EPiC Series in Language and Linguistics
Volume 1, 2016, Pages 396–410
CILC2016. 8th International
Conference on Corpus Linguistics

The Portuguese Vocabulary Profile: Pilot Study

Shintaro Torigoe
Tokyo University of Foreign Studies, Tokyo, Japan.
Gunma Prefectural Women's University, Gunma, Japan.
torigoe.shintaro@tufs.ac.jp

Abstract

This paper reports on a pilot study from the *Portuguese Vocabulary Profile* project. In this pilot study, a vocabulary list for learners of Portuguese was developed by analysing learner corpora, an approach inspired by CEFR-based wordlists, such as the *English Vocabulary Profile*. A draft wordlist was constructed from two learner corpora of L2 Portuguese, the *Corpora do PLE* and the *Corpus de PEAPL2*. The draft wordlist was then compared to the *LMCPC*, a wordlist derived from a million-word native speaker corpus, in order to investigate differences between learners and native speakers and to identify aspects of the wordlist needing improvement. The pilot study indicated that the use of Portuguese by the Intermediate and Advanced learner is quite different from that of native speakers and that learner's language use was affected by data collection tasks and learning environments.

1 Introduction

The aim of the pilot study reported in this paper was to construct the *Portuguese Vocabulary Profile*, a vocabulary list for elementary, intermediate, and advanced learners of Portuguese. The list was derived by analysing learner corpora, an approach inspired by the *English Vocabulary Profile* (Capel 2010, 2012). In Japan, an increasing number of Portuguese language textbooks have been published since 2010 (cf. Torigoe & Yamada 2015), yet there are no guidelines for appropriate tasks, grammar, or vocabulary. In response to this situation, the author proposed developing a prototype wordlist by adopting the *Common European Framework of Reference for Languages* (*CEFR*) levels (Council of Europe 2001) as proficiency criteria and by using learner corpora as the data source. Such a list could be expected to improve the quality of Portuguese language textbooks in Japan and also contribute to the development of textbooks for Intermediate learners (resource which are currently lacking in the Japanese market).

A.Moreno Ortiz and C.Pérez-Hernández (eds.), CILC2016 (EPiC Series in Language and Linguistics, vol. 1), pp. 396–410

2 Background

As mentioned above, there is an increasing number of Portuguese language textbooks in Japan, especially since 2010. Most of these are intended for beginners; however, the goals of the various books differ. Moreover, within these textbooks there is no specification of the amount of vocabulary included, or the criteria adopted for vocabulary selection.

In Japan, there are robust national guidelines, *the Courses of Studies* (Ministry of Education, Culture, Sports, Science and Technology-Japan 2009), which define detailed goals for the learning and teaching of the majority of subjects including English. The guidelines, however, simply describe the goals of foreign languages other than English as they 'should follow the objectives and contents of English instruction' (ibid, p.90). This effectively means that the decisions about the teaching of these subjects are delegated to language teachers and textbook writers.

For English language learning, there are examples. One approach to establishing vocabulary goals is based on the language use of native speakers (NS), with corpora used as the data sources. For example, Tono (2013) reports that the top 100 words of the spoken subcorpora of the *British National Corpus* account for 67 per cent of the total word frequency, the top 1000 words account for 87 per cent, and the top 2000 words account for 92 per cent. A vocabulary target of 1000 words is equivalent to the goal specified for English language learning in the 8th grade in Japan, and 2000 words is equivalent to the expectation for 10th grade. In Japan, there are some learning books for English vocabulary based on corpus information, such as Ishikawa (2006) or Tono (2008).

For Portuguese language learning, some corpus-based teaching materials are also available. They are mainly dictionaries; for example, *Dicionário Verbo Língua Portguesa* (2008), based on the *Corpus de Referência de Português Comtemporâneo* (*CRPC*)[*]; *Dicionário Gramatical de Verbos Portugueses* (2007), based on the *Português Fundamental*[†]; and *A Frequency Dictionary of Portuguese* (Davis & Preto-Bay 2008), based on the *Corpus do Português*[‡]. In Japan, there are some examples of corpus-based teaching materials for Portuguese, such as Aires & Iyanaga (2012), a vocabulary book based totally on the *LMCPC* (a subcorpus of the *CRPC*); and Ichinose et al. (2015), a Portuguese-Japanese dictionary that uses the *Corpus do Português* as one reference for important words.

2.1 Does native speakers' language use correspond to learners' use?

All the resources identified in the previous section are examples of teaching materials based on native speaker corpora. In second language acquisition research, however, there is a suggestion that NSs' use of language may not always correspond to learners' use. The author agrees with this idea and, therefore, adopts data sources that depict learner use of Portuguese as the basis for developing the new vocabulary list. In other words, learner corpora are used to describe learner use. The *Common European Framework of Reference for Languages* (*CEFR*; Council of Europe 2001) is also used to provide criteria for learner proficiency.

2.2 The *Common European Framework of Reference for Languages*

The *Common European Framework of Reference for Languages* (Council of Europe 2001) provides versatile guidelines for language learning in multilingual and multicultural contexts. The *CEFR* is not based on dogmatic notions of 'what learners should learn', but, rather, on actual learner

[*] www.clul.ul.pt/pt/investigacao/183-reference-corpus-of-contemporary-portuguese-crpc
[†] www.clul.ul.pt/pt/recursos/84-spoken-corpus-qportugues-fundamental-pfq-r
[‡] www.corpusdoportugues.org/

behaviour, that is 'what learners can do'. The *CEFR* is applied very widely, including in learner assessment by teachers, learner self-assessment, textbook writing, syllabus development, etc.

The *CEFR* assesses listening, speaking, reading, writing, and oral interaction, using six proficiency levels: A1 and A2 ('basic user'), B1 and B2 ('independent user'), and C1 and C2 ('proficient user'). According to Tono (2013), C-level language learners can perform as well as, or even better than, native speakers. For example, a C-level learner '[c]an hold his/her own in formal discussion of complex issues, putting an articulate and persuasive argument, at no disadvantage to native speakers' (Council of Europe 2001, 78), although, of course, it is difficult to compare these two groups objectively.

For further details of the CEFR, such as its history, the examples of its implementation, and related problems, see Council of Europe (2001).

2.3 CEFR-based wordlist

The *English Vocabulary Profile* (Capel 2010, 2012) and the *CEFR-J Wordlist* (Tono 2013) are two instances of vocabulary lists that have been constructed with reference to learner corpora, which have been divided into subcorpora according to the *CEFR* proficiencies.

According to Capel (2010), the *EVP* is a wordlist initially organised using the headwords of the *Cambridge Learner Dictionary*, which is based on a native speaker corpus (the *Cambridge International Corpus*). However, to enhance its validity, the vocabulary in the *EVP* was compared with a *CEFR*-based learner corpus (the *Cambridge Learner Corpus*) and textbook corpora, and was checked by native speakers. A noteworthy feature of the *EVP* is that the classification of vocabulary into the *CEFR* proficiency levels was done not by lemma, but by meaning.

The *CEFR-J* (Tono 2013) is a modified version of the *CEFR* and was developed to reflect the particular context of English teaching in Japan. It has more proficiency classifications within the A and B levels than the original *CEFR*. The *CEFR-J Wordlist* (ibid) is derived from grading corpora of English textbooks in East Asian countries by *CEFR* proficiencies. Tono (ibid) compares the *EVP* and the *CEFR-J Wordlist* quantitatively (Table 1).

CEFR proficiency levels	Words in *EVP*	Words in *CEFR-J Wordlist*
A1	601	1000
A2	925	1000
B1	1429	2000
B2	1711	2000
C1	2300	n/a

Table 1: Comparison between the English Vocabulary Profile and the CEFR-J wordlist

Although Tono does not compare the content of the two wordlists, he shows that the *CEFR-J Wordlist* has more words in each proficiency level from A1-B2, on the other hand, the *CEFR-J Wordlist* does not have the C-levels.

For Portuguese language learning, there is currently no wordlist available that is based on learner corpora. However, two learner corpora are available that are composed of subcorpora classified by *CEFR* proficiencies (see 4.1 below). These may represent useful source material for developing a wordlist for learners for learners of Portuguese.

3 Research questions

This paper describes a pilot study that involved using *CEFR*-based learner corpora to construct a wordlist for learners of Portuguese. The research questions are below.

1. How much vocabulary should learners learn?
2. Does native speaker's use correspond to learner's use?

The author first developed the new wordlist from learner corpora and then examines the degree of similarity between this list and a wordlist derived from a NS corpus.

4 Methodology

Although this study was inspired by the *EVP* (Capel 2010, 2012), it adopts different methodology. The *EVP* was initially derived from the headwords of a dictionary, and then these were compared to a learner corpus. The wordlist developed in the current study, in contrast, draws on learner corpora first in order to obtain more learner-centric data. Furthermore, the words in the present study are not selected based on simple frequency alone, but using additional statistical techniques.

4.1 Learner data

As the source of the new wordlist, the author uses two learner written corpora: the *Corpora do Português Língua Estrangeira* (*PLE*)[§], collected by the Linguistic Centre of the University of Lisbon; and the *Corpus de Produções Escritas de Aprendentes de PL2* (*PEAPL2*)[**], collected by the General and Applied Linguistics Centre of the University of Coimbra. The writing samples in both corpora were obtained using the same writing task. Summaries of the two corpora are shown below (Table.2).

	PLE	*PEALP2*
Total words	Approx. 70,500	Approx. 120,000
Number of subcorpora	471	546
Number of mother tongues of the informants	27	39
Learning environment	Studying in home country	Studying at University of Coimbra

Table 2: Summaries of the *PLE* and the *PEAPL2*

All the informants for both corpora responded to a can-do self-assessment questionnaire for placement, and then each informant was classified by *CEFR* proficiency. Whereas in the *PEAPL2* the informants were classified into five levels (A1, A2, B1, B2, and C1), in the *PLE* the informants are classified into just three levels (A1-A2, B1-B2, and C1-C2). To allow both these corpora to be used for the present study, the author integrated the A1 and A2 levels, as well as the B1 and B2 levels, of the PEAPL2. Following this, the author integrated the PLE and the re-classified PEAPL2 into a final, three-level learner corpora by integrating the A1-A2, the B1-B2 and the C1-C2 levels from both corpora.. Before integrating the corpora, the author examined the proximity between the subcorpora by clustering[††], in order to validate the integration. The analysis confirmed the proximity between subcorpora of the same proficiency level, with the exception of the B1-B2 level of the *PEAPL2* (Figure 1). This seemed to be because of the larger size of this subcorpus compared to the others. Despite this finding, the author did not discard the B1-B2 subcorpus of the *PEAPL2* from this study.

[§] www.clul.ul.pt/pt/recursos/314-corpora-of-ple
[**] www.uc.pt/fluc/rcpl2
[††] The author used *Seagull-Stat*, a statistical Add-in for Microsoft Excel®.
www7b.biglobe.ne.jp/~hayakari/

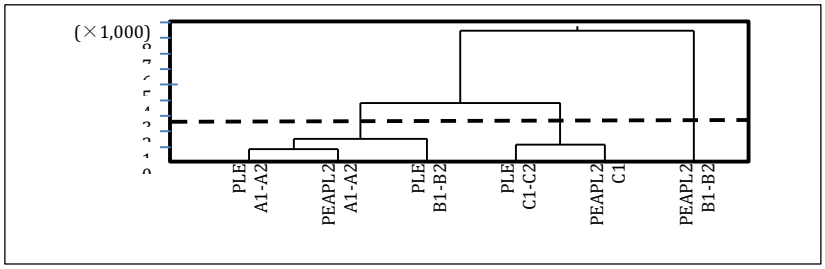

Figure 1: Result of clustering subcorpora

4.2 Results of the learner corpus analysis

The results of the learner corpus analysis are shown below (**Table 3**).

	A1-A2	B1-B2	C1-C2	Total
Words	59932	101758	16793	178483
Lemmas	2198	3961	1286	4865
Including foreign words and orthographical mistakes				

Table 3: Number of words and lemmas in integrated subcorpora

Across the whole learner corpus, the top 100 words accounted for 72 per cent of the total text, the top 1000 words accounted for 95 per cent, and the top 2000 words accounted for 98 per cent. Restricting the analysis to the A-level subcorpora, the top 500 words accounted for 95 per cent of the total text in the remaining subcorpus, the top 600 words account for 96 per cent, and the top 700 account for 97 per cent. Given these results, the author set the cut off points for each proficiency level at the top 500th, 1000th, and 1500th words.

4.3 Determining final for word placement

Through the analysis described in the previous section, the raw frequency of the words was obtaind. However, dividing the wordlist into proficiency levels by simply using the cut off points may not be adequate. For this reason, additional statistical techniques were used to determine the final word placement..

