

SECOND EDITION

a **FREQUENCY** dictionary of
SPANISH

CORE VOCABULARY FOR LEARNERS

**Mark Davies and
Kathy Hayward Davies**

● Practical: the top 5,000 most frequently used Spanish words

● Learner friendly: gives you the core vocabulary for Spanish quickly

● Useful: 30 thematic boxes give the top words for specific topics

A Frequency Dictionary of Spanish

A Frequency Dictionary of Spanish has been fully revised and updated, including over 500 new entries, making it an invaluable resource for students of Spanish.

Based in part on a new web-based corpus containing more than 2 billion words collected from 21 Spanish-speaking countries, the second edition of *A Frequency Dictionary of Spanish* provides the most expansive and up-to-date guidelines on Spanish vocabulary. Each entry is accompanied by an illustrative example and full English translation.

The *Dictionary* provides a rich resource for language teaching and curriculum design, while a separate CD version provides the full text in a tab-delimited format ideally suited for use by corpus and computational linguists.

With entries arranged both by frequency and alphabetically, *A Frequency Dictionary of Spanish* enables students of all levels to get the most out of their study of vocabulary in an engaging and efficient way.

Mark Davies is Professor of Linguistics at Brigham Young University, USA.

Kathy Hayward Davies is Adjunct Professor of Spanish at Brigham Young University, USA.

Routledge Frequency Dictionaries

General Editors

Paul Rayson, *Lancaster University, UK*

Mark Davies, *Brigham Young University, USA*

Editorial Board

Michael Barlow, *University of Auckland, New Zealand*

Sylviane Granger, *Université catholique de Louvain, Belgium*

Barbara Lewandowska-Tomaszczyk, *University of Lodz, Poland*

Josef Schmied, *Chemnitz University of Technology, Germany*

Andrew Wilson, *Lancaster University, UK*

Hongying Tao, *University of California at Los Angeles, USA*

Chris Tribble, *King's College London, UK*

Scott Crossley, *Georgia State University, USA*

Other books in the series

A Frequency Dictionary of Arabic

A Frequency Dictionary of Contemporary American English

A Frequency Dictionary of Czech

A Frequency Dictionary of Dutch

A Frequency Dictionary of French

A Frequency Dictionary of German

A Frequency Dictionary of Japanese

A Frequency Dictionary of Korean

A Frequency Dictionary of Mandarin Chinese

A Frequency Dictionary of Persian

A Frequency Dictionary of Portuguese

A Frequency Dictionary of Russian

A Frequency Dictionary of Spanish (second edition)

A Frequency Dictionary of Turkish

For more titles, please see www.routledge.com/Routledge-Frequency-Dictionaries/book-series/RFD

A Frequency Dictionary of Spanish

Core Vocabulary for Learners

Second Edition

Mark Davies and Kathy Hayward Davies

Second edition published 2018
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

and by Routledge
605 Third Avenue, New York, NY 10017

Routledge is an imprint of the Taylor & Francis Group, an informa business

© 2018 Mark Davies and Kathy Hayward Davies

The right of Mark Davies and Kathy Hayward Davies to be identified as authors of this work has been asserted by them in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

Trademark notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

First edition published by Routledge 2006

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data

A catalog record for this title has been requested

ISBN: 978-1-138-68652-6 (hbk)

ISBN: 978-1-138-68654-0 (pbk)

ISBN: 978-1-315-54263-8 (ebk)

DOI: 10.4324/9781315542638

Supplementary Resources Disclaimer

Additional resources were previously made available for this title on CD. However, as CD has become a less accessible format, all resources have been moved to a more convenient online download option.

You can find these resources available here: <https://resourcecentre.routledge.com/books/9781138686540>

Please note: Where this title mentions the associated disc, please use the downloadable resources instead.

Contents

Thematic vocabulary lists | **vi**

Series preface | **vii**

Acknowledgments | **ix**

Abbreviations | **x**

Preface to the second edition | **xi**

Introduction | **1**

Frequency index | **12**

Alphabetical index | **258**

Part-of-speech index | **310**

Thematic vocabulary lists

- | | |
|---|---|
| 1 Animals 19 | 17 Opposites 142 |
| 2 Body 26 | 18 Nationalities and place adjectives 149 |
| 3 Food 33 | 19 Adjectives with <i>ser/estar</i> 158 |
| 4 Clothing 41 | 20 Adjectives of emotion 166 |
| 5 Transportation 48 | 21 Adjectives—differences across genres 174 |
| 6 Family 56 | 22 Verbs of movement 182 |
| 7 Materials 63 | 23 Verbs of communication 190 |
| 8 Time 71 | 24 Use of the “reflexive marker” <i>se</i> 198 |
| 9 Sports 79 | 25 Preterit/imperfect 206 |
| 10 Natural features and plants 86 | 26 Subjunctive triggers 215 |
| 11 Weather 94 | 27 Verbs—differences across genres 223 |
| 12 Professions 102 | 28 Adverbs—differences across genres 231 |
| 13 Creating nouns 110 | 29 New words since the 1800s 239 |
| 14 Diminutives 118 | 30 Word length (Zipf’s Law) 248 |
| 15 Nouns—differences across genres 126 | |
| 16 Colors 134 | |

Series preface

Frequency information has a central role to play in learning a language. Nation (1990) showed that the 4,000–5,000 most frequent words account for up to 95% of a written text and the 1,000 most frequent words account for 85% of speech. Although Nation's results were only for English, they do provide clear evidence that, when employing frequency as a general guide for vocabulary learning, it is possible to acquire a lexicon which will serve a learner well most of the time. There are two caveats to bear in mind here. First, counting words is not as straightforward as it might seem. Gardner (2007) highlights the problems that multiple word meanings, the presence of multiword items, and grouping words into families or lemmas, have on counting and analysing words. Second, frequency data contained in frequency dictionaries should never act as the only information source to guide a learner. Frequency information is nonetheless a very good starting point, and one which may produce rapid benefits. It therefore seems rational to prioritize learning the words that you are likely to hear and read most often. That is the philosophy behind this series of dictionaries.

Lists of words and their frequencies have long been available for teachers and learners of language. For example, Thorndike (1921, 1932) and Thorndike and Lorge (1944) produced word-frequency books with counts of word occurrences in texts used in the education of American children. Michael West's *General Service List of English Words* (1953) was primarily aimed at foreign learners of English. More recently, with the aid of efficient computer software and very large bodies of language data (called corpora), researchers have been able to provide more sophisticated frequency counts from both written text and transcribed speech. One important feature of the resulting frequencies presented in this series is that they are derived from recently collected language data. The earlier lists for English included samples from, for example, Austen's *Pride and Prejudice* and Defoe's *Robinson Crusoe*, thus they could no longer represent present-day language in any sense.

Frequency data derived from a large representative corpus of a language brings students closer to language as it is used in real life as opposed to textbook language (which often distorts the frequencies of features in a language, see Ljung, 1990). The information in these dictionaries is presented in a number of formats to allow users to access the data in different ways. So, for example, if you would prefer not to simply drill down through the word frequency list, but would rather focus on verbs for example, the part-of-speech index will allow you to focus on just the most frequent verbs. Given that verbs typically account for 20% of all words in a language, this may be a good strategy. Also, a focus on function words may be equally rewarding – 60% of speech in English is composed of a mere 50 function words. The series also provides information of use to the language teacher. The idea that frequency information may have a role to play in syllabus design is not new (see, for example, Sinclair and Renouf, 1988). However, to date it has been difficult for those teaching languages other than English to use frequency information in syllabus design because of a lack of data.

Frequency information should not be studied to the exclusion of other contextual and situational knowledge about language use and we may even doubt the validity of frequency information derived from large corpora. It is interesting to note that Alderson (2007) found

that corpus frequencies may not match a native speaker's intuition about estimates of word frequency and that a set of estimates of word frequencies collected from language experts varied widely. Thus corpus-derived frequencies are still the best current estimate of a word's importance that a learner will come across. Around the time of the construction of the first machine-readable corpora, Halliday (1971: 344) stated that 'a rough indication of frequencies is often just what is needed'. Our aim in this series is to provide as accurate as possible estimates of word frequencies.

Paul Rayson and Mark Davies
Lancaster and Provo, 2008

References

- Alderson, J. C. (2008).** 'Judging the frequency of English words.' *Applied Linguistics*, 28 (3): 383–409.
- Gardner, D. (2007).** 'Validating the construct of Word in applied corpus-based vocabulary research: a critical survey.' *Applied Linguistics*, 28: 241–65.
- Halliday, M. A. K. (1971).** 'Linguistic functions and literary style.' In S. Chatman (ed.) *Style: A Symposium*. Oxford University Press, pp. 330–65.
- Ljung, M. (1990).** *A Study of TEFL Vocabulary*. Almqvist & Wiksell International, Stockholm.
- Nation, I. S. P. (1990).** *Teaching and Learning Vocabulary*. Heinle & Heinle, Boston.
- Sinclair, J. M. and Renouf, A. (1988).** 'A lexical syllabus for language learning.' In R. Carter and M. McCarthy (eds.) *Vocabulary and Language Teaching*. Longman, London, pp. 140–58.
- Thorndike, E. (1921).** *Teacher's Word Book*. Columbia Teachers College, New York.
- Thorndike, E. (1932).** *A Teacher's Word Book of 20,000 Words*. Columbia University Press, New York.
- Thorndike, E. and Lorge, I. (1944).** *The Teacher's Word Book of 30,000 Words*. Columbia University Press, New York.
- West, M. (1953).** *A General Service List of English Words*. Longman, London.

Acknowledgments

We are grateful to Doug Biber, James Jones, and Nicole Tracy from Northern Arizona University, who helped with the part-of-speech tagging and lemmatization for the 20-million-word corpus, as well as Eckhard Bick, who did most of the part-of-speech tagging and lemmatization for the 2-billion-word web-based corpus. We acknowledge the several students at Illinois State University and Brigham Young University, who helped with the sample sentences, revision, and other tasks for the first edition. Finally, we are grateful to SpanishDict for the initial translations of about 4,000 of the 5,000 sample sentences from Spanish to English.

Abbreviations

Meaning

<i>art</i>	article
<i>adj</i>	adjective
<i>adv</i>	adverb
<i>conj</i>	conjunction
<i>f</i>	feminine
<i>+fam</i>	familiar
<i>-fam</i>	formal
<i>interj</i>	interjection
<i>m</i>	masculine
<i>n</i>	neuter
<i>nc</i>	noun: common
<i>nf</i>	noun: feminine
<i>nf (el)</i>	noun: feminine (with el)
<i>nm</i>	noun: masculine
<i>nmf</i>	noun: masc/fem: different meanings
<i>nm/f</i>	noun: masc/fem: msc form given
<i>num</i>	number
<i>obj</i>	object
<i>dir obj</i>	direct object
<i>indir obj</i>	indirect object
<i>pl</i>	plural
<i>prep</i>	preposition
<i>pron</i>	pronoun
<i>sg</i>	singular
<i>subj</i>	subject
<i>v</i>	verb
<i>//</i>	separate speakers in sample sentence

Example

1 el, la *art* the
802 oscuro *adj* dark, obscure
486 apenas *adv* hardly, barely
131 aunque *conj* although, even though
61 la *pron* 3rd person obj f
48 te *pron* you (obj/+fam)
113 usted *pron* you (obj/-fam)
1009 ay *interj* oh no! oh my!
18 lo *pron* 3rd person obj m
111 esto *pron* this (n)
795 estudiante *nc* student
106 casa *nf* house
204 agua *nf (el)* water
230 libro *nm* book
421 order *nmf* sequence (m), religious (f)
513 autor *nm/f* author
1138 doce *num* twelve
48 te *pron* you (obj/+fam)
61 la *pron* (3rd person: dir obj: f)
25 le *pron* (3rd person: indir obj)
4196 lente *nmf* lens (pl) glasses
62 sobre *prep* on top of, over, about
165 nosotros *pron* we (subj)
184 tú *pron* you (subj: sg: +fam)
28 yo *pron* I (subj)
249 sentar *v* to sit (down), seat
3949 espontáneo *adj* spontaneous ¿Fue algo premeditado? // No; fue espontáneo.

