

Artigo / Article

Health literacy and language teaching: data-based host language lexicons

Literacia em saúde e ensino de língua: léxicos de língua de acolhimento baseados em dados

Raquel Amaro 

Universidade NOVA de Lisboa, Portugal

raquelamaro@fcsh.unl.pt

<https://orcid.org/0000-0002-4923-7186>

Recebido em: 31/10/2023 | Aprovado em: 10/02/2023

Abstract

Using a corpus-based approach, this paper presents methods and results for assessing, extracting, and describing the core vocabulary relevant to healthcare access among migrant populations. The aim is to bridge the gap between the basic information conveyed to people arriving in Portugal and the materials as well as other lexicographic resources used in language teaching. The work includes identifying available resources and/or sources for compiling the relevant dataset for healthcare access; selecting available tools for corpus inquiry; testing and comparing results from different functionalities and different lexical statistics measures available in the tools; manual filtering of the data; and analyzing the results and the extracted lexicon. The obtained results reflect the organization of the extracted lexicon in subdomains, the organization of the items within each subdomain, the relationship with common vocabulary, and the extraction of authentic examples from the corpus.

Keywords: Lexicons • Host Language Teaching • Corpus-Based • Health Literacy

Resumo

Utilizando uma abordagem baseada em *corpus*, o presente artigo apresenta métodos e resultados relativos à avaliação, extração e descrição do vocabulário nuclear relevante para o acesso a cuidados de saúde da população migrante. O objetivo é diminuir a distância entre a informação essencial passada às pessoas que chegam a Portugal e os materiais e outros recursos lexicográficos usados

no ensino de língua. O trabalho aqui apresentado inclui a identificação de recursos e/ou fontes para a compilação dos conjuntos de dados relevantes no domínio do acesso a cuidados de saúde; a seleção de ferramentas de exploração de *corpus* disponíveis; o teste e a comparação de resultados de diferentes funcionalidades e diferentes medidas de estatística lexical disponíveis nas ferramentas; a filtragem manual dos dados; e a análise dos resultados e do léxico extraído. Os resultados obtidos refletem a organização do léxico extraído em subdomínios, a organização dos itens em cada subdomínio, a relação com o vocabulário comum e a extração de exemplos autênticos do *corpus*.

Palavras-chave: Léxicos • Ensino de Língua de Acolhimento • Abordagem baseada em corpus • Literacia em saúde

Introduction

Health literacy is a vital component of a person's ability to access, comprehend, and use health-related information. According to current international initiatives (CoE, 2023; Healthy People 2030), one of the major goals of such endeavours is to “Eliminate health disparities, achieve health equity, and attain health literacy to improve the health and well-being of all” (Healthy People 2030¹). In some circumstances, however, eliminating disparities starts with taking care of the basic, which involves providing access to the communication code – the language – in a useful, productive, and inclusive way.

For some time now, global research endeavours have been shedding light on the connection between education, literacy levels, and the overall state of an individual's health. This nexus echoes a fundamental truth: one's ability to comprehend, interrogate, and engage effectively with their health is intrinsically linked to one's literacy proficiency (Davis *et al.*, 2006; Nutbeam, 2008; van der Heide *et al.*, 2013). Health literacy, thus, transcends its individual implications. It emerges as a cornerstone of civic engagement, characterized by informed decision-making and the cultivation of autonomy in health prevention and management.

The reflection presented in this paper adds an extra challenge to the task of improving health literacy, in the sense of enabling people to decipher the discourse of healthcare professionals, grasp medication instructions, interpret diagnostic test results, or give the necessary informed consent prior to medical procedures. Based on the specific case of European Portuguese as the host language, this paper focuses on the gap between the basic information to be conveyed to people arriving in Portugal regarding health issues and the teaching materials as well as other lexicographic resources used in language teaching.

¹ Retrieved from <https://health.gov/healthypeople/priority-areas/health-literacy-healthy-people-2030>; last accessed in October 2023.

1 Host language and integration

The host language is the communication language spoken in the country to which people move to live on a more permanent basis; it is the language used in a specific geopolitical territory for everyday and professional communication. It can correspond to official or national languages and has been recognized for many years as an essential part of receiving newly arrived people, involving the establishment of training structures dedicated to host language training (CoE, 2007, p. 22). Language skills have been considered a form of host-country-specific human capital in economics since the early 1980s (Carliner, 1981; McManus; Gould; Welch, 1983), and several studies have clearly established the negative consequences of language barriers in healthcare (Jaeger *et al.*, 2019). Effective language training is a key factor in promoting socio-economic integration, as well as access to rights and services (CoE, 2018).

As discussed in Amaro *et al.* (2022, p. 185-186), integration consists of a process involving “the host society, which should create the opportunities for the immigrant people full economic, social, cultural, and political participation. It also involves adaptation by the migrant people who are supposed to have rights but also responsibilities in relation to their new country of residence (EC, 2020, p. 1-2)”. Therefore, health literacy is definitely a part of integration. At the same time, almost all action and policies for integrating and including migrant people include language as mediating communication. Thus, learning the language(s) of the host country is a crucial step to successfully integrate and thrive autonomously (Elsod; Marques, 2019, p. 9).

In many contexts, migrant people must quickly adapt linguistically to the new environment. Often, this appropriation takes place outside academic structures and it is promoted by institutions and organizations responsible for managing and assisting in the integration of migrant and refugee people, such as the United Nations High Commissariat for Migrations, or governmental and non-governmental institutions such as Caritas, the National Entity for Health Regulation, and Social Security, to name a few.

For host language teaching to be effective and useful to both the migrant people and the host community, it must correspond to the needs of its target audience and to the goals of successful integration into the host community. This means providing relevant and pragmatic information regarding many aspects of daily life, such as housing, education, employment, or health, which are not universal but depend on specific laws, regulations, and systems.

However, the current state of the art remains that many available language teaching and learning programs and materials are overly broad and general, not covering the specific needs related to host languages (Cooke; Roberts, 2007; Bryers; Winstanley; Cooke, 2014, p. 38). For instance, vocabulary domains such as family, house, leisure, or food are usually covered in general language teaching curricula, and, although relevant for every person, they do not address other immediate and essential needs of migrant or refugee people (Elsod; Marques, 2019). This is also the case for Portuguese as a Foreign Language (PFL), in which the

communicative situations addressed in PFL teaching materials such as handbooks, audio files for training, among others² (e.g., vacations, general cultural trends, cooking recipes, museum visits, etc.) can be an extra obstacle to the integration of people arriving in the country. In these materials, relevant language use is insufficient or unapproachable to migrant learners, who wish to achieve autonomous participation in the community as soon as possible, and there are few or no parts of these materials dedicated to health. In this specific case, health literacy goes hand in hand with language teaching.

2 Goals and methodology

As briefly presented in the previous section, courses for host language should entail specifically designed curricula and materials. Corpus linguistics can make an important contribution to bridging this gap, as it allows for the extraction and analysis of language features from documents envisioning integration. The analysis of real texts targeting migrant people can, thus, pave the way to data-based host language lexicons that can be focused and enhanced for specific purposes, such as health literacy.