First, the top 500 words of the A-level subcorpus were automatically classified as the Elementary level words, because statistical analysis confirmed that there was a high correlation (r=.90) between the top 500 words of the A-level subcorpus and the whole corpus.

Second, to classify the Intermediate and Advanced level words, the remaining 1000 words from the top 1500 words of the whole corpora were analysed using the chi-square test. The words found to be significant for the A-level were added to the existing list of 500 Elementary words; similarly, those significant for the B- and C-levels were added to the existing Intermediate and Advanced lists respectively.

Third, the remaining word (which had not been significant at any proficiency level according to the chi-square test) were classified. Those which had been produced within the whole corpus more than 10 times were classified as Intermediate words, and those produced less than 10 times were classified as Advanced words.

Finally, Arabic numerals (*1*, *2*, *3*, ...), alphabet letters (*a*, *b*, *c*, ...), foreign words not naturalized in Portuguese (*new*, *españa*), and some orthographical mistakes (*tambem*, *portugues*) were all excluded.

4.4 Native speaker baseline data: The LMCPC

For the native speaker baseline data, the author chose the *Léxico Multifuncional Computorizado do Português Contemporâneo* (*LMCPC*), a wordlist derived from the *CORLEX* (a subcorpus of the *CRPC*). This list was selected because of its balance and the number of words included. The *LMCPC* is frequency list of more than 25000 words. The rank and frequency of the words in the *LMCPC* were annotated manually on the learner wordlist obtained as a result of the steps described in in 4.1-4.3.

5 Findings

The final wordlist derived from the two learner corpora is presented in the Appendix. In this section, findings related to this list are detailed.

The correlation (mutual information score) between the word frequencies in the learner corpora and in the *LMCPC* was examined. For the elementary level vocabulary, high mutual information score was obtained (r=.98); lower scores were obtained for the intermediate (r=.30) and the advanced vocabulary (r=.28).

Reviewing the final wordlist qualitatively, the author noted that there were some words that appeared frequently in the learner corpora but much less frequently in the NS corpus. One obvious example concerns the bias in the learner corpora towards European Portuguese vocabulary. For example, *autocarro* (bus), *comboio* (train), and *desporto* (sport) all appeared in the top 500 words of the final word list; meanwhile, their counterparts in Brazilian Portuguese, *ônibus*, *trem*, and *esporte*, did not even appear even in the top 1500 words. Similarly, the frequency of *você* (you)[‡‡], one of the most frequent words in Brazilian Portuguese, was not as highly ranked in the learner use wordlist (n=40, 267th) as in the native speaker corpus. A second type of bias observed in the learner corpus was towards words related to academic life, for example, *universidade* (university), *faculdade* (faculty, college), *república* (a dormitory system in Coimbra), *doutoramento* (doctoral programme), etc. One further bias noted within the final wordlist was caused by the tasks used for data collection. For example, the words *coreano* (Korean), *Edimburgo* (Edinburgh), *Bucareste* (Bucharest), which seemed to have been elicited in the self-presentation task[§§] responses of A1-A2 learners, all appeared in the Elementary level vocabulary list.

Finally, some intuitively basic words such as numerals (*sete* [seven], *nove* [nine]) and months (*janeiro* [January], *outubro* [October]) appeared in the Intermediate or Advanced level. These words would normally be included in textbooks for beginners.

6 Discussion

Let us now consider the second research question: Does native speakers' use of language correspond to learners' use? The high correlation between the frequencies of the Elementary words in the learner corpora and the NS baseline corpus (the *LMCPC*) suggests that certain basic vocabulary is common to both L2 learners and NS. On the other hand, one interpretation of the low correlation between the frequencies of the Intermediate and Advanced words in the learner corpora and the NS baseline corpus suggests that the most important words for these groups of learners may differ; if so, this may directly answer the second research question.

[‡‡] In Brazilian Portuguese usage, the third-person singular *você* is almost exclusively used to indicate interlocutor, whereas in European Portuguese, *você* and the second-person singular *tu* are optional depending on the relationship with interlocutor.

[§§] One of the optional tasks used for corpus data collection.

Another, perhaps more probable, explanation of the low correlations observed for Intermediate and Advanced words, is that these results may have been due to the cut off points selected in developing the learner word list as well as particular characteristics of the corpora used. In terms of the cut off points, compared to the *English Vocabulary Profile* and the *CEFR-J Wordlist* (see **Table 1**), just 500 words at each proficiency level are too few; this could mean that the level of difficulty of some words, which should really be classified as Elementary, is overestimated and the words are placed into the Intermediate or Advanced levels. More words within each proficiency level are needed due to the size of both the *PLE* and the *PEAPL2*. In terms of the characteristics of the corpora used, the discrepancy between the C-level word frequencies for learners and NS, as well as the biases observed in the previous section, seem to be influenced by the data collection task and the informants' learning environment.

For these reasons, the first research question still requires further study: How much vocabulary should learners learn?

7 Concluding remarks

This paper reports on a pilot study designed to develop a prototype of the *Portuguese Vocabulary Profile*, a wordlist for learners of Portuguese and derived from *CEFR*-based learner corpora. As the wordlist reported in this paper is the pilot version, opportunities for further improvement can be identified as well as some limitations resulting from the corpora used. First, the new wordlist was based solely on written corpora, meaning that there was no consideration of learners' use of spoken language. Secondly, constrained by the proficiency classifications used in the *PLE*, this study adopted a three-level classification; however, it would, of course, be more desirable to follow the original, six-level *CEFR* structure. The use of a three-level classification is also in contrast to the recent trends favouring the use of more proficiency levels; for example, the *CEFR-J* (Tono 2013) adopts 12 proficiency levels in order to reflect the distribution of learners of English in Japan. Finally, the points mentioned in the previous sections (that is, the absence of L2 Brazilian Portuguese words within the learner corpora, the biases caused by characteristics of the corpora used, and the adequateness of the cut off points selected for the word list) are also areas to be improved.

In conclusion, some areas for future study are indicated. First, the wordlist developed in this pilot study should be compared with additional corpus-based wordlist other than the *LMCPC*. Second, as Capel (2010) did for the *English Vocabulary Profile*, the learner wordlist should be compared with textbook corpora and should also be assessed by native speakers. Third, a list of *n*-grams should also be obtained and examined through the methods outlined above. The more sophisticated wordlist is expected to help improve the quality of both textbooks for Portuguese language learners and learner dictionaries in Japan.

References

Aires, P., & Iyanaga, S. (2010/2012). *Verbos Fundamentis do Português.* Kyoto: Kyoto University of Foreign Studies.

Baba, J. (2010). *Comparison of French Textbooks Published in Japan and France.* Tokyo: Tokyo University of Foreign Studies.

Capel, A. (2010). A1-B2 Vocabulary: Insights and Issues Arising from the English Profile Wordlists Project. *English Vocabulary Journal*, 1-11.

Capel, A. (2012). Completing the English Vocabulary Profile: C1 and C2 Vocabulary. *English Vocabulary Journal, 5*(1), 1-14.

Council of Europe. (2001). *Commom European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.

Davis, M., & Preto-Bay, A. R. (2008). *A Frequency Dictionary of Portuguese: Core Vocabulary for Learners*. New York: Routledge.

Ichinose, A., Toida, H., Hayashida, M., & Yoshino, T. (2015). *Shogakukan Dicionário da Língua Portuguesa*. Tokyo: Shogakukan.

Ishikawa, S. (2006). *Vital-series CORPUS 4500 words*. Tokyo: Bun-Ei-Do.

Tono, Y. (2008). *Favourite Corpus 1800 words*. Tokyo: Tokyo Syoseki.

Tono, Y. (2013). *CEFR-J Guidebook*. Tokyo: Taisyukan.

Torigoe, S., & Yamada, M. (2015, November 29). Analysis of the Proficiency Goals of Portuguese Language Textbooks Based on the CEFR-J. Paper presented at the 19th anual meeting of The Japan Association of Foreign Language Education. Tokyo, Japan.

Acknowledgement

This study is supported by *KAKENHI* (*Grant-in-Aid for Young Scientists (B)* of the Japan Society for the Promotion of Science), number 15K16792.