Preface to the second edition

The second edition of *A Frequency Dictionary of Spanish* offers several significant additions and improvements from the first edition, which was published in 2006. These include the following:

- 1 *The corpus.* The previous frequency dictionary was based on a corpus of 20 million words from fiction and non-fiction texts, as well as transcripts of spoken language. This has been greatly supplemented in the second corpus by a corpus with the following advantages:
 - a *Size.* The new web-based corpus contains more than 2 billion words of text (in more than 2 million web pages), which is more than 100 times the size of the first corpus.
 - b *Geographical diversity.* The new corpus comes from 21 different Spanish-speaking countries, with tens of millions of words of data for even the smallest countries.
 - c *Recent language.* All of the web pages for the new corpus were collected in 2014–2015, which means that the language is very recent (compared to the corpus for the first edition, which only included texts up to 1999).
 - d *Meaningful language.* The Spanish language that many learners will encounter will be on the web, and so it makes sense to have the web-based corpus be a significant part of the corpus. In addition, nearly 50 percent of this web-based corpus comes from blogs, which means that this is not just stuffy, academic language; rather, it represents “everyday” Spanish as written on tens of thousands of blogs by native speakers of Spanish.
- 2 *New entries.* Because the frequency lists are based on web pages (in addition to spoken, fiction, and non-fiction texts from the first edition), they of course contain many words that are new in this second edition. In total, about 560 of the 5,000 words in the dictionary are new in this edition.
- 3 *Translations.* The previous dictionary included sample sentences in Spanish (from the corpus) for each word, but there was no translation of these sample sentences to English. Because this dictionary is intended primarily for use by language learners (who may not understand some words and phrases in the sample sentences), we have now included translations of all 5,000 sentences.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Introduction

1. The value of a frequency dictionary of Spanish

What is the value of a frequency dictionary for language teachers and learners? Why not simply rely on the vocabulary lists in a course textbook? The short answer is that although a typical textbook provides some thematically related vocabulary in each chapter (foods, illnesses, transportation, clothing, etc.), there is almost never any indication of which of these words the student is most likely to encounter in actual conversation or texts. In fact, sometimes the words are so infrequent in actual texts that the student may never encounter them again in the “real world,” outside of the test for that particular chapter (see Davies and Face, 2006).

While the situation for the classroom learner is sometimes bleak with regard to vocabulary acquisition, it can be equally as frustrating for independent learners. These individuals may pick up a newspaper or go to a web page and begin to work through the text word for word, as they look up unfamiliar words in a dictionary. Yet there is often the uncomfortable suspicion on the part of such learners that their time could be maximized if they could simply begin with the most common words in Spanish, and work progressively through the list.

Finally, frequency dictionaries can be a valuable tool for language teachers. It is often the case that students begin an intermediate language course with deficiencies in terms of their vocabulary. In these cases, the teacher often feels frustrated, because there doesn't seem to be any systematic way to bring less advanced students up to speed. With a frequency dictionary, however, the teacher could assign remedial students to work through the list and fill in gaps in their vocabulary, and they would know that the students are using their time in the most effective way possible.

2. What is in this dictionary?

This frequency dictionary is designed to meet the needs of a wide range of language students and teachers, as well as those who are interested in the

computational processing of Spanish. The main index contains the 5,000 most common words in Spanish, starting with such basic words as *el* and *de*, and quickly progressing through to more intermediate and advanced words. The dictionary is based on the actual frequency of words in more than 2 billion words of text in two different corpora (collections of texts), from many different types of Spanish texts (web pages, fiction, non-fiction, and actual conversations). As a result, the user can feel comfortable that these are words that one is very likely to subsequently encounter in the “real world.”

In addition to providing a list of the most frequent 5,000 words, the entries provide other information that should be of great use to the language learner. Each entry also shows the part-of-speech (noun, verb, etc.), a simple definition of the word in English, an actual example of the word in context (taken from the Corpus del Español, www.corpusdelespanol.org), and a translation of that sample sentence into English (new in this second edition). Finally, the entries show whether the word is more common in spoken, fiction, non-fiction texts, or from the web—so that the learner acquires greater precision in knowing exactly when and where to use the word.

Aside from the main frequency list, there are also indexes that sort the entries by alphabetical order and part-of-speech. The alphabetical index can be of great value to students who want to look up a word from a short story or newspaper article, and see how common the word is in general. The part-of-speech indexes could be of benefit to students who want to focus selectively on verbs, nouns, or some other part-of-speech. Finally, there are a number of thematically related lists and lists related to common grammatical problems for beginning and intermediate students, all of which should enhance the learning experience. The expectation, then, is that this frequency dictionary will significantly maximize the efforts of a wide range of students and teachers who are involved in the acquisition of Spanish vocabulary.

3. Previous frequency dictionaries of Spanish

There have been a number of other frequency dictionaries and lists for Spanish (Buchanan, 1927; Eaton, 1940; Rodríguez Bou, 1952; García Hoz, 1953; Chang-Rodríguez & Juilland, 1964; Alameda & Cueto, 1995; Sebastián, Carreiras, & Cueto, 2000), but all of these suffer from significant limitations. First, all of these frequency dictionaries are based exclusively on written Spanish, and contain no data from the spoken register. Second, five of the dictionaries (Buchanan, 1927; Eaton, 1940; Rodríguez Bou, 1952; García Hoz, 1953; Chang-Rodríguez & Juilland, 1964) are based on texts from the 1950s or earlier, and are now quite outdated. Third, two of the most recent dictionaries both suffer from other important limitations. Alameda and Cueto (1995) only lists exact forms (e.g. *digo*, *dices*, *dijeran*) rather than lemma (e.g. *decir*), and very few of the written texts that it uses are from outside of Spain. The other recent dictionary—Sebastián et al. (2000)—exists only in electronic form and is extremely hard to acquire, especially outside of Spain.

Among the dictionaries just mentioned, most researchers recognize Chang-Rodríguez and Juilland (1964) as the most complete frequency dictionary of Spanish to date. Yet because of its methodological limitations, its list of words is somewhat problematic. As mentioned, all of the texts are from 60–70 years ago (or before), they are nearly all from Spain, and they are all from written texts. In addition, due to limitations in data collection of more than 50 years ago, the corpus is quite small (less than a million words), and is limited just to written texts—spoken Spanish is not represented at all in the word list.

Because of the limitations just mentioned, the vocabulary in Chang-Rodríguez and Juilland is highly skewed. For example, the word *poeta* is word number 309 in their frequency list, with other cases such as *lector* (453), *gloria* (566), *héroe* (601), *marqués* (653), *dama* (696), and *príncipe* (737). This skewing is not limited just to nouns, but also includes words that would be much less frequent in a less skewed corpus, such as *acudir* (498), *figurar* (503), *podar* (1,932) and *malograr* (2,842), and the adjectives *bello* (612), *fecundo* (2,376), and *galán* (2,557).

On the other hand, there are a number of words that we would expect to be in their list that are not. For example, their list of the top 5,000 words of Spanish does not include the following words (the numbers show their placement in our list):

- *Nouns*: equipo 373, película 543, oportunidad 564, control 621, televisión 825, tratamiento 1,003, paciente 1,119, acceso 1,178, marca 1,276, parque 1,354, rama 1,557, experto 1,824
- *Verbs*: enfrentar 897, controlar 994, identificar 1,080, recuperar 1,128, transmitir 1,313, grabar 1,560, investigar 1,571, apretar 1,766, fumar 1,883, distribuir 2,260, fallar 2,355, quebrar 3,113
- *Adjectives*: capaz 645, extraño 736, básico 1,156, temprano 1,578, listo 1,684, latino 1,792, sucio 1,854, japonés 2,019, probable 2,154, feo 2,373, cruel 2,930, moreno 3,304

Thus, while Chang-Rodríguez and Juilland (1964) was quite an achievement for its time, it seems clear that more than 50 years later, it is time for a new frequency dictionary of Spanish, which is based on the more advanced data collection techniques that are now available.

4. The corpus

In order to have an accurate listing of the top 5,000 words in Spanish, the first step is to create robust and representative corpora of Spanish. We have used two different corpora. The first corpus contains 20 million words from the 1900s portion of the genre-oriented Corpus del Español (www.corpusdelespanol.org/hist-gen/), which was created with a grant from the U.S. National Endowment for the Humanities in 2001–2002. These 20 million words of text are about 20 times larger than the corpus used by Chang-Rodríguez and Juilland (1964).

In terms of being representative, the corpus contains a much wider collection of registers and text types than that of any previous frequency dictionary of Spanish. As we see in Table 0.1, two-thirds of the corpus comes from the written register, while a full one-third (6,750,000 words) comes from spoken Spanish.

Table 0.1 Composition of the 20-million-word genre-oriented corpus

	# words	Spain	# words	Latin America
Spoken	1.00	España Oral ¹	2.00	Habla Culta (10 countries)
	0.35	Habla Culta (Madrid, Seville)		
3.35	1.35		2.00	
Transcripts/ plays	1.00	Transcripts/interviews (congresses, press conferences, other)	1.00	Transcripts/interviews (congresses, press conferences, other)
	0.27	Interviews in the newspaper ABC		
	0.40	Plays	0.73	Plays
3.40	1.67		1.73	
Literature	0.06	Novels (BV) ²	1.60	Novels (BV) ²
	0.00	Short stories (BV) ²	0.87	Short stories (BV) ²
	0.19	Three novels (BYU) ³	1.11	Twelve novels (BYU) ³
	2.17	Mostly novels, from LEXESP ⁴	0.18	Four novels from Argentina ⁵
			0.20	Three novels from Chile ⁶
6.38	2.42		3.96	
Texts	1.05	Newspaper ABC	3.00	Newspapers from six different countries
	0.15	Essays in LEXESP ⁴	0.07	Cartas (“letters”) from Argentina ⁵
	2.00	Encarta encyclopedia	0.30	Humanistic texts (e.g. philosophy, history from Argentina) ⁵
			0.30	Humanistic texts (e.g. philosophy, history from Chile) ⁶
6.87	3.20		3.67	
Total	8.64		11.36	

Sources:

- 1 *Corpus oral de referencia de la lengua española contemporánea* (http://elvira.llf.uam.es/docs_es/corpus/corpus.html).
- 2 *The Biblioteca Virtual* (www.cervantesvirtual.com).
- 3 Fifteen recent novels, acquired in electronic form from the Humanities Research Center, Brigham Young University.
- 4 *Léxico informatizado del español* (www.edicionsub.com/coleccion.asp?coleccion=90).
- 5 From the *Corpus lingüístico de referencia de la lengua española en argentina* (www.llf.uam.es/~fmarcos/informes/corpus/coarginl.html).
- 6 From the *Corpus lingüístico de referencia de la lengua española en chile* (www.llf.uam.es/~fmarcos/informes/corpus/cochile.html).

