignorar el léxico desconocido:

“Students should be encouraged to read the text globally first, skipping unknown words, and make preliminary assumptions about the topic of the text. In the second reading, they should be encouraged to guess the meaning of new words from context. The use of the dictionary should not be permitted until the third reading, after students have identified the key words essential to comprehension. The crucial principle here for both teachers and students is that not every word in the text needs to be actively controlled to reach comprehension at the level determined by the teacher” (al-Batal 2006:337).

La idea la repetiría Brustad (2006) en el mismo monográfico. Sin embargo, la problemática de dicha propuesta pedagógica es que no considera el porcentaje de vocabulario desconocido presente en el texto y, por tanto, ignora las dificultades que experimenta el estudiante de ALE durante la lectura de textos que están por encima de su nivel y las posibles consecuencias negativas de dicha experiencia.

A pesar de dicha posición, al-Batal (2006:336), se adscribe, al igual que otros muchos investigadores (Hansen y Fettah 2007, Hansen 2010:568, y Ryding 2013:151-152) a la idea de que la repetición del vocabulario es clave para una rápida adquisición del mismo y para alcanzar el reconocimiento automático que precede a la lectura fluida.

E. CONCLUSIÓN

En el contexto de la EALE es importante tener en cuenta que la primera vez que la expresión “lectura extensiva” fue utilizada tal y como se entiende hoy fue en un artículo de Elsaid Badawi (2002)⁶⁴, aunque lo hizo tres décadas después de implementar lo que, en su momento, fue un pionero programa de lectura. El objetivo primordial de Badawi era que el estudiante mejorara su dominio del léxico, las expresiones idiomáticas y las estructuras lingüísticas en general. Con anterioridad, la expresión “lectura extensiva” no implicaba la lectura sostenida de una gran cantidad de textos a lo largo de cierto periodo de tiempo; sino, más bien, la variación en el tipo de textos elegidos (Bellamy et al. 1963: v, Brinner y Khouri 1971: vii), aunque el objetivo último fuera, tal y como se explicaba, expandir el vocabulario, las expresiones idiomáticas y los recursos estilísticos del estudiante.

También es importante recordar que, tras la aparición de los métodos comunicativos, la EALE comenzó a mostrar predilección por la lectura intensiva. Se favorecía, por ejemplo, la práctica del escaneado, el *skimming*, y la inferencia, especialmente en los niveles iniciales e intermedios (Brustad 2006), en contraposición con la lectura extensiva, que apenas aparece mencionada. Aun cuando la *lectura extensiva* aparece tratada en la literatura científica, el término no se utiliza en el sentido tradicional de la acepción, sino que se incluye como estrategia pedagógica recomendada para el nivel Avanzado (= B1/B2 en el *Marco*) (Brustad 2006) o se muestra como una práctica de lectura más, sin pautas específicas que ayuden a implementarla (Ryding 2013:147-148). En definitiva, la lectura extensiva independiente aparece como una práctica propia de los niveles avanzados y superiores, a diferencia de lo que ocurre en la enseñanza de lenguas europeas.

En cuanto al vocabulario, los especialistas en EALE coinciden en que se trata de un elemento clave en el desarrollo de la competencia lectora (Gamal 1998, al-Batal 2006, Ryding 2013), aunque pocas veces se hace referencia a la forma en que debe integrarse en la didáctica de la lectura. Se menciona de forma somera que la aparición

⁶⁴ Anteriormente, el concepto había aparecido en un artículo del investigador egipcio-estadounidense Adel Gamal (1998) donde ofrecía una panorámica sobre la lectura extensiva en inglés como L2 y el papel de dicha práctica en la adquisición de vocabulario, sin mayores implicaciones para la EALE.

reiterada del vocabulario es importante para la adquisición del léxico, pero no se presentan propuestas pedagógicas concretas. De hecho, una idea que parece estar extendida entre los docentes de ALE es que no es necesario que el estudiante conozca todas las palabras de un texto para entenderlo. Ya lo decía Brustad (2006:344): *“It is possible to understand much of a text without knowing all the words in it”*. El problema de afirmaciones tan genéricas es que no aclara cuál es el porcentaje de palabras desconocidas admisibles para que pueda tener lugar una comprensión adecuada. De hecho, la investigadora señala que los aprendientes de árabe a menudo *“report difficulty in skipping over unfamiliar words and passages”* y añade que la solución se encuentra en la insistencia del profesor en la necesidad de saltarse esas partes difíciles del texto (Brustad 2006:350), práctica que no puede realizarse si el porcentaje de palabras desconocidas es alto, pues la experiencia sería frustrante para el estudiante. La excepción a este discurso es el trabajo de Hansen (2007, 2010, 2014), quien inicia una línea de investigación centrada en el reconocimiento visual automático del léxico como base sobre la que construir una práctica sólida de la lectura fluida.

A continuación, repasaré la literatura científica que se ocupa de las lecturas graduadas como vía para el desarrollo de la lectura extensiva fluida y, después, propondré una ruta similar para la EALE basada en el uso de textos literarios simplificados.

CAPÍTULO IV: LAS LECTURAS GRADUADAS COMO VÍA PEDAGÓGICA

La predilección por un modelo de lectura u otro, el papel otorgado al vocabulario en la lectura, el tipo de actividades y ejercicios utilizados en el aula y nuestra opinión como docentes respecto a todo lo anterior son cuestiones que tienen implicaciones directas en nuestra práctica pedagógica dentro y fuera del aula, y también en la planificación curricular y el diseño de materiales didácticos.

Si creemos que el vocabulario es el elemento central en la comprensión lectora, el currículo de ALE deberá estar diseñado para ayudar al aprendiente a adquirir la mayor cantidad de léxico posible en un plazo de tiempo razonable. Al tratarse de la destreza lectora es aconsejable tener en cuenta los datos empíricos obtenidos a partir de investigaciones en otras lenguas en relación con las prácticas pedagógicas que facilitan la adquisición de vocabulario. A continuación hago un breve repaso al lugar que ocupan las lecturas graduadas en la enseñanza de lenguas extranjeras y las evidencias a favor de su incorporación al currículo.

A. ¿QUÉ SON LAS LECTURAS GRADUADAS?

Una manera de promover la lectura extensiva y la adquisición de vocabulario de forma incidental es a través de las lecturas graduadas (Nation y Waring 1997:11, Schmitt 2010:32). Se trata de libros escritos o adaptados (Hill 2008:185) para el aprendiente de lenguas extranjeras utilizando un vocabulario y unas estructuras gramaticales simplificadas (Nation y Ming-tzu 1999:356, Horst 2009:44) basadas en la frecuencia del léxico (Hill 2008:185). Estas obras pueden pertenecer a diferentes géneros (poesía, prosa, textos científicos, históricos, etc.) y tratar una variedad de temas.

Este tipo de lecturas se utiliza por lo general como material suplementario, pero también puede constituir el pilar central de un curso de lenguas extranjeras. Cuando las lecturas graduadas se leen durante semanas, meses e incluso cursos académicos, se habla de programas de lectura extensiva. En estos casos, se recomienda que el aprendiente tenga acceso a un gran número de libros para que pueda elegir libremente los textos que desea leer (Renandya and Jacobs 2002:296).

La ventaja de este material es que permite controlar y acotar el vocabulario al que se exponen los estudiantes; es decir, los textos pueden nivelarse y ajustarse a la competencia del lector siguiendo una serie de pautas lingüísticas bien definidas (Nation y Ming-tzu 1999:356). De la elaboración de dichas obras hablaré con más detalle en el apartado referente a la lingüística de corpus.

Las lecturas graduadas se caracterizan por utilizar un vocabulario delimitado que aparece de forma reiterada en el texto. Como ya he explicado, desde el punto de vista cognitivo, decodificar (leer) palabras cuyo significado conocemos fortalece la conexión existente entre la forma escrita de dicho vocablo y su significado (Perfetti 2010:291). Por tanto, exponer al estudiante a este tipo de textos significa ofrecerle la oportunidad de reforzar dichas conexiones visuales y semánticas que hacen posible la comprensión eficaz. Desde el punto de vista pedagógico, asignar lecturas graduadas como parte del currículo significa ofrecer al estudiante un vocabulario comprensible (*input*) que puede ser adquirido de forma eficaz, asumiendo que las obras hayan sido elaboradas de forma cuidadosa.

La enseñanza de lenguas europeas como L2 dispone desde hace mucho tiempo de lecturas graduadas a diferentes niveles lingüísticos. En inglés y en francés, por ejemplo, este tipo de material didáctico existe desde al menos los años 20 del siglo XX (Beardsley 1937:396, West 1927:3 y 1937:221), época en la que ya se reconocían las ventajas de ese material. En la actualidad, la variedad de lecturas graduadas disponible en el mercado es inmensa, y los títulos publicados se cuentan por miles (Hill 2013:85).

B. TEXTOS GRADUADOS VERSUS TEXTOS AUTÉNTICOS

Como bien es sabido, en la enseñanza de lenguas extranjeras se utilizan textos auténticos y textos simplificados. A continuación, ofrezco la definición de ambos términos según el especialista en desarrollo de materiales Brian Tomlinson:

- Texto auténtico: “*A text which is not written or spoken for language-teaching purposes. A newspaper article, a rock song, a novel, a radio interview, instructions on how to play a game and a traditional fairy story are examples of authentic texts. A story written to exemplify the use of reported speech, a dialogue scripted to exemplify ways of inviting and a linguistically simplified version of a novel would not be authentic texts*” (Tomlinson 2011: ix).
- Texto simplificado: “*These are texts which have been made simpler so as to make it easier for learners to read them. The usual principles of simplification involve reduction in length of the text, shortening of sentences, omission or replacement of difficult words or structures, omission of qualifying clauses and omission of non-essential detail. Some applied linguists prefer to simplify texts by adding examples, by using repetition and paraphrase and by increasing redundant information. In other words, by lengthening rather than shortening the text, by elaboration rather than reduction*” (Tomlinson 2011: xvi).

Muchos docentes opinan que utilizar *textos auténticos* es imprescindible para enseñar lenguas extranjeras de forma eficaz, pues mostraría la riqueza natural del discurso lingüístico (Nuttall 1996:177) y motivaría a los estudiantes (Tomlinson 2003:5-6). Para este grupo de docentes, utilizar textos simplificados significa ofrecer al estudiante textos de valor lingüístico y cultural inferior, precisamente por ser simplificaciones de textos auténticos (Hill 2001:304).

Sin embargo, algunos conocidos investigadores han criticado el valor de dichas premisas, como el conocido especialista Widdowson (2003)⁶⁵ y Day y Bamford (1998:53-62), quienes argumentan en contra del “culto” a la autenticidad del texto. El

⁶⁵ Citado en Perez-Paredes (2010), *Defining Issues in English Language Teaching*. Oxford.

especialista británico David R. Hill (2013:89), fundador del *Edinburgh Project on Extensive Reading*⁶⁶, asimismo, critica duramente el “fervor religioso” que muestran algunos docentes en su insistencia de utilizar textos auténticos, pues no beneficia al estudiante: “*This insistence on authenticity does not help their students, who need input to be comprehensible and whose reading becomes arduous and slow if they encounter unfamiliar words in complex sentences*” Hill (2013:89). En realidad, esta idea tiene su origen en la aparición del método comunicativo (Day y Bamford 1998:54).

De hecho, los especialistas en la materia denuncian las desventajas pedagógicas de utilizar, muchas veces, textos originales. A continuación, cito a Day y Bamford (1998:55), quienes resumen magistralmente una idea que se ha venido expresando desde los años 80:

“Authentic texts, for all their virtues, can actually set back reading development. Williams refers to the paradox that the use of authentic text with learners often has an effect opposite to that intended; instead of helping the reader to read for the meaning of the message, and authentic text at too difficult a level of language forces the reader to focus on the code (1983, p.175). Wilga Rivers points out that “when average students encounter ungraded material too soon, they are usually forced back into deciphering with the aid of a dictionary, and valuable training in the reading skill is wasted” (1981, pp.37-38). Nuttall, in her discussion of authentic materials, concedes that “linguistically difficult texts are unlikely to be suitable for developing most reading skills” (1996, p.177). In addition, there is the affective toll. Rivers observes that “rushing students too soon into reading material beyond their present capacity for fluent comprehension with occasional contextual guessing [...] destroys confidence” (1981, p.260).

Más recientemente, Nation (2014:9) ha señalado que los textos originales no ofrecen condiciones favorables de lectura que faciliten el aprendizaje de vocabulario de forma incidental, en especial para aquellos aprendientes que conozcan menos de 9,000 familias de palabras; es decir, para la gran mayoría de los estudiantes.

En realidad, el concepto de *autenticidad* del texto y el debate existente acerca de si es preferible utilizar textos auténticos o simplificados existe en nuestro campo

⁶⁶ Puede consultarse en el siguiente enlace: <http://www.er-central.com/the-edinburgh-project-on-extensive-reading/about-eper/>

desde hace casi un siglo, época en la que ya había lecturas graduadas (Beardsley 1937).

C. EVIDENCIAS EMPÍRICAS A FAVOR DE LAS LECTURAS GRADUADAS

Como acabo de señalar, existen miles de lecturas graduadas en inglés, y posiblemente en otros idiomas europeos; y a pesar de esta abundancia, existen algunas críticas tradicionalmente dirigidas a este tipo de material didáctico. Se ha criticado la calidad del texto (Beardsley 1937, Hill 2008:185), la descentralización del rol del profesor en el proceso educativo (Day y Bamford 1998:47), la falta de tiempo en el aula para implementar la lectura extensiva debido a presiones administrativas para cubrir un currículo predeterminado (Renanday y Jacobs 2002:300, Stoller et al. 2013), la inautenticidad del texto graduado (Hill 2013:89), e, incluso, se ha percibido como una pérdida de tiempo, pues algunos docentes opinan que el tiempo dentro del aula podría ser invertido en actividades más provechosas (Renanday y Jacobs 2002:300). La reconocida investigadora en pedagogía de la lectura Fredricka Stoller tilda a tales críticos de “negativistas”⁶⁷. Aunque es cierto que hay lecturas graduadas de baja calidad literaria y pedagógica, existen también numerosos materiales bien escritos (Nation y Ming-tzu 1999:356). Según Hill (2008:185), descalificar así a todo un género de material didáctico de lenguas extranjeras es injusto y poco conveniente para los intereses del estudiante de lenguas extranjeras.

De hecho, existen evidencias empíricas que demuestran la eficacia de utilizar lecturas graduadas como vía para la adquisición del léxico y para el desarrollo de la velocidad lectora; en definitiva, para formar a lectores hábiles. Así, por ejemplo, se sabe que los textos tradicionalmente denominados “auténticos” contienen un vocabulario en su mayor parte inaccesible para el estudiante de lenguas extranjeras (Day y Bamford 1998:55). Por tanto, la simplificación de textos a través de la restricción del vocabulario

⁶⁷Término utilizado durante el *Third World Congress on Extensive Reading*, celebrado en Dubai del 18 al 20 de septiembre de 2015.

contribuye al aprendizaje y construcción del léxico (Nation 2013:223). En otras palabras, acotar y controlar el vocabulario en textos simplificados considerando la frecuencia de uso ha demostrado ser un procedimiento muy eficaz para reforzar y expandir el vocabulario del aprendiente (Nation 2013:247-248).

De igual modo, la lectura de textos graduados es eficaz en el desarrollo de la comprensión lectora (Horst 2009) y de la lectura fluida (Nation 2013:249). Por un lado, esto significa que este tipo de textos ofrece un remedio eficaz a la frustración que sienten muchos estudiantes cuando leen textos *auténticos* inaccesibles, que contienen gran cantidad de léxico desconocido (Hudson 2007:74). Por otro lado, leer textos simplificados es una forma muy eficaz de que los estudiantes adquieran mayor velocidad lectora (Bell 2001, Iwahori 2008) y también parece contribuir al progreso del reconocimiento automático del léxico escrito, lo que constituye un paso previo fundamental para un procesamiento sintáctico adecuado (Horst 2009).

Por consiguiente, y adoptando la opinión de Day y Bamford (1998:53), los textos de lectura más adecuados serán aquellos que sean comprensibles e interesantes, pero escritos a un nivel *i minus I*⁶⁸ de la competencia lingüística del aprendiente si el objetivo es convertir el texto de lectura en una oportunidad para beneficiarse a nivel lingüístico.

D. LA IDONEIDAD DE LA LITERATURA COMO TEXTO GRADUADO

En capítulos anteriores expuse los múltiples beneficios de leer textos literarios y de leer de forma extensiva en el contexto del aprendizaje de lenguas extranjeras (cap. II y III). Puede decirse que leer textos graduados tiene los mismos beneficios, pero también algunas diferentes. Por ejemplo (Simensen 1987:41):

⁶⁸ El concepto *i minus I* es justamente el contrario al que presentó Krashen en el año 1977 para explicar su Hipótesis de *Input i+I*, (siendo *i* el nivel actual del estudiante y el número *I* un nivel justo por encima), según el cual el alumno debe exponerse a textos un poco por encima de su nivel lingüístico para poder desarrollar una mayor competencia lingüística e ir progresando en su conocimiento de la L2. En esta Tesis no se contradice la validez de dicha opinión, sino que se argumenta a favor de un *input i minus I* en el caso particular de los textos de lectura extensiva.

- Ofrecen al docente una serie de material preparado que se adecua a las necesidades del estudiante en diferentes etapas.
- Promueve la independencia y autonomía del estudiante en el proceso de aprendizaje.
- Representan un puente entre el material didáctico y el texto de lectura auténtico.

Aunque las lecturas graduadas pueden pertenecer a diferentes géneros (literatura, historia, ciencia, etc.), los investigadores concuerdan en que el texto literario, y en especial la novela, es el que más ventajas ofrece en términos de repetición de léxico y de atractivo para el estudiante (Hirsh y Nation 1992:689, Ellis 2013:90). De hecho, tras un análisis de 4,000 lecturas graduadas de inglés L2, el especialista británico Hill, llega a la conclusión de que los textos pertenecientes al género de la prosa no literaria son los menos adecuados para la lectura extensiva y que, por eso mismo, existen pocas series de lecturas graduadas basadas en textos no literarios (Hill 2013:90).

E. UNA PROPUESTA PEDAGÓGICA PARA LA EALE

Teniendo en cuenta los beneficios del texto literario, de la lectura extensiva y del papel central que desempeña la repetición del vocabulario en el aprendizaje de lenguas extranjeras, sería conveniente desarrollar una serie de lecturas literarias simplificadas graduadas que puedan utilizarse en el contexto de la enseñanza y aprendizaje de ALE.

Aunque la idea ha comenzado a ser implementada en la EALE con la publicación de algunos materiales didácticos (Khorshid 2009a, 2009b, 2009c, 2011, 2013, Familiar 2013, 2016), aún no existen obras que hayan incorporado entre sus criterios de diseño la frecuencia léxica. Gamal (1998:74) ya hizo notar esa necesidad pedagógica y destacó la importancia de utilizar en los textos de lectura el vocabulario de uso frecuente. También la hizo notar Mughazy (2005), quien señaló que es preciso tener en cuenta la frecuencia léxica, morfológica y sintáctica de los textos auténticos destinados al hablante nativo de árabe al desarrollar materiales didácticos para la EALE. Esta idea

fue años después reiterada por Hansen (2014:58) en relación con los materiales para el desarrollo de la destreza lectora.

Teniendo en cuenta mi definición de “lectura extensiva”⁶⁹, me atrevería a afirmar que este tipo de material didáctico, si tiene en cuenta la frecuencia real del léxico en la narrativa árabe contemporánea, podría generar varios beneficios lingüísticos: 1) adquisición incidental de vocabulario, 2) reconocimiento automático del vocabulario, y 3) incremento de la velocidad lectora. En definitiva, este tipo de material permitirá aumentar la fluidez lectora.

A través de la lectura de estos textos, los estudiantes también pueden aprender nuevas estructuras gramaticales, adquirir un amplio conocimiento cultural y llegar a ser lectores independientes capaces de disfrutar de la lectura literaria.

Sin embargo, antes de desarrollar una oferta curricular basada en textos literarios, se necesita disponer de una lista de frecuencia léxica apropiada que facilite la creación de materiales de lectura atractivos y capaces de preparar al estudiante de ALE para leer de forma fluida. Estos materiales, además, servirían de puente para que el aprendiente pueda enfrentarse con éxito al texto literario una vez alcanzado el nivel B2, que es el nivel mínimo necesario para cursar estudios de posgrado en árabe en la universidad española⁷⁰. Para desarrollar materiales didácticos de calidad es imperativo recurrir a los instrumentos que nos proporciona la lingüística de corpus. En el siguiente capítulo explico en qué consiste la lingüística de corpus y cómo puede utilizarse para llevar a cabo mi anterior propuesta. Además, expondré que uno de los requisitos principales para elaborar lecturas graduadas en árabe es la extracción de una lista de vocabulario de uso frecuente en la narrativa árabe contemporánea, y explicaré cómo puede la lingüística de corpus contribuir a ello.

⁶⁹“*La práctica de leer fuera del aula, narrativa escrita y adaptada al nivel lingüístico de los aprendientes, para promover el placer de la lectura y desarrollar diversas destrezas lingüísticas y culturales*” (cap. I).

⁷⁰ Aunque la mayoría de los programas de posgrado requieren haber alcanzado un B2, el recién nacido Máster en Estudios Árabes Contemporáneos de la UAB requiere un B1 para matricularse (http://www.uab.cat/web/postgrado/master-en-estudios-arabes-contemporaneos/informacion-general-1206597472083.html/param1-3616_es/param2-2002/).

CAPÍTULO V: LINGÜÍSTICA DE CORPUS PARA EL DISEÑO DE MATERIALES DIDÁCTICOS

El establecimiento de puentes entre la docencia de lenguas extranjeras y la investigación en lingüística aplicada es indispensable para la producción de materiales didácticos de calidad y para el aprendizaje efectivo de lenguas extranjeras (Biber et al. 1998, Masuhara y Tomlinson 2010, Reinhardt 2010, Nation y Webb 2011, Grabe y Stoller 2013, Tomlinson 2013).

En el caso particular de la lingüística de corpus, los instrumentos de análisis disponibles son capaces de ofrecer datos empíricos descriptivos acerca del comportamiento de las lenguas (O’Keeffe et al. 2007, McEnery y Hardie 2015). Esto, a su vez, sirve para diseñar y desarrollar materiales didácticos que reflejen de forma fiel la realidad de la lengua (Carter et al. 1998, Sinclair 2004), especialmente si atienden a criterios de frecuencia léxica y gramatical (O’Keeffe 2007 et al.: xiii). Esto es posible porque la lingüística de corpus se vale de la lingüística computacional⁷¹ para obtener datos objetivos acerca del comportamiento del lenguaje, como, por ejemplo, las listas de frecuencia (Sinclair 1991, Flowerdew 2012:9).

Dado que la lingüística de corpus se ocupa, sobre todo, del estudio del vocabulario, es importante definir algunos términos esenciales relativos al ámbito léxico que aparecerán reiterados en las siguientes páginas:

- Corpus⁷²: Colección de textos electrónicos anotados, a partir de los cuales se extrae información lingüística.
- Token: Unidad gráfica u auditiva de lo que, por regla general, denominamos “palabra”. (Nation 2013:9). Por ejemplo, la oración *Me alegro de que me lo hayas dicho* contiene ocho tokens. En el caso del árabe, los tokens no tienen

⁷¹ La lingüística computacional es una “aproximación a la lingüística que utiliza técnicas matemáticas, a menudo con ayuda de un ordenador. Incluye el análisis de datos lingüísticos; p.ej. para establecer el orden de adquisición de ciertas reglas gramaticales por parte de los estudiantes o la frecuencia de aparición de un elemento en particular” (Richards et al. 1997:251)

⁷² El plural de *corpus* es *corpora*.

por qué aparecer como palabras gráficamente independientes. Así, el ejemplo *surirt bi-ma 'rifatik* contiene cuatro tokens, pues la oración contiene dos clíticos que son *bi* y el pronombre final *-k*.

- Tipo: Forma léxica única o caracteres únicos que distinguen una palabra de otra. En el mismo ejemplo anterior, la palabra *me* se repite dos veces; por tanto, decimos que la oración contiene siete formas o tipos diferentes. En el caso del árabe, la forma *mā* en la oración *mā ra 'ayt wa mā sami 't šay'* se considera forma léxica única, puesto que tiene un único significado: partícula negativa
- Lema: Forma léxica troncal de la que derivan otras formas derivadas pertenecientes a la misma categoría léxica. En español, por ejemplo, *libro* y *libros* pertenecen al mismo lema: *libro*. En árabe, sin embargo, los plurales irregulares no pertenecen al mismo lema que el singular; de hecho, las palabras *kitāb* y *kutub* se consideran lemas distintos. Un lema en árabe sería la raíz verbal *ka-taba* cuyas formas derivativas verbales forman parte de la misma raíz verbal.

A. LINGÜÍSTICA DE CORPUS Y LECTURAS GRADUADAS

1. Listas de frecuencia léxica

Una de las aplicaciones más frecuentes de la lingüística de corpus es la extracción de listas de frecuencia léxica⁷³ a partir de una colección de textos electrónicos, o corpus, que puede contener un gran número de palabras (McEnery y Hardie 2015:2). Una lista de frecuencia es una lista de palabras gráficas ordenadas de forma descendente según su repetición de uso en un determinado corpus, ya sea oral o escrito. Dichas listas se pueden extraer de forma rápida a través del uso de programas informáticos que incluso permiten comparar la reiteración del vocabulario en diferentes corpora electrónicos y, por tanto, en distintos contextos (O'Keeffe et al. 2007:11).

⁷³ También denominadas *listas unigramas* o *n-gram*, término utilizado en el ámbito de la lingüística computacional.

Tradicionalmente, las listas de frecuencia han servido a los lexicógrafos para la elaboración de diccionarios (Biber et al. 1998:21, O’Keeffe et al. 2007:17, Plá 2007:15). Matsuoka (2012:160) afirma que los estudios que se ocupan de las frecuencias léxicas se llevan a cabo, ante todo, por razones didácticas. De hecho, la influyente lista de vocabulario inglés *The Teacher’s Word Book* (Thorndike 1921) sirvió para elaborar una multitud de materiales pedagógicos destinados a diferentes públicos y fue, posteriormente, ampliada en *A Teacher’s Word Book of 30,000 Words* (1944), obra basada en un corpus de 18 millones de palabras (Matsuoka 2012:160-162)⁷⁴. Tras estas dos obras de Thorndike, y a medida que los instrumentos informáticos se iban desarrollando y refinando, se sucedieron otras aplicaciones lingüísticas, listas de frecuencia léxica y corpora que no interesan, por el momento, al ámbito de la presente investigación⁷⁵.

Aunque las listas de frecuencia léxica por sí solas ofrecen información lingüística limitada, no debemos olvidar que la repetición de uso de un vocablo determina la rapidez con que dicha palabra es aprendida (Schmitt 2010:63, Kilgariff et al. 2014, Gries y Ellis 2015) y decodificada con posterioridad (Hasher y Zacks 1984:1382-1383, Schmitt 2010:63), tanto en la lengua materna como en el aprendizaje de lenguas extranjeras. Es por esto que datos de dicha naturaleza pueden y deben guiar el diseño curricular de materiales de lectura en la enseñanza de lenguas extranjeras.

En este sentido, es importante distinguir entre dos conceptos esenciales muy relacionados entre sí, a saber: la frecuencia y la distribución del vocablo. Hasta la aparición de Thorndike (1921), solo se había manejado el término “frecuencia” para indicar el número de veces que aparece determinado léxico en un mismo texto o corpus. Sin embargo, Thorndike introduce el concepto de “distribución” (*range*) léxica para referirse a la frecuencia de dicho vocabulario en diferentes tipos de texto y géneros (Thorndike 1921: iii, Nation y Waring 1997:13, Matsuoka 2012:161, Nation 2013:24). La distinción tiene implicaciones importantes en la creación de corpora representativos y, por

⁷⁴ La primera lista de frecuencia léxica data del s.XIX y fue elaborada de forma manual a partir de un corpus de 11 millones de palabras, con la ayuda de 5.000 personas (Matsuoka 2012:159-160, McEnery y Wilson 2012). De igual manera, las obras de Thorndike fueron elaboradas de forma manual y diligente. Más tarde nacieron los primeros programas informáticos para contabilizar el léxico en textos electrónicos. De hecho, el primer corpus digital aparece, en Brown University, en el año 1960 (Leech et al. 2001:x, Baayen et al. 2016:5), más conocida hoy como Brown Corpus.

⁷⁵ Varias publicaciones ofrecen información adicional al respecto (Nation y Waring 1997, Matsuoka 2012, Kilgariff et al. 2014).

tanto, en la selección del vocabulario más útil para el desarrollo de materiales de lectura eficaces, puesto que a mayor “distribución” mayor utilidad léxica⁷⁶.

En definitiva, uno de los beneficios más importantes de utilizar listas de frecuencia es que sirven de guía para decidir qué vocabulario debe presentarse en los materiales didácticos y cómo dosificarlo de forma graduada (O’Keeffe et al. 2007:46-47). De hecho, los especialistas señalan que las listas de frecuencia léxica pueden tener varias utilidades en la enseñanza de lenguas (Leech et al. 2001: ix, Nation 2004:12, Kilgariff et al. 2014):

- Diseño curricular (selección del vocabulario a utilizar en libros infantiles, libros de texto para estudiantes de lenguas extranjeras, diccionarios, etc.).
- Elaboración de lecturas graduadas.
- Diseño de exámenes de L2.

Esto es así porque el vocabulario central (*core vocabulary*) de un idioma es el que ocurre con mayor frecuencia y, por tanto, el más práctico en términos de aprendizaje. En palabras de Schmitt (2010:63):

“The frequency in which a word occurs in language permeates all aspects of vocabulary behavior. It is arguably the single most important characteristic of lexis that researchers must address. It is not difficult to see why, as it affects the acquisition, processing, and use of vocabulary. In terms of acquisition, frequent vocabulary are, by definition, the words most likely to be met in discourse. As a result, learners generally acquire more frequent vocabulary before less frequent lexis”⁷⁷.

Además, el comportamiento del léxico es altamente predictivo, como lo demuestra la ley de Zipf. Según esta fórmula matemática, la distribución del vocabulario en una lengua determinada es regular: $\text{rank} \times \text{frecuencia} = \text{constante}$ (Nation

⁷⁶ Por poner un ejemplo, el vocablo *voltage* es el sustantivo más frecuente en un corpus inglés de electrónica e ingeniería informática compuesto de 425.000 palabras aproximadamente (Khamis y Abdullah 2013). Pero no puede tomarse como referencia de frecuencia léxica si el objetivo es diseñar un curso de inglés periodístico o, para el caso, de lectura extensiva, puesto que el vocablo no tiene una distribución adecuada en diferentes registros de la lengua y, por tanto, tendría poca utilidad.

⁷⁷ Tal y como señalan Milton y Masrai (2016), dicha idea ya había sido expresada por Palmer en 1917, en su libro *The Scientific Study and Teaching of Languages*.

2013:32). En otras palabras, la frecuencia de una palabra es inversamente proporcional a su ranquin en una lista de frecuencia. Es decir, la palabra más reiterada se utiliza el doble de veces que la segunda palabra más frecuente; ésta, a su vez, el doble de veces que la tercera; y así sucesivamente en una relación inversa proporcional. Dicho de otro modo, la ley de Zipf viene a mostrar que existe un pequeño porcentaje de vocabulario que es utilizado con mucha frecuencia y que, por el contrario, existe una gran cantidad de léxico que se utiliza muy poco (Nation y Webb 2011:132). En teoría, la ley de Zipf es susceptible de probarse en todas las lenguas, aunque no todas hayan sido estudiadas desde el punto de vista de la frecuencia léxica. En el caso del árabe, la ley Zipf ha demostrado ser aplicable (Milton y Masrai 2016). Este punto vuelve a tratarse en el cap. XI, apart. B.

2. La elaboración de lecturas graduadas

En la actualidad, la mayoría de libros de texto que se utilizan en la enseñanza de lenguas extranjeras se basa en intuiciones acerca del uso del lenguaje, y no en datos empíricos sustentados en estudios de lingüística de corpus (Boulton 2010:37-39), sobre todo en lo que respecta al lenguaje oral (O’Keeffe et al. 2007:21). A pesar del escaso impacto de la lingüística aplicada en la enseñanza de lenguas extranjeras (Boulton 2010, Reinhardt 2010), una de las ideas centrales de esta tesis es que los instrumentos disponibles en la lingüística de corpus son esenciales para desarrollar materiales didácticos eficaces que contribuyan a la formación de estudiantes capaces de comunicarse e interactuar de forma adecuada en las distintas destrezas que constituyen la L2.

En el ámbito de la competencia lectora, la lingüística de corpus ha desempeñado un papel importante en la elaboración de lecturas graduadas (Hirsh y Nation 1992, Nation y Ming-tzu 1999, Nation 2013, Teng 2014 (a)), gracias a que las herramientas informáticas a nuestro alcance ofrecen datos objetivos y fiables acerca de la frecuencia léxica (McEnery y Hardie 2015:28) y de las estructuras sintácticas y gramaticales más utilizadas.

Una característica fundamental de las lecturas graduadas es que utilizan el vocabulario frecuente de la lengua meta. Dicho de otro modo, las palabras que el estudiante se encuentra en una lectura graduada son, o deberían ser, las que aparecen con mayor frecuencia en la lengua. De hecho, las editoriales que publican lecturas graduadas para la enseñanza de inglés L2 disponen de listas de frecuencia que se emplean como guía en la elaboración de dicho material. La mayoría de esas listas estarían basadas en estudios de corpus (Simensen 1987:49)⁷⁸, aunque son inaccesibles para el público general (Nation y Webb 2011:133), y permiten distinguir los diferentes niveles de lecturas de una misma serie. Schmidt y Schmidt (2012:491), sin embargo, señalan que son pocas las series de lecturas graduadas, en inglés, que disponen de listas de vocabulario frecuente elaboradas de forma científica. De hecho, si bien las lecturas más básicas contienen entre 250 y 300 lemas, y las más difíciles entre 3.000 y 4.000 (Simensen 1987:49, Hill 1997:62-63, Nation y Webb 2011:133)⁷⁹, esas franjas de frecuencia no se corresponden con el vocabulario más frecuente en la lengua meta (Schmidt y Schmidt 2012:491).

El valor de utilizar listas de frecuencia léxica en la elaboración de literatura simplificada ya se observó en los años 20 del siglo XX (West 1927, Darton 1945), cuando Michael West introdujo la, en aquella época, novedosa metodología consistente en crear lecturas graduadas destinadas a los aprendientes de inglés L2 de las colonias inglesas a partir de listas de frecuencia (West 1927, 1937). El vocabulario utilizado en dichos textos estaba extremadamente controlado para que, así, el aprendiente fuera capaz de progresar en la competencia lectora de forma gradual. Por ejemplo, en las lecturas graduadas diseñadas por West, por cada 60 tokens escritos solo aparecía un lema nuevo (West 1927:11), lo cual equivale a un 1,6%. Este porcentaje se corresponde, en realidad, con el hecho de que el alumno necesita conocer el 98% del léxico para alcanzar una adecuada comprensión lectora. Actualmente, la idea de que medir estadísticamente la repetición del vocabulario en las lecturas graduadas es necesario, y está aceptada como método necesario y eficaz (Teng (b) 2014:37).

⁷⁸ Algunas lecturas graduadas, sin embargo, parecen haber sido elaboradas de forma intuitiva (Hill 2013:89).

⁷⁹ La escala léxica que utiliza The Extensive Reading Foundation Graded Reader Scale puede consultarse en http://www.robwaring.org/er/scale/rw_scale.htm

En definitiva, la irrupción de la lingüística de corpus en la enseñanza de lenguas extranjeras muestra datos objetivos que deben considerarse a la hora de ofrecer ejemplos lingüísticos auténticos a los aprendientes. Esto significa que tanto el léxico como las estructuras gramaticales deben presentarse de forma graduada y de acuerdo a la competencia lingüística del estudiante (O’Keeffe et al. 2007:25-27).

B. LINGÜÍSTICA DE CORPUS Y DIDÁCTICA DE ALE

1. Listas de frecuencia léxica

Si bien desde hace décadas se ha extendido cierta conciencia de la importancia de incorporar la lingüística aplicada al diseño curricular, y en particular las herramientas propias de la lingüística de corpus al desarrollo de materiales didácticos, la iniciativa sigue sin calar en la EALE.

Un investigador pionero que llamó la atención sobre el inadecuado diseño didáctico de libros de texto destinados a la EALE fue Jacob Landau, quien observaba que el material de lectura no estaba cuidadosamente elegido y que en particular era deficiente en relación con la selección del léxico (Landau 1959: v-vi):

“One need hardly stress again that vocabulary selection is one of the basic problems connected with the compiling and improvement of textbooks. A careful selection may lead to greater efficiency in language teaching. This may be even more relevant for literary Arabic than for other languages, as Arabic is rich in its choice of words and idioms and is rather difficult to study, apparently more so for Americans and Europeans, who, as Professor Tritton has put it, find “none of the familiar landmarks”.

Landau hizo notar que los libros de texto utilizaban un vocabulario aleatorio y que, además, no se empleaba de forma reiterada (Landau 1959:vi), lo cual implicaba que el alumno no tenía la oportunidad de reencontrarse con el léxico en diferentes contextos, condición necesaria para el aprendizaje.

Las primeras listas de frecuencia que se conocen en el contexto de la EALE fueron elaboradas por el profesor Moshe Brill en los años 30 del siglo XX para el género periodístico (Landau 1959: vii-ix). A ellas les siguieron las realizadas por Landau en el ámbito de la prosa árabe moderna, cuyo objetivo era facilitar al estudiante de ALE la comprensión de textos escritos periodísticos y literarios (Landau 1959: ix). Para el investigador, la creación de listas de frecuencia era “*el método más eficaz*” (Landau 1959: iv) para introducir el vocabulario en el currículum y superar las deficiencias señaladas en los párrafos anteriores.

Tras esas dos obras pioneras, la EALE dio la espalda a la lingüística aplicada y se optó por desarrollar materiales de enseñanza en los que el vocabulario, las estructuras gramaticales y los textos eran elegidos por conveniencia, no por frecuencia de uso. Esta dinámica no se interrumpió hasta 2011, cuando la aparición de *A Frequency Dictionary of Arabic: Core Vocabulary for Learners*, de Buckwalter y Parkinson, volvió a manifestar la necesidad de desarrollar materiales didácticos de mejor calidad, incluidas lecturas graduadas, que se basaran en listas de frecuencia actuales (Buckwalter y Parkinson 2011:1).

La lista de frecuencia de Buckwalter y Parkinson fue extraída de un corpus de 30 millones de palabras y contiene 5000 lemas. Según explican los autores, el corpus tiene la siguiente composición:

Discurso oral	Lenguaje oral	10%
Discurso escrito	Noticias de prensa	18%
	Editoriales y/o columnas de opinión	18%
	Prosa académica	18%
	Discusiones en foros online	18%
	Literatura	18%
TOTAL		100%

Tabla 2: Composición del corpus de Buckwalter y Parkinson (2011)

Por “lenguaje oral” se entiende el registro dialectal del árabe en sus diferentes variantes, y no el lenguaje oral que se produce en árabe estándar moderno. Por “prosa académica” se entienden mayoritariamente textos académicos y científicos, incluidos textos sobre el islam y artículos de publicaciones periódicas; y por “literatura”, cuentos cortos, novelas y teatro. Por otro lado, los autores afirman que en la categoría de lenguaje oral se incluyen muestras de discurso espontáneo en diferentes dialectos, y que las categorías de discurso escrito incluyen textos provenientes de todas las regiones del mundo árabe, con la excepción del género literario que comprende mayoritariamente textos de Egipto y el Levante (Buckwalter y Parkinson 2011:3-4).

En el campo de la EALE española no parecen existir estudios léxicos similares, aunque sí parece haber cierta conciencia de la relevancia de dichas listas de frecuencia. Así, por ejemplo, existe una obra de Corriente, Monferrer y Ould M. Baba titulada *Vocabulario árabe graduado* (2013) que consiste en un listado de vocabulario destinado a que el estudiante de ALE aprenda progresivamente el léxico en los niveles A1-C2, como herramienta necesaria en la superación del principal escollo del árabe, a saber, el léxico (Corriente et al. 2013:5-6). Sin embargo, los autores no especifican cómo seleccionaron dicho vocabulario ni en base a qué criterios lo dividieron en niveles de competencia.

2. *La elaboración de lecturas graduadas en árabe*

Si bien no existen aún lecturas graduadas para propiamente la EALE, los hadices del Profeta sí se han graduado por frecuencia léxica. En un reciente estudio perteneciente al ámbito del árabe clásico, el académico ruso Maxim Romanov (2015) compiló un corpus de 34.000 hadices (2,8 millones de tokens), con el objetivo de facilitar la enseñanza y el aprendizaje de esos textos en el ámbito universitario occidental. Romanov señala que una de las mayores dificultades para aprender árabe clásico es la enorme cantidad de vocabulario y que, por ello, las lecturas graduadas, en particular en el área de los estudios islámicos, son un instrumento útil. En palabras del investigador:

“With a paced increase in difficulty of texts and incremental expansion of vocabulary, students are capable of digesting much larger volumes of text both in class and at home, and such an extended exposure enables students to internalize the authentic language more efficiently” (Romanov 2015: a)

Romanov asegura que, utilizando este método, se pueden enseñar 400 hadices en un semestre de forma gradual y asimilar el vocabulario y las estructuras gramaticales que en ellos aparecen. Esto es así porque los 100 tokens más frecuentes en dichos textos constituyen el 56% de todo el léxico que aparece en el corpus (Romanov 2015: a). A partir de su análisis, Romanov creó un índice de dificultad lectora y logró ordenar 400 hadices en base a la frecuencia del vocabulario y de las estructuras gramaticales utilizadas. Además, describió dos hechos muy reveladores: 1) los hadices más largos incluían el léxico menos frecuente; por tanto, los hadices más fáciles eran los más cortos, y 2) el vocabulario más frecuente tiende también a aparecer en las estructuras gramaticales y sintácticas más frecuentes (Romanov 2015:c). La estrecha relación entre la frecuencia léxica y las estructuras gramaticales frecuentes, probablemente debida a fenómenos de coocurrencia, sugiere que las listas de frecuencia basadas en corpora podrían servir para guiar la graduación de lecturas en muchos otros tipos de textos.

Mi tesis doctoral continúa la línea de investigación iniciada por Brill, Landau, Buckwalter y Parkinson, y Romanov. En el siguiente capítulo explico en detalle los objetivos de mi estudio y cómo pretendo contribuir con él al desarrollo de lecturas literarias graduadas.

CAPÍTULO VI: LA PRESENTE INVESTIGACIÓN

A. OBJETIVOS

El objetivo específico de esta tesis es determinar cuáles son los 3.500 lemas más frecuentes en la novela árabe contemporánea, ya que es esa la franja de frecuencias cuyo conocimiento se exige para alcanzar los niveles A1 a B2 (cap. III, apart. A). Para la elaboración de esa lista se empleará la metodología propia de la lingüística de corpus (cap. V, apart. A.1). La presente investigación es, pues, un estudio empírico.

El objetivo general es proporcionar datos lingüísticos objetivos para el desarrollo de materiales didácticos en la EALE y, en concreto, para la elaboración de lecturas literarias graduadas eficientes capaces de desarrollar habilidades asociadas a la lectura extensiva entre los estudiantes de los niveles A1 a B2, entre otras, automaticidad en el reconocimiento visual del léxico, velocidad lectora y capacidad para adquirir vocabulario de forma incidental. El beneficio para los estudiantes de lenguas extranjeras de exponerse a la lectura extensiva desde los niveles iniciales también ha quedado demostrado (cap. III).

B. PREGUNTAS DE INVESTIGACIÓN

Esta tesis tiene como finalidad responder a las siguientes preguntas:

1. ¿Cuáles son los 3.500 lemas más frecuentemente utilizados en la novela árabe contemporánea?
2. ¿En qué porcentaje coincide esa lista con la elaborada por Buckwalter y Parkinson?

C. JUSTIFICACIÓN

Como hemos podido comprobar, la EALE carece de lecturas graduadas diseñadas en base a datos lingüísticos objetivos, y en particular obtenidos de un corpus literario, y tampoco existen obras literarias adaptadas o simplificadas que se basen en ese tipo de datos (cap. IV, apart. E). Como ha quedado demostrado, la elaboración de listas de frecuencia léxica estadísticamente fundadas contribuye sustancialmente a proporcionar esos datos en el plano léxico, e incluso ofrece pistas sustanciales para la localización de datos lingüísticos en los planos morfológico y sintáctico (Romanov 2015). Por tanto, la elaboración de una lista como la propuesta en el objetivo específico de esta tesis (cap. VI, apart. A) constituye un nicho de investigación que debe ser atendido, puesto que la exposición del estudiante de ALE a esos materiales de lectura extensiva, en particular a los literarios, es muy necesaria (cap. II).

La única lista de frecuencia léxica actualizada, a saber, la de Buckwalter y Parkinson, no está basada exclusivamente en un corpus literario, y las muestras de discurso escrito literario que comprende no es representativa tal y como se verá más adelante (cap. VII, apart. B y C). Por tanto, es preciso determinar en qué medida representa un instrumento adecuado para desarrollar material de lectura graduada, para estudiantes de A1 a B2, especialmente en lo que respecta la novela contemporánea, género que nos ocupa en esta tesis.

La necesidad de desarrollar materiales de lectura extensiva para la EALE, y de basar esos materiales en datos lingüísticos estadísticamente fundados, así como la necesidad de comprobar si los datos léxicos ya existentes son útiles en diferentes tipos de discurso y para el tipo de alumnos señalado, justifican plenamente, a mi juicio, la presente investigación.

D. VIABILIDAD

Para completar con éxito la presente investigación es imprescindible disponer de un corpus literario representativo de la novela árabe contemporánea en formato electrónico y analizarlo utilizando programas informáticos adecuados a los objetivos del estudio. Todo ello es posible actualmente como se explicará más adelante con detalle.

E. DISEÑO DE LA INVESTIGACIÓN

Para responder con éxito a las preguntas de investigación y lograr los objetivos propuestos se seguirán tres pasos principales:

1. Establecer los criterios de representatividad del corpus literario.
2. Reunir un corpus representativo.
3. Analizar el léxico del corpus tras su etiquetado y lematizado.

Cada uno de estos pasos será descrito en detalle en los próximos capítulos.

CAPÍTULO VII: CRITERIOS DE REPRESENTATIVIDAD DEL CORPUS LITERARIO

A. REPRESENTAR LA NOVELA ÁRABE CONTEMPORÁNEA

En el marco de esta investigación, por “novela árabe contemporánea” se entienden las compuestas en lengua árabe estándar en las últimas dos décadas por autores árabes de procedencia nacional diversa. Con arreglo a esa definición habrán de tenerse en cuenta tres criterios fundamentales para identificar la población o universo a partir de la cual se extraerá la muestra que conforme el corpus de la presente investigación. Esos tres criterios son: género, diacronía y registro. A continuación, definimos esas variables independientes.

1. Género

Los diferentes géneros literarios muestran características lingüísticas particulares (Busse 2014:103-104). La composición de un corpus a partir de muestras tomadas exclusivamente de novelas destinadas al público general adulto garantizará la homogeneidad de género (Atkins et al., 1992). Por ello, el corpus excluirá el cuento corto, pues la novela constituye en sí misma un género narrativo con características discursivas propias (Hafez 1993:160, 199)⁸⁰, y la literatura infantil y juvenil, ya que su léxico y estructuras sintácticas difieren de los de la novela destinada al público general (Thompson y Sealey 2007, Stakic 2014).

⁸⁰ Hafez señala, incluso, que el lenguaje utilizado en el cuento corto árabe difiere del que se usa en la novela en que el primero es más poético y muestra un vocabulario cuidadosamente elegido (Sabry Hafez 1993:13), lo cual podría significar que es menos frecuente; aunque, para demostrarlo, esto debe ser estudiado utilizando técnicas propias de la lingüística de corpus.

2. *Diacronía*

Convencionalmente se adoptará como criterio de selección cronológica que las muestras representen a novelas publicadas en los últimos veinte años. Esto garantizará la sincronía lingüística y un uso contemporáneo de la lengua árabe estándar. Por tanto, las obras anteriores a 1991 se excluirán de la población.

3. *Registro*

Esta investigación se centra en el léxico en árabe estándar moderno. En la población, por tanto, si bien se descartan las novelas escritas en su totalidad en árabe dialectal, se aceptarán aquellas con diálogos en ese registro, por ser esta una característica común a la novela árabe contemporánea desde principios del siglo XX (Zakariyya, 1964)⁸¹. Sin embargo, en las muestras se descartarán los fragmentos dialectales.

B. *CORPORA DISPONIBLES*

En la actualidad no existe un corpus accesible exclusivamente de novela árabe contemporánea. Entre los pocos corpora electrónicos de acceso público con muestras literarias árabes disponibles (Zaghouni 2014) conviene señalar los siguientes por su relevancia para la presente investigación:

⁸¹ Me refiero a la publicación de *Zaynab*, considerada por muchos críticos la primera novela en la historia de la literatura árabe. Sin embargo, cabe señalar que existen debates recientes sobre si *Sāq 'alā sāq* (1855), de Fāres Aḥmad al-Šidyāq, sería la verdadera primera novela árabe, en lugar de *Zaynab*, de Muḥammad Ḥusein Haykal: <http://arablit.org/2015/04/27/first-arabic-novel/>, aunque en este contexto no tiene particular importancia por no incluir la primera diálogos en dialecto, y por no ajustarse a las condiciones establecidas para la formación de mi corpus.

- *ArabiCorpus*: Interfaz de consulta *online* basada en un corpus de 173 millones de palabras⁸² compuesto por muestras de diversos géneros, entre otros la novela árabe moderna y, en menor medida, contemporánea. Parkinson, autor de *ArabiCorpus*, informa⁸³ de que en 2012⁸⁴ las muestras literarias eran de obras compuestas en 7 países árabes⁸⁵ y representaban a 32 títulos que, en teoría, son novelas. Una vez analizadas, el listado se reduce a 25 títulos, pues 7 son obras teatrales o colecciones de cuentos cortos⁸⁶. La proporción de autores varones y mujeres (6 mujeres frente a 10 varones) se adecúa con la realidad editorial de obras publicadas por ambos sexos. Sin embargo, algunos de los autores señalados tienen varias novelas en el corpus. Por ejemplo, de las 25 novelas incluidas, 5 pertenecen al egipcio Naguib Mahfuz y 3 a la argelina Ahlam Mustaghanmi. Otro aspecto importante de esta muestra literaria es que abarca obras publicadas en una franja temporal muy amplia (de 1959 a 2005), por lo que podría ponerse en duda que represente a la “novela contemporánea” según quedó definida más arriba. El contenido de este corpus, por último, no es de acceso público.
- *International Corpus of Arabic*: Proyecto fundado por la Biblioteca de Alejandría en Egipto cuyo objetivo es compilar un corpus representativo de la lengua árabe. El corpus está en fase de desarrollo y, según la página web del proyecto⁸⁷, se está recopilando con artículos de prensa, artículos en línea, libros y obras académicas. La muestra contiene textos pertenecientes a 11 géneros y 24 subgéneros distintos. Entre los géneros se incluye la literatura, dividida en dos subgéneros: prosa y poesía. La prosa ha sido subdividida en cuatro sub-subgéneros: novela, cuento corto, literatura infantil y teatro. En total, el género literario ocupa un 12% del corpus, aunque no se especifica qué obras en concreto. El corpus no es de acceso público.

⁸² Consulta: 5 de noviembre de 2016.

⁸³ <http://arabicorpus.byu.edu/search.php> Bajo el epígrafe “Detailed corpus information”. Consulta: 5 de noviembre de 2016.

⁸⁴ La muestra literaria no ha cambiado desde entonces.

⁸⁵ Egipto, Palestina, Argelia, Arabia Saudí, Sudán, Siria y Líbano.

⁸⁶ Para un listado completo de las obras, véase Apéndice A

⁸⁷ <http://www.bibalex.org/unl/frontend/Project.aspx?id=9> Página web consultada por última vez el 23 de enero de 2017.

Los dos corpora que he descrito son inadecuados para los objetivos de la presente investigación por dos razones fundamentales:

- 1) *ArabiCorpus* no contiene una muestra representativa de la novela árabe contemporánea, y
- 2) el *International Corpus of Arabic* no especifica la composición de la muestra literaria y, por tanto, los resultados de su explotación no pueden justificarse como representativos.

C. ADECUACIÓN DE LA POBLACIÓN

Teniendo en cuenta las tres variables dependientes rectoras de la selección de mi población he considerado conveniente adoptar las novelas ganadoras del *Naguib Mahfouz Medal for Literature* (NMML) - *yā`izat nayīb mahfūz*- y las finalistas y/o ganadoras del *International Prize for Arabic Fiction* (IPAF)⁸⁸ *al-yā`iza al-`ālamīyya li-l-riwāya al-`arabīyya* como población, para seguidamente tomar de ellas muestras representativas.

A continuación, ofrezco una breve descripción de los dos premios:

- 1) NMML: Es un premio que la Universidad Americana de El Cairo (AUC) otorga desde 1996 a la mejor novela árabe contemporánea no traducida al inglés. Según su página web oficial, se trata de un premio que busca fomentar la literatura árabe contemporánea y su traducción para el público angloparlante⁸⁹. El premio es anual y consiste en la entrega de 1.000 dólares al novelista ganador, además de la firma de un contrato con la editorial de la AUC para la traducción y publicación de la obra ganadora.
- 2) IPAF: Es un premio que la Autoridad de Turismo y Cultura de Abu Dhabi concede desde 2008 a la mejor novela árabe del año. Según su página web oficial, tiene como objetivo apoyar la ficción árabe de calidad, promover la lectura y fomentar la traducción de literatura árabe⁹⁰. El premio es anual y consiste en la

⁸⁸ A partir de aquí llamaré a ambos premios por su correspondiente acrónimo.

⁸⁹ http://www.aucpress.com/t-nmmdescription.aspx?template=template_naguibmahfouz

⁹⁰ <http://www.arabicfiction.org/>

entrega de 10.000 dólares a cada uno de los 6 autores finalistas y de 50.000 dólares al autor de la novela ganadora.

El criterio principal que me ha llevado a la elección de estos dos premios es que están formados por novelas contemporáneas que disfrutaban de un alto grado de reconocimiento en los círculos académicos y en la prensa, y que son populares entre los lectores árabes. Es importante recordar que los textos que forman parte de estos dos premios son novelas escritas en árabe estándar moderno y han sido publicadas en los últimos veinte años. Es decir, respetan las tres variables dependientes que guían la selección de mi población.

De algún modo, utilizar las novelas de la NMML y del IPAF como criterio principal en la composición del corpus conlleva un alejamiento de los cánones literarios oficiales. Si entendemos por “canon literario” aquellas obras reconocidas por los críticos como obras magistrales, influyentes y representativas de la novela árabe en general, entonces sería más adecuado formar un corpus narrativo a partir de los autores y novelas que se repiten en las antologías literarias y en las historias de la literatura más importantes (en esencia, las obras de Mohammed Badawi 1993, Roger Allen 1995, Sabry Hafez 1993, Jayyusi 2005, Somekh 1991 y 2012, y Sakkut 2000). Sin embargo, dichas antologías suelen centrarse en la novela árabe moderna y no en la contemporánea.

Por el contrario, si se considera nuestra definición de “novela árabe contemporánea”, las novelas de la NMML y del IPAF constituyen una población adecuada a los objetivos de esta investigación. De hecho, comienzan a aparecer artículos científicos que legitiman el premio IPAF como un nuevo canon literario contemporáneo (McManus 2016) y se publican numerosas reseñas de dichas obras, de entre las que destacamos los artículos del crítico literario egipcio Şalāḥ Faḍl (2015).