Appendix: Wordlist derived from the learner corpora

elementary level									rank	lemma	English	A1–A2	B1–B2	C1–C2	total	LMCPC rank	LMCPC freq.
rank	lemma	English	A1–A2	B1–B2	C1–C2	total	LMCPC rank	LMCPC freq.	114	ler	to read	74	160	6	240	340	4409
1	o	the	6125	11892	2432	20449	1	2324605	115	conhecer	to know	72	172	15	259	148	8329
2	de	of, from	3425	6271	1228	10924	2	1454101	116	sair	to leave	72	95	9	176	204	6656
3	e	and	2778	3469	575	6822	4	435092	117	pensar	to think	71	110	33	214	151	8140
4	ser	to be	2065	2786	486	5337	6	334978	118	praia	beach	71	56	41	168	928	1693
5	em	in	2031	3237	498	5766	3	490549	119	noite	night	69	142	5	216	161	7790
6	eu	I	1767	1049	113	2929	34	27963	120	comida	food	69	95	26	190	2496	527
7	muito	very, many	1697	2039	289	4025	31	36522	121	castanho	brown	67	5	0	72	3449	401
8	um	a	1450	2522	411	4383	7	323755	122	livro	book	66	198	22	286	355	4182
9	meu	my	1377	1286	128	2791	53	20991	123	então	then, so	66	128	16	210	115	10661
10	que	that, who	1254	3003	555	4812	5	403899	124	há	there is	66	119	25	210		
11	ter	to have	1185	1506	263	2954	13	117212	125	gente	people	65	156	17	238	196	6927
12	gostar	to like	894	872	82	1848	217	6243	126	autocarro	bus	65	38	0	103	2124	653
13	com	with	789	1047	191	2027	11	137880	127	voltar	to return	64	89	10	163	208	6565
14	ir	to go	774	934	119	1827	23	45510	128	até	until, even	63	131	22	216	49	24580
15	não	not, no	731	1329	244	2304	12	149709	129	durante	during	63	124	7	194	119	10209
16	a	the	713	1486	288	2487	8	384393	130	ouvir	to hear	63	64	7	134	213	6448
17	para	to, for	655	1233	196	2084	10	153883	131	comboio	train	63	25	1	89	1296	1197
18	mas	but	534	780	121	1435	20	55588	132	aprender	to learn	62	97	0	159	950	1653
19	estar	to be	429	739	117	1285	18	65520	133	transporte	transportat	62	28	14	104	955	1648
20	mais	more	423	922	171	1516	17	76321	134	sobre	on, about	61	136	29	226	47	21893
21	porque	why, because	402	507	70	979	70	19138	135	encontrar	to meet	60	186	32	278	90	12785
22	fazer	to do, to mak	384	733	74	1191	22	47273	136	preferir	to prefer	60	77	8	145	1178	1326
23	me	me	369	460	40	869	73	15822	137	dizer	to say	59	171	26	256	26	40117
24	ano	year	366	301	36	703	27	36764	138	chegar	to arrive	59	92	10	161	97	12042
25	casa	house	360	370	29	759	92	12731	139	ainda	still, yet	58	148	26	232	38	26128
26	por	for, by	356	593	138	1087	9	174329	140	magro	thin	58	0	0	58	2640	488
27	também	also, too	350	660	60	1070	35	27763	141	melhor	better	57	180	23	260	219	9494
28	amigo	friend	334	360	37	731	354	4510	142	importante	important	56	118	12	186	231	6188
29	ele	he	319	349	56	724	29	31193	143	espanhol	Spanish	56	46	0	102	937	2581
30	morar	to live	291	241	12	544	2206	627	144	avião	airplane	55	11	1	67	1364	1128
31	se	oneself	287	746	201	1234	16	111682	145	tu	you	52	57	8	117	506	2947
32	tempo	time, weather	285	565	54	904	68	17878	146	mundo	world	51	115	37	203	162	7765
33	poder	can	265	716	149	1130	25	47602	147	beber	to drink	50	66	6	122	908	1733
34	cidade	city	253	779	121	1153	130	9258	148	cinema	movie	50	69	3	122	933	1686
35	querer	to want	246	269	54	569	55	20593	149	filho	son	50	44	1	95	171	7644
36	todo	all	244	474	74	792	33	30694	150	preto	black	50	12	0	62	1998	1208
37	como	how, like	237	518	94	849	19	69312	151	banho	bath	50	4	0	54	1335	1158
38	pessoa	person	237	444	98	779	79	14421	152	cultura	culture	49	167	18	234	82	13809
39	chamar	to call	229	91	13	333	178	7518	153	mês	month	49	72	8	129	199	6862
40	este	this	219	444	76	739	15	80789	154	andar	to walk, floo	49	57	20	126	234	6596
41	quando	when	216	495	43	754	58	23460	155	escola	school	49	75	2	126	149	8296
42	ver	to see	204	280	47	531	43	23398	156	féria	holiday	49	50	14	113	1282	1211
43	viajar	to travel	204	120	8	332	1988	709	157	madrid	Madrid	49	17	0	66		
44	portugal	Portugal	191	193	19	403			158	assim	thus, like th	48	122	21	191	144	18717
45	agora	now	186	237	23	446	78	14684	159	esperar	to hope	48	111	11	170	287	5057
46	isso	that	185	385	74	644	59	19657	160	jantar	dinner, to h	48	53	1	102	1059	1699
47	dia	day	184	337	53	574	46	21896	161	inglês	English	48	45	1	94	1519	1885
48	português	Portuguese	181	186	13	380	118	12656	162	lugar	place	47	158	23	228	170	7645
49	ficar	to stay	180	294	35	509	66	18031	163	fim	end, purpos	47	81	5	133	175	7560
50	coisa	thing	179	449	53	681	77	14657	164	hora	hour	47	64	10	121	111	10776
51	pai	father	178	140	14	332	233	5932	165	itália	Italy	47	51	0	98		
52	haver	to have	177	483	61	721	24	42560	166	querido	dear	47	33	1	81	3485	595
53	outro	other	172	410	68	650	39	26075	167	primeiro	first	46	86	23	155	32	29577
54	coimbra	Coimbra	172	222	15	409			168	filme	film	46	59	8	113	536	2776
55	estudar	to study	169	132	7	308	688	2229	169	esse	that	45	97	24	166	30	28960
56	grande	big	167	219	51	437	57	21363	170	desporto	sport	45	49	0	94	109	10871
57	ou	or	161	479	53	693	21	51034	171	problema	problem	44	117	20	181	155	8050
58	falar	to speak	161	206	32	399	120	10260	172	tão	so, as	44	92	19	155	133	9049
59	sempre	always	159	290	52	501	76	14837	173	vir	to come	44	81	6	131	67	17884
60	vez	time	157	325	51	533	42	23535	174	comigo	with me	44	69	8	121	1288	1203
61	irmão	brother	146	33	0	179	554	2718	175	verão	Summer	44	65	8	117	1080	1439
62	vida	life	142	430	85	657	84	13495	176	curso	course	44	48	5	97	590	2537
63	bom	good	142	260	52	454	117	11731	177	apanhar	to catch	44	33	7	84	903	1745
64	dois	two	139	163	16	318	37	27440	178	rapaz	young man	44	27	1	72	639	2375
65	estudo	study	139	71	0	210	386	3828	179	vinte	twenty	44	4	0	48	814	2083
66	bem	well	131	203	31	365	71	18785	180	junto	together	43	66	4	113	400	4531
67	achar	to think	127	298	31	456	281	5181	181	faculdade	faculty, coll	43	38	0	81	3280	360
68	visitar	to visit	127	117	8	252	1104	1413	182	momento	moment	42	86	14	142	239	5817
69	aqui	here	126	205	12	343	98	11903	183	antes	before	42	54	13	109	173	7623
70	universidade	university	123	116	10	249	1746	835	184	desejo	desire	42	29	5	76	968	1631
71	nos	us	122	188	32	342	207	6566	185	saudade	longing	41	55	7	103	2083	671
72	depois	after	121	167	21	309	48	21772	186	três	three	41	62	0	103	102	12064
73	bonito	beautiful	118	79	15	212	1118	1492	187	beijo	kiss	41	51	5	97	2952	419
74	estudante	student	116	109	2	227	969	1665	188	lisboa	Lisbon	41	47	1	89		
75	já	already	114	230	28	372	28	35541	189	estrangeiro	foreign	41	20	5	66	1085	2467
76	comer	to eat	114	115	22	251	583	2603	190	mesmo	same	40	220	33	293	101	14886
77	trabalhar	to work	112	77	7	196	293	4986	191	rua	road	40	105	11	156	373	3961
78	país	country	110	192	45	347	91	12737	192	escrever	to write	40	75	9	124	326	4533
79	seu	his, her, their	108	325	79	512	14	86106	193	lindo	beautiful	40	25	8	73	1910	740
80	cabelo	hair	106	13	0	119	998	1563	194	veneza	Venice	40	1	0	41		
81	passar	to pass	105	185	36	326	72	15936	195	lá	there	39	113	10	162	131	13188
82	livre	free	105	190	1	296	686	2354	196	aula	class, lesso	39	97	8	144	1412	1084
83	pequeno	little, small	105	163	19	287	168	8605	197	cozinha	kitchen	39	50	1	90	1188	1315
84	perto	close	100	163	6	269	562	2666	198	apartamento	apartment	39	29	7	75	3555	318
85	alto	high	99	47	8	154	319	5752	199	tomar	to take	38	38	9	85	251	5624
86	olho	eye	97	26	0	123	228	6030	200	italiano	Italian	38	42	0	80	1683	1355
87	saber	to know	96	175	18	289	50	21882	201	sala	room	38	4	0	42	479	3040
88	nós	we	96	95	44	235	108	11030	202	lembrar	to remembe	37	88	7	132	366	4033
89	mãe	mother	96	39	0	135	336	4439	203	simpático	kind	37	21	1	59	3150	398
90	semana	week	92	138	7	237	221	6141	204	normalmente	normally	36	58	5	99	1208	1296
91	viver	to live	90	362	57	509	186	7163	205	adorar	to adore	36	49	9	94	2552	512
92	trabalho	work	89	127	11	227	99	11675	206	feliz	happy	36	45	8	89	1114	1393
93	música	music	89	89	0	178	600	2491	207	professor	proffessor	36	41	2	79	473	3082
94	só	only, just	88	278	27	393	52	23104	208	sem	without	35	143	41	219	54	20764
95	onde	where	88	221	36	345	61	19296	209	exemplo	example	35	94	18	147	163	7761
96	nosso	our	87	168	88	343	203	6672	210	restaurante	restraurant	35	82	2	119	1669	884
97	pouco	little	86	146	18	250	267	11972	211	acabar	to finish	35	60	4	99	176	7532
98	viagem	trip	86	48	5	139	616	2449	212	espanha	Spain	35	60	0	95		
99	família	family	85	59	15	159	260	5442	213	jardim	garden	35	52	2	89	815	1928
100	quarto	room, fourth	83	42	0	125	523	4071	214	tarde	afternoon	35	39	3	77	401	5880
101	algum	some	82	198	26	306	51	21397	215	olá	hi	35	11	3	49	7369	99
102	língua	language	82	90	16	188	736	2119	216	sol	sun	34	24	1	59	729	2131
103	tudo	all	81	196	31	308	75	15482	217	nadar	to swim	34	10	7	51	4847	198
104	mim	me	81	150	7	238	305	4816	218	entre	between	33	104	20	157	40	25655
105	carro	car	80	85	13	178	508	2926	219	isto	that	33	101	17	151	125	9989
106	dinheiro	money	79	50	6	135	399	3696	220	dar	to give	33	84	27	144	36	27265
107	comprar	to buy	77	48	15	140	559	2641	221	nunca	never	33	75	7	115	141	8493
108	centro	centre	76	96	17	189	381	3898	222	mar	sea	33	42	14	89	389	3770
109	meio	means, middle	76	65	14	155	187	11058	223	verde	green	33	21	2	56	1087	1840
110	novo	new	75	191	22	288	65	18929	224	gordo	fat	33	2	0	35	3310	570
111	jogar	to play	75	49	13	137	603	2489	225	teu	your	32	104	8	144	421	3529
112	diferente	different	74	189	21	284	242	5764	226	velho	old	32	63	1	96	338	5597
113	te	you	74	152	16	242	417	3549	227	divertir	to have fun	32	60	3	95	3424	338