Approximately one-half of the spoken corpus comes from transcriptions of natural speech, including 2,300,000 words in the *Habla Culta* corpus of conversations with speakers from 11 different countries, and 1,000,000 words from the *Corpus Oral de Referencia*, which contains transcripts of conversations, lectures, sermons, sports broadcasts, and many other types of spoken Spanish. The written corpus is divided in half between literature and non-literary texts,

including newspaper articles, essays, encyclopedias, letters, and humanistic texts. In addition to having a good selection of different genres, this corpus is the first to have a good balance of texts from both Latin America and Spain—approximately 43 percent of the texts come from Spain, while 57 percent come from Latin America. In terms of the time period represented, virtually all of the texts are from 1970–2000, with the clear majority being from the 1990s.

Table 0.2 Composition of the 2-billion-word web-based corpus

Country	General (may include blogs)			(Only) blogs			Total		
	Words	Pages	Sites	Words	Pages	Sites	Words	Pages	Sites
Argentina	93,195,550	82,512	6,791	89,509,348	97,682	7,869	182,704,898	180,194	14,660
Bolivia	19,970,008	21,843	2,528	23,372,055	22,253	920	43,342,063	44,096	3,448
Chile	36,316,495	32,691	5,480	34,281,784	39,925	4,249	70,598,279	72,616	9,729
Colombia	84,285,729	84,349	5,064	95,859,929	104,658	5,855	180,145,658	189,007	10,919
Costa Rica	13,354,311	13,424	3,056	18,338,981	20,242	1,799	31,693,292	33,666	4,855
Cuba	35,842,945	21,666	1,209	31,812,745	32,028	677	67,655,690	53,694	1,886
Dominican Republic	18,085,484	20,972	3,032	19,136,866	26,907	1,967	37,222,350	47,879	4,999
Ecuador	24,378,016	27,559	3,217	32,538,210	36,591	2,102	56,916,226	64,150	5,319
El Salvador	16,546,308	18,266	2,102	22,600,727	20,538	2,076	39,147,035	38,804	4,178
Guatemala	23,636,182	25,231	2,751	36,254,690	37,613	1,683	59,890,872	62,844	4,434
Honduras	16,420,245	18,129	1,653	22,512,838	26,109	787	38,933,083	44,238	2,440
Mexico	132,651,925	135,344	6,713	127,946,347	156,877	8,379	260,598,272	292,221	15,092
Nicaragua	15,257,754	15,528	1,506	19,379,274	21,226	635	34,637,028	36,754	2,141
Panama	11,766,575	13,364	2,728	12,932,194	16,355	1,147	24,698,769	29,719	3,875
Paraguay	17,223,991	16,102	2,200	15,766,153	17,791	1,017	32,990,144	33,893	3,217
Peru	47,120,271	53,169	4,262	68,204,165	71,162	4,275	115,324,436	124,331	8,537
Puerto Rico	14,744,742	13,340	2,769	20,911,898	20,946	1,545	35,656,640	34,286	4,314
Spain	208,808,667	209,247	10,475	250,504,154	218,067	15,051	459,312,821	427,314	25,526
Uruguay	17,419,641	15,418	3,390	23,966,953	21,214	2,051	41,386,594	36,632	5,441
USA	94,603,634	78,225	9,044	85,324,555	87,070	23,008	179,928,189	165,295	32,052
Venezuela	41,672,159	45,060	3,331	66,306,730	71,045	3,577	107,978,889	116,105	6,908
TOTAL	983,300,632	961,439	83,301	1,117,460,596	1,166,299	90,669	2,100,761,228	2,127,738	173,970

For this second edition of the dictionary, we have supplemented the 20-million-word corpus with a 2-billion-word corpus of Spanish, which we created with funding from the U.S. National Endowment for the Humanities (NEH) from 2015 to 2016. This new corpus—which is about 100 times larger than the previous corpus—is taken from web pages that were collected in 2014–2015.

This corpus is taken from about 2 million web pages from 21 Spanish-speaking countries. To create the list of URLs for the corpus, we ran high-frequency three-word strings against Google (e.g. *y de los, pero es una*). Because these strings aren’t really related to specific content (e.g. shoes, cars, baseball), Google doesn’t know what to do, and it just returns essentially random URLs, which is exactly what we wanted. We used this same process for each of the 21 countries, for both general web pages as well as for blogs (it was possible to limit searches in Google to just blogs at the time that the corpus was created). The end result is that almost half of the corpus comes from blogs, which ensures that there is a good “informal” character to the corpus, in addition to the more formal texts in the “general” corpus (e.g. product descriptions, news, or things such as Wikipedia pages). In addition to size, one significant advantage of this newer corpus is that it represents very recent Spanish—from just the past two or three years.

Using the data from the web resulted in a frequency list that is somewhat different from the frequency list that was created for the first edition of this book, which was based on just the 20 million words from non-fiction, fiction, and spoken texts. There are about 560 new words in this list, and conversely about 560 words that were there before but have now been dropped from the dictionary (in other words, approximately 10 percent of the words in in this second edition are new).

5. Annotating the data from the corpus

In order to create a useful and accurate listing of the top 5,000 words in Spanish, the corpora needed to be tagged and lemmatized. Tagging means that we assign a part-of-speech to each word in the corpus. In order to do this, we created a lexicon of Spanish, which contained more than 400,000 separate word forms, with their part-of-speech and lemma (where lemma refers to the “base word” or “dictionary headword” to which each individual form belongs). For example, the following are five word forms from the 400,000-word lexicon:

<i>word form / lemma / part-of-speech (pos):</i>	
lápices / lápiz / noun_ masc_pl	pronto / pronto / adverb
tengo / tener / verb_ present_1pers_sg	doscientas / doscientos / number_fem_pl
francesa / francés / adjective_fem_sg	

In cases where there is just one lexicon entry for a given word form, then that form is easy to annotate (e.g. *tengo* = tener / verb: present: 1sg). For many other word forms, however, a given word form has to have more than one entry in the lexicon. For example, *trabajo* “(the) work, I work” can either be [lemma = trabajo, pos = noun_masc_sg] or [lemma = trabajar, pos = verb_present_1sg]. Another example would be *limpia* “clean, 3sg cleans,” which can be either [lemma = limpio, pos = adjective_fem_sg] or [lemma = limpiar, pos = verb_present_3sg]. Such is the case for thousands of different word forms. In these cases, we used rules to tag the text. For example, in the case of *trabajo*, the tagger uses the preceding definite article [*e*] to tag [*el trabajo*] as [lemma = trabajo, pos = noun_masc_sg], whereas it would use the preceding subject pronoun [*yo*] to tag [*yo trabajo*] as [lemma = trabajar, pos = verb_present_1pers_sg].

In many other cases, it is even more difficult than simply using rules to disambiguate the different lemmas and parts of speech of a given word form, and in these cases we have used probabilistic information. For example, one of the most difficult classes of words to tag are past participles (e.g. *dicho, controlado, apagado*). The “rule-based” component of the tagger looks for a preceding form of *haber* “to have” and identifies the word as the form of a verb (e.g. *he [escrito]* “I have written” is [lemma = escribir, pos = verb_pp_masc_sg]). In a case such as [*periódico escrito*], however, *escrito* can either be a past participle of the verb *escribir* (leí el periódico escrito ayer “I read the newspaper (that was) written yesterday”) or it can have a more adjectival-like sense (“the *written* newspaper, as opposed to the electronic newspaper”). In cases such as these, we looked at the total number of cases where the past participle was preceded in the corpus by *ser* (which suggests a passive/verbal reading) or by *estar* (which

suggests a resultative/adjectival reading). If the cases with *ser* were more common with this particular past participle, then ambiguous cases such as [n + past part] (*periódico escrito*) would be marked as passive/verb. The fact that all of the data were stored in a relational database made this type of probabilistic tagging and lemmatization much easier to carry out than may have been possible with linear, word-by-word annotation.

The tagging for the 20-million-word genre-oriented corpus was carried out at Northern Arizona University under the direction of Prof. Douglas Biber in 2003–2005 as part of a grant that we had received from the U.S. National Science Foundation to analyze syntactic variation in Spanish. The tagging of the 2-billion-word web-based corpus was done by Eckhard Bick in 2015 as part of the grant from the U.S. National Endowment for the Humanities. In both cases, this preliminary tagged and lemmatized information was input into a MS SQL Server database, where the data was cleaned up using the rule-based and probabilistically-based procedures mentioned above.

We have not carried out formal tests to determine the accuracy of the part-of-speech tagging and lemmatization, but we have examined the annotation in detail at many different stages of the project. After the preliminary tagging, we determined which word forms belonged to two or more lemmas that were within the 20,000 most frequent lemmas in the corpus (i.e. *limpia* or *trabajo*, as mentioned above). For each one of these forms, we examined the collocations (words to the left and right) to make sure that we had annotated these forms correctly, and made any necessary adjustments. Later, we went through each of the 6,000 most frequent lemmas, and again looked for any form for any of these lemmas that also appeared as a member of another lemmas, and again checked the collocations and made the appropriate adjustments. Finally, we continually compared our list to that in Chang-Rodríguez and Juilland, and carefully examined all of the forms of any word that was in our list but was not in Chang-Rodríguez and Juilland, or any word that was in their list but was not in our top 5,000 words.

6. Organizing and categorizing the data

Even after annotating the corpus for part-of-speech and lemma—as described in the previous section—there remained a number of difficult decisions

regarding how the lemma should be grouped together. In most cases, we have followed the parts of speech from Chang-Rodríguez and Juilland (1964). In some cases, however, we have conflated categories that Chang-Rodríguez and Juilland kept distinct. The three primary areas of difference are the following:

(a) Noun/adjective

In many cases, there is only minor syntactic and semantic difference between nouns and adjectives in Spanish, as in the case of *ella es católica* “she is (a) Catholic.” This holds true not only for religions and nationalities (*él es ruso / italiano* “he is (a) Russian / (an) Italian”), but also cases such as *los ricos no ayudan a los pobres* “the rich don’t help the poor” or *los últimos recibieron más que los primeros* “the ones who came last got more than those who came early.” In most cases, these were assigned a final part-of-speech of [adjective], and learners can easily apply this information to these cases where there is a more nominal sense.

(b) Past participle

It is often very hard to disambiguate between the [passive/verbal] and [adjectival/resultative] senses of the past participle, as shown above with the example of *periódico escrito*. One solution would be to simply include all past participles as part of the verbal lemma, so that *organizado* is listed with *organizar*, *descrito* is listed with *describir*, and so on. Yet there are other cases where the past participle has a clearly adjectival sense, as in *los niños cansados* “the tired children,” *un libro pesado* “a heavy book,” or *unos casos complicados* “some complicated cases.” Our approach has been to manually check each of the adjective entries in the dictionary, which have the form of a past participle. When the majority of the occurrences of this initially tagged form have a strongly agentive reading, then that past participle would be reassigned to the verbal lemma.

(c) Determiner/pronoun/adjective/adverb

Many frequency lists and dictionaries create fine-grained distinctions between these categories, which may be of minimal use to language learners. For example, some frequency lists and dictionaries distinguish between determiner and adjective. Yet it is probably impossible to say where the category [determiner] ends and [adjective] starts, as in cases

such as *varios*, *algunos*, *cuyos* “several, some, whose.” As a result, we follow the lead of Chang-Rodríguez and Juilland, and assign all determiners (except the articles *el* and *la*) to the category [adjective].