Aunque este premio, en particular, ha provocado controversia entre periodistas, novelistas y críticos literarios desde su inicio (se ha criticado la falta de “neutralidad y objetividad” del premio⁹¹, la novelista egipcia fallecida Radwa Ashour se negó a presentar sus obras al premio⁹², y el novelista egipcio Gamal al-Ghitani llegó a tildar a los

⁹¹ <http://bit.ly/1crAYng>

⁹² El proceso de selección de obras participantes en el IPAF pasa por la propuesta de la editorial y la firma de un documento de aceptación por parte del autor de la novela en cuestión.

miembros del jurado del premio de “caprichosos”⁹³, por citar solo algunos ejemplos), muchas de las críticas dirigidas al IPAF parecen provocadas por la gran suma de dinero que se ofrece al ganador del año o se producen porque autores árabes de renombre (casi siempre pertenecientes a generaciones menos jóvenes) son “excluidos” del premio, es decir, no alcanzan los puestos finalistas o no consiguen estar entre los premiados⁹⁴. En realidad, la gran mayoría de artículos y críticas literarias publicadas en la prensa árabe consideran positivamente el premio y sus novelas finalistas. A pesar de ello, cada año, el comité de jurados parece obligado a aclarar públicamente que la selección de obras se hace únicamente en base a la relevancia y calidad literaria del texto.

Otro factor que contribuye al reconocimiento de las novelas de la NMML y del IPAF por parte del público árabe es la publicación de noticias referentes a estos premios literarios, la votación popular en algunas plataformas literarias⁹⁵ y el número de ediciones que al parecer se imprime de esas novelas⁹⁶. En otras palabras, la presencia mediática de esas obras en los círculos literarios y de lectores, sobre todo en lo que respecta al IPAF, contribuye a que las obras se conviertan rápidamente en éxitos de ventas.

Otro factor que contribuye de forma indirecta al reconocimiento de las novelas de nuestra población es que, si reciben el premio, son traducidas ipso facto. En el caso de las obras ganadoras de la NMML, las novelas se traducen al inglés por imperativo

⁹³ <http://bit.ly/1GO40Kt>

⁹⁴ En una entrevista personal con uno de los editores árabes que más premios IPAF ha recibido, el editor afirmaba que muchos autores ponían como condición para publicar su novela con determinadas editoriales el que la novela fuera presentada como candidata al IPAF, lo cual parece confirmar indirectamente la relevancia del premio (Feria del Libro de Abu Dhabi, mayo 2015). Esto ha sido, también, publicado en prensa (<http://www.aljazeera.com/indepth/opinion/2014/05/arabic-sci-fi-other-literary-revo-2014587172235514.html>)

⁹⁵ Resulta indicativa la lista de Mejores Cien Novelas Árabes aparecida en *Goodreads*, basada en la votación popular de los lectores árabes (<https://www.goodreads.com/list/show/8012>), y tan diferente de la lista hecha pública por la Unión de Escritores Árabes (<http://arablit.org/for-readers/top-105/>). La primera incluía en 2015 entre las 100 mejores novelas árabes 15 ganadoras en la NMML y el IPAF, y 10 en 2016, mientras que la segunda es más bien reflejo de lo que viene considerándose canon narrativo árabe. En esta misma línea conviene señalar que los editores árabes que han publicado novelas ganadoras y finalistas del IPAF admiten que el premio ha contribuido en gran medida a fomentar la lectura entre los árabes y que las obras, una vez anunciadas en la lista de finalistas, se convierten inmediatamente en éxitos de ventas (entrevistas realizadas durante la Feria del Libro de Abu Dhabi, en mayo de 2015).

⁹⁶ En ocasiones, el número de ediciones impreso no es muy indicativo del número de ejemplares vendidos, puesto que algunos editores lo utilizan simplemente como estrategia de marketing. De hecho, la tirada de cada edición puede variar desde cientos hasta miles de copias, un dato que nunca se indica en el libro. Sin embargo, por regla general, el número de ediciones sí es indicativo de la popularidad de la obra entre los lectores y/o de los intentos editoriales de popularizar el texto (entrevistas realizadas durante la Feria del Libro de Abu Dhabi, en mayo de 2015).

contractual, mientras que las novelas ganadoras en el IPAF son generalmente traducidas a diversos idiomas, al igual que algunas de las finalistas, si suscitan interés⁹⁷.

En resumen, estos dos premios son muy prestigiosos en los círculos literarios y gozan de gran reconocimiento entre los lectores árabes⁹⁸, especialmente el IPAF, que parece estar convirtiéndose en un canon literario pan-árabe en formación. Por tanto, su selección como población validará, a mi juicio, los resultados de esta investigación, especialmente si se tiene en cuenta que la reputación puede a ser un factor relevante en la selección de un corpus (Miller 2015:128).

D. DESCRIPCIÓN DE LA POBLACIÓN

Tras descartar dos obras que no se ajustan al arco temporal⁹⁹, la población se compone de 145 novelas. A estas hay que restar, además, otra novela no fue posible localizar en ningún formato¹⁰⁰. En la tabla y el gráfico siguientes se describe la población.

Nº de novelas	NMML	18
	IPAF	129
	Novelas repetidas	3

⁹⁷ En el caso particular del español, cabe señalar que la editorial Turner inició una serie de literatura árabe contemporánea traducida que refleja la importancia del IPAF, puesto que, hasta el momento, ha publicado varias de sus obras ganadoras: *El oasis*, del autor egipcio Bahaa Taher (ganadora en 2008), *El arco y la mariposa*, del marroquí Mohammed Achaari (ganadora en 2011), y *Los drusos de Belgrado*, del libanés Rabee Jaber (ganadora en 2012). En la colección figura también una obra de Sinan Antoon, autor finalista en el IPAF (2013).

⁹⁸ Conviene señalar la existencia de un tercer y reciente premio literario que se ha descartado en la presente investigación. Se trata del premio Katara, establecido en 2014 y ofrecido por Qatar. Las obras pertenecientes a ese premio se descartan porque no puede medirse aún su reconocimiento entre críticos, académicos y lectores. El premio ofrece como primer premio la casi millonaria suma de 650.000 dólares.

⁹⁹ *Qiṣat Ḥubb*, de Yūsuf 'Idrīs (1957), y *Rāma wa al-Tinnīn*, de 'Edwar al-Jarrāṭ (1979).

¹⁰⁰ Se trata de *al-Maḥbūbāt* de la autora iraquí 'Āliya Mamdūh.

	Novelas diferentes	144 ¹⁰¹
Nº de autores	NMML	18
	IPAF	101
	Autores repetidos	10
	Total autores diferentes	109 ¹⁰²
Sexo de los autores	Mujeres	30
	Hombres	79
Nacionalidad de los autores	Egipto	35
	Líbano	24
	Palestina	15
	Siria	14
	Irak	13
	Marruecos	12
	Arabia Saudí	7
	Argelia	7
	Kuwait	4
	Túnez	4
	Sudán	4
Yemen	3	

¹⁰¹ Las novelas *Brūklīn hāits*, *Šawq al-darwīs* y *Lā sakākīn fī maṭābij hādīhi al-madīna* se repiten en ambas listas.

¹⁰² Mirāl al-Ṭaḥāwī, Ḥammūr Ziyāda y Jālid Jalīfa se repiten en ambas listas con las mismas obras. Asimismo, existen autores que aparecen en las dos listas como Ibrāhīm ‘Abd al-Majīd, Hudā Barakāt, Binsālim Ḥammīš, Jālid Jalīfa, Saḥar Jalīfa, Jairī Šalabī y Ḥasan Dāwūd, aunque con obras diferentes.

	Jordania	1
	Libia	1
	TOTAL	144
Fecha de publicación	NMML	1996 a 2015
	IPAF	2008 a 2016

Tabla 3: Descripción de la población

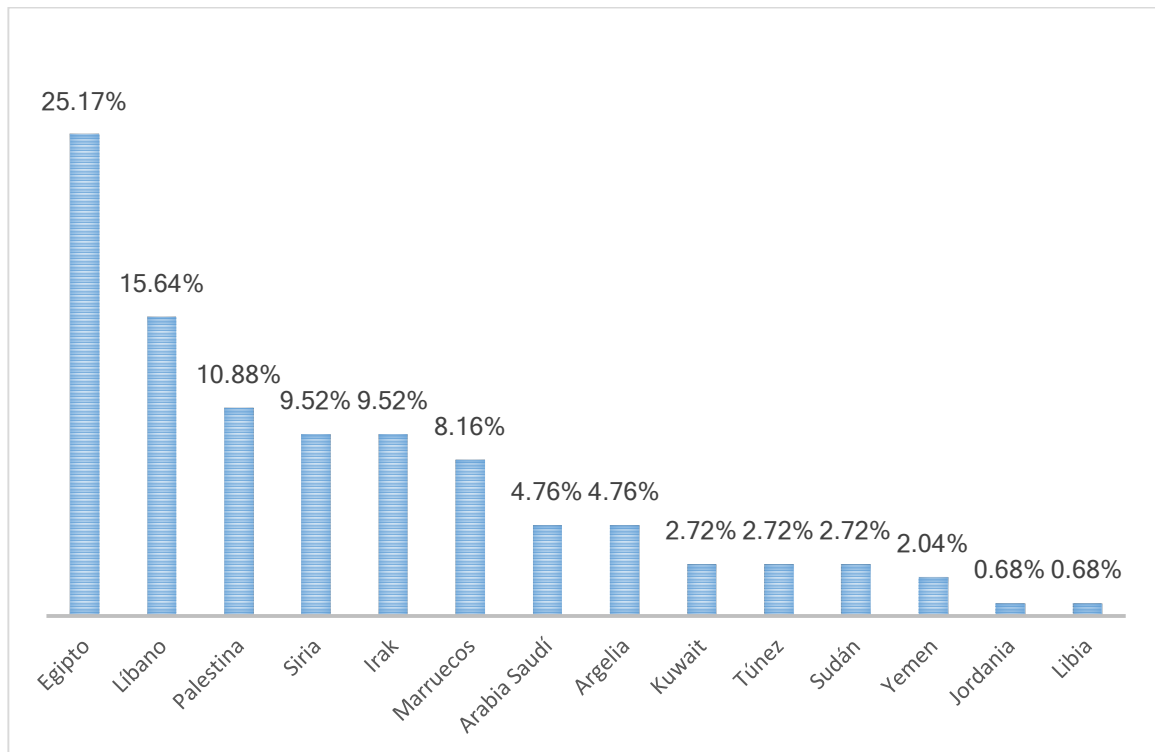


Gráfico 1: Procedencia nacional de los autores

E. POBLACIÓN y VARIABLES

La diversidad que puede observarse en las novelas que componen la población es notable y deriva en una ventaja esencial: permite formar un corpus literario representativo de la novela árabe contemporánea. De este modo, la población, en su totalidad, podría permitir el estudio de las variables que se describen a continuación.

1. Autoría

La inclusión de muestras lingüísticas escritas por una gran variedad de autores contribuye a reducir la idiosincrasia estilística del corpus en su totalidad y, por tanto, garantiza una mayor representatividad (Atkins et al., 1992). Esta variable, además, podría permitir estudiar el léxico utilizado por diferentes autores si la muestra estudiada tiene el tamaño y el diseño adecuados. Los objetivos de mi investigación, sin embargo, permiten eliminar esta variable.

2. Edad

La población incluye a autores bien establecidos en la escena literaria árabe y a autores noveles. Aunque los estudios no son concluyentes respecto a si existe una correlación entre la edad de la persona y su uso del lenguaje (Newman 2008), la población en su conjunto podría en el futuro permitir análisis centrados en esta variable. Para el objetivo de la presente investigación la diferencia generacional no supone desventaja o ventaja alguna, por lo que esta variable puede ser desestimada.

3. *Variación diatópica*

El vocabulario y las estructuras lingüísticas utilizadas por los hablantes nativos de árabe difieren de una región a otra en su forma escrita (Badawi et al., 2004; Ibrahim, 2008). Ha quedado empíricamente demostrado, por ejemplo, que los egipcios y los libaneses eligen sistemáticamente vocablos y expresiones léxicas diferentes al redactar artículos de prensa en árabe moderno estándar (Parkinson y Ibrahim 1999), así como que muchos egipcios tienen dificultad para entender los titulares de noticias publicadas en periódicos libaneses debido al uso de lemas diferentes (Parkinson y Ibrahim 1999). De hecho, parece existir una interrelación entre la variedad dialectal materna del autor y el léxico utilizado en la escritura, fruto de la diglosia que caracteriza al árabe desde hace siglos (Versteegh 2001, Ibrahim, 2008). Así pues, la incorporación de novelas escritas por autores que pertenecen a diversos países del mundo árabe enriquecerá los resultados, ya que serán diatópicamente más representativos.

Este tipo de estudios empíricos podría igualmente aplicarse a la población. Sin embargo, esta variable diatópica se diría irrelevante para mi investigación, puesto que se pretende extraer léxico frecuente para unos fines didácticos muy concretos y para unos niveles lingüísticos en los que las diferencias diatópicas son aún poco rentables.

4. *Variación temática*

Aunque la novela árabe contemporánea presenta, en ocasiones, ciertas tendencias, como puede ser la reciente vuelta a la novela histórica y el tratamiento de temas sociopolíticos, como la opresión del ciudadano¹⁰³, considero que el contexto social y el bagaje cultural-vivencial del autor se reflejará en los temas abordados por diferentes textos y, por tanto, en el vocabulario utilizado para tales fines. La importancia de esta variable radica en que, cuanto mayor sea la riqueza temática, mayor variedad encontraremos en la elección y uso del vocabulario por parte de los autores y, por tanto, se obtendrá una mayor representatividad lingüística (Kilgariff et al. 2014). Además, la

¹⁰³ Ver un artículo reciente acerca de la dirección que ha tomado la literatura árabe a raíz de la Primavera Árabe: <http://www.thenational.ae/opinion/comment/how-the-arab-spring-has-altered-the-direction-of-regional-literature>

variable puede ser controlada y, por tanto, podría estudiarse el comportamiento léxico en diferentes contextos temáticos.

5. Sexo

La inclusión en nuestra población de textos escritos por hombres y por mujeres promueve, sin duda, su representatividad. Aunque no puede asegurarse que las proporciones de representación masculina y femenina en la población se correspondan con las de la realidad editorial o de ventas, ya que no disponemos de datos al respecto¹⁰⁴, parece evidente que el predominio de autores varones en mi población es realista.

Téngase en cuenta que la variable de sexo tiene un reflejo diferente en el lenguaje que utilizan hombres y mujeres, tal y como han demostrado diversos investigadores para el inglés (Baker 2014:22-24). De hecho, la psicología social –apoyada por instrumentos de lingüística de corpus– ha logrado aportar pruebas suficientes que demuestran que el vocabulario y las estructuras lingüísticas utilizadas por hombres y mujeres difieren en muchos aspectos, al menos en inglés (Mehl y Pennebaker 2003; Newman et al. 2008; Pennebaker 2011). En la lengua escrita en particular, ya sea en textos literarios o no, las palabras funcionales aparecen utilizadas de forma distinta en hombres y mujeres (Koppel et al. 2002; Newman et al. 2008), al igual que el vocabulario relativo a procesos psicológicos y contextos sociales (Newman et al. 2008).

Existen pocas investigaciones sobre la variable sociolingüística de sexo en lengua árabe (Bassiouney 2009:128-197). Sabry Hafez sugiere que la irrupción de la mujer en la escena narrativa árabe generó “mutaciones sintácticas” y “un nuevo modo de expresión lingüística” (Hafez 1995:157), y afirma que un estudio detenido lo demostraría; sin embargo, años después, la sociolingüista egipcia Reem Bassiouney (2009:154) afirma: “*None of the studies discussed demarcates a clear and crucial difference in the way men and women use language in narration. Perhaps there is none. More studies are needed, however, before one can reach this conclusion*”.

¹⁰⁴ M. Lynx Qualey señala que, en el campo de la traducción, un tercio de las obras literarias árabes traducidas pertenece a mujeres (<http://arablit.org/2013/05/17/where-are-the-women-in-arabic-translation/>). De forma similar, la organización estadounidense VIDA, recoge en su página web datos anuales que muestran la menor presencia femenina en la escena literaria (publicaciones, reseñas, traducciones, etc.) en comparación con los hombres (<http://www.vidaweb.org/the-count/>).

Un corpus representativo de la novela árabe contemporánea, como el mío, podría servir para indagar en el futuro en esta variable independiente mediante métodos propios de la lingüística de corpus y descubrir si también en árabe existen diferencias discursivas y léxicas entre la escritura literaria femenina y masculina (Stubbs 2014). Para los fines de la actual investigación, sin embargo, esta variable es prescindible.

E. CONCLUSIÓN

La población es, a mi juicio, representativa. La ventaja de que esté conformada por las novelas participantes en los premios de la NMML y del IPAF radica en que esas obras son homogéneas desde el punto de vista del género literario (novelas contemporáneas), del registro lingüístico (están escritas en árabe moderno estándar) y de la variable diacrónica, al tiempo que tienen acreditado su reconocimiento académico y popular. Por otro lado, la población de muestras debe ser suficientemente representativa en relación con las variables de sexo y de diatopía.

Por todo ello, si la extracción de las muestras se realiza en base a criterios fundados, el corpus sería representativo de un corpus novelístico panárabe contemporáneo y permitirá alcanzar los objetivos de esta investigación.

CAPÍTULO VIII: COMPOSICIÓN Y CONSTRUCCIÓN DEL CORPUS

A. DISPONIBILIDAD DE OBRAS

La disponibilidad de los elementos conformantes de la población es una condición esencial para reunir el corpus en la práctica. Aunque las novelas IPAF y NMML son generalmente fáciles de localizar en formato impreso, es imprescindible contar con ellas en formato electrónico por razones de gestión informática obvias.

Otro requisito fundamental para la viabilidad de la investigación es obtener permiso para el uso de los textos de sus titulares de copyright. Para lograrlo me puse en contacto con las editoriales y los autores de la población, les expliqué los objetivos de la investigación y les ofrecí garantías del uso de las copias para fines estrictamente académicos. En ocasiones se remitió una carta oficial¹⁰⁵ avalando el estudio. También se les aseguró que los textos no serían puestos en circulación y que se protegerían los derechos de autoría (Atkins et al., 1992).

Acceder a toda la población resultó tarea ardua. El proceso de recopilación se extendió entre noviembre de 2014 y enero de 2017 a través de contactos con novelistas y editoriales. El resultado de estos esfuerzos es la recopilación de textos que se describe a continuación.

¹⁰⁵ Procedente de la UAM y la UGR y firmada por los directores de la presente investigación. Ver Apéndice B.

B. TAMAÑO DE POBLACIÓN

La población está formada por 147 novelas IPAF y NMML, de las cuales logré obtener 47 en formato Word, 41 en formato PDF (original o escaneado) y 59 en formato impreso, es decir, el 32%, el 28% y el 40%, respectivamente. Las obras en formato Word las obtuve directamente de los autores en la gran mayoría de los casos, mientras que las obras en PDF las conseguí a través de fuentes diversas: los autores, los editores, en Internet, o las escaneé. Para estimar el tamaño de la población seguí estos pasos:

- 1) Sumé el número de palabras de las novelas disponibles en formato Word. En el cálculo incluí: título de la obra, nombre del autor, texto completo de la novela, dedicatorias y agradecimientos, citas que aparecen al principio de la obra o que preceden los capítulos, notas a pie de página, índices y referencias bibliográficas históricas utilizadas por el autor. No incluí información perteneciente a la editorial (suele aparecer en las primeras páginas de las obras), ni listas de obras pertenecientes a los autores o publicadas por la editorial, las cuales suelen aparecer en las páginas finales del libro.
- 2) Hice un recuento del número de páginas impresas que aparece en cada una de las novelas que forman parte de la población. Para ello recurrí a las novelas que poseo en formato impreso y a los datos editoriales ofrecidos por el mayor proveedor de libros online: www.neelwafurat.com. Tomé como referencia las obras impresas disponibles y así pude estimar que, al número de páginas ofrecido por las editoriales, hay que restarle, de media, diez páginas (las cinco primeras y las cinco últimas), porque suelen ser la portada, la contraportada, páginas en blanco, información referente a la editorial y al autor y, en ocasiones, una introducción a la novela redactada por un crítico literario. Finalmente, estimé que la población incluye 44.223 páginas en total.
- 3) Calculé el promedio de palabras por página basándome en las 43 novelas disponibles en formato Word¹⁰⁶. Dichos textos suman 2.462.030 palabras distribuidas en 12.784 páginas impresas, lo cual significa que el promedio de palabras por página impresa es de 192.

¹⁰⁶ Esto se hizo a fecha de 1 de diciembre de 2016. La cifra exacta ha cambiado ligeramente, pues fui recibiendo los textos en formato Word de forma gradual.

Por razones de conveniencia, la unidad de página, equivalente a 192 palabras, se ha utilizado como unidad de recuento a la hora de seleccionar las muestras que compondrán el corpus.

- 4) Para calcular el tamaño de los elementos de la población en PDF se multiplicó el promedio de palabras por página por el número de páginas en ese formato, lo que dio como resultado una estimación de 6.036.288 palabras.

Como conclusión, la población total estimada asciende a 8.498.318 palabras.

C. TAMAÑO DE LA MUESTRA

Conocido el tamaño de la población pasé a calcular la extensión en palabras de la muestra estadísticamente representativa. Esa muestra es el corpus utilizado en esta presente. Para obtener la representatividad se fijó el porcentaje de error en el 1% y el nivel de confianza en el 100%, cifras extremadamente conservadoras. Con esas consideraciones, la muestra es de 89.057 palabras.

Si bien un corpus de ese tamaño es más bien pequeño si se trata de representar la lengua general (Koester, 2010), no lo es para un corpus especializado, de un único género o registro (Biber 1993), y que sirve a un objetivo muy concreto (Nelson, 2010). Como afirma Koester: “*When analyzing high-frequency items, a relatively modest corpus may still yield robust and powerful findings*” (2010: 67). La calidad de la muestra, recordemos, debe primar sobre su cantidad (Carter y McCarthy 2001).

Por otro lado, el objetivo de la investigación es extraer el léxico frecuente de la novela árabe contemporánea, y para eso, en principio, no se necesita un corpus demasiado amplio (siempre que la población sea representativa). En inglés, por ejemplo, el 80% del léxico (tokens) que aparece en cualquier texto pertenece a la franja de las 4.500 palabras más frecuentes (Nation 2013:21)¹⁰⁷.

¹⁰⁷ Nation utiliza la familia de palabras como unidad de recuento y señala que el 80% de los tokens más frecuentes pertenecen a la primera franja de las 1.000 familias de palabras, lo cual equivale a unos 4.500 lemas aproximadamente (Schmitt y Schmitt 2014).

En conclusión, siendo la población representativa y robusta, y habida cuenta los fines de la investigación y los criterios estadísticos adoptados para calcular la muestra, un corpus de 89.057 palabras es muy representativo.

D. MÉTODO DE MUESTREO

Para garantizar la representatividad de la muestra, he realizado un muestreo estratificado, no un muestreo aleatorio simple.

Un muestreo aleatorio simple consistiría en elegir aleatoriamente las 89.057 palabras requeridas para la muestra, ya sea de una novela o de varias. El inconveniente de este método es que la temática de la novela y la idiosincrasia del autor afectarían enormemente a la calidad de la lista de frecuencia léxica y, por tanto, los resultados de la investigación quedarían invalidados. En otras palabras, la muestra elegida no sería representativa de la población (Biber et al. 1998).

Por el contrario, un muestreo estratificado selecciona muestras proporcionales de cada subgrupo (o estrato) de la población, en este caso, de cada obra. Por poner un ejemplo, si un estrato representa el 1% de la población, el número de palabras seleccionadas de dicha obra corresponderá al 1% de la muestra. De este modo se garantiza que la muestra sea representativa de la variación léxica existente en la población (Biber 1990, Biber et al. 1998:247-248) y de sus características lingüísticas discursivas idiosincrásicas (Stubbs, 1996).

Para calcular el tamaño de la muestra correspondiente a cada estrato se utilizó una sencilla regla de tres. Como puede observarse en el Apéndice C, el tamaño de los textos que conforman la muestra oscila entre 217 y 1835 palabras.

E. SELECCIÓN DE LA MUESTRA

Una vez determinado el tamaño de la muestra apliqué un muestreo aleatorio simple para extraer el corpus. El muestreo fue de carácter transversal y basado en la

página como unidad de recuento. La ventaja del muestreo transversal es que garantiza una mayor representatividad léxica. Dicho de otro modo, en lugar de seleccionar un texto de oraciones y páginas contiguas compuesto por el número de palabras necesarias en cada novela, seleccioné distintas partes en cada obra: en la novela *Ma 'bad anāmil al-ḥarīr*, por ejemplo, se necesita una muestra de 1.097 palabras o el equivalente a 6 páginas aproximadamente; en lugar de seleccionar al azar 6 páginas seguidas, utilicé un randomizador en línea ¹⁰⁸ para elegir 6 páginas no sucesivas, evitando con ello la coincidencia temática.

Respecto al uso del randomizador en línea, conviene aclarar lo siguiente:

- Los parámetros de selección incluyeron solo el cuerpo textual de la novela. En ningún momento se incluyeron la portada, contraportada, biografía del autor, o información referente a la editorial.
- El *output* del randomizador es una serie de páginas seleccionadas aleatoriamente que no suelen ser contiguas. Las páginas que pasaron a formar parte de la muestra fueron seleccionadas en orden de aparición en dicha serie. En las ocasiones en que el *output* del randomizador incluía páginas en blanco, seleccioné de forma automática la página siguiente a la indicada.
- Cuando la página seleccionada incluía solo un título, una dedicatoria, un agradecimiento o unas pocas líneas, el texto se seleccionó y se pasó a la siguiente página indicada por el randomizador.

En cuanto a la selección de la muestra lingüística en las páginas indicadas por el randomizador, quiero señalar lo siguiente:

- Cuando las páginas seleccionadas contienen texto que forma parte del cuerpo central de la novela, el número de palabras por página oscila entre las 100 y 260 palabras, dependiendo de la naturaleza del texto (narración pura o diálogo).
- Se estableció que la muestra lingüística extraída de cada página debía contener 192 palabras +/-10%. Este porcentaje se ha estimado oportuno, considerando la naturaleza del lenguaje literario y la dificultad de ajustar el tamaño de la muestra lingüística por página.

¹⁰⁸ www.randomizer.org

- Las muestras literarias elegidas son coherentes y cohesivas desde el punto de vista sintáctico. Todas las muestras seleccionadas comienzan tras un punto final o punto y seguido, y terminan igualmente en un punto final o punto y seguido. Sin embargo, en las ocasiones en que el estilo del autor se caracteriza por utilizar frases muy largas y no hacer uso de la puntuación para indicar el fin de la idea, se ha interpretado la coma o el punto y coma como un punto y seguido, siempre que el significado de la frase quedara completo. La importancia de que las muestras lingüísticas estén formadas por palabras contiguas que forman oraciones completas radica en que, en el futuro, el corpus permitirá analizar variaciones lingüísticas morfológicas y sintácticas (Biber 1993:248-249).
- Cuando el texto seleccionado empieza por una frase que tiene su inicio en la página anterior, solo se incorporó si la oración comenzaba en la última línea. De igual modo, cuando la muestra seleccionada finaliza en una frase que acaba en la página siguiente, esta se incluyó siempre que no excediera de una línea en la nueva página.
- En las muestras se eliminaron los fragmentos dialectales. En muchas ocasiones esto implicó eliminar todo un diálogo.
- Se eliminaron también todas las palabras o frases escritas en grafía no árabe. Los títulos de todas las novelas y sus correspondientes autores se incluyeron en el corpus final, por entender que forman parte esencial de la muestra literaria.
- El texto final seleccionado fue limpiado de dígitos y los guiones, para que no contabilizaran como palabras.

Estas pautas de recogida de la muestra se aplicaron a todas las novelas, independientemente del formato en que estuvieran disponibles. En el caso de las obras disponibles en formato impreso o PDF, la muestra se mecanografió y se pasó a formato Word para computar con exactitud el número de palabras incluidas en cada muestra.

En el caso de las obras disponibles en Word, estas se formatearon¹⁰⁹ antes de seleccionar la muestra, para garantizar que las páginas contuvieran una media de 192 palabras.

¹⁰⁹ Se utilizó la fuente Times New Roman, en tamaño 18, y en espacio y medio. En cuanto al formato de la página se utilizaron los siguientes parámetros: Formato Carta, cero indente, cero espaciado anterior, 10pt espaciado posterior, margen de 3'81 cms en la parte inferior de la página, 1'54 cms en la parte superior, y 3'15 cms a la izquierda y a la derecha.

Finalmente, y como ya se indicó, se decidió excluir de la muestra dos obras que fueron ganadoras del NMML a título póstumo. Las obras son *Qiṣṣat ḥubb*, de Yūsuf 'Idrīs (1957), y *Rāma wa al-Tinnīn*, de Edwar al-Jarrāṭ (1979). Además, no se pudo incluir una muestra de *al-maḥbūbāt* de 'āliya Mamdūḥ por la dificultad que supuso acceder a una copia de la novela. Las muestras compiladas pertenecen, así, a 144 novelas.

CAPÍTULO IX: ANÁLISIS

A. HERRAMIENTAS DE ANÁLISIS

1. Segmentación, tokenización y lematización

Para analizar morfológicamente la muestra (89.732 palabras gráficas) utilicé MADAMIRA 1.0 Beta, un programa desarrollado por Pasha et al. (2014). El programa es gratuito y de acceso público¹¹⁰. Según sus desarrolladores, MADAMIRA tiene un acierto del 96%, lo que lo convierte en un lematizador automático muy fiable (Pasha et al. 2014).

En esta investigación se ordenó a MADAMIRA segmentar y diacritizar el *input*, salvo los enclíticos listados en su manual (Pasha et al 2014:25). Como resultado se obtuvieron 114.629 tokens.

Para comprobar que el margen de error en el *output* de MADAMIRA no es superior al señalado por los autores del programa se revisó manualmente el 1,47% de sus tokens, es decir, 1683. En ellos se encontraron 30 errores imputables a MADAMIRA y uno imputable al mecanografiado de la muestra; es decir, el 1,8% de los tokens analizados eran erróneos. Los errores localizados se pueden dividir en tres categorías: de segmentación, de tokenización y de lematización.

- a) Errores de segmentación: Se producen mayoritariamente en palabras gráficas no diacritizadas cuyos componentes MADAMIRA pueda interpretar como prefijos o sufijos. Por ejemplo, **b* *īd* fue segmentada erróneamente en dos tokens: *bi* + *īd* (*bi* [preposición] + *īd* [sustantivo “fiesta”]), en lugar de *ba* *īd* [adjetivo “lejos”]).

¹¹⁰ Agradezco a Nizar Habash y Hind Saddiki de la New York University Abu Dhabi su ayuda durante la instalación del programa y en el diseño de programas Java para la presente investigación.

- b) Errores de tokenización: Como hemos podido comprobar, los errores de segmentación conducen a errores de tokenización (en el ejemplo anterior, dos tokens errores en lugar de uno acertado).
- c) Errores de lematización: asignación al token de signos vocálicos incorrectos como consecuencia del alto porcentaje de homógrafos del árabe¹¹¹.

El repaso manual permitió corregir 26 lemas¹¹², es decir, el 1,5% del *output* repasado. Esos márgenes de error en el *output* de MADAMIRA son perfectamente asumibles.

2. Elaboración de una lista de frecuencia

Para calcular la frecuencia de los lemas utilicé el programa WordSmith7. El *output* resultante incluye, entre otra, la siguiente información:

- El ranquin de frecuencia de los lemas en orden descendente.
- El número de ocurrencias de cada lema.
- El porcentaje de ocurrencia de cada lema en relación con el conjunto de lemas.
- El porcentaje de dispersión de las ocurrencias de cada lema calculado dividiendo el *input* en 8 partes iguales¹¹³.

La segmentación y tokenización del corpus resultó en una lista de frecuencia formada por 10.247 lemas. Un repaso manual parcial de los datos me permitió eliminar más errores de MADAMIRA no localizados anteriormente, entre otros el símbolo #, abreviaciones de etiquetas morfológicas en grafía latina y transliteraciones de palabras inglesas (como *yū* = “*you*”). El repaso desveló, asimismo, que MADAMIRA trata la palabra gráfica **lkn* como dos lemas distintos: *lākin* y *lākinna*. Considerando que esa tokenización no es justificable lingüísticamente, reduje ambos lemas a uno (*lākin*). Una

¹¹¹ En relación con la opacidad ortográfica y la homografía, véase el cap. I apart. C.2, y el cap. I. apart. C.3, respectivamente.

¹¹² Ver Apéndice D.

¹¹³ Abordo esta cuestión con detalle en el apartado C de este capítulo.

vez introducidas esas nuevas correcciones finales, la lista de frecuencia se redujo a 10.241 lemas.

Conviene señalar que la lista final de frecuencias tuvo que ser revisada y corregida manualmente de nuevo para eliminar dos errores tipográficos, ya que la *hamzat al-waṣl* y la *alif jan̄ariyya* habían sido eliminadas por WordSmith7 y sustituidos por un signo de interrogación. Además, durante el proceso de revisión, se descubrieron nuevos errores en segmentación, tokenización y lematización que obligó a eliminar o corregir algunos lemas y reajustar su correspondiente ranquin en la lista¹¹⁴.

B. FRECUENCIA Y DISTRIBUCIÓN LÉXICA

Como quedó dicho en el cap. V, apart. A.1 de esta tesis, la distribución del léxico es altamente predictiva y se rige por leyes estadísticas conocidas, en particular por la Ley de Zipf. A continuación, se repasa esta ley y se describe someramente la distribución de lemas en mi corpus.

La Ley de Zipf describe la distribución de un conjunto de datos mediante una constante exponencial inversa. La Ley de Zipf, aplicada a la distribución léxica, afirma que la frecuencia de una palabra es inversamente proporcional a su ranquin en una lista de frecuencia; dicho de otro modo, el lema más frecuente en un corpus debe aparecer el doble de veces que el segundo y este, a su vez, el doble de veces que el tercero, y los happax legómena¹¹⁵ deben constituir en torno al 50% de los datos, los dislegómena¹¹⁶ en torno al 16% y los trislegómena¹¹⁷ en torno al 8% (lo que sumaría aproximadamente el 75% de los lemas). Por tanto, en una muestra lingüística representativa el léxico se distribuiría gráficamente en un eje de coordenadas cartesianas con una recta de tendencia con pendiente cercana a -1. Sin embargo, como se ha podido observar en todas las lenguas, la Ley de Zipf no refleja con total exactitud la distribución léxica en ciertas

¹¹⁴ Para una lista de los errores, ver Apéndice E.

¹¹⁵ Palabras que ocurren una sola vez en un texto o corpus.

¹¹⁶ Palabras que ocurren dos veces.

¹¹⁷ Palabras que ocurren tres veces.

frangas del conjunto de datos, en particular en relación con las palabras de mayor y menor frecuencia.

La Ley de Zipf es tan predictiva de la distribución del léxico que, si bien su cumplimiento no indica necesariamente suficiente representatividad del léxico, su incumplimiento indica sin duda carencia de representatividad. A continuación, se presenta la distribución de frecuencia léxica de mi corpus con una gráfica logarítmica en base 10, como es común, ya que con gráficas lineares no se apreciarían las diferencias en las frecuencias más bajas.

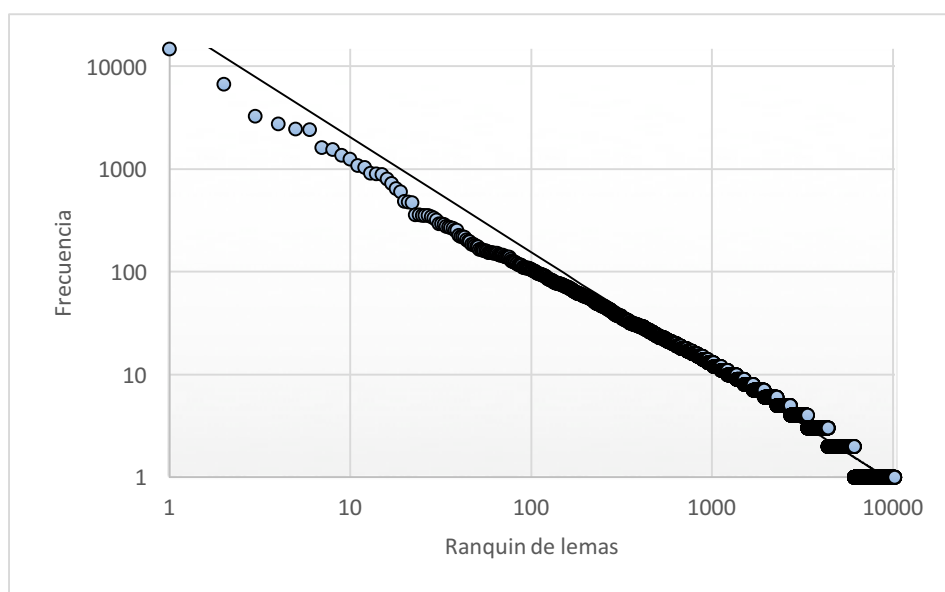


Gráfico 2: Distribución Léxica Zipf en mi corpus (10.241 lemas)

Como en cualquier otra distribución léxica representativa, los lemas de mayor ocurrencia descienden de forma muy acusada al principio (parte superior izquierda del gráfico), y dicho descenso se ralentiza a medida que nos acercamos a los elementos de menor frecuencia (Baroni 2009:11). Mis datos, aunque muestran una distribución regular, se distribuyen con una pendiente de $-1,52$, es decir, más acusada que la ideal zipfiana, lo que podría explicarse por ser el árabe una lengua altamente derivativa. Este fenómeno se ha comprobado igualmente en idiomas como el irlandés y el latín, donde la pendiente se acerca a -2 , especialmente en los ránquines más bajos (Hu et al. 2006).

Los tres lemas más frecuentes de nuestra lista son el artículo determinante *al*, la conjunción *wa* y la preposición *bi* con el 12,95%, el 5,84% y el 2,87% de frecuencia,

respectivamente. Los 10 lemas más frecuentes suman, pues, el 33,4% de los tokens. Este porcentaje coincide con lo comprobado en el léxico general de todas las lenguas (Baayern 2001, Baroni 2009:5).

En el otro extremo de la lista de frecuencia, los hápax legómena, los dislegómena y los trislegómena constituyen el 40,2%, el 16,9% y el 10%, respectivamente, de los lemas de mi corpus. Este grupo de vocabulario constituye por tanto el 67,1% de lemas del corpus, lo que coincide igualmente, de forma general, con lo que se observa en el léxico general de otras lenguas (Baayern 2001:8, Baroni 2009:5).

Por otra parte, en relación con las medidas de tendencia central conviene señalar que los lemas de mi lista de frecuencia muestran una frecuencia media de 11,9, lo que se explica si consideramos que casi el 70% del vocabulario únicamente aparece de 1 a 3 veces. La distribución léxica de mi corpus, por otro lado, presenta una mediana de 2 (lo que significa que el 50% de los lemas ocurre 2 o más veces, y el otro 50% ocurre 2 o menos veces) y una moda de 1 (por ser esta la frecuencia más repetida). Estas cifras son, en efecto, características de cualquier lengua y corpus representativo. De hecho, Baroni (2009:6, 8) muestra que esa distribución léxica es tan predictiva que puede llegar a cumplirse incluso en una sola novela.

El siguiente gráfico representa el perfil ranquin-frecuencia de los lemas de mi lista. Como puede observarse, su curva muestra la característica forma de L. La línea vertical de la izquierda representa los lemas más frecuentes en el corpus, mientras que las líneas horizontales que aparecen sobre el eje x son los hápax legómena (línea inferior), seguida de los dislegómena, y así sucesivamente. El punto más alto en la gráfica, pues, representa el lema *al*.

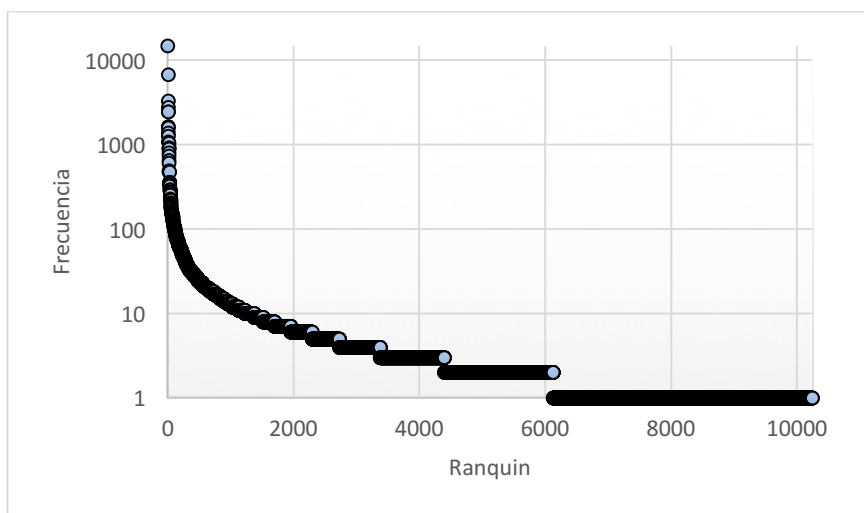


Gráfico 3: Perfil ranquin/frecuencia de lemas en mi corpus (10.241 lemsa)

Para clarificar esta información aún en mayor medida, la gráfica siguiente despliega a la izquierda los lemas de menor frecuencia, que son los más numerosos; así pues, el punto que aparece justo sobre el eje de ordenadas representa los 4115 hápax legómena y el último punto a la derecha, prácticamente sobre el eje de abscisas, el lema de mayor frecuencia.

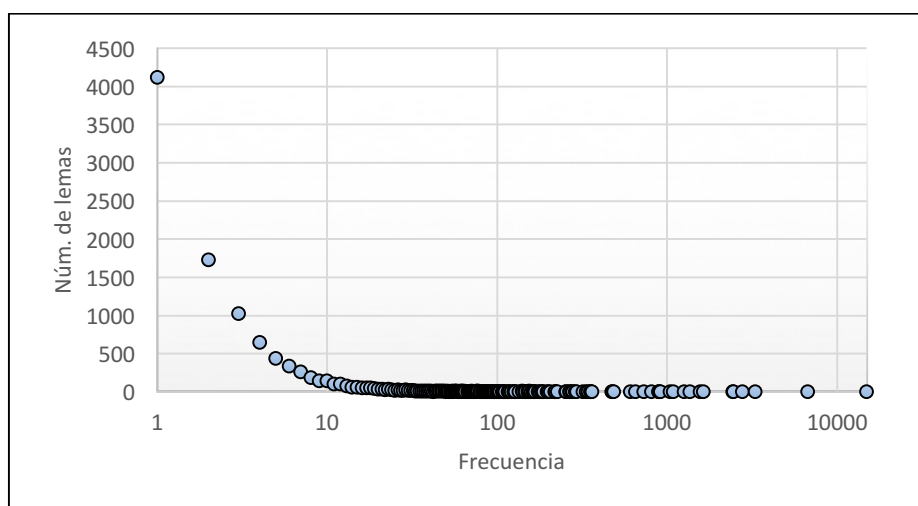


Gráfico 4: Espectro de frecuencia de lemas en mi corpus (10.241 lemas)

Este perfil de distribución léxica también es típico y puede observarse en todas las lenguas y en corpora muy diferentes (Baayen 2001, Baroni 2009).

Todos los datos anteriores confirman que la muestra empleada es representativa.

C. FRECUENCIAS ABSOLUTA Y RELATIVA

Es necesario, para interpretar datos referentes a la frecuencia léxica, distinguir entre frecuencias absoluta y relativa. Los valores de frecuencia analizados en el apartado anterior reflejan el número de veces que cada vocablo aparece en la muestra, es decir, su frecuencia absoluta; sin embargo, no indican el grado de uniformidad de la distribución del léxico en los distintos textos que forman la muestra. Dicho de otro modo, al afirmar que una determinada palabra tiene frecuencia absoluta n no estamos aclarando si lo hace n veces en el mismo texto o fragmento, o en contextos diversos. Para solventar esta cuestión se recomienda aplicar fórmulas de dispersión (Gries 2008).

WordSmith7 incorpora, como ya he indicado (cap. IX, apart. A.2), una medida de dispersión a partir de la división del corpus (que en este caso es la lista de lemas, recordemos, no la muestra analizada con MADAMIRA) en 8 partes iguales. La cifra resultante, expresada porcentualmente, indica el grado de distribución de los lemas: un alto porcentaje de dispersión indica una gran uniformidad en la distribución del lema y, por el contrario, un porcentaje bajo indicaría que el lema aparece muy concentrado (aunque la cifra absoluta de ocurrencias sea la misma).

En principio, mi objetivo era utilizar estos porcentajes de dispersión para afinar la lista de frecuencia en su totalidad. Sin embargo, al analizar los resultados se observaron fenómenos preocupantes, entre otros, que el lema *wa* no presentara una distribución del 100%, lo que lo colocó en la novena posición del ranquin (mientras que el lema *'an* pasó a ocupar la segunda). Ante la sospecha de que, por razones que no se han podido determinar, los porcentajes de dispersión proporcionados por WordSmith7 no sean del todo exactos, y al no disponer de un medio mejor para calcular la dispersión en este momento, opté por hacer el uso muy conservador de esos porcentajes que se describe a continuación.

D. ELECCIÓN Y AFINACIÓN DE UNA FRANJA LÉXICA

Para garantizar la elección de una franja de la lista completa de frecuencias que permita alcanzar los objetivos fijados para esta investigación se decidió utilizar la franja 1-2292 de la lista de frecuencias absolutas. Entre las ventajas de la franja seleccionada conviene señalar las siguientes:

- 1) Incluye el 65% del número de lemas previsto en los objetivos de la investigación, lo que puede considerarse un porcentaje admisible en la práctica;
- 2) la frecuencia absoluta mínima de ocurrencia en esa franja es 6, es decir, el mínimo necesario para poder realizar pruebas de χ^2 en un futuro (William y Cso-may 2016: 131-132);
- 3) el porcentaje de distribución medio según Wordsmith7 asciende a aproximadamente el 50%, y
- 4) en las franjas subsiguientes los datos no muestran similar grado de sistematicidad, probablemente debido al tamaño de la muestra.

Para que la lista final tuviera el mayor grado de representatividad posible, y una vez comprobado que el porcentaje medio de distribución ascendía al 50%, se comprobó manualmente que una distribución inferior al 20% implicaba que el lema ocurría en entre 1 y 3 de los textos que conforman la muestra. Ante ello, resolví eliminar de esa primera franja de 1 a 2292 los lemas cuyo porcentaje de distribución fuera inferior al 20%, independientemente de su frecuencia absoluta. En definitiva, eso supuso la eliminación de 33 lemas, por lo que la lista quedó reducida a 2259 lemas.

Esa, que en principio parecía ser la lista definitiva, volvió a precisar reajustes cuando se intentó comparar con la lista de Buckwalter y Parkinson. Como se verá a continuación, el proceso de limpieza de ambas listas fue circular y prolongado.

E. AJUSTES INTRODUCIDOS EN LA LISTA DE BUCKWALTER Y PARKINSON Y AJUSTES FINALES EN MI LISTA

La comparación de mi lista con la de Buckwalter y Parkinson, que es uno de los objetivos fijados en la investigación, resultó más difícil de lo esperado debido a las diferencias en la lematización, en los métodos matemáticos utilizados para calcular el ranquin, en el concepto mismo de lema (es decir, en el tipo de palabras incluidas en las listas) y en la ortotipografía. En la lista de Buckwalter y Parkinson se hubieron de introducir los siguientes ajustes para que la comparación fuera posible:

1. Se eliminó todo el léxico dialectal, puesto que en mi muestra no se incluían fragmentos dialectales (cap. VIII, apart. E). Según Buckwalter y Parkinson (2011:144-146), su diccionario incluye 102 lemas dialectales; sin embargo, un repaso manual desveló que en realidad son 125.
2. Los lemas *ḥabba* y *quddām*, listados como léxico propio del árabe moderno estándar y de algunas variantes dialectales, fueron eliminados porque se consideraron sincrónicamente dialectales¹¹⁸.
3. Ante dos ortografías posibles de un mismo lema incluidas en Buckwalter y Parkinson, se optó por la coincidente con la lematización de MADAMIRA.
4. Si un lema en Buckwalter y Parkinson con dos ortografías posibles no aparece en mi lista, se optó por la ortografía más común.
5. Si de las dos ortografías posibles una puede considerarse homógrafo de otro lema, esa fue descartada. Por ejemplo, entre **klwa* y **klya* (“riñón”) se descartó el segundo para evitar la coincidencia homográfica con *kulliya* (“facultad”).
6. Si Buckwalter y Parkinson no distinguen entre dos homógrafos, les añadí signos diacríticos. Por ejemplo, al lema *ḥadīt*, listado dos veces con dos significados distintos (“moderno” y “conversación”), le añadí *tanwīn* para marcar uno de ellos (“conversación”).

¹¹⁸ De hecho, comprobé el uso de ambos lemas en el buscador online www.arabicorpus.byu.edu (elaborado por el propio Parkinson) y aparece mayoritariamente en textos dialectales (búsqueda realizada el 25 de marzo de 2017).

Otras características que hubieron de tenerse en cuenta para comparar la lista de Buckwalter y Parkinson con la mía es que la primera:

1. No incluye nombres propios (personas, ciudades, países). Esos lemas aparecen en tablas independientes.
2. Recoge separadamente las variantes del pronombre relativo *allaḏī* y del demostrativo *hāḏā*.
3. Recogen algunos compuestos léxicos, como *kaḏālika*, *liḏālika*, *li`anna*, *laḡad*, *bimā*, *mimmā* e *indamā*, mientras que MADAMIRA no los trata como tal, sino que los divide en sus distintos componentes.
4. No incluye el compuesto *lawlā*, que, sin embargo, aparece en el *output* de MADAMIRA como lema.
5. No distingue entre el lema *ājar* y *ujrā*, mientras que MADAMIRA sí lo hace.
6. No distingue entre el lema *taraka* y *tark*. Ambos parecen haber sido unificados en *taraka*.

Teniendo en cuenta las características de la lista de Buckwalter y Parkinson, introduje los correspondientes ajustes en mi lista de lemas:

1. Todos los nombres propios fueron eliminados.
2. Por razones pedagógicas relacionadas con los objetivos de la presente investigación, decidí contabilizar las distintas variantes del pronombre relativo *allaḏī* y del demostrativo *hāḏā* en el *output* de MADAMIRA e incorporarlos como lemas separados en mi lista¹¹⁹. Para ello se eliminaron primero esos dos lemas de la lista y, a continuación, se contabilizaron nuevamente en la muestra los distintos tokens que representan y se insertaron en mi lista de frecuencias con su correspondiente número de ocurrencias¹²⁰.
3. Por las mismas razones, contabilicé en mi corpus las ocurrencias de los compuestos que aparecen como lemas independientes en Buckwalter y Parkinson:

¹¹⁹ Aunque conceptualmente todas las variantes de esos lemas pueden reducirse a uno, la decodificación visual de esas formas por parte de un estudiante de ALE supone distintos grados de dificultad por la enorme discrepancia entre sus frecuencias. Como indicaré en el cap. X, una lista de frecuencia léxica puede servir como guía también en la selección de los lemas cuyo contenido contiene información gramatical relevante.

¹²⁰ Para una lista completa de esos lemas, véase Apéndice F

kaḍālika, liḍālika, li'anna, laqad, bimā, mimmā e 'indamā. Esos tokens, pues, pasaron a engrosar mi lista de lemas¹²¹.

4. Dado que Buckwalter y Parkinson incluyen *lawlā* bajo el lema *law*, procedí a eliminarlo de mi lista y sumar sus ocurrencias a las del lema *law*.
5. Aunque Buckwalter y Parkinson listan *ājar* y *ujrā* bajo el lema *ājar*, en mi lista opté por dejar ambos porque gráficamente son distintos y, por tanto, su procesamiento visual por parte del estudiante de ALE puede suponer cierta dificultad.
6. Al no distinguir MADAMIRA tampoco entre *taraka* y *tark*, se colapsaron ambos lemas en *taraka*¹²².

El último paso antes de comparar las dos listas fue eliminar los signos diacríticos en ambas listas, ya que no eran coincidentes, y dejarlos solo en el caso de homógrafos, para diferenciarlos. En total, sin embargo, hay 4 homógrafos que no pudieron distinguirse en Buckwalter y Parkinson: *mā, qad, li, y la*¹²³. Esos homógrafos aparecen en Familiar unificados debido a la naturaleza del *output* de MADAMIRA y la dificultad de contabilizarlos manualmente. Finalmente, los lemas *ḥīn* y *ḥīna*, en Buckwalter y Parkinson, aparecen unificados en Familiar.

Tras este laborioso proceso de limpieza y preparación de las listas, la composición de las dos listas léxicas comparables fue la siguiente:

- Mi lista de frecuencia: 2151 lemas.
- Lista de Buckwalter y Parkinson (2011): 4851 lemas.

¹²¹ No creí necesario modificar las ocurrencias de los lemas que forman parte de dichos compuestos, por tratarse de lemas de muy alta frecuencia. Dicho de otro modo, aunque el compuesto *'indamā* sea segmentado por MADAMIRA en dos lemas (*'inda + mā*), la inclusión del compuesto requeriría revisar de nuevo la etiquetación de todos los *mā* que ocurren en árabe y asignarles, desglosados por tipo, un ranquin distinto, lo que, habida cuenta su frecuencia, resultaría extremadamente lento. Por razones prácticas de falta de tiempo, dicha tarea no se llevó a cabo.

¹²² Cabe señalar que en mi corpus la gran mayoría de las veces se trata del lema *taraka* (verbo).

¹²³ En los 2.166 primeros lemas de Buckwalter y Parkinson *mā* ocurre 8 veces, *qad, li, y la* dos veces. Al tratarse de lemas funcionales, se consideró aceptable no distinguirlos, aunque una distinción detallada de ellos sería interesante para analizar el discurso literario a nivel sintáctico.

F. NORMALIZACIÓN DE LAS LISTAS DE FRECUENCIA

Una vez preparadas las dos listas de frecuencia (en adelante, Buckwalter y Parkinson y Familiar) y para garantizar que los resultados fueran comparables, se procedió a normalizar los resultados considerando ocurrencias sobre 1000 y no número absoluto de ocurrencias, puesto que el tamaño de los respectivos corpus es muy dispar (Crawford y Csomay 2016:148-149)¹²⁴. Para ello, se utilizaron las cifras de ocurrencia absoluta que aparecen en Buckwalter y Parkinson, y no su ranquin, que está calculado aplicando una cifra de dispersión (Buckwalter y Parkinson 2011:5). Recordemos, además, que la cifra de dispersión fue descartada para mi lista de frecuencia.

En el próximo capítulo, presento los resultados más importantes tras comparar ambas listas.

¹²⁴ En la práctica se trata de dividir la frecuencia por el tamaño del corpus y multiplicar el resultado por 1.000. La cifra resultante indica el número de ocurrencia del lema por cada 1.000 palabras. Véase los resultados de la normalización de frecuencias en los 2.000 primeros lemas en Apéndice G.

CAPÍTULO X: RESULTADOS

Recordemos las dos preguntas de investigación planteadas en cap. VI, apart. B:

1. ¿Cuáles son los 3.500 lemas más frecuentemente utilizados en la novela árabe contemporánea?
2. ¿En qué porcentaje coincide esa lista con la elaborada por Buckwalter y Parkinson?

Como se ha visto, la lista de frecuencia extraída a partir del corpus literario utilizada en el análisis no cubrió finalmente 3.500 lemas, ya que el tamaño de la muestra no lo permitió. Por tanto, el primer resultado de esta investigación es que una muestra de 90.000 palabras en lengua árabe, con las características de la utilizada, no permite extraer una lista de frecuencia compuesta por 3.500 lemas.

Considerando, pues, que mi lista final está compuesta por 2.151 lemas, responderé a las preguntas de investigación limitándome a dicha franja léxica. Comenzaré, sin embargo, por responder a la segunda pregunta, ya que, como se verá, los resultados de la comparación de ambas listas tienen implicaciones importantes a la hora de elaborar y dar por válida una lista de frecuencia léxica basada en la novela árabe contemporánea que permita elaborar lecturas graduadas para los niveles A1 a B2.