rank	lemma	English	A1–A2	B1–B2	C1–C2	total	LMPCP rank	LMCPC freq.	rank	lemma	English	A1–A2	B1–B2	C1–C2	total	LMPCP rank	LMCPC freq.
228	futebol	football	32	29	2	63	556	2652	342	domingo	Sunday	17	10	1	28	941	1673
229	dançar	to dance	32	26	1	59	2539	516	343	calor	hot	17	7	2	26	1191	1311
230	porto	port	32	15	5	52	1693	868	344	roupa	cloth	17	7	0	24	1308	1185
231	loja	shop	31	74	7	112	1160	1343	345	figueira	Figueira	17	6	0	23	11471	45
232	nada	nothing	31	65	7	103	138	8921	346	aquele	that	16	82	10	108	56	20526
233	bastante	enough	31	54	11	96	517	3381	347	causa	cause	16	59	18	93	333	4449
234	televisão	television	31	28	0	59	794	1984	348	cozinhar	to cook	16	55	0	71	3423	338
235	avô	grandfather	31	26	1	58	1499	1018	349	mau	bad	16	42	6	64	456	3228
236	azul	blue	31	2	0	33	1074	1680	350	correr	to run	16	41	3	60	358	4158
237	menos	less	30	107	18	155	232	10674	351	antigo	old	16	31	10	57	258	5594
238	tanto	so much	30	90	21	141	288	6431	352	produto	product	16	13	25	54	371	3970
239	preciso	precise	30	60	11	101	469	3102	353	aí	there	16	16	0	32	205	6642
240	interessante	interesting	30	47	14	91	1477	1033	354	foz	Foz	16	4	0	20	14524	28
241	dormir	to sleep	30	36	0	66	826	1901	355	quem	who	15	31	20	66	88	12978
242	festa	party	29	125	11	165	695	2204	356	cheio	full	15	38	12	65	611	2538
243	desde	since	29	72	6	107	136	8788	357	levar	to take	15	36	12	63	122	10145
244	talvez	maybe, perha	29	46	2	77	268	5316	358	agradável	comfortable	15	40	1	56	2404	555
245	fora	outside	29	35	8	72	570	5207	359	segundo	second, acc	15	30	11	56	172	14335
246	concerto	concert	28	39	3	70	1553	980	360	capital	capital	15	28	3	46	412	3635
247	futuro	future	28	32	8	68	539	4108	361	bicicleta	bicycle	15	24	0	39	4817	200
248	parque	park	28	36	1	65	1284	1209	362	logo	soon	15	17	1	33	272	5297
249	quatro	four	28	26	1	55	224	6445	363	enorme	enormous	15	10	5	30	758	2080
250	erasmus	Erasmus	27	54	0	81			364	ilha	island	15	8	7	30	680	2263
251	frio	cold	27	30	6	63	1152	2319	365	literatura	literature	15	12	3	30	1717	853
252	rio	river	27	24	7	58	494	2983	366	normal	normal	15	15	0	30	1070	1470
253	comercial	comercial	27	14	2	43	782	2016	367	nome	name	15	13	1	29	215	6353
254	parte	part	26	106	20	152	81	14010	368	familiar	familiar	15	11	1	27	1549	1482
255	quase	almost	26	79	5	110	166	7698	369	namorar	to date stea	15	4	0	19	3728	299
256	alemanha	Germany	26	67	0	93			370	sul	South	14	43	11	68	689	2224
257	ajudar	help	26	58	7	91	642	2369	371	tranquilo	calm	14	43	3	60	2906	430
258	verdade	truth	26	51	8	85	344	4338	372	contar	to tell, to co	14	34	5	53	256	5521
259	mulher	woman	26	25	1	52	121	10168	373	decidir	to decide	14	26	9	49	466	3122
260	cão	dog	26	13	12	51	1298	1194	374	aproveitar	to make us	14	25	2	41	703	2193
261	namorado	boyfriend	26	17	1	44	4379	234	375	caro	expensive	14	21	2	37	1544	989
262	bairro	neighborhood	25	174	0	199	1407	1087	376	cerveja	beer	14	19	2	35	3443	335
263	actividade	activity	25	55	3	83	316	4713	377	breve	brief	14	17	2	33	764	2057
264	tipo	type	25	47	7	79	218	6230	378	doce	sweet	14	17	1	32	2129	816
265	alemão	German	25	29	1	55	1048	2100	379	amar	to love	14	17	0	31	1271	1216
266	frança	France	25	21	0	46	21759	10	380	bolo	cake	14	14	0	28	2996	410
267	você	you	25	15	0	40	533	2787	381	cansar	to get tired	14	12	0	26	3971	273
268	mudar	to move	24	107	12	143	525	2823	382	gato	cat	14	6	0	20	2290	593
269	além	beyond	24	83	30	137	194	7059	383	aldeia	village	13	95	3	111	802	1949
270	maior	larger	24	72	8	104	128	10487	384	café	coffee	13	67	1	81	1165	1339
271	norte	North	24	54	10	88	683	2249	385	oportunidade	opportunity	13	52	15	80	876	1802
272	ti	you	24	41	7	72	1586	947	386	história	history, sto	13	48	16	77	276	5251
273	parar	to stop	24	42	2	68	648	2359	387	possibilidade	possibility	13	50	11	74	607	2476
274	europa	Europe	24	41	2	67			388	acontecer	to happen	13	46	12	71	257	5519
275	jogo	game	24	36	6	66	179	7416	389	animal	animal	13	23	35	71	662	2584
276	casar	to marry	24	10	0	34	890	1764	390	sim	yes	13	38	10	61	235	5922
277	coreano	Korean	24	1	0	25	11445	88	391	compra	purchase	13	45	1	59	1146	1357
278	usar	to use	23	58	36	117	435	3394	392	inverno	Winter	13	39	1	53	1194	1306
279	começar	to begin	23	65	10	98	123	10132	393	baixa	downtown	13	31	0	44	2469	534
280	inglaterra	England	23	48	1	72			394	vivo	alive	13	26	5	44	743	2317
281	tocar	to touch	23	46	0	69	734	2126	395	passear	to go for a	13	28	1	42	2277	598
282	teatro	theatre	23	45	0	68	1029	1512	396	barato	cheap	13	25	0	38	2807	513
283	francês	French	23	27	1	51	749	2968	397	saudável	healthy	13	13	7	33	3652	308
284	si	himself	23	23	0	46	359	4179	398	frente	front	13	16	1	30	291	5050
285	colega	colleague	23	22	0	45	1332	1160	399	bucareste	Bucharest	13	15	0	28		
286	paris	Paris	23	9	0	32			400	setembro	September	13	11	2	26	785	2009
287	cada	each	22	151	32	205	112	10689	401	favorito	favourite	13	11	0	24	3203	497
288	lado	side	22	91	23	136	116	10288	402	metro	metro, subv	13	7	3	23	540	2752
289	conseguir	to succeed	22	49	12	83	159	7872	403	felicidade	happiness	13	7	2	22	2202	628
290	entrar	to enter	22	22	8	52	195	6931	404	janela	window	13	8	0	21	763	2060
291	manhã	morning	22	28	1	51	440	3355	405	julho	July	13	2	1	16	1261	1228
292	brasil	Brazil	22	11	0	33			406	turquia	Turkey	13	3	0	16		
293	curto	short	22	10	0	32	595	2516	407	comprido	long	13	2	0	15	2281	596
294	pusan	Busan	22	0	9	31			408	qual	which	12	55	24	91	104	11247
295	hoje	today	21	68	22	111	93	12443	409	embora	although	12	68	4	84	274	6134
296	costumar	to have the h	21	63	2	86	1628	910	410	nem	nor	12	43	11	66	95	13296
297	bar	bar	21	60	2	83	2249	608	411	opinião	opinion	12	25	10	47	457	3215
298	computador	computer	21	22	5	48	961	1637	412	nascer	to be born	12	29	4	45	497	2965
299	terminar	to finish	21	17	1	39	599	2496	413	longo	long	12	17	11	40	220	6170
300	cinco	five	21	13	0	34	285	5351	414	publicidade	advertising	12	28	0	40	1941	726
301	dever	must, should	20	68	35	123	45	22034	415	tentar	try	12	28	0	40	341	4403
302	sentir	to feel	20	60	21	101	182	7271	416	brasileiro	Brazilian	12	23	0	35	1518	1418
303	jovem	young	20	51	14	85	418	4693	417	carne	meat	12	18	0	30	949	1658
304	lhe	to him	20	39	9	68	85	16416	418	praticar	to practice	12	18	0	30	1256	1235
305	exame	exam	20	16	0	36	1554	979	419	londres	London	12	16	0	28		
306	cantar	to sing	20	41	3	64	926	1756	420	unidos	united	12	8	7	27		
307	barco	ship, boat	20	6	2	28	834	1884	421	levantar	to raise	12	12	0	24	480	3037
308	berlim	Berlin	20	7	0	27			422	branco	white	12	9	1	22	507	3962
309	mora	?	20	3	0	23	10995	49	423	porquê	why	12	8	0	20	1948	1097
310	campo	field	19	342	67	428	349	4238	424	coreia	Korea	12	6	0	18		
311	sonho	dream	19	76	0	95	884	1779	425	taxi	taxi	12	1	0	13		
312	maneira	way, manner	19	43	18	80	337	4439	426	existir	to exist	11	83	16	110	181	7303
313	ali	there	19	33	0	52	209	6483	427	próprio	own	11	43	15	69	310	7372
314	estado	state	19	24	8	51	290	5050	428	pois	for, becaus	11	33	15	59	210	10009
315	chinês	Chinese	19	16	1	36	1579	1436	429	típico	tipical	11	31	10	52	2176	635
316	ás	ace	19	6	1	26	13142	34	430	cá	here	11	35	2	48	416	3684
317	coser	boil	19	5	0	24	5398	166	431	calmo	calm	11	30	1	42	3078	396
318	marido	husband	19	4	0	23	832	1888	432	ideia	idea	11	21	8	40	315	4721
319	solteiro	bachelor, sing	19	4	0	23	6510	235	433	república	republic	11	29	0	40	2817	450
320	último	last, latest	18	59	12	89	69	23762	434	vinho	wine	11	23	2	36	659	2320
321	difícil	difficult	18	49	14	81	432	3465	435	sítio	site	11	18	6	35	1109	1406
322	monumento	monument	18	36	3	57	2219	622	436	medo	fear	11	22	0	33	653	2331
323	brincar	to play, to jok	18	37	1	56	1877	758	437	quente	hot	11	16	2	29	978	1628
324	famoso	famous	18	24	11	53	1627	924	438	desejar	to desire	11	16	1	28	1020	1525
325	ganhar	to get, to win	18	20	9	47	279	5183	439	tirar	to take out,	11	14	2	27	470	3101
326	quanto	how much	18	21	6	45	275	6974	440	pão	bread	11	11	1	23	1147	1357
327	forte	strong	18	25	1	44	410	3938	441	seguir	to follow	11	11	0	22	211	6482
328	menino	boy	18	26	0	44	657	2703	442	alegre	happy	11	10	0	21	3505	325
329	final	final	18	17	0	35	309	6399	443	algarve	Algarve	11	10	0	21		
330	cedo	soon, early	18	13	0	31	1143	1359	444	nacionalidade	nationality	11	10	0	21	4145	256
331	alcalá	Alcalá	18	8	0	26			445	característica	characteris	11	8	0	19	622	2431
332	criança	child	17	72	12	101	249	5652	446	estranho	strange	11	8	0	19	1171	1556
333	especialmente	especially	17	45	12	74	1223	1281	447	sistema	system	11	8	0	19	247	5690
334	longe	far	17	49	2	68	532	3089	448	morada	residence	11	3	0	14	3860	284
335	possível	possible	17	38	11	66	261	5600	449	primo	cousin	11	3	0	14	2185	651
336	próximo	next	17	29	2	48	246	7133	450	almoçar	to have lun	11	2	0	13	3422	338
337	pé	foot	17	22	7	46	348	4263	451	busan	Busan	11	0	0	11		
338	cor	color	17	19	1	37	529	2799	452	henares	Henares	11	0	0	11		
339	baixo	low, short	17	17	0	34	582	4480	453	instituto	institute	11	0	0	11	2999	410
340	olhar	to look	17	14	2	33	254	7043	454	liso	straight	11	0	0	11	3951	276
341	plano	plan	17	10	2	29	484	3244	455	telefonar	to telephon	11	0	0	11	2591	504

rank	lemma	English	A1–A2	B1–B2	C1–C2	total	LMPCP rank	LMPCP freq.
456	montanha	mountain	10	34	41	85	1483	1031
457	natureza	nature	10	50	19	79	667	2291
458	uso	use	10	42	18	70	1023	1521
459	natal	Christmas	10	52	5	67	5491	264
460	sobretudo	above all	10	46	7	63	411	3717
461	cultural	cultural	10	45	5	60	544	2728
462	facto	fact	10	42	6	58	165	7722
463	único	unique	10	35	10	55	317	5558
464	portanto	therefore	10	35	8	53	320	4609
465	são	sound, safe	10	33	2	45	5475	163
466	principal	principal	10	23	4	37	298	4943
467	rápido	rapid	10	20	7	37	873	1855
468	sorte	luck	10	24	0	34	1323	1178
469	prazer	pleasure	10	20	1	31	1464	1058
470	preparar	to prepare	10	13	8	31	491	2991
471	palavra	word	10	19	1	30	244	5786
472	uma	a	10	20	0	30	304	4844
473	estação	station, seasc	10	12	3	25	580	2569
474	japonês	Japanese	10	14	1	25	2823	690
475	seis	six	10	12	1	23	441	3556
476	almoço	lunch	10	8	0	18	1688	871
477	feira	fair	10	7	0	17	1970	716
478	nova	new	10	2	3	15	1026	1518
479	mesa	table	10	4	0	14	565	2633
480	conduzir	to lead	10	1	1	12	868	1819
481	escuro	dark	10	1	0	11	1374	1348
482	biológico	biological	10	0	0	10	3350	350
483	louro	blond	10	0	0	10	3952	714
484	ondular	to wave	10	0	0	10	6502	123
485	york		10	0	0	10		
486	agricultura	agriculture	9	4	8	21	1229	1272
487	minha	my	9	3	2	14		
488	ténis	tennis	9	4	0	13	3243	366
489	economia	economy	9	3	0	12	100	11607
490	publicar	to publish	9	3	0	12	753	2091
491	sábado	Saturday	9	3	0	12	742	2114
492	escócia	Scotland	9	2	1	12		
493	particular	particular	8	3	0	11	821	2189
494	new		9	1	1	11		
495	discoteca	nightclub, disc	7	2	0	9	3332	352
496	direito	right, law	8	1	0	9	114	11788
497	doutoramento	doctorial degr	8	1	0	9	6764	114
498	vermelho	red	6	0	3	9	1187	1489
499	tio	uncle	8	0	1	9	881	1788
500	ana	Ana	9	0	0	9		
501	loiro	blond	9	0	0	9		
502	reciclagem	recycle	9	0	0	9	6810	113
503	vampiro	vampire	9	0	0	9	12782	37
504	corredor	corridor, runn	7	1	0	8	1790	909
505	elegante	elegant	7	1	0	8	3723	299
506	falador	speaker	7	1	0	8	17927	24
507	junho	June	7	1	0	8	1279	1212
508	voleibol	volleyball	7	1	0	8	5064	184
509	pasta	folder	6	1	0	7	2153	642
510	sapato	shoe	6	1	0	7	2069	677
511	inteligente	intelligent	6	1	0	7	2877	437
512	óculo	glass	7	0	0	7	3336	352
513	basquetebol	basketball	7	0	0	7	4043	266
514	navio	ship	7	0	0	7	1654	891
515	az	?	6	0	0	6		
516	desarrumar	disorganize	6	0	0	6	14733	27
517	lestat	Lestat	6	0	0	6		
518	letónia	Latvia	6	0	0	6		
519	vicenza	Vicenza	6	0	0	6		
520	actor	actor	5	0	0	5	1098	1419
521	cm		5	0	0	5	3202	372
522	encaracolar	to curl	5	0	0	5	13582	32
523	indonésia	Indonesia	5	0	0	5		
524	luca	Luca	5	0	0	5		
525	queens	Queens	5	0	0	5		
526	sílvia	Silvia	5	0	0	5		
527	tavira	Tavira	5	0	0	5		
528	tibet	Tibet	5	0	0	5		
529	ucrânia	Ucrine	5	0	0	5		
530	vietnam	Vietname	5	0	0	5		
531	vivenda	dwelling	5	0	0	5	6279	130