Yet we also depart from Chang-Rodríguez and Juilland on a number of points, primarily with regard to the categorization of pronouns, adjectives, and adverbs. For example, they distinguish between the adjectival use of *temprano* “early” (*fue un verano temprano* “it was an early summer”) and the adverbial use (*el verano llegó temprano* “summer arrived early”). While they list the word twice in the dictionary, we assume that a learner can easily apply the meaning to both cases, and simply list it once under [adjective]. Similarly, Chang-Rodríguez and Juilland distinguish between the adjectival use of *todo* “all/every” (*están todos los hombres* “all the men are here”) and the putative pronominal use (*están todos* “everyone is here”), whereas we list *todo* just once—again as an [adjective]. In fact, with an atomistic division of part-of-speech categories, the same word can theoretically span three different parts of speech—noun, adjective, and adverb—and the question is whether to list them all separately in the dictionary. For example, Chang-Rodríguez and Juilland list *menos* “less/least” three times in the dictionary—as noun (*había menos de lo que queríamos* “there was less than we wanted”), adjective (*había menos dinero del que queríamos* “there was less money than we wanted”), and adverb (*cobraron menos que nosotros* “they charged less than us”). In our dictionary, we assume that the

learner can easily apply the one meaning to the three contexts, and we accordingly conflate the three uses to the [adjective] category. Finally, we group together the masculine and feminine forms of the definite article (*el/la*), as both we and Chang-Rodríguez and Juilland have done for all other determiners (*ese/esa, otro/otra*, etc.).

Finally, we should note that there is one category of words with which we separate more lemmas than is typically done in other frequency dictionaries. Other dictionaries will often include all of the forms of a pronoun under the masculine/singular/subjective case form of the pronoun. For example, Chang-Rodríguez and Juilland group together under the one entry *yo* “I” the following pronouns: *me* “me,” *nos* “us,” *nosotros* “we,” *le/les* “to 3sg/3pl,” and even *se* (the “reflexive marker” in Spanish). Because they are morphologically distinct forms that would not be readily recognized as forms that are related to *yo*, we include them (and similar pronouns) as their own entries.

7. Range, frequency, and weighting

At this point, each of the 2 billion words of text had been assigned to a lemma and part-of-speech, and with some lemmas these categories were conflated, as discussed in the previous section. The final step was to determine exactly which of these words would be included in the final list of 5,000 words. One approach would be to simply use frequency counts. For example, all lemmas that occur *n* times or more in the corpus would be included in the dictionary. Imagine, however, a case where a particular scientific

Good dispersion					Poor dispersion				
Freq	Spanish	POS	English	Disp	Freq	Spanish	POS	English	Disp
400	debilidad	n	weakness	0.87	414	sandinista	n	Sandinista	0.10
415	requerir	v	to require	0.85	434	gasoducto	n	pipeline	0.12
415	inevitable	j	unavoidable	0.85	417	califato	n	caliphate	0.22
415	humilde	j	humble, modest	0.85	415	visigodo	n	visigothic	0.23
415	doloroso	j	painful	0.85	412	superintendencia	n	supervision	0.29
416	ajustar	v	to adjust, tighten	0.84	423	glucosa	n	glucose	0.30
410	iluminar	v	to illuminate	0.84	418	génico	adj	gene	0.30
411	sembrar	v	to sow, plant	0.84	405	anglicano	n	Anglican	0.31
411	desviar	v	to deviate, change	0.84	440	ultraviolet	adj	ultraviolet	0.32
410	choque	n	crash, collision	0.83	423	precolumbino	adj	pre-Columbian	0.38

+ range in spoken					+ range in non-fiction				
Spanish	POS	English	Spok	NF	Spanish	POS	English	Spok	NF
hola	i	hello, hi	102.3	2.1	monarca	n	monarch	0.5	46.1
tuyo	p	yours (sg/+fam)	57.8	2.6	hidrógeno	n	hydrogen	0.8	34.1
horrible	j	horrible	39.8	1.9	yacimiento	n	deposit, site	1.3	47.4
platicar	v	talk, chat	34.0	1.7	petrolero	j	oil, petroleum	0.8	28.4
señorita	n	young woman	43.0	4.4	trono	n	throne	1.6	45.3
contento	j	happy, content	54.8	5.7	bacteria	n	germ, bacteria	1.6	41.5
inclusive	r	including, even	85.4	9.4	costero	j	coastal	1.5	37.0
tonto	j	stupid, dumb	83.2	9.3	canciller	n	chancellor	2.0	37.6
cariño	n	love, affection	44.3	5.9	oscilar	v	oscillate	1.9	35.5
feo	j	ugly, nasty	28.7	5.0	órbita	n	orbit	2.2	37.5

term was used repeatedly in 11 of the 2 million web pages, but appears almost nowhere else in the corpus. Alternatively, suppose that a given word is spread throughout an entire register (spoken, fiction, non-fiction, or web), but that it is still limited almost exclusively to that register. Should the word still be included in the frequency dictionary? The argument could be made that we should look at more than just raw frequency counts in cases such as this, and that we ought to include some measure of how well the word is “spread across” all of the registers in the entire corpus.

As a clear example of the contrast between “frequency” and “dispersion,” consider the following table. The words to the left have a Juilland d “dispersion” of more than 0.80, meaning that the word is spread fairly evenly across the corpus. The words to the right, on the other hand, have much worse dispersion, and occur in just a few sections of the corpus (e.g. only a few academic or web domains). Most would easily agree that the words shown on the left would be more useful in a frequency dictionary, because they represent a wide range of texts and text types in the corpus.

A second issue deals with the relative weights assigned to the four main registers—spoken, fiction, non-fiction, and web pages. Is one register more “important” in terms of how well it represents what we perceive to be the most “useful” variety of Spanish? Consider the following table, which shows words with a frequency of between 400 and 600 in the 20-million-word “genre-based” corpus. The words to the left are much more frequent in spoken than in non-fiction texts (newspapers and academic

texts) (the numbers refer to the frequency in each genre times in the dispersion for that genre; see Section 8 below). It seems fairly uncontroversial that the “spoken” list on the left represents more basic vocabulary, and so we would argue that a higher weight should be given to words that occur more in the spoken register than in the non-fiction register.

Consider also the types of words that we get from the 2 billion words of data from the web. The following table shows some of the words that are much more common on the web than in spoken, fiction, and non-fiction texts. Note how many deal with technology, sports, politics, and science, which are things that one would expect to find lots of on the web (note that in this table we have only included nouns and adjectives).

+Web
(noun) blog, usuario, web, video, Internet, foro, navegador, gol, liga, torneo, parlamento, euro, ciudadanía, infraestructura, distrito (adjective) digital, virtual, integral, federal, legislativo, gubernamental, ambiental, monetario

After comparing the four different genres, we decided to assign the following weightings: 20 percent of the score comes from spoken, 30 percent from fiction, 20 percent from non-fiction (newspapers and academic texts), and 30 percent from the web corpus.

8. The final calculation

After looking at the issue of frequency, dispersion, and the weighting for the different registers, we created the following formula:

$$x = .3 * (Fic_{fpm} * Fic_{disp}) + .3 * (Web_{fpm} * Web_{disp}) + .2 * (NF_{fpm} * NF_{disp}) + .2 * (Spok_{fpm} * Spok_{disp})$$

where:

Fic_{fpm} , Web_{fpm} , NF_{fpm} , $Spok_{fpm}$ is the frequency (per million words) in the four genres of fiction, web pages, non-fiction (newspaper and academic), and spoken.

Fic_{disp} , Web_{disp} , NF_{disp} , $Spok_{disp}$ is the Juillard d dispersion in the four genres of fiction, web pages, non-fiction (newspaper and academic), and spoken.

As a concrete example, let's take the word *digno* "worthy":

$$\begin{aligned} 34.31 = & .3 * ((93 / 4.604) * 0.882) + \\ & .3 * ((90,566 / 1,985.133) * 0.993) + \\ & .2 * ((292 / 9.755) * 0.894) + \\ & .2 * ((73 / 3.352) * 0.784) \end{aligned}$$

The word *digno* occurs 93 times in fiction, 90,566 in web pages, 292 times in newspapers and academic, and 73 times in spoken. In each case, the raw frequency is divided by the size of that genre (e.g. 4.604 for fiction, because there are about 4,604,000 words in fiction), and this gives us the frequency per million words. The fact that we use the corpus size to "normalize" these numbers is why we can use genres from the 20-million-word genre-based corpus as well as the much larger 2-billion-word web-based corpus, because in the latter case the raw frequency is divided by a much larger number. For each of the four genres, we then multiply the normalized frequency (per million words) by the dispersion for that genre (e.g. 0.882 for fiction). Finally, we assign a "weighting" to each genre (either 0.2 or 0.3).

It is this figure of [34.31] for *digno* that represents its score and determines whether the word is included in the dictionary. We simply take the top 5,000 scores, and these words are the ones that are included in the dictionary.

While the actual formula may seem complicated, hopefully the general criteria for the inclusion of a word in the dictionary are somewhat easier to understand. First, weighting is given to all four registers—spoken, fiction, non-fiction, and the web—and it is unlikely that a word will be included if it is common in only one of these four registers. Second, the score for any genre is a function of both frequency and dispersion. In other words, a word must not only occur many times in the

corpus; it must also be "spread out" well throughout the entire corpus. Third, there is a slight weighting advantage given to the fiction and web genres, although the final weighting is still relatively equal across the four genres in the corpus.

9. The main frequency index

The second chapter contains the main index in this dictionary—a rank-ordered listing of the top 5,000 words (lemmas) in Spanish, starting with the most frequent word (the definite article *el, la*) and progressing through to *armario* "closet, cupboard," which is number 5,000. The following information is given for each entry:

rank frequency (1, 2, 3, . . .), headword, part-of-speech, English equivalent

- sample sentence, translation to English
raw frequency for fiction + spoken + non-fiction, raw frequency from the web corpus, indication of major register variation

As a concrete example, let us look at the entry for *colchón* "mattress"

4240 *colchón nm* mattress

- le ofrecemos una cama en el suelo y un colchón viejo – we offer you a bed on the floor and an old mattress
212 10915 +f

This entry shows that word number 4,240 in our rank order list is *colchón*, which is a masculine noun [nm] that can be translated as "mattress" in English. We then see an actual sentence or phrase that shows the word in context, as well as the translation of that sentence to English. The two following numbers show that the word occurs 212 times in non-fiction, fiction, and spoken from the 20-million-word genre-based corpus, and 10,915 times in the 2-billion-word web-based corpus. Finally, the [+f] indicates that the word is much more common in the fiction register than would otherwise be expected.

Let us briefly add some additional notes to the explanation just given.

- (a) *The part-of-speech.* Remember that some categories have been conflated, such as noun/adjective with religions and nationalities

(*católico, americano*) or adjective/pronoun (*todos*). With nouns, there are several different markings for gender. Most nouns are either *nm* (masculine; *año, libro*) or *nf* (feminine; *tierra, situación*). Nouns that are feminine but are preceded by the articles *el* and *un* are marked *nm (el)* (*agua, alma*), while nouns that have the same form for masculine or feminine are marked *nc* (*joven, artista*). In most cases, professions and other nouns relating to people are marked *nm/f* (*maestro, autor*), which means that only the masculine form appears in the dictionary, but the frequency statistics have been grouped together with a possible feminine form (*maestra, autora*). Finally, a few nouns have both masculine and feminine forms (*nmf*), but these have different meanings (*order* = sequence (*m*), religious (*f*); *radio* = radio set (*m*), means of communication (*f*)).