A. SEMEJANZAS Y DISCREPANCIAS ENTRE MI LISTA DE FRECUENCIA Y LA DE BUCKWALTER Y PARKINSON

Una vez preparadas ambas listas de frecuencia como quedó explicado (cap. IX, apart. E y F) calculé el número total de lemas coincidentes, que resultó ser 1895; por

tanto, el 88% de los lemas de la lista Familiar figura en la lista de Buckwalter y Parkinson. Dicho de otro modo, el 12% de los aproximadamente 2.000 lemas más frecuentes en la novela contemporánea árabe no se encuentran entre los aproximadamente 5.000 lemas más frecuente de la lengua árabe general estándar contemporánea, si consideramos que tanto el corpus de Buckwalter y Parkinson como el de Familiar son, efectivamente, representativos de esos ámbitos; o, quizás, podríamos aventurar que esa diferencia es la que media entre el léxico de la novela árabe contemporánea y el del ámbito general, que es el mejor representado en Buckwalter y Parkinson. El dato, en cualquier caso, indica, como se verá más adelante, que el vocabulario utilizado en la novela árabe contemporánea muestra rasgos distintivos.

Sin embargo, considerando la muy dispar extensión de ambas listas, resulta mucho más conveniente calcular el grado de coincidencia por franjas léxicas para confirmar esas peculiaridades. Para ello, y guiándome por los resultados preliminares obtenidos por Ricks (2015) referente al ritmo de adquisición léxica en estudiantes de ALE¹²⁵ (cap. III, apart. A), opté por comparar franjas de 500 lemas. Al ser preciso comparar franjas completas, esto significó que fue necesario limitar el análisis a los 2.000 primeros lemas. La ventaja de fijar las franjas en grupos de 500 lemas es, fundamentalmente, que permite estudiar las características semánticas del léxico que podría aprender un estudiante de ALE por cuatrimestres de 6 créditos ECTS¹²⁶.

La siguiente tabla muestra los resultados de comparar ambas listas por franjas de 500 lemas¹²⁷:

Franja	Ranquin	Nº de lemas coincidentes	% de coincidencia
1	1-500	317	63%
2	501-1.000	116	23%
3	1.001-1.500	72	14%
4	1.501-2.000	58	12%

Tabla 4: Coincidencia del léxico literario y el léxico general en lengua árabe (2.000 lemas desglosados en franjas de 500)

¹²⁵ 800-1.000 lemas por año académico.

¹²⁶ Según la normativa del Espacio Europeo de Educación Superior, 6 créditos son 60 horas de clase por semestre (que en la realidad son cuatro meses) e implica unas 150 horas de estudio por parte del alumno. Ver <http://eees.universia.es/preguntas-frecuentes/conceptos-basicos/>

¹²⁷ La comparación fue posible gracias al diseño de una tabla Excel dinámica que permite comparar elementos por franjas léxicas.

Como puede observarse, la primera es la franja con mayor coincidencia (63%). Esto no es de extrañar, pues el vocabulario más frecuente en una lengua, con independencia del género textual de la muestra, debe ser en gran medida coincidente, puesto que para empezar incluye un porcentaje muy importante de palabras formato o “lemas funcionales” (Nation 2013:18).

A partir de la segunda franja, sin embargo, el grado de coincidencia decrece de forma muy acusada (23%, 15% y 12%, respectivamente). Estos datos parecen confirmar que el léxico utilizado en la novela árabe contemporánea es, efectivamente, idiosincrásico y difiere, al menos en las franjas estudiadas, del vocabulario frecuente de la lengua general o, en todo caso, del corpus de Buckwalter y Parkinson.

Un análisis minucioso de los lemas más frecuentes en ambas listas desvela datos muy interesantes acerca del contenido semántico del léxico representado por ambos corpora, lo que demuestra su distinta naturaleza. A continuación, destaco algunos resultados del análisis semántico de ambas listas en los lemas más y menos frecuentes y en distintas categorías léxicas.

1. Los 10 lemas más frecuentes

Como suele ocurrir en cualquier lista de frecuencia elaborada a partir de un corpus representativo, el léxico más frecuente está compuesto, en su mayoría, por lemas funcionales¹²⁸. En los dos corpus comparados se trata exactamente de los mismos lemas, aunque ocupen puestos ligeramente distintos en el ranquin:

Ranquin	Buckwalter y Parkinson	Familiar
1	<i>al</i>	<i>al</i>
2	<i>wa</i>	<i>wa</i>
3	<i>fī</i>	<i>bi</i>

¹²⁸ Para los objetivos del presente estudio, este grupo de lemas incluirá todos aquellos lemas que no pertenezcan a las categorías de sustantivo, verbo, adjetivo y adverbio, y también algunos lemas que sintácticamente tienen la función de preposiciones, aunque en árabe sean técnicamente sustantivos (como *ba'da*, *fawqa*, etc.). Así lo hacen también Buckwalter y Parkinson (2011:515-518).

4	<i>Min</i>	<i>fī</i>
5	<i>Li</i>	<i>li</i>
6	<i>bi</i>	<i>min</i>
7	<i>‘alà</i>	<i>‘alà</i>
8	<i>anna</i>	<i>kāna</i>
9	<i>ilà</i>	<i>anna</i>
10	<i>kāna</i>	<i>ilà</i>

Tabla 5: Los 10 lemas más frecuentes (Buckwalter y Parkinson vs. Familiar)

Un dato que destaca es que el lema *kāna* (“ser”) supera en dos puestos en mi lista a su posición Buckwalter y Parkinson. Esto pudiera estar relacionado con el hecho de que el discurso literario es narrativo y requiere de una mayor cantidad de verbos que en otros discursos, como el argumentativo. Recordemos, además, que una diferencia de dos puestos en ranquin significa una diferencia en miles de ocurrencias; de hecho, una vez normalizadas las listas, el lema mostró un 53% más de frecuencias en Familiar, lo que resulta estadísticamente muy significativo y revelador del carácter más narrativo de mi corpus y más argumentativo del corpus de Buckwalter y Parkinson. Como se verá, esto será confirmado más adelante.

En esta misma línea, es interesante observar que las preposiciones ocupan ranquins distintos en ambas listas. El hecho podría estar relacionado con la naturaleza de los verbos utilizados en ambos corpora y sus correspondientes frecuencias, puesto que, recordemos, en árabe muchos verbos requieren complementos preposicionales.

En cualquier caso, lo más interesante en la Tabla 5 es que demuestra que los lemas más frecuentes en cualquier corpus representativo son coincidentes.

2. Los 30 sustantivos más frecuentes

Para elaborar esta lista se eliminaron algunos lemas que forman parte de la categoría nominal en árabe como, por ejemplo, los cuantificadores¹²⁹ y sustantivos con

¹²⁹ Por ejemplo, *kull* (“todo”), *yamī‘* (“todos”), *ba‘ d* (“algunos”).

uso frecuente como partículas temporales¹³⁰. El objetivo, pues, es comparar sustantivos con un marcado contenido semántico para desvelar, a grandes rasgos, algunos de los temas comúnmente tratados en ambos corpora.

Nº ¹³¹	Buckwalter y Parkinson	Familiar
1	<i>allāh</i> ¹³²	<i>šay'</i>
2	<i>yawm</i>	<i>nafs</i>
3	<i>šay'</i>	<i>yawm</i>
4	<i>nafs</i>	<i>marra</i>
5	<i>ra'īs</i>	<i>'ayn</i>
6	<i>'amal</i>	<i>aḥad</i>
7	<i>dawla</i>	<i>bayt</i>
8	<i>'ām</i>	<i>yad</i>
9	<i>aḥad</i>	<i>šawt</i>
10	<i>ajj</i>	<i>yimm</i>
11	<i>sana</i>	<i>waḡh</i>
12	<i>amr</i>	<i>abb</i>
13	<i>quwwa</i>	<i>raḡul</i>
14	<i>marra</i>	<i>amr</i>
15	<i>abb</i>	<i>makān</i>
16	<i>mintaqā</i>	<i>bāb</i>
17	<i>ḥaqq</i>	<i>ḥayāt</i>
18	<i>mawdū'</i>	<i>ra's</i>
19	<i>wazīr</i>	<i>ibn</i>
20	<i>ism</i>	<i>ism</i>
21	<i>raḡul</i>	<i>waqt</i>

¹³⁰ Por ejemplo, *ḥīn* (“momento”) y *ba'd* (“posterioridad”) utilizados con el significado de “cuando” y “después”, respectivamente. Aunque algunos diccionarios categorizan estos lemas como preposición, morfológicamente son sustantivos.

¹³¹ Esta numeración no refleja el ranquin de los lemas en sus correspondientes listas de frecuencia. Su único propósito es aclarar su orden relativo de aparición en las listas. Aplíquese lo mismo a las tablas siguientes.

¹³² Aunque el lema *'Allāh* pierde en muchas expresiones orales su contenido semántico literal, su inclusión en la lista de Buckwalter y Parkinson no sorprende si se tiene en cuenta que su corpus contiene textos religiosos, además lenguaje oral y foros online, en un porcentaje que puede llegar a sumar el 30% aproximadamente (Buckwalter y Parkinson 2011:3).

22	<i>maʿyilis</i>	<i>ān</i>
23	<i>ʿālam</i>	<i>allāh</i>
24	<i>ḥayāt</i>	<i>gurfa</i>
25	<i>balad</i>	<i>arḍ</i>
26	<i>waqt</i>	<i>ṣūra</i>
27	<i>ḥukūma</i>	<i>lahza</i>
28	<i>bayt</i>	<i>qalb</i>
29	<i>qalb</i>	<i>ṭarīq</i>
30	<i>nās</i>	<i>sayyāra</i>

Tabla 6: Los 30 sustantivos más frecuentes (Buckwalter y Parkinson vs. Familiar) con resaltado de los lemas coincidentes

Lo primero que se observa es una serie de lemas de alta frecuencia que son comunes a ambas listas. Algunos, como *marra*, *naḥs*, *aḥad* y *ṣayʿ*, tienen un contenido semántico general. Otros lemas tienen un contenido más específico como *waqt*, *bayt* y *qalb*, aunque en el caso del lema *bayt* es claro que en mi corpus tiene una frecuencia muchísimo más alta.

Sin embargo, existe otro grupo de lemas que refleja un contenido realmente muy distinto. Por ejemplo, entre los sustantivos más frecuentes en la novela árabe contemporánea aparecen lemas que hacen referencia al cuerpo humano (*ʿayn*, *waʿyḥ*, *ṣawt*, *yad* y *raʿs*) y a la vida cotidiana (*bayt*, *bāb*, *gurfa*, *arḍ*, *ṣūra*, *sayyāra*, *makān*). Por el contrario, entre el vocabulario más frecuente recogido por Buckwalter y Parkinson, se encuentran lemas de naturaleza política (*raʿīs*, *dawla*, *ḥaqq*, *ʿālam*, *mauwdūʿ*, *wazīr*, *ḥukūma*, *balad*, *maʿyilis*). Aunque el diccionario de Buckwalter y Parkinson pretende ser reflejo de la lengua árabe en general, el vocabulario más frecuente incluido en dicha lista muestra las consecuencias de que el 36% de los textos utilizados como muestra sean de naturaleza periodística, como afirman los propios autores (Buckwalter y Parkinson 2011:3).

Otro grupo de lemas interesante es el utilizado para referirse al concepto de tiempo, pues ambas listas de frecuencia comparten *waqt*, un término muy genérico. Sin embargo, entre los lemas más frecuentes de Familiar aparecen *ān* y *lahza*, más típicos de la narración literaria, mientras que en Buckwalter y Parkinson aparecen *sana* y *ʿām*, generalmente utilizados en el discurso argumentativo para describir acontecimientos.

3. Los 30 verbos más frecuentes

La siguiente tabla muestra, resaltados, los verbos comunes a ambas listas, junto a su correspondiente ranquin:

Nº	Buckwalter y Parkinson	Ranquin	Familiar	Ranquin
1	<i>kāna</i>	10	<i>kāna</i>	8
2	<i>qāla</i>	15	<i>qāla</i>	16
3	<i>ʿarafa</i>	48	<i>ʿarafa</i>	34
4	<i>tamma</i>	74	<i>ra`ā</i>	42
5	<i>ra`ā</i>	76	<i>āda</i>	45
6	<i>amkana</i>	93	<i>laysa</i>	63
7	<i>qāma</i>	105	<i>arāda</i>	67
8	<i>yā`a</i>	110	<i>waýada</i>	68
9	<i>waýada</i>	119	<i>naẓara</i>	81
10	<i>šā`a</i>	122	<i>ša`ara</i>	82
11	<i>arāda</i>	124	<i>sa`ala</i>	86
12	<i>āda</i>	125	<i>badā</i>	89
13	<i>anā</i>	137	<i>sami`a</i>	90
14	<i>amila</i>	139	<i>ajaḍa</i>	97
15	<i>sami`a</i>	156	<i>istaṭā`a</i>	98
16	<i>akkada</i>	164	<i>bada`a</i>	99
17	<i>bada`a</i>	169	<i>fa`ala</i>	105
18	<i>ašbaḥa</i>	170	<i>jaraýa</i>	108
19	<i>šāra</i>	171	<i>taraka</i>	111
20	<i>wašala</i>	179	<i>wašala</i>	119
21	<i>ḥakā</i>	184	<i>aḥabba</i>	120
22	<i>ajaḍa</i>	197	<i>šāra</i>	123
23	<i>istaṭā`a</i>	212	<i>dajala</i>	125
24	<i>zāla</i>	215	<i>yā`a</i>	130
25	<i>waýaba</i>	230	<i>ṭalaba</i>	132

26	<i>qaddama</i>	246	<i>aṣbaḥa</i>	135
27	<i>aḍāfa</i>	253	<i>ḍahaba</i>	136
28	<i>baqiya</i>	262	<i>baqiya</i>	145
29	<i>ḍakara</i>	265	<i>ḡalasa</i>	146
30	<i>ašāra</i>	280	<i>zāla</i>	147

Tabla 7: Los 30 verbos más frecuentes (Buckwalter y Parkinson vs. Familiar), con resaltado de los lemas coincidentes

El primer dato que se obtiene de comparar los 30 verbos más frecuentes en ambas listas es que los cuatro primeros (*kāna*, *qāla*, *ʿarifa* y *raʿā*) se repiten, e incluso su ranquin es casi idéntico en ambas listas, lo que significa que se trata de lemas de muy alta frecuencia en la mayoría de los textos de producción escrita en MSA, con independencia de su género. También se observa que algunos verbos auxiliares son de alta frecuencia, como, por ejemplo, *zāla*, *aṣbaḥa*, *badaʿa*, *ajaḍa*, *istaṭāʿa*. En total, se observa que solo el 56% de los 30 lemas verbales más frecuentes en ambas listas es coincidente.

Un análisis detallado muestra que la lista Familiar contiene verbos que se utilizan para describir las interacciones que ocurren en la vida cotidiana (*naḡara*, *ḍahaba*, *dajala*, *šaʿara*, *saʿala*, *ṭalaba*, *aḡabba*, entre otros). Dichos lemas no se encuentran entre los verbos más frecuentes de Buckwalter y Parkinson, que, sin embargo, incluye otro tipo de verbos típicos de los géneros argumentativos, como *waḡaba*, *tamma*, *amkana* y *ʿanā*. Además, en la lista de Buckwalter y Parkinson encontramos un grupo de lemas referentes a la comunicación propios del lenguaje periodístico, como *akkada*, *ašāra*, *aḍāfa* y *ḍakara*, mientras que en Familiar no aparecen. De nuevo, parece haber indicios de que, o el género literario recurre a lemas de una naturaleza particular, o el listado de Buckwalter y Parkinson no es de naturaleza tan general como se le supone, o ambas cosas.

Finalmente, cabe señalar que el verbo núm. 30 en Buckwalter y Parkinson ocupa el ranquin 280 en la lista general, mientras que su equivalente en Familiar ocupa el ranquin 147. Tal diferencia se observa, asimismo, en otras franjas léxicas. De hecho, en los 1.000 primeros lemas de Buckwalter y Parkinson ocurren 136 verbos en total: 64 en la primera franja de 500 lemas y 72 en la segunda franja. Por el contrario, en los

1.000 primeros lemas de Familiar ocurren 254 verbos en total: 117 verbos en la primera franja y 137 en la segunda franja. Es decir, casi el doble. Todo esto apunta a que los verbos son mucho más comunes en el género narrativo que en los géneros informativos y argumentativos que están principalmente representados en Buckwalter y Parkinson.

4. Los 30 adjetivos más frecuentes

En la elaboración de esta lista se incluyeron, además de los adjetivos comunes, los comparativos, los colores, y el ordinal “primero”. Sin embargo, se eliminó el lema *dā'im*, ya que generalmente se utiliza en la forma adverbial *dā'iman*, y se descartaron los numerales ordinales a partir de *tānī*. Asimismo, eliminé los lemas *ḥuluw* y *tayyeb* en Buckwalter y Parkinson porque pertenecen al árabe moderno estándar y también al árabe dialectal. Ciertamente, este razonamiento debiera habernos inducido, quizá, a eliminar algunos otros lemas; sin embargo, en este caso concreto, su aparición inicial en los lemas más frecuentes, columna de Buckwalter y Parkinson, me pareció consecuencia de su altísima frecuencia en la mayoría de los dialectos árabes y, por tanto, de la inclusión de un porcentaje de lenguaje hablado en su corpus, lo que no es el caso en Familiar¹³³.

La siguiente tabla muestra, resaltados, los adjetivos comunes a ambas listas, junto a su correspondiente ranquin:

Nº	Buckwalter y Parkinson	Familiar
1	<i>ājar</i>	<i>ājar</i>
2	<i>awwal</i>	<i>akṭar</i>
3	<i>‘arabī</i>	<i>awwal</i>
4	<i>kaṭīr</i>	<i>kaṭīr</i>
5	<i>ḡadīd</i>	<i>ṣagīr</i>
6	<i>akṭar</i>	<i>ṭawīl</i>

¹³³ Recordemos, además, que en el corpus Familiar se eliminaron todos los diálogos escritos en dialectal.

7	<i>kabīr</i>	<i>kabīr</i>
8	<i>amrīkī</i>	<i>lālīl</i>
9	<i>‘āmm</i>	<i>ŷadīd</i>
10	<i>filīṣṭīnī</i>	<i>ba ‘īd</i>
11	<i>siyāsī</i>	<i>ajīr</i>
12	<i>dawlī</i>	<i>abyaḍ</i>
13	<i>waṭanī</i>	<i>akbar</i>
14	<i>jāṣ</i>	<i>garīb</i>
15	<i>‘irāqī</i>	<i>waḥīd</i>
16	<i>islāmī</i>	<i>qadīm</i>
17	<i>isrā ‘īlī</i>	<i>jāṣ</i>
18	<i>lubnānī</i>	<i>ŷamīl</i>
19	<i>akbar</i>	<i>a ‘lā</i>
20	<i>māḍī</i>	<i>‘arabī</i>
21	<i>muttaḥīd</i>	<i>qarīb</i>
22	<i>ajīr</i>	<i>aḥmar</i>
23	<i>muslim</i>	<i>qādim</i>
24	<i>ṣagīr</i>	<i>‘ālī</i>
25	<i>muhimm</i>	<i>šāb</i>
26	<i>mujtalīf</i>	<i>aqall</i>
27	<i>sābiq</i>	<i>bārid</i>
28	<i>iqtisādī</i>	<i>ṭaqīl</i>
29	<i>ḥālī</i>	<i>ŷayyīd</i>
30	<i>‘ālamī</i>	<i>māḍī</i>

Tabla 8: Los 30 adjetivos más frecuentes (Buckwalter y Parkinson vs. Familiar) con resal-
tado de los lemas coincidentes

Lo primero que llama la atención es que el primer lema en ambas listas es *ājar*¹³⁴ y que hay un grupo de lemas comunes a ambas listas (el 40%).

¹³⁴ El dato es interesante, pues este mismo lema constituye el adjetivo más frecuente en el British National Corpus (http://ucrel.lancs.ac.uk/bncfreq/lists/5_3_all_rank_adjective.txt), algo que podría ser común a otras lenguas.

El dato más revelador, sin embargo, es que la lista de Buckwalter y Parkinson incluye muchos adjetivos *nisba* referentes al concepto de identidad (nacionalidad y religión), además de lemas típicos de géneros argumentativos, como *muttaḥid*, *dawlī*, *islāmī*, *iqtiṣādī*, y *siyāsī*, mientras que lemas de esta naturaleza están ausentes en Familiar, salvo *‘arabī*.

Mi lista, por su parte, incluye adjetivos que pueden considerarse de uso cotidiano, incluidos numerosos pares de antónimos comunes, como, por ejemplo, “más-menos”, “mucho-poco”, “grande-pequeño”, “viejo-nuevo”, “lejos-cerca”, “presente-pasado” y “primero-último”. En mi lista también aparecen dos colores: *abyaḍ* y *aḥmar*. Todo ello parece indicar, de nuevo, que mi corpus recurre a un léxico más descriptivo y típico de la vida cotidiana.

5. *Lemas exclusivos del corpus Familiar*

Al comparar las dos listas de frecuencia se observa que 256 de los 2.151 lemas de Familiar no figuran entre los 4.852 lemas de Buckwalter y Parkinson. La mayoría de esos lemas se sitúan en la tercera franja léxica de Familiar, es decir, a partir del ranquin 1000¹³⁵.

En la primera franja léxica de Familiar, por ejemplo, solo hay un lema que no ocurre en Buckwalter y Parkinson, a saber, el verbo auxiliar *rāḥa* que en el discurso literario se utiliza para señalar el comienzo de una acción. Un ejemplo tomado del corpus analizado es *tumma rāḥat taḥki* (“empezó a narrar”). El dato es interesante porque hay otros lemas que indican el comienzo de una acción en árabe (*bada’a* y *ajada*), que aparecen en ambas listas, pero cuyo ranquin es mucho más alto en Familiar¹³⁶. En cualquier caso, el ejemplo ofrecido refuerza los indicios existentes de que el corpus de Familiar tiene una naturaleza semántica más cotidiana que el de Buckwalter y Parkinson por el tipo de vocabulario frecuente que presenta.

¹³⁵ Véase una lista completa de esos lemas en el Apéndice H.

¹³⁶ Esto puede verse en Tabla 7. Sería interesante averiguar, en un futuro, qué porcentaje exacto de esos lemas tiene un uso auxiliar y cuántos tienen un contenido semántico completo.

Entre el vocabulario propio de la lista Familiar, destacan 8 lemas referentes a la vestimenta, que son, en orden de frecuencia, *nazzāra*, *anīq*, *šāl*, *ŷilbāb*, *mandīl*, *badla*, *‘imāma*, *fustān* y *qubba ‘a*.

Asimismo, destaca que haya tres lemas referentes a colores que no aparecen en la lista de Buckwalter y Parkinson: *dākin*, *ramādī* y *sumra*. Se trata de un vocabulario que se utiliza para describir la complejión de las personas y los objetos; lo cual parece ser típico del discurso narrativo literario.

Destaca también que entre esos 256 lemas haya varios que hacen referencia a partes del cuerpo humano, con una frecuencia relativamente alta si se compara con el ranquin de esos mismos lemas en Buckwalter y Parkinson. La siguiente tabla compara el ranquin de esos lemas:

Lema	Ranquin en Familiar	Ranquin en Buckwalter y Parkinson ¹³⁷
<i>rukba</i>	793	6780
<i>fajd</i>	913	(-)
<i>jadd</i>	1038	6053
<i>sāq</i>	1139	(-)
<i>lihya</i>	1421	(-)
<i>šārib</i>	1740	(-)
<i>daqn</i>	1983	(-)

Tabla 9: Partes del cuerpo en Familiar vs. Buckwalter y Parkinson

Como puede verse en Apéndice H, muchos de esos 256 lemas hacen referencia a acciones u objetos propios de la vida cotidiana y de las interacciones humanas.

¹³⁷ (-) indica que el lema no se encuentra en Buckwalter y Parkinson. Los lemas con cifra, aunque no llegaron a figurar en el diccionario, figuran en una tabla donde Buckwalter y Parkinson (2011:16) listan todos los lemas referentes a partes del cuerpo.

6. Lemas propios del lenguaje periodístico en Buckwalter y Parkinson

Que el corpus de Buckwalter y Parkinson contenga un 36% de textos periodísticos se evidencia al comparar la suya con mi lista de frecuencia; así, los 400 lemas más frecuentes en Buckwalter y Parkinson se caracterizan por su alto contenido político y mediático. La siguiente tabla muestra un conjunto de lemas de alta frecuencia en Buckwalter y Parkinson que no figuran en Familiar:

Lema	Ranquin en Buckwalter y Parkinson	Ranquin en Familiar
<i>ḥukūma</i>	104	(-)
<i>šArika</i>	114	(-)
<i>duwalī</i>	115	(-)
<i>islāmī</i>	175	(-)
<i>isrā`ilī</i>	190	(-)
<i>muttaḥid</i>	204	(-)
<i>idāra</i>	224	(-)
<i>natīya</i>	227	(-)
<i>milyūn</i>	248	(-)
<i>wilāya</i>	256	(-)
<i>qiṭā`</i>	261	(-)
<i>ittiḥād</i>	279	(-)
<i>iqtiṣādī</i>	293	(-)
<i>muwātin</i>	319	(-)
<i>mušāraka</i>	321	(-)
<i>islām</i>	346	(-)
<i>mu`tamar</i>	366	(-)
<i>da`m</i>	374	(-)
<i>dūlār</i>	376	(-)
<i>intijāb</i>	380	(-)
<i>ta`āwun</i>	397	(-)

Tabla 10: Lemas de alta frecuencia en Buckwalter y Parkinson que no figuran en Familiar

En total, más de 700 lemas incluidos en las 4 primeras franjas léxicas (2.000 lemas) de Buckwalter y Parkinson se podrían considerar propios del género argumentativo de tipo periodístico.

7. *Discrepancias léxicas recurrentes*

Las secciones anteriores destacan algunas semejanzas y discrepancias léxicas en la franja de los 500 lemas más frecuentes, pero también pueden observarse patrones de comportamiento léxico similares en otras franjas. Por ejemplo, la siguiente tabla ofrece el ranquin de algunos lemas que en Buckwalter y Parkinson son de frecuencia media, y su correspondiente ranquin en Familiar.

Lema	Ranquin en Buckwalter y Parkinson	Ranquin en Familiar
<i>ibtāsama</i>	1600	246
<i>ašbaha</i>	1524	271
<i>gādara</i>	1652	285
<i>šaraja</i>	1760	296
<i>qahwa</i>	1814	321
<i>malbas</i>	1697	328
<i>šifa</i>	1876	360
<i>ibitsāma</i>	1925	363
<i>sarīr</i>	1893	372
<i>ajrayā</i>	1991	396
<i>taqīl</i>	1648	420
<i>ḥabb</i>	1620	422
<i>nādir</i>	1926	459
<i>uḍun</i>	1751	464
<i>raddada</i>	1934	474
<i>iḥsās</i>	1592	484

<i>arsala</i>	1588	485
<i>ṭāwila</i>	1899	493
<i>manaḥa</i>	1918	520
<i>zāra</i>	1637	529
<i>zāwiya</i>	1638	530
<i>‘aḡūz</i>	1775	539
<i>ḡiwār</i>	1685	556
<i>ḥāra</i>	1985	557
<i>šāhada</i>	1771	564
<i>gāba</i>	1585	567

Tabla 11: Lemas de uso cotidiano en Buckwalter y Parkinson vs. Familiar

Como puede observarse, Tabla 11 incluye, por un lado, lemas que podrían decirse de uso cotidiano y propios de las interacciones personales. Por ejemplo, *sarīr*, *‘aḡūz*, *ṭāwila*, *ibtāsama*, *qahwa*, *iḥsās*, *malbas*, entre otros. Se trata, en realidad, de un vocabulario previsible en la narrativa.

Por otro lado, hay otro grupo de lemas cuyo contenido semántico es más general, como en el caso de *arsala*, *ḡiwār*, *zāra*, *manaḥa*, y *gādar*. Este tipo de palabras puede utilizarse, en principio, en contextos y discursos de diversa naturaleza. Sin embargo, su mayor frecuencia en Familiar indicaría que se trata de lemas de un uso muy frecuente en la novela árabe contemporánea. De hecho, si observamos algunos ejemplos tomados del corpus Familiar, encontramos que este vocabulario se utiliza para describir acciones cotidianas:

<i>lam yazur ‘ā’ilatahu</i>
<i>wa lākinnaḥu zaarahā yawm l-jamīs</i>
<i>tarāhu kullamā zāra l-diyār</i>
<i>rakibat l-qitār ‘ilā nyūyurk li tazūr šarīkahā</i>
<i>wa kuntu zurtuhā marra wāḥida qabla zawāyina</i>

Tabla 12: Uso del lema *zāra* en el corpus Familiar

<i>wa ba 'da 'an gādarat arsalat lanā wajbat ṭa 'ām</i>
<i>al- ṭuyūr al-katīfa l-latī gādarat a 'shāshahā</i>
<i>qabla 'an yugādira 'alā matn bājira imlāqa</i>
<i>tumma taqtari hu 'an nugādira l-maqhà</i>
<i>gādarnā l-tā 'ira</i>

Tabla 13: Uso del lema *gādara* en el corpus Familiar

<i>faḍḍaltu l-ḍahāb li-l- 'ayš 'ilā yiwār wālidī</i>
<i>yālasa fī l- yiwār</i>
<i>tasallaltu 'ilā l-firāš biyiwāriha</i>
<i>fa iṣṭahabahā fī nuzha fī l-yiwār</i>
<i>yālasat 'ilā yiwār 'ammī 'āṭif</i>

Tabla 14: Uso del lema *yiwār* en el corpus Familiar

9. Conclusión

Las comparaciones anteriores apuntan a dos conclusiones principales:

- El léxico utilizado en la novela árabe contemporánea difiere en gran medida del vocabulario incluido en el diccionario de Buckwalter y Parkinson en todas las franjas léxicas. El léxico utilizado en la novela árabe contemporánea parece ser idiosincrásico y de un contenido semántico que gira en torno a la vida cotidiana en mayor medida que el incluido en Buckwalter y Parkinson.
- El diccionario de Buckwalter y Parkinson contiene lemas de alta frecuencia propios del género argumentativo, particularmente del lenguaje periodístico y político, en mayor medida que Familiar, debido a la composición de su corpus.

B. LOS LEMAS MÁS FRECUENTES

El segundo objetivo de la presente investigación era elaborar una lista que incluyera los 3.500 lemas más frecuentes en la novela árabe contemporánea. El tamaño de la muestra analizada, como ya se ha indicado, no permitió extraer una lista de frecuencia fiable de tal tamaño, aunque sí pudo extraerse una lista de 2.000 lemas representativos de la novela árabe contemporánea, por las razones expuestas en el capítulo anterior.

Tal y como quedó demostrado en el apartado anterior, los lemas más frecuentes en Familiar están compuestos por lemas funcionales, tal y como cabría esperar en cualquier lista de frecuencia extraída a partir de un corpus representativo.

Además, a lista de frecuencia está compuesta, en su gran mayoría, por lemas cuyo contenido hace referencia a la vida cotidiana, en todas las categorías léxicas. Por ejemplo, es frecuente el vocabulario que hace referencia a la familia, el hogar, las interacciones humanas, el cuerpo humano, los colores, e incluso los verbos auxiliares que indican comienzo de acción, típicos del discurso narrativo.

Por poner un último ejemplo, destaca que entre los 270 lemas más frecuentes haya palabras que pertenecen a este tipo de vocabulario, y que en Buckwalter y Parkinson sean de menor frecuencia. La siguiente tabla ofrece algunos ejemplos¹³⁸:

Lema	Ranquin en Familiar	Ranquin en Buckwalter y Parkinson
<i>sha'ara</i>	82	505
<i>gurfa</i>	96	534
<i>yasad</i>	129	914
<i>yalasa</i>	146	876
<i>intazara</i>	149	671
<i>nazra</i>	151	1011
<i>'ā'ila</i>	196	870
<i>šamt</i>	201	812

¹³⁸ La lista completa de lemas en Familiar y su correspondiente ranquin en Buckwalter y Parkinson se encuentra en Apéndice I.

<i>imra`a</i>	204	1019
<i>nawm</i>	213	947
<i>ḥulm</i>	223	638
<i>ḍaḥika</i>	224	1017
<i>aḡāba</i>	226	1065
<i>aḥassa</i>	227	1400
<i>fuḡ`a</i>	229	1166
<i>kāda</i>	233	4718
<i>ibtasama</i>	246	1600
<i>rā`iḥa</i>	256	1464
<i>tawb</i>	267	1454

Tabla 15: Frecuencia de lemas de contenido semántico cotidiano en Familiar

Finalmente, cabe señalar que, en el ámbito del registro literario, se observa que entre el vocabulario más frecuente en Familiar se encuentran algunos lemas propios del lenguaje figurativo. Destacan, por ejemplo, lemas típicamente utilizados en el símil árabe (*tašbīh*):

Lema	Ranquin en Familiar	Ranquin en Buckwalter y Parkinson
<i>ka</i>	24	829
<i>miṭla</i>	72	89
<i>ka`anna</i>	114	314
<i>ašbaha</i>	274	1525

Tabla 16: Ranquin de lemas típicos del lenguaje figurativo en Familiar vs. Buckwalter y Parkinson

CAPÍTULO XI: CONCLUSIONES, APLICACIONES DIDÁCTICAS Y FUTURAS LÍNEAS DE INVESTIGACIÓN

A. CONCLUSIONES

La presente investigación se llevó a cabo con el ánimo de contribuir a la mejora del material didáctico disponible en el ámbito de la lectura extensiva del género narrativo en la EALE. Con vistas a lograr dicho propósito utilicé métodos propios de la lingüística de corpus para elaborar una lista de frecuencia léxica (cap. V) que pueda guiar a los especialistas en el desarrollo de lecturas literarias graduadas (cap. IV) destinadas a distintos niveles de competencia lingüística.

El objetivo principal de esta investigación pudo lograrse parcialmente. Por una parte, ha quedado demostrado que el corpus recopilado para la extracción de mi lista de frecuencia es sin duda representativo de la novela árabe contemporánea (cap. VII y VIII). Por otra, aunque no pudo obtenerse una lista fiable de los 3.500 lemas más frecuentes, como inicialmente pretendí, se han tenido por válidos los primeros 2.000 lemas.

El segundo objetivo era averiguar el grado de similitud existente entre una lista de frecuencia extraída a partir de un corpus literario representativo y la lista seminal de Buckwalter y Parkinson. Como ha quedado demostrado, las diferencias léxicas son importantes y han desvelado dos aspectos relevantes: 1) el léxico utilizado en la novela árabe contemporánea es idiosincrásico, y 2) todo parece indicar que el corpus utilizado por Buckwalter y Parkinson para elaborar su diccionario de frecuencia no es del todo representativo de la lengua general, ya que contiene un alto porcentaje de léxico argumentativo propio de textos periodísticos y políticos.

Los resultados obtenidos de este estudio apuntan a que, para desarrollar lecturas graduadas de naturaleza literaria narrativa, es recomendable utilizar la lista Familiar o

cualquier otra lista extraída a partir de un corpus literario representativo, ya que dicho vocabulario está compuesto por lemas de un contenido semántico idiosincrásico. Este tipo de vocabulario propio del género novelístico se asocia a ámbitos semánticos de la vida cotidiana en mayor medida que el vocabulario recogido en Buckwalter y Parkinson. Así pues, la afirmación de Buckwalter y Parkinson (2011:1) de que su diccionario de frecuencia puede utilizarse como referencia para el desarrollo de materiales propios de la lectura extensiva debe relativizarse.

Sería recomendable utilizar la lista de Buckwalter y Parkinson, sin embargo, en el desarrollo de materiales dedicados a la enseñanza del lenguaje periodístico o, en general, de géneros más argumentativos e informativos que narrativos. Por tanto, para hacer un uso consecuente de esa lista de frecuencia en la selección de y/o elaboración de textos apropiados para el desarrollo de la destreza lectora el profesor de ALE debe tener claros los fines que persigue y el perfil de su alumnado.

En cualquier caso, la selección y la dosificación apropiada del léxico en las lecturas graduadas destinadas a la lectura extensiva permitiría desarrollar y fortalecer la destreza lectora de los estudiantes de ALE en su sentido más amplio, siempre y cuando los textos de lectura no incluyan más de un 5% de vocabulario desconocido (Nation 2013:206). Como quedó dicho (caps. III y IV), visualizar el vocabulario más frecuente de forma reiterada contribuiría a que el aprendiente: 1) desarrolle automaticidad en el reconocimiento visual del léxico, 2) incremente su velocidad lectora, y 3) continúe adquiriendo nuevo vocabulario de forma incidental a través de la lectura. En la actualidad, el material de ALE disponible no permite reconocer y comprender el *sight vocabulary* (cap. III, apart. C) típicamente utilizado en la narrativa árabe. En consecuencia, sería recomendable desarrollar textos graduados que se basen en la lista de frecuencia Familiar o en cualquier otra lista elaborada de forma similar.

Es más, el desarrollo y el uso de lecturas extensivas literarias en prosa desde los niveles iniciales podría constituir una estrategia didáctica eficaz en la superación de los obstáculos típicos que dificultan la decodificación visual del léxico árabe (cap. I), y en la formación de un léxico mental de uso frecuente (cap. III) que, más adelante, capacitaría a los estudiantes de ALE para leer novelas de forma eficaz. No hay que olvidar, tampoco, que el uso de material didáctico que contenga un *input* comprensible puede

contribuir a rebajar la ansiedad que suele acompañar al aprendiente de lenguas extranjeras (Horwitz et al, 1986).

Conviene añadir que, además de todos los beneficios que conlleva leer textos literarios en prosa (cap. II, apart. A), la integración del texto literario a través de narrativa graduada elaborada a partir de la lista de frecuencia Familiar (o cualquier otra lista elaborada de forma similar), tendría un impacto directo sobre la capacidad del aprendiente de ALE de comunicarse y mantener interacciones que giran en torno a la vida cotidiana, ya que, como quedó demostrado en el capítulo anterior, el léxico que aparece en la novela árabe contemporánea es generalmente de uso cotidiano y tiene una relación directa con las interacciones humanas, el hogar, la familia, el cuerpo y otras temáticas muy prácticas para el fomento de la capacidad comunicativa cotidiana.

Para que esto pueda ocurrir, habrá de integrarse en el currículum del ALE el vocabulario típico del género novelístico que ya se vio en el capítulo anterior: adjetivos comunes y sus correspondientes antónimos, los colores, un alto porcentaje de verbos referentes a interacciones humanas cotidianas, lemas propios del lenguaje figurativo, las partes del cuerpo humano, y una variedad de verbos auxiliares que indican el comienzo de acciones.

Finalmente, quisiera destacar que la extracción de datos referentes a la frecuencia léxica y su naturaleza semántica en la novela árabe contemporánea no habría sido posible si no fuera porque el corpus compuesto para la presente investigación es extremadamente representativo y fue diseñado siguiendo una metodología científica estricta, tal y como dictan las normas de la lingüística de corpus. De hecho, se trata posiblemente del primer corpus árabe cuya composición se describe en detalle en un estudio de esta naturaleza.

B. APLICACIONES DOCENTES

1. *Elaborar un diccionario de frecuencia*

La lista de frecuencia léxica obtenida como fruto de esta tesis puede utilizarse en la elaboración de un diccionario de frecuencia léxica que sirva como base para desarrollar lecturas narrativas graduadas destinadas a estudiantes de niveles A1-B1. La publicación de un diccionario de 2.000 lemas sería una contribución importante en el campo de la EALE, en particular para el desarrollo de material didáctico destinado a la lectura extensiva.

2. *Desarrollar y adaptar lecturas graduadas*

Disponer de un diccionario de frecuencia especializado en el género narrativo permitiría, por su parte, desarrollar lecturas graduadas interesantes y de calidad que emulen el lenguaje utilizado en la narrativa árabe contemporánea. Sería aconsejable que ese material cubriera fundamentalmente los niveles A1 y A2, debido a la gran escasez de material de lectura apropiado para los estudiantes de ALE en esos niveles, que además son los más usuales entre los estudiantes de árabe en España.

Mi propia experiencia en la adaptación de novelas árabes contemporáneas destinadas a la enseñanza de ALE (Familiar 2013, 2016) me reafirma en la convicción de que el uso de listas de frecuencia es un instrumento esencial en la elaboración de lecturas narrativas graduadas eficientes. A partir de ahora, el docente y el experto en el desarrollo de material didáctico podrá guiarse por datos léxicos fiables, en lugar de recurrir a su intuición lingüística, al elaborar lecturas graduadas.

La actual lista de frecuencia podría, por ejemplo, dividirse en franjas léxicas para desarrollar lecturas niveladas dirigidas a estudiantes de ALE e incluso para hablantes nativos. En cuanto a la simplificación de textos literarios originales, la lista podría servir de guía para adaptar dichos textos al nivel del aprendiente, aunque la dificultad radicaría en que el vocabulario habría de ser cotejado de forma manual para comprobar si el léxico utilizado en el texto original se encuentra en la lista de frecuencia. Esto podría solucionarse, sin embargo, si se desarrollara un interfaz de consulta

online donde el docente pudiera introducir el texto original con el propósito de analizar la frecuencia del léxico. Esto ya se ha hecho para el inglés y puede verse en Lextutor, una página web destinada a dicho objetivo¹³⁹.

C. FUTURAS LÍNEAS DE INVESTIGACIÓN EN LA EALE

1. Elaborar listas de frecuencia que cubran los niveles A1-C1

Los resultados obtenidos en esta investigación indican que es necesario utilizar una muestra lingüística mucho mayor que la utilizada en el presente estudio (90.000 palabras) para elaborar una lista de frecuencia estadísticamente fiable que permita desarrollar también materiales de lectura extensiva en el nivel B2 y C1.

Aunque los datos existentes acerca del tamaño mínimo requerido de un corpus para elaborar listas de frecuencia fiables son muy vagos (Miller 2015), es posible que sea necesario utilizar varios millones de palabras a partir de una muestra representativa de la población identificada. Un dato importante a tener en cuenta en este sentido es que, según el lexicógrafo Sinclair (2005), para estudiar el comportamiento de las palabras en todas sus facetas semánticas, gramaticales y sintácticas, estas deben aparecer en el corpus al menos 20 veces.

Para el propósito original marcado en la presente investigación, pues, sería conveniente pasar todas las novelas escaneadas y en formato PDF pertenecientes a mi población a formato .txt¹⁴⁰ para que en un futuro cercano sea posible seleccionar una muestra representativa a partir de la cual extraer una lista de frecuencia estadísticamente fiable con los 5.000 lemas más frecuentes en la novela árabe contemporánea. Esa cifra se considera relevante desde el punto de vista pedagógico, pues representa una barrera funcional que implica la capacidad de leer textos originales para una comprensión adecuada (Martinez y Schmitt 2015:451). Una vez elaborada una lista léxica de ese tamaño

¹³⁹ <http://www.lex Tutor.ca/vp/eng/>

¹⁴⁰ Esta tarea ya ha sido iniciada con la ayuda de una empresa especializada.

podrán desarrollarse también lecturas graduadas de narrativa árabe destinadas a los niveles de competencia B2-C1.

2. Elaborar listas de frecuencia especializadas

Otra línea de investigación interesante para desarrollar lecturas graduadas enfocadas a subgéneros narrativos concretos consistiría en agrupar las novelas de la población por temática y extraer listas de frecuencia que permitan elaborar novelas históricas, de guerra, de ciencia-ficción, etc. Como quedó dicho (cap.VII, apart.E.4), la variedad existente en la población permite manipular esa variable independiente e investigar el comportamiento léxico en contextos temáticos diferentes.

3. Aplicar fórmulas estadísticas para cuantificar el comportamiento léxico

En un contexto de investigación ideal, el desarrollo de listas de frecuencia fiables debe ir acompañado de métodos cuantitativos capaces de ofrecer datos objetivos acerca del comportamiento léxico (Cantos Gómez 2002, Gries 2015). Por un lado, sería muy recomendable aplicar fórmulas de dispersión léxica para cuantificar la distribución del vocabulario en el corpus de Familiar (Gries 2008, 2015), de modo que se neutralizara el efecto de ciertas muestras particulares en el corpus. Por otro lado, a la hora de comparar la lista Buckwalter y Parkinson con Familiar (o con cualquier otra lista similar), sería beneficioso aplicar alguna fórmula estadística que permita cuantificar las divergencias léxicas (Kilgarriff 2001).

4. Estudiar características gramaticales

El desarrollo de lecturas graduadas requiere tener en cuenta múltiples planos lingüísticos, no solo el plano léxico (cap. IV, apart. A). Por tanto, para elaborar materiales didácticos eficientes deben analizarse otras características del lenguaje literario utilizando metodología propia de la lingüística de corpus. Algunos aspectos gramaticales que se suelen estudiar en las lenguas europeas, y que sería beneficioso conocer en árabe también, es la frecuencia de los tiempos verbales, la oración relativa, la condicio-

nal, los verbos auxiliares, etc. Esto es importante por dos razones: a) es un tipo de investigación completamente ausente en el género narrativo árabe, y b) los datos obtenidos permitirían guiar el desarrollo de lecturas graduadas y otro tipo de material didáctico destinado a la EALE, e incluso cambiar el orden en que se introducen en la actualidad ciertas estructuras gramaticales en los libros de texto. El corpus utilizado para la presente investigación, sirve, en mi opinión, para realizar dicho estudio, debido a la estricta metodología seguida en la selección de las 144 muestras lingüísticas que lo componen (cap. VIII, apart. E).

5. *Estudiar aspectos léxicos propios del árabe*

Al ser el árabe una lengua semítica que utiliza la grafía árabe, el desarrollo de lecturas graduadas debe tener en cuenta algunas características idiosincrásicas de esa lengua que, de ser estudiadas, podrían beneficiar al aprendiente de ALE en el desarrollo eficiente y eficaz de la destreza lectora. A continuación, listo algunos aspectos que sería recomendable investigar a partir del corpus actualmente disponible:

- Frecuencia y uso de los homógrafos: Como ya se indicó (cap. I, apart. C.3), la homografía es uno de los mayores obstáculos en el desarrollo de la destreza lectora en árabe. Un aspecto que pude observar durante la elaboración de la lista de frecuencia Familiar y su comparación con Buckwalter y Parkinson es que la mayor parte de los homógrafos son lemas derivados de la misma raíz y representan el verbo y su correspondiente *maşḍar*. Sería interesante estudiar, por ejemplo, qué categoría léxica es más frecuente en el lenguaje narrativo (verbo o *maşḍar*), puesto que tiene implicaciones directas en la enseñanza y aprendizaje de las formas verbales y su frecuencia.
- Las raíces: El código java que se diseñó para MADAMIRA (cap. IX) en esta investigación extrae las raíces de los lemas, aunque su porcentaje de acierto no pudo ser comprobado por falta de tiempo. Sería deseable investigar la frecuencia de las raíces árabes en el género narrativo y su grado de productividad para, de este modo, elaborar una lista de raíces frecuentes.
- Aspectos morfológicos: Como ya se indicó (cap. IX, apart. A), los comandos ordenados a MADAMIRA para el análisis de mi corpus no incluyeron los enclíticos. Sería necesario incluir dichos morfemas en un futuro para estudiar, por

ejemplo, la frecuencia con que los novelistas árabes contemporáneos utilizan aspectos gramaticales como el dual, el plural femenino y sus correspondientes pronombres sufijos.

Todo lo anterior podría ofrecer información relevante acerca de cómo introducir, dosificar y priorizar distintos aspectos morfológicos propios del árabe en las lecturas graduadas y otro tipo de material didáctico. De hecho, la naturaleza derivativa del léxico árabe y la frecuencia de diferentes aspectos morfológicos debería ocupar un lugar central en la ecuación enseñanza-aprendizaje de ALE, pues un conocimiento morfológico sólido por parte de los aprendientes de ALE podría compensar el conocimiento *a priori* de que disponen los hablantes nativos como estrategia para decodificar los homógrafos y reconocer de forma efectiva y rápida una gran cantidad de léxico (cap. I, apart. C.3) y contribuirá a que el estudiante de ALE construya un léxico mental robusto (cap. III, apart. B).

6. *Unidades multiléxicas*

El concepto de unidades multiléxicas hace referencia a combinaciones de dos lemas o más (Nation 2016:71), y engloba distintos tipos de *enigramas*: colocaciones, coligaciones, concordancias, modismos, grupos léxicos, etc. Por poner un ejemplo, investigar la frecuencia de las colocaciones en la novela árabe contemporánea sería muy conveniente si se quiere desarrollar lecturas graduadas que emulen fehacientemente la novela árabe original. Un estudio de esta naturaleza podría beneficiarse enormemente de la línea de investigación iniciada por Santillán (2015) para clasificar los distintos tipos de colocaciones propios del discurso narrativo. La obtención de datos de ese tipo podría, incluso, facilitar el desarrollo de materiales dedicados a la enseñanza de la escritura creativa a estudiantes de ALE y a los propios hablantes nativos.

D. *FUTURAS LÍNEAS DE INVESTIGACIÓN EN OTRAS ÁREAS*

La disponibilidad de 8.5 millones de palabras provenientes de la novela árabe contemporánea (la población) en un formato que posibilite su análisis permitirá llevar

a cabo estudios en áreas interdisciplinarias aún sin explorar para el árabe. En realidad, un corpus especializado de este tamaño será el primero de tal naturaleza en árabe, y esto en sí constituirá un logro si se tiene en cuenta la dificultad de acceder a obras literarias originales para propósitos de investigación en lingüística de corpus.

Teniendo en cuenta las distintas variables que se identificaron en la población identificada para la presente investigación (cap. VII, apart. E), las posibles perspectivas de investigación incluyen, entre otras, las siguientes disciplinas:

1. Sociolingüística y discurso de género

Como quedó dicho (cap. VII, apart. E.5), la sociolingüista egipcia Reem Basiouney (2009) hizo un llamamiento al estudio del discurso narrativo escrito por hombres y mujeres para poder confirmar o rechazar la teoría de que el sexo del autor influye en las características léxico-gramaticales del texto producido. Teniendo en cuenta que los estudios realizados en inglés han aportado pruebas de la existencia de diferencias lingüísticas entre el lenguaje de hombres y mujeres, un corpus en árabe compuesto por 30 novelas escritas por mujeres y 79 novelas por hombres permitirá obtener datos muy fiables acerca del modo en que hombres y mujeres utilizan el léxico y distintas estructuras gramaticales y sintácticas en la novela contemporánea¹⁴¹. Sería interesante, por otra parte, analizar el discurso narrativo femenino y masculino a la hora de tratar ciertas temáticas, como pueden ser la muerte, el amor, el sexo, la política, etc. y observar si existen diferencias léxicas.

2. Variación diatópica y diglosia

Disponer de decenas de obras provenientes de distintos países árabes permitirá comparar el discurso narrativo producido en varias regiones árabes y estudiar el posible efecto de los dialectos sobre las estructuras lingüísticas del árabe moderno estándar en el género de la novela contemporánea. De confirmarse esta hipótesis, los hallazgos podrían tener implicaciones importantes en la enseñanza del árabe en todas sus variantes y, en consecuencia, en el desarrollo de materiales didácticos apropiados. Por ejemplo,

¹⁴¹ Un corpus de menor tamaño, incluso, podría permitir el estudio de este aspecto.

sería interesante cuantificar el léxico y las estructuras sintáctico-gramaticales que coinciden en ambos registros de la lengua (escrito vs. oral) e, incluso, el coincidente entre diversos registros novelísticos regionales, pues hasta el momento no existen datos acerca de esta cuestión.

3. *Estilometría*

El campo de la estilometría está estrechamente relacionado con varias disciplinas y, como se verá, tiene muchas aplicaciones. Una vez que se disponga de toda la población en formato .txt, se podrán estudiar, por ejemplo, marcadores léxicos de estilo que permitirían comparar a diversos autores incluidos en la población, pues algunos de ellos disponen de varias obras en la lista IPAF y NMML. Esta línea de investigación podría ser relevante para el campo de los estudios literarios.

También se podrían comparar algunos aspectos léxico-gramaticales propios de distintos novelistas y continuar la línea de investigación iniciada por García-Barrero et al (2016) en atribución de autoría. Esta área está estrechamente relacionada con la disciplina de lingüística de corpus forense y la posibilidad de poder descubrir casos de plagio en el ámbito de la novela árabe¹⁴².

El campo de la estilometría también permite estudiar las diferencias léxico-gramaticales y sintácticas a nivel diacrónico. Esto es interesante, pues permite comparar, por ejemplo, los cambios que han experimentado distintas estructuras en la novela árabe a lo largo de las décadas, y estudiar características propias de distintas épocas (Biber y Conrad 2009). Por poner un ejemplo, se podría comparar mi corpus con otro compuesto por la novela árabe moderna del siglo pasado.

¹⁴² En el ámbito de la literatura árabe contemporánea, los últimos años han sido testigo de varias acusaciones de plagio dirigidas a distintos tipos de autores (<http://intihalat.blogspot.ae/>).

4. Variación entre géneros

Si consideramos que la variación lingüística entre géneros es sistemática (Biber y Conrad 2009:4), y teniendo en cuenta que la EALE está siendo sometida a un alto grado de especialización, sería conveniente estudiar las características que distinguen el discurso narrativo de otros géneros como el periodístico, la prosa académica, la correspondencia formal, etc. De este modo, dispondríamos de datos que harían más eficiente la EALE, pues los aprendientes alcanzarían la destreza lingüística marcada como meta mucho más rápido y, por tanto, serían capaces de operar en árabe en distintos ámbitos profesionales de forma más eficaz.

5. Listas de frecuencia en otras lenguas y lectura extensiva

Finalmente, los hallazgos de la presente investigación tienen consecuencias para la disciplina de la enseñanza de lenguas extranjeras en su conjunto, y en especial en lo referente a la lectura extensiva. Como quedó dicho (cap. V, apart. A.2), la mayoría de las lecturas graduadas que existen en inglés se basan en listas de frecuencia que no han sido elaboradas de forma científica y que tampoco se corresponden con las listas extraídas a partir de corpora de la lengua general. Sería, pues, recomendable investigar si el léxico propio del género narrativo escrito en inglés (y otras lenguas), por ejemplo, difiere del léxico que forma parte de las listas de frecuencia generales más conocidas y de las listas utilizadas por las editoriales especializadas en enseñanza de lenguas extranjeras, y en qué medida. Esto permitirá conocer si puede mejorarse la calidad de las lecturas graduadas literarias existentes para otras lenguas, especialmente si el objetivo último es que el estudiante adquiera las destrezas necesarias para leer literatura original en la lengua meta en un tiempo razonable.

CONCLUSIONS

This study was undertaken with the aim of improving the quality of the teaching materials available for Teaching Arabic as a Foreign Language (TAFL) in the area of literary extensive reading. To accomplish this goal, I followed corpus linguistics methods in order to create a frequency list (Chapter V) that can guide specialists in the design of literary Graded Readers (Chapter IV) geared towards students of Arabic at different proficiency levels.

The main objective of this investigation was partially fulfilled. On one hand, it was proven that the compiled corpus used to create my frequency list is representative of contemporary Arabic fiction (Chapters VII and VIII). On the other hand, although it was not possible to create a reliable list of the top 3,500 lemmas, as initially planned, the top 2,000 lemmas of the frequency list generated from my corpus were considered valid for the purposes of this study.

The second objective was to investigate the extent to which a frequency list based on a representative literary corpus is similar to the frequency list of Buckwalter & Parkinson (2011). As I showed, the lexical differences are important and revealed two relevant aspects: 1) the lexicon used in contemporary Arabic novels is idiosyncratic, and 2) there is lexical evidence that the corpus used by Buckwalter & Parkinson for developing their frequency dictionary is not fully representative of the general Arabic language because it contains a high percentage of journalistic, argumentative and political vocabulary.

The results obtained from this study indicate that, in order to develop Graded Readers of literary nature, it is recommended to use the frequency list of Familiar, or any other list based on a representative literary corpus, because its lemmas have an idiosyncratic semantic content. This type of vocabulary is particular to the narrative genre of novels and it is associated to themes related to daily life in a greater extent than the vocabulary included in the Buckwalter & Parkinson list. Therefore, the affirmation of Buckwalter & Parkinson (2011:1) that their dictionary can be used as a reference to develop teaching materials geared towards extensive reading should be relativized.