intermediate level								
rank	lemma	English	A1–A2	B1–B2	C1–C2	total	LMPCP rank	LMPCP freq.
532	região	region	5	78	9	92	188	7154
533	desvantagem	disadvantag	0	52	12	64	4348	236
534	qualquer	any	9	46	7	62	105	11236
535	fácil	easy	2	48	10	60	684	2246
536	forma	form	5	44	9	58	96	12252
537	situação	situation	8	42	8	58	191	7129
538	barulho	noise	3	51	2	56	2951	419
539	tradicional	traditional	6	42	7	55	1044	1488
540	zona	zone	3	45	6	54	237	5843
541	escolher	to choose	5	48	0	53	552	2659
542	apesar	despite	5	42	4	51	269	5290
543	claro	clear	8	40	3	51	490	4873
544	deixar	to leave, to	1	40	10	51	87	12979
545	relação	relation	9	34	8	51	183	7221
546	povo	people	4	42	4	50	431	3418
547	diferença	difference	7	34	8	49	717	2168
548	geral	general	4	43	1	48	433	4327
549	certeza	certainty	7	38	2	47	817	1924
550	maioria	majority	7	31	9	47	460	3198
551	realidade	reality	3	37	6	46	610	2469
552	certo	certain	3	34	9	46	584	5567
553	espaço	space, roon	2	34	9	45	83	13522
554	habitante	inhabitant	0	34	10	44	1193	1306
555	mal	badly	7	31	2	40	321	4847
556	faltar	to lack	6	26	8	40	613	2451
557	abraço	hug	5	31	3	39	3844	285
558	algo	something	7	30	2	39	1150	1381
559	prato	plate	6	27	6	39	1609	933
560	melhorar	improve	2	32	4	38	1317	1180
561	igreja	church	6	31	0	37	975	1615
562	pessoal	personal	3	26	8	37	772	2699
563	esquecer	to forget	4	29	3	36	519	2858
564	social	social	4	30	1	35	200	6858
565	nenhum	no	1	33	0	34	628	4719
566	tradição	tradition	0	32	2	34	973	1619
567	razão	reason	0	28	6	34	303	4846
568	especial	special	4	28	1	33	579	2792
569	encontro	meeting, da	3	22	8	33	403	3657
570	oferecer	to offer	1	29	2	32	634	2392
571	geralmente	generally	6	26	0	32	1260	1228
572	organizar	to organize	5	26	1	32	783	2012
573	demais	too mach	6	25	0	31	3705	638
574	permitir	to allow	2	21	6	29	227	6046
575	necessidade	necessity	3	19	7	29	407	3615
576	infelizmente	unfortunate	1	27	0	28	2597	502
577	clima	clime	1	24	3	28	1158	1345
578	criar	to create	3	19	6	28	214	6439
579	imaginar	to imagine	0	27	0	27	1203	1298
580	depender	to depend	2	25	0	27	971	1628
581	sé	See	3	23	1	27	12587	38
582	paisagem	landscape	4	21	2	27	1450	1054
583	tranquilidade	tranquility	4	21	2	27	3711	301
584	hábito	habit	0	23	3	26	1468	1039
585	dentro	within	3	22	1	26	240	5790
586	estrangeiro	foreigner	0	22	4	26		
587	considerar	to consider	0	20	6	26	140	8509
588	contrário	contrary	2	18	6	26	643	2880
589	participar	to participa	9	15	2	26	674	2279
590	comum	common	1	23	1	25	631	2641
591	dificuldade	difficulty	2	20	3	25	575	2589
592	minuto	minute	2	22	0	24	426	3478
593	realizar	to make ha	0	22	2	24	295	4951
594	vizinho	neighbor	2	21	1	24	1447	1536
595	vontade	willingness	2	21	1	24	737	2119
596	ensinar	to teach	1	21	2	24	1529	1000
597	bocado	piece	5	18	1	24	640	2373
598	bélgica	Belguim	0	23	0	23		
599	perceber	to understa	2	20	1	23	592	2527
600	resto	rest	3	19	1	23	545	2728
601	urbano	urban	1	19	3	23	1108	1434
602	nível	level	1	18	4	23	226	6051
603	monte	mount	2	16	5	23	1503	1018
604	acreditar	to believe	0	21	1	22	649	2358
605	crescer	to grow	1	20	1	22	784	2009
606	complicar	to complica	2	19	1	22	4333	237
607	estilo	style	1	19	2	22	1025	1519
608	ocupar	to occupy	2	17	3	22	661	2305
609	discoteca	nightclub, d	5	16	1	22		
610	respeito	respect	1	18	2	21	636	2384
611	situar	to situate	2	16	3	21	655	2325
612	século	century	1	18	1	20	270	5289
613	floresta	forest	2	17	1	20	1829	786
614	horóscopo	horoscope	0	19	0	19	13623	32
615	económico	economic	0	18	1	19	306	4798
616	significar	to mean	1	17	1	19	919	1710
617	relaxar	to relax	2	16	1	19	12759	37
618	preocupar	to worry ab	1	15	3	19	1610	933
619	caminhar	to walk	6	13	0	19	1955	720
620	apenas	only	2	16	0	18	113	10825
621	positivo	positive	1	16	1	18	1219	1529
622	turista	tourist	1	15	2	18	2123	654
623	mensagem	message	0	13	5	18	1263	1226
624	mudança	change	0	17	0	17	865	1823
625	graça	grace	0	14	3	17	982	1594
626	comunidade	community	5	12	0	17	774	2041
627	autor	author	1	15	0	16	553	2655
628	distinto	distinct	0	15	1	16	2152	642
629	fogo	fire	0	15	1	16	867	1822
630	vila	town	1	14	1	16	1127	1381
631	marcar	to mark	0	14	2	16	395	3733
632	facilidade	happiness	0	11	5	16	2062	679
633	par	pair	8	8	0	16	983	1592
634	verdadeirame	truly	0	15	0	15	2351	573
635	emigrante	emmigrant	0	15	0	15	3230	427
636	paixão	passion	0	15	0	15	1512	1014
637	aberto	open	1	13	1	15	550	2854
638	concordar	to agree	1	13	1	15	2029	692
639	sentido	meaning	1	13	1	15	302	4884
640	local	place	0	12	3	15	44	25119
641	celebrar	to celebrat	0	14	0	14	1894	747
642	desenvolver	to develop	0	12	2	14	478	3049
643	ritmo	rhythm	0	11	3	14	1270	1217
644	bulgária	Bulgary	0	13	0	13		

rank	lemma	English	A1–A2	B1–B2	C1–C2	total	LMPCP rank	LMCPC freq.	rank	lemma	English	A1–A2	B1–B2	C1–C2	total	LMPCP rank	LMCPC freq.
645	crime	crime	0	13	0	13	974	1616	759	entender	to understa	5	17	1	23	500	2963
646	real	real	0	13	0	13	745	2926	760	conhecimento	knowledge	5	16	2	23	586	2547
647	valor	value	0	13	0	13	167	7691	761	emprego	job	2	15	6	23	696	2200
648	preço	price	5	8	0	13	443	3345	762	macau	Macau	9	9	5	23		
649	arquitectura	architecture	0	12	0	12			763	museu	Museum	4	18	0	22	1795	806
650	companhia	company	0	12	0	12	987	1583	764	piscina	pool	6	16	0	22	2652	484
651	contrariar	to contradict	0	12	0	12	2746	463	765	completamen	completely	5	15	2	22	1053	1481
652	evento	event	0	12	0	12	2612	496	766	crer	to believe	8	14	0	22	1213	1292
653	sentimento	feeling	0	12	0	12	1170	1335	767	ligar	to link	5	16	0	21	430	3421
654	partilhar	to share	0	11	1	12	2028	693	768	flor	flower	4	14	3	21	1013	1544
655	morrer	to die	2	10	0	12	362	4110	769	junta	council	3	14	4	21	1376	1119
656	contudo	however	0	10	2	12	806	1944	770	cama	bed	8	12	1	21	843	1863
657	população	population	0	10	2	12	330	4501	771	triste	sad	9	11	1	21	1283	1258
658	variedade	variety	0	10	2	12	1952	722	772	pena	shame	3	16	1	20	1048	2786
659	alimentar	to feed, relate	0	11	0	11	1446	1408	773	cabeça	head	4	15	1	20	318	4663
660	biblioteca	library	0	11	0	11	2828	447	774	suficiente	sufficient	6	14	0	20	1299	1375
661	pesar	to weigh	0	11	0	11	2253	664	775	fresco	fresh, cool	5	14	1	20	2002	914
662	provocar	to provoke	0	11	0	11	465	3130	776	aborrecer	to bore	7	13	0	20	4544	219
663	simplesmente	simply	0	11	0	11	1568	957	777	trocar	to exchang	5	13	2	20	1605	935
664	tecnologia	technology	0	11	0	11	1397	1099	778	diverso	diverse	2	13	5	20	728	2132
665	arranjar	to arrange, to	0	10	1	11	1024	1520	779	rir	to laugh	6	12	2	20	948	1661
666	exposição	exposition	0	10	1	11	779	2029	780	maravilhoso	marvelous	6	10	4	20	3591	402
667	apreciar	to appreciate	1	9	1	11	1787	810	781	alguém	someone	3	16	0	19	711	2175
668	ponte	bridge	1	9	1	11	1845	778	782	prédio	building	3	16	0	19	1703	862
669	stress	stress	0	10	0	10	5351	169	783	legume	vegetable	4	15	0	19	4395	233
670	alimentação	food	0	10	0	10	2022	694	784	edifício	building	3	15	1	19	690	2221
671	avançar	to advance	0	10	0	10	730	2130	785	recordar	to recall	3	14	2	19	654	2325
672	guitarra	guitar	0	10	0	10	4007	271	786	amor	love	2	14	3	19	455	3237
673	manchester	Manchester	0	10	0	10			787	alta	high	7	12	0	19	1459	1046
674	movimento	movement	0	10	0	10	452	3267	788	experimentar	to try	6	12	1	19	2256	606
675	nacional	national	0	10	0	10	255	5745	789	físico	physical	4	11	4	19	966	1632
676	obter	to obtain	0	10	0	10	409	3593	790	jornal	newspaper	4	14	0	18	135	8800
677	particularmen	particularly	0	10	0	10	1293	1199	791	pronto	ready	4	14	0	18	1144	1717
678	pra	for	0	10	0	10	4287	241	792	simples	simple	2	14	2	18	566	2677
679	rapidamente	rapidly	0	10	0	10	1384	1112	793	fechar	to close	5	13	0	18	621	2432
680	torre	tower	0	10	0	10	3270	362	794	copo	cup	4	13	1	18	1742	837
681	época	period, epoch	0	10	0	10	446	3314	795	sentar	to sit	4	13	1	18	633	2393
682	negro	black	6	4	0	10	962	2182	796	foto	photo	5	12	1	18	1996	705
683	encanto	fascination	1	7	1	9	3172	379	797	costa	coast	3	12	3	18	2901	430
684	área	area	1	7	1	9	229	5972	798	receber	to receive	2	12	4	18	248	5682
685	emigrar	to emigrate	0	8	0	8	5759	150	799	antigamente	anciently	4	11	3	18	3107	391
686	gordura	fat	0	8	0	8	3359	349	800	humano	human	4	11	3	18	392	3825
687	musical	musical	1	7	0	8	1980	737	801	janeiro	January	9	9	0	18	461	3185
688	carácter	caracter	2	6	0	8	854	1840	802	américa	America	7	8	3	18		
689	romance	novel	2	6	0	8	1830	786	803	semestre	semester	2	15	0	17	3710	301
690	virgem	virgin	0	7	0	7	5137	198	804	chover	to rain	3	14	0	17	2723	470
691	espectacular	spectacular	0	6	0	6	3425	338	805	seguro	safe	3	14	0	17	1606	1351
692	pertencer	to belong to	0	6	0	6	841	1872	806	mandar	to send, to	4	13	0	17	526	2821
693	hobby	hobby	1	5	0	6	15581	24	807	fruta	fruit	5	12	0	17	2741	464
694	linha	line	1	5	0	6	363	4105	808	rock	rock	5	12	0	17	4583	217
695	braga	Braga	2	4	0	6			809	mão	hand	4	12	1	17	160	7806
696	cruz	cross	2	4	0	6	5783	149	810	fantástico	fantastic	2	11	4	17	3958	306
697	atividade	activity	0	5	0	5			811	demasiado	too much	7	10	0	17	1587	1102
698	doze	twelve	0	5	0	5	2020	695	812	piano	piano	7	10	0	17	2848	442
699	isolar	to isolate	0	5	0	5	2101	661	813	aeroporto	airport	8	9	0	17	2169	636
700	salsa	parsley	0	5	0	5	2691	476	814	iorque	York	7	9	1	17		
701	traje	suit	0	5	0	5	4725	207	815	cima	top	6	9	2	17	475	3062
702	conta	against	1	4	0	5	277	5193	816	ontem	yesterday	6	9	2	17	126	9733
703	telemóvel	mobile phone	5	113	59	177	6045	139	817	médico	medical doc	9	8	0	17	428	3655
704	vantagem	advantage	2	82	25	109	946	1666	818	objecto	object	5	2	10	17	551	2661
705	parecer	to seem	3	58	20	81	124	10365	819	impossível	impossible	2	14	0	16	1505	1066
706	homem	man	6	61	11	78	80	14031	820	interessar	to interest	2	14	0	16	932	1689
707	vário	various	5	51	21	77	376	3930	821	aparecer	to appear	3	13	0	16	347	4309
708	tornar	to become	8	40	18	66	184	7182	822	barcelona	Barcelona	3	13	0	16		
709	principalment	mainly	1	18	46	65	954	1649	823	fotografia	photograph	3	13	0	16	1234	1268
710	precisar	to need	7	40	16	63	495	2976	824	condição	condition	2	13	1	16	266	5321
711	caso	case	1	36	20	57	107	12009	825	serviço	service	2	12	2	16	212	6474
712	ar	air	6	27	18	51	346	4313	826	tua	your	2	12	2	16		
713	qualidade	quality	0	28	14	42	439	3367	827	rural	rural	1	12	3	16	1735	862
714	continuar	to continue	4	26	12	42	142	8488	828	idade	age	5	11	0	16	380	3905
715	entanto	however	5	24	11	40	283	5159	829	perigoso	dangerous	4	11	1	16	1794	807
716	sociedade	society	1	24	13	38	86	13184	830	dúvida	doubt	3	10	3	16	520	2850
717	perder	to lose	2	20	12	34	245	5750	831	alar	relating to	7	9	0	16	8029	85
718	tal	such	2	18	10	30	150	10011	832	lago	lake	6	9	1	16	2562	509
719	papel	paper, role	1	14	14	29	313	4743	833	combinar	to combine	5	9	2	16	2246	609
720	descobrir	to discover	2	20	6	28	652	2339	834	mondego	Mondego	4	9	3	16		
721	necessário	necessary	1	15	8	24	327	4563	835	felizmente	fortunately	2	13	0	15	2974	414
722	ponto	point	2	18	3	23	157	7998	836	questão	question	2	13	0	15	262	5391
723	manter	to maintain	1	13	9	23	216	6301	837	histórico	historical	3	12	0	15	797	2004
724	porém	however	0	12	10	22	492	2990	838	largo	large	3	12	0	15	1125	2153
725	electrónico	electronic	0	14	7	21	3199	373	839	presente	present	3	12	0	15	645	2990
726	tratar	to treat	0	11	10	21	241	5774	840	árvore	tree	2	12	1	15	811	1934
727	desenvolvime	development	0	11	6	17	311	4753	841	início	beginning	1	12	2	15	390	3764
728	gastar	to waste	0	10	7	17	1368	1127	842	actualmente	currently	4	11	0	15	936	1684
729	experiência	experience	8	30	2	40	571	2604	843	conjunto	set	4	11	0	15	413	4062
730	vista	view	8	30	2	40	378	3922	844	texto	text	4	11	0	15	505	2953
731	trazer	to bring	6	25	9	40	367	4000	845	ninguém	no one	4	10	1	15	335	4441
732	altura	height	9	26	4	39	252	5623	846	popular	popular	2	10	3	15	1231	1748
733	grupo	group	9	29	0	38	152	8138	847	ajuda	help	6	9	0	15	1257	1233
734	falta	lack	9	21	8	38	405	3623	848	facilitar	to facilitate	1	12	1	14	1749	834
735	descansar	to rest	6	23	8	37	2669	480	849	imenso	immense	1	12	1	14	1788	1307
736	praça	square	9	25	0	34	2078	672	850	pessoalmente	personally	1	12	1	14	4336	237
737	variar	to vary	8	25	1	34	1437	1066	851	brincadeira	joke	3	10	1	14	3456	333
738	partir	to depart	7	25	2	34	169	7650	852	igual	equal	3	10	1	14	698	2373
739	moderno	modern	8	23	3	34	1148	1392	853	representar	to represen	1	10	3	14	328	4520
740	rico	rich	8	24	1	33	1139	1755	854	comunicar	to communi	5	9	0	14	2040	689
741	modo	mode	6	20	7	33	253	5619	855	pizza	pizza	6	8	0	14		
742	etcétera	et cetera	8	23	1	32			856	portar	to carry	5	8	1	14	5348	169
743	carta	letter	8	22	1	31	573	2596	857	russo	Russian	6	7	1	14	1743	1163
744	assunto	topic	8	21	1	30	701	2194	858	gozar	to enjoy	8	6	0	14	2166	638
745	realmente	really	8	19	2	29	705	2191	859	nata	cream	8	6	0	14	3708	301
746	dança	dance	7	21	0	28	2018	696	860	bruxelas	Brussels	1	12	0	13		
747	procurar	to look for	5	22	0	27	375	3938	861	capaz	capable	1	12	0	13	658	2320
748	passeio	walk	6	21	0	27	1532	999	862	notícia	news	2	11	0	13	892	1763
749	informação	information	6	20	0	26	353	4188	863	fixe	good, cool	1	11	1	13	12351	39
750	supermercad	supermarket	6	20	0	26	4743	206	864	valer	to be worth	1	11	1	13	702	2194
751	internacional	international	7	19	0	26	94	12409	865	ambos	both	2	10	1	13	609	2552
752	cara	face	5	19	1	25	879	1794	866	hospital	hospital	1	10	2	13	985	1585
753	arte	art	7	15	3	25	510	2907	867	castelo	castle	4	9	0	13	2074	674
754	frequentar	to attend	7	15	3	25	2415	551	868	chocolate	chocolate	4	9	0	13	4562	218
755	volta	return	6	18	0	24	332	4461	869	juntar	to gather	4	9	0	13	568	2617
756	calhar	to happen	3	17	4	24	2336	580	870	europeu	European	0	9	4	13	420	3963
757	peixe	fish	8	16	0	24	1003	1561	871	festival	festival	0	9	4	13	2533	518
758	caminho	path, way	6	17	0	23	396	3726	872	oeste	West	0	9	4	13	2435	574