- (b) *English equivalent*. Only the most basic translations for the word are given. This is not a bilingual dictionary, which lists all possible meanings of a given word, and intermediate to advanced users will certainly want to consult such a dictionary for additional meanings. Also note that high-frequency phrases in which a given word occurs are not given, except when the vast majority of all occurrences of that word occur within such a phrase. There are a handful of such words in the dictionary, and they are marked as such (e.g. 203 *sin [embargo]*, 270 *a [través]*, 663 *a [medida] que*, 958 *a [menudo]*, 1233 *no [obstante]*, etc.). Finally, in most cases, we have not given the special senses that the word acquires when used pronominally (i.e. with *se*), although this is noted in a handful of cases where a very high percentage of the occurrences are with *se*, as in *rendirse* “to give in” or *colarse* “to slip in.”
- (c) *Phrase in context*. All of these phrases and sentences come from the Corpus del Español (www.corpusdelespanol.org). The goal has been to choose phrases whose meaning reflects well the basic meaning of the word with the minimal number of words, and this has been more possible in some cases than in others. With “invented” sentences, it would have certainly been possible to have concise sentences that express the core meaning very clearly, but this

would have been at the expense of less authentic examples. Finally, in some cases, the original sentence has been shortened by taking out some words whose absence does not affect the basic meaning of the phrase as a whole. Note that in this second edition, all of the 5,000 sentences in Spanish are also translated to English.

- (d) *Register variation*. The symbols [o f nf w] show that the word in question has a high score (a combination of frequency and range) in the indicated register (oral, fiction, non-fiction, web). These symbols appear only when the word is in the top 300 words in that register, in terms of its relative frequency to the overall frequency of the word in the entire corpus.

10. Thematic vocabulary (“call-out boxes”)

Placed throughout the main frequency-based index are approximately 30 “call-out boxes” that serve to display in one list a number of thematically related words. These include lists of words related to the body, food, family, weather, professions, nationalities, colors, emotions, verbs of movement and communication, and several other semantic domains. In addition, however, we have focused on several topics in Spanish grammar that are often difficult for beginning and intermediate students. For example, there are lists that show the most common diminutives, superlatives, and derivational suffixes to form nouns, the most common verbs and adjectives that take the subjunctive, which verbs most often take the “reflexive marker” *se*, which verbs most often occur almost exclusively in the imperfect and preterit, and which adjectives occur almost exclusively with the two copular verbs *ser* and *estar*. Finally, there are even more advanced lists that compare the use of nouns, verbs, adjectives, and adverbs across registers, and show which words are used primarily in spoken, fiction, and non-fiction texts, as well as from the 2-billion-word corpus from the web. Related to this is a list showing which are the most frequent words that have entered the language in the past 100–200 years. Note that in most cases, the words in the thematic lists are based on just the 20-million-word genre-oriented corpus (as in the first edition of the dictionary), although the rank order numbers (e.g. that a word is the 3,837th most common word in the dictionary) are based on the entire corpus—including the 2 billion words of data from the web.

11. Alphabetical and part-of-speech indexes

The third chapter contains an alphabetical listing of all words listed in the frequency index. Each entry includes the following information: (1) lemma; (2) part-of-speech; (3) a basic English translation; and (4) rank order frequency. The fourth chapter contains “part-of-speech” listings of the 5,000 words in the frequency and alphabetical indexes. Within each of the categories (noun, verb, adjective, etc.), the lemmas are listed in order of descending frequency. Because each entry is linked to the other two indexes via the rank frequency number, each of the entries in this index contains only the rank frequency and lemma.

12. More data

Note that if you are interested in even more frequency data from Spanish—beyond the 5,000 words in this dictionary—you might want to look at the data from www.wordfrequency.info/spanish. That site contains the top 40,000 words from Spanish—or about eight times more than what is in this book. In addition, you can browse all of these words at www.wordandphrase.info/span/. For each word, you can see on one page the definition, several synonyms, the top 20–30 collocates (nearby words, which give great insight into meaning and usage), 20–30 related topics from the 2-billion-word corpus, and 200 sortable concordance lines—all with many links to other related words.

References

- Alameda, J.R. & Cuetos, F. (1995)** *Diccionario de Frecuencias de las Unidades Lingüísticas del Castellano*. Oviedo: Universidad de Oviedo.
- Buchanan, M.A. (1927)** *A Graded Spanish Word Book*. Toronto: University of Toronto Press.
- Chang-Rodríguez, E. & Juilland, A. (1964)** *Frequency Dictionary of Spanish Words*. The Hague: Mouton.
- Davies, M. and Face, T.L. (2006)** “Vocabulary Coverage in Spanish Textbooks: How Representative Is It?” In *Selected Proceedings from the Conference on the Acquisition of Spanish and Portuguese as First and Second Languages*, ed. Jacqueline Toribio. Cascadilla. 132–143.
- Eaton, H. (1940)** *An English – French – German – Spanish Word Frequency Dictionary*. New York: Dover Publications.
- García Hoz, V. (1953)** *Vocabulario Usual, Vocabulario Común y Vocabulario Fundamental*. Madrid: CSIC.
- Rodríguez Bou, L. (1952)** *Recuento de Vocabulario Español*. Río Piedras: Universidad de Puerto Rico.
- Sebastián, N., Martí, M.A., Carreiras, M.F., & Cuetos, F. (2000)** *LEXESP, Léxico Informatizado del Español*. Barcelona: Ediciones de la Universitat de Barcelona. (CD-ROM only)

Frequency index

ID L1 *c1* defin

- span — eng
freq_gen | freq_web | genres

1 el, la *art* the (+m, f)

- el diccionario tenía también frases útiles – the dictionary also had useful phrases
2055835 | 201481381

2 de *prep* of, from

- es el hijo de un amigo mío – he is the son of a friend of mine
1104364 | 133341468

3 que *conj* that, which

- dice que no quiere estudiar – he says that he doesn't want to study
543408 | 70319525

4 y *conj* and

- saben leer y escribir – they know how to read and write
524574 | 56511956

5 en *prep* in, on

- vivo en el segundo piso – I live on the second floor
496295 | 49340069

6 un *art* a, an

- era un hombre simpático – he was a nice man
381976 | 37389547

7 ser *v* to be (norm)

- la primera persona narrativa puede ser por el protagonista – the first-person narrative can be through the protagonist
324390 | 40017214

8 a *prep* to, at

- me voy al extranjero por el verano – I am going abroad for the summer
328540 | 48746618

9 él *pron* he, [ellos] them (m)

- él es bastante simpático – he is very nice
428383 | 2743194

10 lo *art* the (+neut)

- lo mejor es estudiar mucho – the best thing is to study a lot
346026 | 31016059

11 no *adv* no

- no tiene dinero – he has no money
198473 | 21970515 +o

12 su *adj* his/her/their/your (–fam)

- ¿Quién era? ¿Su hermana? ¿Su amiga? – Who was it? His sister? Her friend?
173372 | 17883988

13 haber *v* to have (+Ved)

- no ha dicho nada – he hasn't said a thing
155853 | 15650271

14 con *prep* with

- hay un hombre con ella – there is a man with her
163958 | 18316135

15 por *prep* by, for, through

- caminamos por el parque – we walked through the park
174450 | 17512892

16 para *prep* for, to, in order to

- ¡Tengo una sorpresa sensacional para ti! – I have a wonderful surprise for you!
112694 | 16947562

17 mí *pron* me (obj *prep*)

- el regalo era para mí – the gift was for me
114124 | 608449 +o

18 lo *pron* [3rd person] (dir obj-m)

- lo compré en la tienda – I bought it at the store
101112 | 5126699 +o

19 tener *v* to have

- Mi padre tiene varios relojes antiguos. – My father has several antique watches.
85865 | 11087983 +o

20 como *conj* like, as

- ser compositor en España es como ser torero en Finlandia – being a composer in Spain is like being a bullfighter in Finland
98660 | 12480595

21 estar *v* to be (location, change from norm)

- él está en el trabajo – he is at work
88421 | 9574000

22 me *pron* me (obj)

- ¿Cuándo me va a llamar? – When is he going to call me?
64024 | 8211349 +o

23 más *adj* more

- él necesitaba más dinero – he needed more money
86998 | 9003898

24 este *adj* this (m) [esta (f)]

- es igual en el este como en el oeste – it is the same in the east as in the west
73920 | 13208363

25 le *pron* [3rd person] (indir obj)

- nunca le dijo la verdad – she never told him the truth
64833 | 7787199

26 hacer *v* to do, make

- he podido hacer lo que me gusta – I have been able to do what I like
64199 | 8808576

27 se *pron* [“reflexive” marker]

- el saber de los geógrafos griegos se difundió por Europa – Greek knowledge of geography spread throughout Europe
60393 | 28984360 +w

28 yo *pron* I (subj)

- ¡Yo soy el padre! – I am the father!
45822 | 10894793 +o +w

29 o *conj* or

- sí esperaba uno o dos muertos – yes, he expected one or two deaths
71054 | 9100644

30 pero *conj* but, yet, except

- no significa nada para mí, pero no puedo olvidarla – it means nothing to me, but I can't forget it
66748 | 6871867

31 decir *v* to tell, say

- Parece que dice la verdad. – It seems like he is telling the truth.
61678 | 6016427 +o

32 poder *v* to be able to, can

- Con estos datos se puede tener una primera idea del mecanismo. – With these facts we can begin to understand the mechanism
61177 | 9248731

33 ir *v* to go

- ¿Quieres ir a la playa este fin de semana? – Do you want to go to the beach this weekend?
52054 | 4435468 +o

34 ese *adj* that (m) [esa (f)]

- ¿Dónde viven esas mujeres? – Where do these women live?
48535 | 4683205 +o

35 otro *adj* other, another

- No comparte su taxi nunca con otros viajeros a menos que les conoce. – Don't share your taxi with other travelers unless you know them.
54762 | 5431865

36 si *conj* if, whether

- si quiere cazar, vamos a cazar – if he wants to hunt, let's go hunting
38328 | 6519137

37 mi *adj* my

- mi casa es su casa – my house is your house
34310 | 4766226

38 ver *v* to see

- había que subir para ver las ruinas – you have to climb up to see the ruins
39015 | 4122615

39 ya *adv* already, still

- su marido ya ha dicho todo – her husband has already said it all
37257 | 3961260 +o

40 porque *conj* because

- lo vendo sólo porque tengo un problema muy grande – I am only selling it because I am in a predicament
34109 | 3154550 +o

41 mucho *adj* much, many, a lot (ADV)

- por lo visto tienen mucho dinero – apparently they have a lot of money
35069 | 4459289 +o

42 dar *v* to give

- me dio esta carta para ud – he gave me this letter for you
32999 | 3900029 +o

43 muy *adv* very, really

- está muy contento con mi trabajo – she is very happy with my work
32854 | 3217244 +o

44 saber *v* to know (a fact), find out

- se necesita humildad para saber reconocer las propias faltas – you need humility to know how to recognize your own faults
27238 | 2642079 +o

45 sí *adv* yes

- quiero una respuesta concreta: sí o no – I need a solid answer: yes or no
31716 | 1100318 +o

46 año *nm* year

- no lo supo hasta casi un año después – he didn't find out until almost a year later
37168 | 3792004

47 ti *pron* you (obj prep-sg/+fam)

- no, la carta no es para ti – no, the letter is not for you
29283 | 366540 +o

48 te *pron* you (obj/+fam)

- ¿No te han hablado? – They haven't spoken with you?
19928 | 4017417 +o

49 también *adv* also

- también habla italiano – she speaks Italian as well
31172 | 3209504

50 qué *pron* what?, which?, how (+ADJ)!