It would be advisable, though, to use the list of Buckwalter & Parkinson in the development of materials geared towards teaching the language of the media and genres that are argumentative and informative in nature. Therefore, to make a responsible use of this frequency list in the selection or development of appropriate texts that can help develop the reading skills of learners of Arabic as a Foreign Language (AFL), practitioners should have in mind clear pedagogical objectives and take in consideration the profile of their learners.

In any case, an appropriate selection and use of frequent vocabulary in Graded Readers, for the purpose of extensive reading, would help developing and strengthening the reading skills of AFL learners, as long as texts do not include more than 5% of unfamiliar vocabulary (Nation 2013:206). As said in Chapters III and IV, visualizing frequently occurring vocabulary in a systematic way can contribute to 1) develop learners' automaticity in lexical visual recognition, 2) increase their reading speed, and 3) help them acquire new incidental vocabulary through reading. Currently, teaching materials available for TAFL do not allow learners to recognize and understand the *sight vocabulary* typically used in Arabic fiction. Consequently, it would be advisable to develop Graded Readers based on the frequency list of Familiar or any other list that was built in a similar way.

Furthermore, the development and use of literary Graded Readers from early stages can constitute an efficient strategy to help learners overcome the typical obstacles that accompany the visual decoding of Arabic lexicon (Chapter I) and a way to build a robust mental lexicon of frequently used words (Chapter III) that can, later on, enable AFL learners to read novels efficiently. It is important to remember, also, that the use of materials that contain comprehensible input can contribute to lower the anxiety that usually accompanies learners of foreign languages (Horowitz et al, 1989).

In addition to all the benefits of reading literary prose (Chapter II.A), the integration of literature through Graded Readers built around the frequency list of Familiar (or any similar list), can have a direct impact on the ability of AFL learners to communicate and sustain interactions that revolve around daily life. As I showed in the previous chapter, the lexicon used in contemporary Arabic novels is generally used in daily life and it relates to themes like: human interactions, home, family, the human

body, and other practical topics, all of which can promote the ability of learners to communicate in everyday situations.

For this to happen, we would need to integrate in the AFL curriculum the vocabulary typically used in fiction (Chapter X): Common adjectives and their corresponding antonyms, colors, verbs that refer to everyday human interactions, lemmas used in figurative language, parts of human body, and a variety of auxiliary verbs that indicate the commencement of actions.

Finally, I would like to highlight that obtaining data related to lexical frequency and its semantic nature from contemporary Arabic novels was possible because the corpus constructed for this study is extremely representative, since the corpus was designed following a strict scientific methodology, in accordance to corpus linguistics recommendations. In fact, it is possibly the first Arabic corpus which composition is described in such detail in the context of this type of research.

APÉNDICE A

Composición de la muestra literaria en el corpus de Buckwalter y Parkinson

Los datos se recogen según aparece en <http://arabicorpus.byu.edu/>¹⁴³

- Modern Literature
 - The next section of corpora comes from Modern Literature, primarily from novels. The countries represented with the novels are Egypt, Palestine, Algeria, Saudi Arabia, the Sudan, Syria, and Lebanon. Currently about half of the material in the Modern Literature corpus is from Egypt, about a fourth from Algeria, and lesser amounts from the rest of the countries. These are the novels found in the corpus as of Feb 2012:
 - The 'Modern Literature' corpus has 1,026,171 words and consists of the following texts:
 - رائحة البحر ريم بسيوني: with 38,485 words
 - مدبولي ريم بسيوني: with 40,042 words
 - إبراهيم عبد المجيد: لا أحد ينام في الاسكندرية with 98,580 words
 - عمارة يعقوبيان: علاء الأسواني: with 55,191 words
 - شيكاجو علاء الأسواني: with 77,026 words
 - خالد الخميسي: تاكسي with 28,391 words
 - نجيب محفوظ: ميرامار with 35,579 words
 - نجيب محفوظ: الكرنك with 13,742 words
 - نجيب محفوظ: صدى النسيان with 6,276 words
 - نجيب محفوظ: أصداء السيرة الذاتية with 11,542 words
 - نجيب محفوظ: أولاد حارتنا with 104,684 words
 - أحلام مستغانمي: ذاكرة الجسد with 71,065 words
 - أحلام مستغانمي: عابر سرير with 53,925 words
 - أحلام مستغانمي: فوضى الحواس with 57,418 words

¹⁴³ Consulta realizada el 5 de noviembre de 2016.

- رجاى عبءالله الصانع: بنات الرياض with 56,173 words
- الولى الطاهر وطار: الولى الطاهر يعوء إلى مقامه الزكى with 19,271 words
- الولى الطاهر وطار: الولى الطاهر ىرفع ىءىءه بالءعاء with 19,957 words
- التظاهر وطار: الحوات والقصر with 26,756 words
- الطىب صالح: عرس الزىن with 16,283 words
- إءوار الخرات: ترابها زعفران with 20,475 words
- لطىفة الزىات: الشىخوخة وقصص أءرى with 19,556 words
- ىحىى حقى: قصص لىحىى حقى with 9,409 words
- إلباس خورى: مملكة الغرباء with 19,790 words
- غسان كنفانى: أم سعد with 6,623 words
- غسان كنفانى: عاءء إلى حىفا with 11,326 words
- غسان كنفانى: مسرحة الباب with 9,263 words
- نءاة حالو: سر الحىاة with 11,211 words
- سعد الله ونوس: مءامرة رأس المملوك ءابر with 17,659 words
- ءمىم صائب: لا ءفقا عىنك ىا أوءىب with 7,607 words
- ءءم لءاكرة بالشمع الأحمر with 4,070 words
- على سالم: أولاءنا فى لئءن with 16,865 words
- أمءء ناصر: حىء لا ءسقط الأمطار with 41,931 words

APÉNDICE B

Carta oficial para acceder al corpus literario



Facultad de Filosofía y Letras
Departamento de Estudios Árabes e Islámicos y Estudios Orientales

10 نوفمبر 2014

إلى من يهيمه الأمر،
تحية طيبة وبعد،

نتقدم إليكم، الدكتورة الأستاذة Nieves Paradela أستاذة اللغة العربية والأدب العربي بجامعة مدريد المستقلة، والدكتور الأستاذ Manuel Feria أستاذ الترجمة بجامعة غرناطة، لنفدكم بأن Laila Familiar طالبة رسالة الدكتوراه في قسم الدراسات العربية والإسلامية بجامعة مدريد المستقلة وبأننا نشرف على بحثها في برنامج الدراسات العليا المذكور أعلاه.

تقوم رسالة الطالبة Laila Familiar للدكتوراه على تحليل لغة الروايات العربية التي نالت بالجائزة العالمية للرواية العربية (البوكر) وأيضا تلك التي وصلت إلى القائمة القصيرة بهدف إخراج قائمة تواتر مفردات يصلح استخدامها كأحد المعايير اللغوية التي يقوم عليه تدريس الأدب العربي المعاصر للطلاب الأجانب في جامعاتنا.

وعليه نتقدم لكم لنستأذنكم بمنح نسخ إلكترونية (على هيئة Word أو Rtf) لروايات البوكر للطالبة لإجراء بحث الدكتوراه ونتعهد لسيادتكم بعدم تداول النسخة الإلكترونية خارج هذا النطاق البحثي ونتعهد أيضا بالحفاظ على جميع حقوق الناشر والكتّاب بجميع الطرق كما تنص عليه لوائح البحث العلمي.

وتقبلوا منا فائق التقدير والاحترام وشكرا لحسن تعاونكم معنا

الدكتورة الأستاذة Nieves Paradela - قسم الدراسات العربية والإسلامية - كلية الآداب - جامعة مدريد المستقلة

الدكتور الأستاذ Manuel Feria - كلية الترجمة - جامعة غرناطة



APÉNDICE C

Composición del corpus literario Familiar

Autor	Obra	País	Tamaño de obra completa aproximada (palabras)	Tamaño de obra en Población (%)	Pág. necesarias para el corpus sometido a análisis	Palabras necesarias para el corpus sometido a análisis	Palabras recogidas	Cobertura de muestra requerida (%)	Páginas seleccionadas en la obra
إلياس فرح	أرض اليمبوس	الأردن	48115	0.60	2.92	530.72	516	97.23	9-161-259
بشير مفتي	دمية النار	الجزائر	35372	0.44	2.14	390.16	420	107.65	37-20-116-120
سمير قسيبي	يوم رائع للموت	الجزائر	20020	0.25	1.21	220.83	218	98.72	101-52
واسيني الأعرج	أصابع لوليتا	الجزائر	89726	1.11	5.44	989.70	1088	109.93	388-320-35-423-437-85
أحلام مستغانمي	ذاكرة الجسد	الجزائر	67400	0.84	4.08	743.44	779	104.78	13-139-314-70
أمين الزاوي	حادي الثبوس	الجزائر	30212	0.37	1.83	333.25	325	97.53	61-18

واسيني الأعرج	رماد الشرق	الجزائر	176448	2.19	10.69	1,946.26	2120	108.93	398-199-455-77-64-192 (Parte 1) + 295-59-311-378-299 (Parte 2)
واسيني الأعرج	البيت الأندلسي	الجزائر	79534	0.99	4.82	877.28	871	99.28	230-70-176-99-97
مقبول موسى العلوي	فتنة جدّة	السعودية	55874	0.69	3.39	616.30	615	99.79	46-77-234
رجاء عالم	طوق الحمام	السعودية	101192	1.25	6.13	1,116.17	1091	97.74	437-279-561-382-452-428
بدرية البشر	غراميات شارع الأعشى	السعودية	51110	0.63	3.10	563.76	611	108.38	104-292-177-198
محمد حسن علوان	القندس	السعودية	60823	0.75	3.69	670.89	686	102.25	223-57-35-250
عبد خال	ترمي بشرر	السعودية	65156	0.81	3.95	718.69	709	98.65	226-407-34-198
عبد الله بن بخيت	شارع العطايف	السعودية	70616	0.88	4.28	778.91	804	103.22	336-225-179-69-391
أميمة الخميس	الوارفة	السعودية	57089	0.71	3.46	629.71	632	100.36	60-266-113-54
أمير تاج السر	366	السودان	39754	0.49	2.41	438.50	445	101.48	135-83-165
حامد الناظر	نبوءة السقا	السودان	50549	0.63	3.06	557.57	558	100.08	153-308-189-200-83
أمير تاج السر	صائد اليرقات	السودان	29743	0.37	1.80	328.07	320	97.54	118-97
حمور زيادة	شوق الدرويش	السودان	81900	1.01	4.96	903.38	864	95.64	334-177-162-312-461-437-289-192

محسن الرملي	تمر الأصابع	العراق	28392	0.35	1.72	313.17	315	100.58	18-65
هدية حسين	ريام وكفى	العراق	37128	0.46	2.25	409.53	439	107.20	51-208-101-17
عبد الخالق الركابي	ليل علي بابا الحزين	العراق	36764	0.46	2.23	405.52	427	105.30	238-56-168
علي بدر	ملوك الرمال	العراق	43680	0.54	2.65	481.80	479	99.42	23-105-24-202
علي بدر	حارس التبغ	العراق	63518	0.79	3.85	700.62	708	101.05	306-241-131-270-84
محسن الرملي	حدائق الرئيس	العراق	63856	0.79	3.87	704.35	718	101.94	147-68-161-264
إنعام كجه جي	طشاري	العراق	43862	0.54	2.66	483.81	464	95.91	34-203-57
إنعام كجه جي	الحفيدة الأميركية	العراق	33670	0.42	2.04	371.39	368	99.09	42-146-150
جنان جاسم حلاوي	أهل النخيل	العراق	67886	0.84	4.11	748.80	758	101.23	305-177-377-51-210
حوراء النداوي	تحت سماء كوبنهاغن	العراق	69524	0.86	4.21	766.87	769	100.28	45-34-108-46-208-338-57
حازم كمال الدين	مياه متصحرة	العراق	24611	0.30	1.49	271.47	276	101.67	149-78-163
أحمد سداوي	فرانكشتاين في بغداد	العراق	62244	0.77	3.77	686.57	714	104.00	237-238-303-21
سنان أنطون	يا مريم	العراق	27300	0.34	1.65	301.13	318	105.60	24-145-56

إسماعيل فهد إسماعيل	في حضرة العنقاء والخل الوفي	الكويت	98332	1.22	5.96	1,084.63	1101	101.51	319-360-81-16-212-323
سعود السنعوسي	ساق البامبو	الكويت	71074	0.88	4.31	783.96	797	101.66	499-53-262-214-40-491
عبد الوهاب الحمادي	لا تقصص رويك	الكويت	42553	0.53	2.58	469.37	483	102.90	136-167-9
طالب الرفاعي	في الهنا	الكويت	29848	0.37	1.81	329.23	350	106.31	127-171-20
أحمد المديني	ممر الصفصاف	المغرب	77674	0.96	4.71	856.76	832	97.11	271-2-313-6
عبد الكريم جويطي	كتيبة الخراب	المغرب	28202	0.35	1.71	311.07	338	108.66	41-26
محمد الأشعري	القوس والفراشة	المغرب	59150	0.73	3.58	652.44	712	109.13	137-227-67-147
عبد الرحيم لحبيبي	تغريبة العبدى المشهور بولد الحميرية	المغرب	68053	0.84	4.12	750.64	773	102.98	165-275-269-147-43
طارق بكارى	نوميديا	المغرب	85081	1.05	5.16	938.46	937	99.84	282-289-226-197-388-267
محمد برادة	بعيدا من الضوضاء، قريبا من السكات	المغرب	45388	0.56	2.75	500.64	486	97.08	183-190-143
يوسف فاضل	طائر أزرق نادر يخلق معي	المغرب	57512	0.71	3.49	634.37	699	110.19	118-73-105-313-252

بنسالم حميش	هذا الأندلسي	المغرب	89908	1.11	5.45	991.71	1020	102.85	400-310-435-54-303-219
بنسالم حميش	العلامة	المغرب	44044	0.55	2.67	485.82	491	101.07	216-165-204-48
بنسالم حميش	معدّتي	المغرب	51870	0.64	3.14	572.14	558	97.53	83-27-174-72
إسماعيل غزالي	موسم صيد الزنجور	المغرب	77168	0.96	4.68	851.18	854	100.33	349-52-367-166-317-356-206
عبد النور مزين	رسائل زمن العاصفة	المغرب	53690	0.67	3.25	592.21	594	100.30	133-198-300-100
حبيب عبد الرب السروري	ابنة سوسلوف	اليمن	38948	0.48	2.36	429.61	424	98.70	146-87-196-131
علي المقري	اليهودي الحالي	اليمن	25844	0.32	1.57	285.07	264	92.61	30-85-55
علي المقري	طعم أسود.. رائحة سوداء	اليمن	19838	0.25	1.20	218.82	214	97.80	101-75
الحبيب السالمي	نساء البساتين	تونس	35854	0.44	2.17	395.48	415	104.94	189-29
الحبيب السالمي	روائح ماري كلير	تونس	38766	0.48	2.35	427.60	459	107.34	118-151-186
شكري المبخوت	الطلياني	تونس	60788	0.75	3.68	670.51	685	102.16	65-18-255-249
حسين الواد	سعادته السيد الوزير	تونس	57500	0.71	3.48	634.24	699	110.21	103-69-229-258-85
فادي عزام	سرمدة	سورية	47120	0.58	2.86	519.74	554	106.59	60-172-156-38

ابتسام إبراهيم ترسي	عين الشمس	سورية	63461	0.79	3.85	699.99	713	101.86	262-267-41-139
روزا ياسين حسن	حرّاس الهواء	سورية	52780	0.65	3.20	582.18	593	101.86	267-255-128-205
لينا هويان الحسن	ألماس ونساء	سورية	35062	0.43	2.12	386.74	387	100.07	229-2-157
خالد خليفة	لا سكاكين في مطابخ هذه المدينة	سورية	44590	0.55	2.70	491.84	484	98.41	8-88-34
محمود حسن الجاسم	نزوح مريم	سورية	69342	0.86	4.20	764.86	779	101.85	234-180-13-187-169
مها حسن	الراويات	سورية	32760	0.41	1.99	361.35	371	102.67	22-76
خالد خليفة	في مديح الكراهية	سورية	80444	1.00	4.88	887.32	912	102.78	407-59-206-252-295
فواز حداد	جنود الله	سورية	67683	0.84	4.10	746.56	729	97.65	289-171-231-300-91
مها حسن	حبل سري	سورية	74336	0.92	4.51	819.94	821	100.13	219-350-228-54-305-273
فواز حداد	المتزجم الخائن	سورية	91108	1.13	5.52	1,004.94	961	95.63	118-458-435-430-236-209
محمد أبو معتوق	القمقم والجني	سورية	67886	0.84	4.11	748.80	791	105.64	32-352-101-28-22
شهلا العجيلي	سماة قريبة من بيتنا	سورية	60424	0.75	3.66	666.49	700	105.03	97-163-214-71-283
خليل صويلح	وراق الحب	سورية	27702	0.34	1.68	305.56	306	100.14	63-90

عاطف أبو سيف	حياة معلقة	فلسطين	72436	0.90	4.39	798.99	842	105.38	11-229-199-182
إبراهيم نصرالله	شرفة الهاوية	فلسطين	57330	0.71	3.47	632.36	663	104.84	192-279-71-233
إبراهيم نصرالله	قناديل ملك الجليل	فلسطين	99918	1.24	6.06	1,102.12	1126	102.17	494-468-434-9-193-496
إبراهيم نصرالله	زمن الخيول البيضاء	فلسطين	91182	1.13	5.53	1,005.76	1021	101.52	464-489-440-409-254-285
ربيعي المدهون	مصائر : كونشرتو الهولوكوست والنكبة	فلسطين	46774	0.58	2.83	515.93	479	92.84	249-233-35-160
أنور حامد	يافا تعد قهوة الصباح	فلسطين	26068	0.32	1.58	287.54	289	100.51	77-56
ربيعي المدهون	السيدة من تل ابيب	فلسطين	57694	0.71	3.50	636.38	573	90.04	180-66-246-272
سحر خليفة	أصل وفصل	فلسطين	80444	1.00	4.88	887.32	977	110.11	364-43-144-225-173-446-438
سحر خليفة	صورة وأيقونة وعهد قديم	فلسطين	46228	0.57	2.80	509.91	509	99.82	230-34-216
يحي خلف	ماء السماء	فلسطين	50050	0.62	3.03	552.06	544	98.54	48-165-70-5-158
مريد البرغوثي	رأيت رام الله	فلسطين	46046	0.57	2.79	507.90	506	99.63	154-199-155
إبراهيم زعرور	رحلة خير الدين بن زرد العجيبة	فلسطين	36946	0.46	2.24	407.52	405	99.38	9 -- 24 -- 78

محمود الريماوي	من يونس السيدة؟	فلسطين	28173	0.35	1.71	310.75	321	103.30	162-141
ليلي الأطرش	ترانيم الغواية	فلسطين	52202	0.65	3.16	575.80	569	98.82	241-90-221
محمود شقير	مديح لنساء العائلة	فلسطين	35672	0.44	2.16	393.47	423	107.50	9-78-191
أنطوان الدويهي	غريقة بحيرة موريه	لبنان	32942	0.41	2.00	363.36	392	107.88	81-176-73
أنطوان الدويهي	حامل الوردة الأرجوانية	لبنان	32760	0.41	1.99	361.35	374	103.50	71-61-57
جنى فواز الحسن	أنا، هي والأخريات	لبنان	41269	0.51	2.50	455.21	461	101.27	39-109-48
ربيع جابر	دروز بلغراد	لبنان	41678	0.52	2.53	459.72	471	102.45	141-63-100
رينيه الحايك	حياة قصيرة	لبنان	31668	0.39	1.92	349.31	350	100.20	109-133-122
ربيع جابر	أميركا	لبنان	76622	0.95	4.64	845.16	855	101.16	24-195-380-298-45
رينيه الحايك	صلاة من أجل العائلة	لبنان	28210	0.35	1.71	311.16	339	108.95	162-146
إلياس خوري	سينالكول	لبنان	91000	1.13	5.52	1,003.75	938	93.45	331-434-75-71-137-373
هدى بركات	ملكوت هذه الأرض	لبنان	62062	0.77	3.76	684.56	655	95.68	70-139-84-108-279
علوية صبح	اسمه الغرام	لبنان	63950	0.79	3.88	705.38	660	93.57	294-202-194
هدى بركات	حارث المياه	لبنان	30940	0.38	1.88	341.28	354	103.73	89-171-19

ربيع جابر	طيور الهوليداي إن	لبنان	73346	0.91	4.45	809.02	891	110.13	428-152-376-285-175
جبور الدويهي	حي الأميركان	لبنان	27300	0.34	1.65	301.13	321	106.60	41-9
جبور الدويهي	شريد المنازل	لبنان	62043	0.77	3.76	684.35	683	99.80	14-27-177-371-305
حسن داوود	مئة وثمانون غروبا	لبنان	50232	0.62	3.04	554.07	608	109.73	7-154-97
جبور الدويهي	مطر حزيران	لبنان	62244	0.77	3.77	686.57	729	106.18	75-258-316-130
حسن داوود	لا طريق إلى الجنة	لبنان	63897	0.79	3.87	704.80	708	100.45	289-85-5-6
فاتن المر	الخطايا الشائعة	لبنان	32157	0.40	1.95	354.70	346	97.55	81-189-112
رشيد الضعيف	تلبيط البحر	لبنان	40498	0.50	2.45	446.70	471	105.44	197-124-12
مي منسي	أنتعل الغبار وأمشي	لبنان	44590	0.55	2.70	491.84	487	99.02	89-248-91-179
جورج يرق	حارس الموتى	لبنان	53508	0.66	3.24	590.21	591	100.13	293-26-240-260
جنى فواز الحسن	طابق 99	لبنان	49333	0.61	2.99	544.15	519	95.38	101-228-168-
أحمد محسن	وارسو قبل قليل	لبنان	20566	0.25	1.25	226.85	226	99.63	5-71-20
شربل قطان	حقائب الذاكرة	لبنان	40040	0.50	2.43	441.65	486	110.04	43-18-95
إبراهيم الكوني	الورم	ليبيا	32760	0.41	1.99	361.35	363	100.46	79-181-128

منى الشيمي	بحجم حبة عنب	مصر	86868	1.08	5.26	958.17	928	96.85	287-285-366-413-139-7
أشرف الخمايسي	مناقي الرب	مصر	64344	0.80	3.90	709.73	685	96.52	234-326-292-414-120-351
جمال ناجي	عندما تشيخ الذئاب	مصر	43134	0.53	2.61	475.78	501	105.30	244-224-175
أشرف الخمايسي	انحراف حاد	مصر	50621	0.63	3.07	558.36	536	96.00	91-231-257-8-228-141
أشرف العشماوي	تويا	مصر	50596	0.63	3.07	558.09	564	101.06	45-124-244
ناصر عراق	العاطل	مصر	62790	0.78	3.81	692.59	700	101.07	166-20-76-248-263
إبراهيم عيسى	مولانا	مصر	104104	1.29	6.31	1,148.29	1049	91.35	577-281-81-10-427-454
لطيفة الزيات	الباب المفتوح	مصر	62972	0.78	3.82	694.60	709	102.07	406-363-141-401-266
محمد البساطي	جوع	مصر	22750	0.28	1.38	250.94	253	100.82	70-72
بهاء طاهر	واحة الغروب	مصر	69160	0.86	4.19	762.85	782	102.51	98-247-195-154
مكاوي سعيد	تغريدة البجعة	مصر	54782	0.68	3.32	604.26	663	109.72	357-361-18-49
عزت القمحاوي	بيت الديب	مصر	54690	0.68	3.31	603.24	621	102.94	137-48-27-221
محمد ربيع	عطارد	مصر	53508	0.66	3.24	590.21	606	102.68	55-133-204
حمدي أبو جليل	الفاعل	مصر	25438	0.32	1.54	280.59	287	102.29	55-115-83

محمد المنسي قنديل	كتيبة سوداء	مصر	108170	1.34	6.56	1,193.14	1182	99.07	146-208-217-248-77-479
أحمد مراد	الفيل الأزرق	مصر	60332	0.75	3.66	665.48	629	94.52	71-223-338-148-130
إبراهيم عبد المجيد	الاسكندرية في غيمة	مصر	82264	1.02	4.99	907.39	912	100.51	258-68-227-216-102-412-253
يوسف زيدان	النبطي	مصر	69524	0.86	4.21	766.87	743	96.89	247-301-129-105-69
ميرال الطحاوي	بروكلين هايتس	مصر	47102	0.58	2.85	519.55	543	104.51	115-18-193
خيري شلبي	إسطاسية	مصر	44044	0.55	2.67	485.82	484	99.63	128-250-28-62
محمد المنسي قنديل	يوم غائم في البر الغربي	مصر	110618	1.37	6.70	1,220.14	1226	100.48	561-310-498-445-473-259-326
يوسف زيدان	عزازيل	مصر	83356	1.03	5.05	919.44	919	99.95	116-225-248-425-266-233
عز الدين شكري فشير	غرفة العناية المركزة	مصر	52780	0.65	3.20	582.18	601	103.23	94-54-276
خيري شلبي	وكالة عطية	مصر	94276	1.17	5.71	1,039.89	1122	107.90	160-153-391-411-220-360
إبراهيم عبد المجيد	البلدة الأخرى	مصر	68068	0.84	4.13	750.81	738	98.29	14-165-12-193-214
عز الدين شكري فشير	عناق عند جسر بروكلين	مصر	38220	0.47	2.32	421.58	416	98.68	147-80-75

خالد البري	رقصة شرقية: طقوس ميلاد وغياب ابن الرفاعية المنتظر	مصر	101729	1.26	6.17	1,122.09	1156	103.02	295-419-357-105-312-368
منصورة عز الدين	وراء الفردوس	مصر	38766	0.48	2.35	427.60	409	95.65	56-188-76
محمد عبد النبي	رجوع الشيخ	مصر	42318	0.52	2.56	466.78	464	99.40	176-37-130
يوسف أبو رية	ليلة عرس	مصر	28028	0.35	1.70	309.16	312	100.92	148-101-57 - 78
هشام الخشن	جرافيت	مصر	38948	0.48	2.36	429.61	424	98.70	54-172-67-97
سمية رمضان	أوراق النرجس	مصر	24388	0.30	1.48	269.01	265	98.51	87-25
إبراهيم فرغلي	معبد أنامل الحرير	مصر	100373	1.24	6.08	1,107.14	1106	99.90	354-119-147-150-310-479
أمينة زيدان	نبيذ أحمر	مصر	47502	0.59	2.88	523.96	544	103.83	102-169-120-167
محمد الرفاعي	كائنات الحزن الليلية	مصر	36036	0.45	2.18	397.49	417	104.91	44-41-134
			8.070.278	100.00	489.10	89.017	89.732		

APÉNDICE D

26 lemas corregidos en el *output* de MADAMIRA tras una revisión manual de 1683 tokens

<i>Output</i> MADAMIRA	Corrección	Tipo de error
نوبيّ	نُوبِيَّة	Tokenización + Lematización
عَلَامَة	عَلَامَة	Lematización
جَذَب	جَذَب	Lematización
نَفْس	نَفْس	Lematización
أثار	آثار	Tokenización + Lematización
قَدْر	قَدْر	Lematización
أَنَّ	أَنَّ	Lematización
بِ + عيد	بَعِيد	segmentación
كُرَّاس	كُرَّاس	Lematización
مَرَض	مَرَضِي	Segmentación + Tokenización
وَجَّه	وَجَّه	Lematización
بُعْد	بُعْد	Lematización
أَحْيَى	أَحْيَاء	Segmentación + Tokenización
خَرَب	خَرِبَ	Lematización
ت ¹⁴⁴ + نِير	أَنَار	Segmentación + Tokenización
شِعْر	شِعْر	Lematización
بُلْعَة	بُلْعَ	Tokenización
طَرِيق	طَرَقَ	Tokenización
رُكْن	رُكْن	Lematización

¹⁴⁴ Se observó que el *output* de MADAMIRA contenía muchas ocurrencias de la forma ت y se descubrió que suele ser prefijo de tercera persona femenina singular, especialmente cuando el verbo va seguido de pronombre sufijo

حَيَّ	حَيَّا	Segmentación + Tokenización
أَمْرَ	أَمَرَ	Lematización
سُمُو	سَمَى	Segmentación + Tokenización
تُرْكِي	تَرَكَ	Segmentación
بِ	بِهِمْ	Segmentación
حُسْنِ	حَسَنَ	Lematización
أَلَقِ	أَلَقَى	Segmentación + Tokenización

APÉNDICE E

Errores de MADAMIRA descubiertos durante la elaboración de la lista de frecuencia Familiar

<i>Output</i> MAD-AMIRA	Ocurrencias ¹⁴⁵	Lema correcto y observaciones	Tipo de error
صَرَ	12	صار	Lematización
وَحَى	12	Distintos lemas que, al ser de baja frecuencia, no aparecen en la lista final	Segmentación
أَسَّ	7	Distintos lemas que, al ser de baja frecuencia, no aparecen en la lista final	Segmentación
أَبِيهِ	6	أَب	Segmentación ¹⁴⁶
أَبِيَّ	5	أَب	Segmentación
وَنَى	6	Distintos lemas que, al ser de baja frecuencia, no aparecen en la lista final	Palabras dialectales no identificadas
صُبَّاح	7	صَبَّاح	Tokenización + Lematización
حَلَّوْ	13	Distintos lemas que, al ser de baja frecuencia, no aparecen en la lista final	Segmentación+Tokenización
جَدِّي	14	Distintos lemas que, al ser de baja frecuencia, fueron eliminados en su	Segmentación

¹⁴⁵ Es importante señalar que el número de ocurrencias no se refiere al número de veces que MADAMIRA se equivoca, sino al número de ocasiones en que el *output* se observa. En la mayoría de los casos, el margen de error de MADAMIRA es mínimo. Sin embargo, con el ánimo de ofrecer unos datos léxicos lo más exactos posible, hubo de realizarse una revisión manual, especialmente en el caso de algunos homógrafos y lemas extraños,

¹⁴⁶ Este tipo de error en segmentación se observa en varias ocasiones cuando el lema va acompañado de sufijos posesivos. No puedo determinarse si el error se debe a la exclusión de los enclíticos del análisis y si dichos errores desaparecerían si los comandos dados a MADAMIRA hubieran sido distintos.

		mayoría excepto algunos que se sumaron al lema جَدَّ	
أَجَل	35	Homógrafo que resultó en tres lemas, uno de los cuales aparece en la lista final (أَجَل), pues los otros dos eran de baja frecuencia ¹⁴⁷ .	Tokenización + Lematización
مَثَل	154	Homógrafo que resultó en tres lemas, uno de los cuales aparece en la lista final (مَثَل con 145 ocurrencias), pues los otros dos ¹⁴⁸ eran de baja frecuencia.	Tokenización + Lematización
رَدَّ ¹⁴⁹	82	Homógrafo que resultó en dos lemas distintos: رَدَّ (46 ocurrencias) y رَدُّ ¹⁵⁰ (26 ocurrencias)	Tokenización + Lematización
حَدَّث	70	Homógrafo que resultó en tres lemas distintos: حَدَّث (37 ocurrencias), حَدَّثُ (18 ocurrencias), حَدَّتْ (14 ocurrencias) ¹⁵¹	Tokenización + Lematización
أَسْرَعَ	15	Homógrafo que resultó en dos lemas distintos, uno de los cuales aparece en la lista final (أَسْرَعَ con 12 ocurrencias)	Tokenización + Lematización
كَفَّ	22	Este homógrafo no siempre fue etiquetado de forma correcta ¹⁵² . Resultó finalmente en dos lemas distintos: كَفَّ – كَفُّ	Tokenización + Lematización

¹⁴⁷ Se trata de los lemas أَجَل – أَجَل.

¹⁴⁸ Se trata de los lemas مَثَل و مَثَل.

¹⁴⁹ Este *output* descubrió algunos serios problemas MADAMIRA a la hora de segmentar, tokenizar y lematizar verbos irregulares que contienen en la raíz semiconsonantes. Por ejemplo, verbos como وَرَدَ و يُرَدُ (لم) – أَرَادَ – أَرَادَ – fueron en ocasiones lematizados como رَدَّ. En otra ocasión, el verbo conjugado أَرَدْتُمْ fue lematizado de dos formas distintas: una vez como رَدَّ y otra como أَرَدَ (que no existe en árabe).

¹⁵⁰ Cabe señalar que el lema رَدَّ (sustantivo) es asimismo un homógrafo, pues no solo aparece con la acepción de “respuesta” sino también en distintas colocaciones, como: رد الجميل – رد فعل.

¹⁵¹ En una ocasión se identificó el verbo تَحَدَّثَ lematizado como حَدَّتْ.

¹⁵² Este error desveló que el diccionario que acompaña a MADAMIRA debiera ser revisado y actualizado.

طلب	95	Este homógrafo no siempre fue etiquetado de forma correcta. Resultó finalmente en dos lemas distintos: طَلَّب (12 ocurrencias) y طَلَب (83 ocurrencias)	Tokenización + Lematización
أُبْعِدَ	20	Homógrafo que result en dos lemas distintos: أُبْعِدَ (9 ocurrencias) y أَبْعَدَ (9 ocurrencias)	Tokenización + Lematización
وَجْهِي	18	وَجْه	Segmentación
عَمَّة	14	عَمَّ	Segmentación
خَزَّان	12	خَزَّان – خزانة	Segmentación+ Tokenización + Lematización
حَبِيبِي	11	حَبِيب	Segmentación
كَذَّ	11	كَادَ	Segmentación+ Tokenización
أَسْمَع	10	سَمِعَ	Segmentación+ Tokenización
حَامِل	10	Homógrafo que en la lista final quedó reducido al lema حَامِل (embarazada), pues el otro lema (حَامِلٌ participio activo de حَمَلَ) era de baja frecuencia.	Lematización
شِفَّت	10	شِفَّة	Segmentación+ Tokenización
ظَهْرِي	9	ظَهْر	Segmentación
شَعُورِي	8	شَعُور	Segmentación
نَحْوِي	13	نَحْو	Segmentación
أَجْد	7	وَجَدَ	Segmentación
أَغَار	5	Distintos lemas que, al ser de baja frecuencia, no aparecen en la lista final.	Segmentación+ Tokenización
أُمِّي	7	أَم	Segmentación

صاحي	6	صاح – صاح	Tokenización + Lematización
مأني	7	مات	Segmentación
أنام	6	نام	Segmentación
حلمي	6	حلم – حلمي	Tokenización + Lematización
صوّة	6	صوت	Segmentación
اذن	50	Homógrafo ¹⁵³ que en la lista final result en tres lemas distintos: أُذُن (24 ocurrencias), إِذْن (16 ocurrencias), y إِذْن (6 ocurrencias).	Tokenización + Lematización
كوان	16	كوب	Segmentación
ذريّ	14	بدرية – ذرية	Segmentación
جهد	12	جهد – جهود	Lematización
بيد	11	ب + يد	Segmentación
لكن	74	لكن ¹⁵⁴	Tokenización
لكن	343	لكن	Tokenización
ت	51	(-)	Segmentación
بك	23	ب	Segmentación
بي	11	ب	Segmentación
كمن	10	ك + من	Segmentación+ Tokenización + Lematización
خالة	9	خال	Segmentación
فواد	8	فواد	Ortografía
عمري	7	عمر	Segmentación
عقلي		عقل	Segmentación

¹⁵³ Este es uno de los homógrafos más complejos de procesar en árabe posiblemente. En mi corpus, este homógrafo resultó en 6 lemas distintos (muchas veces con errores), mientras que en la lista Buckwalter y Parkinson solo se recogen 3 lemas en total.

¹⁵⁴ MADAMIRA procesa el lema لكن de dos formas distintas en cuanto a su vocalización. Puesto que ambas lematizaciones se refieren, en realidad, al mismo lema, se colapsaron.

APÉNDICE F

Variantes del pronombre relativo *alladī* y del demostrativo *hādā*, sus respectivas ocurrencias en el corpus Familiar, y su presencia en el diccionario de frecuencia de Buckwalter y Parkinson (2011)

Lema	Ocurrencias en Familiar	Presencia en Buckwalter y Parkinson
هذا	414	Sí
هذان	0	Sí
هذه	284	Sí
هاتان	0	Sí
هؤلاء	29	Sí
هذين	1	No
هاتين	1	No
أولئك	6	Sí
الذي	447	Sí
اللذان	1	Sí
التي	519	Sí
اللتان	1	Sí
الذين	53	Sí
اللواتي	5	Sí
اللاتي	2	Sí
اللائي	0	No
اللتين	5	No

APÉNDICE G

Dos listas de frecuencia: Buckwalter y Parkinson vs. Familiar (2.000 lemas)

Ranquin	Buckwalter y Parkinson	Frecuencia Absoluta	Frecuencia normalizada	Familiar	Frecuencia absoluta	Frecuencia normalizada
1	ال	5004793	166.83	ال	14851	165.01
2	و	1110144	37.00	و	6699	74.43
3	في	924823	30.83	ب	3334	37.04
4	مِنْ	745190	24.84	في	2783	30.92
5	لِ	584786	19.49	لِ	2462	27.36
6	ب	553234	18.44	مِنْ	2429	26.99
7	على	518692	17.29	على	1631	18.12
8	أَنَّ	303942	10.13	كَانَ	1568	17.42
9	إلى	299648	9.99	أَنَّ	1365	15.17
10	كَانَ	281097	9.37	إلى	1264	14.04
11	لا	247225	8.24	لا	1086	12.07
12	الله	211480	7.05	ما	914	10.16
13	أَنْ	192807	6.43	ف	899	9.99
14	عن	180308	6.01	أَنْ	889	9.88
15	قال	170778	5.69	لَمْ	807	8.97
16	هذا	155507	5.18	قال	649	7.21
17	التي	155316	5.18	عن	605	6.72
18	مع	152096	5.07	التي	519	5.77
19	كل	145875	4.86	ذلك	487	5.41
20	ما	127514	4.25	كل	477	5.30
21	ف	117124	3.90	مع	472	5.24
22	هو	115797	3.86	الذي	447	4.97
23	هذه	110618	3.69	لكن	417	4.63

24	أو	108848	3.63	هذا	414	4.60
25	الذي	108664	3.62	ك	361	4.01
26	أنا	104898	3.50	هو	361	4.01
27	يا	96575	3.22	بَعْدَ	355	3.94
28	ما	95564	3.19	س	354	3.93
29	لَمْ	94821	3.16	قَدْ	354	3.93
30	يوم	93393	3.11	أو	335	3.72
31	إِنَّ	86705	2.89	شيء	321	3.57
32	بَيْنَ	81509	2.72	أَخْرَ	315	3.50
33	هي	79477	2.65	أنا	293	3.26
34	بَعْدَ	76319	2.54	عَرَفَ	293	3.26
35	ذلك	75488	2.52	نَفْسُ	289	3.21
36	قَدْ	73404	2.45	هذه	284	3.16
37	أَخْرَ	63343	2.11	ثُمَّ	278	3.09
38	شيء	63264	2.11	حتى	275	3.06
39	عند	60994	2.03	هي	270	3.00
40	أول	59417	1.98	يوم	267	2.97
41	غَيْرَ	56329	1.88	عند	255	2.83
42	عربي	53325	1.78	رأى	254	2.82
43	إذا	52854	1.76	جِئْنَا	227	2.52
44	نَفْسُ	52702	1.76	بَيْنَ	221	2.46
45	رئيس	52478	1.75	عادَ	220	2.44
46	أَيَّ	51568	1.72	مرة	218	2.42
47	عَمَلِ	50613	1.69	عَيْنِ	210	2.33
48	عَرَفَ	50348	1.68	أَحَدٌ	205	2.28
49	بعض	49902	1.66	بيت	200	2.22
50	يعني	49368	1.65	يد	197	2.19
51	دولة	49360	1.65	صوت	193	2.14
52	كما	48078	1.60	يا	182	2.02
53	إِلَّا	46733	1.56	قَبْلَ	178	1.98
54	لكن	45859	1.53	أَمْ	172	1.91
55	ليس	45416	1.51	وَجْهَ	170	1.89

56	أنت	45267	1.51	غَيْر	167	1.86
57	كثير	44541	1.48	إِنْ	164	1.82
58	لِأَنَّ	43701	1.46	أب	163	1.81
59	واحد	43466	1.45	رَجُل	162	1.80
60	إِنْ	40464	1.35	أَمْرٌ	159	1.77
61	جديد	40375	1.35	إِلَّا	154	1.71
62	خلال	39028	1.30	أَيَّ	154	1.71
63	عام	38637	1.29	ليس	154	1.71
64	أَحَدٌ	37897	1.26	أَكْثَر	153	1.70
65	أَكْثَر	37538	1.25	أول	153	1.70
66	كبير	37236	1.24	بعض	153	1.70
67	أخ	36766	1.23	أراد	152	1.69
68	كيف	36412	1.21	وَجَدَ	150	1.67
69	قَبْلَ	35711	1.19	كثير	147	1.63
70	هل	34854	1.16	أمام	146	1.62
71	سَنَةَ	34782	1.16	مِثْلَ	145	1.61
72	أَمْرٌ	34681	1.16	مكان	145	1.61
73	قوة	34642	1.15	باب	142	1.58
74	تم	34557	1.15	حياة	142	1.58
75	حيث	34098	1.14	رَأْس	140	1.56
76	رأى	34062	1.14	هناك	140	1.56
77	هناك	33870	1.13	لِأَنَّ	138	1.53
78	مرة	33861	1.13	لو	137	1.52
79	أب	33766	1.13	عندما	135	1.50
80	لكن	33727	1.12	هل	134	1.49
81	أمريكي	33325	1.11	نَظَرَ	130	1.42
82	ثاني	33321	1.11	شَعَرَ	128	1.42
83	حتى	33051	1.10	واحد	127	1.41
84	ثُمَّ	33045	1.10	ابن	126	1.40
85	منطقة	32675	1.09	دون	125	1.39
86	جميع	32553	1.09	سَأَلَ	125	1.39
87	إِنَّ	32395	1.08	اسم	121	1.34

88	حَقَّ	32127	1.07	صغير	121	1.34
89	مَنْ	31648	1.05	بدا	118	1.31
90	مِثْل	31527	1.05	سَمِعَ	118	1.31
91	لو	31350	1.05	وقت	117	1.30
92	عام	30916	1.03	أَنْ	116	1.29
93	أمكن	30678	1.02	طويل	116	1.23
94	موضوع	30657	1.02	إذا	111	1.23
95	وزير	30636	1.02	الله	111	1.23
96	اسم	30452	1.02	غرفة	111	1.22
97	رَجُل	30264	1.01	أَخَذَ	110	1.21
98	مجلس	30206	1.01	استطاع	109	1.21
99	عالم	30204	1.01	بَدَأَ	109	1.21
100	حياة	30100	1.00	منذ	109	1.20
101	نحن	29658	0.99	أرض	108	1.20
102	بلد	29658	0.99	صورة	108	1.19
103	وقت	29493	0.98	لحظة	107	1.16
104	حكومة	29332	0.98	لن	104	1.14
105	قام	29003	0.97	فَعَلَ	103	1.14
106	فلسطيني	28613	0.95	كبير	103	1.13
107	سياسي	28526	0.95	أيضا	102	1.13
108	بيت	27875	0.93	خرج	102	1.11
109	دون	27400	0.91	أنت	100	1.10
110	جاء	26234	0.87	قَلْب	99	1.10
111	قَلْب	26075	0.87	ترك ¹⁵⁵	99	1.08
112	ألف	26022	0.87	طريق	97	1.08
113	ناس	25207	0.84	قليل	97	1.07
114	شركة	24875	0.83	كأن	96	1.07
115	دولي	24817	0.83	كي	96	1.06
116	طريق	24751	0.83	جديد	95	1.06
117	عدد	24721	0.82	هنا	95	1.03
118	أرض	24627	0.82	سيارة	93	1.03

¹⁵⁵ Este lema no distingue entre verbo y sustantivo (*maşdar*). Se optó por colapsarlos en uno debido a que Buckwalter y Parkinson solo recogen un lema, etiquetado como verbo.

119	وَجَدَ	24528	0.82	وَصَلَ	93	1.02
120	سَبَب	24502	0.82	أَحَب	92	1.02
121	شعب	24222	0.81	أُم	92	1.01
122	شاء	24097	0.80	إِنَّ	91	1.01
123	صورة	24052	0.80	صار	91	1.00
124	أراد	23965	0.80	سَنَّة	90	0.99
125	عادَ	23911	0.80	دَخَلَ	89	0.97
126	أَنِ	23885	0.80	تحت	87	0.96
127	س	23562	0.79	بلد	86	0.96
128	أهل	23479	0.78	فوق	86	0.94
129	الذين	23446	0.78	جَسَد	85	0.93
130	شَكَلَ	23021	0.77	جاء	84	0.92
131	عَيْنُ	22982	0.77	جميع	83	0.92
132	عملية	22811	0.76	طَلَبَ	83	0.92
133	أمام	22799	0.76	مدينة	83	0.92
134	وطني	22704	0.76	عَمَلَ	82	0.89
135	خاص	22528	0.75	أصبح	80	0.89
136	أما	22527	0.75	ذَهَبَ	80	0.89
137	عنى	22428	0.75	نَحَوَ	80	0.89
138	أيضا	22381	0.75	كلام	80	0.88
139	عَمِلَ	22316	0.74	كيف	79	0.86
140	حَوْلَ	22294	0.74	أخ	77	0.86
141	بل	22123	0.74	بعيد	77	0.86
142	خير	22099	0.74	ثلاث	77	0.86
143	عندما	21928	0.73	ساعة	77	0.86
144	مدينة	21718	0.72	كلمة	77	0.84
145	ابن	21714	0.72	بقي	76	0.84
146	طَيِّبَ	21587	0.72	جلس	76	0.84
147	مشروع	21424	0.71	زَالَ	76	0.84
148	يد	21407	0.71	والد	76	0.83
149	ثلاث	21101	0.70	انتظر	75	0.83
150	منذ	21011	0.70	داخل	75	0.83

151	صوت	20881	0.70	نظرة	75	0.82
152	نسبة	20802	0.69	حيث	74	0.82
153	لدى	20719	0.69	ماء	73	0.81
154	لن	20626	0.69	وَقَفَ	73	0.80
155	برنامج	20293	0.68	ظَلَّ	72	0.79
156	سَمِعَ	20116	0.67	أتى	71	0.79
157	حال	20003	0.67	حاول	71	0.79
158	شهر	19967	0.67	حَوْلَ	71	0.79
159	هنا	19821	0.66	موت	71	0.79
160	عراقي	19794	0.66	نساء	71	0.78
161	حرب	19739	0.66	جِدَّ	70	0.78
162	وَضَعَ	19640	0.65	رغم	70	0.77
163	أم	19626	0.65	ماذا	69	0.76
164	أكد	19464	0.65	أخير	68	0.76
165	هُم	19420	0.65	جَعَلَ	68	0.76
166	جِدَّ	19258	0.64	طفل	68	0.74
167	دور	19242	0.64	أمكن	67	0.74
168	وَجْهَ	19108	0.64	ربما	67	0.74
169	بَدَأَ	19025	0.63	سَبَبَ	67	0.72
170	أصبح	18931	0.63	دار	65	0.72
171	صار	18833	0.63	سؤال	65	0.71
172	سلام	18801	0.63	جانب	64	0.71
173	كلمة	18787	0.63	نَكَرَ	64	0.71
174	طفل	18707	0.62	صباح	64	0.71
175	إسلامي	18613	0.62	قوة	64	0.71
176	مكان	18520	0.62	ليلة	64	0.70
177	مشكلة	18438	0.61	عُمَرَ	63	0.70
178	عَشْرَ	18273	0.61	خارج	63	0.70
179	وَصَلَ	18236	0.61	فَتَحَ	63	0.70
180	أَفَدَ	18150	0.61	ناس	63	0.69
181	جانب	18004	0.60	أبيض	62	0.69
182	شيخ	17884	0.60	أكبر	62	0.69

183	ساعة	17755	0.59	حَمَلٌ	62	0.69
184	حكى	17673	0.59	خَبَّرَ	62	0.69
185	سيد	17543	0.58	شارع	62	0.69
186	نظام	17514	0.58	عاش	62	0.69
187	عدم	17472	0.58	مات	62	0.69
188	صاحب	17450	0.58	هُمُ	62	0.68
189	ها	17352	0.58	تمام	61	0.68
190	إسرائيلي	17343	0.58	ثاني	61	0.68
191	لبناني	17309	0.58	زوج	61	0.68
192	رب	17092	0.57	فقط	61	0.67
193	حركة	16997	0.57	لون	60	0.67
194	كتاب	16965	0.57	وَلَدٌ	60	0.66
195	أكبر	16943	0.56	بل	59	0.66
196	حَدٌّ	16894	0.56	عائلة	59	0.66
197	أَخَذَ	16858	0.56	ليل	59	0.64
198	ماضي	16829	0.56	أبد	58	0.64
199	نَعَم	16762	0.56	بِلا	58	0.64
200	ذات	16679	0.56	ذات	58	0.64
201	قضية	16601	0.55	صَمَّتْ	58	0.64
202	إنسان	16596	0.55	فَهَمَ	58	0.64
203	رغم	16548	0.55	قام	58	0.63
204	مُتَّجِدٌ	16529	0.55	امرأة	57	0.63
205	أخير	16524	0.55	حال	57	0.63
206	حالة	16488	0.55	حَقَّ	57	0.63
207	أخت	16432	0.55	سوى	57	0.63
208	جامعة	16373	0.55	شيخ	57	0.62
209	قرار	16341	0.54	صديق	56	0.62
210	و	16314	0.54	غريب	56	0.62
211	حُبٌّ	16244	0.54	فَكَرَّ	56	0.62
212	استطاع	16240	0.54	لدى	56	0.62
213	لجنة	16153	0.54	نوم	56	0.61
214	وزارة	16143	0.54	انتهى	55	0.61

215	زَالٌ	16026	0.53	أَمَّا	55	0.60
216	مُجْتَمِعٌ	15867	0.53	خَلْفٌ	54	0.60
217	رَأْسٌ	15864	0.53	رَاحٌ	54	0.60
218	أَمْسٌ	15763	0.53	وَحِيدٌ	54	0.60
219	هَدَفٌ	15682	0.52	وَرَقٌ	54	0.59
220	تَحْتَ	15655	0.52	الَّذِينَ	53	0.59
221	حَلْوٌ	15620	0.52	شَهْرٌ	53	0.59
222	شُكْرٌ	15492	0.52	عَالَمٌ	53	0.58
223	مَرْكَزٌ	15483	0.52	حُلْمٌ	52	0.58
224	إِدَارَةٌ	15434	0.51	ضَجَجَكَ	52	0.58
225	عِلَاقَةٌ	15425	0.51	طَرَفٌ	52	0.57
226	مَا	15396	0.51	أَجَابٌ	51	0.57
227	نَتِيجَةٌ	15357	0.51	أَحْسٌ	51	0.57
228	مَسْلَمٌ	15307	0.51	بِمَا	51	0.57
229	حِينٌ	15052	0.50	فَجَاءَ	51	0.57
230	وَجِبٌ	15042	0.50	مَرٌّ	51	0.56
231	رَأْيٌ	15012	0.50	خُبٌّ	50	0.56
232	وَحْدَةٌ	14958	0.50	زَوْجَةٌ	50	0.56
233	قَانُونٌ	14954	0.50	كَادٌ	50	0.56
234	فَقَطٌ	14945	0.50	كَتَبٌ	50	0.54
235	صَغِيرٌ	14834	0.49	أَهْلٌ	49	0.54
236	أَمَةٌ	14824	0.49	حَيٌّ	49	0.54
237	قَدٌّ	14795	0.49	دَائِمٌ	49	0.54
238	بِمَا	14771	0.49	قَدِيمٌ	49	0.54
239	عَضْوٌ	14748	0.49	كِتَابٌ	49	0.54
240	أَمْنٌ	14715	0.49	كَمَا	49	0.54
241	مَاءٌ	14641	0.49	وَحْدٌ	49	0.53
242	فَرِيقٌ	14555	0.49	عَادَةٌ	48	0.53
243	شَأْنٌ	14518	0.48	لِمَاذَا	48	0.53
244	دَكْتُورٌ	14494	0.48	مَرْأَةٌ	48	0.53
245	مَوْقِعٌ	14428	0.48	مَنْ	48	0.52
246	قَدَّمَ	14411	0.48	ابْتَسَمَ	47	0.52

247	شخص	14382	0.48	شِعْر	47	0.52
248	مليون	14317	0.48	لَفَذٌ	47	0.52
249	مئة	14276	0.48	وَضَعَ	83	0.51
250	سيارة	14274	0.48	قرر	47	0.51
251	كلام	14240	0.47	إذ	46	0.51
252	داخل	14216	0.47	أعاد	46	0.51
253	أضاف	14215	0.47	حالة	46	0.51
254	فعل	14189	0.47	خَدِيبٌ	46	0.51
255	ماذا	14181	0.47	رَدَّ	46	0.51
256	ولاية	14150	0.47	رائحة	46	0.51
257	بنت	14110	0.47	شَكَّل	46	0.51
258	مهم	14105	0.47	شمس	46	0.51
259	مختلف	14095	0.47	ظَنَّ	46	0.50
260	مجال	14094	0.47	هكذا	46	0.50
261	قطاع	14028	0.47	بينما	45	0.50
262	بقي	13980	0.47	تَحَدَّثَ	45	0.50
263	نَظَرٌ ¹⁵⁶	13973	0.47	طبيب	45	0.50
264	مِمَّا	13967	0.47	عبد	45	0.50
265	ذَكَرَ	13960	0.47	وراء	45	0.49
266	وجود	13951	0.47	وَضَعَ	45	0.49
267	فترة	13943	0.46	ثوب	44	0.49
268	تحقيق	13875	0.46	خاص	44	0.49
269	شَب	13854	0.46	علاقة	44	0.48
270	حتى	13792	0.46	فِعْلٌ	44	0.48
271	أم	13775	0.46	أَشْبَهَ	43	0.48
272	سابق	13759	0.46	أمل	43	0.48
273	سؤال	13720	0.46	بنت	43	0.48
274	حزب	13699	0.46	جميل	43	0.48
275	مستوى	13696	0.46	حَدَّ	43	0.48
276	طالِبٌ	13692	0.46	رَفَعَ	43	0.48
277	حَلَّ	13635	0.45	صَدَّرَ	43	0.48

¹⁵⁶ El lema se refiere al sustantivo نَظَرٌ. Sin embargo, los signos diacríticos se eligieron para que Excel pudiera distinguir entre este lema y su homógrafo referente al verbo.