Starting from the assumption that teaching a host language for integration requires the assessment, extraction, and description of the relevant core vocabulary, and that corpus exploitation is a widely accepted practice in the development of structured lexical resources (Beloso, 2015; Lindemann, 2013), this paper presents work on data-based host language lexicons for health literacy.

Our goal is to present the methods and results of extracting lexical information from health-related texts targeting migrant people and how this data can inform lexicons for host language teaching that contribute to health literacy, eliminate disparities, and improve the health and well-being of all.

2.1 General methodology

Lexicon extraction relies on several principles, such as standards of frequency of occurrence, ranges of likelihood of co-occurrence, application of linguistic filters for isolating surface forms, application of exclusion lists, meaning condensation, among others. When dealing with large amounts of data, linguists and lexicographers welcome automation (Perez; Rizzo, 2014; Lang; Schneider; Suchowolec, 2018). Regarding the treatment of specific domains, frequency is relevant only when contrasted with common language: frequencies of specialized and general language corpora are compared, and only items with significant relative frequency differences are considered (Drouin, 2003; Barbero; Amaro, 2020), or wordlists extracted from specialized corpora are contrasted with pre-existing terminological databases,

² See MATERIALS_PFL_2023 corpus description in section 3.1.

used as a “golden standard” (Lang; Schneider; Suchowolec, 2018). This means that, in addition to data for specific purposes, reference data are necessary.

Considering the goals aimed at the extraction of the relevant lexicon involves the following general tasks:

(1) Compilation and/or selection of corpora: identifying available resources and/or sources for compiling the relevant dataset for the issue at hand, namely information on health issues targeting migrant people and PFL teaching materials (handbooks, exercises, etc.).

(2) Tool selection and data extraction: selecting available tools for corpus exploitation that allow for concordances, collocations, and corpora comparisons.

(3) Data extraction and analysis: testing and comparing results from different functionalities and from different lexical statistics measures available in the tools.

(4) Data manual filtering: selecting the relevant items from the lists obtained automatically using native speaker knowledge and external knowledge sources.

(5) Analysis of the results and of the extracted lexicon.

The next sections present these steps in more detail.

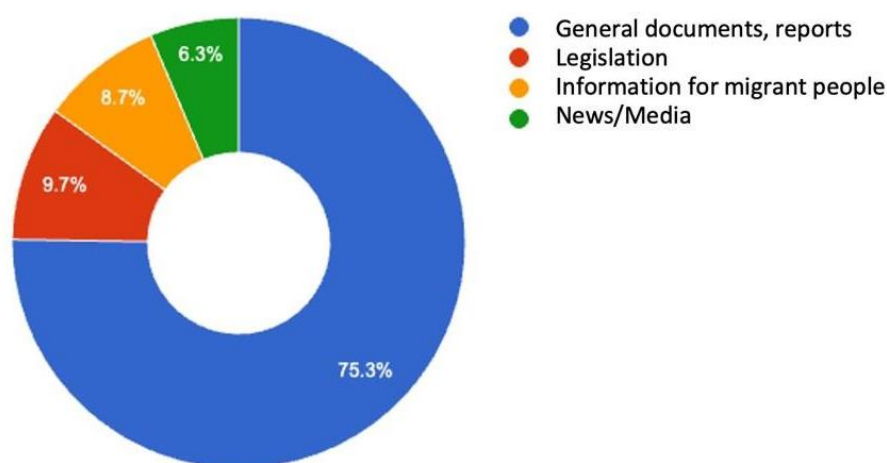
3. Data extraction and analysis

3.1 Corpora selection

The base corpus used in the work depicted in this paper was MIGRANTE.PT (Amaro; Correia; Gonçalves, 2021), a European Portuguese corpus for specific purposes, composed of texts relevant to hosting and integrating migrant people in Portugal. The corpus, comprising 1,435,551 tokens, includes texts from the following Portuguese institutions: the UN Portuguese High Commissariat for Migrations, the Portuguese Refugees Council, Caritas, the National Entity for Health Regulation, the Migrations Observatory, the Portuguese Foreign and Borders Services, the Portuguese Ministry of Internal Administration, the Portuguese Social Security, the Portuguese Parliament, and several Portuguese municipalities. It also includes media texts from various sources collected by these organizations, aimed at migrant people for their informative and/or promotional character.³

³ The full details on the constitution of the corpus can be consulted at <https://clunl.fcsh.unl.pt/en/online-resources/corpora/migrante-pt/>.

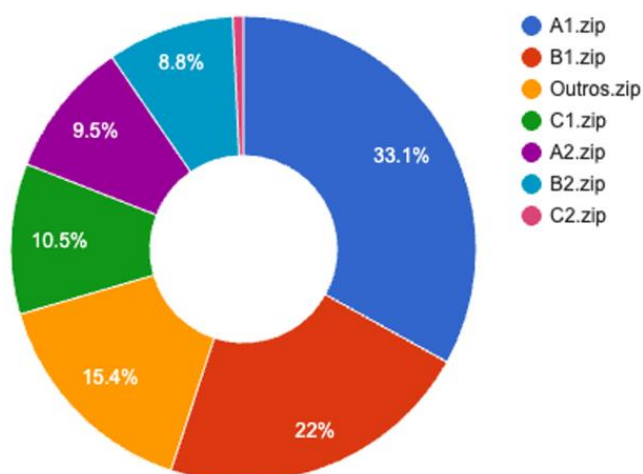
Figure 1. MIGRANTE.PT constitution



Source: Retrieved from Sketch Engine, MIGRANTE.PT corpus info.

The MATERIALS_PFL_2023 corpus is a Portuguese corpus comprising 1,110,948 tokens. It is composed of handbooks and teaching materials sourced from well-established publishing houses such as Lidel, Porto Editora, as well as national and international institutions dedicated to teaching PFL, including the Portuguese Ministry of Education, PPPLE (Portal of Portuguese as Foreign/Non-native Language Teacher), CPLP (Community of Portuguese Language Countries), FLAD (Luso-American Development Foundation), ACM, and the European Commission. The corpus is organized by proficiency levels, as illustrated in Figure 2.

Figure 2. MATERIALS_PFL_2023 corpus constitution



Source: Retrieved from Sketch Engine, MATERIALS_PFL_2023 corpus info

These two corpora will serve as the focus corpus and reference corpus for keyword extraction, respectively, as explained in the following subsection.

3.2 Extraction of candidates

The exploitation of the corpora for the extraction of lexicon unit candidates to feature in relevant PLE teaching/learning materials for migrant people was done using Sketch Engine⁴ (Kilgarriff *et al.*, 2004; Kilgarriff *et al.*, 2014a). Two main functionalities were utilised to extract relevant candidates: keyword extraction (simple word items and multiword expressions) and collocations.