- no sé qué voy a hacer – I don't know what I'm going to do
29722 | 2317175 +o

51 alguno *adj* some, a few

- habló algunas palabras con el agente de negocios – he spoke a few words with the business agent
27224 | 3037373

52 nos *pron* us (obj)

- nos vio en la calle – he saw us on the street
17300 | 4445709 +o +w

53 tu *adj* your (+fam)

- ésta es tu casa y ésta es tu cama – this is your house and this is your bed
8068 | 7260841 +w

54 sin *prep* without

- se habían quedado sin dinero – they were left without money
25332 | 2499874

55 mismo *adj* same

- pronunciando el mismo discurso en siete idiomas – giving the same speech in seven languages
23973 | 3339470

56 eso *pron* that (n)

- y eso no es todo – and that isn't everything
20972 | 2339488 +o

57 cuando *adv* when

- aquel libro es de mis nietos, cuando eran bebés todavía – that book belonged to my grandchildren, when they were still babies
32690 | 71759

58 querer *v* to want, love

- quiero que este proceso salga con limpieza – I want this process to go cleanly
20091 | 2897092

59 vez *nf* time (specific occurrence); en v. de: instead of

- es la primera vez que estoy junto al mar – this is the first time that I have been near to the sea
27629 | 1558559

60 hasta *prep* until, up to, even (ADV)

- toda la noche, hasta las tres de la mañana – all night long, until three in the morning
27638 | 2258791

61 la *pron* [3rd person] (dir obj-f)

- la puso en su bolsillo – he put it in his pocket
8190 | 7941346 +o +w

62 sobre *prep* on top of, over, about

- dejó el papel sobre la mesa y se fue – he left the paper on top of the table and left
26590 | 3030210

63 entre *prep* between, among

- la cosa es entre tú y yo – the matter is between you and me
32379 | 2607736

64 dos *num* two

- la familia se reúne cada dos años – the family gets together every two years
27487 | 2017874

65 día *nm* day

- cada día hay más problemas – every day there are more problems
20735 | 2577473

66 grande *adj* large, great, big

- tiene los grandes ojos negros y bonitos – she has the biggest and most beautiful black eyes
29813 | 2030849

67 así *adv* like that

- la vida es así – life is like that
21331 | 2321804

68 pasar *v* to pass, spend (time)

- no ha pasado ni un bus – not a single bus has passed by
20946 | 1905621

69 cosa *nf* thing

- estoy ya interesado en otra cosa – I am interested in a different matter
17977 | 1903169 +o

70 desde *prep* from, since

- Lo había pensado desde el principio que era una mal idea. – I thought from the beginning that this was a bad idea.
24743 | 2424665

- 71 deber** *v* should, ought to; to owe
- la cosa está en el lugar donde debe estar – the thing is in the place where it should be
19965 | 3248053
- 72 ella** *pron* she, [ellas] them (f)
- ella es muy estudiosa – she is very studious
18720 | 1491405
- 73 pues** *conj* then, well then
- pues, venga usted cuando quiera – well then, come whenever you want
18535 | 1210256 +o
- 74 entonces** *adv* so, then
- vamos entonces a cambiarlo – then we are going to change it
19759 | 888406 +o
- 75 llegar** *v* to arrive
- por fin llegamos al fondo de la catarata – we finally arrived at the bottom of the waterfall
18877 | 1848469
- 76 poco** *adj* little, few, a little bit (*adv*)
- trabajó poco tiempo con él – he worked a little while with him
18209 | 1845555
- 77 nuestro** *adj* our
- no servía para nuestro país porque somos diferentes – it was no good for our country because we're different
13411 | 3496070 +w
- 78 bien** *adv* well
- el número de bienes que pueda poseer en un momento dado – the number of goods one possesses in a given moment
16402 | 1959560 +o
- 79 ni** *conj* not even, neither, nor
- pero ni eso me tranquilizó – but not even that reassured me
16340 | 1927472
- 80 tiempo** *nm* time (general)
- ha estado mucho tiempo con ella – he has been with her a long time
18408 | 1921017
- 81 ahora** *adv* now
- lo importante ahora es contarte lo que le pasó – the important thing now is to tell you what happened
17072 | 1591817 +o
- 82 primero** *adj* first
- Primero, vamos a ir al supermercado. – First, we are going to go to the grocery store.
20758 | 2484553
- 83 creer** *v* to believe, think
- creo en la justicia de Dios – I believe in the justice of God
15928 | 1685972 +o
- 84 donde** *adv* where
- no sé donde está la llave – I don't know where the key is
17601 | 2151872
- 85 vida** *nf* life
- ha dedicado toda la vida a la música – she has dedicated her life to music
15149 | 2424426
- 86 dejar** *v* to let, leave
- ella no dejó que yo lo olvidara – she didn't let me forget it
14241 | 1881239
- 87 nada** *pron* nothing, (not) at all
- no hay nada que hacer – there is nothing to do
14024 | 1498787 +o
- 88 tanto** *adj* so much, so many
- no podía creer que haya tanta gente junta – I couldn't believe there were so many people together
16618 | 1960237
- 89 parecer** *v* to seem, look like
- Mi burra es tan flaca parece un esqueleto. – My donkey is so thin he looks like a skeleton.
15096 | 1330117
- 90 hablar** *v* to speak, talk
- algunos hablan quechua, español, e inglés – some of them speak Quechua, Spanish, and English
14325 | 1432852 +o
- 91 poner** *v* to put (on), get (+ADJ)
- puso el cuchillo en la mano – he put the knife in his hand
14619 | 1657298
- 92 parte** *nf* part, portion
- el arte es una parte tan importante de la cultura – art is such an important part of culture
20938 | 1744336
- 93 eh** *interj* eh
- Te interesa esta bufanda, ¿eh? – You are interested in this scarf, eh?
16011 | 33653 +o
- 94 nuevo** *adj* new
- marcaba el comienzo de una nueva vida para mí – it marked the beginning of a new life for me
19641 | 1964400

95 sólo *adv* only, just

- no siempre vestía ropa de mujer; sólo en carnaval – I don't always wear women's clothes; only during carnival
15824 | 2043480

96 siempre *adv* always, forever

- siempre ha sido así – she has always been like that
13141 | 1671829

97 hombre *nm* man, mankind, husband

- yo soy un hombre de pocas necesidades – I am a man of few needs
14773 | 1240825

98 bueno *adv* well . . .

- Bueno, ¿y ahora qué hacemos? – Well, so what are we going to do now?
15729 | 1108838 +o

99 seguir *v* to follow, keep on

- no la puedo seguir porque habla demasiado rápido – I couldn't keep up because she was speaking so quickly
14081 | 1824937

100 quedar *v* to remain, stay

- El aire empezó a quedarse quieto. – The air began to remain still.
14040 | 1348837

101 llevar *v* to take, carry

- colgado al hombro llevaba un bolso – the man carried a bag over his shoulder
14452 | 1607827

102 encontrar *v* to find

- encontró un lugar apropiado para instalar la sucursal – he found a suitable place to set up the office
14919 | 1855163

103 bueno *adj* good

- Bueno, ¿y ahora qué hacemos? – Well, so what are we going to do now?
11348 | 2387499

104 tan *adv* such, as, too, so

- ¡Tiene unas orejas tan grandes! – His ears are so big!
13498 | 1325191

105 pensar *v* to think

- trataba de pensar de dónde podría sacar el dinero – he tried to think of where he could get some money
12014 | 1417538

106 casa *nf* house

- bienvenido a mi casa – welcome to my home
14237 | 862495

107 cada *adj* each, every

- aprendo cada día un poco más – I learn a little more each day
14313 | 1896200

108 persona *nf* person

- era una persona de mucho talento – she was a very talented person
11205 | 2605899 +o +w

109 país *nm* country

- ha viajado por todos los países de las Américas – he has traveled through all the countries in the Americas
19994 | 1937398

110 algo *pron* something, somewhat

- tengo algo que decirte – I have something to tell you
10802 | 1698004

111 esto *pron* this (n)

- quiero esto o aquello – I want this or that
11112 | 1932241 +o

112 volver *v* to return, to V again

- Los habitantes afectados pueden volver a sus casas. – Affected people can return to their homes.
12732 | 991596

113 usted *pron* you (subj./fam)

- ¿Cómo está usted? – How are you?
11604 | 934651 +o

114 salir *v* to leave, go out

- ¿A qué hora sale el tren? – What time does the train leave?
11954 | 1136211

115 después *adv* after

- vamos a dar un paseo después de comer – we are going to go for a walk after eating
13264 | 1375185

116 mejor *adj* best, better (ADV)

- yo, al fin y al cabo, soy el mejor amigo – I, after all, am the best friend
11220 | 1955684

117 menos *adj* less, fewer

- no vale más ni menos que la anterior – it is worth no more nor less than the previous one
12484 | 1480647

118 venir *v* to come

- unos afirman que el jazz viene de África – some say jazz comes from Africa
11538 | 1031314 +o

119 forma *nf* form, shape, way

- necesito estar en buena forma física – I need to be in good physical shape
16411 | 1627720

120 mujer *nf* woman, wife

- su mamá es una mujer muy enferma y no puede trabajar – your mother is a very sick woman and she cannot work
11127 | 1222695

121 momento *nm* moment, time

- estaba pasando uno de los peores momentos de mi vida – I was having one of the worst moments of my life
11664 | 1345941

122 llamar *v* to call, name

- el juez se llamaba Pedro Barreda – the judge was named Pedro Barreda
12663 | 1006445

123 mundo *nm* world

- no se parecía a ninguna ciudad del mundo – it didn't look like any other city in the world
12289 | 1696206

124 claro *adj* clear

- Eres genial por explicarlo de forma tan clara. – It is so nice of you to explain it in such a clear way.
12245 | 500115 +o

125 mirar *v* to look, watch

- miraba distraída a las estrellas – I gazed absentmindedly at the stars
11499 | 473946

126 medios *nm* means, resources

- No cuenta con los medios suficientes para comprar los medicamentos que necesita. – He doesn't have the means to buy the medicine he needs.
10198 | 1034356

127 mientras *conj* while, whereas, as long as

- nos miró mientras comía un gran trozo de pan – he looked at us while eating a big piece of bread
10233 | 1900487

128 conocer *v* to know (someone or place)

- ella conoce muy bien a Nicolás – she knows Nicholas very well
11312 | 1347692

129 aquel *adj* that (over there)

- unos veinte millones de pesetas (de aquella época) – about twenty million pesetas (at that time)
10264 | 1194884

130 aquí *adv* here

- ¡Aquí hay dos también! – There are two here as well!
10408 | 987546 +o

131 aunque *conj* although, even though

- aunque hay problemas, muchos se pueden evitar – even though there are problems, many can be avoided
13708 | 1343325

132 quien *pron* who, whom

- ¿Quién sabe el título de la película? – Who knows the name of the movie?
11451 | 1811913

133 tomar *v* to take, drink

- mi padre tomó un sorbo de agua y continuó hablando – my father drank a sip of water and continued speaking
10273 | 1245473

134 tres *num* three

- hay tres direcciones: diagonal, horizontal, y vertical – there are three directions: diagonal, horizontal, and vertical
13375 | 904878

135 mano *nf* hand

- se pueden contar con los dedos de la mano – you can count with the fingers on one hand
10282 | 699443

136 sentir *v* to feel, regret

- es una expresión personal de un sentir personal de él – it is a personal expression of his personal feeling
8436 | 1165631

137 gente *nf* people

- la gente tiene miedo de perder su identidad – people are scared of losing their identity
8869 | 1156740 +o

138 cuándo *adv* when?