278	فيما	13495	0.45	سيد	43	0.48
279	اتحاد	13364	0.45	قَتَل	43	0.47
280	أشار	13362	0.45	نحن	43	0.47
281	نَحَوَ	13340	0.44	أدرك	42	0.47
282	حَمَدَ	13315	0.44	عَشْرَ	42	0.46
283	مؤسسة	13224	0.44	كلما	42	0.46
284	عُمِرَ	13094	0.44	حقيقة	41	0.46
285	دائم	13082	0.44	غادر	41	0.46
286	دراسة	13037	0.43	قرأ	41	0.46
287	أجل	13027	0.43	مكتب	41	0.46
288	خَبِرَ	13016	0.43	نسي	41	0.44
289	تاريخ	13012	0.43	نصف	41	0.44
290	موقف	12978	0.43	بَحَثَ	40	0.44
291	جَعَلَ	12881	0.43	حركة	40	0.44
292	مجموعة	12856	0.43	رغبة	40	0.44
293	اقتصادي	12838	0.43	مِمَّا	40	0.43
294	سَلَّمَ	12735	0.42	ظَهَرَ	39	0.43
295	سوق	12718	0.42	أين	39	0.43
296	حالي	12597	0.42	صرخ	39	0.43
297	دَخَلَ	12590	0.42	صلاة	39	0.43
298	باب	12585	0.42	عَلِمَ	39	0.43
299	طلع	12562	0.42	كتف	31	0.43
300	عالمي	12522	0.42	قُرْبَ	39	0.42
301	نائب	12514	0.42	نافذة	39	0.42
302	خدمة	12380	0.41	مدرسة	38	0.42
303	مدرسة	12353	0.41	أطلق	38	0.42
304	نوع	12318	0.41	أعلى	38	0.42
305	حُكْمَ	12301	0.41	بحر	38	0.42
306	كَذَلِكَ	12279	0.41	تذكر	38	0.42
307	أتى	12231	0.41	جبل	38	0.42
308	جميل	12198	0.41	دم	38	0.42
309	ضد	12137	0.40	رواية	38	0.42

310	جهة	12124	0.40	شقة	38	0.42
311	سأل	12075	0.40	عربي	38	0.41
312	قول	12036	0.40	مال	38	0.41
313	سلطة	11978	0.40	وسط	37	0.41
314	طويل	11914	0.40	أعطى	37	0.41
315	كان	11773	0.39	ألم	37	0.41
316	حقيقة	11773	0.39	حدّث	37	0.41
317	حاول	11772	0.39	روح	37	0.41
318	عسكري	11770	0.39	سرعة	37	0.41
319	مواطن	11765	0.39	فتاة	37	0.41
320	مرأة	11747	0.39	قريب	37	0.41
321	مشاركة	11721	0.39	قهوة	37	0.41
322	حتى	11707	0.39	مضى	37	0.40
323	لما	11598	0.39	منزل	37	0.40
324	بدأ	11547	0.38	نزل	36	0.40
325	أحب	11536	0.38	أحمر	36	0.40
326	حمل	11514	0.38	حبيب	36	0.40
327	جيش	11497	0.38	دفع	36	0.40
328	عاش	11475	0.38	ملبس	36	0.39
329	لماذا	11470	0.38	نام	36	0.39
330	خمس	11406	0.38	وقع	35	0.39
331	جهاز	11379	0.38	اتجاه	35	0.39
332	بحث	11374	0.38	أسد	35	0.39
333	دين	11346	0.38	بد	35	0.39
334	خاصة	11333	0.38	جدار	35	0.39
335	وضع	11331	0.38	طريقة	35	0.39
336	جرى	11306	0.38	عم	35	0.39
337	طبع	11291	0.38	قادم	35	0.38
338	اعتبر	11281	0.38	نهر	35	0.38
339	مسؤول	11182	0.37	وجهة	34	0.38
340	كم	11171	0.37	أستاذ	34	0.38
341	مثل	11170	0.37	تأمل	34	0.38

342	حبيب	11167	0.37	حكاية	34	0.38
343	اجتماعي	11095	0.37	خطوة	34	0.38
344	خرج	11082	0.37	شجرة	34	0.38
345	أعطى	11077	0.37	شخص	34	0.38
346	إسلام	11031	0.37	شعور	34	0.38
347	رقم	11013	0.37	عالي	34	0.38
348	أ	11013	0.37	فكر	34	0.38
349	فوق	10908	0.36	كامل	34	0.38
350	خارج	10903	0.36	نار	34	0.37
351	سياسة	10867	0.36	نعم	34	0.37
352	أربع	10823	0.36	نهاية	33	0.37
353	بعيد	10807	0.36	التفت	33	0.37
354	مرحلة	10806	0.36	جهة	33	0.37
355	كتب	10747	0.36	خزن	33	0.37
356	جين	10724	0.36	حقيقية	33	0.37
357	كي	10724	0.36	خروج	33	0.37
358	مصدر	10701	0.36	رسالة	33	0.37
359	وطن	10693	0.36	سجن	33	0.37
360	وَأَلَد	10683	0.36	شفة	33	0.36
361	حَدِيثٌ	10667	0.36	صاحب	33	0.36
362	إضافة	10665	0.36	قَبِلَ	32	0.36
363	دم	10621	0.35	ابتنسامة	32	0.36
364	صحيح	10605	0.35	تاريخ	32	0.36
365	ثالث	10560	0.35	تَوَقَّفَ	32	0.36
366	مؤتمر	10549	0.35	حَلٌّ	32	0.36
367	حَدَّثَ	10546	0.35	خير	32	0.36
368	ربما	10486	0.35	درى	32	0.36
369	مدير	10445	0.35	ذراع	32	0.36
370	مصري	10434	0.35	رأي	32	0.36
371	طرف	10407	0.35	زيارة	32	0.36
372	عَبْرَ	10358	0.35	سرير	32	0.36
373	قليل	10329	0.34	عام	32	0.36

374	دَعْم	10286	0.34	عودة	32	0.36
375	عَلِمَ	10266	0.34	شَاب	40	0.34
376	دولار	10209	0.34	فيما	32	0.34
377	واقع	10193	0.34	قَدْر	32	0.34
378	نهاية	10187	0.34	نَظَر	32	0.34
379	لقاء	10174	0.34	أَجَل	31	0.34
380	انتخاب	10144	0.34	بقية	31	0.34
381	تعليم	10122	0.34	جرى	31	0.34
382	شرق	10112	0.34	حرب	31	0.34
383	كامل	10092	0.34	حكى	31	0.34
384	دعا	10054	0.34	دين	31	0.34
385	ظَلَّ	10006	0.33	رؤية	31	0.34
386	إذ	10000	0.33	سَارَ	31	0.34
387	أفضل	9989	0.33	سر	31	0.34
388	مباراة	9985	0.33	طوال	31	0.34
389	أسبوع	9972	0.33	عدد	31	0.34
390	مال	9951	0.33	قَدَّمَ	31	0.34
391	أعلن	9920	0.33	معرفة	31	0.34
392	روح	9915	0.33	نهار	31	0.34
393	بناء	9894	0.33	هواء	31	0.34
394	أعلى	9854	0.33	واقع	31	0.33
395	ليل	9832	0.33	وجب	31	0.33
396	ممکن	9784	0.33	أخرج	30	0.33
397	تعاون	9755	0.33	خلال	30	0.33
398	صديق	9751	0.33	زمن	30	0.33
399	نادي	9733	0.32	زواج	30	0.33
400	رَدَّ	9732	0.32	عدم	30	0.33
401	طريقة	9727	0.32	عَقَلَ	30	0.33
402	لاعب	9647	0.32	عَلِمَ	30	0.33
403	هيئة	9595	0.32	عَمِلَ	30	0.33
404	لحظة	9522	0.32	قَدَّمَ	30	0.33
405	أدى	9519	0.32	قرية	30	0.33

406	معلومة	9494	0.32	قضى	30	0.33
407	بلغ	9476	0.32	مجرد	30	0.33
408	سوف	9471	0.32	منطقة	30	0.32
409	فرصة	9470	0.32	موضوع	30	0.32
410	جنوب	9470	0.32	ارتدى	29	0.32
411	أين	9321	0.31	انتبه	29	0.32
412	فَعَلَ	9309	0.31	أخبر	29	0.32
413	قل	9296	0.31	أسبوع	29	0.32
414	عديد	9291	0.31	أصاب	29	0.32
415	درى	9258	0.31	أفضل	29	0.32
416	اعتقد	9201	0.31	أقل	29	0.32
417	قصة	9186	0.31	إنسان	29	0.32
418	بداية	9185	0.31	بارد	29	0.32
419	مَلِك	9153	0.31	تَحَوَّلَ	29	0.32
420	تعالى	9087	0.30	ثقیل	29	0.32
421	إجراء	9084	0.30	جيد	29	0.32
422	حَسَبَ	9079	0.30	حَبَّ	29	0.32
423	هؤلاء	9044	0.30	سَبَقَ	29	0.32
424	مادة	9026	0.30	ظَهَرَ	29	0.32
425	سعودي	9006	0.30	قادر	29	0.32
426	قيمة	9002	0.30	كفى	29	0.32
427	دقيقة	8932	0.30	كَمْ	29	0.32
428	رَدَّ	8870	0.30	ماضي	29	0.32
429	نَظَرَ	8856	0.30	مستشفى	29	0.32
430	تَحَدَّثَ	8852	0.30	هدوء	29	0.32
431	رسالة	8844	0.29	هَزَّ	29	0.31
432	اجتماع	8843	0.29	هؤلاء	29	0.31
433	خط	8830	0.29	احتاج	28	0.31
434	سعر	8806	0.29	التقى	28	0.31
435	عَرَضَ	8798	0.29	تَقَدَّمَ	28	0.31
436	حوار	8792	0.29	حاجة	28	0.31
437	حَصَلَ	8769	0.29	خوف	28	0.31

438	نقطة	8743	0.29	دعا	28	0.31
439	ل	8710	0.29	دولة	28	0.31
440	ذَهَبَ	8703	0.29	سماء	28	0.31
441	تَنْفِيذَ	8675	0.29	شاي	28	0.31
442	منظمة	8647	0.29	ضوء	28	0.31
443	قَوِيَّ	8637	0.29	عام	28	0.31
444	قادم	8633	0.29	عنى	28	0.31
445	معنى	8629	0.29	قول	28	0.30
446	تمنى	8580	0.29	ل	28	0.30
447	لغة	8559	0.29	أخت	27	0.30
448	لازم	8549	0.28	أسود	27	0.30
449	موجود	8547	0.28	أصل	27	0.30
450	أَهَمَّ	8546	0.28	دقيقة	27	0.30
451	بد	8543	0.28	سوف	27	0.30
452	صحيفة	8543	0.28	شك	27	0.30
453	مشكور	8543	0.28	صَدَّقَ	27	0.30
454	صباح	8524	0.28	صندوق	27	0.30
455	تمام	8523	0.28	قصة	27	0.30
456	مرض	8490	0.28	لغة	27	0.30
457	أقل	8481	0.28	محل	27	0.30
458	تالي	8475	0.28	مَلَكَ	27	0.30
459	لِدَلِكْ	8472	0.28	نادر	27	0.30
460	خصوص	8455	0.28	نوع	27	0.30
461	زوج	8453	0.28	واضح	27	0.29
462	صحة	8421	0.28	واقف	27	0.29
463	وَقَعَ	8415	0.28	اقترب	26	0.29
464	صلى	8403	0.28	أذن	26	0.29
465	طَلَبَ	8391	0.28	أربع	26	0.29
466	منتدى	8390	0.28	أشار	26	0.29
467	شَهِدَ	8378	0.28	تَحَرَّكَ	26	0.29
468	شارع	8360	0.28	جيش	26	0.29
469	نار	8346	0.28	حضر	26	0.29

470	سوى	8312	0.28	خط	26	0.29
471	ترك	8295	0.28	داخلي	26	0.29
472	درجة	8282	0.28	دخول	26	0.29
473	قائد	8278	0.28	رَدِّ	26	0.29
474	أخر	8267	0.28	ردد	26	0.29
475	فكرة	8227	0.27	زجاج	26	0.29
476	مصلحة	8227	0.27	سِنِّ	26	0.29
477	مرور	8201	0.27	عامِلٌ	26	0.29
478	إنما	8192	0.27	عَبْرَ	26	0.29
479	قريب	8185	0.27	عَلَقَ	26	0.29
480	وسط	8164	0.27	فترة	26	0.29
481	موت	8147	0.27	قاد	26	0.29
482	أساس	8141	0.27	لاحظ	26	0.28
483	واضح	8125	0.27	مَنَعَ	26	0.28
484	قيادة	8047	0.27	إحساس	25	0.28
485	زيارة	8046	0.27	أرسل	25	0.28
486	رائع	8028	0.27	أهم	25	0.28
487	بيان	8020	0.27	ثالث	25	0.28
488	جيد	8006	0.27	دنيا	25	0.28
489	عودة	7992	0.27	رَفَضَ	25	0.28
490	اتصال	7989	0.27	سلام	25	0.28
491	احتلال	7979	0.27	سمح	25	0.28
492	أمني	7967	0.27	سياسي	25	0.28
493	وَقَفَ	7924	0.26	طاولة	25	0.28
494	تقرير	7919	0.26	غَضَبَ	25	0.28
495	سَلِمَ	7888	0.26	فَرَحَ	25	0.28
496	أكيد	7875	0.26	فرصة	25	0.28
497	دنيا	7872	0.26	قراءة	25	0.28
498	كرة	7866	0.26	كَذَلِكَ	25	0.28
499	رسول	7845	0.26	كلب	25	0.28
500	زوجة	7840	0.26	لقاء	25	0.27
501	نساء	7826	0.26	مهم	25	0.27

502	مقاومة	7766	0.26	اثن	24	0.27
503	صَحَّ	7762	0.26	أقام	24	0.27
504	عدة	7758	0.26	إِيًّا	24	0.27
505	شَعَرَ	7748	0.26	تمنى	24	0.27
506	إعادة	7694	0.26	تفصيل	24	0.27
507	سلاح	7684	0.26	جمال	24	0.27
508	ثقافة	7674	0.26	درجة	24	0.27
509	قديم	7669	0.26	سوق	24	0.27
510	لون	7668	0.26	شخصية	24	0.27
511	تنمية	7662	0.26	صَعَب	24	0.27
512	زاد	7657	0.26	طعام	24	0.27
513	عالي	7576	0.25	غطى	24	0.27
514	بحر	7560	0.25	كتابة	24	0.27
515	مَرَّ	7524	0.25	مُعَلِّم	24	0.27
516	عِلْم	7502	0.25	مفتوح	24	0.27
517	جزء	7495	0.25	مفهى	24	0.27
518	فَنَحَّ	7487	0.25	ملا	24	0.27
519	داخلي	7466	0.25	ملمح	24	0.27
520	تقديم	7457	0.25	مَنَحَ	24	0.26
521	محكمة	7435	0.25	وجود	24	0.26
522	زيادة	7424	0.25	اهتمام	23	0.26
523	تطوير	7422	0.25	أغنية	23	0.26
524	حربة	7416	0.25	بحيرة	23	0.26
525	نَقَلَ	7415	0.25	بنو	23	0.26
526	ليلة	7411	0.25	جالس	23	0.26
527	هكذا	7408	0.25	جماعة	23	0.26
528	ضرورة	7401	0.25	رحلة	23	0.26
529	أوضح	7400	0.25	زار	23	0.26
530	مقبل	7399	0.25	زاوية	23	0.26
531	فِكْر	7397	0.25	زميل	23	0.26
532	انتهى	7390	0.25	ساعد	23	0.26
533	وراء	7372	0.25	ساكن	23	0.26

534	غرفة	7358	0.25	شَرِبَ	23	0.26
535	غريب	7346	0.24	صعد	23	0.26
536	وسيلة	7325	0.24	صف	23	0.26
537	فني	7291	0.24	طَبَعَ	23	0.26
538	حقيقي	7272	0.24	عبارة	23	0.26
539	بينما	7250	0.24	عجوز	23	0.26
540	احتاج	7243	0.24	عمق	23	0.26
541	مستقبل	7237	0.24	غياب	23	0.26
542	دفاع	7233	0.24	قائل	23	0.26
543	غربي	7217	0.24	قرار	23	0.26
544	أساسي	7216	0.24	قطار	23	0.26
545	جمعية	7203	0.24	مساء	23	0.26
546	علمي	7188	0.24	مشهد	23	0.26
547	منزل	7183	0.24	مثنى	23	0.26
548	سبيل	7153	0.24	مَقْبَر	23	0.26
549	أبد	7129	0.24	مقعد	23	0.26
550	أثر	7125	0.24	ميت	23	0.24
551	محلي	7121	0.24	وصول	23	0.24
552	مواجهة	7117	0.24	ارتفع	22	0.24
553	ثقافي	7056	0.24	أخفى	22	0.24
554	سوري	7026	0.23	أوقف	22	0.24
555	إعلام	7023	0.23	جندي	22	0.24
556	أصل	7020	0.23	جوار	22	0.24
557	رسمي	7019	0.23	حارة	22	0.24
558	عِلْم	7014	0.23	حجر	22	0.24
559	أوروبي	7003	0.23	خَمْس	22	0.24
560	زمن	6996	0.23	خيمة	22	0.24
561	أستاذ	6991	0.23	زاد	22	0.24
562	كَرِيم	6989	0.23	سعيد	22	0.24
563	فرنسي	6976	0.23	سَمَى	22	0.24
564	أهمية	6963	0.23	شاهدَ	22	0.24
565	فيلم	6929	0.23	صامت	22	0.24

566	مكتب	6924	0.23	عزيز	22	0.24
567	مدى	6917	0.23	غاب	22	0.24
568	ورق	6916	0.23	فكرة	22	0.24
569	والد	6914	0.23	قَبْل	22	0.24
570	خارجية	6914	0.23	قَصْر	22	0.24
571	قاعدة	6878	0.23	لسان	22	0.24
572	شِعْر	6856	0.23	معنى	22	0.23
573	أمير	6843	0.23	ها	22	0.23
574	طبيعي	6829	0.23	ابتعد	21	0.23
575	منتخب	6828	0.23	أ	21	0.23
576	أبيض	6824	0.23	أظهر	21	0.23
577	عظيم	6811	0.23	أقرب	21	0.23
578	استخدام	6785	0.23	أَكْمَل	21	0.23
579	جندي	6781	0.23	ألقى	21	0.23
580	دخول	6772	0.23	تَبِعَ	21	0.23
581	بطولة	6755	0.23	جَدَّ	21	0.23
582	دورة	6738	0.22	جسم	21	0.23
583	اتفاق	6737	0.22	حاج	21	0.23
584	قَبْل	6731	0.22	حديقة	21	0.23
585	وصول	6728	0.22	حَسَنَ	21	0.23
586	خطوة	6719	0.22	حَصَلَ	21	0.23
587	تجربة	6687	0.22	حَمَام	21	0.23
588	دعوة	6683	0.22	رب	21	0.23
589	طَلَبُ	6681	0.22	سابق	21	0.23
590	حَدَّثَ	6675	0.22	ساحة	21	0.23
591	قدرة	6671	0.22	سفر	21	0.23
592	شاب	6663	0.22	سقط	21	0.23
593	قيام	6658	0.22	سَلَّمَ	21	0.23
594	مدة	6655	0.22	شاء	21	0.23
595	صراحة	6651	0.22	شب	21	0.23
596	ضِمْنٌ	6646	0.22	شديد	21	0.23
597	وَحْدَ	6640	0.22	عملية	21	0.23

598	بِلا	6635	0.22	عميق	21	0.23
599	فَهْمٌ	6620	0.22	قدرة	21	0.23
600	صَعَبٌ	6619	0.22	كامرا	21	0.23
601	مات	6614	0.22	مَدَّ	21	0.22
602	تغيير	6605	0.22	مفتاح	21	0.22
603	ست	6601	0.22	اعتبر	20	0.22
604	شرطة	6594	0.22	اعتقد	20	0.22
605	نشاط	6586	0.22	انتقل	20	0.22
606	وحيد	6577	0.22	اهتم	20	0.22
607	نصف	6574	0.22	إصبع	20	0.22
608	دارٌ	6566	0.22	أغلق	20	0.22
609	أوكي	6505	0.22	أنهى	20	0.22
610	بَعْدُ	6475	0.22	برد	20	0.22
611	إِيَّا	6471	0.22	بكى	20	0.22
612	بنك	6456	0.22	تابع	20	0.22
613	أوسط	6448	0.21	تَعَلَّمَ	20	0.22
614	جماعة	6437	0.21	تم	20	0.22
615	خطة	6433	0.21	جو	20	0.22
616	رجع	6428	0.21	حائط	20	0.22
617	مُحَافَظَةٌ	6386	0.21	خال	20	0.22
618	معرفة	6380	0.21	خَطَرَ	20	0.22
619	نجاح	6377	0.21	خفيف	20	0.22
620	شاعر	6364	0.21	ذا	20	0.22
621	ظَهَرَ	6363	0.21	رجع	20	0.22
622	إطلاق	6322	0.21	سَحَبَ	20	0.22
623	قَتَلَ	6311	0.21	سَكَنَ	20	0.22
624	كافة	6301	0.21	سلاح	20	0.22
625	عشرون	6296	0.21	شَقَّ	20	0.22
626	وكالة	6274	0.21	ضَحَّكَ	20	0.22
627	فرد	6261	0.21	صَرَبَ	20	0.22
628	شمال	6260	0.21	عادي	20	0.22
629	مدني	6252	0.21	عربة	20	0.22

630	عامِلٌ	6247	0.21	عشرون	20	0.22
631	عنوان	6247	0.21	فرنسي	20	0.22
632	حاجة	6242	0.21	فم	20	0.22
633	محاولة	6239	0.21	قصير	20	0.22
634	لعب	6236	0.21	محطة	20	0.22
635	مساعدة	6221	0.21	مختلف	20	0.22
636	يومي	6214	0.21	مدة	20	0.22
637	سواء	6208	0.21	واسع	20	0.21
638	حلم	6192	0.21	اختار	19	0.21
639	مملكة	6184	0.21	اختلفي	19	0.21
640	شديد	6170	0.21	اشترى	19	0.21
641	فكر	6168	0.21	أثار	19	0.21
642	قدم	6167	0.21	ألف	19	0.21
643	عزيز	6161	0.21	أمريكي	19	0.21
644	مزيد	6154	0.21	بداية	19	0.21
645	عادي	6124	0.20	بكاء	19	0.21
646	إن	6119	0.20	جار	19	0.21
647	اهتمام	6117	0.20	جاز	19	0.21
648	شعبي	6114	0.20	جلسة	19	0.21
649	طبيب	6108	0.20	جنة	19	0.21
650	صلاة	6100	0.20	حرية	19	0.21
651	يلا	6096	0.20	حقيقي	19	0.21
652	يهودي	6093	0.20	دام	19	0.21
653	إطار	6079	0.20	دور	19	0.21
654	عيد	6068	0.20	رسم	19	0.21
655	مساء	6066	0.20	رئيس	19	0.21
656	سرعة	6061	0.20	سطح	19	0.21
657	أمل	6059	0.20	طول	19	0.21
658	أسرة	6043	0.20	ظلام	19	0.21
659	أزمة	6034	0.20	عضو	19	0.21
660	استثمار	6008	0.20	غير	19	0.21
661	عقل	6004	0.20	قاتل	19	0.21

662	مَثَلٌ	6003	0.20	قَصَدَ	19	0.21
663	اِثْنٌ	6001	0.20	قَطَعَ	19	0.21
664	حُضُورٌ	5974	0.20	كَرِيمٌ	19	0.21
665	دَفَعَ	5963	0.20	كَنِيسَةٌ	19	0.21
666	حَقَّقَ	5946	0.20	لِذَلِكَ	19	0.21
667	أَمِينٌ	5939	0.20	مَشْرُوعٌ	19	0.21
668	جَهْدٌ	5918	0.20	مَطْبِخٌ	19	0.21
669	تَنْظِيمٌ	5914	0.20	نَسْبَةٌ	19	0.20
670	مِقَابِلٌ	5905	0.20	هَبِطَ	19	0.20
671	اِنْتَظَرَ	5903	0.20	اِتَّفَقَ	18	0.20
672	سَهْمٌ	5902	0.20	اِسْتَعَادَ	18	0.20
673	نُورٌ	5892	0.20	اِسْتَيْقِظَ	18	0.20
674	شَارَكَ	5880	0.20	اِنْتَظَرَ	18	0.20
675	أَثْنَاءَ	5868	0.20	اِنْطَلَقَ	18	0.20
676	قَرْيَةٌ	5843	0.19	أَخَذَ	18	0.20
677	تَحْدِيدٌ	5839	0.19	آخِرٌ	18	0.20
678	ظَرْفٌ	5834	0.19	أَمْسَ	18	0.20
679	مَشِيرٌ	5830	0.19	إِنْمَا	18	0.20
680	رَفَضَ	5802	0.19	أَيُّ	18	0.20
681	عَبْدٌ	5795	0.19	بَلَغَ	18	0.20
682	حَاجَةٌ	5784	0.19	بَنَاءٌ	18	0.20
683	قَائِلٌ	5777	0.19	تَزَوَّجَ	18	0.20
684	عَاصِمَةٌ	5764	0.19	تَصَوَّرَ	18	0.20
685	شَهِيدٌ	5762	0.19	تَعَرَّفَ	18	0.20
686	صَحِيٌّ	5759	0.19	تَغَيَّرَ	18	0.20
687	جَوْ	5758	0.19	تَكَلَّمَ	18	0.20
688	رِيَاضَةٌ	5746	0.19	ثِقَةٌ	18	0.20
689	نَزَلَ	5743	0.19	جِزْءٌ	18	0.20
690	اِسْتَمَرَ	5737	0.19	حَدَّثَ	18	0.20
691	تَعَرَّضَ	5689	0.19	حِذَاءَ	18	0.20
692	بَحَثَ	5653	0.19	حِسَابٌ	18	0.20
693	مَلَفٌ	5643	0.19	حُضُورٌ	18	0.20

694	خَافَ	5628	0.19	خيظ	18	0.20
695	كذا	5614	0.19	دكتور	18	0.20
696	فَضَّلَ	5613	0.19	راقب	18	0.20
697	إنساني	5603	0.19	رَسَمَ	18	0.20
698	قَدَرَ	5602	0.19	رغب	18	0.20
699	قسم	5602	0.19	زعيم	18	0.20
700	ديني	5598	0.19	سريع	18	0.20
701	عارف	5577	0.19	سقف	18	0.20
702	أسود	5560	0.19	صاح	18	0.20
703	إرهاب	5556	0.19	طابق	18	0.20
704	اعتبار	5516	0.18	عَدَّ	18	0.20
705	سبع	5505	0.18	فجر	18	0.20
706	عنصر	5500	0.18	فراش	18	0.20
707	سعادة	5499	0.18	فقد	18	0.20
708	شمس	5481	0.18	قضية	18	0.20
709	ما	5477	0.18	كاتب	18	0.20
710	عَرَبَ	5465	0.18	مرور	18	0.20
711	مضبوط	5462	0.18	مواجهة	18	0.20
712	رياضي	5453	0.18	موعد	18	0.20
713	وَقَّعَ	5446	0.18	مئة	18	0.20
714	إيراني	5424	0.18	نادى	18	0.20
715	اللهم	5404	0.18	نظام	18	0.20
716	جمهور	5401	0.18	نَفَسَ	18	0.20
717	فوز	5387	0.18	نهض	18	0.20
718	مستشفى	5386	0.18	هادئ	18	0.20
719	رئاسة	5379	0.18	واصل	18	0.20
720	شَكَّلَ	5374	0.18	وَعَدَ	18	0.19
721	محل	5373	0.18	ودع	18	0.19
722	جمهورية	5372	0.18	اعتاد	17	0.19
723	طاقة	5368	0.18	اكتشف	17	0.19
724	تعلق	5362	0.18	انتباه	17	0.19
725	أجنبي	5360	0.18	انفجر	17	0.19

726	ما	5345	0.18	أثناء	17	0.19
727	كاتب	5344	0.18	أزرق	17	0.19
728	زين	5343	0.18	إشارة	17	0.19
729	مالية	5342	0.18	أضاف	17	0.19
730	شخصية	5340	0.18	أكد	17	0.19
731	ذو	5336	0.18	بُعد	17	0.19
732	حجم	5334	0.18	بقاء	17	0.19
733	تَطَوَّر	5334	0.18	تركبي	17	0.19
734	صناعة	5316	0.18	تلقى	17	0.19
735	عافية	5312	0.18	ثمن	17	0.19
736	إصابة	5301	0.18	جُرْح	17	0.19
737	نَشْر	5291	0.18	جلد	17	0.19
738	اتصل	5289	0.18	جلوس	17	0.19
739	نهائي	5285	0.18	جَمَعَ	17	0.19
740	مَنْ	5271	0.18	خالي	17	0.19
741	نسي	5270	0.18	راهب	17	0.19
742	ضَغَط	5269	0.18	رصاص	17	0.19
743	قرأ	5267	0.18	رقم	17	0.19
744	غد	5254	0.18	ست	17	0.19
745	معروف	5252	0.18	سلطة	17	0.19
746	سجن	5246	0.17	سيده	17	0.19
747	متى	5240	0.17	صبي	17	0.19
748	فتاة	5229	0.17	صحة	17	0.19
749	تحية	5215	0.17	صراخ	17	0.19
750	جريمة	5204	0.17	صيف	17	0.19
751	رابع	5202	0.17	عَدَّ	17	0.19
752	سماء	5198	0.17	عَرَضَ	17	0.19
753	طول	5194	0.17	فَتَحَ	17	0.19
754	نووي	5190	0.17	فور	17	0.19
755	عدو	5189	0.17	قتيل	17	0.19
756	قمة	5186	0.17	قَطَعَ	17	0.19
757	سمو	5183	0.17	قِطْعَة	17	0.19

758	شُغِلَ	5144	0.17	لعل	17	0.19
759	مسلح	5138	0.17	لهجة	17	0.19
760	مسألة	5136	0.17	مسألة	17	0.19
761	كُون	5136	0.17	مشعر	17	0.19
762	إقامة	5135	0.17	مطر	17	0.19
763	زمان	5131	0.17	مطعم	17	0.19
764	أيها	5122	0.17	منتصف	17	0.19
765	مَلَك	5118	0.17	هَرَبَ	17	0.19
766	تَقَدَّمَ	5117	0.17	واجب	17	0.19
767	حماية	5112	0.17	وَصَفَّ	17	0.19
768	رَفَعَ	5104	0.17	وَقَّفَ	17	0.19
769	ساكن	5093	0.17	يمين	17	0.18
770	علاج	5073	0.17	يومي	17	0.18
771	عَمَّ	5063	0.17	إِذْنَ	16	0.18
772	طالِبَ	5063	0.17	استمر	16	0.18
773	عادة	5056	0.17	استمع	16	0.18
774	حُزْنَ	5054	0.17	أسفل	16	0.18
775	قَدَّرَ	5050	0.17	أَكَلَ	16	0.18
776	خارجي	5049	0.17	بَحَثَ	16	0.18
777	جسم	5040	0.17	تجاه	16	0.18
778	زواج	5033	0.17	تقريب	16	0.18
779	شبكة	5022	0.17	تمكن	16	0.18
780	قرر	5019	0.17	تَنَاولَ	16	0.18
781	كَشَفَ	5018	0.17	جسر	16	0.18
782	حَدِيثَ	5014	0.17	جيب	16	0.18
783	حساب	5004	0.17	حارس	16	0.18
784	ارتفاع	5004	0.17	حجرة	16	0.18
785	حصول	4989	0.17	حصان	16	0.18
786	أطلق	4988	0.17	حَمَى	16	0.18
787	إمارة	4969	0.17	خشبي	16	0.18
788	مجرد	4961	0.17	خشي	16	0.18
789	ديمقراطي	4961	0.17	خيال	16	0.18

790	شخصي	4953	0.17	دهشة	16	0.18
791	مهمة	4947	0.16	ذاك	16	0.18
792	أحسن	4942	0.16	رحل	16	0.18
793	مسؤولية	4941	0.16	ركبة	16	0.18
794	أصاب	4927	0.16	سعادة	16	0.18
795	نصريح	4918	0.16	سيجارة	16	0.18
796	مسجد	4912	0.16	شتاء	16	0.18
797	صهيوني	4882	0.16	صحيح	16	0.18
798	عدّ	4879	0.16	ضخم	16	0.18
799	مُشترِك	4868	0.16	طالب	16	0.18
800	بال	4864	0.16	ظِلّ	16	0.18
801	بريطاني	4863	0.16	عاري	16	0.18
802	أردني	4853	0.16	عدة	16	0.18
803	صدر	4847	0.16	عَرَض	16	0.18
804	سعى	4841	0.16	غد	16	0.18
805	شرط	4834	0.16	غطاء	16	0.18
806	فن	4834	0.16	فرد	16	0.18
807	نفس	4822	0.16	فيلم	16	0.18
808	تجاري	4819	0.16	قائد	16	0.18
809	إنتاج	4817	0.16	قَتْل	16	0.18
810	استمرار	4816	0.16	قَدَر	16	0.18
811	غالي	4811	0.16	قَوِيّ	16	0.18
812	صمّت	4809	0.16	كرسي	16	0.18
813	مليار	4803	0.16	كوب	16	0.18
814	اتجاه	4800	0.16	لحم	16	0.18
815	أقام	4800	0.16	لَعِبَ	16	0.18
816	قَتْل	4789	0.16	مبنى	16	0.18
817	مناسبة	4785	0.16	متعب	16	0.18
818	حوالي	4772	0.16	مريض	16	0.18
819	خطأ	4767	0.16	مزيد	16	0.18
820	إعلان	4765	0.16	مسافة	16	0.18
821	عَقَدَ	4748	0.16	مصري	16	0.18

822	دائرة	4742	0.16	مَلِك	16	0.17
823	تربية	4731	0.16	نُور	16	0.17
824	مطلوب	4730	0.16	ارتاح	15	0.17
825	نبي	4728	0.16	التقط	15	0.17
826	كأس	4718	0.16	انسحب	15	0.17
827	رئيسي	4715	0.16	أسرة	15	0.17
828	ظَلَّ	4710	0.16	أسعد	15	0.17
829	ك	4708	0.16	إيقاع	15	0.17
830	سمح	4704	0.16	إيمان	15	0.17
831	قرن	4700	0.16	بات	15	0.17
832	سر	4699	0.16	بنى	15	0.17
833	اشتغل	4697	0.16	تَأَكَّدَ	15	0.17
834	ناحية	4693	0.16	تراب	15	0.17
835	جلسة	4691	0.16	تساءل	15	0.17
836	خروج	4688	0.16	جامعة	15	0.17
837	حُرَّ	4684	0.16	جثة	15	0.17
838	نפט	4684	0.16	جَمَعَ	15	0.17
839	مناسب	4681	0.16	حجم	15	0.17
840	تَوَقَّعَ	4678	0.16	حُسْن	15	0.17
841	ماشني	4671	0.16	خجل	15	0.17
842	ممثل	4670	0.16	دائرة	15	0.17
843	قرآن	4667	0.16	دعوة	15	0.17
844	سارَ	4666	0.16	دَفَعَ	15	0.17
845	حكومي	4659	0.16	دمع	15	0.17
846	قناة	4655	0.16	دوري	15	0.17
847	ساحة	4652	0.16	ذكري	15	0.17
848	حاكم	4646	0.15	رفيق	15	0.17
849	أغنية	4645	0.15	سافر	15	0.17
850	صراع	4642	0.15	سهولة	15	0.17
851	صحفي	4642	0.15	شَعَلَ	15	0.17
852	عهد	4637	0.15	صَنَعَ	15	0.17
853	عالم	4624	0.15	ضد	15	0.17

854	مُعْظَم	4617	0.15	ضَبِيق	15	0.17
855	رُؤْيَة	4610	0.15	طَيِّب	15	0.17
856	حَضْر	4598	0.15	ظَرْف	15	0.17
857	تَحْرِير	4588	0.15	عَابِر	15	0.17
858	رَفَع	4577	0.15	عَنْف	15	0.17
859	مَرِيض	4575	0.15	عَنْق	15	0.17
860	وَاسِع	4568	0.15	عَهْد	15	0.17
861	تَعَامَل	4559	0.15	غَبَار	15	0.17
862	عَانِي	4558	0.15	فَرْقَة	15	0.17
863	حَمَلَة	4553	0.15	قِتَال	15	0.17
864	نَصَّ	4548	0.15	كَشَفَ	15	0.17
865	عَنْف	4537	0.15	لَعْنَة	15	0.17
866	حَيَّ	4532	0.15	لَمَّا	15	0.17
867	قَادِر	4522	0.15	مَتَى	15	0.17
868	قَانِم	4520	0.15	مَجْمُوعَة	15	0.17
869	فَتَّحَ	4517	0.15	مَجْنُون	15	0.17
870	عَائِلَة	4514	0.15	مَحَاوَلَة	15	0.17
871	مَضَى	4502	0.15	مَحِيط	15	0.17
872	جَبَل	4501	0.15	مَدخَل	15	0.17
873	لُعْب	4501	0.15	نَاحِيَة	15	0.17
874	طَبِي	4497	0.15	نَائِم	15	0.17
875	شَاكَ	4487	0.15	نِظَارَة	15	0.17
876	جَلَسَ	4487	0.15	هَدِيَة	15	0.16
877	مَعِين	4469	0.15	وَادِي	15	0.16
878	إِنشَاء	4444	0.15	اتَّصَلَ	14	0.16
879	رِيَال	4440	0.15	اِحْتِمَال	14	0.16
880	ضَمَّ	4427	0.15	ازْدَاد	14	0.16
881	ضَوء	4426	0.15	اسْتَحْدَم	14	0.16
882	إِصْلَاح	4424	0.15	اسْتَدَارَ	14	0.16
883	نَقَلَ	4421	0.15	انْتَابَ	14	0.16
884	بَسِيط	4418	0.15	أَجْمَلَ	14	0.16
885	وَقَّفَ	4417	0.15	أَحَاطَ	14	0.16

886	أداء	4402	0.15	أَصْر	14	0.16
887	سلامة	4387	0.15	أَمْسَكْ	14	0.16
888	بطل	4386	0.15	بَاع	14	0.16
889	ألم	4385	0.15	بَشِّرْ	14	0.16
890	رجا	4383	0.15	تَلَّ	14	0.16
891	تلفون	4380	0.15	تَوَقَّفْ	14	0.16
892	تشكيل	4379	0.15	جَدَلْ	14	0.16
893	دمع	4377	0.15	جَنَسْ	14	0.16
894	سريع	4375	0.15	حَدَّثْ	14	0.16
895	لِما	4361	0.15	حَقْلْ	14	0.16
896	وفد	4355	0.15	خَافْ	14	0.16
897	تدريب	4334	0.14	خَشَبْ	14	0.16
898	بَدَلْ	4315	0.14	خَطَأْ	14	0.16
899	خوف	4297	0.14	نَرَجْ	14	0.16
900	فنان	4297	0.14	نَرَسْ	14	0.16
901	ثمانية	4296	0.14	ذَاكِرَة	14	0.16
902	اختيار	4296	0.14	رَكُضْ	14	0.16
903	صالح	4284	0.14	سَهْلْ	14	0.16
904	سَجَّلْ	4277	0.14	شَارِكْ	14	0.16
905	فرقة	4274	0.14	صَخْرْ	14	0.16
906	كتابة	4267	0.14	صَفْحَة	14	0.16
907	تجارة	4265	0.14	ضَعَطْ	14	0.16
908	عَصْرُ	4253	0.14	طَالْ	14	0.16
909	ممارسة	4246	0.14	عَاشِقْ	14	0.16
910	قاعد	4229	0.14	عَائِدْ	14	0.16
911	اقتصاد	4227	0.14	فَاجَأْ	14	0.16
912	كلية	4225	0.14	فَتْحَة	14	0.16
913	طائرة	4225	0.14	فَخَذْ	14	0.16
914	جَسَدْ	4224	0.14	فَصَّلْ	14	0.16
915	صفحة	4220	0.14	فِضَاءْ	14	0.16
916	تطبيق	4213	0.14	فَنَدِيقْ	14	0.16
917	سيدة	4204	0.14	قَلِقْ	14	0.16

918	ديمقراطية	4193	0.14	كَلَمَ	14	0.16
919	مباشر	4192	0.14	كيس	14	0.16
920	شَكَرَ	4188	0.14	لَيْسَ	14	0.16
921	قصيدة	4188	0.14	مَرْكَزَ	14	0.16
922	تَابِعَ	4186	0.14	مُعَلَّقَ	14	0.16
923	خَطَرَ	4179	0.14	مُقَدَّم	14	0.16
924	عبارة	4174	0.14	ممکن	14	0.16
925	سمى	4171	0.14	مهمة	14	0.16
926	تمكن	4171	0.14	ناعم	14	0.16
927	لعل	4167	0.14	نَصَّ	14	0.16
928	توفير	4160	0.14	نقطة	14	0.16
929	ضَبَطَ	4157	0.14	نَقَلَ	14	0.16
930	شَعَرَ	4154	0.14	نَقَلَ	14	0.16
931	كلما	4144	0.14	هَمَّ	14	0.16
932	مُعَلِّم	4143	0.14	هما	14	0.14
933	بارك	4141	0.14	وَعِي	14	0.14
934	صف	4139	0.14	ادعى	13	0.14
935	تقريب	4137	0.14	امتد	13	0.14
936	ثقة	4126	0.14	انبعى	13	0.14
937	تَوَقَّفَ	4121	0.14	أَثَرَ	13	0.14
938	إفريقي	4118	0.14	إجابة	13	0.14
939	دوري	4115	0.14	أحرق	13	0.14
940	تقدير	4111	0.14	أَحْسَنَ	13	0.14
941	موظف	4098	0.14	أحضر	13	0.14
942	طبيعة	4090	0.14	أداة	13	0.14
943	عقوبة	4089	0.14	أعد	13	0.14
944	دليل	4078	0.14	أقبل	13	0.14
945	وَجَّهَ	4075	0.14	أيقظ	13	0.14
946	كاذب	4069	0.14	بال	13	0.14
947	نوم	4064	0.14	بطء	13	0.14
948	باقي	4053	0.14	بيع	13	0.14
949	تكلم	4053	0.14	تَأَخَّرَ	13	0.14

950	تاريخي	4053	0.14	تبادل	13	0.14
951	شامل	4050	0.14	تَجَاوَزَ	13	0.14
952	تابع	4036	0.13	تَرَدَّدَ	13	0.14
953	معركة	4030	0.13	تفكير	13	0.14
954	قومي	4019	0.13	تَوَقَّعَ	13	0.14
955	قراءة	4015	0.13	ثلج	13	0.14
956	سلطان	4015	0.13	ثورة	13	0.14
957	سفير	4009	0.13	جرؤ	13	0.14
958	ذكرى	4001	0.13	حُرِّ	13	0.14
959	محطة	3999	0.13	حَسَبَ	13	0.14
960	رغبة	3997	0.13	حَلَمَ	13	0.14
961	مرحبا	3987	0.13	خطاب	13	0.14
962	بقية	3986	0.13	خلاص	13	0.14
963	شَمِلَ	3985	0.13	خلفي	13	0.14
964	استقرار	3976	0.13	خيل	13	0.14
965	مالي	3972	0.13	دعاء	13	0.14
966	دينار	3968	0.13	دليل	13	0.14
967	أحمر	3966	0.13	ذهن	13	0.14
968	خلاف	3966	0.13	راضي	13	0.14
969	واجه	3934	0.13	رجا	13	0.14
970	إعلامي	3929	0.13	رعب	13	0.14
971	موسم	3914	0.13	ساد	13	0.14
972	مبلغ	3913	0.13	سفينة	13	0.14
973	جمعة	3911	0.13	سواد	13	0.14
974	مقالة	3911	0.13	شاطئ	13	0.14
975	شهادة	3904	0.13	شدة	13	0.14
976	ثورة	3893	0.13	صحيفة	13	0.14
977	نجح	3891	0.13	ضحكة	13	0.14
978	قُرْبَ	3891	0.13	ضرورة	13	0.14
979	كمبيوتر	3888	0.13	ضَغَطَ	13	0.14
980	نجم	3877	0.13	ظَهَرَ	13	0.14
981	جُرْحَ	3874	0.13	عائش	13	0.14

982	تَغَيَّرَ	3855	0.13	عرس	13	0.14
983	جَنَّة	3847	0.13	عرق	13	0.14
984	بدء	3839	0.13	عشاء	13	0.14
985	ذاك	3833	0.13	علا	13	0.14
986	كويتي	3833	0.13	عنوان	13	0.14
987	كفى	3827	0.13	عَيْش	13	0.14
988	استخدم	3825	0.13	غامض	13	0.14
989	ذَكَرَ	3824	0.13	غرض	13	0.14
990	سَبَقَ	3819	0.13	فراغ	13	0.14
991	نَشَرَ	3812	0.13	قبيلة	13	0.14
992	حضرة	3812	0.13	قَلْبَ	13	0.14
993	رحمة	3796	0.13	قميص	13	0.14
994	ضابط	3796	0.13	كَفَّ	13	0.14
995	خليج	3795	0.13	كأنما	13	0.14
996	صَدَّقَ	3768	0.13	لبناني	13	0.14
997	دام	3752	0.13	مزاج	13	0.14
998	غاية	3743	0.12	مقابل	13	0.14
999	مبادرة	3719	0.12	موجود	13	0.14
1000	مُؤَكَّد	3718	0.12	مؤخر	13	0.14
1001	نهار	3717	0.12	موقف	13	0.14
1002	نفسي	3715	0.12	نيرة	13	0.14
1003	أقصى	3713	0.12	نَقَدَ	13	0.13
1004	تعزير	3713	0.12	هدأ	13	0.13
1005	متابعة	3712	0.12	اتجه	12	0.13
1006	ألقي	3711	0.12	احتفظ	12	0.13
1007	أسلوب	3701	0.12	استحق	12	0.13
1008	تَحَوَّلَ	3697	0.12	امتلاً	12	0.13
1009	بيئة	3696	0.12	انتهاء	12	0.13
1010	شجرة	3691	0.12	انحنى	12	0.13
1011	نظرة	3687	0.12	أتم	12	0.13
1012	صَدَرَ	3682	0.12	أثبت	12	0.13
1013	طعام	3681	0.12	أخضر	12	0.13

1014	تأثير	3676	0.12	أدخل	12	0.13
1015	مستمر	3673	0.12	أسرع	12	0.13
1016	أخضر	3670	0.12	أصفر	12	0.13
1017	ضحك	3667	0.12	إطلاق	12	0.13
1018	مسابقة	3659	0.12	أعلن	12	0.13
1019	امرأة	3649	0.12	أقنع	12	0.13
1020	صندوق	3648	0.12	بائع	12	0.13
1021	تعبير	3645	0.12	بسيط	12	0.13
1022	معرض	3629	0.12	بندقية	12	0.13
1023	تأكيد	3627	0.12	تحرك	12	0.13
1024	عكس	3626	0.12	ترجع	12	0.13
1025	رعاية	3617	0.12	تطلع	12	0.13
1026	بات	3601	0.12	تعليم	12	0.13
1027	سقط	3600	0.12	ثمة	12	0.13
1028	موعد	3593	0.12	جامع	12	0.13
1029	نحو	3590	0.12	جملة	12	0.13
1030	إقليمي	3577	0.12	جهد	12	0.13
1031	جار	3576	0.12	جيل	12	0.13
1032	سبحان	3571	0.12	حاد	12	0.13
1033	قلم	3559	0.12	حادث	12	0.13
1034	قضاء	3557	0.12	حاكم	12	0.13
1035	ما	3557	0.12	حدق	12	0.13
1036	احترام	3551	0.12	حظ	12	0.13
1037	قائمة	3550	0.12	خائف	12	0.13
1038	ظن	3526	0.12	خد	12	0.13
1039	ظاهرة	3511	0.12	دخان	12	0.13
1040	ارتفع	3510	0.12	دخن	12	0.13
1041	لسان	3508	0.12	ذهاب	12	0.13
1042	رحلة	3507	0.12	رابع	12	0.13
1043	مشارك	3495	0.12	رمى	12	0.13
1044	اتفاقية	3491	0.12	ريح	12	0.13
1045	مشعر	3489	0.12	زحام	12	0.13

1046	خامس	3477	0.12	زينة	12	0.13
1047	إعداد	3477	0.12	سبع	12	0.13
1048	قانوني	3475	0.12	سكين	12	0.13
1049	تَقَدَّمَ	3471	0.12	سلطان	12	0.13
1050	إشارة	3465	0.12	سياسة	12	0.13
1051	خاف	3463	0.12	سيرة	12	0.13
1052	سفر	3461	0.12	شاحب	12	0.13
1053	تلفزيون	3461	0.12	شأن	12	0.13
1054	أظهر	3453	0.12	شبكة	12	0.13
1055	إمام	3450	0.12	شر	12	0.13
1056	نمو	3447	0.11	شَرَحَ	12	0.13
1057	معارضة	3442	0.11	شعاع	12	0.13
1058	تعديل	3441	0.11	شهادة	12	0.13
1059	أي	3436	0.11	شهوة	12	0.13
1060	خَلَقَ	3434	0.11	صراحة	12	0.13
1061	معدل	3433	0.11	صليب	12	0.13
1062	بشري	3413	0.11	صوب	12	0.13
1063	عَقْدَ	3412	0.11	ضابط	12	0.13
1064	فَقَّطَعَ	3403	0.11	طَلَّبَ	12	0.13
1065	أجاب	3397	0.11	طَقَّسَ	12	0.13
1066	إرهابي	3388	0.11	عثر	12	0.13
1067	جائزة	3386	0.11	عَكَّسَ	12	0.13
1068	دار	3384	0.11	عَلَامَة	12	0.13
1069	صفة	3378	0.11	غداء	12	0.13
1070	هجوم	3376	0.11	غسل	12	0.13
1071	فجر	3368	0.11	غناء	12	0.13
1072	ههه	3366	0.11	فارسي	12	0.13
1073	مفتوح	3365	0.11	فارغ	12	0.13
1074	مشهد	3365	0.11	فرحة	12	0.13
1075	مؤمن	3361	0.11	فريق	12	0.13
1076	مشى	3357	0.11	فنجان	12	0.13
1077	التقى	3356	0.11	قابَلْ	12	0.13

1078	ساعد	3354	0.11	قاعدة	12	0.13
1079	تحد	3349	0.11	قائمة	12	0.13
1080	زعيم	3346	0.11	قرن	12	0.13
1081	صادر	3334	0.11	قسم	12	0.13
1082	عامل	3329	0.11	قَفَر	12	0.13
1083	إيجابي	3322	0.11	قيمة	12	0.13
1084	جدار	3316	0.11	كابوس	12	0.13
1085	تَجَاوَزَ	3316	0.11	كأس	12	0.13
1086	عما	3313	0.11	كرة	12	0.13
1087	دَرَسَ	3312	0.11	كِلَا	12	0.13
1088	بَشَر	3311	0.11	لف	12	0.13
1089	أنتم	3308	0.11	لمح	12	0.13
1090	مبدأ	3303	0.11	مادة	12	0.13
1091	سوداني	3291	0.11	مثير	12	0.13
1092	سِنَّ	3282	0.11	مرض	12	0.13
1093	سنوي	3275	0.11	ممر	12	0.13
1094	إنجاز	3274	0.11	نبيذ	12	0.13
1095	قمر	3270	0.11	نهائي	12	0.13
1096	فئة	3267	0.11	هب	12	0.13
1097	تفكير	3265	0.11	هَمَسَ	12	0.12
1098	خميس	3265	0.11	واجه	12	0.12
1099	حَيَّ	3254	0.11	اتسع	11	0.12
1100	جَمَعَ	3254	0.11	اجتماع	11	0.12
1101	بيع	3252	0.11	احتضن	11	0.12
1102	صفر	3252	0.11	استسلم	11	0.12
1103	قصير	3248	0.11	أبدى	11	0.12
1104	مَنَع	3244	0.11	إجازة	11	0.12
1105	أدرك	3240	0.11	أدب	11	0.12
1106	شغلة	3239	0.11	أدبي	11	0.12
1107	لذا	3235	0.11	أغلب	11	0.12
1108	رواية	3233	0.11	أغض	11	0.12
1109	صناعي	3228	0.11	إقامة	11	0.12

1110	جزى	3222	0.11	أقصى	11	0.12
1111	لَ	3220	0.11	أقوى	11	0.12
1112	ثمن	3219	0.11	أَل	11	0.12
1113	حاضر	3215	0.11	أوشك	11	0.12
1114	مهندس	3209	0.11	بَدَل	11	0.12
1115	ضحية	3199	0.11	برنامج	11	0.12
1116	أعاد	3190	0.11	برهان	11	0.12
1117	مدرب	3189	0.11	تاجر	11	0.12
1118	جيل	3185	0.11	تالي	11	0.12
1119	آه	3182	0.11	تَجَمَّعَ	11	0.12
1120	ضرب	3180	0.11	تحية	11	0.12
1121	دق	3178	0.11	تَدَخَّلَ	11	0.12
1122	دستور	3176	0.11	تَدَخَّلَ	11	0.12
1123	أفاد	3174	0.11	تَصَرَّفَ	11	0.12
1124	مسرح	3167	0.11	تعبير	11	0.12
1125	دفع	3166	0.11	تغيير	11	0.12
1126	قاضي	3165	0.11	جبهة	11	0.12
1127	استعداد	3158	0.11	جريح	11	0.12
1128	قَبِلَ	3151	0.11	جلب	11	0.12
1129	أحلى	3150	0.11	جنب	11	0.12
1130	شعور	3147	0.10	جنيه	11	0.12
1131	أسف	3147	0.10	حصة	11	0.12
1132	غضب	3145	0.10	حُكْم	11	0.12
1133	شرقي	3140	0.10	خاصة	11	0.12
1134	صعيد	3138	0.10	خصوص	11	0.12
1135	انطلق	3135	0.10	خفي	11	0.12
1136	ضفة	3130	0.10	خَلَعَ	11	0.12
1137	اختلف	3129	0.10	دير	11	0.12
1138	مهرجان	3129	0.10	رقبة	11	0.12
1139	نبأ	3128	0.10	ساق	11	0.12
1140	مجلة	3115	0.10	سرقة	11	0.12
1141	تفصيل	3114	0.10	شخصي	11	0.12

1142	إداري	3114	0.10	شد	11	0.12
1143	فائدة	3113	0.10	شَعْر	11	0.12
1144	نصر	3112	0.10	شُكْر	11	0.12
1145	حرف	3111	0.10	شَيْطَان	11	0.12
1146	وَصَفَ	3111	0.10	صحراء	11	0.12
1147	هما	3107	0.10	صعوبة	11	0.12
1148	ثلاثون	3095	0.10	صلى	11	0.12
1149	خبرة	3090	0.10	ضَبُط	11	0.12
1150	اعتمد	3088	0.10	ضَرَب	11	0.12
1151	محدد	3086	0.10	ضعيف	11	0.12
1152	عَبَّرَ	3084	0.10	ضَيِّق	11	0.12
1153	خطاب	3084	0.10	ظالم	11	0.12
1154	أَلَّا	3080	0.10	عَبَّرَ	11	0.12
1155	انتظار	3078	0.10	عتبة	11	0.12
1156	مساحة	3077	0.10	عظيم	11	0.12
1157	فساد	3074	0.10	عنيف	11	0.12
1158	عامل	3072	0.10	غرق	11	0.12
1159	خسارة	3068	0.10	فرج	11	0.12
1160	حَطَّ	3065	0.10	فريد	11	0.12
1161	عدوان	3064	0.10	فساد	11	0.12
1162	سبب	3060	0.10	فَهْم	11	0.12
1163	سيما	3060	0.10	قاعة	11	0.12
1164	خطير	3056	0.10	قَدَّرَ	11	0.12
1165	مخيم	3053	0.10	قَلَق	11	0.12
1166	فجأة	3052	0.10	قوم	11	0.12
1167	لقي	3049	0.10	لاجئ	11	0.12
1168	تسعة	3049	0.10	لذة	11	0.12
1169	رَفَضَ	3048	0.10	لفت	11	0.12
1170	تضمن	3048	0.10	مُجْتَمَع	11	0.12
1171	ألماني	3041	0.10	مجلس	11	0.12
1172	أعظم	3035	0.10	مرأة	11	0.12
1173	شراء	3033	0.10	مرحلة	11	0.12

1174	صيف	3032	0.10	مسلم	11	0.12
1175	تَنَاوَل	3027	0.10	مصباح	11	0.12
1176	غياب	3023	0.10	مصدر	11	0.12
1177	نَمَّ	3018	0.10	مطار	11	0.12
1178	أثار	3015	0.10	معركة	11	0.12
1179	طير	3014	0.10	مفاجئ	11	0.12
1180	بنو	3011	0.10	ملون	11	0.12
1181	مثال	3008	0.10	نجح	11	0.12
1182	أحد	3004	0.10	نسخة	11	0.12
1183	قضى	2985	0.10	وافق	11	0.12
1184	مركزي	2981	0.10	وحدة	11	0.12
1185	رَجَمَ	2967	0.10	وضوح	11	0.11
1186	بلدية	2965	0.10	وقوف	11	0.11
1187	خبير	2963	0.10	استسلام	10	0.11
1188	مباشرة	2961	0.10	استقر	10	0.11
1189	عموم	2958	0.10	امتلك	10	0.11
1190	متوسط	2957	0.10	انتزع	10	0.11
1191	حاج	2956	0.10	انتشر	10	0.11
1192	شريف	2947	0.10	انتمى	10	0.11
1193	رضي	2944	0.10	إبرة	10	0.11
1194	جلالة	2944	0.10	أدار	10	0.11
1195	وظيفة	2939	0.10	أدى	10	0.11
1196	عَدَل	2939	0.10	إصرار	10	0.11
1197	سَهَّل	2922	0.10	أطل	10	0.11
1198	أجمل	2919	0.10	أمضى	10	0.11
1199	كهرياء	2915	0.10	إمكان	10	0.11
1200	ما	2915	0.10	بادر	10	0.11
1201	بكى	2907	0.10	باقي	10	0.11
1202	عمق	2904	0.10	بَشَّرَ	10	0.11
1203	جولة	2891	0.10	بطل	10	0.11
1204	حَسَّ	2887	0.10	بني	10	0.11
1205	نهر	2886	0.10	بوابة	10	0.11