According to Kilgarriff *et al.* (2014b), the keyword extraction functionality in Sketch Engine allows for easy and almost immediate identification of candidate terms in corpora. The candidates are selected by calculating the frequency of items (tokens, lemmas or lemma + part of speech) in a domain corpus compared to a reference corpus. Thus, keyword extraction not only serves as a tool for terminology extraction, but also allows us to easily discern what is specific to a given corpus in comparison to another. In the specific case presented here, keyword extraction was used to list single word candidates and multiword candidates, considering MIGRANTE.PT as the focus corpus, and MATERIALS_PFL_2023 as the reference corpus, to isolate occurrences in the texts directed to migrant people that are not considered or deemed relevant in Portuguese teaching materials.

The objective was to extract items specific to texts targeting migrant people and, from that list, select candidates related to health. After extracting the keyword lists, we obtained 46 single words candidates and 47 multiword candidates from two lists of 1,000 keywords each.

In addition to relevant cases in the lists of single keywords, we also used the collocations functionality. Collocations refer to sets of two or more words that co-occur with statistical relevance in a specific corpus, in different relative positions and distances. Co-occurrence data enable the identification of words that tend to co-occur more frequently in a specific corpus than expected based on the frequency of each element in that same corpus. This allows us to identify nominal compounds, idioms, formulae, proverbs, light verb constructions, amongst others (see typologies discussed in Sag *et al.* (2002) or Cowie (1994; 2001)). For our purposes, and for language teaching, the modelling and/or classification of different degrees of fixedness and/or idiomatic meaning (Mel'čuk, 1998; Sinclair, 1991; Fonseca; Sadat; Lareau, 2017) are not relevant. However, acknowledging the existence of these phenomena is important, as multiword expressions are quite common in specific domains. For these reasons, collocations of the items *saúde* (health), *doença* (illness), *doente* (patient), *medicamento* (medicine) and *médico* (doctor) – keywords for the healthcare domain, with sufficient frequency to produce collocates – were extracted from the MIGRANTE.PT corpus and analysed. Table 2 below presents the extracted results. The complete list of items is reproduced in Appendix 1.

⁴ Available at: <http://www.sketchengine.eu>.

Table 1. Collocates and candidates related to health extracted from MIGRANTE.PT

Base element of the collocation	Number of collocates	Number of candidates extracted
<i>saúde</i> (health)	828	94
<i>médico</i> (doctor)	258	28
<i>doença</i> (illness)	193	36
<i>doente</i> (patient)	76	10
<i>medicamento</i> (medicine)	28	7
Total	1383	175

Source: Own elaboration.

As expected, the initial data obtained showed repetitions, as collocates are often calculated based on wordforms rather than lemmas. Additionally, 17 of the multiword keywords extracted coincided with collocates. The treatment of the extracted lists (removal of duplicates, normalization, lemmatisation) resulted in a list of 207 single and multiword items related to health (see Appendix 1).

3.3 Results for Portuguese as host language teaching

A mere list of words and/or expressions does not constitute a lexicon for language teaching/learning, let alone for promoting health literacy. For this reason, the analysis of the results obtained included organizing the extracted lexicon into subdomains, arranging the items within each subdomain, establishing relationships with common vocabulary typically addressed in PFL materials, and extracting authentic examples from the corpus.

3.3.1 Subdomains

The organization of the extracted items into subdomains considered two main aspects:

- i. Communication goals, i.e., the relevance for the target audience and the goals of the hosting institutions, which are pertinent for determining the relevant subdomains.
- ii. Lexical organization, i.e., lexical-conceptual relations such as synonymy (or near-synonymy), hypernymy/hyponymy, meronymy, etc.

Given the meaning of the items listed, verified by the analysis of their concordances and the analysis of the semantic relations between them, it appeared that the first need to be satisfied is access to healthcare. This is because the topics referred to concern rights and access to healthcare, health system organization, and healthcare providers. Based on this, and according

to the topics covered in the corpus and the items extracted, the candidates were organized into 10 subdomains⁵:

- 1) Health rights and access to healthcare
- 2) Portuguese health system organization
- 3) Portuguese health areas
- 4) Health providers/institutions
- 5) Health professionals
- 6) Health status/condition
- 7) Diagnosis and treatment processes
- 8) Illnesses
- 9) Health-related documentation
- 10) Other

Health rights and access to healthcare domain, for instance, encompasses the following candidates:

- 1) Health rights and access to healthcare
 - acesso à saúde/acesso a cuidados de saúde* (access to healthcare)
 - acesso ao SNS* (access to the national health service)
 - acordo de saúde* (health agreement)
 - acordo internacional de saúde* (international health agreement)
 - atribuição* (assignment)
 - beneficiar* (to benefit)
 - beneficiário* (beneficiary)
 - Cartão Europeu de Seguro de Doença* (European Health Insurance Card)
 - participação de medicamentos* (medicines reimbursement)
 - convenção* (convention/concord)
 - despesas de saúde* (health expenses)
 - direito à proteção da saúde* (right to health protection)
 - direito à saúde* (right to health)
 - direitos básicos de saúde* (basic health rights)
 - direitos do doente/dos doentes* (patient rights)
 - mobilidade de doentes* (patient mobility)
 - mobilidade internacional de doentes* (patient international mobility)
 - prestação por doença/ subsídio de doença* (sickness benefit)
 - proteção* (protection)
 - proteção da doença* (sickness protection)
 - proteção da saúde* (health protection)
 - protocolo de saúde* (health protocol)
 - seguro de saúde* (health insurance)
 - subsistema de saúde* (health sub system)
 - utilização dos serviços de saúde* (use of health services)
 - vacinação* (vaccination)

⁵ The domains proposed here emerged from the vocabulary lists. Further validation is required from hosting institutions, on the one hand, and from healthcare professionals and target users (migrant people), on the other.

Health providers/institutions, yet another example, include the following items.

