

WHAT THE TOP SPANISH WORDS CAN TEACH US

or, Wiktionary Mostly Resolves the Targeted/Non-Targeted Debate

By Bryce Hedstrom

Source: http://en.wiktionary.org/wiki/Wiktionary:Frequency_lists/Spanish1000

This is an analysis of the top 505 words on the Wiktionary word frequency list for Spanish. The Wiktionary lists are compiled by analyzing the subtitles from movies, television shows and other videos. The corpus size is millions upon millions of words and growing weekly as more subtitles are added and analyzed by the program. Since screen writers and TV producers attempt to communicate in natural-sounding language, the Wiktionary list may be a close approximation to actual spoken language. Is it a perfect representation of spoken language? Nope. But it may be one of the best we have available. Writers, after all, have to write scripts that sound like normal language to viewers.

Why 505 words? The number is arbitrary, but I would be pleased if my students could understand and use all of these top 505 words, in all of the forms, by level 3 or 4. Analyzing the top 800 or 2,000 words might give a slightly more precise depiction of actual language use, but I am willing to bet that 505 comes close.

The Wiktionary list is different than other word frequency lists because it contains ALL forms of words, including the conjugated forms of verbs, in all tenses and moods, which can sound very different to novices. This analysis can help language teachers in several ways. It can help us to:

- 1. Gain insight into how language is used.*
- 2. Understand which verbs are the most valuable for students to learn.*
- 3. Understand which tenses and moods are most valuable for students.*
- 4. Help us to relax and use natural language with our students.*

The Vocabulary Shelters Itself

Verbs play an outsized role in the most frequently used words on the Wiktionary list, but there is not a lot of variety among them. In the top 505 words on Wiktionary there are just 51 different verbs, but these verbs appear in many different forms. Notice the similarities between the top 12 verbs on these two lists:

Wiktionary Top 12 Verbs:

hacer	estar	ver
ir	haber	deber
ser	poder	saber
decir	tener	querer

Frequency Dictionary of Spanish (Mark Davies) Top 12 Verbs:

ser	hacer	ver
haber	poder	dar
estar	decir	saber
tener	ir	querer

Among the most-used words on these and other frequency lists the same verbs keep showing up. The naturally occurring vocabulary seems to be sheltering itself.

The Grammar Shelters Itself

Specific forms of the top 8 verbs alone account for 86/505 words on the Wiktionary list (17%); the top 20 verbs, 137/505 (27%). In all, **the various conjugations of verbs account for 174 of the 505 most frequently used words in Spanish (34%)**.

This limited set of verbs appear in many different grammatical forms, but some forms are much more common. Here is the breakdown of tenses, moods and forms among those most frequently used verbs.

<i>Tense/Mood</i>	<i># of Occurrences</i>	<i>% of this Tense/Mood Is Used</i>
present	91	52% of all verbs on this list
infinitive	29	16%
preterit	15	9%
imperfect	9	6%
present subjunctive	8	5%
past participle	6	3.4%
conditional	5	2.8%
imperative	4	2.2%
past subjunctive	3	1.7%
future	2	1.1%
present participle	2	1.1%

Many verb forms, some of which world language teachers would consider advanced, are used in this basic set, but the vocabulary is limited. If the size of the Wiktionary corpus were smaller it would be more suspect, but each Wiktionary list is based on such an immense corpus (millions upon millions of words) that it may approach an accurate depiction of natural language use.

The implications for teaching are that if we want our students to develop fluency, we need use natural language. We need to use language that focuses on high-frequency vocabulary, particularly the top verbs, and use the grammar that fits the moment.

Susan Gross taught, **“Shelter vocabulary, not grammar.”** This list indicates that the idea of sheltering vocabulary, but not grammar, occurs naturally. On it we see all kinds of grammar in the form of conjugations, but the vocabulary is limited. If we use natural language, we will automatically shelter vocabulary but not grammar. Natural language use will give students the vocabulary they need (the highest frequency words) as well as the grammatical forms (the most used verb conjugations) that students most need to learn.

Given enough input, the acquisition of the most important vocabulary and grammar seems to take care of itself. With time and comprehended input, students will acquire what they need to acquire. We can only screw it up if we overthink the process and focus too much on one particular aspect of language to the neglect of natural language use and real communication. Still, these questions are worth pondering:

- Are we using these verb forms proportionally and appropriately with our students?
- Can we extrapolate to all of the language based on the top 505 most frequently used words?
- Does the speech that non-native teachers use in the classroom alter natural language acquisition?
- Are textbooks and other curricular choices giving input that is less valuable to students?
- If we cannot consciously and spontaneously use these words, can we find materials that purposefully do?