1206	عميق	2885	0.10	بياض	10	0.11
1207	معالجة	2885	0.10	تأمل	10	0.11
1208	ريح	2883	0.10	تبيين	10	0.11
1209	جمال	2880	0.10	تجربة	10	0.11
1210	شر	2876	0.10	تحديد	10	0.11
1211	صبر	2875	0.10	تحرير	10	0.11
1212	شمالي	2875	0.10	تحسس	10	0.11
1213	عَفْدٌ	2874	0.10	تحقيق	10	0.11
1214	مناقشة	2863	0.10	تعلق	10	0.11
1215	أدب	2860	0.10	تمتم	10	0.11
1216	تصميم	2859	0.10	تَمَدَّدَ	10	0.11
1217	طَرَحَ	2858	0.10	تنظيم	10	0.11
1218	طَرَحَ	2857	0.10	تَوَجَّهَ	10	0.11
1219	سيطرة	2856	0.10	ثَقُلَ	10	0.11
1220	تَعَلَّمَ	2851	0.10	جاف	10	0.11
1221	سليبي	2851	0.10	جحيم	10	0.11
1222	مقر	2844	0.09	جر	10	0.11
1223	فَرَحَ	2841	0.09	جنازة	10	0.11
1224	ممن	2840	0.09	جواب	10	0.11
1225	سوء	2839	0.09	حاضر	10	0.11
1226	داع	2837	0.09	حرارة	10	0.11
1227	ضروري	2834	0.09	حَرَجَ	10	0.11
1228	جَدَّ	2834	0.09	حَرَمَ	10	0.11
1229	ضمان	2834	0.09	حزين	10	0.11
1230	تيار	2834	0.09	حقق	10	0.11
1231	مفهوم	2833	0.09	حمار	10	0.11
1232	لُعْبَةٌ	2832	0.09	حماية	10	0.11
1233	انتهاء	2831	0.09	حَمَلَ	10	0.11
1234	ميروك	2831	0.09	حوالي	10	0.11
1235	أسير	2827	0.09	خزانة	10	0.11
1236	أربعاء	2817	0.09	حيرة	10	0.11
1237	لوحة	2816	0.09	دافئ	10	0.11

1238	باحث	2815	0.09	دراسة	10	0.11
1239	قاد	2810	0.09	دراسي	10	0.11
1240	فقد	2804	0.09	دق	10	0.11
1241	تجاه	2800	0.09	ذَكَرَ	10	0.11
1242	مغربي	2798	0.09	ذئب	10	0.11
1243	حفل	2793	0.09	رَبَطَ	10	0.11
1244	برلمان	2793	0.09	رَجَمَ	10	0.11
1245	توفيق	2789	0.09	رضي	10	0.11
1246	شرعي	2785	0.09	رفيع	10	0.11
1247	بعدها	2779	0.09	روى	10	0.11
1248	انبعى	2779	0.09	زجاجة	10	0.11
1249	قَدَّرَ	2770	0.09	زمان	10	0.11
1250	حلقة	2764	0.09	ساخن	10	0.11
1251	هاتف	2764	0.09	سارع	10	0.11
1252	قبول	2760	0.09	سبيل	10	0.11
1253	توزيع	2757	0.09	سجاد	10	0.11
1254	بقي	2750	0.09	سَحَبَ	10	0.11
1255	كلب	2743	0.09	سَلَّمَ	10	0.11
1256	تذكر	2742	0.09	سواء	10	0.11
1257	اختار	2736	0.09	شاشة	10	0.11
1258	فور	2736	0.09	شبح	10	0.11
1259	تَحَقَّقَ	2731	0.09	شبه	10	0.11
1260	ذاكرة	2728	0.09	شرشف	10	0.11
1261	نام	2727	0.09	شرفة	10	0.11
1262	معهد	2725	0.09	شَكَّلَ	10	0.11
1263	خيار	2724	0.09	شَهِدَ	10	0.11
1264	ملعب	2722	0.09	صالون	10	0.11
1265	كمية	2721	0.09	صحبة	10	0.11
1266	حَفِظَ	2717	0.09	صحن	10	0.11
1267	دواء	2716	0.09	صداقة	10	0.11
1268	خَلَّصَ	2715	0.09	صوفي	10	0.11
1269	بالغ	2710	0.09	ضاق	10	0.11

1270	ضُعْف	2708	0.09	ضحية	10	0.11
1271	إيمان	2703	0.09	ضربة	10	0.11
1272	سامع	2703	0.09	ضفة	10	0.11
1273	شعار	2699	0.09	ضَمَّ	10	0.11
1274	سادس	2697	0.09	ضيعة	10	0.11
1275	هواء	2694	0.09	طائرة	10	0.11
1276	مادي	2691	0.09	طَبَّق	10	0.11
1277	هَدَفَ	2690	0.09	طن	10	0.11
1278	التزام	2685	0.09	ظهور	10	0.11
1279	جنوبي	2683	0.09	عاصفة	10	0.11
1280	إله	2681	0.09	عَبَّرَ	10	0.11
1281	خليجي	2680	0.09	عجيب	10	0.11
1282	اشترى	2677	0.09	عدا	10	0.11
1283	هدوء	2677	0.09	عش	10	0.11
1284	وجهة	2676	0.09	عَصُرَ	10	0.11
1285	متوقع	2676	0.09	عمارة	10	0.11
1286	ثمة	2675	0.09	عَيَّنَ	10	0.11
1287	تأمين	2673	0.09	غاية	10	0.11
1288	حَمَلَ	2668	0.09	غربة	10	0.11
1289	مسيرة	2668	0.09	غربي	10	0.11
1290	فَرَضَ	2666	0.09	فَرِحَ	10	0.11
1291	نظم	2666	0.09	فَضَّلَ	10	0.11
1292	تنفيذي	2664	0.09	قاسي	10	0.11
1293	منصب	2663	0.09	قديس	10	0.11
1294	معني	2653	0.09	قناة	10	0.11
1295	حرام	2652	0.09	قيام	10	0.11
1296	اتخاذ	2652	0.09	لباس	10	0.11
1297	إِمَّا	2647	0.09	لَعِبَ	10	0.11
1298	إيجاد	2646	0.09	مالك	10	0.11
1299	إلكتروني	2644	0.09	مباشرة	10	0.11
1300	اعتماد	2642	0.09	مُحامي	10	0.11
1301	اعترف	2636	0.09	مدى	10	0.11

1302	عَرَضَ	2624	0.09	مَسَحَ	10	0.11
1303	تسجيل	2623	0.09	مسرح	10	0.11
1304	جاز	2623	0.09	مشكلة	10	0.11
1305	عاد	2621	0.09	مصلحة	10	0.11
1306	واصل	2620	0.09	معروف	10	0.11
1307	ندوة	2616	0.09	مُعْظَم	10	0.11
1308	متر	2611	0.09	مغادرة	10	0.11
1309	هام	2606	0.09	مقدار	10	0.11
1310	غاز	2602	0.09	مكتوب	10	0.11
1311	نسخة	2601	0.09	مَنْع	10	0.11
1312	ابنة	2593	0.09	نبيل	10	0.11
1313	صعوبة	2586	0.09	نسيان	10	0.11
1314	بلدة	2586	0.09	نشاط	10	0.11
1315	دَرْس	2579	0.09	هاتف	10	0.11
1316	قَصْرٌ	2572	0.09	وعى	10	0.10
1317	حكاية	2569	0.09	يسار	10	0.10
1318	متعدد	2569	0.09	استغرق	9	0.10
1319	حادثة	2566	0.09	استقبل	9	0.10
1320	حارس	2565	0.09	اضطر	9	0.10
1321	حدد	2561	0.09	افترض	9	0.10
1322	تقنية	2553	0.09	امتحان	9	0.10
1323	إِذْن	2552	0.09	أَبْعَدَ	9	0.10
1324	متعلق	2544	0.08	أَبْعَدُ	9	0.10
1325	مكافحة	2541	0.08	أساس	9	0.10
1326	مؤشر	2537	0.08	أسمر	9	0.10
1327	كافي	2536	0.08	أَغْضَبَ	9	0.10
1328	حضارة	2528	0.08	أفق	9	0.10
1329	قَلَق	2521	0.08	أَكَل	9	0.10
1330	طال	2517	0.08	آلة	9	0.10
1331	قَصَدَ	2515	0.08	إله	9	0.10
1332	فكري	2512	0.08	أَمَّنَ	9	0.10
1333	أَثَّرَ	2509	0.08	أَنْجَبَ	9	0.10

1334	ازداد	2508	0.08	أنف	9	0.10
1335	عَقِبَ	2508	0.08	آه	9	0.10
1336	سَكَنَ	2504	0.08	أيها	9	0.10
1337	بُعِدَ	2504	0.08	بالغ	9	0.10
1338	داخِليّة	2496	0.08	برود	9	0.10
1339	ضعيف	2495	0.08	بِضْعٍ	9	0.10
1340	حيوان	2495	0.08	بطن	9	0.10
1341	أهلي	2487	0.08	بلدة	9	0.10
1342	أَكَلَ	2486	0.08	تَحَمَّلَ	9	0.10
1343	أجرى	2485	0.08	تَفَحَّصَ	9	0.10
1344	قوم	2484	0.08	تكرر	9	0.10
1345	أصدر	2481	0.08	تلا	9	0.10
1346	اثن	2480	0.08	ثانوي	9	0.10
1347	تعليق	2477	0.08	جرس	9	0.10
1348	اقترب	2475	0.08	جلدي	9	0.10
1349	عماني	2475	0.08	جمعة	9	0.10
1350	كيفية	2471	0.08	حدد	9	0.10
1351	طوال	2470	0.08	جِدِّي	9	0.10
1352	ظلم	2469	0.08	حديد	9	0.10
1353	آية	2467	0.08	حرف	9	0.10
1354	سعيد	2466	0.08	حرك	9	0.10
1355	مطار	2466	0.08	جِسِّ	9	0.10
1356	جبهة	2465	0.08	حِضْنٍ	9	0.10
1357	رَسَمَ	2458	0.08	حفيد	9	0.10
1358	انجليزي	2454	0.08	حَلَقَ	9	0.10
1359	مرشح	2452	0.08	حيوان	9	0.10
1360	نموذج	2450	0.08	خاطب	9	0.10
1361	آلية	2446	0.08	خبز	9	0.10
1362	رُبع	2441	0.08	خدم	9	0.10
1363	نافذة	2427	0.08	داكن	9	0.10
1364	مُشْتَرِك	2426	0.08	درب	9	0.10
1365	روسي	2425	0.08	دِقَّة	9	0.10

1366	صحافة	2423	0.08	ذهني	9	0.10
1367	قَصْد	2419	0.08	ذو	9	0.10
1368	بديل	2419	0.08	راحة	9	0.10
1369	مفاوضة	2418	0.08	ربيع	9	0.10
1370	أَمَن	2417	0.08	رَتَّبَ	9	0.10
1371	فقير	2413	0.08	رصيف	9	0.10
1372	مقارنة	2413	0.08	رف	9	0.10
1373	اتفق	2412	0.08	رَفَضَ	9	0.10
1374	تعليمي	2404	0.08	رَفَع	9	0.10
1375	مميز	2403	0.08	رمادي	9	0.10
1376	تهديد	2403	0.08	زيت	9	0.10
1377	سَهْلٌ	2403	0.08	ساخر	9	0.10
1378	إمكانية	2400	0.08	سادس	9	0.10
1379	سابع	2399	0.08	سائق	9	0.10
1380	فعالية	2399	0.08	سخرية	9	0.10
1381	يمين	2397	0.08	سرق	9	0.10
1382	متميز	2396	0.08	سري	9	0.10
1383	جريدة	2392	0.08	سلامة	9	0.10
1384	فندق	2391	0.08	سماع	9	0.10
1385	جُهْد	2390	0.08	سمعة	9	0.10
1386	مبكر	2390	0.08	سوري	9	0.10
1387	طائفي	2389	0.08	سير	9	0.10
1388	عفو	2388	0.08	شَكَرَ	9	0.10
1389	لاحظ	2383	0.08	شوق	9	0.10
1390	حفاظ	2368	0.08	صالة	9	0.10
1391	جوي	2367	0.08	صبر	9	0.10
1392	معقول	2363	0.08	صرخة	9	0.10
1393	فَهْم	2363	0.08	ضاع	9	0.10
1394	مَنَعَ	2363	0.08	طاقة	9	0.10
1395	وَلِيّ	2362	0.08	طبيعة	9	0.10
1396	توقيع	2361	0.08	طَرَحَ	9	0.10
1397	مقتل	2355	0.08	طرق	9	0.10

1398	محيط	2351	0.08	طير	9	0.10
1399	حجر	2350	0.08	طين	9	0.10
1400	أحس	2350	0.08	عجلة	9	0.10
1401	طيران	2342	0.08	عدو	9	0.10
1402	انتخابي	2341	0.08	عديد	9	0.10
1403	غالب	2334	0.08	عسكري	9	0.10
1404	فَقِيلَ	2332	0.08	عمى	9	0.10
1405	صحافي	2329	0.08	عناية	9	0.10
1406	صاروخ	2328	0.08	غمر	9	0.10
1407	شبيعة	2328	0.08	غيبوبة	9	0.10
1408	جهاد	2321	0.08	غيمة	9	0.10
1409	مفيد	2312	0.08	فارس	9	0.10
1410	مستقل	2308	0.08	فحص	9	0.10
1411	أقرب	2307	0.08	فلاح	9	0.10
1412	شدة	2302	0.08	قاطع	9	0.10
1413	مارس	2298	0.08	قبالة	9	0.10
1414	فَصْل	2297	0.08	قبر	9	0.10
1415	ضخم	2297	0.08	قسوة	9	0.10
1416	حرارة	2295	0.08	قطيع	9	0.10
1417	سير	2295	0.08	قلم	9	0.10
1418	واجب	2295	0.08	كَفَّ	9	0.10
1419	ساهم	2291	0.08	كاهن	9	0.10
1420	زراعة	2287	0.08	كائن	9	0.10
1421	ثلاثاء	2280	0.08	لحية	9	0.10
1422	ليبي	2279	0.08	لمس	9	0.10
1423	وادي	2276	0.08	ليلي	9	0.10
1424	هوية	2274	0.08	مباشر	9	0.10
1425	حادث	2273	0.08	متأخر	9	0.10
1426	بارد	2270	0.08	متأكد	9	0.10
1427	فضاء	2269	0.08	متر	9	0.10
1428	استحق	2265	0.08	متعة	9	0.10
1429	اتخذ	2265	0.08	مجاور	9	0.10

1430	زراعي	2265	0.08	مجلة	9	0.10
1431	جواب	2264	0.08	مجهول	9	0.10
1432	فم	2264	0.08	محدد	9	0.10
1433	سياق	2264	0.08	مخافة	9	0.10
1434	لقب	2263	0.08	مخيم	9	0.10
1435	تنسيق	2263	0.08	مساحة	9	0.10
1436	فاز	2258	0.08	مستعد	9	0.10
1437	متخصص	2258	0.08	مسجد	9	0.10
1438	إستراتيجي	2258	0.08	مصنع	9	0.10
1439	تحسين	2250	0.08	مصيبة	9	0.10
1440	عدالة	2249	0.07	مظهر	9	0.10
1441	تحوّل	2248	0.07	معرض	9	0.10
1442	شوق	2248	0.07	معلق	9	0.10
1443	تحرّك	2245	0.07	مفكر	9	0.10
1444	أفق	2245	0.07	ممرض	9	0.10
1445	جنيه	2245	0.07	مناسب	9	0.10
1446	صادق	2243	0.07	مناسبة	9	0.10
1447	تواصل	2243	0.07	منظر	9	0.10
1448	مقعد	2242	0.07	موسيقى	9	0.10
1449	انتشار	2241	0.07	موضع	9	0.10
1450	جثة	2240	0.07	موقع	9	0.10
1451	ظهر	2238	0.07	ميدان	9	0.10
1452	أبرز	2238	0.07	ميل	9	0.10
1453	جمع	2236	0.07	نجا	9	0.10
1454	ثوب	2235	0.07	نحيل	9	0.10
1455	دستوري	2233	0.07	نصر	9	0.10
1456	إستراتيجية	2232	0.07	هتف	9	0.10
1457	وعي	2231	0.07	وطن	8	0.10
1458	دعاء	2229	0.07	جنون	9	0.09
1459	مؤخر	2228	0.07	ودّ	9	0.09
1460	مصاب	2220	0.07	وزير	9	0.09
1461	ركن	2217	0.07	اتخذ	8	0.09

1462	روعة	2217	0.07	اتصال	8	0.09
1463	تمويل	2213	0.07	اجتاح	8	0.09
1464	رائحة	2204	0.07	احتل	8	0.09
1465	ميت	2199	0.07	احتمل	8	0.09
1466	موسيقى	2199	0.07	ارتباط	8	0.09
1467	وصف	2198	0.07	ارتبك	8	0.09
1468	توفر	2198	0.07	ارتد	8	0.09
1469	مبنى	2195	0.07	استقبال	8	0.09
1470	قاعة	2194	0.07	اشتعل	8	0.09
1471	احتمال	2193	0.07	اطمان	8	0.09
1472	أثبت	2189	0.07	افتقد	8	0.09
1473	طالما	2187	0.07	اكتفى	8	0.09
1474	وثيقة	2184	0.07	أبقى	8	0.09
1475	وافق	2182	0.07	أتاح	8	0.09
1476	انكليزي	2176	0.07	أتي	8	0.09
1477	زميل	2175	0.07	أجبر	8	0.09
1478	ممتاز	2175	0.07	أجرى	8	0.09
1479	شيعي	2175	0.07	أداء	8	0.09
1480	يمني	2174	0.07	أربعون	8	0.09
1481	حينما	2171	0.07	أسوأ	8	0.09
1482	تركي	2170	0.07	أسى	8	0.09
1483	قعد	2170	0.07	أشدّ	8	0.09
1484	جرّص	2166	0.07	أطول	8	0.09
1485	ظهور	2161	0.07	إعادة	8	0.09
1486	دل	2156	0.07	أفاق	8	0.09
1487	جزائري	2156	0.07	أمّر	8	0.09
1488	مطر	2153	0.07	أمّن	8	0.09
1489	مفروض	2152	0.07	أمّن	8	0.09
1490	قبيلة	2150	0.07	أمّني	8	0.09
1491	استفادة	2150	0.07	أمير	8	0.09
1492	ظُهر	2147	0.07	أنيق	8	0.09
1493	ثانوي	2144	0.07	أهلي	8	0.09

1494	جسر	2143	0.07	أيسر	8	0.09
1495	ملاحظة	2141	0.07	بَرَكَة	8	0.09
1496	اغتيال	2141	0.07	بساطة	8	0.09
1497	وزن	2140	0.07	بطاقة	8	0.09
1498	جملة	2138	0.07	تارك	8	0.09
1499	هدية	2134	0.07	تبقى	8	0.09
1500	أكل	2134	0.07	تخلى	8	0.09
1501	تلقي	2134	0.07	ترتيب	8	0.09
1502	سيادة	2133	0.07	تسليم	8	0.09
1503	كتلة	2132	0.07	تَعَامَلْ	8	0.09
1504	خفيف	2127	0.07	تفسير	8	0.09
1505	انتقل	2125	0.07	تلفزيون	8	0.09
1506	مسيحي	2124	0.07	تهمة	8	0.09
1507	سقوط	2123	0.07	تيار	8	0.09
1508	قتيل	2122	0.07	ثابت	8	0.09
1509	محافظ	2122	0.07	ثقافة	8	0.09
1510	إمكان	2120	0.07	جامعي	8	0.09
1511	همّ	2118	0.07	جرجير	8	0.09
1512	مصير	2117	0.07	جمهور	8	0.09
1513	تَمَتَّعَ	2117	0.07	جهاز	8	0.09
1514	تطلب	2116	0.07	جَهْلٌ	8	0.09
1515	عاشق	2116	0.07	جَهْلٌ	8	0.09
1516	مطلق	2114	0.07	حافظ	8	0.09
1517	قيد	2114	0.07	حجة	8	0.09
1518	احتفال	2113	0.07	حَدْرٌ	8	0.09
1519	وفاة	2112	0.07	حرس	8	0.09
1520	رَسْمٌ	2112	0.07	حريق	8	0.09
1521	نقد	2106	0.07	حضارة	8	0.09
1522	فتنة	2104	0.07	حَفِظْ	8	0.09
1523	هوى	2103	0.07	حُمَى	8	0.09
1524	أشبه	2102	0.07	حوض	8	0.09
1525	ضيف	2097	0.07	خدمة	8	0.09

1526	درهم	2097	0.07	خضرة	8	0.09
1527	موافقة	2094	0.07	خطيئة	8	0.09
1528	فَرْق	2092	0.07	خَلْق	8	0.09
1529	حَسَب	2088	0.07	خمن	8	0.09
1530	بعث	2087	0.07	داعب	8	0.09
1531	توجيه	2083	0.07	دافع	8	0.09
1532	حظ	2082	0.07	دَرَسَ	8	0.09
1533	فَضَّلَ	2080	0.07	دفع	8	0.09
1534	فَرَضَ	2080	0.07	دفتر	8	0.09
1535	إصدار	2076	0.07	دَقَّة	8	0.09
1536	بطاقة	2074	0.07	دورة	8	0.09
1537	سياحة	2071	0.07	ذنب	8	0.09
1538	تفضل	2067	0.07	راكب	8	0.09
1539	متقدم	2063	0.07	رث	8	0.09
1540	استهدف	2060	0.07	رحيل	8	0.09
1541	تدمير	2059	0.07	رَكِبَ	8	0.09
1542	بنية	2058	0.07	رسول	8	0.09
1543	صفقة	2058	0.07	زيون	8	0.09
1544	فقر	2056	0.07	زَحَفَ	8	0.09
1545	ظلام	2055	0.07	زقاق	8	0.09
1546	ثابت	2053	0.07	سارَ	8	0.09
1547	بقاء	2049	0.07	سرعان	8	0.09
1548	دَقِيقٌ	2049	0.07	سيف	8	0.09
1549	قَطَعَ	2046	0.07	سينما	8	0.09
1550	هائل	2044	0.07	شال	8	0.09
1551	شيطان	2044	0.07	شرق	8	0.09
1552	قطري	2044	0.07	شَمَّ	8	0.09
1553	وقوف	2043	0.07	شَمال	8	0.09
1554	درب	2043	0.07	شهيد	8	0.09
1555	وكيل	2041	0.07	صابون	8	0.09
1556	راحة	2039	0.07	صبية	8	0.09
1557	تبادل	2038	0.07	صحا	8	0.09

1558	استقبال	2034	0.07	صَدَرَ	8	0.09
1559	معيار	2033	0.07	صدفة	8	0.09
1560	مُقَدِّمَة	2031	0.07	صِدْق	8	0.09
1561	تَشْرِيْعِي	2030	0.07	صِعْر	8	0.09
1562	غرض	2024	0.07	صراع	8	0.09
1563	خائف	2023	0.07	صوتي	8	0.09
1564	وَرْد	2022	0.07	ضاحك	8	0.09
1565	انفجار	2021	0.07	ضُعْف	8	0.09
1566	حِفْظ	2017	0.07	طالِب	8	0.09
1567	شرف	2016	0.07	عانى	8	0.09
1568	سليم	2013	0.07	عاود	8	0.09
1569	أمانة	2013	0.07	عروس	8	0.09
1570	جامعي	2009	0.07	عزاء	8	0.09
1571	متهم	2009	0.07	عَشِيق	8	0.09
1572	اعتداء	2009	0.07	عصبي	8	0.09
1573	محور	2007	0.07	عظم	8	0.09
1574	هجمة	2004	0.07	عيد	8	0.09
1575	شبه	2003	0.07	فات	8	0.09
1576	إبداع	2003	0.07	فاكهة	8	0.09
1577	طب	2001	0.07	فر	8	0.09
1578	لحم	1999	0.07	فضح	8	0.09
1579	تحليل	1998	0.07	فك	8	0.09
1580	سيء	1995	0.07	قصيدة	8	0.09
1581	ترتيب	1993	0.07	قماش	8	0.09
1582	أمان	1991	0.07	قنديل	8	0.09
1583	سياحي	1991	0.07	كُبْر	8	0.09
1584	إقليم	1988	0.07	كتلة	8	0.09
1585	غاب	1984	0.07	كثرة	8	0.09
1586	اتهام	1982	0.07	كرر	8	0.09
1587	فرع	1977	0.07	كوخ	8	0.09
1588	أرسل	1976	0.07	لعبة	8	0.09
1589	رمز	1973	0.07	لمع	8	0.09

1590	امتحان	1970	0.07	مارس	8	0.09
1591	ذهاب	1970	0.07	مأساة	8	0.09
1592	إحساس	1968	0.07	مبلغ	8	0.09
1593	مقرر	1964	0.07	متابعة	8	0.09
1594	حكمة	1963	0.07	متقدم	8	0.09
1595	انسحاب	1962	0.07	متوسط	8	0.09
1596	مسافة	1960	0.07	محكمة	8	0.09
1597	مثلما	1959	0.07	مدفون	8	0.09
1598	اعتراف	1954	0.07	مراقبة	8	0.09
1599	ملح	1952	0.07	مرح	8	0.09
1600	ابتنم	1951	0.07	مساعدة	8	0.09
1601	تقليدي	1945	0.06	مَسَاك	8	0.09
1602	ما	1944	0.06	مسؤول	8	0.09
1603	متحدث	1943	0.06	مسيح	8	0.09
1604	أهمّ	1941	0.06	مصائب	8	0.09
1605	لواء	1940	0.06	مطلوب	8	0.09
1606	ميدان	1938	0.06	معتاد	8	0.09
1607	تحويل	1936	0.06	معدة	8	0.09
1608	دوم	1934	0.06	مكتبة	8	0.09
1609	مقدس	1929	0.06	ممتلئ	8	0.09
1610	مراقبة	1927	0.06	موظف	8	0.09
1611	تهمة	1927	0.06	موكب	8	0.09
1612	أجلّ	1926	0.06	نائب	8	0.09
1613	خمسون	1925	0.06	نسوة	8	0.09
1614	تَدَخَّل	1910	0.06	نسيم	8	0.09
1615	جار	1903	0.06	نظيف	8	0.09
1616	حَرَم	1903	0.06	نما	8	0.09
1617	جماعي	1902	0.06	نوى	8	0.09
1618	عَلَامَة	1897	0.06	هَدَف	8	0.09
1619	سُنَّة	1896	0.06	هن	8	0.09
1620	حَبّ	1894	0.06	هوى	8	0.09
1621	امتد	1894	0.06	هيا	8	0.09

1622	تركيز	1893	0.06	هيئة	8	0.09
1623	منتصف	1887	0.06	وائق	8	0.09
1624	انطلاق	1886	0.06	واضيع	8	0.09
1625	اعتقال	1886	0.06	وزارة	8	0.09
1626	حاد	1885	0.06	وفاة	8	0.08
1627	تَعَرَّفَ	1884	0.06	يقين	8	0.08
1628	صرَّحَ	1884	0.06	ابتلع	7	0.08
1629	حصار	1882	0.06	اثنين	7	0.08
1630	تَنَاولَ	1878	0.06	اجتاز	7	0.08
1631	منهج	1876	0.06	اجتماعي	7	0.08
1632	دراسي	1872	0.06	اجتمع	7	0.08
1633	ثامن	1870	0.06	احتفال	7	0.08
1634	إجمالي	1870	0.06	احتوى	7	0.08
1635	شاشة	1860	0.06	اختلف	7	0.08
1636	شاهد	1858	0.06	استبدل	7	0.08
1637	زار	1857	0.06	استشعر	7	0.08
1638	زاوية	1855	0.06	استعمل	7	0.08
1639	حزين	1855	0.06	استغرب	7	0.08
1640	أصلي	1854	0.06	اشند	7	0.08
1641	سينما	1854	0.06	اعتذر	7	0.08
1642	حَلَّ	1851	0.06	اعترض	7	0.08
1643	ملكي	1849	0.06	اقتراب	7	0.08
1644	محدود	1848	0.06	اكتمل	7	0.08
1645	افتتاح	1846	0.06	انتصب	7	0.08
1646	كثرة	1843	0.06	انشغال	7	0.08
1647	راهن	1840	0.06	أجرة	7	0.08
1648	ثقل	1838	0.06	أخذ	7	0.08
1649	صحراء	1838	0.06	أخلى	7	0.08
1650	حديقة	1836	0.06	إذن	7	0.08
1651	ذاتي	1831	0.06	أريكة	7	0.08
1652	غادر	1826	0.06	أشعل	7	0.08
1653	تام	1823	0.06	إصابة	7	0.08

1654	خشي	1822	0.06	أعلم	7	0.08
1655	بريء	1821	0.06	إقناع	7	0.08
1656	لفت	1821	0.06	ألماني	7	0.08
1657	بيّن	1820	0.06	أمامي	7	0.08
1658	قاتل	1820	0.06	أمانة	7	0.08
1659	مُنْتَج	1820	0.06	أمة	7	0.08
1660	مورد	1819	0.06	إنساني	7	0.08
1661	تَراجُع	1818	0.06	أنصت	7	0.08
1662	شريك	1813	0.06	أهمية	7	0.08
1663	وَلِي	1813	0.06	أوصل	7	0.08
1664	هجرة	1812	0.06	بدء	7	0.08
1665	قبر	1811	0.06	بدائي	7	0.08
1666	ضرر	1810	0.06	بدن	7	0.08
1667	مَدَّ	1809	0.06	برج	7	0.08
1668	منطق	1807	0.06	بشري	7	0.08
1669	اختلاف	1807	0.06	بناية	7	0.08
1670	هندي	1800	0.06	بيرا	7	0.08
1671	إجابة	1799	0.06	تام	7	0.08
1672	هادئ	1798	0.06	تجاري	7	0.08
1673	ذهبي	1797	0.06	تَحَكَّم	7	0.08
1674	وَلَدَ	1795	0.06	تخيل	7	0.08
1675	تَصْرَفُ	1795	0.06	تَرَدُّد	7	0.08
1676	مسلسل	1795	0.06	تسلل	7	0.08
1677	سلسلة	1794	0.06	تَصْرَفُ	7	0.08
1678	ليت	1791	0.06	تَعَامُل	7	0.08
1679	تَدَخَّل	1790	0.06	تعذيب	7	0.08
1680	جنسية	1788	0.06	تعمد	7	0.08
1681	تفجير	1788	0.06	تَقَدَّمَ	7	0.08
1682	استقبل	1787	0.06	تلفون	7	0.08
1683	سَبَب	1783	0.06	تمثال	7	0.08
1684	تزوج	1781	0.06	تَنَاول	7	0.08
1685	جوار	1781	0.06	توتر	7	0.08

1686	مناقسة	1779	0.06	توفيق	7	0.08
1687	نَقْد	1778	0.06	ثلاثون	7	0.08
1688	شريط	1777	0.06	جذع	7	0.08
1689	سلطنة	1776	0.06	جزيرة	7	0.08
1690	ضربة	1775	0.06	جلباب	7	0.08
1691	سلوك	1775	0.06	جولة	7	0.08
1692	هنالك	1772	0.06	حاجز	7	0.08
1693	سكر	1771	0.06	حالي	7	0.08
1694	مستحيل	1766	0.06	حان	7	0.08
1695	باع	1766	0.06	حَذَرَ	7	0.08
1696	تاسع	1765	0.06	حراسة	7	0.08
1697	ملبس	1764	0.06	حَطَّ	7	0.08
1698	تَجَاوَز	1764	0.06	حطب	7	0.08
1699	تخطيط	1763	0.06	حفرة	7	0.08
1700	حَسَن	1761	0.06	حَكَمَ	7	0.08
1701	بحري	1761	0.06	حكيم	7	0.08
1702	أدخل	1760	0.06	حلو	7	0.08
1703	مساعد	1760	0.06	حماس	7	0.08
1704	اهتم	1759	0.06	حملة	7	0.08
1705	أربعون	1759	0.06	حنان	7	0.08
1706	تفعيل	1756	0.06	حَوَّلَ	7	0.08
1707	صِدْق	1755	0.06	خادم	7	0.08
1708	واسطة	1755	0.06	خاض	7	0.08
1709	معلاة	1755	0.06	خافت	7	0.08
1710	صلة	1752	0.06	خبرة	7	0.08
1711	سني	1752	0.06	خسارة	7	0.08
1712	اكتشف	1751	0.06	خصم	7	0.08
1713	شقيق	1750	0.06	خطة	7	0.08
1714	انتصار	1750	0.06	خف	7	0.08
1715	غذائي	1749	0.06	خميس	7	0.08
1716	موحد	1747	0.06	دابة	7	0.08
1717	جراء	1746	0.06	دَفَنَ	7	0.08

1718	تشغيل	1746	0.06	دوم	7	0.08
1719	معتقل	1744	0.06	ذاب	7	0.08
1720	أداة	1743	0.06	ذبح	7	0.08
1721	ثروة	1737	0.06	ذهبي	7	0.08
1722	صنَع	1732	0.06	رامي	7	0.08
1723	تساءل	1730	0.06	رائع	7	0.08
1724	عَيَّرَ	1727	0.06	رخيص	7	0.08
1725	سافر	1727	0.06	رخيم	7	0.08
1726	رابطة	1726	0.06	رطوبة	7	0.08
1727	صبح	1723	0.06	رفقة	7	0.08
1728	رائد	1721	0.06	رقص	7	0.08
1729	لاحق	1721	0.06	رمل	7	0.08
1730	ثَرَدَدَ	1717	0.06	زر	7	0.08
1731	أملَ	1716	0.06	زهور	7	0.08
1732	أعرب	1716	0.06	زيتون	7	0.08
1733	مجهول	1715	0.06	سحر	7	0.08
1734	كارثة	1706	0.06	سراي	7	0.08
1735	حفل	1706	0.06	سرى	7	0.08
1736	هيا	1706	0.06	سعى	7	0.08
1737	احتوى	1705	0.06	سفح	7	0.08
1738	أقوى	1704	0.06	سيلم	7	0.08
1739	مسكين	1702	0.06	سور	7	0.08
1740	تمثل	1701	0.06	شارب	7	0.08
1741	مساهمة	1700	0.06	شرط	7	0.08
1742	حَكَمَ	1695	0.06	شرطة	7	0.08
1743	زيت	1695	0.06	شرقي	7	0.08
1744	أدنى	1695	0.06	شريط	7	0.08
1745	سَحَبَ	1694	0.06	صَبَاجِي	7	0.08
1746	حُسْنُ	1693	0.06	صدور	7	0.08
1747	عيب	1692	0.06	صفة	7	0.08
1748	أَخَذَ	1692	0.06	صَفَع	7	0.08
1749	طائفة	1692	0.06	صَمَّتْ	7	0.08

1750	مولى	1692	0.06	صنف	7	0.08
1751	أُنْ	1688	0.06	ضاحكة	7	0.08
1752	تاجر	1685	0.06	ضجيج	7	0.08
1753	أمر	1684	0.06	ضيمَنَ	7	0.08
1754	خالص	1684	0.06	ضمير	7	0.08
1755	سفينة	1683	0.06	ضيف	7	0.08
1756	استقلال	1680	0.06	طاف	7	0.08
1757	مرتبط	1679	0.06	طبي	7	0.08
1758	طبقة	1675	0.06	طبيعي	7	0.08
1759	مُحْتَلَّ	1674	0.06	عالج	7	0.08
1760	صرخ	1671	0.06	عالم	7	0.08
1761	شعري	1671	0.06	عثماني	7	0.08
1762	إنقاذ	1670	0.06	عَجَزَ	7	0.08
1763	عيط	1670	0.06	عَجَزَ	7	0.08
1764	جزيرة	1669	0.06	عذاب	7	0.08
1765	مزرع	1667	0.06	عريض	7	0.08
1766	تونسي	1667	0.06	عزلة	7	0.08
1767	قنبلة	1666	0.06	عسكر	7	0.08
1768	نَقْص	1664	0.06	عصا	7	0.08
1769	استفاد	1664	0.06	عصفور	7	0.08
1770	مجاور	1664	0.06	عطر	7	0.08
1771	شاهد	1662	0.06	علاج	7	0.08
1772	حفلة	1661	0.06	علبة	7	0.08
1773	دمار	1661	0.06	عيادة	7	0.08
1774	تراث	1661	0.06	عيال	7	0.08
1775	عجوز	1661	0.06	غاية	7	0.08
1776	ودع	1660	0.06	غروب	7	0.08
1777	مبارك	1659	0.06	غلب	7	0.08
1778	مُقَدَّم	1655	0.06	غنى	7	0.08
1779	آل	1654	0.06	فاتح	7	0.08
1780	صَرْف	1649	0.05	فتى	7	0.08
1781	بأس	1648	0.05	فَرْش	7	0.08

1782	عجيب	1647	0.05	فريسة	7	0.08
1783	سطح	1642	0.05	فَصَلَّ	7	0.08
1784	كرسي	1639	0.05	قاضي	7	0.08
1785	تأسيس	1639	0.05	قاوم	7	0.08
1786	أشد	1638	0.05	قبضة	7	0.08
1787	خَلَقَ	1636	0.05	قرآن	7	0.08
1788	تَبِعَ	1635	0.05	قعد	7	0.08
1789	صيني	1635	0.05	فيد	7	0.08
1790	إيطالي	1633	0.05	كافي	7	0.08
1791	خدم	1631	0.05	كثيف	7	0.08
1792	تبيين	1630	0.05	كِرَّة	7	0.08
1793	جَلَّ	1630	0.05	كَسَرَ	7	0.08
1794	شدد	1630	0.05	كلف	7	0.08
1795	أَكْمَلَ	1621	0.05	كلية	7	0.08
1796	أزرق	1621	0.05	كَمَلَ	7	0.08
1797	مُشْرِف	1621	0.05	كنب	7	0.08
1798	نية	1620	0.05	كومة	7	0.08
1799	شناء	1620	0.05	ليث	7	0.08
1800	حامِلٌ	1619	0.05	لجنة	7	0.08
1801	تشجيع	1619	0.05	لوح	7	0.08
1802	قد	1618	0.05	لوحة	7	0.08
1803	خَلَّصَ	1615	0.05	مائدة	7	0.08
1804	إرادة	1614	0.05	متعدد	7	0.08
1805	إرسال	1613	0.05	مجال	7	0.08
1806	مراجعة	1611	0.05	محيء	7	0.08
1807	دوام	1610	0.05	محبة	7	0.08
1808	مطلب	1607	0.05	محتمل	7	0.08
1809	راتب	1606	0.05	محفوظ	7	0.08
1810	بكاء	1603	0.05	مدير	7	0.08
1811	ناجح	1602	0.05	مذكرة	7	0.08
1812	هدد	1602	0.05	مذهب	7	0.08
1813	حديد	1601	0.05	مَرَّرَ	7	0.08

1814	قهوة	1601	0.05	مرسوم	7	0.08
1815	نقاش	1600	0.05	مُرْهَق	7	0.08
1816	نال	1599	0.05	مس	7	0.08
1817	تراب	1597	0.05	مستحيل	7	0.08
1818	مجاهد	1597	0.05	مستسلم	7	0.08
1819	مُحامي	1594	0.05	مستوى	7	0.08
1820	مواصلة	1590	0.05	مسكين	7	0.08
1821	وضوح	1588	0.05	مسموع	7	0.08
1822	مظهر	1587	0.05	مُعَن	7	0.08
1823	مسار	1587	0.05	مقاتل	7	0.08
1824	جاد	1585	0.05	مقارنة	7	0.08
1825	رغب	1585	0.05	مقاومة	7	0.08
1826	مضيف	1584	0.05	مقر	7	0.08
1827	قضائي	1581	0.05	مندبل	7	0.08
1828	عاجل	1578	0.05	مؤسسة	7	0.08
1829	تربوي	1578	0.05	نادي	7	0.08
1830	تَوَجُّه	1576	0.05	نال	7	0.08
1831	سلمي	1576	0.05	نبات	7	0.08
1832	رئاسي	1576	0.05	نبي	7	0.08
1833	مقدار	1575	0.05	نجمة	7	0.08
1834	تَحْرُك	1575	0.05	نداء	7	0.08
1835	إنهاء	1573	0.05	نَشْر	7	0.08
1836	تحالف	1573	0.05	نَطَق	7	0.08
1837	حاجز	1572	0.05	نقيب	7	0.08
1838	مستشار	1571	0.05	هجرة	7	0.08
1839	بريد	1569	0.05	هرع	7	0.08
1840	سباق	1569	0.05	هَمَس	7	0.08
1841	حَلَم	1569	0.05	هنالك	7	0.08
1842	سري	1567	0.05	هندي	7	0.08
1843	حصاري	1567	0.05	واجهه	7	0.08
1844	كيان	1566	0.05	وثيقة	7	0.08
1845	مهني	1563	0.05	وسيلة	7	0.08

1846	تصوير	1562	0.05	وضعية	7	0.08
1847	نطاق	1559	0.05	وطأة	7	0.08
1848	نيابي	1558	0.05	وطني	7	0.08
1849	تَوَصَّل	1556	0.05	وظيفة	7	0.08
1850	قلة	1555	0.05	ولادة	7	0.08
1851	ذهن	1553	0.05	يمني	7	0.07
1852	أهه	1552	0.05	يهودي	7	0.07
1853	نعمة	1549	0.05	اتفاق	6	0.07
1854	كَلَّا	1547	0.05	احتجاب	6	0.07
1855	جدول	1546	0.05	احتلال	6	0.07
1856	خليفة	1546	0.05	اختبأ	6	0.07
1857	أغلب	1545	0.05	اختفاء	6	0.07
1858	تداول	1545	0.05	اختلط	6	0.07
1859	ميزان	1542	0.05	ارتجف	6	0.07
1860	غني	1540	0.05	استحم	6	0.07
1861	تَرَاجَعَ	1539	0.05	استرسل	6	0.07
1862	خُلِقَ	1537	0.05	استعداد	6	0.07
1863	كرامة	1537	0.05	استمرار	6	0.07
1864	فاعل	1536	0.05	استولى	6	0.07
1865	مُؤَلَّف	1536	0.05	اعترف	6	0.07
1866	ضَيِّقَ	1534	0.05	اقتحم	6	0.07
1867	خسر	1533	0.05	التزام	6	0.07
1868	أولئك	1532	0.05	التصق	6	0.07
1869	جاهز	1524	0.05	انتظام	6	0.07
1870	سفارة	1524	0.05	انتقام	6	0.07
1871	مستثمر	1522	0.05	اندفاع	6	0.07
1872	تكنولوجيا	1521	0.05	اندفع	6	0.07
1873	وَدَّ	1520	0.05	انسحاب	6	0.07
1874	مختص	1520	0.05	انصرف	6	0.07
1875	مهنة	1518	0.05	انفجار	6	0.07
1876	شفة	1515	0.05	انقطع	6	0.07
1877	زهر	1514	0.05	انكسر	6	0.07

1878	مُوَجَّه	1514	0.05	انهار	6	0.07
1879	خص	1513	0.05	إجراء	6	0.07
1880	ضَرْب	1513	0.05	أجلس	6	0.07
1881	فوضى	1513	0.05	أجنبي	6	0.07
1882	مقابلة	1512	0.05	أخاف	6	0.07
1883	مؤقت	1512	0.05	أدهش	6	0.07
1884	وَعَدَ	1509	0.05	إرادة	6	0.07
1885	حجة	1509	0.05	أرملة	6	0.07
1886	رَبِحَ	1507	0.05	أزعج	6	0.07
1887	تَمَيَّرَ	1505	0.05	أسلوب	6	0.07
1888	مفاجأة	1504	0.05	أشرفَ	6	0.07
1889	تَأَكَّدَ	1504	0.05	أصدر	6	0.07
1890	معاناة	1503	0.05	إضافة	6	0.07
1891	روى	1503	0.05	أضفى	6	0.07
1892	ماما	1503	0.05	أعجب	6	0.07
1893	سرير	1502	0.05	إعلان	6	0.07
1894	حصاة	1501	0.05	أفاد	6	0.07
1895	قاسي	1499	0.05	أفرغ	6	0.07
1896	كردي	1499	0.05	أفريقي	6	0.07
1897	متقف	1498	0.05	أفلت	6	0.07
1898	تكوين	1497	0.05	أقدم	6	0.07
1899	طاولة	1495	0.05	ألح	6	0.07
1900	دبلوماسي	1495	0.05	ألغى	6	0.07
1901	عناية	1494	0.05	إمام	6	0.07
1902	فيروس	1494	0.05	أمسية	6	0.07
1903	تقبل	1493	0.05	أنتم	6	0.07
1904	توتر	1493	0.05	إنجليزي	6	0.07
1905	ظاهر	1492	0.05	أنزل	6	0.07
1906	علا	1492	0.05	أنوثة	6	0.07
1907	تسوية	1492	0.05	أهيف	6	0.07
1908	وَرَدَ	1490	0.05	أوحى	6	0.07
1909	سهولة	1488	0.05	أوصى	6	0.07

1910	مكتبة	1488	0.05	أولئك	6	0.07
1911	مطعم	1487	0.05	باحث	6	0.07
1912	تعامَل	1486	0.05	باشا	6	0.07
1913	سائق	1486	0.05	بدلة	6	0.07
1914	أعجب	1486	0.05	برهة	6	0.07
1915	استغلال	1484	0.05	بري	6	0.07
1916	مُحَافَظَة	1483	0.05	بريء	6	0.07
1917	موضح	1482	0.05	بريق	6	0.07
1918	مَنَحَ	1481	0.05	بقعة	6	0.07
1919	ضريبة	1477	0.05	بنك	6	0.07
1920	اختبار	1477	0.05	بؤس	6	0.07
1921	نقابة	1474	0.05	بيض	6	0.07
1922	حَوْلَ	1472	0.05	تاسع	6	0.07
1923	دِقَّة	1470	0.05	تسلق	6	0.07
1924	جريح	1470	0.05	تسمية	6	0.07
1925	ابتهامة	1470	0.05	تعالى	6	0.07
1926	نادر	1469	0.05	تعثر	6	0.07
1927	مصنع	1468	0.05	تعليق	6	0.07
1928	نادى	1468	0.05	تقرير	6	0.07
1929	فعلي	1468	0.05	تمرد	6	0.07
1930	ناقش	1468	0.05	تنظيف	6	0.07
1931	فراغ	1467	0.05	تنفس	6	0.07
1932	رصاص	1466	0.05	ثامن	6	0.07
1933	حكيم	1466	0.05	ثري	6	0.07
1934	ردد	1465	0.05	ثكنة	6	0.07
1935	فَصَلَّ	1461	0.05	جدوى	6	0.07
1936	فَصَّفَ	1457	0.05	جذب	6	0.07
1937	مذكور	1456	0.05	جريدة	6	0.07
1938	تقييم	1456	0.05	جريمة	6	0.07
1939	عَكَسَ	1452	0.05	جَلَّ	6	0.07
1940	مكتوب	1451	0.05	جوهر	6	0.07
1941	غذاء	1450	0.05	حاصر	6	0.07

1942	ذنب	1448	0.05	حافة	6	0.07
1943	خيال	1445	0.05	حتم	6	0.07
1944	مَنَح	1444	0.05	حجاب	6	0.07
1945	تَغَيَّر	1443	0.05	جِدَّة	6	0.07
1946	مستعد	1442	0.05	حَزَّ	6	0.07
1947	خطورة	1441	0.05	حزب	6	0.07
1948	مخاطر	1440	0.05	حشد	6	0.07
1949	فاق	1439	0.05	حصول	6	0.07
1950	نصيحة	1436	0.05	حَضَنَ	6	0.07
1951	سعي	1436	0.05	حليب	6	0.07
1952	تَحَمَّلَ	1433	0.05	حليف	6	0.07
1953	طفولة	1432	0.05	حماسة	6	0.07
1954	متواصل	1430	0.05	حماقة	6	0.07
1955	استعمال	1430	0.05	حمرة	6	0.07
1956	كَشَفَ	1430	0.05	حميم	6	0.07
1957	رصيد	1430	0.05	حنون	6	0.07
1958	كنيسة	1430	0.05	حوار	6	0.07
1959	تسليم	1429	0.05	خاتم	6	0.07
1960	سمك	1426	0.05	خارجي	6	0.07
1961	صدد	1426	0.05	خاطر	6	0.07
1962	قَدَّرَ	1424	0.05	خامس	6	0.07
1963	نهض	1423	0.05	خدر	6	0.07
1964	مفتاح	1421	0.05	خروف	6	0.07
1965	وارد	1420	0.05	خطا	6	0.07
1966	امتلك	1419	0.05	خطف	6	0.07
1967	نيابة	1419	0.05	خفر	6	0.07
1968	مصرف	1416	0.05	خلا	6	0.07
1969	نزاع	1415	0.05	خَلَقَ	6	0.07
1970	بنى	1413	0.05	خُلِقَ	6	0.07
1971	بِرَكَّة	1412	0.05	دافع	6	0.07
1972	اقتراح	1412	0.05	داهم	6	0.07
1973	جنرال	1411	0.05	دراجة	6	0.07

1974	أصغر	1405	0.05	نُزج	6	0.07
1975	عاشر	1405	0.05	دفاع	6	0.07
1976	جنس	1405	0.05	دلال	6	0.07
1977	تسبب	1405	0.05	دماغ	6	0.07
1978	مدخل	1404	0.05	دوام	6	0.07
1979	رقابة	1404	0.05	دينار	6	0.07
1980	شرعية	1403	0.05	ديني	6	0.07
1981	ختام	1402	0.05	ديوان	6	0.07
1982	كهربائي	1402	0.05	ذعر	6	0.07
1983	عذاب	1402	0.05	ذقن	6	0.07
1984	وَقَوْ	1401	0.05	ذهول	6	0.07
1985	حارة	1400	0.05	رائد	6	0.07
1986	فتى	1397	0.05	رحب	6	0.07
1987	لاجئ	1397	0.05	رحمة	6	0.07
1988	تفسير	1396	0.05	رزق	6	0.07
1989	ذهب	1395	0.05	رضا	6	0.07
1990	كَبْر	1395	0.05	ركن	6	0.07
1991	أخرج	1394	0.05	رن	6	0.07
1992	اجتمع	1393	0.05	روائي	6	0.07
1993	انتشر	1392	0.05	زبالة	6	0.07
1994	تولى	1392	0.05	زَحْف	6	0.07
1995	ثنائي	1391	0.05	سابع	6	0.07
1996	حَدَرَ	1390	0.05	سائح	6	0.07
1997	ناتج	1390	0.05	سائل	6	0.07
1998	فاضل	1387	0.05	ستون	6	0.07
1999	ولادة	1385	0.05	سِجِلّ	6	0.07
2000	عسل	1383	0.05	سخر	6	0.07

APÉNDICE H

Lemas exclusivos al corpus Familiar

Lema	Ranquin en Familiar
راح	217
ملمح	519
صاح	702
راهب	741
ركبة	793
كوب	813
التقط	825
أسعد	828
نظارة	875
استدار	882
انتاب	883
تَلَّ	890
فخذ	913
أحرق	939
أيقظ	945
عائش	981
عَيْش	987
نيرة	1002
انحنى	1010
أفنع	1019
حذق	1035
خد	1038
نَحْنُ	1040
شاحب	1052

شهوة	1059
نبيذ	1094
استسلم	1102
أغمض	1108
أوشك	1113
برهان	1116
خفي	1135
خَلَع	1136
دير	1137
ساق	1139
مصباح	1175
انتزع	1190
إبرة	1193
تحسس	1212
تمتم	1215
تَمَدَّد	1216
ثَقَلَ	1219
سجاد	1253
شرشف	1260
صوفي	1268
ضيعة	1274
عش	1283
قاسي	1292
قديس	1293
أَغْضَبَ	1327
أَنْجَبَ	1333
برود	1338
تَفَحَّصَ	1343
جلدي	1348
حفيد	1357
داكن	1363
رَتَّبَ	1370

رف	1372
رمادي	1375
عمى	1404
غمر	1406
قطيع	1416
كاهن	1419
لحية	1421
ليلي	1423
ميل	1452
نحيل	1454
هتف	1456
ارتبك	1467
ارتد	1468
اطمان	1471
افتقد	1472
أجبر	1477
أشدّ	1483
أفاق	1486
أنيق	1492
جرجير	1511
خضرة	1526
خطيئة	1527
خمن	1529
داعب	1530
رث	1539
زَحَفَ	1544
زقاق	1545
سارّ	1546
شال	1550
شَمّ	1552
صابون	1555
صحا	1557

عاود	1568
فضح	1578
قنديل	1582
كوخ	1587
لعبة	1588
لمع	1589
مدفون	1597
مرح	1599
ممتلئ	1609
هوى	1620
واضع	1624
ابتلع	1628
اجتاز	1630
استبدل	1636
استشعر	1637
انتصب	1645
انشغال	1646
أجرة	1647
أخذ	1648
أخلى	1649
أريكة	1651
أشعل	1652
أنصت	1661
أوصل	1663
بدائي	1665
بدن	1666
بيرا	1670
تَحَكَّمَ	1673
تسلل	1676
تعمد	1680
جذع	1688
جلباب	1690

حطب	1698
خافت	1709
دابة	1716
ذفن	1717
ذبح	1720
رخيم	1725
رقص	1728
زهور	1731
سراي	1734
سبح	1737
شارب	1740
صباحي	1745
صفع	1748
صمت	1749
ضاحكة	1751
طاف	1756
عيال	1774
قرش	1781
كمن	1796
كنب	1797
كومة	1798
مجيء	1806
محفوظ	1809
مرر	1813
مُرهِق	1815
مستسلم	1818
مسموع	1821
منديل	1827
هرع	1839
احتجاب	1854
اختبأ	1856
اختلط	1858

ارتجف	1859
استحم	1860
استرسل	1861
استولى	1864
التصق	1868
اندفاع	1871
اندفع	1872
انصرف	1874
انكسر	1877
أجلس	1880
أخاف	1882
أدهش	1883
أرملة	1885
أزعج	1886
أشرف	1888
أضفى	1891
أفرغ	1895
أفلت	1897
أقدم	1898
ألح	1899
أمسية	1902
أنوثة	1906
أهيف	1907
بدلة	1913
بؤس	1920
تسلق	1923
تعثر	1926
ثري	1933
ثكنة	1934
جَلّ	1939
حاصر	1941
حَصَنَ	1950

حماسة	1953
حماقة	1954
حمرة	1955
حميم	1956
خدر	1963
خروف	1964
خطا	1965
خفر	1967
داهم	1972
دراجة	1973
دُرُج	1974
دلال	1976
ديوان	1981
ذعر	1982
ذفن	1983
زبالة	1993
زَخْف	1994
سائل	1997
سخر	2000
سخيف	2001
سكينة	2003
سمرة	2004
شارد	2009
شاهق	2011
شتم	2012
شتيمة	2013
شحنة	2014
شرع	2015
شفي	2017
صافح	2021
صافي	2022
ضايق	2029

طَبَعَ	2033
طَبَل	2035
طَرَدَ	2037
طلعة	2039
طموح	2040
طوفان	2041
عمامة	2054
غليظ	2063
غنم	2064
فستان	2067
فضول	2068
قَاطَعَ	2073
قبض	2075
قبعة	2076
قبو	2077
قبيل	2078
لاحقَ	2087
مأخوذ	2092
مبعثر	2095
متجه	2098
مُخَزَّن	2103
مدام	2104
مرار	2106
مراهق	2107
مرعب	2108
مريب	2109
مسمع	2112
مضطرب	2113
معبد	2116
معجب	2117
ناول	2130
نَشِيط	2136

نفع	2137
نَقَش	2140
نُوبَة	2141
ورث	2146
وَلِي	2149

APÉNDICE I

Lemas en Familiar y su correspondiente ranquin en Buckwalter y Parkinson (2151 lemas)