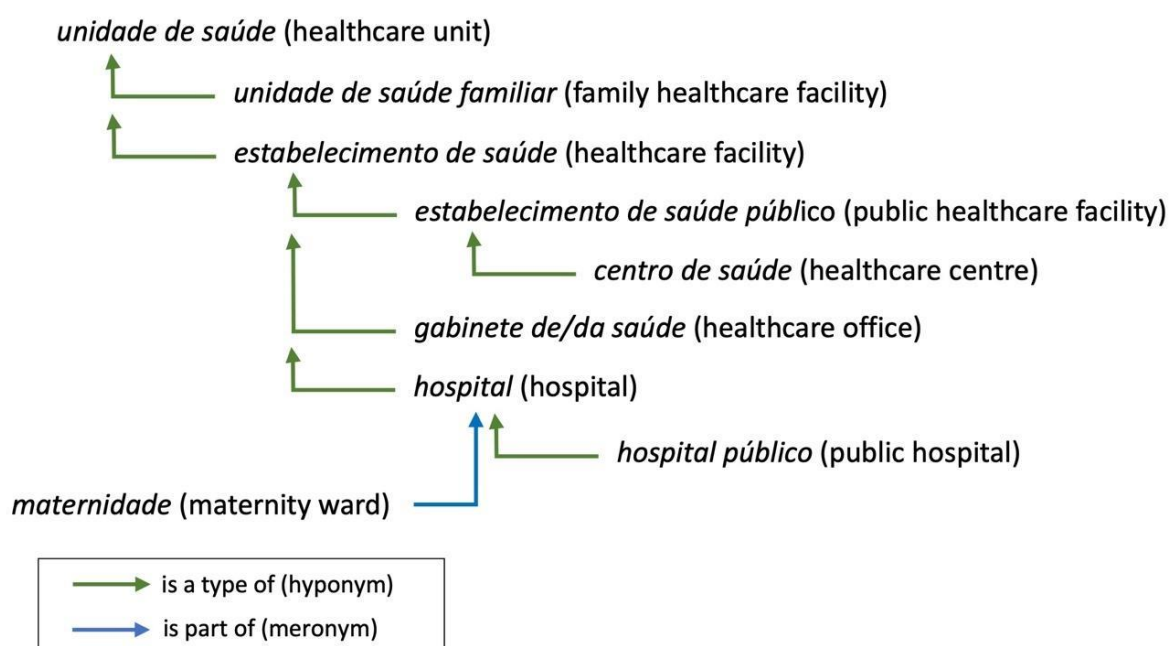
- 4) Health providers/institutions
 - área da saúde* (healthcare area)
 - berçário* (nursery)
 - centro de saúde* (healthcare centre)
 - estabelecimento de saúde público* (public healthcare facility)
 - estabelecimento de saúde/instituição de saúde* (healthcare facility/healthcare institution)
 - extensão de saúde* (healthcare extension)
 - gabinete de/da saúde* (healthcare office)
 - hospital* (hospital)
 - hospital público* (public hospital)
 - prestador de saúde* (healthcare provider)
 - unidade de saúde* (healthcare unit)
 - unidade de saúde familiar* (family healthcare unit)

The full list of domains and items organized by domains can be consulted in Appendix 2.

3.3.2 Lexical organization

An additional analysis of these sets allows us to further organize them. Considering lexical-conceptual relations within the framework of WordNet (Fellbaum, 1998; Vossen, 2002; Marrafa, 2002), it is possible to establish hypernymy/hyponymy and meronymy/holonymy relations that help us to structure the set. Figure 3 presents an example of the list of items in the health providers/institutions.

Figure 3. Organization of health providers/institutions



Source: Own elaboration.

With this structuring step, it is possible to easily observe that *maternity ward* is part of a *hospital*, for instance, and that *hospital* is a type of *healthcare facility*, which in turn is a type of *healthcare unit*. Besides contributing to the organization of the lexicon, this structure also conveys information on inference processes, thus being helpful for vocabulary acquisition (Haastrup; Henriksen, 2000; Crossley; Salsbury; McNamara, 2009; Eguchi; Kyle, 2020). The hypernymy/hyponymy relation can be described as a lexical-conceptual relation that concerns both world knowledge and linguistic knowledge. This can be verified in anaphoric constructions such as the ones presented below, where the hypernym is used to refer to a more specific referent (the hyponym) previously introduced (lexical anaphora) (1a), whereas the opposite is not possible (1b).

- (1)
- a. He went to the hospital, but the healthcare facility was not admitting patients.
 - b. #He went to the healthcare facility, but the hospital was not admitting patients.

The hierarchical nature of the hyponymy relation can be further tested in contrastive contexts and simple coordination structures, showing the meaning differences between the hyponym and hypernym, as shown in (2).

- (2)
- a. #He went to a hospital and to a healthcare facility.
 - b. #The hospital is more effective than a healthcare facility.
 - c. #He went to a hospital but not to a healthcare facility.

Hyponymy/hyponymy relation also incorporates a monotonic inheritance device (see Miller, 1990) that allows for describing lexical items in an economic way. Since hyponyms inherit the semantic and conceptual properties of their hypernym, a sentence such as “migrant people in Portugal are entitled to health care provided in healthcare facilities” entails that migrant people can be treated in hospitals, healthcare centres, healthcare offices, etc. All these properties facilitate vocabulary acquisition.

This organization of the lexicon also allows for easily noticing vocabulary gaps, i.e., identifying relevant missing items. Focusing, again, on the previous example, and considering this is a corpus-based approach⁶, it is easily noticeable that items such as *clínica* (clinic), as a type of healthcare facility, or *maternidade* (maternity ward), *serviço de pediatria/pediatria* (paediatrician unit), *serviço de ortopedia/ortopedia* (orthopaedics unit), *serviço de cirurgia/cirurgia* (surgery unit) or *serviço de urgência/urgência* (emergency unit), as parts of a hospital, are missing. Therefore, the step of organizing the lexicon is useful for the lexicographer as well as for the language learner.

⁶ A corpus-based approach refers to “a methodology that avails itself of the corpus mainly to expound, test or exemplify theories and descriptions that were formulated before large corpora became available to inform language study” (Tognini-Bonelli, 2001, p. 65). It differs from corpus-driven approaches in which the model and/descriptions of the phenomena emerge only from the corpus data.

3.3.3 Relation to common vocabulary

Producing language teaching/learning materials more attuned to the specific needs of target audiences does not mean ignoring or dismissing existing ones. In fact, existing or general language teaching/learning materials cover basic, familiar vocabulary that is naturally acquired by native speakers in the first years of language acquisition, usually in family and informal contexts. Words such as mother, father, son, sister, grandmother, etc., are part of the basic or fundamental lexicons, word lists that contain frequent vocabulary (Dottrens; Massarenti, 1948; Gougenheim *et al.*, 1954; Bacelar do Nascimento; Rivenc; Segura da Cruz, 1984; Lopez Morales, 1986). These are consensually included in language learning/teaching materials as they are essential to communication, but the lexicon covered has a lower probability of occurrence in reference corpora.

The proposal presented here is to combine both lists in a complementary way, thus boosting vocabulary learning. This can be achieved by relating items extracted from the health domain with items from fundamental lexicons. The examples in (3) illustrate this.

- (3) a. *saúde da criança* (child's health)
criança (child) is near synonym of *menor* (minor) ♦ see *cuidados de saúde a menores* (child healthcare)
is related to *infantil* (of child) ♦ see *saúde infantil* (child health)
saúde materno-infantil (mother-child health)
- b. *saúde materna* (mother's health)
materna (of mother, maternal) is related to *mãe* (mother) ♦ see *saúde materno-infantil* (mother-child health)

3.3.4 Authentic examples

As described earlier, the use of real texts targeting migrant people related to health helps ensure that the language learning/teaching materials will meet the communicative and informational needs of this target group. One way to further ensure this is to use the corpus also as the source for examples. Instead of constructing sentences from scratch to illustrate the use of a given word, it is possible to search for examples in the corpus. This enriches the final materials in two ways:

- Examples can serve to further clarify the meaning of a lexical item⁷.
- Authentic examples help to build collocational/distributional paradigms more effectively.