rank	lemma	English	A1–A2	B1–B2	C1–C2	total	LMPCP rank	LMCPC freq.
873	limpo	clean	5	8	0	13	2421	549
874	contactar	to contact	4	8	1	13	1784	812
875	suíça	Switzerland	7	6	0	13		
876	destino	destination	6	6	1	13	965	1633
877	vestir	to wear	5	5	3	13	1079	1440
878	pôr	to put	1	10	1	12	174	7613
879	habituar	to inhabit	3	9	0	12	1808	798
880	mestrado	master's degr	3	9	0	12	10003	58
881	pobre	poor	3	9	0	12	704	2312
882	aventura	adventure	2	9	1	12	2491	528
883	mil	thousand	2	9	1	12	134	8882
884	naturalmente	naturally	2	9	1	12	1433	1072
885	dez	ten	4	8	0	12	498	3171
886	dezembro	December	3	8	1	12	406	3619
887	resolver	to resolve	3	8	1	12	477	3051
888	canadá	Canada	5	7	0	12		
889	confortável	Comfotable	5	7	0	12	4166	254
890	fado	fado	4	7	1	12	2175	635
891	alugar	to rent	6	6	0	12	3482	329
892	checa	Czech	6	6	0	12		
893	estrela	star	5	6	1	12	1111	1397
894	pastelaria	pastry	7	5	0	12	7676	92
895	queima	burning	4	5	3	12	6276	130
896	artigo	article	1	10	0	11	201	6837
897	personalidade	personally	1	10	0	11	1702	862
898	todavia	however	1	10	0	11	1456	1050
899	buscar	to search	2	9	0	11	1449	1055
900	canção	song	2	9	0	11	2302	588
901	banda	band	1	9	1	11	1770	821
902	causar	to cause	1	9	1	11	1155	1347
903	desfrutar	to enjoy	1	9	1	11	8039	85
904	mostrar	to show	1	9	1	11	278	5186
905	nó	knot	0	9	2	11	887	1774
906	empresa	company, firm	3	8	0	11	164	7752
907	engraçado	funny	3	8	0	11	4658	326
908	sete	seven	1	8	2	11	589	2544
909	pedir	to ask	3	7	1	11	294	4953
910	romano	Roman	3	7	1	11	1783	1333
911	sensação	sensation	3	7	1	11	1885	753
912	santo	saint	5	6	0	11	1536	1110
913	medicina	medicine	6	5	0	11	2291	593
914	roma	Rome	6	5	0	11		
915	quinta	Thursday, fan	5	5	1	11	2675	480
916	cerca	about	5	4	2	11	605	2979
917	alimento	food	1	9	0	10	1338	1156
918	aumentar	to feed	1	9	0	10	436	3377
919	conclusão	conclusion	1	9	0	10	1091	1426
920	corpo	body	1	9	0	10	282	5167
921	espectáculo	spectacle, sh	1	9	0	10	759	2077
922	fronteira	border, frontie	0	9	1	10	1772	820
923	tema	theme	2	8	0	10	651	2344
924	após	after	1	8	1	10	351	4269
925	distância	distance	1	8	1	10	981	1596
926	acordar	to wake up	0	8	2	10	1300	1192
927	rodear	to spin	0	8	2	10	1658	889
928	memória	memory	3	7	0	10	910	1731
929	escada	stair	2	7	1	10	1721	850
930	mochila	backpack	2	7	1	10	14140	30
931	opção	option	2	7	1	10	1487	1026
932	praga	Prague	2	7	1	10	5543	160
933	turismo	tourism	2	7	1	10	1636	907
934	atrás	behind, ago	1	7	2	10	687	2230
935	bebida	drink	4	6	0	10	3349	350
936	depressa	quickly	4	6	0	10	1616	926
937	parir	to give birth	4	6	0	10	7417	98
938	verdadeiro	true	2	6	2	10	660	2437
939	letras	science of hu	5	5	0	10		
940	áustria	Austria	5	5	0	10		
941	serra	moutain	4	5	1	10	1681	876
942	escolha	choice	3	5	2	10	1042	1490
943	seguinte	following	3	5	2	10	434	4065
944	construir	to construct	5	4	1	10	485	3020
945	descrever	to describe	5	4	1	10	1303	1191
946	regressar	to return	5	4	1	10	707	2188
947	rússia	Russia	6	3	1	10		
948	vento	wind	6	2	2	10	723	2150