Núm.	Lemas en Familiar	Ranquin en Buckwalter y Parkinson	Ranquin en Familiar
1	ال	1	1
2	و	2	2
3	ب	6	3
4	في	3	4
5	لِ	5	5
6	مِنْ	4	6
7	على	7	7
8	كَانَ	10	8
9	أَنَّ	8	9
10	إلى	9	10
11	لا	11	11
12	ما	20	12
13	ف	21	13
14	أَنَّ	13	14
15	أَلَمْ	29	15
16	قال	15	16
17	عن	14	17
18	التي	17	18
19	ذلك	35	19
20	كل	19	20
21	مع	18	21
22	الذي	25	22
23	لكن	54	23

24	هذا	16	24
25	ك	829	25
26	هو	22	26
27	بَعْدَ	34	27
28	س	127	28
29	قَدْ	36	29
30	أو	24	30
31	شيء	38	31
32	آخِرَ	37	32
33	أنا	26	33
34	عَرَفَ	48	34
35	نَفْسٌ	44	35
36	هذه	23	36
37	ثُمَّ	84	37
38	حتى	83	38
39	هي	33	39
40	يوم	30	40
41	عند	39	41
42	رأى	76	42
43	جِئْنَا	229	43
44	بَيْنَ	32	44
45	عَادَ	125	45
46	مرة	78	46
47	عَيْنٌ	131	47
48	أَحَدًا	64	48
49	بيت	108	49
50	يد	148	50
51	صوت	151	51
52	يا	27	52
53	قَبْلَ	69	53
54	أم	163	54
55	وَجْهَ	168	55

56	غَيْر	41	56
57	إِنْ	60	57
58	أَب	79	58
59	رَجُل	97	59
60	أَمْر	72	60
61	إِلَّا	53	61
62	أَيَّ	46	62
63	لَيْسَ	55	63
64	أَكْثَرُ	65	64
65	أَوَّلُ	40	65
66	بَعْضُ	49	66
67	أَرَادَ	124	67
68	وَجَدَ	119	68
69	كَثِيرٌ	57	69
70	أَمَامَ	133	70
71	مِثْلُ	90	71
72	مَكَانَ	176	72
73	بَابَ	298	73
74	حَيَاةَ	100	74
75	رَأْسَ	217	75
76	هَنَّاكَ	77	76
77	لِأَنَّ	58	77
78	لَوْ	91	78
79	عِنْدَمَا	143	79
80	هَلْ	70	80
81	نَظَرَ	429	81
82	شَعَرَ	505	82
83	وَاحِدٌ	59	83
84	ابْنَ	145	84
85	دُونَ	109	85
86	سَأَلَ	311	86
87	اسْمَ	96	87

88	صغير	235	88
89	بدا	324	89
90	سمع	156	90
91	وقت	103	91
92	أن	126	92
93	طويل	314	93
94	إذا	43	94
95	الله	12	95
96	غرفة	534	96
97	أخذ	197	97
98	استطاع	212	98
99	بدأ	169	99
100	منذ	150	100
101	أرض	118	101
102	صورة	123	102
103	لحظة	404	103
104	لن	154	104
105	فعل	412	105
106	كبير	66	106
107	أيضا	138	107
108	خرج	344	108
109	أنت	56	109
110	قلب	111	110
111	ترك	471	111
112	طريق	116	112
113	قليل	373	113
114	كان	315	114
115	كي	357	115
116	جديد	61	116
117	هنا	159	117
118	سيارة	250	118
119	وصل	179	119

120	أحب	325	120
121	أم	271	121
122	إن	31	122
123	صار	171	123
124	سنة	71	124
125	دخل	297	125
126	تحت	220	126
127	بلد	102	127
128	فوق	349	128
129	جسد	914	129
130	جاء	110	130
131	جميع	86	131
132	طلب	465	132
133	مدينة	144	133
134	عمل	47	134
135	أصبح	170	135
136	ذهب	440	136
137	نحو	281	137
138	كلام	251	138
139	كيف	68	139
140	أخ	67	140
141	بعيد	353	141
142	ثلاث	149	142
143	ساعة	183	143
144	كلمة	173	144
145	بقي	262	145
146	جلس	876	146
147	زال	215	147
148	والد	569	148
149	انتظر	671	149
150	داخل	252	150
151	نظرة	1011	151

152	حيث	75	152
153	ماء	241	153
154	وَقَفَّ	493	154
155	ظَلَّ	385	155
156	أتى	307	156
157	حاول	317	157
158	حَوْلَ	140	158
159	موت	481	159
160	نساء	501	160
161	جَدَّ	166	161
162	رغم	203	162
163	ماذا	255	163
164	أخير	205	164
165	جَعَلَ	291	165
166	طفل	174	166
167	أمكن	93	167
168	ربما	368	168
169	سَبَبَ	120	169
170	دارٌ	608	170
171	سؤال	273	171
172	جانب	181	172
173	ذَكَرَ	265	173
174	صَبَّاحَ	454	174
175	قوة	73	175
176	ليلة	526	176
177	عُمُرَ	284	177
178	خارج	350	178
179	فَتَّحَ	518	179
180	ناس	113	180
181	أبيض	576	181
182	أكبر	195	182
183	حَمَلَ	326	183

184	خَبِرَ	288	184
185	شَارَعَ	468	185
186	عَاشَ	328	186
187	مَاتَ	601	187
188	هُمَّ	165	188
189	تَمَامَ	455	189
190	ثَانِي	82	190
191	زَوْجَ	461	191
192	فَقَطَ	234	192
193	لَوْنَ	510	193
194	وَأَدَ	360	194
195	بَلَ	141	195
196	عَائِلَةً	870	196
197	لَيْلَ	395	197
198	أَبَدَ	549	198
199	بَلَا	598	199
200	ذَاتَ	200	200
201	صَمَّتْ	812	201
202	فَهَمَّ	599	202
203	قَامَ	105	203
204	امْرَأَةً	1019	204
205	حَالَ	157	205
206	حَقَّ	88	206
207	سَوَى	470	207
208	شَيْخَ	182	208
209	صَدِيقَ	398	209
210	غَرِيبَ	535	210
211	فَكَّرَ	641	211
212	لَدَى	153	212
213	نَوْمَ	947	213
214	انْتَهَى	532	214
215	أَمَّا	136	215

216	خَلَفَ	694	216
217	راح		217
218	وحيد	606	218
219	ورق	568	219
220	الذين	129	220
221	شهر	158	221
222	عالم	99	222
223	خُلِمَ	638	223
224	ضَحِكَ	1017	224
225	طرف	371	225
226	أجاب	1065	226
227	أحس	1400	227
228	بِما	238	228
229	فجأة	1166	229
230	مَرَّ	515	230
231	حُبَّ	211	231
232	زوجة	500	232
233	كاد	4718	233
234	كتب	355	234
235	أهل	128	235
236	حَيَّ	866	236
237	دائم	285	237
238	قديم	509	238
239	كتاب	194	239
240	كما	52	240
241	وَحَدَّ	597	241
242	عادة	773	242
243	لماذا	329	243
244	مَرَأة	320	244
245	مَنْ	89	245
246	ابتسم	1600	246
247	شِعْر	572	247

248	أَفَدَّ	180	248
249	وَضَعَّ	335	249
250	قَرَّرَ	780	250
251	إِذْ	386	251
252	أَعَادَ	1116	252
253	حَالَةً	206	253
254	حَدِيثٌ	361	254
255	رَدُّ	400	255
256	رَائِحَةٌ	1464	256
257	شَكْلٌ	130	257
258	شَمْسٌ	708	258
259	ظَنَّ	1038	259
260	هَكَذَا	527	260
261	بَيْنَمَا	539	261
262	تَحَدَّثَ	430	262
263	طَيِّبٌ	649	263
264	عَبْدٌ	681	264
265	وَرَاءَ	533	265
266	وَضَعَّ	162	266
267	ثَوْبٌ	1454	267
268	خَاصٌ	135	268
269	عِلَاقَةٌ	225	269
270	فِعْلٌ	254	270
271	أَشْبَهَ	1524	271
272	أَمَلٌ	657	272
273	بِنْتٌ	257	273
274	جَمِيلٌ	308	274
275	حَدَّ	196	275
276	رَفَعَ	768	276
277	صَدْرٌ	803	277
278	سَيِّدٌ	185	278
279	قَتَلَ	623	279

280	نحن	101	280
281	أدرك	1105	281
282	عَشْر	178	282
283	كلما	931	283
284	حقيقة	316	284
285	غادر	1652	285
286	قرأ	743	286
287	مكتب	566	287
288	نسي	741	288
289	نصف	607	289
290	بَحَثْ	692	290
291	حركة	193	291
292	رغبة	960	292
293	مِمَّا	264	293
294	ظَهَرَ	1451	294
295	أَيْنَ	411	295
296	صرخ	1760	296
297	صلاة	650	297
298	عَلِمَ	375	298
299	كَتَفَ	2997	299
300	قُرْبَ	978	300
301	نافذة	1363	301
302	مدرسة	303	302
303	أطلق	786	303
304	أعلى	394	304
305	بحر	514	305
306	تذكر	1256	306
307	جبل	872	307
308	دم	363	308
309	رواية	1108	309
310	شقة	2329	310
311	عربي	42	311

312	مال	390	312
313	وسط	480	313
314	أعطى	345	314
315	ألم	889	315
316	حَدَّث	367	316
317	روح	392	317
318	سرعة	656	318
319	فتاة	748	319
320	قريب	479	320
321	قهوة	1814	321
322	مضى	871	322
323	منزل	547	323
324	نزل	689	324
325	أحمر	967	325
326	حبيب	342	326
327	دَفَع	665	327
328	ملبس	1697	328
329	نام	1261	329
330	وَقَعَ	463	330
331	اتجاه	814	331
332	أسد	3121	332
333	بد	451	333
334	جدار	1084	334
335	طريقة	401	335
336	عَم	771	336
337	قادم	444	337
338	نهر	1205	338
339	وَجَّه	945	339
340	أستاذ	561	340
341	تَأَمَّل	3055	341
342	حكاية	1317	342
343	خطوة	586	343

344	شجرة	1010	344
345	شخص	247	345
346	شعور	1130	346
347	عالي	513	347
348	فكر	531	348
349	كامل	383	349
350	نار	469	350
351	نعم	199	351
352	نهاية	378	352
353	التفت	2198	353
354	جهة	310	354
355	حزن	774	355
356	حقيبة	2954	356
357	خروج	836	357
358	رسالة	431	358
359	سجن	746	359
360	شفة	1876	360
361	صاحب	188	361
362	قيل	1128	362
363	ابتنسامة	1925	363
364	تاريخ	289	364
365	توقف	937	365
366	حل	277	366
367	خير	142	367
368	درى	415	368
369	ذراع	3251	369
370	رأي	231	370
371	زيارة	485	371
372	سرير	1893	372
373	عام	63	373
374	عودة	489	374
375	شباب	592	375

376	فيما	278	376
377	قَدَّرَ	775	377
378	نَظَرَ	263	378
379	أَجَلَ	287	379
380	بقية	962	380
381	جرى	336	381
382	حرب	161	382
383	حكى	184	383
384	دين	333	384
385	رؤية	855	385
386	سارَ	844	386
387	سر	832	387
388	طوال	1351	388
389	عدد	117	389
390	قَدَّمَ	246	390
391	معرفة	618	391
392	نهار	1001	392
393	هواء	1275	393
394	واقع	377	394
395	وجب	230	395
396	أخرج	1991	396
397	خلال	62	397
398	زمن	560	398
399	زواج	778	399
400	عدم	187	400
401	عَقَلَ	661	401
402	عِلْمٌ	558	402
403	عَمِلَ	139	403
404	قَدَّمَ	642	404
405	قرية	676	405
406	قضى	1183	406
407	مجرد	788	407

408	منطقة	85	408
409	موضوع	94	409
410	ارتدى	2734	410
411	انتبه	2523	411
412	أخبر	2510	412
413	أسبوع	389	413
414	أصاب	794	414
415	أفضل	387	415
416	أقل	457	416
417	إنسان	202	417
418	بارد	1426	418
419	تحوّل	1008	419
420	ثقل	1648	420
421	جيد	488	421
422	حبّ	1620	422
423	سبق	990	423
424	ظهر	621	424
425	قادر	867	425
426	كفى	987	426
427	كم	340	427
428	ماضي	198	428
429	مستشفى	718	429
430	هدوء	1283	430
431	هزّ	2101	431
432	هؤلاء	423	432
433	احتاج	540	433
434	التقى	1077	434
435	تقدّم	1049	435
436	حاجة	632	436
437	خوف	899	437
438	دعا	384	438
439	دولة	51	439

440	سماء	752	440
441	شاي	2053	441
442	ضوء	881	442
443	عام	92	443
444	عنى	137	444
445	قول	312	445
446	ل	1111	446
447	أخت	207	447
448	أسود	702	448
449	أصل	556	449
450	دقيقة	427	450
451	سوف	408	451
452	شك	875	452
453	صدق	996	453
454	صندوق	1020	454
455	قصة	417	455
456	لغة	447	456
457	محل	721	457
458	ملك	765	458
459	نادر	1926	459
460	نوع	304	460
461	واضح	483	461
462	واقف	2013	462
463	اقترب	1348	463
464	أذن	1751	464
465	أربع	352	465
466	أشار	280	466
467	تحرك	1443	467
468	جيش	327	468
469	حضر	856	469
470	خط	433	470
471	داخلي	519	471

472	دخول	580	472
473	رَدَّ	428	473
474	ردد	1934	474
475	زجاج	3001	475
476	سِنَّ	1092	476
477	عَامِلٌ	630	477
478	عَبَّرَ	372	478
479	عَلَّقَ	2194	479
480	فترة	267	480
481	قاد	1239	481
482	لاحظ	1389	482
483	مَنَعَ	1394	483
484	إحساس	1592	484
485	أرسل	1588	485
486	أَهَمَّ	450	486
487	ثالث	365	487
488	دنيا	497	488
489	رَفَضَ	680	489
490	سلام	172	490
491	سمح	830	491
492	سياسي	107	492
493	طاولة	1899	493
494	غَضِبَ	1132	494
495	فَرَحَ	1223	495
496	فرصة	409	496
497	قراءة	955	497
498	كَذَلِكَ	306	498
499	كلب	1255	499
500	لقاء	379	500
501	مهم	258	501
502	اثن	663	502
503	أقام	815	503

504	إِيَّا	611	504
505	تَمْنَى	446	505
506	تَفْصِيل	1141	506
507	جَمَال	1209	507
508	دَرَجَة	472	508
509	سُوق	295	509
510	شَخْصِيَّة	730	510
511	صَنْعَب	600	511
512	طَعَام	1013	512
513	غَطَى	2252	513
514	كِتَابَة	906	514
515	مُعَلِّم	932	515
516	مَفْتُوح	1073	516
517	مَقْهَى	2157	517
518	مَلَأ	4511	518
519	مَلْمَح		519
520	مَنْحَ	1918	520
521	وَجُود	266	521
522	اِهْتِمَام	647	522
523	أَغْنِيَة	849	523
524	بَحِيرَة	3554	524
525	بَنُو	1180	525
526	جَالِس	2316	526
527	جَمَاعَة	614	527
528	رِحْلَة	1042	528
529	زَار	1637	529
530	زَاوِيَة	1638	530
531	زَمِيل	1477	531
532	سَاعِد	1078	532
533	سَاكِن	769	533
534	شَرِبَ	2089	534
535	صَعَد	2462	535

536	صف	934	536
537	طُبْع	337	537
538	عبارة	924	538
539	عجوز	1775	539
540	عمق	1202	540
541	غياب	1176	541
542	قائل	683	542
543	قرار	209	543
544	قطار	2365	544
545	مساء	655	545
546	مشهد	1074	546
547	مشى	1076	547
548	مَقْبِر	2215	548
549	مقعد	1448	549
550	ميت	1465	550
551	وصول	585	551
552	ارتفع	1040	552
553	أخفى	2719	553
554	أوقف	3260	554
555	جندي	579	555
556	جوار	1685	556
557	حارة	1985	557
558	حجر	1399	558
559	خَمْس	330	559
560	خيمة	3292	560
561	زاد	512	561
562	سعيد	1354	562
563	سمى	925	563
564	شاهد	1771	564
565	صامت	2405	565
566	عزيز	643	566
567	غاب	1585	567

568	فكرة	475	568
569	قَبِلَ	584	569
570	قَصُرَ	3208	570
571	لسان	1041	571
572	معنى	445	572
573	ها	189	573
574	ابتعد	2338	574
575	أ	348	575
576	أظهر	1054	576
577	أقرب	1411	577
578	أَكْمَلَ	1795	578
579	ألقى	1006	579
580	تَبِعَ	1788	580
581	جَدَّ	1228	581
582	جسم	777	582
583	حاج	1191	583
584	حديقة	1650	584
585	حَسَنَ	1700	585
586	حَصَلَ	437	586
587	حَمَّام	2336	587
588	رب	192	588
589	سابق	272	589
590	ساحة	847	590
591	سفر	1052	591
592	سقط	1027	592
593	سَلَّمَ	294	593
594	شاء	122	594
595	شب	269	595
596	شديد	640	596
597	عملية	132	597
598	عميق	1206	598
599	قدرة	591	599

600	كامرا	2742	600
601	مَدَّ	1667	601
602	مفتاح	1964	602
603	اعتبر	338	603
604	اعتقد	416	604
605	انتقل	1505	605
606	اهتم	1704	606
607	إصبع	2624	607
608	أغلق	2525	608
609	أنهى	2399	609
610	برد	2160	610
611	بكى	1201	611
612	تَابِعَ	922	612
613	تَعَلَّمَ	1220	613
614	تم	74	614
615	جو	687	615
616	حائط	2824	616
617	خال	2842	617
618	خَطَرَ	923	618
619	خفيف	1504	619
620	ذا	2082	620
621	رجع	616	621
622	سَحَبَ	1745	622
623	سَكَنَ	1336	623
624	سلاح	507	624
625	شَقَّ	3934	625
626	ضَحَّكَ	2392	626
627	ضَرَبَ	1120	627
628	عادي	645	628
629	عربة	2818	629
630	عشرون	625	630
631	فرنسي	563	631

632	فم	1432	632
633	قصير	1103	633
634	محطة	959	634
635	مختلف	259	635
636	مدة	594	636
637	واسع	860	637
638	اختار	1257	638
639	اختلفى	2056	639
640	اشترى	1282	640
641	أثار	1178	641
642	ألف	112	642
643	أمريكي	81	643
644	بداية	418	644
645	بكاء	1810	645
646	جار	1615	646
647	جاز	1304	647
648	جلسة	835	648
649	جنة	983	649
650	حرية	524	650
651	حقيقي	538	651
652	دام	997	652
653	دور	167	653
654	رَسَم	1520	654
655	رئيس	45	655
656	سطح	1783	656
657	طول	753	657
658	ظلام	1545	658
659	عضو	239	659
660	غَيَّرَ	1724	660
661	قاتل	1658	661
662	قَصَدَ	1331	662
663	قَطَعَ	1064	663

664	كريم	562	664
665	كنيسة	1958	665
666	لذلك	459	666
667	مشروع	147	667
668	مطبخ	2778	668
669	نسبة	152	669
670	هبط	3158	670
671	اتفق	1373	671
672	استعاد	3267	672
673	استيقظ	3074	673
674	انتظار	1155	674
675	انطلق	1135	675
676	أخذ	1748	676
677	آخر	474	677
678	أمس	218	678
679	إنما	478	679
680	أي	1059	680
681	بلغ	407	681
682	بناء	393	682
683	تزوج	1684	683
684	تَصَوَّرَ	2429	684
685	تَعَرَّفَ	1627	685
686	تَغَيَّرَ	982	686
687	تكلم	949	687
688	ثقة	936	688
689	جزء	517	689
690	حَدَّثَ	590	690
691	حذاء	3458	691
692	حساب	783	692
693	حضور	664	693
694	خيط	2756	694
695	دكتور	244	695

696	راقب	3470	696
697	رَسَمَ	1357	697
698	رغب	1825	698
699	زعيم	1080	699
700	سريع	894	700
701	سقف	2442	701
702	صاح		702
703	طابق	3856	703
704	عَقَدَ	1063	704
705	فجر	1071	705
706	فراش	2640	706
707	فقد	1240	707
708	قضية	201	708
709	كاتب	727	709
710	مرور	477	710
711	مواجهة	552	711
712	موعد	1028	712
713	مئة	249	713
714	نادى	1928	714
715	نظام	186	715
716	نَفَسَ	2173	716
717	نهض	1963	717
718	هادئ	1672	718
719	واصل	1306	719
720	وَعَدَ	2770	720
721	ودع	1776	721
722	اعتاد	3624	722
723	اكتشف	1712	723
724	انتباه	2511	724
725	انفجر	2363	725
726	أثناء	675	726
727	أزرق	1796	727

728	إشارة	1050	728
729	أضاف	253	729
730	أكد	164	730
731	بُعد	1337	731
732	بقاء	1547	732
733	تركي	1482	733
734	تلقى	1501	734
735	ثمن	1112	735
736	جُرِح	981	736
737	جلد	2143	737
738	جلوس	3082	738
739	جَمَعَ	1100	739
740	خالي	2076	740
741	راهب		741
742	رصاص	1932	742
743	رقم	347	743
744	ست	603	744
745	سلطة	313	745
746	سيدة	917	746
747	صبي	2499	747
748	صحة	462	748
749	صراخ	3321	749
750	صيف	1174	750
751	عَدَّ	798	751
752	عَرَضَ	1302	752
753	فَنَح	869	753
754	فور	1258	754
755	قتيل	1508	755
756	قَطَعَ	1549	756
757	قِطْعَة	2003	757
758	لعل	927	758
759	لهجة	2192	759

760	مسألة	760	760
761	مشعر	1045	761
762	مطر	1488	762
763	مطعم	1911	763
764	منتصف	1623	764
765	هَرَبَ	2272	765
766	واجب	1418	766
767	وَصَفَ	1146	767
768	وَقَفَ	885	768
769	يمين	1381	769
770	يومي	636	770
771	إِذْنَ	646	771
772	استمر	690	772
773	استمع	2171	773
774	أسفل	2292	774
775	أَكَلَ	1342	775
776	بَحَثَ	332	776
777	تجاه	1241	777
778	تقريب	935	778
779	تمكن	926	779
780	تَنَاولَ	1175	780
781	جسر	1494	781
782	جيب	2576	782
783	حارس	1320	783
784	حجرة	2775	784
785	حصان	3310	785
786	حَمَى	2618	786
787	خشبي	3768	787
788	خشي	1654	788
789	خيال	1943	789
790	دهشة	2710	790
791	ذاك	985	791

792	رحل	2142	792
793	ركبة		793
794	سعادة	707	794
795	سيجارة	3131	795
796	ثتاء	1799	796
797	صحيح	364	797
798	ضخم	1415	798
799	طالب	276	799
800	ظَلّ	828	800
801	عاري	3823	801
802	عدة	504	802
803	عَرَض	435	803
804	غد	744	804
805	غطاء	2743	805
806	فرد	627	806
807	فيلم	565	807
808	قائد	473	808
809	قَتَلَ	816	809
810	قَدَرَ	698	810
811	قَوِيّ	443	811
812	كرسي	1784	812
813	كوب		813
814	لحم	1578	814
815	لَعِبَ	634	815
816	مبنى	1469	816
817	متعب	3736	817
818	مريض	859	818
819	مزيد	644	819
820	مسافة	1596	820
821	مصري	370	821
822	مَلِك	419	822
823	نُور	673	823

824	ارتاح	4364	824
825	النقط		825
826	انسحب	2785	826
827	أسرة	658	827
828	أسعد		828
829	إيقاع	4204	829
830	إيمان	1271	830
831	بات	1026	831
832	بنى	1970	832
833	تَأَكَّدَ	4049	833
834	تراب	1817	834
835	تساءل	1723	835
836	جامعة	208	836
837	جثة	1450	837
838	جَمَعَ	1453	838
839	حجم	732	839
840	حُسِّنَ	1746	840
841	خجل	3049	841
842	دائرة	822	842
843	دعوة	588	843
844	دَفَعَ	1125	844
845	دمع	893	845
846	دوري	939	846
847	ذكري	958	847
848	رفيق	2018	848
849	سافر	1725	849
850	سهولة	1909	850
851	شَغَلَ	2267	851
852	صَنَعَ	1722	852
853	ضد	309	853
854	ضيق	2975	854
855	طَيَّبَ	146	855

856	ظرف	678	856
857	عابر	2707	857
858	عنف	865	858
859	عنق	3153	859
860	عهد	852	860
861	غبار	2649	861
862	فرقة	905	862
863	قتال	2031	863
864	كشفت	781	864
865	لعنة	2868	865
866	أما	323	866
867	متى	747	867
868	مجموعة	292	868
869	مجنون	2315	869
870	محاولة	633	870
871	محيط	1398	871
872	مدخل	1978	872
873	ناحية	834	873
874	نائم	2394	874
875	نظارة		875
876	هدية	1499	876
877	وادي	1423	877
878	اتصل	738	878
879	احتمال	1471	879
880	ازداد	1334	880
881	استخدم	988	881
882	استدار		882
883	انتاب		883
884	أجمل	1198	884
885	أحاط	2636	885
886	أصر	2467	886
887	أمسك	2223	887

888	باع	1695	888
889	بَشَّرَ	1088	889
890	تَلَّ		890
891	تَوَقَّفَ	2596	891
892	جدل	2139	892
893	جنس	1976	893
894	حَدَّثَ	3078	894
895	حقل	1735	895
896	خاف	1051	896
897	خشب	4533	897
898	خطأ	819	898
899	دَرَجَ	4375	899
900	دَرَسَ	1315	900
901	ذاكرة	1260	901
902	ركض	3682	902
903	سَهَّلَ	1197	903
904	شارك	674	904
905	صخر	2133	905
906	صفحة	915	906
907	ضَنَعَطَ	2326	907
908	طال	1330	908
909	عاشق	1515	909
910	عائد	2086	910
911	فاجأ	3197	911
912	فتحة	4429	912
913	فخذ		913
914	فَصَّلَ	1935	914
915	فضاء	1427	915
916	فندق	1384	916
917	فَلِقَ	3661	917
918	كَلَّمَ	2152	918
919	كيس	2605	919

920	أَيْسَ	2606	920
921	مَرْكَز	223	921
922	مُعَلَّق	3427	922
923	مُقَدَّم	1778	923
924	ممكن	396	924
925	مهمة	791	925
926	ناعم	2552	926
927	نَص	864	927
928	نقطة	438	928
929	نَقَلَ	883	929
930	نَقَلَ	525	930
931	هَمَّ	1511	931
932	هما	1147	932
933	وعى	1457	933
934	ادعى	2108	934
935	امتد	1621	935
936	انبعى	1248	936
937	أثر	550	937
938	إجابة	1671	938
939	أحرق		939
940	أحسن	792	940
941	أحضر	4755	941
942	أداة	1720	942
943	أعد	2855	943
944	أقبل	3184	944
945	أيقظ		945
946	بال	800	946
947	بطء	3275	947
948	بيع	1101	948
949	تأخَّرَ	2356	949
950	تبادل	1557	950
951	تجاوزَ	1085	951

952	تَرَدَّدَ	1730	952
953	تَفْكَير	1097	953
954	تَوَقَّعَ	840	954
955	تَلَجَّ	3489	955
956	ثَوْرَة	976	956
957	جَرَوْ	4013	957
958	حُرَّ	837	958
959	حَسَبَ	422	959
960	حَلَّمَ	1841	960
961	خَطَاب	1153	961
962	خِلاص	3484	962
963	خَلْفِي	4203	963
964	خَيْل	2443	964
965	دَعَاء	1458	965
966	دَلِيل	944	966
967	ذَهْن	1851	967
968	رَاضِي	3212	968
969	رَجَا	890	969
970	رَعَب	2601	970
971	سَاد	2204	971
972	سَفِينَة	1755	972
973	سَوَاد	3028	973
974	شَاطِئ	2185	974
975	شِدَة	1412	975
976	صَحِيفَة	452	976
977	ضَحْكَة	3738	977
978	ضَرُورَة	528	978
979	ضَنَعَط	742	979
980	ظَهْر	1492	980
981	عَانِش		981
982	عَرَس	2100	982
983	عَرَق	3142	983

984	عَشَاء	2371	984
985	علا	1906	985
986	عنوان	631	986
987	عَيْش		987
988	غامض	2654	988
989	غرض	1562	989
990	فراغ	1931	990
991	قبيلة	1490	991
992	قَلْب	4237	992
993	قميص	4620	993
994	كَف	2410	994
995	كأنما	2434	995
996	لبناني	191	996
997	مزاج	3796	997
998	مقابل	670	998
999	موجود	449	999
1000	مؤخر	1459	1000
1001	موقف	290	1001
1002	نيرة		1002
1003	نَقْد	1687	1003
1004	هدأ	3037	1004
1005	اتجه	2033	1005
1006	احتفظ	3007	1006
1007	استحق	1428	1007
1008	امتأ	4734	1008
1009	انتهاء	1233	1009
1010	انحنى		1010
1011	أتم	3482	1011
1012	أثبت	1472	1012
1013	أخضر	1016	1013
1014	أدخل	1702	1014
1015	أسرع	3277	1015

1016	أَصْفَر	2183	1016
1017	إِطْلَاق	622	1017
1018	أَعْلَن	391	1018
1019	أَفْنَع		1019
1020	بَائِع	3045	1020
1021	بَسِيط	884	1021
1022	بِنْدَقِيَّة	4273	1022
1023	تَحْرُك	1834	1023
1024	تَرَاجَع	1861	1024
1025	تَطَّلَع	2562	1025
1026	تَعْلِيم	381	1026
1027	ثَمَّة	1286	1027
1028	جَامِع	3728	1028
1029	جَمَلَة	1498	1029
1030	جَهْد	668	1030
1031	جِيل	1118	1031
1032	حَاد	1626	1032
1033	حَادِث	1425	1033
1034	حَاكِم	848	1034
1035	حَدَق		1035
1036	حَظ	1532	1036
1037	خَائِف	1563	1037
1038	خَد		1038
1039	دَخَان	2382	1039
1040	دَخَّنَ		1040
1041	ذَهَاب	1591	1041
1042	رَابِع	751	1042
1043	رَمَى	2589	1043
1044	رِيح	1208	1044
1045	زَحَام	4348	1045
1046	زِينَة	4045	1046
1047	سَبْع	705	1047

1048	سكين	4662	1048
1049	سلطان	956	1049
1050	سياسة	351	1050
1051	سيرة	2199	1051
1052	شاحب		1052
1053	شأن	243	1053
1054	شبكة	779	1054
1055	شر	1210	1055
1056	شَرَحَ	2767	1056
1057	شعاع	2671	1057
1058	شهادة	975	1058
1059	شهوة		1059
1060	صراحة	595	1060
1061	صليب	3807	1061
1062	صوب	2547	1062
1063	ضابط	994	1063
1064	طَلَّبَ	589	1064
1065	طَقَسَ	2546	1065
1066	عثر	2673	1066
1067	عَكَسَ	1024	1067
1068	عَلَامَة	1618	1068
1069	غداء	2900	1069
1070	غسل	2872	1070
1071	غناء	2145	1071
1072	فارسي	3643	1072
1073	فارغ	2531	1073
1074	فرحة	2212	1074
1075	فريق	242	1075
1076	فنجان	3979	1076
1077	قَابَل	4356	1077
1078	قاعدة	571	1078
1079	قائمة	4243	1079

1080	قرن	831	1080
1081	قسم	699	1081
1082	قَفَزَ	3032	1082
1083	قيمة	426	1083
1084	كابوس	4738	1084
1085	كأس	826	1085
1086	كرة	498	1086
1087	كِلَا	2361	1087
1088	لف	3486	1088
1089	لمح	4108	1089
1090	مادة	424	1090
1091	مثير	2238	1091
1092	مرض	456	1092
1093	ممر	4423	1093
1094	نبيذ		1094
1095	نهائي	739	1095
1096	هب	2465	1096
1097	هَمَسَ	2887	1097
1098	واجه	969	1098
1099	اتسع	2945	1099
1100	اجتماع	432	1100
1101	احتضن	4427	1101
1102	استسلم		1102
1103	أبدى	2151	1103
1104	إجازة	2128	1104
1105	أدب	1215	1105
1106	أدبي	2070	1106
1107	أغلب	1857	1107
1108	أغمض		1108
1109	إقامة	762	1109
1110	أقصى	1003	1110
1111	أقوى	1738	1111

1112	آل	4283	1112
1113	أوشك		1113
1114	بَدَل	898	1114
1115	برنامج	155	1115
1116	برهان		1116
1117	تاجر	1752	1117
1118	تالي	458	1118
1119	تَجَمَّعَ	4165	1119
1120	تحية	749	1120
1121	تَدَخَّلَ	1679	1121
1122	تَدَخَّلَ	1614	1122
1123	تَصَرَّفَ	3261	1123
1124	تعبير	1021	1124
1125	تغيير	602	1125
1126	جبهة	1356	1126
1127	جريح	1924	1127
1128	جلب	2938	1128
1129	جنب	2078	1129
1130	جنيه	1445	1130
1131	حصة	1894	1131
1132	حُكْم	305	1132
1133	خاصة	334	1133
1134	خصوص	460	1134
1135	خفي		1135
1136	خَلَعَ		1136
1137	دير		1137
1138	رقبة	3772	1138
1139	ساق		1139
1140	سرقة	2789	1140
1141	شخصي	790	1141
1142	شد	2359	1142
1143	شعر	930	1143

1144	شُكْر	222	1144
1145	شَيْطَان	1551	1145
1146	صحراء	1649	1146
1147	صعوبة	1313	1147
1148	صلى	464	1148
1149	ضَبْط	929	1149
1150	ضَرْب	1880	1150
1151	ضعيف	1339	1151
1152	ضَيِّق	1866	1152
1153	ظالم	2553	1153
1154	عَبَّرَ	1152	1154
1155	عتبة	4409	1155
1156	عظيم	577	1156
1157	عنيف	2168	1157
1158	غرق	3040	1158
1159	فرج	3759	1159
1160	فريد	2759	1160
1161	فساد	1157	1161
1162	فَهْم	1393	1162
1163	قاعة	1470	1163
1164	قَدَّرَ	1249	1164
1165	قَلَّق	1329	1165
1166	قوم	1344	1166
1167	لاحق	1729	1167
1168	لذة	4307	1168
1169	لفت	1656	1169
1170	مُجْتَمِع	216	1170
1171	مجلس	98	1171
1172	مِرْأَة	2764	1172
1173	مرحلة	354	1173
1174	مسلم	228	1174
1175	مصباح		1175

1176	مصدر	358	1176
1177	مطار	1355	1177
1178	معركة	953	1178
1179	مفاجئ	2551	1179
1180	ملون	3536	1180
1181	نجح	977	1181
1182	نسخة	1311	1182
1183	وافق	1475	1183
1184	وحدة	232	1184
1185	وضوح	1821	1185
1186	وقوف	1553	1186
1187	استسلام	4323	1187
1188	استقر	2610	1188
1189	امتلك	1966	1189
1190	انتزع		1190
1191	انتشر	1993	1191
1192	انتمى	2058	1192
1193	إبرة		1193
1194	أدار	2923	1194
1195	أدى	405	1195
1196	إصرار	2763	1196
1197	أطل	2353	1197
1198	أمضى	4671	1198
1199	إمكان	1510	1199
1200	بادر	2951	1200
1201	باقي	948	1201
1202	بَشَرَ	4479	1202
1203	بطل	888	1203
1204	بني	4036	1204
1205	بوابة	2482	1205
1206	بياض	3616	1206
1207	تَأَمَّل	4444	1207

1208	تبيين	1792	1208
1209	تجربة	587	1209
1210	تحديد	677	1210
1211	تحرير	857	1211
1212	تحسس		1212
1213	تحقيق	268	1213
1214	تعلق	724	1214
1215	تمتم		1215
1216	تَمَدَّدَ		1216
1217	تنظيم	669	1217
1218	تَوَجَّهَ	2387	1218
1219	ثَقُلَ		1219
1220	جاف	3517	1220
1221	جحيم	3053	1221
1222	جر	4426	1222
1223	جنازة	4789	1223
1224	جواب	1431	1224
1225	حاضر	1113	1225
1226	حرارة	1416	1226
1227	خَرَجَ	3191	1227
1228	خَرَمَ	1616	1228
1229	حزين	1639	1229
1230	حقق	666	1230
1231	حمار	2494	1231
1232	حماية	767	1232
1233	حَمَلَ	1288	1233
1234	حوالي	818	1234
1235	خزانة	3733	1235
1236	حيرة	2565	1236
1237	دافئ	2955	1237
1238	دراسة	286	1238
1239	دراسي	1632	1239

1240	دق	1121	1240
1241	ذَكَرَ	3333	1241
1242	ذئب	2566	1242
1243	رَبَطَ	2264	1243
1244	رَجِمَ	1185	1244
1245	رضي	1193	1245
1246	رفيع	2229	1246
1247	روى	1891	1247
1248	زجاجة	4389	1248
1249	زمان	763	1249
1250	ساخن	2391	1250
1251	سارع	4050	1251
1252	سبيل	548	1252
1253	سجاد		1253
1254	سَخَب	3537	1254
1255	سَلَّمَ	3437	1255
1256	سواء	637	1256
1257	شاشة	1635	1257
1258	شبح	3908	1258
1259	شبهه	1575	1259
1260	شُرشف		1260
1261	شرفه	3433	1261
1262	شَكَلَ	720	1262
1263	شَهِد	467	1263
1264	صالون	3276	1264
1265	صحبة	3921	1265
1266	صحن	4251	1266
1267	صداقة	2332	1267
1268	صوفي		1268
1269	ضاق	2473	1269
1270	ضحية	1115	1270
1271	ضربة	1690	1271

1272	ضفة	1136	1272
1273	ضَمَّ	880	1273
1274	ضيعة		1274
1275	طائرة	913	1275
1276	طَبَّق	4063	1276
1277	طن	2406	1277
1278	ظهور	1485	1278
1279	عاصفة	2571	1279
1280	عَبَّرَ	3339	1280
1281	عجيب	1782	1281
1282	عدا	2271	1282
1283	عش		1283
1284	عَصَّرَ	2539	1284
1285	عمارة	3002	1285
1286	عَيَّنَ	3706	1286
1287	غاية	2125	1287
1288	غرابة	2588	1288
1289	غربي	543	1289
1290	فَرَّحَ	3246	1290
1291	فَضَّلَ	696	1291
1292	قاسي		1292
1293	قديس		1293
1294	قناة	846	1294
1295	قيام	593	1295
1296	لباس	3941	1296
1297	لعب	873	1297
1298	مالك	4020	1298
1299	مباشرة	1188	1299
1300	مُحامي	1819	1300
1301	مدى	567	1301
1302	مَسَّحَ	3255	1302
1303	مسرح	1124	1303

1304	مشكلة	177	1304
1305	مصلحة	476	1305
1306	معروف	745	1306
1307	مُعْظَم	854	1307
1308	مغادرة	2720	1308
1309	مقدار	1833	1309
1310	مكتوب	1940	1310
1311	مَنْع	1104	1311
1312	نبيل	3114	1312
1313	نسيان	2874	1313
1314	نشاط	605	1314
1315	هاتف	1251	1315
1316	وعى	3660	1316
1317	يسار	2448	1317
1318	استغرق	2463	1318
1319	استقبل	1682	1319
1320	اضطر	2699	1320
1321	افترض	2643	1321
1322	امتحان	1590	1322
1323	أَبْعَدَ	4100	1323
1324	أَبْعُدُ	2630	1324
1325	أساس	482	1325
1326	أسمر	4508	1326
1327	أَعْضَبَ		1327
1328	أفق	1444	1328
1329	أَكُلُ	1500	1329
1330	آلة	2060	1330
1331	إله	1280	1331
1332	أَمَّنَ	3855	1332
1333	أَنْجَبَ		1333
1334	أنف	3062	1334
1335	آه	1119	1335

1336	أيها	764	1336
1337	بالغ	1269	1337
1338	برود		1338
1339	بِضْع	2104	1339
1340	بطن	2096	1340
1341	بلدة	1314	1341
1342	تَحَمَّلَ	1952	1342
1343	تَفَحَّصَ		1343
1344	تكرر	2540	1344
1345	تلا	4106	1345
1346	ثانوي	1493	1346
1347	جرس	3190	1347
1348	جلدي		1348
1349	جمعة	973	1349
1350	حدد	1321	1350
1351	جَدِّي	2557	1351
1352	حديد	1813	1352
1353	حرف	1145	1353
1354	حرك	2470	1354
1355	جِسِّ	3503	1355
1356	حِضْن	3227	1356
1357	حفيد		1357
1358	حَلَقَ	3840	1358
1359	حيوان	1340	1359
1360	خاطب	4557	1360
1361	خبز	2246	1361
1362	خدم	1791	1362
1363	داكن		1363
1364	درب	1554	1364
1365	دِقَّة	1923	1365
1366	ذهني	4566	1366
1367	ذو	731	1367

1368	راحة	1556	1368
1369	ربيع	2025	1369
1370	رَتَّبَ		1370
1371	رصيف	3864	1371
1372	رف		1372
1373	رَفُضَ	1169	1373
1374	رَفَعَ	858	1374
1375	رمادي		1375
1376	زيت	1743	1376
1377	ساخر	3283	1377
1378	سادس	1274	1378
1379	سائق	1913	1379
1380	سخرية	2819	1380
1381	سرق	3063	1381
1382	سري	1842	1382
1383	سلامة	887	1383
1384	سماع	2944	1384
1385	سمعة	3200	1385
1386	سوري	554	1386
1387	سير	1417	1387
1388	شَكَرَ	920	1388
1389	شوق	1442	1389
1390	صالة	2665	1390
1391	صبر	1211	1391
1392	صرخة	3674	1392
1393	ضاع	2110	1393
1394	طاقة	723	1394
1395	طبيعة	942	1395
1396	طَرَحَ	1218	1396
1397	طرق	4079	1397
1398	طير	1179	1398
1399	طين	3224	1399

1400	عجلة	4164	1400
1401	عدو	755	1401
1402	عديد	414	1402
1403	عسكري	318	1403
1404	عمى		1404
1405	عناية	1901	1405
1406	غمر		1406
1407	غيبوبة	4660	1407
1408	غيمة	4317	1408
1409	فارس	2660	1409
1410	فحص	2057	1410
1411	فلاح	3180	1411
1412	قاطع	3019	1412
1413	قبالة	4313	1413
1414	قبر	1665	1414
1415	قسوة	2821	1415
1416	قطيع		1416
1417	قلم	1033	1417
1418	كفّ	3288	1418
1419	كاهن		1419
1420	كائن	2474	1420
1421	لحية		1421
1422	لمس	3666	1422
1423	ليلي		1423
1424	مباشر	919	1424
1425	متأخر	2141	1425
1426	متأكد	3633	1426
1427	متر	1308	1427
1428	متعة	2929	1428
1429	مجاور	1770	1429
1430	مجلة	1140	1430
1431	مجهول	1733	1431

1432	محدد	1151	1432
1433	مخافة	2805	1433
1434	مخيم	1165	1434
1435	مساحة	1156	1435
1436	مستعد	1946	1436
1437	مسجد	796	1437
1438	مصنع	1927	1438
1439	مصيبة	2507	1439
1440	مظهر	1822	1440
1441	معرض	1022	1441
1442	مغلق	2368	1442
1443	مفكر	2685	1443
1444	ممرض	4397	1444
1445	مناسب	839	1445
1446	مناسبة	817	1446
1447	منظر	2619	1447
1448	موسيقى	1466	1448
1449	موضع	2731	1449
1450	مَوْقِع	245	1450
1451	ميدان	1606	1451
1452	ميل		1452
1453	نجا	4610	1453
1454	نحيل		1454
1455	نصر	1144	1455
1456	هتف		1456
1457	وطن	359	1457
1458	جنون	2408	1458
1459	وَدَّ	1873	1459
1460	وزير	95	1460
1461	اتخذ	1429	1461
1462	اتصال	490	1462
1463	اجتاح	4253	1463

1464	احتل	2241	1464
1465	احتمل	3086	1465
1466	ارتباط	2948	1466
1467	ارتبك		1467
1468	ارتد		1468
1469	استقبال	1558	1469
1470	اشتغل	833	1470
1471	اطمأن		1471
1472	افتقد		1472
1473	اكتفى	2932	1473
1474	أبقى	4697	1474
1475	أتاح	2560	1475
1476	أتي	2230	1476
1477	أجبر		1477
1478	أجرى	1343	1478
1479	أداء	886	1479
1480	أربعون	1705	1480
1481	أسوأ	2389	1481
1482	أسى	2774	1482
1483	أشدّ		1483
1484	أطول	2832	1484
1485	إعادة	506	1485
1486	أفاق		1486
1487	أمرَ	2225	1487
1488	أمن	240	1488
1489	أمنَ	1370	1489
1490	أمني	492	1490
1491	أمير	573	1491
1492	أنيق		1492
1493	أهلي	1341	1493
1494	أيسر	2617	1494
1495	بركة	1971	1495

1496	بساطة	2255	1496
1497	بطاقة	1536	1497
1498	تارك	4309	1498
1499	تبقى	2191	1499
1500	تخلى	2808	1500
1501	ترتيب	1581	1501
1502	تسليم	1959	1502
1503	تَعَامَل	1912	1503
1504	تفسير	1988	1504
1505	تلفزيون	1053	1505
1506	تهمة	1611	1506
1507	تيار	1230	1507
1508	ثابت	1546	1508
1509	ثقافة	508	1509
1510	جامعي	1570	1510
1511	جرجير		1511
1512	جمهور	716	1512
1513	جهاز	331	1513
1514	جَهْل	4816	1514
1515	جَهْل	2485	1515
1516	حافظ	2690	1516
1517	حجة	1885	1517
1518	حَدْر	2291	1518
1519	حرس	2381	1519
1520	حريق	2341	1520
1521	حضارة	1328	1521
1522	حَفِظْ	1266	1522
1523	حُمَى	3566	1523
1524	حوض	2845	1524
1525	خدمة	302	1525
1526	خضرة		1526
1527	خطيئة		1527

1528	خَلَقَ	1060	1528
1529	خَمَنَ		1529
1530	دَاعَبَ		1530
1531	دَافِعَ	2304	1531
1532	دَرَسَ	1087	1532
1533	دَفَّءَ	3549	1533
1534	دَفْتَرَ	3974	1534
1535	دَفَّقَ	4561	1535
1536	دَوَّرَ	582	1536
1537	ذَنَّبَ	1942	1537
1538	رَاكَبَ	2533	1538
1539	رَثَ		1539
1540	رَحِيلَ	2023	1540
1541	رَكِبَ	3156	1541
1542	رَسُولَ	499	1542
1543	زَبُونَ	3858	1543
1544	زَحَفَ		1544
1545	زَقَّاقَ		1545
1546	سَارَ		1546
1547	سَرَعَانَ	2302	1547
1548	سَيْفَ	2303	1548
1549	سَيْمًا	1641	1549
1550	شَالَ		1550
1551	شَرَّقَ	382	1551
1552	شَمَّ		1552
1553	شَمَالَ	628	1553
1554	شَهِيدَ	685	1554
1555	صَابُونَ		1555
1556	صَبِيَّةَ	3777	1556
1557	صَحَا		1557
1558	صَدَّرَ	1012	1558
1559	صَدَفَةَ	2584	1559

1560	صِدْق	1707	1560
1561	صِغَر	3730	1561
1562	صراع	850	1562
1563	صوتي	2798	1563
1564	ضاحك	3872	1564
1565	ضَعْف	1270	1565
1566	طالِب	772	1566
1567	عانى	862	1567
1568	عاود		1568
1569	عروس	2528	1569
1570	عزاء	3357	1570
1571	عَشِيق	4101	1571
1572	عصبي	2383	1572
1573	عظم	2534	1573
1574	عيد	654	1574
1575	فات	2423	1575
1576	فاكهة	2965	1576
1577	فر	4780	1577
1578	فضح		1578
1579	فك	3455	1579
1580	قصيدة	921	1580
1581	قماش	3364	1581
1582	قنديل		1582
1583	كَبِر	1990	1583
1584	كتلة	1503	1584
1585	كثرة	1646	1585
1586	كرر	2701	1586
1587	كوخ		1587
1588	لعبة		1588
1589	لمع		1589
1590	مارس	1413	1590
1591	مأساة	2320	1591

1592	مبلغ	972	1592
1593	متابعة	1005	1593
1594	متقدم	1539	1594
1595	متوسط	1190	1595
1596	محكمة	521	1596
1597	مدفون		1597
1598	مراقبة	1610	1598
1599	مرح		1599
1600	مساعدة	635	1600
1601	مَسَاك	3472	1601
1602	مسؤول	339	1602
1603	مسيح	2936	1603
1604	مصاب	1460	1604
1605	مطلوب	824	1605
1606	معناد	3521	1606
1607	معدة	3387	1607
1608	مكتبة	1910	1608
1609	ممنلى		1609
1610	موظف	941	1610
1611	موكب	4170	1611
1612	نائب	301	1612
1613	نسوة	4249	1613
1614	نسيم	3782	1614
1615	نظيف	2503	1615
1616	نما	2166	1616
1617	نوى	3229	1617
1618	هَدَف	219	1618
1619	هن	3113	1619
1620	هوى		1620
1621	هيا	1736	1621
1622	هيئة	403	1622
1623	واثق	3003	1623

1624	واضع		1624
1625	وزارة	214	1625
1626	وفاة	1519	1626
1627	يفين	2314	1627
1628	ابتلع		1628
1629	اثنين	1346	1629
1630	اجتاز		1630
1631	اجتماعي	343	1631
1632	اجتمع	1992	1632
1633	احتفال	1518	1633
1634	احتوى	1737	1634
1635	اختلف	1137	1635
1636	استبدل		1636
1637	استشعر		1637
1638	استعمل	2881	1638
1639	استغرب	4210	1639
1640	اشند	3740	1640
1641	اعتذر	2981	1641
1642	اعترض	3102	1642
1643	اقترب	3134	1643
1644	اكتمل	3477	1644
1645	انتصب		1645
1646	انشغال		1646
1647	أجرة		1647
1648	أخذ		1648
1649	أحلى		1649
1650	إذن	1323	1650
1651	أريكة		1651
1652	أشعل		1652
1653	إصابة	736	1653
1654	أعلم	3564	1654
1655	إقناع	3348	1655

1656	ألماني	1171	1656
1657	أمامي	4138	1657
1658	أمانة	1569	1658
1659	أمة	236	1659
1660	إنساني	697	1660
1661	أنصت		1661
1662	أهمية	564	1662
1663	أوصل		1663
1664	بدء	984	1664
1665	بدائي		1665
1666	بدن		1666
1667	برج	2750	1667
1668	بشري	1062	1668
1669	بناية	4515	1669
1670	بيرا		1670
1671	تام	1653	1671
1672	تجاري	808	1672
1673	تَحَكَّم		1673
1674	تخيل	3757	1674
1675	تَرَدُّد	4227	1675
1676	تسلل		1676
1677	تَصْرُف	1675	1677
1678	تَعَامَل	861	1678
1679	تعذيب	2169	1679
1680	تعمد		1680
1681	تَقَدُّم	766	1681
1682	تلفون	891	1682
1683	تمثال	4301	1683
1684	تَنَاوُل	1630	1684
1685	توتر	1904	1685
1686	توفيق	1245	1686
1687	ثلاثون	1148	1687

1688	جذع		1688
1689	جزيرة	1764	1689
1690	جلباب		1690
1691	جولة	1203	1691
1692	حاجز	1837	1692
1693	حالي	296	1693
1694	حان	2729	1694
1695	حَدَّرَ	1996	1695
1696	حراسة	3620	1696
1697	حَطَّ	4120	1697
1698	حطب		1698
1699	حفرة	3997	1699
1700	حَكَمَ	1742	1700
1701	حكيم	1933	1701
1702	حلو	221	1702
1703	حماس	3571	1703
1704	حملة	863	1704
1705	حنان	2280	1705
1706	حَوَّلَ	2451	1706
1707	خادم	2061	1707
1708	خاض	3041	1708
1709	خافت		1709
1710	خبرة	1149	1710
1711	خسارة	1159	1711
1712	خصم	3874	1712
1713	خطة	615	1713
1714	خف	4528	1714
1715	خميس	1098	1715
1716	دابة		1716
1717	دَفَنَ		1717
1718	دوم	1608	1718
1719	ذاب	3760	1719

1720	ذبح		1720
1721	ذهبي	1673	1721
1722	رامي	3778	1722
1723	رائع	486	1723
1724	رخيص	2815	1724
1725	رخيم		1725
1726	رطوبة	4573	1726
1727	رفقة	2992	1727
1728	رقص		1728
1729	رمل	2095	1729
1730	زر	4042	1730
1731	زهور		1731
1732	زيتون	2411	1732
1733	سحر	2519	1733
1734	سراي		1734
1735	سرى	3122	1735
1736	سعى	804	1736
1737	سفع		1737
1738	سيلم	2308	1738
1739	سور	2385	1739
1740	شارب		1740
1741	شرط	805	1741
1742	شرطة	604	1742
1743	شرقي	1133	1743
1744	شريط	1688	1744
1745	صباحي		1745
1746	صدر	2409	1746
1747	صفة	1069	1747
1748	صَفَع		1748
1749	صَمَت		1749
1750	صنف	3990	1750
1751	ضاحكة		1751

1752	ضجيج	3818	1752
1753	ضِمن	596	1753
1754	ضمير	2481	1754
1755	ضيف	1525	1755
1756	طاف		1756
1757	طبي	874	1757
1758	طبيعي	574	1758
1759	عالج	2976	1759
1760	عالم	853	1760
1761	عثماني	4163	1761
1762	عَجَزَ	2621	1762
1763	عَجَزَ	2165	1763
1764	عذاب	1983	1764
1765	عريض	2103	1765
1766	عزلة	3391	1766
1767	عسكر	4450	1767
1768	عصا	3561	1768
1769	عصفور	3640	1769
1770	عطر	2146	1770
1771	علاج	770	1771
1772	علبة	2924	1772
1773	عيادة	4443	1773
1774	عيال		1774
1775	غاية	998	1775
1776	غروب	3356	1776
1777	غلب	3094	1777
1778	غَنَى	2355	1778
1779	فاتح	2480	1779
1780	فتى	1986	1780
1781	فَرَشَ		1781
1782	فريسة	4779	1782
1783	فَصَلَ	2037	1783

1784	قاضي	1126	1784
1785	قاوم	3739	1785
1786	قبضة	3960	1786
1787	قرآن	843	1787
1788	قعد	1483	1788
1789	قيد	1517	1789
1790	كافي	1327	1790
1791	كثيف	3052	1791
1792	كِرَة	2478	1792
1793	كَسْرَ	2218	1793
1794	كلف	2322	1794
1795	كلية	912	1795
1796	كَمَل		1796
1797	كنب		1797
1798	كومة		1798
1799	لبث	3555	1799
1800	لجنة	213	1800
1801	لوح	4136	1801
1802	لوحة	1237	1802
1803	مائدة	2811	1803
1804	متعدد	1318	1804
1805	مجال	260	1805
1806	مجيء		1806
1807	محبة	2457	1807
1808	محتمل	2404	1808
1809	محفوظ		1809
1810	مدير	369	1810
1811	مذكرة	2176	1811
1812	مذهب	2377	1812
1813	مَرَر		1813
1814	مرسوم	3093	1814
1815	مُرْهِق		1815

1816	مس	2893	1816
1817	مستحيل	1694	1817
1818	مستسلم		1818
1819	مستوى	275	1819
1820	مسكين	1739	1820
1821	مسموع		1821
1822	مُعلن	3071	1822
1823	مقاتل	3683	1823
1824	مقارنة	1372	1824
1825	مقاومة	502	1825
1826	مقر	1222	1826
1827	مندبل		1827
1828	مؤسسة	283	1828
1829	نادي	399	1829
1830	نال	1816	1830
1831	نيات	2730	1831
1832	نبي	825	1832
1833	نجمة	3209	1833
1834	نداء	2350	1834
1835	نَشَرَ	991	1835
1836	نَطَقَ	2715	1836
1837	نقيب	3422	1837
1838	هجرة	1664	1838
1839	هرع		1839
1840	هُمَسَ	4197	1840
1841	هنالك	1692	1841
1842	هندي	1670	1842
1843	واجهه	3480	1843
1844	وثيقة	1474	1844
1845	وسيلة	536	1845
1846	وضعية	3679	1846
1847	وطأة	3689	1847