By presenting real examples related to specific needs or communicative contexts, we increase the likelihood of people recognizing the words. The sentences in (4) demonstrate this.

⁷ In lexicography, examples are traditionally used to convey information on syntactic properties such as subcategorization properties, as well as semantic properties, like argument structure and semantic domain, and register (formal vs. informal contexts).

(4)

a. É uma consulta destinada à vigilância, manutenção e promoção da **saúde da criança** e do jovem. (It is an appointment intended for the surveillance, maintenance, and promotion of the health of children and young people).

b. Caso esteja grávida, tem à sua disposição consultas de **saúde materna**, gratuitas, que efetuam o acompanhamento da gravidez e preparação para o parto. (If you are pregnant, you have access to free **maternal health** consultations that monitor your pregnancy and prepare you for childbirth.)

c. As várias **unidades de saúde familiar**, Serra da Lousã e Trevim Sol, UCC Arouce, funcionam, desde o início de 2015, em instalações novas. (The various **family healthcare facilities**, Serra da Lousã and Trevim Sol, UCC Arouce, have been operating in new facilities since the beginning of 2015.)

d. Tenho de pagar as consultas ou o **internamento** para o parto? (Do I have to pay for appointments or **hospitalization** for childbirth?)

e. Assim, quando surgirem os sinais de parto, basta ir ao serviço de urgência do hospital ou **maternidade** da sua área de residência. (Therefore, when signs of labour occur, simply go to the emergency unit of the hospital or **maternity ward** in your area of residence.)

Collocational/distributional paradigms can be easily understood by sorting concordances from the corpus. Figure 4 illustrates this for the word *maternidade* (maternity/maternity ward), sorted by the first and second words to the left of the keyword in context (KWIC). In lines 1 to 10 (except for line 5), the concordances show the item *maternity* = condition/situation of pregnancy, childbirth, or the first month of motherhood. Lines 11 to 21 show concordances relating to social benefits related to maternity. In these cases, the word is part of the multiword expressions *prestação de maternidade/subsídio de maternidade* (maternity benefits). The sense ‘maternity ward’ appears in lines 5, 32, 33 and 34.

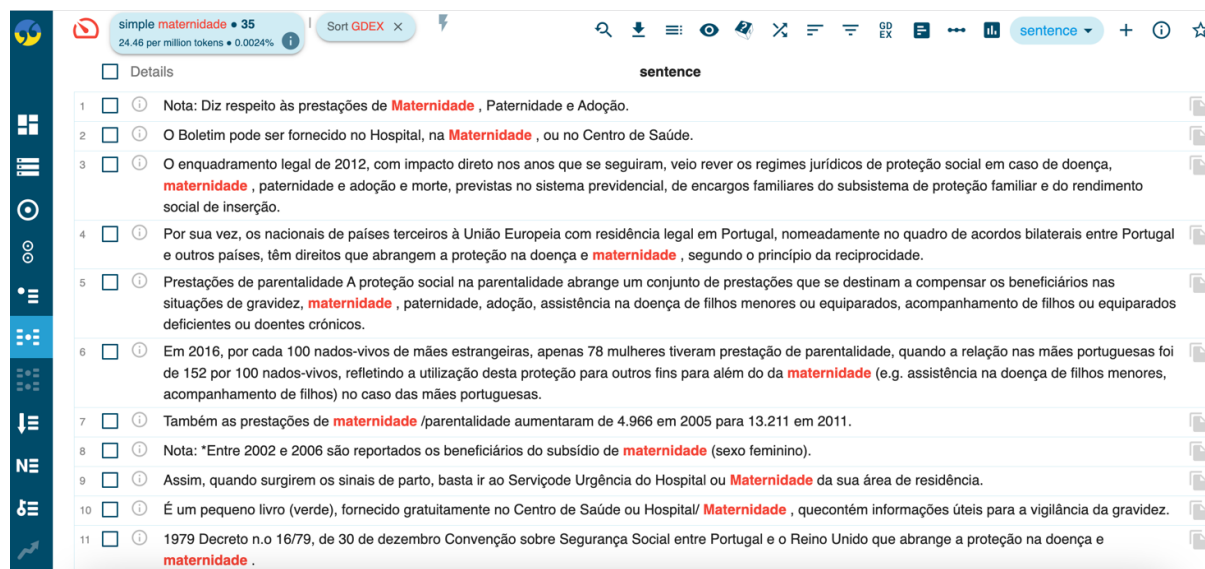
Figure 4. Concordances for *maternidade* in the MIGRANTE.PT corpus



Source: Retrieved from Sketch Engine, MIGRANTE.PT corpus

The examples can be selected from the concordances available in the corpus, as those displayed in Figure 4, or extracted through the Sketch Engine function Good Dictionary Examples (GDEX). GDEX is presented as a system for evaluating sentences regarding their suitability to serve as dictionary examples. It considers sentence length, presence of complex vocabulary, controversial topics (politics, religion...), anaphoric chains pointing outside of the retrieved context (e.g., pronouns pointing to previous sentences), proper names, among other criteria (see Kilgarriff *et al.*, 2008). In practical terms, it allows us to immediately access full, and not too long sentences from the corpus, instead of KWIC-centred lines of context (concordances). Figure 5 presents the results of GDEX for *maternidade*.

Figure 5. First results for GDEX for *maternidade* in the MIGRANTE.PT corpus



Source: Retrieved from Sketch Engine, MIGRANTE.PT corpus

3.4 Evaluation of results

The final step of the work presented in this paper concerns the evaluation of the proposed methodology with regard to the intended goals. For this purpose, and considering the corpus-based approach followed, the evaluation of results is done through comparison with results from other corpora (Schäfer; Bildhauer, 2013) and with other lexical resources with similar goals (Strandqvist *et al.*, 2018; Killgarriff *et al.*, 2014c).

To perform the initial evaluation, we compiled a very specific and small corpus of 124,872 tokens, composed of two institutional publications targeting migrant populations and focusing on their access to healthcare in Portugal: the *Manual de Acolhimento no Acesso ao Sistema de Saúde de Cidadãos Estrangeiros* (Handbook for Hosting Foreign Citizens in the Access to the Healthcare System), from the Ministry of Health, 2022, and the *Direitos e Deveres dos Utentes dos Serviços de Saúde*, (Rights and Duties of Health Service Users), from the Entidade Reguladora da Saúde, 2023. Both these documents were published after the compilation of the MIGRANTE.PT corpus.

We replicated the process presented in this paper and extracted both single keywords and multiword expression keywords using MATERIALS_PLE_2023 as the reference corpus. As expected, the 1,000 candidate lists extracted from the evaluation corpus included more candidates from the health domain since the corpus was domain-specific. In the case of MIGRANTE.PT, the corpus included other areas: employment, education, housing, etc. As expected, the set of candidates retrieved from MIGRANTE.PT was covered by the keyword

lists extracted from the evaluation corpus⁸. This means that the enlargement of the corpus with domain-specific documents would improve the lexicon coverage. However, it does not invalidate the methodology proposed.