The Top Spanish Verbs

The frequency lists from Wiktionary show that natural language use tends to limit vocabulary, but not grammar. The verbs are in a relatively small set. They tend toward a set of highly used ones, while the grammar includes a diversity of forms. *If we want students to become fluent, we need to focus on the high frequency vocabulary, the top verbs.* Susie Gross always said, "Shelter vocabulary, not grammar." This list confirms that idea. On it we see *all kinds of conjugations, but limited vocabulary.*

1. Hacer (14)

hacer _____
 hace _____
 hecho _____
 haciendo _____
 haces _____
 hacerlo _____
 hizo _____
 hice _____
 haré _____
 hiciste _____
 hago _____
 hagas _____
 haga _____
 hacen _____

2. Ir (13)

vamos _____
 fue _____
 voy _____
 va _____
 ir _____
 vas _____
 fuera _____
 van _____
 vaya _____
 iba _____
 vete _____
 fui _____
 fueron _____

3. Ser (13)

es _____
 soy _____
 ser _____
 son _____
 era _____
 eres _____
 sea _____
 fuera _____
 será _____
 sido _____
 somos _____
 sería _____
 eran _____

4. Decir (10)

decir _____
 dijo _____
 dije _____
 dice _____
 digo _____
 dicho _____
 dijiste _____
 dices _____
 dicen _____
 dime _____

5. Estar (10)

está _____
 estoy _____
 estás _____
 estaba _____
 están _____
 estar _____
 estamos _____
 estado _____
 estabas _____
 esté _____

6. Haber (10)

hay _____
 ha _____
 he _____
 has _____
 han _____
 había _____
 haber _____
 hemos _____
 haya _____
 hubiera _____

8. Tener (8)

tengo _____
 tiene _____
 tienes _____
 tenemos _____
 tener _____
 tenía _____
 tienen _____
 tuve _____

10. Deber (6)

debe _____
 debo _____
 debería _____
 debes _____
 debemos _____
 deberías _____

13. Creer (4)

creo _____
 crees _____
 cree _____
 creer _____

7. Poder (8)

puedo _____
 puede _____
 puedes _____
 podemos _____
 podría _____
 pueden _____
 poder _____
 pueda _____

9. Ver (7)

ver _____
 ven _____
 visto _____
 ve _____
 veo _____
 ves _____
 vi _____

11. Saber (6)

sé _____
 sabes _____
 sabe _____
 saber _____
 sabía _____
 saben _____

14. Necesitar (4)

necesitamos _____
 necesita _____
 necesitas _____
 necesito _____

12. Querer (5)

quiero _____
 quieres _____
 quiere _____
 quería _____
 quieren _____

15. Pasar (4)

pasó _____
 pasa _____
 pasar _____
 paso _____

- | | | | |
|--|---|--|---|
| 16. <u>Venir</u> (4)
viene _____
venir _____
venga _____
vino _____ | 23. <u>Pensar</u> (2)
pensé _____
pensar _____ | 32. <u>Entender</u> (1)
entiendo _____ | 42. <u>Preocupar</u> (1)
preocupes _____ |
| 17. <u>Dar</u> (3)
da _____
dame _____
dar _____ | 24. <u>Tomar</u> (2)
toma _____
tomar _____ | 33. <u>Entrar</u> (1)
entrar _____ | 43. <u>Quedar</u> (1)
queda _____ |
| 18. <u>Dejar</u> (3)
deja _____
déjame _____
dejar _____ | 25. <u>Vivir</u> (2)
vivar _____
vivo _____ | 34. <u>Escuchar</u> (1)
escucha _____ | 44. <u>Salir</u> (1)
salir _____ |
| 19. <u>Hablar</u> (3)
hablar _____
hablando _____
habla _____ | 26. <u>Amar</u> (1)
amo _____ | 35. <u>Importar</u> (1)
importa _____ | 45. <u>Seguir</u> (1)
sigue _____ |
| 20. <u>Esperar</u> (2)
espera _____
espero _____ | 27. <u>Ayudar</u> (1)
ayuda _____ | 36. <u>Llamar</u> (1)
llama _____ | 46. <u>Sentir</u> (1)
siento _____ |
| 21. <u>Gustar</u> (2)
gusta _____
gustaría _____ | 28. <u>Comer</u> (1)
comer _____ | 37. <u>Llegar</u> (1)
llegar _____ | 47. <u>Significar</u> (1)
significa _____ |
| 22. <u>Morir</u> (2)
muerto _____
morir _____ | 29. <u>Conocer</u> (1)
conozco _____ | 38. <u>Mirar</u> (1)
mira _____ | 48. <u>Suponer</u> (1)
supongo _____ |
| | 30. <u>Contar</u> (1)
cuenta _____ | 39. <u>Matar</u> (1)
matar _____ | 49. <u>Trabajar</u> (1)
trabajar _____ |
| | 31. <u>Encontrar</u> (1)
encontrar _____ | 40. <u>Oír</u> (1)
oye _____ | 50. <u>Valer</u> (1)
vale _____ |
| | | 41. <u>Parecer</u> (1)
parece _____ | 51. <u>Volver</u> (1)
volver _____ |

There are spaces on the list above to check student comprehension from time-to-time—not necessarily as a test, but as a measure of which verbs and forms that YOU, the teacher, need to focus on a bit more. If several students do not know a verb or form, you will need to purposefully use it more often.