1848	وطني	134	1848
1849	وظيفة	1195	1849
1850	ولادة	1999	1850
1851	يمني	1480	1851
1852	يهودي	652	1852
1853	اتفاق	583	1853
1854	احتجاب		1854
1855	احتلال	491	1855
1856	اختبأ		1856
1857	اختفاء	4618	1857
1858	اختلط		1858
1859	ارتجف		1859
1860	استحم		1860
1861	استرسل		1861
1862	استعداد	1127	1862
1863	استمرار	810	1863
1864	استولى		1864
1865	اعترف	1301	1865
1866	اقتحم	4125	1866
1867	التزام	1278	1867
1868	التصق		1868
1869	انتظام	4785	1869
1870	انتقام	4025	1870
1871	اندفاع		1871
1872	اندفع		1872
1873	انسحاب	1595	1873
1874	انصرف		1874
1875	انفجار	1565	1875
1876	انقطع	2251	1876
1877	انكسر		1877
1878	انهار	4260	1878
1879	إجراء	421	1879

1880	أجلس		1880
1881	أجنبي	725	1881
1882	أخاف		1882
1883	أدهش		1883
1884	إرادة	1804	1884
1885	أرملة		1885
1886	أزعج		1886
1887	أسلوب	1007	1887
1888	أشرف		1888
1889	أصدر	1345	1889
1890	إضافة	362	1890
1891	أضفى		1891
1892	أعجب	1914	1892
1893	إعلان	820	1893
1894	أفاد	1123	1894
1895	أفرغ		1895
1896	أفريقي	938	1896
1897	أفلت		1897
1898	أقدم		1898
1899	ألح		1899
1900	ألغى	4111	1900
1901	إمام	1055	1901
1902	أمسية		1902
1903	أنتم	1089	1903
1904	إنجليزي	1358	1904
1905	أنزل	3951	1905
1906	أنوثة		1906
1907	أهيف		1907
1908	أوحى	3140	1908
1909	أوصى	4670	1909
1910	أولئك	1868	1910
1911	باحث	1238	1911

1912	باشا	2735	1912
1913	بدلة		1913
1914	برهة	3659	1914
1915	بري	2575	1915
1916	بريء	1655	1916
1917	بريق	4565	1917
1918	بقعة	2928	1918
1919	بنك	612	1919
1920	بوس		1920
1921	بيض	2573	1921
1922	تاسع	1696	1922
1923	تسلق		1923
1924	تسمية	2697	1924
1925	تعالى	420	1925
1926	تعثر		1926
1927	تعليق	1347	1927
1928	تقرير	494	1928
1929	تمرد	3178	1929
1930	تنظيف	3223	1930
1931	تنفس	3653	1931
1932	ثامن	1633	1932
1933	ثري		1933
1934	ثكنة		1934
1935	جدوى	2269	1935
1936	جذب	3380	1936
1937	جريدة	1383	1937
1938	جريمة	750	1938
1939	جَلّ		1939
1940	جوهر	3230	1940
1941	حاصر		1941
1942	حافة	3352	1942
1943	حتم	2903	1943

1944	حجاب	2296	1944
1945	جِدَّة	2432	1945
1946	حَرَ	3238	1946
1947	حزب	274	1947
1948	حشد	3220	1948
1949	حصول	785	1949
1950	حَصَنَ		1950
1951	حليب	2214	1951
1952	حليف	3204	1952
1953	حماسة		1953
1954	حماقة		1954
1955	حمرة		1955
1956	حميم		1956
1957	حنون	3877	1957
1958	حوار	436	1958
1959	خاتم	4723	1959
1960	خارجي	776	1960
1961	خاطر	2020	1961
1962	خامس	1046	1962
1963	خدر		1963
1964	خروف		1964
1965	خطا		1965
1966	خطف	3270	1966
1967	خفر		1967
1968	خلا	2156	1968
1969	خَلَقَ	1787	1969
1970	خُلِقَ	1862	1970
1971	دافع	2611	1971
1972	داهم		1972
1973	دراجة		1973
1974	دُرَج		1974
1975	دفاع	542	1975

1976	دلال		1976
1977	دماغ	3514	1977
1978	دوام	1807	1978
1979	دينار	966	1979
1980	ديني	700	1980
1981	ديوان		1981
1982	ذعر		1982
1983	ذقن		1983
1984	ذهول	3962	1984
1985	رائد	1728	1985
1986	رحب	2035	1986
1987	رحمة	993	1987
1988	رزق	3182	1988
1989	رضا	2614	1989
1990	ركن	1461	1990
1991	رن	3771	1991
1992	روائي	3449	1992
1993	زبالة		1993
1994	زَحْف		1994
1995	سابع	1379	1995
1996	سائح	3259	1996
1997	سائل		1997
1998	ستون	2337	1998
1999	سِجَل	3066	1999
2000	سخر		2000
2001	سخيف		2001
2002	سكت	2941	2002
2003	سكينة		2003
2004	سمرة		2004
2005	سَمْع	3183	2005
2006	سهر	4350	2006
2007	سهرة	3080	2007

2008	سيء	1580	2008
2009	شارد		2009
2010	شاهد	1636	2010
2011	شاهق		2011
2012	شتم		2012
2013	شنيمة		2013
2014	شحنة		2014
2015	شرع		2015
2016	شعب	121	2016
2017	شفي		2017
2018	شقيق	1713	2018
2019	شهري	2318	2019
2020	صادق	1446	2020
2021	صافح		2021
2022	صافي		2022
2023	صب	2362	2023
2024	صحفي	851	2024
2025	صعود	2550	2025
2026	صلة	1710	2026
2027	صيفي	2327	2027
2028	صيني	1789	2028
2029	ضايق		2029
2030	ضحى	4326	2030
2031	ضعف	3460	2031
2032	طار	2354	2032
2033	طبع		2033
2034	طبق	2733	2034
2035	طبل		2035
2036	طرح	1217	2036
2037	طرّد		2037
2038	طلع	299	2038
2039	طلعة		2039

2040	طموح		2040
2041	طوفان		2041
2042	ظهيرة	4678	2042
2043	عاجل	1828	2043
2044	عارف	701	2044
2045	عاشر	1975	2045
2046	عالمي	300	2046
2047	عتمة	3677	2047
2048	عجيبة	4590	2048
2049	عراقي	160	2049
2050	عزز	2692	2050
2051	عشوق	2373	2051
2052	عقد	821	2052
2053	عقل	661	401
2054	عمامة		2054
2055	عمود	3136	2055
2056	عون	2535	2056
2057	غارق	4295	2057
2058	غالب	1403	2058
2059	غالي	811	2059
2060	غائب	3724	2060
2061	غرب	710	2061
2062	غفر	3459	2062
2063	غليظ		2063
2064	غنم		2064
2065	فاح	4713	2065
2066	فائدة	1143	2066
2067	فستان		2067
2068	فضول		2068
2069	فضي	4747	2069
2070	فكري	1332	2070
2071	فلسطيني	106	2071

2072	فلسفة	2455	2072
2073	قَاطِع		2073
2074	قانون	233	2074
2075	قَبْض		2075
2076	قَبْعة		2076
2077	قَبو		2077
2078	قَبيل		2078
2079	قَرُب	3253	2079
2080	قَط	2917	2080
2081	قل	413	2081
2082	قمر	1095	2082
2083	كَذَب	3118	2083
2084	كِذِب	2237	2084
2085	كراهية	2972	2085
2086	كفر	2921	2086
2087	لاحق		2087
2088	لاعب	402	2088
2089	لص	4145	2089
2090	لعن	3715	2090
2091	لعين	4337	2091
2092	مأخوذ		2092
2093	مالي	965	2093
2094	مبتسم	3842	2094
2095	مبعثر		2095
2096	مبكر	1386	2096
2097	متبادل	2689	2097
2098	متجه		2098
2099	متفرق	3268	2099
2100	متوقع	1285	2100
2101	محقق	2645	2101
2102	مُخْرِج	2210	2102
2103	مَخْرَن		2103

2104	مدام		2104
2105	مراجعة	1806	2105
2106	مرار		2106
2107	مراهق		2107
2108	مرعب		2108
2109	مريب		2109
2110	مستشار	1838	2110
2111	مسلح	759	2111
2112	مسمع		2112
2113	مضطرب		2113
2114	مطمئن	4071	2114
2115	معاملة	2072	2115
2116	معبد		2116
2117	معجب		2117
2118	مُعلَق	2963	2118
2119	معلوم	2888	2119
2120	مغرب	4617	2120
2121	مفهوم	1231	2121
2122	مقبل	530	2122
2123	مقيم	2087	2123
2124	مكاملة	2158	2124
2125	مكانة	2132	2125
2126	مل	4645	2126
2127	مليء	2666	2127
2128	موافقة	1527	2128
2129	مؤلم	2908	2129
2130	ناول		2130
2131	نجاح	619	2131
2132	نجم	980	2132
2133	نزع	3366	2133
2134	نزف	3953	2134
2135	نسائي	2270	2135

2136	نَشِيط		2136
2137	نفع		2137
2138	نفق	3669	2138
2139	نفي	2059	2139
2140	نقش		2140
2141	نَوْبَة		2141
2142	هاتفِي	2527	2142
2143	هاجم	4074	2143
2144	هرم	3081	2144
2145	وجع	4266	2145
2146	ورث		2146
2147	وَعْد	1884	2147
2148	وفر	2216	2148
2149	وَلِي		2149
2150	حَامِلٌ	3944	2150

BIBLIOGRAFÍA

- ‘aṭīfī, ‘amr ‘aḥmad. 2015. *ṣinā‘at al-mu‘yam al-‘arabī al-ḥadīṭ: dirāsa taṭbīqiyya*. Al-qāhira: ‘ālam al-kutub.
- Abu-Rabia, S. 1998. «Reading Arabic Texts: Effects of Text Type, Reader Type and Vowelization.» *Reading and Writing: An Interdisciplinary Journal* 10: 105-119.
- Abu-Rabia, S. 2002. «Reading in a Root-Based-Morphology language: The Case of Arabic.» *Journal of Research in Reading* 25 (3): 299-309.
- Abu-Rabia, S. 1997. «Reading in Arabic Orthography: The Effect of Vowels and Context on Reading Accuracy of Poor and Skilled Native Arabic Readers in Reading Paragraphs, Sentences and Isolated Words.» *Journal of Psycholinguistic Research* 26 (4): 465-482.
- Abu-Rabia, S. 1999. «The Effect of Arabic Vowels on the Reading Comprehension of Second- and Sixth-Grade Native Arab Children.» *Journal of Psycholinguistic Research* 28 (1): 93-101.
- Abu-Rabia, S. 2001. «The Role of Vowels in Reading Semitic Scripts: Data from Arabic and Hebrew.» *Reading and Writing: An Interdisciplinary Journal* 14: 39-59.
- Abu-Rabia, S., y J. Awwad. 2004. «Morphological Structures in Visual Word Recognition: The Case of Arabic.» *Journal of Research in Reading* 27 (3): 321-336.

- Acquaroni, R. 2007. *Las palabras que no se lleva el viento: literatura y enseñanza de español como LE/L2*. Madrid: Santillana y Universidad de Salamanca.
- Aguilar, V. 2011. «Enseñanza del árabe en España.» *Afkar/Ideas*, Otoño: 82-84.
- Aguilar, V., P. Santillán Grimm, y L.M.P. Cañada. 2010. «Introducción de ARABELE 2009: Enseñanza y aprendizaje de la lengua árabe.» *ARABELE 2009*. España: Universidad de Murcia, Casa Árabe-IEM, Escuela de Traductores de Toledo. 9-14.
- Al-Batal, Mahmoud. 2006. «Playing with Words: Teaching Vocabulary in the Arabic Curriculum.» En *Handbook for Arabic Language Teaching Professionals in the 21st Century*, de Kassem M. Wahba, Zeinab A. Taha y Liz England, 331-340. New York: Routledge.
- Alcántara Plá, Manuel. 2007. *Introducción al análisis de estructuras lingüísticas en corpus*. Madrid : UAM Ediciones.
- Alderson, J. C. 2000. *Assessing Reading*. Cambridge: Cambridge University Press.
- Alderson, J. C. 1984. «Reading in a Foreign Language: A Reading Problem or a Language Problem?» En *Reading in a Foreign Language*, de J. C. Alderson y A. H. Urquhart, 1-24. London: Longman.
- Allen, R. 2000. *An Introduction to Arabic Literature*. Cambridge: Cambridge University Press.
- Allen, R. 1995. «The Arabic Short Story and the Status of Women.» En *Love and Sexuality in Modern Arabic Literature*, de Allen et al, 77-90. London: Saqi Books.
- Allen, Roger. 1985. «Arabic Proficiency Guidelines.» *al-‘Arabiyya* 18 (1/2): 45-70.
- Allen, Roger. 1990. «Proficiency and the Teacher of Arabic: Curriculum, Course, and Classroom.» *al-‘Arabiyya* 23 (1/2): 1-30.
- Allen, Roger. 1984. «Proficiency Guidelines for Arabic.» *al-‘Arabiyya* 17 (1/2): 119-123.

- . 1995. *The Arabic Novel: An Historical and Critical Introduction*. Syracuse, N.Y.: Syracuse University Press.
- al-Mahrooqi, Rahma, y Christopher Denman. 2016. «Arab Identity and Literature in Translation: The Politics of Selection and Representation.» *Arab World English Journal* 5: 5-20.
- Altantawy, Mohamed, Nizar Habash, Owen Rambow, y Ibrahim Saleh. 2010. «Morphological Analysis and Generation of Arabic Nouns: A Morphemic Functional Approach.» *Language Resources and Evaluation Conference*.
- and, William Brinner. s.f. «Readings in Modern Arabic Literature.» Brill.
- Anderson, Neil J. 1999. «Improving Reading Speed: Activities for the Classroom.» *English Teaching Forum* 2-5.
- Arvide Cambra, L. M. 1997. «Apuntes sobre métodos de enseñanza y técnicas de aprendizaje de la lengua árabe (I).» *Anaquel de Estudios Árabes* VIII: 41-56.
- Arvide Cambra, L. M. 1998. «Apuntes sobre métodos de enseñanza y técnicas de aprendizaje de la lengua árabe (II).» *Anaquel de Estudios Árabes* X: 9-17.
- Arvide Cambra, L. M. 2000. «Apuntes sobre métodos de enseñanza y técnicas de aprendizaje de la lengua árabe (III).» *Anaquel de Estudios Árabes* XI: 109-122.
- Ashour, R., M. Berrada, F. J. Ghazoul, y A. Rachid. 2009. «Arab Women Writers.» *Southwest Review* 94 (1): 9-18.
- Asín Palacios, M. 1959. *Crestomatía de árabe literal con glosario y elementos de gramática. Edición corregida*. Madrid: Editorial Maestre.
- Atkins, S., J. Clear, y N. Ostler. 1992. «Corpus Design Criteria.» *Literary and Linguistic Computing* 7 (1): 1-16.
- Baayen, R. Harald, Petar Milin, y Michael Ramscar. 2016. «Frequency in Lexical Processing.» *Aphasiology* 1-47.
- Baayen, R.H. 2001. *Word Frequency Distributions*. Dordrecht, Netherlands: Springer.

- Badawi, E. (1975). *Mustawayaat al-luġa al-‘arabiyya fii miṣr: baḥth fii ‘alaaqaat al-luġa bi-l-ḥa ḍaara*. Daar al-ma‘aarif bi-miṣr.
- Badawi, E. 2002. «In the Quest for the Level 4+ in Arabic.» En *Developing Professional-Level Language Proficiency*, de Betty Lou Leaver y Boris Shekhtman, 156-176. Cambridge: Cambridge University Press.
- Badawi, E., M. G. Carter, y A. Gully. 2004. *Modern Written Arabic: A Comprehensive Grammar*. London: Routledge.
- Badawi, M. M. 1993. *A Short History of Modern Arabic Literature*. Oxford: Clarendon Press.
- Badry, F. 2009. «Milestones in Arabic Language Development. .»
https://www.academia.edu/4657235/Milestones_in_Arabic_Language_Development.
- Badry, Fatima. 2004. *Acquiring the Arabic Lexicon*. Bethesda: Academica Press, LLC.
- Baker, P. 2014. *Using Corpora to Analyze Gender*. London: Bloomsbury.
- Barnett, M. A. 1989. *More Than Meets the Eye. Foreign Language Reading: Theory and Practice*. New Jersey: Center for Applied Linguistics.
- Baroni, Marco. 2009. «Distributions in Text.» En *Corpus Linguistics: An International Handbook*, de Anke Lüdeling y Meja Kytö, 777-802. Berlin: Mouton de Gruyter.
- Barreda Sureda, J, C. García Cecilia, y F. Ramos López. 2007. *La lengua árabe en el espacio europeo de educación superior. Consideraciones para la enseñanza y el aprendizaje del árabe como lengua extranjera en la universidad*. Vol. 2, de *Modelos de organización de profesores en la educación universitaria. Redes de investigación docente.*, de M. J. Frau y N. Sauleda Parés, 253-269. Alicante: Universidad de Alicante-Editorial Marfil.
- Bassiouney, Reem. 2009. *Arabic Sociolinguistics: Topics in Diglossia, Gender, Identity, and Politics*. Washington, DC: Georgetown University Press.

- Bate, J., y P. Byrne. 2016. *Literature and Mental Health*. Prod. The University of Warwick. <https://www.futurelearn.com/courses/literature>.
- Bauer, Laurie, y Paul Nation. 1993. «Word Families.» *International Journal of Lexicography* 6 (4): 253-279.
- Bell, Timothy. 2001. «Extensive reading: Speed and Comprehension.» *The Reading Matrix* 1 (1).
- Belnap, Kirk. 1987. «Who's Taking Arabic and What on Earth For? A Survey of Students in Arabic Language Programs.» *Al-Arabiyya* 20: 29-42.
- Bernhardt, E., y M. Kamil. 2010. «Reading in a Second Language.» En *Concise Encyclopedia of Applied Linguistics*, de M. Berns, 179-186. Oxford: Elsevier.
- Biber, Douglas, y Susan Conrad. 2009. *Register, Genre, and Style*. Cambridge: Cambridge University Press.
- Biber, Douglas. 1990. «Methodological Issues Regarding Corpus-based Analyses of Linguistic Variation.» *Literary and Linguistic Computing* 5 (4): 257-269.
- Biber, Douglas. 1993. «Representativeness in Corpus Design.» *Literary and Linguistic Computing* 8 (4): 243-257.
- Biber, Douglas, Susan Conrad, y Randi Reppen. 1998. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.
- Boudelaa, S. 2014. «Is the Arabic Mental Lexicon Morpheme-Based or Stem-Based? Implications for Spoken and Written Word Recognition.» En *Handbook of Arabic Literacy: Insights and Perspectives*, de Saiegh-Haddad y Joshi, 31-54. London: Springer.
- Boulton, Alex. 2010. «Data-Driven Learning: On Paper, in Practice.» En *Corpus Linguistics in Language Teaching*, de Tony Harris y María Moreno Jaén, 17-51. Bern: Peter Lang.

- Brinner, William M., y Mounah A. Khouri. 1971. *Readings in Modern Arabic Literature: Selected and edited with vocabularies and notes. Part One: The short story and the novel.* Leiden: E. J. Brill.
- Brosh, H, y L. Mansur. 2013. *Arabic Stories for Language Learners: Traditional Middle Eastern Tales in Arabic and English.* Tuttle Publishing.
- Brustad, Kristen. 2006. «Reading Fluently in Arabic.» En *Handbook of Arabic Language Teaching Professionals in the 21st Century*, de K. Wahba, Z. Taha y L. England, 341-352. New Jersey: Routledge.
- Buckwalter, Tim, y Dilworth Parkinson. 2011. *A Frequency Dictionary of Arabic Core Vocabulary for Learners.* Abingdon: Routledge.
- Busse, Beatrix. 2014. «Genre.» En *The Cambridge Handbook of Stylistics*, de Peter Stockwell y Sara Whiteley, 103-116. Cambridge: Cambridge University Press.
- Calfee, R. C., y R. L. Venezky. 1969. «Component Skills in Beginning Reading.» En *Psycholinguistics and the Teaching of Reading*, de Goodman y Fleming, 91-110. Delaware: International Reading Association.
- Cantos Gómez, Pascual. 2002. «Do we need statistics when we have linguistics?» *DELTA* 233-271.
- Carter, R. 2007. «Literature and Language Teaching 1986-2006: A Review.» *International Journal of Applied Linguistics* 17 (1): 3-13.
- Carter, R., y M. J. McCarthy. 2001. «Size Isn't Everything: Spoken English, Corpus and the Classroom.» *TESOL Quarterly* 35 (2): 337/340.
- Carver, R. 1992. «Reading Rate: Theory, Research, and Practical Implications.» *Journal of Reading* 36: 84-95.
- Colla, Elliott. 2009. «How Zaynab Became the First Arabic Novel.» *History Compass* 7 (1): 214-225.
- Collie, J., y S. Slater. 1987. *Literature in the Language Classroom: A Resource Book of Ideas and Activities.* Cambridge : Cambridge University Press.

- Cook, V., y B. Bassetti. 2010. «An Introduction to Researching Second Language Writing Systems.» En *Second Language Writing Systems*, de V. Cook y B. Bassetti, 1-67. England: Viva Books.
- Corriente, Federico. 1988. *Gramática árabe - Segunda edición*. Barcelona: Herder.
- Corriente, Federico. 1999. «Las etimologías árabes en la obra de Joan Corimines.» En *L'obra de Joan Coromines: Cicle d'estudi i homenatge*, de Joan Solá, 67-87. Sabadell: Fundació Caixa de Sabadell.
- Corriente, Federico, J. P. Monferrer, y A. S. Ould M. Baba. 2013. *Vocabulario Árabe Graduado*. Barcelona: Herder.
- Crawford, Willima J., y Eniko Csomay. 2016. *Doing Corpus Linguistics*. New York: Routledge.
- Darton, G. C. 1945. «The "New Method" of Teaching English.» *African Studies* 4 (1): 41-44.
- Day, Richard R., y Julian Bamford. 1998. *Extensive Reading in the Second Language Classroom*. Cambridge: Cambridge University Press.
- Diab, M., y N. Habash. 2014. «Natural Language Processing of Arabic and its Dialects (a tutorial).» *Proceedings EMNLP*. Doha, Qatar.
<http://www.nizarhabash.com/tutorials/EMNLP-2014-Diab+Habash-Tutorial.pdf>.
- Díaz-Pérez, F.J. et al. 2013. *Global Issues in the Teaching of Language, Literature and Linguistics*. Bern: Peter Lang.
- Doss, M. y Davies, H. T. (2013) *al 'ammiyya l-maSSriyya l-maktuuba: muhtaaraat min 1401 'ilaa 2009*. al-hay'a l-maSriyya l-'aama li-l-kitaab.
- Duff, A., y A. Maley. 1990. *Literature*. Oxford: Oxford University Press.
- Elyasi, Javid. 2013. «Teaching language through literature: A Content-based Instruction model.» *Journal of Academic and Applied Studies* 3 (10): 9-18.

- Eviatar, Z., y R. Ibrahim. 2014. «Why is it Hard to Read Arabic?» En *Handbook of Arabic Literacy: Insights and Perspectives*, de Saiegh-Haddad y Joshi, 77-96. London: Springer.
- Faḍl, Ṣ. (2015). *Sardiyyāt al-qarn al-ŷadīd*. Al-Qāhira: Al-daar al-Maṣriyya al-lubnāniyya.
- Familiar, Laila. 2013. *Hoda Barakat's Sayyidi wa Habibi: The Authorized Abridged Edition for Students of Arabic*. Washington, DC: Georgetown University Press.
- . 2017. *Saud al-Sanousi's Saaq al-Bambu: The Authorized Abridged Edition for Students of Arabic*. Washington, DC: Georgetown University Press.
- . 2008. «The role of the Arabic root in guessing the meaning of unfamiliar nominals. M.A. Thesis.»
https://www.academia.edu/7042882/The_Role_of_the_Arabic_Root_in_Guessing_the_Meaning_of_Unfamiliar_Nominals_-_M.A._Thesis_2008.
- Farghaly, A., y K. Shaalan. 2009. «Arabic Natural Language Processing: Challenges and Solutions.» *ACM Transactions on Asian Language Information Processing* 8 (4): 1-22.
- Ferguson, C. A. 1959. «Diglossia.» *Word* 15: 325-340.
- Fernández-Parrilla, Gonzalo. 2012. «Reseña a "Aproximación al cuento narrado en árabe marroquí.» *Anaquel de Estudios Árabes* 23: 155-156.
- Frangieh, B. K. 2005. *Anthology of Arabic Literature, Culture and Thought from Pre-Islamic Times to the Present*. Yale University Press.
- Gamal, Adel. 1998. «Vocabulary Studies from Arabic and Western Perspectives: Theory and Practice-Modern Standard Arabic. The Intermediate Level.» *al-ʿArabiyya* 55-87.
- Gamel, M. K. 1984. «Language and Literature: Allies Not Enemies.» *ADFL Bulletin* 15 (3): 8-11.

- García-Barrero, David, Manuel Fera, y Maria Teresa Turell. 2012. «International Association of Forensic Linguistics.» Porto, Portugal. 42-56.
- Gass, Susan. 2013. *Second Language Acquisition: An Introductory Course*. New York: Routledge.
- Gass, Susan, Jennifer Behney, y Luke Plonsky. 2013. *Second Language Acquisition: An Introductory Course*. New York: Routledge.
- Giménez Reillo, Antonio. 2014. «Amplia docencia anónima.» *Anís del Moro*. 19 de agosto. <https://anisdelmoro.blogspot.com/2014/08/amplia-docencia-anonima.html>.
- . 2014. «Especialistas de irreconocible competencia.» *Anís del Moro*. 24 de octubre. <https://anisdelmoro.blogspot.com/2014/10/especialistas-de-irreconocible.html>.
- Goodman, K. S. 1969. «Analysis of Oral Language Miscues. Applied Psycholinguistics.» *Reading Research Quarterly* 5: 9-30.
- Goodman, K. S. 1979. «Foreword.» En *Language and Literacy: the Selected Writings of Kenneth S. Goodman*, de Gollasch, vii-ix. Londong: Routledge y Kegan Paul.
- Goodman, K. S. 1967. «Reading: A Psycholinguistic Guessing Game.» *Journal of the Reading Specialist* 6: 126-135.
- Goodman, K. S. 1968. «The Psycholinguistic Nature of the Reading Process.» En *The Psycholinguistic Nature of the Reading Process*, de K. S. Goodman, 15-26. Detroit: Wayne State University Press.
- Grabe, W. 1988. «Reassessing the Term "Interactive" .» En *Interactive Approaches to Second Language Reading*, de P. L. Carrell, J. Devine y D. E. Eskey, 56-70. Cambridge: Cambridge University Press.
- Grabe, W., y F. R. Stoller. 2013. *Teaching and Researching Reading*. Abingdon: Routledge.

- Gries, Stefan Th. 2008. «Dispersions and Adjusted Frequencies in Corpora.» *International Journal of Corpus Linguistics* 13 (4): 403-437.
- Gries, Stephan. 2015. «Quantitative designs and statistical techniques.» En *The Cambridge Handbook of English Corpus Linguistics*, de Douglas Biber y Randi Reppen, 50-71. Cambridge: Cambridge University Press.
- Gries, Stephan, y Nick Ellis. 2015. «Statistical Measures for Usage-Based Linguistics.» *Language Learning* 65: 228-255.
- Gómez Camarero, C. 1994. «La producción científica española en la literatura árabe contemporánea (II).» *Miscelánea de Estudios Árabes y Hebraicos XLII-XLIII*: 97-111.
- Ha, L.Q., D.W. Stewart, P. Hanna, y F.J. Smith. 2006. «Zipf and Type-Token Rules for the English, Spanish, Irish and Latin Languages.» *Web Journal of Formal, Computational y Cognitive Linguistics* 8.
https://www.researchgate.net/publication/228794781_Zipf_and_Type-Token_rules_for_the_English_Spanish_Irish_and_Latin_languages.
- Habash, Nizar. 2010. *Introduction to Arabic Natural Language Processing*. Morgan y Claypool Publishers.
- Habash, Nizar, y Owen Rambow. 2006. «MAGEAD: a morphological analyzer and generator for the Arabic dialects.» *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics. Association for Computational Linguistics*.
- Haeri, N. 2009. «The Elephant in the Room: Language and Literacy in the Arab World.» En *The Cambridge Handbook of Literacy*, de D. Olson y N. Torrance, 418-430. Cambridge: Cambridge University Press.
- Hafez, Sabry. 1993. «The Genesis of Arabic Narrative Discourse: A Study in the Sociology of Modern Arabic Literature.» London: Saqi Books.
- Hafez, Sabry. 1995. «Women's Narrative in Modern Arabic Literature: A Typology.» En *Love and Sexuality in Modern Arabic Literature*, de Allen et al, 244-247. London: Saqi Books.

- Hansen, G. F. 2010. «Word Recognition in Arabic as a Foreign Language.» *The Language Modern Journal* 94 (4): 567-581.
- Hansen, G. F. 2014. «Word Recognition in Arabic: Approaching a Language-Specific Reading Model.» En *Handbook of Arabic Literacy: Insights and Perspectives*, de Saiegh-Haddad y Joshi, 55-76. London: Springer.
- Hansen, Gunna Funder, y Saliha Marie Fettah. 2007. *Read Arabic: An Arabic Reader for Arabic Students at all levels*. Odense: University Press of Southern Denmark.
- Heatly, A, I.S.P. Nation, y A Coxhead. 2002. «RANGE and FREQUENCY programs.» 1-15. http://www.vuw.ac.nz/lals/staff/Paul_Nation.
- Hermena, E. W., D. Drieghe, S. Hellmuth, y S. P. Liversedge. 2015. «Processing of Arabic Diacritical Marks: Phonological-Syntactic Disambiguation of Homographic Verbs and Visual Crowding Effects.» *Journal of Experimental Psychology: Human Perception and Performance* 41 (2): 494-507.
- Hill, David R. 2013. «Graded readers.» *ELT Journal* (Oxford University Press) 67/1: 85-125.
- Hill, David R. 2008. «Graded readers in English.» *ELT Journal* 62 (2): 184-204.
- Hirsh, David, y Paul Nation. 1992. «What Vocabulary Size is Needed to Read Unsimplified Texts for Pleasure?» *Reading in a Foreign Language* 8 (2): 689-696.
- Horst, Marlise. 2009. «Developing Definitional Vocabulary Knowledge and Lexical Access Speed through Extensive Reading.» *Second Language Reading: Research and Instruction* 40-64.
- Horwitz, Elaine, Michael Horwitz, y Joann Cope. 1986. «Foreign Language Classroom Anxiety.» *The Modern Language Journal* 125-132.
- Hudson, T. 2007. *Teaching Second Language Reading*. Oxford: Oxford University Press.
- Husni, R., y D. L. Newman. 2008. *Modern Arabic Short Stories*. London: Saqi Books.

- Ibrahim, Z. s.f. «Egyptian and Lebanese Written Modern Standard Arabic: Are They One and the Same? Unpublished PhD dissertation.» Georgetown Univeresity.
- Ibrahim, Z. 2008. «Lexical Variation in Modern Standard Arabic.» En *The Encyclopedia of Arabic Language and Linguistics (EALL)*, 13-21. Leiden: Brill Publishers.
- Ismail, Sherif H. 2015. «Arabic Literature Into English: The (Im)possibility of Understanding.» *Interventions* 17 (6): 916-931.
- Iwahori, Yurika. 2008. «Developing reading fluency: A study of extensive reading in EFL.» *Reading in a Foreign Language* 70-91.
- Jayyusi, Salma Khadra. 2005. *Modern Arabic Fiction: An Anthology*. New York: Columbia University Press.
- Johnson, H. L. 1956. «A Review of the Teaching of Modern Languages (Problems in Education-X), a UNESCO Publication.» *American Association of Teachers of Spanish and Portuguese* 39 (2): 234-236.
- Just, M. A., y P. A. Carpenter. 1987. *The Psychology of Reading and language Comprehension*. Boston, MA: Allyn y Bacon.
- Keen, S. 2007. *Empathy and the Novel*. Oxford: Oxford University Press.
- Khaldieh, S. 1999. «An Investigation into Reader-Text Interaction: Native, Proficient, and Less Proficient Non-Native Readers of Arabic.» *Al-Arabiyya* 32: 119-159.
- Khaldieh, S. 2001. «The Relationship Between Knowledge of i'raab, Lexical Knowledge, and Reading Comprehension of Nonnative Readers of Arabic.» *The Modern Language Journal* 85 (iii): 416-431.
- Khaldieh, S. 1996. «Word Recognition of Arabic as a Foreign Language by American Learners.» *Al-Arabiyya* 29: 129-152.
- Khamis, Noorli, y Imran Ho Abdullah. 2013. «Wordlists Analysis: Specialised Language Categories.» *Pertanika Journal of Social Sciences y Humanities* 21 (4): 1563-1581.

- Khorshid, Ahmed H. 2009. *Sahlawayhi 1: Graded Stories - Beginners - Level 1 (350 words)*. Publisher: Author.
- . 2009. *Sahlawayhi 2: Graded Stories - Beginners - Level 2 (450 words)*. Publisher: Author.
- . 2009. *Sahlawayhi 3: Graded Stories - Beginners - Level 3 (550 words)*. Publisher: Author.
- . 2011. *Sahlawayhi 4: Graded Stories - Intermediate - Level 4 (650 words)*. Publisher: Author.
- . 2013. *Sahlawayhi 5: Graded Stories - Intermediate - Level 5 (750 words)*. Publisher: Author.
- Kilgariff, Adam, Frieda Charalabopoulou, Maria Gavrilidou, Janne Bondi Johannessen, Saussan Khalil, Sofie Johansson Kokkinakis, Robert Lew, Serge Sharoff, Ravikiran Vadlapudi, y Elena Volodina. 2014. «Corpus-based vocabulary lists for language learners for nine languages.» *Language Resources y Evaluation* 48 (1): 121-163.
- Kilgariff, Adam. 2001. «Comparing Corpora.» *International Journal of Corpus Linguistics* 6 (1): 97-133.
- Koda, K. 2007. «Reading and Language Learning: Crosslinguistic Constraints on Second Reading Development.» *Language Learning* (University of Michigan) 57 (1): 1-44.
- Koester, A. 2010. «Building Small Specialized Corpora.» En *The Routledge Handbook of Corpus Linguistics*, de O'Keeffe y McCarthy, 66-79. London: Routledge.
- Koppel, M., S. Argamon, y A. R. Shimoni. 2002. «Automatically Categorizing Written Texts by Author Gender.» *Literary and Linguistic Computing* 17 (4): 401-412.
- Kramsch, C., y O. Kramsch. 2000. «The Avatars of Literature in Language Study.» *The Modern Language Journal* 84 (iv): 553-573.

- Krashen, S. D. 2004. *The Power of Reading: Insights from the Research*. USA: Heinemann.
- Laufer, Batia. 1992. «How Much Lexis is Necessary for Reading Comprehension?» En *Vocabulary and Applied Linguistics*, de P. J. L. Arnaud y H. Bejoint, 126-132. London: Mcmillan.
- Lazar, G. 1993. *Literature and Language Teaching: A Guide for Teachers and Trainers*. Cambridge: Cambridge University Press.
- Lee, Y., y D. Hirsh. 2012. «Quality and Quantity of Exposure in L2 Vocabulary Learning.» En *Current Perspectives in Second Language Vocabulary Research*, de David Hirsh, 79-116. Bern: Peter Lang.
- Leech, Geoffrey, Paul Rayson, y Andrew Wilson. 2001. *Word Frequencies in Written and Spoken English*. Harlow: Longman.
- . 2001. *Word Frequencies in Written and Spoken English*. London: Longman.
- Lewis, M. 1997. *Implementing the Lexical Approach*. Hove: Language Teaching Publications.
- Maamouri, M., A. Biens, T. Buckwalter, y W. Mekki. 2004. «The Penn Arabic Treebank: Building a Large-Scale Annotated Arabic Corpus.» *NEMLAR Conference on Arabic Language Resources and Tools*.
- Macdonald, A., y G. Macdonald. 1983. «Teaching Metaphor in a Second Language.» *Exercise Exchange* 28 (2): 7-11.
- Martí Tormo, V. 2012. «Metodología pedagógica e investigación en el campo de la enseñanza y aprendizaje de árabe como segunda lengua: Elementos para un estado de la cuestión.» *MEAH: Sección árabe-Islam* 61: 37-60.
- Martínez Castro, A. 2014. «El teatro por escenas en la adquisición del árabe desde el nivel básico del MCER.» En *ARABELE 2012: Enseñanza y aprendizaje de la lengua árabe*, 239-250. Murcia: Editum, Ediciones de la Universidad de Murcia.

- Martinez, Ron, y Norbert Schmitt. 2015. «Vocabulary.» En *Cambridge Handbook of English Corpus Linguistics*, de Douglas Biber y Randi Reppen, 439-459. Cambridge: Cambridge University Press.
- Matsuoka, Warren. 2012. «Searching for the Right Words: Creating Word Lists to Inform EFL Learning.» En *Current Perspectives in Second Language Vocabulary Research*, de David Hirsh, 151-177. Bern: Peter Lang.
- McEnery, Tony, y Andrew Hardie. 2012. *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.
- McEnery, Tony, y Andrew Wilson. 1997. «Teaching and Language Corpora (TALC).» *ReCALL* 9 (1): 5-14.
- McKay, S. 1982. «Literature in the ESL Classroom.» *TESOL Quarterly* 16 (4): 529-536.
- McManus, Anne-Marie E. 2016. «Scale In The Balance: Reading With The International Prize for Arabic Fiction ("The Arabic Booker").» *International Journal of Middle East Studies* 48: 217-241.
- Mejdell, Gunvor. 2006. «The Use of Colloquial in Modern Egyptian Literature - A Survey.» En *Current Issues in the Analysis of Semitic Grammar and Lexicon II*, de Lutz Edzard y Jan Retso, 195-213. Harrassowitz Verlag.
- Milton, James. 2013. «Measuring the contribution of vocabulary knowledge to proficiency in the four skills.» En *L2 vocabulary acquisition, knowledge and use. New perspectives on assessment and corpus analysis.*, de C. Bardel, C. Lindqvist y B. Laufer, 57-78. Amsterdam: EUROSLA.
- Milton, James, y Ahmed Masrai. 2016. «How Different is Arabic from Other Languages? The Relationship between Word Frequency and Lexical Coverage.» *Journal of Applied Linguistics and Language Research* 3 (1): 15-35.
- Moscoso García, F. 2004. *Cuentos en dialecto árabe de Ceuta*. Toledo: Universidad de Castilla-La Mancha y Escuela de Traductores de Toledo.

- Mughazy, M. 2005. «Reading Despite Ambiguity: The Role of Metacognitive Strategies in Reading Arabic Authentic Texts.» *Al-Arabiyya* 38/39: 57-74.
- Myhill, J. 2014. «The Effect of Diglossia on Literacy in Arabic and Other Languages.» En *Handbook of Arabic Literacy: Insights and Perspectives*, de Saiegh-Haddad y Joshi, 197-223. Dordrecht: Springer.
- Nagy, Willima E. 2010. «The Word Games.» En *Bringing Reading Research to Life*, de M. G. y Kucan, L. McKeown, 72-91. New York: The Guilford Press.
- Nation, Paul, y Karen Wang Ming-tzu. 1999. «Graded Readers and Vocabulary.» *Reading in a Foreign Language* 12 (2): 355-380.
- Nation, Paul. 2004. «A study of the most frequent word families in the British National Corpus.» En *Vocabulary in a Second Language. Selection, Aquisition, and Testing.*, de Paul Bogaards y Batia Laufer, 3-13. Amsterdam: John Benjamins Publishing Company.
- Nation, Paul. 2006. «How Large a Vocabulary is Needed for Reading and Listening?» *Canadian Modern Language Review* 63: 59-82.
- Nation, Paul. 2014. «How much input do you need to learn the most frequent 9,000 words.» *Reading in a Foreign Language* 26 (2): 1-16.
- . 2013. *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- . 2016. *Making and Using Word Lists for Language Learning and Testing*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Nation, Paul. 2015. «Principles guiding vocabulary learning through extensive reading.» *Reading in a Foreign Language* 27 (1): 136-145.
- Nation, Paul, y Karen Wang Ming-tzu. 1999. «Graded Readers and Vocabulary.» *Reading in a Foreign Language* 12 (2): 355-380.
- Nation, Paul, y Robert Waring. 1997. «Vocabulary size, text coverage and word lists.» En *Vocabulary: Description, Acquisition and Pedagogy*, de Norbert Schmitt y Michael McCarthy, 6-19. Cambridge: Cambridge University Press.

- Nation, Paul, y Stuart Webb. 2011. *Researching and Analyzing Vocabulary*. Boston: Heinle, Cengage Learning.
- Nelson, M. 2010. «Building a Written Corpus: What are the Basics?» En *The Routledge Handbook of Corpus Linguistics*, de O'Keeffe y McCarthy, 53-65. London: Routledge.
- Newman, M. L., C. J. Groom, L. D. Handelman, y J. W. Pennebaker. 2008. «Gender Differences in Language Use: An Analysis of 14,000 Text Samples.» *Discourse Processes* 45: 211-236.
- Nuttall, Christine. 1996. *Teaching Reading Skills in a Foreign Language*. Oxford: Heinemann.
- O'Keeffe, Anne, Michael McCarthy, y Ronald Carter. 2007. *From Corpus to Classroom: Language Use and Language Teaching*. Cambridge: Cambridge University Press.
- Papadopoulos, P. M., A. Karatsolis, y Z. Ibrahim. 2013. «Learning Activities, Educational Games, and Tangibles: Arabic Language Learning in the ALADDIN Project.» *Proceedings of the 17th Panhellenic Conference on Informatics (PCI 13)*. New York, NY. 98-105.
<http://doi.acm.org/10.1145/2491845.2491852>.
- Papadopoulos, P. M., Z. Ibrahim, y A. Karatsolis. 2015. «Educational Games for Early Childhood: Using Tabletop Surface Computers for Teaching the Arabic Alphabet.» Editado por M. Helfert, M. T. Restivo, S. Zvacek y J. Uhomoihi. *Proceedings of 7th International Conference on Computer Supported Education - CSEDU 2015*. 130-138.
- . 2014. «Teaching the Arabic Alphabet to Kindergarteners: Writing Activities on Paper and Surface Computers.» Editado por M. Svacek, T. Restivo y J. Uhomoihi. *Proceedings of the 6th International Conference on Computer Supported Education, CSEDU 2014*. Barcelona: SCITEPRESS. 433-439.
- Paradela, Nieves. 2000. «La literatura árabe moderna en el arabismo español.» *Awraq XXI*: 221-250.

- Parkinson, D. B. 1985. «Proficiency to Do What? Developing Oral Proficiency in Students of Modern Standard Arabic.» *Al-Arabiyya* 18: 11-43.
- Parkinson, D., y Z. Ibrahim. 1999. «Testing Lexical Differences in Regional Standard Arabics.» En *Perspectives on Arabic Linguistics XII*, de Benmamoun, 183-202. Philadelphia: John Benjamins.
- Pasha, Arfath, Mohammad Al-Badrashiny, Mona Diab, Nizar Habash, Manoj Pooleery, Owen Rambow, y Ryan Roth. s.f.
- . 2014. «MADAMIRA v1.0 Beta User Manual.» Center for Computational Learning Systems Columbia University.
- . 2014. «MADAMIRA v1.0 Beta User Manual.» *MADAMIRA v1.0 Beta User Manual*. Center for Computational Learning Systems Columbia University, January.
- Pasha, Arfath, Mohammad Al-Badrashiny, Mona Diab, Nizar Habash, Manoj Pooleery, Owen Rambow, y Ryan Roth. 2014. *MADAMIRA v1.0 Beta User Manual*. Center for Computational Learning Systems Columbia University.
- Pennebaker, J. W. 2011. *The Secret Life of Pronouns*. New York: Bloomsbury Press.
- Pérez-Paredes, Pascual. 2010. «Corpus Linguistics and Language Education in Perspective: Appropriation and the Possibilities Scenario.» En *Corpus Linguistics in Language Teaching*, de Tony Harris y María Moreno Jaén, 53-73. Bern: Peter Lang.
- Perfetti, Charles. 2010. «Decoding, Vocabulary, and Comprehension: The Golden Triangle of Reading Skill.» En *Bringing Reading Research to Life*, de M. G. y Kucan, L. McKeown, 291-302. New York: The Guilford Press.
- Phillips, B. 2013. «Global Citizenship, Ecocriticism and the Teaching of Literature.» En *Global Issues in the Teaching of Language, Literature and Linguistics*. Bern: Peter Lang.
- Rayner, K., y A. Pollatsek. 1989. *The Psychology of Reading*. New Jersey: Prentice-Hall International.

- Reinhardt, Jonathon. 2010. «The Potential of Corpus-Informed L2 Pedagogy.» *Studies in Hispanic and Lusophone Linguistics* 3 (1): 239-251.
- Renandya, W. A. y Jacobs, G. M. 2002. «Extensive reading: Why aren't we all doing it?» En *Methodology in language teaching: An anthology of current practice*, de J. C. Richards y W. A. Renandya, 295-302. Cambridge: Cambridge University Press.
- Reppen, R. 2010. «Building a Corpus: What are the Key Considerations?» En *The Routledge Handbook of Corpus Linguistics*, de O'Keeffe y McCarthy, 31-37. London: Routledge.
- Reynolds, Barry Lee. 2014. «A Mixed-Methods Approach to Investigating First- and Second-Language Incidental Vocabulary Acquisition Through the Reading of Fiction.» *Reading Research Quarterly* 50 (1): 111-127.
- Richards, J. C., y T. S. Rodgers. 1986. *Enfoques y métodos en la enseñanza de idiomas*. Cambridge: Cambridge University Press.
- Richards, Jack C., John Platt, y Heidi Platt. 1997. *Diccionario de lingüística aplicada y enseñanza de lenguas*. Barcelona: Ariel.
- Ricks, Robert Stephen. 2015. «The development of frequency-based assessments of vocabulary breadth and depth for L2 Arabic. Ph.D. dissertation.» Washington, DC: Georgetown University.
- Rivers, W. 1981. *Teaching Foreign Language Skills*. University of Chicago Press.
- Roman, G., y B. Pavard. 1987. «A Comparative Study: How We Read Arabic and French.» Editado por J. K. O'Regan y A. Levy-Schoen. *Eye Movements from Physiology to Cognition: Selected/Edited Proceedings of the Third European Conference on the Eye Movements, Dourdan, France*. Amsterdam: Elsevier.
- Romanov, Maxim. 2015. *Classic Arabic Through the Words of the Prophet: A Frequency-Based Hadith Reader*. Perseus Project y Department of Classics: Tufts University.
- Rossner, R. 1983. «Talking Shop: H. G. Widdowson on literature and ELT.» *ELT Journal* 37 (1): 30-35.

- Ryding, Karin C. 2013. *Teaching and Learning Arabic as a Foreign Language*. Georgetown: Georgetown University Press.
- Sabatini, J. P., T. O'Reilly, L. K. Halderman, y K. Bruce. 2014. «Integrating Scenario-Based and Component Reading Skill Measures to Understand the Reading Behavior of Struggling Readers.» *Learning Disabilities Research y Practice* 29 (1): 36-43.
- Saiegh-Haddad, E., y B. Spolsky. 2014. «Acquiring Literacy in a Diglossic Context: Problems and Prospects.» En *Handbook of Arabic Literacy: Insights and Perspectives*, de Saiegh-Haddad y Joshi, 225-240. Dordrecht: Springer.
- Saiegh-Haddad, E., y Joshi, . 2014. *Handbook of Arabic Literacy: Insights and Perspectives*. Dordrecht: Springer.
- Sakkut, Hamdi. 2000. *The Arabic Novel: Bibliography and Critical Introduction 1865-1995*. Cairo: The American University in Cairo Press.
- Saleh, W. 2010. «استثمار القصة في تعليم العربية للناطقين بغيرها.» *ARABELE 2009: Enseñanza y aprendizaje de la lengua árabe*. Universidad de Murcia, Casa Árabe-IEM, Escuela de Traductores de Toledo. 219-229.
- . 2014. «العناصر الاجتماعية والثقافية في تعليم العربي.» *ARABELE 2012: Enseñanza y aprendizaje de la lengua árabe*. Editum, Ediciones de la Universidad de Murcia. 99-114.
- . 2003. *Cuentos tradicionales árabes: Antología didáctica y bilingüe*. Madrid: Ibersaf.
- . 2001. *Curso práctico de lengua árabe I*. Madrid: Ibersaf Editores.
- . 2002. *Curso práctico de lengua árabe II*. Madrid: Ibersaf Editores.
- Saleh, W., y T. Gallega Ortega. 2010. *Historias jocosas de Yuha: Selección comentada y bilingüe*. Madrid: Ibersaf y Casa Árabe.
- Samuels, S. Jay. 1979. «The method of repeated readings.» *The Reading teacher* 322.
- Santillán, Paula. 2015. «Nueva taxonomía de las colocaciones en árabe: Aspectos sintácticos, léxicos y semánticos.» *Tesis Doctoral*. Universidad de Granada.

- Schmitt, Norbert, y Diane Schmitt. 2014. «A Reassessment of Frequency and Vocabulary Size in L2 Vocabulary Teaching.» *Language Teaching* 47 (4): 484-503.
- Schmitt, Norbet. 2010. *Researching Vocabulary: A Vocabulary Research Manual*. New York: Palgrave Mcmillan.
- Scott, C. T. 1964. «Literature and the ESL Program.» *The Modern Language Journal* 48 (8): 489-493.
- Shimron, J., y T. Sivan. 1994. «Reading Proficiency and Orthography: Evidence from Hebrew and English.» *Language Learning* 44 (1): 2-27.
- Simensen, Aud Marit. 1987. «Adapted Readers: How are they Adapted?» *Reading in a Foreign Language* 4 (1): 41-57.
- Sinclair, John. 2004. «New evidence, new priorities, new attitudes.» En *How to Use Corpora in Language Teaching*, de John McHardy Sinclair, 271-299. Amsterdam: John Benjamins Publishing Company.
- Smith, F. 1971. *Understanding Reading: A Psycholinguistic Analysis of Reading and Learning to Read*. New York: Holt, Rinehart and Winston.
- Soliman, I., y S. Alwakeel. 2014. *Mastering Arabic through Literature: The Short Story al-Rubaa Volume I*. Cairo: American University in Cairo.
- Somekh, Sasson. 1991. *Genre and Language in Modern Arabic Literature*. Wiesbaden: Harrassowitz.
- Sasson Somekh. 2012. *New Trends in the Language of Contemporary Arabic Literature*. Haifa: The Arabic Language Academy.
- Sorell, C. Joseph. 2013. «Zipf's Law and Vocabulary.» En *The Encyclopedia of Applied Linguistics*, de Carol A. Chapelle. Oxford: Wiley-Blackwell.
- Stakic, Mirjana. 2014. «The Specifics of Children's Literature in the Context of Genre Classification.» *Mediterranean Journal of Social Sciences* 5 (19): 243-251.
- Stetkevych, J. 1966. «Review of the Book A Reader in Modern Literary Arabic, by Ziadeh, F. J.» *Journal of the American Oriental Society* 86 (3): 350-353.

- Stockwell, Peter, y Sara Whiteley. 2014. «Introduction.» En *The Cambridge Handbook of Stylistics*, de Peter Stockwell y Sara Whiteley, 1-10. Cambridge: Cambridge University Press.
- Stoller, Fredricka, Neil J. Anderson, William Grabe, y Reiko Komiyama. 2013. «Instructional Enhancements to Improve Student's Reading Abilities.» *Contemporary Foreign Language Studies* 51 (1): 2-11, 33.
- Stubbs, Michael. 2014. «Quantitative Methods in Literary Linguistics.» En *The Cambridge Handbook of Stylistics*, de Peter Stockwell y Sara Whiteley, 46-62. Cambridge: Cambridge University Press.
- Tannen, D. 1990. *You Just Don't Understand: Women and Men in Conversation*. London: Virago.
- Teng, Feng. 2014 - a. «Incidental Vocabulary Learning by Assessing Frequency of Word Occurrence in a Graded Reader: Love or Money.» *Language Education and Acquisition Research Network (LEARN) Journal* 7 (2): 36-50.
- Teng, Feng. 2014 (b). «Strategies for Teaching and Learning Vocabulary.» *Beyond Words: A Journal on Language Education, Applied Linguistics and Curriculum y Instructions* 41-57.
- Thompson, Paul, y Alison Sealey. 2007. «Through Children's Eyes? Corpus Evidence of the Features of Children's Literature.» *International Journal of Corpus Linguistics* 12 (1): 1-23.
- Thorndike, Edward L. 1921. *The Teacher's Word Book*. New York City: Teachers College, Columbia University.
- Thorndike, Edward L., y Irving Lorge. 1944. *A Teacher's Word Book of 30,000 Words*. New York: Bureau of Publications. Teachers College. Columbia University.
- Tomlinson, Brian. 2013. *Applied Linguistics and Materials Development*. London: Bloomsbury.

- Tomlinson, Brian. 2011. «Glossary of basic terms for materials development in language teaching.» En *Materials Development in Language Teaching*, de Brian Tomlinson, ix-xviii. Cambridge: Cambridge University Press.
- Tomlinson, Brian. 2003. «Introduction: Are Materials Developing?» En *Developing Materials for Language Teaching*, de Brian Tomlinson, 1-11. London: Continuum.
- Tresilian, D. 2008. *Brief Introduction to Modern Arabic Literature*. London: Saqi.
- Urquhart, S., y C. Weir. 1998. *Reading in a Second Language: Process, Product and Practice*. London: Longman.
- Van Mol, Mark. 1999. «Arabic Language and Vocabulary Acquisition.» 434-440.
- Versteegh, Kees. 2006. «History of Arabic Language Teaching.» En *Handbook for Arabic Language Teaching Professionals in the 21st Century*, de K.M. Wahba, Z. A. Taha, England y L., 3-12. New York: Routledge.
- . 2001. *The Arabic Language*. Edinburgh: Edinburgh University Press.
- Waleed, S. 2002. *Curso práctico de lengua árabe II*. Madrid: Ibersaf Editores.
- Wan-a-rom, Udorn. 2008. «Comparing the vocabulary of different graded-reading schemes.» *Reading in a Foreign Language* 20 (1): 43-69.
- Webb, Stuart. 2007. «The Effects of Repetition on Vocabulary Knowledge.» *Applied Linguistics* 28 (1): 46-65.
- West, Michael. 1937. «The "Reading Approach" and "The New Method System".» *The Modern Language Journal* 22 (3): 220-222.
- West, Michael. 1927. «The Construction of Reading Material for Teaching a Foreign Language.» *Dacca University Bulletin* (Oxford University Press) 13: 1-32.
- Wightwick, J., y M. Gaafar. 2011. *Easy Arabic Reader*. McGraw-Hill Books.
- Younes, M. A. 2001. *Tales from Kalila wa Dimna for Students of Arabic*. Ithaca, NY: Spoken Language Services, Inc.

- Zaghouani, W. 2014. «Critical Survey of the Freely Available Arabic Corpora.» *Proceedings of the International Conference on Language Resources and Evaluation (LREC'2014), OSACT Workshop*. Rejkavik, Iceland.
<http://www.qatar.cmu.edu/~wajdiz/corpora.html> .
- Zakariyya, N. 2006. *taarii h ad-da 'wa 'ilaa l-'aammiyya wa 'aa taaruha fii miSr*. El Cairo: daar ad-da' wa l-'islaamiyya li-n-na šr wa-t-tawzii'. (Primera publicación: 1964)
- Zhong, H. 2012. «Multidimensional Vocabulary Knowledge: Development from Receptive to Productive Use.» En *Current Perspectives in Second Language Vocabulary Research*, de D. Hirsh, 23-55. Bern: Peter Lang.