- ¿Cuándo vuelven? preguntó – When do they return? he asked
1059 | 3891764 +o +w

139 durante *prep* during

- durante una semana traté de no verla – during the week I tried not to see her
15759 | 1261080

140 hijo *nm* son, [pl] children

- cada padre espera que su hijo sea más que él – every father hopes that his child will be more than him
10211 | 920802

141 tratar *v* to try, treat, deal with

- Traté de aprender español, pero es muy difícil. – I tried to learn Spanish, but it is very difficult.
11217 | 1207426

142 vivir *v* to live

- vamos a Buenos Aires, a vivir en la miseria – I am going to Buenos Aires to live in poverty
9590 | 1243358

143 *sino conj* but, except, rather

- no somos nosotros el centro sino ella – we are not the focus, but rather she is
9899 | 1244242

144 *lugar nm* place, position

- es el lugar más lindo de la costa – it is the most beautiful place on the coast
11400 | 1243840

145 *problema nm* problem

- el problema es que no nos queda tiempo – the problem is that we don't have enough time
9816 | 1277381 +o

146 *hacia prep* toward, towards

- sus sentimientos hacia él son todavía tan firmes – her feelings for him are still so strong
11866 | 741099

147 *presidente nm/f* president

- el presidente es la máxima autoridad del país – the president is the one with the most authority in the country
8921 | 620111

148 *cual pron* which, who, whom

- el amor fue el motivo por el cual lo hizo – love was the reason for which he did it
9747 | 1589043

149 *junto adj* together

- la tierra giraba en torno al sol junto con cinco planetas más – the earth revolved around the sun together with five more planets
11077 | 568036

150 *luego adv* later, afterwards

- primero está ella y luego yo – she's first then I'm next
9231 | 946952 +o

151 *cómo adv* how?

- ¿Cómo se hace un pastel de chocolate? – How do you make a chocolate cake?
10654 | 751068 +o

152 *trabajo nm* work, job, effort

- pasó su vida entera dedicado a su trabajo – she spent her entire life dedicated to her work
9316 | 1369672

153 *caso nm* case, occasion

- luego, en este caso está plenamente justificada – then, in this case it is fully justified
10589 | 1275188

154 *mayor adj* larger, older, main

- yo trabajé la mayor parte del tiempo – I worked the majority of the time
12862 | 1146715

155 *además adv* also, as well, besides

- además de español, ¿hablas algún otro idioma? – besides Spanish, do you speak any other language?
9862 | 1287844

156 *acercar v* to come near

- ella me miró, se acercó, me besó – she looked at me, came up to me, and kissed me
9202 | 235097 +f

157 *esperar v* to wait, hope (for), expect

- te esperamos en el restaurante – we'll wait for you at the restaurant
8192 | 1047055

158 *nunca adv* never, ever

- se aprecia ahora más que nunca la libertad – you appreciate freedom now more than ever
8018 | 885725

159 *cierto adj* certain, sure, true

- no sé si esto es cierto o no – I don't know if this is true or not
9361 | 801421

160 *hora nf* hour, time (specific)

- te veo dentro de media hora – I will see you in half an hour
8778 | 939823

161 *dónde adv* where?

- ¿Dónde está escrito esto en la historia? – Where is that written in history?
3088 | 2385406 +o +w

162 *padre nm* father

- ¿Es el padre de mi futuro hijo? – Is he the father of my future child?
8732 | 798637

163 *gustar v* to be pleasing to

- ¡Me gusta mucho la música folklórica! – I really like country music!
6726 | 1015967 +o

164 *noche nf* night, evening

- trabaja por la mañana, por la tarde, y por la noche – he works in the morning, the afternoon, and the evening
8409 | 453279

165 *nosotros pron* we (subj)

- ¡ábreos! somos nosotros – open up! it's us
6985 | 944327 +o

166 *casi adv* almost, nearly

- estamos casi siempre solos – we are almost always alone
9678 | 707650

167 *hoy adv* today, nowadays

- ayer estuvo aquí; hoy puede estar allá – yesterday he was here; today he might be there
9156 | 1226935

1. Animals (animal ₃₂₂)

Top 50 words

perro 888-M dog	pez 3579-M fish	tortuga 5064-F turtle
caballo 907-M horse	serpiente 3662-F snake	abeja 5144-F bee
pájaro 1607-M bird	cerdo 3720-M pig	sapo 5828-M toad
gato 1728-M cat	tigre 3813-M tiger	cabra 5863-F goat
ave 2086-M bird	lobo 3816-M wolf	mosquito 5931-M mosquito
león 2107-M lion	mono 3835-M monkey	rana 6326-F frog
toro 2280-M bull	oveja 4040-F sheep	pajarito 6480-M little bird
vaca 2775-F cow	gusano 4217-M worm	víbora 6670-F viper
insecto 3006-M insect	araña 4427-F spider	zorro 6815-M fox
gallo 3025-M rooster	oso 4449-M bear	ballena 6911-F whale
gallina 3153-F hen	hormiga 4482-F ant	grillo 7012-M cricket
paloma 3182-F dove	burro 4626-M donkey	cuervo 7109-M raven
mariposa 3359-F butterfly	conejo 4728-M rabbit	murciélago 7143-M bat
ratón 3415-M mouse	buey 4769-M ox	cisne 7366-M swan
mosca 3432-F fly	pato 4857-M duck	ganado 8001-M livestock
pollo 3577-M chicken	elefante 5001-M elephant	gaviota 8872-F seagull
rata 3596-F rat	águila 5027-M eagle	

168 tipo *nm* type, kind

- reconozco tu oposición a este tipo de ideología – I recognize your opposition to this kind of ideology
9812 | 1114288

169 ojo *nm* eye

- los grandes ojos verdes la miraban con ternura – his large green eyes looked at her with tenderness
7668 | 407074 +f

170 manera *nf* way, manner

- lo hizo de manera muy informal – he did it in a very informal manner
7278 | 1167687

171 importante *adj* important

- el momento más importante de la ceremonia – the most important moment of the ceremony
10834 | 1182061

172 contar *v* to tell, count

- me contó una larga y estúpida historia – he told me a long and silly story
7680 | 1058235

173 niño *nm* child, little boy

- he tenido suerte desde que era niño – I have been lucky since I was a child
7839 | 951541 +o

174 trabajar *v* to work

- comenzaban a trabajar de camareros – they began to work as waiters
7870 | 993836 +o

175 empezar *v* to begin, start

- lentamente su auto empieza a avanzar – his car slowly began to move forward
7692 | 768959

176 verdad *nf* truth

- diga la verdad! – tell the truth!
7268 | 838669 +o

177 contra *prep* against, opposite

- hermanos contra hermanos y padres contra hijos – brothers against brothers and fathers against sons
9716 | 1039671

178 ciudad *nf* city

- quieren estar en la ciudad y no salir al campo – they want to be in the city and not leave for the country
11294 | 816256

179 buscar *v* to look for

- decidí volver a la calle y buscar un empleo – I decided to return to the streets and look for a job
6677 | 1028159

180 ahí *adv* there

- debe estar por ahí – it must be around there
7524 | 509440 +o

181 solo *adj* lonely, alone

- se sentía solo, espantosamente solo – he felt alone, terrifyingly alone
6415 | 1004305

182 fin *nm* end; a f. de: in order to

- puse fin a la discusión – I put an end to the discussion
8165 | 804876

183 propio *adj* own, proper, typical

- le doy dinero de mi propio bolsillo – I give him money from my own pocket
8448 | 1087489

184 tú *pron* you (subj.-sg/+fam)

- ella es más alta que tú – she is taller than you
8267 | 396838 +o

185 gobierno *nm* government

- el feudalismo fue la forma de gobierno – feudalism was the form of government
13260 | 1084108 +nf

186 historia *nf* history, story

- aquello ya pasó a la historia – that's ancient history now
8769 | 1003036

187 ejemplo *nm* example

- no des mal ejemplo a los niños – don't set a bad example for the kids
10173 | 397593 +o

188 último *adj* last, final

- se opuso al viaje hasta el último momento – he was against the trip until the last minute
12625 | 1282453

189 punto *nm* point, dot, period

- este punto de vista es mejor – this point of view is better
8516 | 906040

190 antes *adv* before

- quería verla antes de salir – I would like to see her before leaving
5870 | 1222943

191 mil *num* thousand

- unos mil años han pasado desde entonces – about a thousand years have passed since then
7149 | 626797

192 palabra *nf* word

- jamás dijo una palabra en inglés – he never said a word in English
6607 | 847879

193 obra *nf* work, book, deed

- la primera obra literaria de Dante fue La Vida Nueva – the first literary work of Dante was *The New Life*
12665 | 671187

194 existir *v* to exist

- ¿Puede existir alguien que sea perfecto? – Is it possible that a perfect person exists?
7661 | 1090490

195 perder *v* to lose, miss

- quiere perder diez libras antes de la boda – she wants to lose ten pounds before the wedding
6864 | 753523

196 cuenta *nf* bill, account

- sería bueno sacar la cuenta y calcular gastos – it would be good to take inventory and calculate the costs
6185 | 947138

197 allí *adv* there, over there

- allí están, junto a la playa – there they are, on the beach
7391 | 395400

198 escribir *v* to write

- mi padre escribe unas cartas cortas – my father wrote a few short letters
7776 | 913546

199 cualquiera *adj* any, anyone

- son patrocinadas con lo que cualquiera puede donar – they are paid for by donations from anyone
5996 | 1222552

200 grupo *nm* group

- ¿Cuál será su papel en ese grupo de trabajo? – What will be your role in this work group?
9482 | 902746

201 señor *nm* sir, Mr, lord

- ¿Puedo pagar con dólares? // Sí, señor. – Can I pay with dollars? // Yes, sir.
8548 | 433994 +o

202 pequeño *adj* little, small, young

- a pesar de su pequeño tamaño, estaba llena de detalles – in spite of its small size, it was full of details
8077 | 657629

203 embargo *adv* nevertheless

- sin embargo, es uno de mis mejores amigos – nevertheless, he is one of my best friends
8903 | 825538

204 agua *nf (el)* water

- Déme un vaso de agua; que tengo mucha sed. – Give me a glass of water; I am very thirsty.
8249 | 707456

205 realizar *v* to fulfill, carry out

- el viaje se realizó en noviembre – the trip took place in November
8318 | 1236630

206 nadie *pron* nobody, anybody

- nadie sabe nada – nobody knows anything
5832 | 640822

207 entrar *v* to enter

- ¡Este maldito no puede entrar en nuestra casa! – This devil cannot enter into our house!
6423 | 566171

208 arte *nm* art, skill

- la creación de un museo de arte moderno y contemporáneo – the creation of a museum of modern and contemporary art
6827 | 289249 +f

209 leer *v* to read

- prefiero leer a los autores de moda – I prefer to read the popular authors
4607 | 1055077 +o +w

210 amigo *nm* friend

- ¿No tienes un amigo a quien me podrías presentar? – You don't have a friend that you could introduce me to?
5887 | 711272 +o

211 recordar *v* to remember, remind

- prefiero no recordar esa época de mi vida – I prefer not to remember that time of my life
5736 | 737507

212 morir *v* to die

- los pobres murieron de pulmonía meses después – the poor people died of pneumonia months later
6594 | 497833

213 único *adj* only, unique, sole

- era el único hijo varón de la familia – he was the only male child in the family
6677 | 735788