To compare the results obtained here with vocabulary covered in existing resources with similar goals, we compared the lists in the *Glossário sobre Migração 2009* (Glossary on Migration), from the International Migration Organization (IOM), with 376 entries. From these, only one entry refers to health, as replicated below:

- (5)
saúde Bem-estar físico, mental e social e não a mera ausência de doença ou de enfermidade. (*health: Physical, mental and social well-being and not the mere absence of illness or infirmity.*) (*Glossário sobre Migração, 2009: 69*).

This small evaluation exercise allows us to i) confirm that the proposed methodology is capable of covering the relevant lexicon, and ii) highlight the necessity of the results to bridging the gap between available resources and needs, especially concerning information on accessing healthcare.

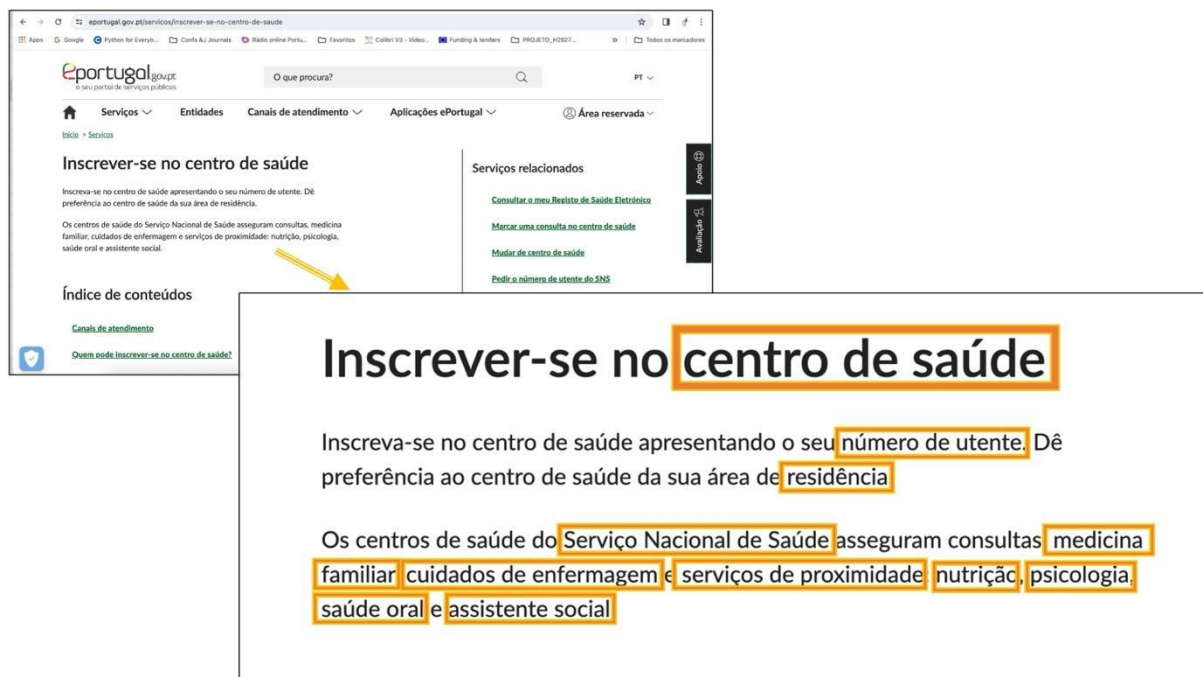
Final remarks

This paper presents a modest contribution to health literacy for a specific target audience: recently arrived migrant people, potentially in vulnerable situations, who must learn Portuguese as a host language. Focusing on vocabulary, we presented methods and results obtained from extracting lexical information from a specialized corpus, along with ways of better organise and present these results to the target audience, including the creation of host language materials based on corpus, i.e., authentic texts.

A proposal to include the identified vocabulary in classes and/or didactic materials (such as handbooks or class activities) is to establish an explicit connection between these specific words and common/general language ones, using authentic texts. Figure 6 illustrates such a case:

⁸ Specific numbers and lists can be provided with the final version of the paper.

Figure 6. Vocabulary activity based on authentic texts.



Source: Own elaboration, based on data from <https://eportugal.gov.pt/servicos/inscrever-se-no-centro-de-saude>, last accessed on February 25th, 2024.

The existence of dedicated dictionaries or glossaries, which establish bridges between vocabulary items, would allow for autonomous vocabulary discovery activities. Additionally, these can facilitate other vocabulary acquisition activities such as the identification of related words: *família* (family_{Noun}) - *familiar* (of.family_{Adj}); *cuidados* (care_{Noun}) - *cuidar* (care_{Verb}).

While demonstrating the validity of the methods and the relevance of the results, further investment in compiling these materials is required to create a comprehensive lexical resource ready for dissemination. Involving individuals and institutions working in the hosting and integration processes, as well as in healthcare access, is beneficial. However, as observed from available resources, obtaining such involvement can be challenging due to limited resources in these fields. Nevertheless, a collaborative effort involving linguists, host and integration specialists and healthcare professionals, supported by sound and testes methodologies, could effectively contribute to health literacy through host language teaching.

Funding

This research is supported by the Portuguese national funding through the FCT – Portuguese Foundation for Science and Technology, I.P. as part of the project UIDB/LIN/03213/2020; 10.54499/UIDB/03213/2020 and UIDP/LIN/03213/2020; 10.54499/UIDP/03213/2020 – Linguistics Research Centre of NOVA University Lisbon (CLUNL).