advanced level								
rank	lemma	English	A1–A2	B1–B2	C1–C2	total	LMPCP rank	LMCPC freq.
949	terra	land, earth	6	24	33	63	206	6585
950	ambiente	environmen	6	25	24	55	720	2283
951	comunicação	communica	9	23	12	44	624	2426
952	planeta	planet	0	0	44	44	1963	718
953	contacto	contact	2	14	26	42	731	2130
954	as	the	9	15	14	38		
955	aspecto	aspect	5	22	10	37	464	3134
956	natural	natural	3	15	16	34	511	3046
957	subir	to go up	7	14	11	32	472	3090
958	utilizar	to utilize	8	10	14	32	301	4887
959	lixo	trash	5	4	23	32	2118	655
960	finalmente	finally	8	15	7	30	902	1747
961	enquanto	while, wher	3	16	8	27	202	6837
962	frequentemen	frequently	2	15	9	26	1771	820
963	saúde	health	4	13	9	26	543	2738
964	japão	Japan	2	12	12	26		
965	água	water	6	6	12	24	146	8340
966	público	public	2	13	8	23	110	18339
967	mercado	market	4	11	7	22	300	4890
968	útil	useful	2	10	10	22	1798	817
969	polícia	police	1	7	14	22	830	1895
970	compreender	to compreh	6	7	8	21	542	2740
971	puro	pure	1	11	8	20	1414	1135
972	china	China	6	5	9	20	13809	31
973	solução	solution	3	6	10	19	635	2392
974	quantidade	quantity	1	7	10	18	914	1719
975	imigrante	imigrate	1	6	11	18	5382	185
976	deitar	to lie down	2	5	10	17	744	2111
977	reduzir	to reduce	0	0	17	17	951	1653
978	através	through	1	9	6	16	236	5867
979	transformar	to transforr	1	9	6	16	606	2484
980	conversar	to talk	5	7	4	16	1903	745
981	energia	energy	5	5	6	16	576	2588
982	temperatura	temperatur	5	4	7	16	1090	1428
983	contente	content	5	8	2	15	2736	465
984	produzir	to produce	0	7	7	14	630	2409
985	respirar	to breathe	0	7	7	14	2344	576
986	geografia	geography	1	6	7	14	3858	284
987	vício	addiction	0	6	6	12	4059	265
988	lavar	to wash	2	5	5	12	1631	908
989	aparelho	device	0	0	12	12	1427	1074
990	estimação	estimation	0	0	11	11	10849	50
991	apresentar	to present	0	6	4	10	145	8422
992	negativo	negative	0	5	5	10	1038	1553
993	material	material	2	4	4	10	656	2603
994	globalização	globalizatio	0	0	10	10	7112	105
995	atenção	attention	1	3	5	9	646	2363
996	plantar	to plant	0	1	8	9	4026	269
997	fome	hunger	4	1	3	8	1634	907
998	templo	temple	1	1	6	8	1537	996
999	mgs	?	0	0	8	8		
1000	ah	ah	0	6	1	7	641	2372
1001	polaco	Polish	2	4	1	7	7327	145
1002	extraordinário	extraordina	0	2	5	7	1722	874
1003	ouro	gold	3	1	3	7	938	1682
1004	símbolo	symbol	1	1	5	7	1796	805
1005	amazonas	Amazon	0	1	6	7		
1006	proteger	to protect	0	1	6	7	1556	978
1007	rompuy	Rompuy	0	1	6	7		
1008	varsóvia	Warsaw	2	0	5	7		
1009	expressão	expression	1	2	3	6	770	2048
1010	cortejo	parade	2	1	3	6	4887	195
1011	colaborar	to collabora	0	1	5	6	2872	437
1012	descrição	description	0	1	5	6	2627	491
1013	universo	universe	0	1	5	6	2070	676
1014	viciar	to addict	1	0	5	6	8592	76
1015	extraterrestre	extraterres	0	0	6	6	13401	33
1016	guam	Guam	0	0	6	6		
1017	lata	can, tin	0	0	6	6	3224	369
1018	objectivo	objective	0	4	1	5	393	4010
1019	proximidade	nearness	0	4	1	5	2629	491
1020	hospitaleiro	hospitable	0	2	3	5	13881	31
1021	imaginação	imagination	0	2	3	5	2406	555
1022	ocidental	ocidental	0	2	3	5	1463	1044
1023	planear	to plan	0	1	4	5	3806	291
1024	spray	spray	0	0	5	5	17710	18
1025	utilização	utilization	0	0	5	5	851	1847
1026	mente	mind	0	9	0	9	3871	283
1027	moda	fashion	0	9	0	9	1410	1086
1028	residência	residence	0	9	0	9	1766	824
1029	signo	zodiac sign	0	9	0	9	4266	243
1030	telefone	telephone	0	9	0	9	1492	1024
1031	varna	Varna	0	9	0	9		
1032	voz	voice	0	9	0	9	408	3614
1033	boi	bull	1	8	0	9	3098	393
1034	clube	club	1	8	0	9	488	3001
1035	contaminação	contaminati	1	8	0	9	6053	138
1036	estádio	studium	1	8	0	9	1889	751
1037	granada	Granada	1	8	0	9	8566	76
1038	membro	member	1	8	0	9	458	3202
1039	menor	smaller	1	8	0	9	859	2105
1040	agitação	agitation	0	8	1	9	4127	257
1041	importância	importance	0	8	1	9	574	2596
1042	passo	step	0	8	1	9	549	2720
1043	pensamento	thought	0	8	1	9	1258	1231
1044	convidar	to invite	2	7	0	9	1250	1242
1045	era	era	2	7	0	9	3146	383
1046	iniciar	to start	2	7	0	9	602	2490
1047	pior	worse, wors	2	7	0	9	1408	1695
1048	segurança	security, sa	2	7	0	9	546	2708
1049	passado	past	1	7	1	9	292	6300
1050	procura	search	1	7	1	9	754	2090
1051	consistir	to consist c	0	7	2	9	1458	1048
1052	religião	religion	0	7	2	9	1343	1155
1053	companheiro	companion	3	6	0	9	1426	1076
1054	programa	programme	3	6	0	9	388	3787
1055	cadeira	chair	2	6	1	9	1355	1138
1056	excelente	excellent	2	6	1	9	1438	1065
1057	frequência	frequency	2	6	1	9	1121	1387
1058	raro	rare	2	6	1	9	1275	1288
1059	surf	surf	2	6	1	9	6817	113
1060	chuva	rain	1	6	2	9	855	1839
1061	nota	note	1	6	2	9	1017	1530

rank	lemma	English	A1–A2	B1–B2	C1–C2	total	LMPCP rank	LMCPC freq.	rank	lemma	English	A1–A2	B1–B2	C1–C2	total	LMPCP rank	LMCPC freq.
1062	activo	active	4	5	0	9	1728	1319	1176	interior	interior, ins	0	6	1	7	748	2680
1063	profissão	profession	4	5	0	9	1763	827	1177	cuidar	to take car	2	5	0	7	1824	790
1064	internet	Internet	3	5	1	9	2171	636	1178	leite	milk	2	5	0	7	1517	1009
1065	refeição	refreshment	3	5	1	9	2265	603	1179	aperfeiçoar	to perfect	1	5	1	7	5685	153
1066	tarefa	task	3	5	1	9	1134	1368	1180	dividir	divide	1	5	1	7	1117	1390
1067	aquilo	that	2	5	2	9	263	5383	1181	terror	terror	1	5	1	7	2676	480
1068	neve	snow	2	5	2	9	1732	844	1182	benefício	to benefit	0	5	2	7	2088	668
1069	bacalhau	cod fish	6	3	0	9	2813	450	1183	deus	God	0	5	2	7	1174	1329
1070	bilhete	ticket	6	3	0	9	2047	684	1184	disciplina	discipline	0	5	2	7	2121	654
1071	edimburgo	Edinburgh	6	3	0	9			1185	hotel	hotel	2	5	0	7	1309	1184
1072	massa	mass, paste	6	3	0	9	829	1895	1186	aniversário	anniversary	3	4	0	7	2447	539
1073	chorar	to cry	5	3	1	9	1232	1270	1187	bolonha	Bologna	3	4	0	7		
1074	reciclar	to recycle	0	0	9	9	8886	72	1188	checo	Czech	3	4	0	7	6759	205
1075	aprendizagem	learning	0	8	0	8	3893	280	1189	materno	maternal	3	4	0	7		
1076	batata	potato	0	8	0	8	2343	576	1190	moto	motorbike	3	4	0	7	4909	194
1077	concluir	to conclude	0	8	0	8	587	2540	1191	partida	departure, g	3	4	0	7	878	1796
1078	dieta	diet	0	8	0	8	5664	154	1192	comércio	commerce	1	4	2	7	880	1788
1079	emigração	emigration	0	8	0	8	4553	219	1193	elemento	element	1	4	2	7	391	3759
1080	esquiar	to ski	0	8	0	8			1194	perguntar	to ask	1	4	2	7	482	3028
1081	exagerar	to exagerate	0	8	0	8	5265	173	1195	primavera	Spring	0	4	3	7	2779	456
1082	extremo	extreme	0	8	0	8	2136	933	1196	força	force, powe	3	4	0	7	264	5356
1083	finlândia	Finland	0	8	0	8			1197	jogging	jogging	3	4	0	7	25564	6
1084	fogueira	bonfire	0	8	0	8	4719	207	1198	fato	fact, suit	0	4	3	7	2482	530
1085	goa	Goa	0	8	0	8			1199	céu	sky, heaver	4	3	0	7	1036	1502
1086	imagem	image	0	8	0	8	238	5822	1200	munique	Munich	4	3	0	7		
1087	inscrever	to register	0	8	0	8	2107	660	1201	tese	thesis	4	3	0	7	2677	480
1088	instrumento	instrument	0	8	0	8	1012	1545	1202	belo	beautiful	3	3	1	7	1422	1573
1089	poeta	poet	0	8	0	8	1211	1294	1203	chá	tea	2	3	2	7	2339	579
1090	responsabilida	responsibility	0	8	0	8	870	1817	1204	coração	heart	2	3	2	7	691	2213
1091	acostumar	to get used to	1	7	0	8	10609	52	1205	investigação	investigatio	4	3	0	7	889	1766
1092	disco	disk	1	7	0	8	1849	776	1206	interesse	interest	3	3	1	7	427	3467
1093	ideal	ideal	1	7	0	8	2401	1003	1207	importar	to import, t	1	3	3	7	1221	1286
1094	luz	light	1	7	0	8	415	3553	1208	aguar	to water	5	2	0	7	6281	129
1095	obviamente	obviously	1	7	0	8	2241	612	1209	cambiar	to exchange	5	2	0	7	23074	8
1096	reparar	to fix, to atte	1	7	0	8	1451	1054	1210	vos	(OE)Iyou	5	2	0	7	3037	404
1097	tento	judgement, go	1	7	0	8	4032	268	1211	fevereiro	February	5	2	0	7	1008	1550
1098	visita	visit	1	7	0	8	677	2269	1212	grego	Greek	5	2	0	7	1706	1745
1099	assistir	to attend, to	0	7	1	8	812	1932	1213	faro	Faro	5	1	1	7		
1100	conveniente	convenient	0	7	1	8	3261	363	1214	actual	actual	0	6	0	6	402	3665
1101	número	number	0	7	1	8	180	7406	1215	afastar	to withdraw	0	6	0	6	808	1941
1102	trânsito	traffic	0	7	1	8	1439	1065	1216	básico	basic	0	6	0	6	1618	922
1103	ai	oh	2	6	0	8	1292	1285	1217	catedral	cathedral	0	6	0	6	5127	180
1104	madeira	tree	2	6	0	8	1073	1444	1218	dono	owner	0	6	0	6	1101	1414
1105	marca	brand name, r	2	6	0	8	1167	1339	1219	espada	sword	0	6	0	6	3100	393
1106	natação	swimming	2	6	0	8	5717	152	1220	flandres	Flanders	0	6	0	6		
1107	nove	nine	2	6	0	8	1119	1401	1221	heidelberg	Heidelberg	0	6	0	6		
1108	subida	rise	2	6	0	8	1491	1024	1222	holanda	Holland	0	6	0	6		
1109	estreito	straight	1	6	1	8	2053	853	1223	iaido	iai–do	0	6	0	6		
1110	infância	childhood	1	6	1	8	1822	791	1224	impresso	printed	0	6	0	6	6671	210
1111	lembrança	memory, souv	1	6	1	8	3682	304	1225	ligação	connection	0	6	0	6	700	2196
1112	prático	practice, prac	1	6	1	8	2262	773	1226	passatempo	hobby	0	6	0	6	8643	75
1113	geográfico	geographic	0	6	2	8	2479	532	1227	perigo	danger	0	6	0	6	1192	1308
1114	influência	influence	0	6	2	8	801	1954	1228	peça	piece, play	0	6	0	6	596	2512
1115	rede	net, network	0	6	2	8	581	2558	1229	princípio	beginning	0	6	0	6	437	3373
1116	aveiro	Aveiro	3	5	0	8			1230	protagonista	protagonist	0	6	0	6	3526	323
1117	equipa	team	3	5	0	8	222	6138	1231	revista	journal	0	6	0	6	1246	1244
1118	pagar	to pay	3	5	0	8	342	4388	1232	secundário	secondary	0	6	0	6	2007	841
1119	polónia	Poland	3	5	0	8			1233	seja	whether	0	6	0	6	3548	320
1120	duro	hard	2	5	1	8	1209	1344	1234	somente	only	0	6	0	6	2453	539
1121	festejar	to celebrate	2	5	1	8	4073	263	1235	óbvio	obvious	0	6	0	6	2674	480
1122	responder	to respond	2	5	1	8	414	3557	1236	adulto	adult	1	5	0	6	2316	774
1123	totalmente	totally	2	5	1	8	1582	948	1237	alma	soul	1	5	0	6	719	2162
1124	ultimamente	lately	2	5	1	8	4437	229	1238	apreender	to compreh	1	5	0	6	2717	471
1125	cujo	whose	1	5	2	8	612	2460	1239	colocar	to put, to c	1	5	0	6	361	4138
1126	beleza	beauty	0	5	3	8	1389	1107	1240	coro	chorus	1	5	0	6	3138	384
1127	tokyo	Tokyo	0	5	3	8			1241	costume	custom	1	5	0	6	1615	926
1128	alentejo	Alentejo	4	4	0	8			1242	desculpa	excuse	1	5	0	6	3099	433
1129	amanhã	tomorrow	4	4	0	8	665	2404	1243	linguagem	language	1	5	0	6	1490	1024
1130	clara	Clara	4	4	0	8	3079	396	1244	lingüístico	linguist	1	5	0	6		
1131	equipar	to equip	4	4	0	8	3809	290	1245	malta	group, gang	1	5	0	6	3572	317
1132	habitualmente	usually	4	4	0	8	2497	527	1246	motivo	motive	1	5	0	6	907	1734
1133	limpar	to clear	4	4	0	8	1861	769	1247	acompanhar	to accompa	0	5	1	6	449	3293
1134	sujo	dirty	4	4	0	8	3136	452	1248	atrair	to attract	0	5	1	6	1978	713
1135	cumprimento	greeting, fulfil	3	4	1	8	2327	582	1249	ave	bird	0	5	1	6	2585	504
1136	quer	whether	2	4	2	8	496	2969	1250	caracterizar	to characte	0	5	1	6	1329	1167
1137	escritor	writer	5	3	0	8	1545	988	1251	comparação	comparison	0	5	1	6	2396	557
1138	namorada	girlfriend	4	3	1	8			1252	dispor	to arrange	0	5	1	6	800	1954
1139	campismo	camp	6	2	0	8	8430	78	1253	erro	error	0	5	1	6	1034	1505
1140	®ord®	?	5	2	1	8			1254	estudantil	student	0	5	1	6	6433	125
1141	abrir	to open	0	7	0	7	273	5271	1255	factor	factor	0	5	1	6	663	2300
1142	admirar	to admire	0	7	0	7	1977	713	1256	inteiro	entire	0	5	1	6	1201	1313
1143	arrumar	to organize	0	7	0	7	2417	550	1257	nomeadament	namely, esp	0	5	1	6	708	2187
1144	artista	artist	0	7	0	7	1129	1378	1258	paz	peace	0	5	1	6	905	1742
1145	bastar	to be enough	0	7	0	7	1081	1438	1259	pergunta	question	0	5	1	6	1069	1447
1146	chato	bothersome	0	7	0	7	5502	193	1260	profissional	professiona	0	5	1	6	916	2420
1147	doença	illness	0	7	0	7	486	3015	1261	sinaai	Sinaai	0	5	1	6		
1148	escutar	to listen	0	7	0	7	2329	582	1262	termo	term	0	5	1	6	345	4317
1149	ladrão	thief	0	7	0	7	3674	329	1263	treinar	to train	0	5	1	6	2618	494
1150	louco	mad, crazy	0	7	0	7			1264	amizade	friendship	2	4	0	6	2317	584
1151	mexicano	Mexican	0	7	0	7	6770	185	1265	brno	Brno	2	4	0	6		
1152	novembro	November	0	7	0	7	601	2491	1266	gritar	to shout	2	4	0	6	1313	1181
1153	pintar	to paint	0	7	0	7	1484	1030	1267	ida	outbound tr	2	4	0	6	3852	285
1154	porta	door	0	7	0	7	296	4921	1268	muro	wall	2	4	0	6	1862	769
1155	recente	recent	0	7	0	7	805	1945	1269	semelhante	similar	2	4	0	6	943	1720
1156	rematar	to conclude	0	7	0	7	3184	376	1270	som	sound	2	4	0	6	1186	1317
1157	universitário	Relating to ur	0	7	0	7	2440	602	1271	terça	Tuesday	2	4	0	6	6818	113
1158	galiza	Galicia	0	7	0	7			1272	americano	american	1	4	1	6	836	2288
1159	arco	arch	1	6	0	7	2455	538	1273	notar	to notice	1	4	1	6	1137	1366
1160	contra	against	1	6	0	7	137	9158	1274	separar	to separate	1	4	1	6	1452	1054
1161	demorar	to delay	1	6	0	7	1748	834	1275	similar	similar	1	4	1	6	5234	175
1162	desportivo	athletic	1	6	0	7	1281	1211	1276	ensino	teaching	0	4	2	6	850	1851
1163	diferenciar	to differentiat	1	6	0	7	4702	208	1277	mentalidade	mentality	0	4	2	6	4074	263
1164	estágio	stage	1	6	0	7	3362	348	1278	poluição	pollution	0	4	2	6	2488	529
1165	explicar	to explain	1	6	0	7	379	3919	1279	certamente	certainly	3	3	0	6	1670	883
1166	km		1	6	0	7	2162	640	1280	corrida	race	3	3	0	6	1640	902
1167	natura	nature	1	6	0	7	25701	6	1281	cómodo	comfort	3	3	0	6	6076	160
1168	noticiar	to announce	1	6	0	7	3335	352	1282	giro	turn, fun	3	3	0	6	5312	196
1169	origem	origin	1	6	0	7	509	2926	1283	sumo	juice	3	3	0	6	3987	303
1170	portagem	toll	1	6	0	7	2547	514	1284	video	video	3	3	0	6		
1171	acesso	access	0	6	1	7	712	2174	1285	colina	hill	2	3	1	6	3567	317
1172	diversidade	diversity	0	6	1	7	2947	420	1286	evitar	to avoid	2	3	1	6	626	2421
1173	nocturno	nocturnal	0	6	1	7	2816	467	1287	fascinar	to charm	2	3	1	6	5444	164
1174	residente	resident	0	6	1	7	2758	739	1288	industrial	industrial	2	3	1	6	1032	1771
1175	horário	schedule	1	6	0	7	2145	725	1289	outubro	October	2	3	1	6	534	2786