214 lado *nm* side

- está situada al lado derecho – it is situated on the right side
6891 | 412948

215 nombre *nm* name, noun

- Susana no es mi nombre verdadero – Susana is not my real name
7223 | 803817

216 recibir *v* to receive

- Pavlov recibe el Premio Nobel por sus estudios – Pavlov receives the Nobel Prize for his studies
7213 | 870744

217 pedir *v* to ask for, request

- me atreví a pedirle que bailara con migo – I dared to ask if he would dance with me
5954 | 772822

218 permitir *v* to allow, permit

- no se me permitió jugar allí por varios días – I wasn't allowed to play there for several days
7076 | 1042896

219 preguntar *v* to ask (a question)

- les voy a preguntar si quieren comer – I am going to ask them if they want to eat
5369 | 556854

220 social *adj* social

- no me gusta la vida social, no voy a ninguna fiesta – I don't like social life, I'm not going to any party
7085 | 1152802

221 sistema *nm* system

- el corazón es parte del sistema circulatorio – the heart is part of the circulatory system
8999 | 1000105

222 tal *adj* such (a)

- bien sabía que tal cosa era imposible – it was well known that such a thing was impossible
5867 | 872650

223 segundo *adj* second

- espérame un segundo – wait just a second
7979 | 694262

224 ante *prep* before, in the presence of

- no sé qué hacer y cómo actuar ante estas circunstancias – I don't know what to do and how to act faced with these circumstances
7035 | 838595

225 viejo *adj* old, aged

- el hombre viejo tiene como ochenta años, y está loco – the old man is about eighty years old, and he is crazy
6198 | 290993

226 madre *nf* mother

- ahora está casada y es madre de familia – now she is married and the mother of a family
5970 | 444891

227 siglo *nm* century, age

- los poetas más populares del siglo 20 – the most popular poets of the twentieth century
14056 | 356541 +nf

228 nacional *adj* national

- le dieron el premio nacional de poesía – they gave him the national award in poetry
11350 | 626349

229 entender *v* to understand

- todavía no entiendo demasiado bien lo que pasó – I still don't understand very well what happened
5086 | 734733 +o

230 libro *nm* book

- ¿Está usted escribiendo un libro sobre la política? – You are writing a book about politics?
6394 | 852313

231 alto *adj* tall, high

- estás tan alto como yo – you are as tall as I am now
7796 | 641429

232 cuerpo *nm* body

- el cuerpo humano se compone de millones de células – the human body is composed of millions of cells
6243 | 591728

233 familia *nf* family

- la familia es la unidad social más significativa – family is the most significant social unit
6333 | 785152

234 comenzar *v* to begin, start

- a los veintidós años comenzó a aprender castellano – at twenty-two years old, he began to learn Spanish
7122 | 682837

235 presentar *v* to introduce, present

- Mi madre me presentó a mi futuro esposo. – My mother introduced me to my future husband.
6988 | 880911

236 bajo *prep* under, underneath

- encontraron las sepulturas bajo tierra – they found the tombs under the ground
7992 | 445300

237 según *prep* according to

- Según su experiencia, ésa era la pista correcta. – According to his experience, this was the right trail.
7719 | 942482

238 frente *adv* front

- si uno situa un espejo frente a otro, crea un portal infinito – if you place one mirror in front of another, it creates an infinite loop
6880 | 519496

239 crear *v* to create

- su actitud crea muchos problemas – your attitude creates a lot of problems
6162 | 1001336

240 tema *nm* theme, subject, topic

- es un tema delicado – it's a delicate topic
5482 | 1050279

241 cuatro *num* four

- tienen cuatro patas y cuernos – they have four legs and horns
7543 | 418165

242 ah *interj* oh

- Ah, bueno . . . ¿Ya movieron los materiales? – Oh, good . . . You already moved the materials?
5730 | 70565 +o

243 humano *adj* human

- El sistema nervioso humano tiene sistemas neurales complejos. – The human nervous system has complex neural systems.
6771 | 916865

244 pueblo *nm* people, village

- el libro no representa al pueblo portugués – the book did not represent the Portuguese people
6356 | 737451

245 producir *v* to produce, cause

- tus acciones me producen tristeza – your actions cause me grief
8884 | 651823

246 abrir *v* to open

- Abre la ventana. Hace calor. – Open the window. It's hot.
5314 | 559877

247 idea *nf* idea

- no podemos renunciar a la idea de la igualdad – we can't give up on the idea of equality
5459 | 775327

248 millón *num* million, fortune

- le ha proporcionado un cheque de un millón de pesetas – he has given him a check for a million pesetas
9748 | 729051 +nf

249 sentar *v* to sit (down), seat

- los niños pasaban a sentarse en sus escritorios – the children went to sit at their desks
5661 | 184493

250 aún *adv* still, yet

- ni aún así me quedaría – not even then would I stay
5253 | 910663

251 caer *v* to fall

- no quiero caer en mi tumba siendo joven – I don't want to fall down dead at a young age
5582 | 394933

252 voz *nf* voice

- un hombre de voz vibrante y rica empezó a cantar – a man with a rich and vibrant voice began to sing
5705 | 310208

253 terminar *v* to finish, end

- cuelga sin terminar la conversación – he hangs up without finishing the conversation
5396 | 592044

254 oír *v* to hear

- no se oía ningún sonido – not a sound was heard
5816 | 201839 +o

255 cambiar *v* to change

- no van a cambiar la disposición de invitados – they are not going to change the seating arrangement
4790 | 775358

256 ninguno *adj* no, none, nobody (PRON)

- No tengo ningún lugar adonde ir. – I have no place to go.
4185 | 947315

257 considerar *v* to consider

- no podía considerar la primera hipótesis con seriedad – I couldn't seriously consider the first hypothesis
7026 | 829130

258 político *adj* political

- la acción política debe respaldarse en la económica – political action should be supported economically
8383 | 779108

259 tierra *nf* earth, land, ground

- la tierra aquí es oscura y rica – the land here is dark and rich
7254 | 495369

260 realidad *nf* reality, actuality

- le habló a Rosabel, que en realidad se llamaba María – he spoke to Rosabel, who in reality is named María
5665 | 487074

261 convertir *v* to convert, change, become

- el café llegó a convertirse en la bebida más popular de Cuba – coffee went on to become the most popular drink in Cuba
7415 | 667497

262 español *adj* Spanish

- la lengua española de hoy es el castellano – the Spanish language of today is Castilian
9074 | 497237

263 igual *adj* equal, same (as)

- los indios son seres humanos iguales que nosotros – Indians are human beings, the same as us
5066 | 688505

264 todavía *adv* still, yet

- los animales todavía están vivos – the animals are still alive
5554 | 378047

265 cabeza *nf* head (part of body)

- planteaba muchas preguntas en mi cabeza – he planted a lot of questions in my head
5247 | 326564

266 derecho *nm* right, justice, law

- Tengo derecho a opinar. – I have the right to express my opinion.
5448 | 1049701

267 mantener *v* to keep, maintain

- Luis cruzó los brazos y se mantuvo sereno – Luis crossed his arms and kept calm
6465 | 753601

268 utilizar *v* to use, utilize

- sería una buena idea utilizar colores pálidos – it would be a good idea to use pale colors
7696 | 840710

269 calle *nf* street

- llegó a la esquina de la calle La Merced y La Esperanza – she arrived at the street corner of La Merced and La Esperanza
5221 | 373582

270 través *prep* a t. de: through, across

- se enteró a través de un amigo – she found out about it through a friend
5745 | 802538

271 estado *nm* state, condition, status

- hay un estado de incertidumbre que flota en el ambiente – there was a state of uncertainty hanging in the air
8304 | 614614

272 relación *nf* relationship, relation

- he tenido una relación muy estrecha con los compositores – I have had a very close relationship with the composers
6009 | 820343

273 sacar *v* to take out

- sacó un bolígrafo de su bolso y firmó el cheque – she took a pen out of her purse and signed the check
4607 | 482163

274 puerta *nf* door

- ¿Por dónde estará la puerta de salida? – Where would the exit door be?
5095 | 264046

275 gracias *interj* thanks, thank you

- Doy gracias a Dios por haberme permitido alcanzar mi objetivo. – I thank God for letting me achieve my goal.
4005 | 914764 +o

276 necesitar *v* to need

- vas a necesitar un trabajo nuevo – you are going to need a new job
4079 | 798163

277 resultar *v* to result, turn out

- el viaje resultó perfecto – the trip turned out perfectly
6069 | 570630

278 luz *nf* light

- desaparece la luz hasta el oscuro – the light disappeared into the darkness
5375 | 430773

279 paso *nm* step, pace

- di dos pasos adelante y uno hacia atrás – I took two steps forward and one backwards
4987 | 551882

280 general *adj* general

- se realiza en el pabellón de operaciones, bajo anestesia general – it is carried out in the operation room, under general anesthesia
8390 | 352757

281 escuchar *v* to listen to

- tuvo oportunidad de escuchar comentarios de viajeros – he had the opportunity of hearing the travelers' comments
4335 | 515444

282 guerra *nf* war, warfare

- qué pronto se olvida a las víctimas de una guerra – how soon we forget the victims of war
9125 | 360439

283 amor *nm* love

- ¡Tengo el amor más grande del mundo! – I have the greatest love in the world!
3972 | 674603

284 cinco *num* five

- mi hija solamente tiene cinco años – my daughter is only five years old
6056 | 333208

285 situación *nf* situation

- ¿Cómo es la situación allí actualmente? – What is the situation there at present?
5135 | 769236

286 conseguir *v* to get, acquire, obtain

- te puedo conseguir su número – I can get you her phone number
5909 | 529879

287 lograr *v* to achieve, get, manage

- un director no puede lograr buenos resultados sin ayuda – a director can't get good results without help
5306 | 772702

288 mes *nm* month

- la boda tendrá lugar el próximo mes de noviembre – the wedding will take place next month in November
4259 | 892040

289 quién *pron* who?, whom?

- ¿Quién es el padre de su hijo? – Who is the father of your child?
4314 | 389795

290 fuerza *nf* strength, force, power

- necesitaba recuperar fuerzas – I needed to recover my strength
6667 | 477593

291 ocurrir *v* to happen, occur

- el miedo a lo que pudiera ocurrir me inmovilizó – the fear of what could happen immobilized me
4786 | 524809

292 aparecer *v* to appear

- detrás suyo, Teresa apareció por la puerta de la cocina – behind him, Teresa appeared at the kitchen door
5703 | 521233

293 diferente *adj* different, separate

- en un entorno urbano el reto, en cambio, es diferente – in an urban environment, however, the challenge is different
5217 | 782724

294 muerte *nf* death

- sólo la muerte no tiene remedio – only death cannot be cured
5489 | 503586

295 ganar *v* to win, earn

- después de diez años ganó el Premio Nobel – after ten years he won the Nobel Prize
5287 | 620106

296 estudio *nm* study, learning

- he colaborado en un estudio de psicología descriptiva – I have participated in a descriptive psychology study
5844 | 643053

297 nivel *nm* level

- el nivel del mar empezaba a aumentar – the sea level began to rise
5110 | 773448

298 proceso *nm* process, procedure

- el proceso electoral fue justo – the electoral process was fair
5786 | 901705

299 fuera *adv* out, outside, away

- hace como diez años que vivo fuera del Paraguay – I have lived outside of Paraguay for about ten years
4746 | 419264

300 largo *adj* long

- Salgo de pronto a un largo viaje. – I'm leaving soon for a long trip.
5321 | 398575