LINHA D'ÁGUA

References

- AMARO, R.; CORREIA, S.; GONÇALVES, M.; BARBERO, C.; MAGALHÃES, M. Inclusive Host Language Teaching: Official Texts for Migrant and Refugee People. In: MELETIADOU, E. (Ed.) *Policies and Practices for Assessing Inclusive Teaching and Learning*. United Kingdom: IGI Global, 2022, p. 183-210. DOI: <https://doi.org/10.4018/978-1-7998-8579-5.ch009>.
- AMARO, R.; CORREIA, S.; GONÇALVES, M. *MIGRANTE.PT. 2021*. Lisboa: CLUNL-NOVA FCSH, 2021. Available at: <https://clunl.fcsh.unl.pt/recursos-em-linha/corpora/migrante-pt/>. DOI: <https://doi.org/10.34619/ctho-dgqk>.
- BACELAR DO NASCIMENTO, M. F.; RIVENC, P.; SEGURA DA CRUZ, M. L. *Português Fundamental, Vocabulário e Gramática*, tomo 1, Vocabulário. Lisboa: INIC, 1984.
- BARBERO, C.; AMARO, R. Exploração de corpora para extração e descrição de léxico de especialidade. *Linha d'Água*, v. 33, n. 1, p. 69-104, 2020. DOI: <http://dx.doi.org/10.11606/issn.2236-4242.v33i1p69-104>.
- BELOSO, B. S. Designing, Describing and Compiling a Corpus of English for Architecture. *Procedia – Social Behaviour Sciences*, v. 198, p. 459-464, 2015. DOI: <https://doi.org/10.1016/j.sbspro.2015.07.466>.
- BRYERS, D.; WINSTANLEY, B.; COOKE, M. The power of discussion. In: MALLOWS, D. (Ed.) *Language issues in migration and integration: perspectives from teachers and learners*. London: British Council, 2014, p. 35-54.
- CARLINER, G. Wage differences by language group and the market for language skills in Canada. *The Journal of Human Resources*, v. 16, n. 3, p. 384-399, 1981. DOI: <https://doi.org/10.2307/145627>.
- COE - Council of Europe, *Guide to Health Literacy. Contributing to Trust Building and Equitable Access to Healthcare*. Steering Committee for Human Rights in the fields of Biomedicine, 2023. Available at: <https://www.coe.int/en/web/bioethics/guide-to-health-literacy>.
- COE - Council of Europe, *From Linguistic Diversity to Plurilingual Education: Guide for the Development of Language Education Policies in Europe*. Language Policy Division Council of Europe, Strasbourg, 2007. Available at: <https://rm.coe.int/16802fc1c4>.
- COE – Council of Europe. *Employment and Social Developments in Europe 2018*. Publications Office of the European Union. 2018.
- COOKE, M.; ROBERTS, C. *Developing adult teaching and learning: practitioner guides: ESOL*. London: NIACE/NRDC, 2007.
- COWIE, A. P. Phraseology. In: ASHER, R. E. (Ed.) *Encyclopedia of Language and Linguistics*. Oxford: Pergamon, 1994, p. 3168-3171.
- COWIE, A. P. Speech formulae in English: Problems of analysis and dictionary treatment. In: VAN DER MEER, G.; TER MEULEN, A. G. B. (Eds.). *Making Senses: From Lexeme to Discourse Honour of Werner Abraham*. Groningen: Center for language and Cognition, 2001, p. 1–12.
- CROSSLEY, S.; SALSBURY, T.; MCNAMARA, D. Measuring L2 lexical growth using hypernymic relationships. *Language Learning*, v. 59, n. 2, p. 307-334, 2009. DOI: <https://doi.org/10.1111/j.1467-9922.2009.00508.x>.
- DAVIS, T. C.; WOLF, M. S.; BASS, P. F.; THOMPSON, J. A.; TILSON, H. H.; NEUBERGER, M.; PARKER, R. M. Literacy and misunderstanding prescription drug labels. *Annals of Internal Medicine*, v. 19, p. 887-894, 2006.
- DOTTRENS, R.; MASSARENTI, D. *Vocabulaire fondamental du français*. Neuchâtel: Delachaux et Niestlé, 1948.

DROUIN, P. Term extraction using non-technical corpora as a point of leverage. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, v. 1, n. 9, p. 99-115, 2003.

EC – European Commission. *Action plan on Integration and Inclusion 2021-2027*. 2020. Available at: <https://ec.europa>.

EGUCHI, M.; KYLE, K. Continuing to explore the multidimensional nature of lexical sophistication: the case of oral proficiency interviews. *The Modern Language Journal*, v. 104, p. 381-400, 2020. DOI: <https://doi.org/10.1111/modl.12637>.

ELSOD, A.; MARQUES, M. (Coord.). *Ask the People: a consultation of migrants and refugees*. EMAB, Open Society Foundations, 2019. Available at: https://www.urbanagenda.urban-initiative.eu/sites/default/files/migration_files/ask_the_people_european_migrant_advisory_board_report_final.pdf.

ERS - Entidade Reguladora da Saúde. *Direitos e Deveres dos Utentes dos Serviços de Saúde*. Entidade Reguladora da Saúde, 2023. Available at: https://www.ers.pt/media/sfbd4x2h/publica%C3%A7%C3%A3o-ers_direitos-e-deveres.pdf.

FELLBAUM, C. (Ed.). *WordNet. An Electronic Lexical Database*, MA: The MIT Press, 1998.

FONSECA, A.; SADAT, F.; LAREAU, F. Combining Dependency Parsing and a Lexical Network Based on Lexical Functions for the Identification of Collocations. *Computational and Corpus-Based Phraseology*, v. 10596, p. 447-461, 2017. DOI: https://doi.org/10.1007/978-3-319-69805-2_31.

GOUGENHEIM, G.; MICHEA, R.; RIVENC, P.; SAUVAGEOT, A. *L'élaboration du français élémentaire: étude sur l'établissement d'un vocabulaire et d'une grammaire de base*. Paris: Didier, 1954.

HAASTRUP, K.; HENRIKSEN, B. Vocabulary acquisition: Acquiring depth of knowledge through network building. *International Journal of Applied Linguistics*, v. 10, n. 2, p. 221-240, 2000. DOI: <https://doi.org/10.1111/j.1473-4192.2000.tb00149.x>.

JAEGER, F. N.; PELLAUD, N.; LAVILLE, B.; KLAUSER, P. The migration-related language barrier and professional interpreter use in primary health care in Switzerland. *BMC Health Services Research*, v. 19, n. 1, p. 1-10, 2019. DOI: <https://doi.org/10.1186/12913-019-4164-4> PMID:31248420.

KILGARRIFF, A.; BAISA, V.; BUŠTA, J.; JAKUBÍČEK, M.; KOVÁŘ, V.; MICHELFEIT, J.; RYCHLÝ, P.; SUCHOMEL, V. The Sketch Engine: ten years on. *Lexicography*, v. 1, p. 7-36, 2014a.

KILGARRIFF, A.; JAKUBÍČEK, M.; KOVÁŘ, V.; RYCHLÝ P.; SUCHOMEL, V. Finding Terms in Corpora for Many Languages with the Sketch Engine. *Proceedings of the Demonstrations at the 14th Conference the European Chapter of the Association for Computational Linguistics*, Sweden, p. 53-56, 2014b.

KILGARRIFF, A.; RYCHLÝ, P.; JAKUBÍČEK, M.; KOVÁŘ, V.; BAISA, V.; KOCINCOVÁ, L. Extrinsic Corpus Evaluation with a Collocation Dictionary Task. In: CALZOLARI, N.; CHOUKRI, K.; DECLERCK, T.; LOFTSSON, H.; MAEGAARD, B.; MARIANI, J.; MORENO, A.; ODIJK, J.; PIPERIDIS, S. *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*. Reykjavik, p. 545-552, 2014c. Available at: http://www.lrec-conf.org/proceedings/lrec2014/pdf/52_Paper.pdf.

KILGARRIFF, A.; HUSÁK, M.; MCADAM, K.; RUNDELL, M.; RYCHLÝ, P. GDEX: Automatically finding good dictionary examples in a corpus. *Proceedings of the 13th EURALEX International Congress*, Spain, p. 425-432, 2008.