rank	lemma	English	A1–A2	B1–B2	C1–C2	total	LMPCP rank	LMCPC freq.
1290	província	province	1	3	2	6	1555	979
1291	milhão	million	0	3	3	6	198	6956
1292	agosto	August	4	2	0	6	548	2688
1293	euro	Euro	4	2	0	6	1633	907
1294	investigador	investigator	4	2	0	6	2233	615
1295	polo	polo, pole	4	2	0	6	13940	31
1296	roménia	Romania	4	2	0	6		
1297	sobremesa	dessert	4	2	0	6	5679	154
1298	irlanda	Ireland	3	2	1	6		
1299	acordo	agreement	5	1	0	6	280	5182
1300	caçador	hunter	5	1	0	6	3514	343
1301	delicioso	delicious	5	1	0	6	4841	198
1302	detestar	to detest	5	1	0	6	3779	293
1303	gira	crazy	5	1	0	6		
1304	grécia	Greece	5	1	0	6		
1305	sopa	soup	5	1	0	6	1239	1256
1306	preguiçoso	lazy	4	1	1	6	10666	52
1307	abusar	to abuse	0	5	0	5	5752	150
1308	adaptar	to adapt	0	5	0	5	2041	688
1309	adversário	opponent	0	5	0	5	1383	1199
1310	afectar	to affect	0	5	0	5	898	1750
1311	alpes	Alps	0	5	0	5		
1312	animar	to encourage	0	5	0	5	2349	573
1313	baviera	Bavaria	0	5	0	5		
1314	chatear	to annoy	0	5	0	5	6568	120
1315	chefe	chief	0	5	0	5	768	2050
1316	cidadão	citizen	0	5	0	5	931	1689
1317	claramente	clearly	0	5	0	5	1797	804
1318	contradição	contradiction	0	5	0	5	3552	319
1319	dependente	dependent	0	5	0	5	2430	547
1320	dialecto	dialect	0	5	0	5	9796	60
1321	discutir	to discuss	0	5	0	5	1005	1559
1322	educar	to educate	0	5	0	5	3016	408
1323	frase	phrase	0	5	0	5	1430	1074
1324	galinha	hen	0	5	0	5	2882	436
1325	galo	cock	0	5	0	5	5446	164
1326	governo	government	0	5	0	5	620	2437
1327	índia	India	0	5	0	5		
1328	intensivo	intensive	0	5	0	5	4660	211
1329	limitar	to limit	0	5	0	5	1050	1484
1330	luxemburgo	Luxembourg	0	5	0	5		
1331	marítimo	maritime	0	5	0	5	2514	707
1332	mouro	moor	0	5	0	5	4016	287
1333	mundial	worldwide	0	5	0	5	923	1719
1334	oxalá	I hope	0	5	0	5	7885	88
1335	político	political, politi	0	5	0	5	62	18819
1336	porco	pig, pork	0	5	0	5	2109	685
1337	previsão	forecast	0	5	0	5	2324	583
1338	prisão	prison	0	5	0	5	1089	1428
1339	quotidiano	day-to-day	0	5	0	5	3593	511
1340	relacionar	to relate	0	5	0	5	1375	1121
1341	resposta	response	0	5	0	5	567	2630
1342	sanar	to cure	0	5	0	5	13288	34
1343	sobreviver	to survive	0	5	0	5	2216	623
1344	tráfico	trafficking	0	5	0	5	3339	352
1345	umbria	Umbria	0	5	0	5		
1346	violência	violence	0	5	0	5	1351	1144
1347	aceitar	to accept	1	4	0	5	512	2902
1348	botânico	botanical	1	4	0	5	8828	92
1349	canto	corner	1	4	0	5	1122	1384
1350	celas	Celas	1	4	0	5		
1351	continente	continent	1	4	0	5	1623	917
1352	culpa	fault	1	4	0	5	1809	796
1353	curtir	to tan, to enjo	1	4	0	5	6729	115
1354	elevador	elevator, lift	1	4	0	5	5091	182
1355	escalada	escalation	1	4	0	5	6925	110
1356	esforço	effort	1	4	0	5	709	2184
1357	espera	wait, expectat	1	4	0	5	710	2184
1358	esqui	ski	1	4	0	5	10728	51
1359	fonte	fountain, sour	1	4	0	5	424	3484
1360	mexer	to mix, to sha	1	4	0	5	1773	820
1361	numeroso	numerous	1	4	0	5	1777	817
1362	onda	wave	1	4	0	5	1311	1182
1363	reconstrução	reconstructio	1	4	0	5	4568	218
1364	regularmente	regularly	1	4	0	5	4504	223
1365	relativamente	relatively	1	4	0	5	692	2212
1366	teste	test	1	4	0	5	2115	657
1367	treino	training	1	4	0	5	1506	1017
1368	académico	academic	0	4	1	5	4052	416
1369	apetecer	to appetize	0	4	1	5	2238	613
1370	dedicar	to dedicate	0	4	1	5	849	1851
1371	dominar	to dominate	0	4	1	5	1071	1445
1372	emocionante	exciting	0	4	1	5	9633	62
1373	espírito	spirit	0	4	1	5	765	2055
1374	obrigar	to obligate	0	4	1	5	513	2887
1375	oito	eight	0	4	1	5	721	2485
1376	rei	king	0	4	1	5	791	1995
1377	senhorio	landlord	0	4	1	5	5678	154
1378	arroz	rice	2	3	0	5	2544	514
1379	cogumelo	mushroom	2	3	0	5	4097	260
1380	doente	sick	2	3	0	5	1099	2039
1381	frequente	frequent	2	3	0	5	1713	855
1382	madrugada	early morning	2	3	0	5	1674	880
1383	meia	sock	2	3	0	5	4484	269
1384	nervoso	nervous	2	3	0	5	2309	607
1385	partido	political party	2	3	0	5	394	3944
1386	piso	floor	2	3	0	5	2483	530
1387	regular	regular	2	3	0	5	2242	897
1388	renovar	to renew	2	3	0	5	2792	454
1389	tailândia	Thailand	2	3	0	5		
1390	trabalhador	worker	2	3	0	5	681	2301
1391	carnaval	carnival	1	3	1	5	5482	162
1392	continental	continental	1	3	1	5	4100	260
1393	data	date	1	3	1	5	761	2066
1394	dilema	dilemma	1	3	1	5	5617	156
1395	facilmente	easily	1	3	1	5	1859	770
1396	leitura	reading	1	3	1	5	1356	1134
1397	lista	list	1	3	1	5	625	2423
1398	raramente	rarely, seldom	1	3	1	5	2940	422
1399	sucesso	success	1	3	1	5	1162	1341
1400	alegria	pleasure	0	3	2	5	1352	1142
1401	central	central	0	3	2	5	842	2570
1402	fugir	to flee	0	3	2	5	750	2094
1403	encantar	to fascinate	3	2	0	5	4676	210
1404	turma	class	3	2	0	5	4095	261
1405	automóvel	automobile	2	2	1	5	1153	1730
1406	doméstico	demostic	2	2	1	5	1806	799
1407	educação	education	2	2	1	5	139	8606
1408	fitas	ribbon, tape	2	2	1	5	2842	444
1409	linda	Linda	2	2	1	5	19377	14
1410	oceano	ocean	2	2	1	5	3931	277
1411	quilómetro	kilometre	2	2	1	5	816	1928
1412	serenata	serenade	2	2	1	5	15141	26
1413	sobrinho	nephew	2	2	1	5	3499	327
1414	sonhar	to dream	2	2	1	5	1695	867
1415	via	way	2	2	1	5	531	2793
1416	oriental	Oriental	1	2	2	5	1419	1108
1417	eslováquia	Slovakia	4	1	0	5		
1418	horrível	horrible	4	1	0	5	3869	283
1419	tímido	shy	4	1	0	5	4319	271
1420	comentar	to comment	3	1	1	5	1423	1080