KILGARRIFF, A.; RYCHLÝ, P.; SMRŽ, P.; TUGWELL, D. The Sketch Engine. *Proceedings of the 11th EURALEX International Congress*, p. 105-116, 2004.

LANG, C.; SCHNEIDER, R.; SUCHOWOLEC, K. Extracting Specialized Terminology from Linguistic Corpora. In: FUB, E.; KONOPKA, M.; TRAWIŃSKI, B.; WAßNER, U. (Eds.). *Grammar and Corpora 2016*. Heidelberg: Heidelberg University Publishing, 2018, p. 425-434. DOI: <https://doi.org/10.17885/heiup.361.509>.

LINDEMANN, D. Bilingual Lexicography and Corpus Methods. The Example of German-Basque as Language Pair. *Procedia - Social and Behavioral Sciences*, v. 95, p. 249-257, 2013. DOI: <https://doi.org/10.1016/j.sbspro.2013.10.645>.

LÓPEZ MORALES, H. *La enseñanza de la lengua materna. Lingüística para maestros de español*. Madrid: Playor, 1986.

MARRAFA, P. The Portuguese WordNet: General Architecture and Semantic Internal Relations. *DELTA*, 2002.

MCMANUS, W.; GOULD, W.; WELCH, F. Earnings of Hispanic men: The role of English language proficiency. *Journal of Labor Economics*, v. 1, n. 2, p. 101-130, 1983. DOI: <https://doi.org/10.1086/298006>.

MEL'ČUK, I. Collocations and lexical functions. Phraseology. In: COWIE, A. P. (Ed.) *Theory, Analysis, and Applications*. Oxford: Oxford University Press, 1998, p. 23-53.

MILLER, G. Nouns in WordNet: a lexical inheritance system. *International Journal of Lexicography*, v. 3, n. 4, p. 245-264, 1990.

NUTBEAM D. The evolving concept of health literacy. *Social Science & Medicine*, v. 67, n. 12, p. 2072-2078, 2008. DOI: <https://doi.org/10.1016/j.socscimed.2008.09.050>.

OIM - Organização Internacional para as Migrações. *Glossário sobre Migração*, OIM Organização Internacional para as Migrações, 2009. Available at: <https://publications.iom.int/system/files/pdf/iml22.pdf>.

PEREZ, M. J. M.; RIZZO, C. R. Assessing four automatic term recognition methods: Are they domain-dependent? *English for Specific Purposes World*, v. 42, n. 15, 2014.

SAG, I.; BALDWIN, T.; BOND, F.; COPESTAKE, A.; FLICKINGER, D. Multiword Expressions: A Pain in the Neck for NLP. In: GELBUKH, A. (Ed). *Computational Linguistics and Intelligent Text Processing*. Berlin, Heidelberg: Springer, 2002, p. 1-15.

SINCLAIR, J. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press, 1991

SCHÄFER, R.; BILDHAUER, F. *Web Corpus Construction*. Switzerland: Morgan & Claypool Publishers, 2013.

STRANDQVIST, W.; SANTINI, M.; LIND, L.; JONSSON, A. Towards a quality assessment of web corpora for language technology applications. In: READ, T.; MONTANER, S.; SEDANO, B. (Ed.). *Technological Innovation for Specialized Linguistic Domains: Languages for Digital Lives and Cultures*. Belgium: Éditions universitaires européennes, 2018, p. 137-150.

TOGNINI-BONELLI, E. *Corpus Linguistics at Work*. Amsterdam: John Benjamins, 2001.

U.S DEPARTMENT of Health and Human Services, Office of Disease Prevention and Health Promotion. Healthy People 2030. Available at: <https://health.gov/healthypeople>.

VAN DER HEIDE, I.; WANG, J.; DROOMERS, M.; SPREEUWENBERG, P.; RADEMAKERS, J.; UITERS, E. The relationship between health, education, and health literacy: results from the Dutch Adult Literacy and Life Skills Survey. *Journal of health communication*, v. 18, n. 1, p. 172-184, 2013. DOI: <https://doi.org/10.1080/10810730.2013.825668>.

VOSSSEN, P. (Ed.) *EuroWordNet General Document*. University of Amsterdam, 2002. Available at: <http://www.vossen.info/docs/2002/EWNGeneral.pdf>.

Appendix 1 - List of candidates

acesso a cuidados de saúde
acesso à saúde
acesso ao SNS
acidente
acordo de saúde
acordo internacionais de Saúde
administração de medicamentos
admissão
análise
aparelho circulatório
aquisição de medicamentos
área da saúde
ARS
assistência medicamentosa
assistir doentes
atestados
atribuição
autoapreciação do estado de saúde
auxiliares de enfermagem
beneficiar
beneficiário
boa saúde
boletim de saúde
boletim de vacinas
carência
catar
causa de doença
causa de morte
centro de saúde
comparticipação de medicamentos
condição de saúde
consulta de saúde
consulta de rotina
convenção
cuidado de saúde
cuidados básicos de saúde
cuidados de saúde a menores
cuidados de saúde continuados
cuidados de saúde de emergência
cuidados de saúde materno-infantil
cuidados de saúde paliativos
cuidados de saúde primários
cuidados de saúde secundários
cuidados de saúde urgentes
declaração de doença
deficiência
desigualdade em saúde
despesas de saúde
determinante da/de saúde
determinantes estruturais da saúde
determinantes individuais da saúde
determinantes sociais da/de saúde
diabetes
hospital
hospital público
incidência
indicador de saúde
Inquérito Europeu de Saúde
Inquérito Nacional de Saúde
instituição de saúde
Instituto Nacional de Saúde Dr. Ricardo Jorge
instrumento médico
internamento
ISS
Lei de Bases da Saúde
maternidade
medicamento
medicamento prescrito
medicamentoso
médico
médico estrangeiro
medidas de saúde
Ministério da Saúde
ministrar
mobilidade de doentes
mobilidade internacional de doentes
mortal
mortalidade
mortalidade infantil
morte
muito boa saúde
nado-vivo
necessidade médica
necessidade de saúde
óbito
Organização Mundial de Saúde
padrão de saúde
parteira
pessoal de saúde
Plano Nacional de Saúde
política de saúde
precoce
prescrição
prescrição de medicamentos
prescrito
prestação de cuidados de saúde
prestação por doença
prestador de saúde
prevalência da doença
prevenção da doença
preventivo
privação
problema de saúde
problema de saúde prolongado
profissional de saúde
proteção

Direção Geral de Saúde	proteção da doença
direito à proteção da saúde	proteção da saúde
direito à saúde	protocolo de saúde
direitos básicos de saúde	queixa de saúde
direitos do doente/dos doentes	rastreio da doença
doença	razão de saúde
doença cardíaca	receita
doença cardiovascular	risco agravado de saúde
doença cerebrovascular	risco de saúde
doença crónica	risco para a saúde
doença de coração	saúde
doença do aparelho circulatório	saúde da área de residência
doença do sistema nervoso	saúde da criança
doença grave	saúde das populações
doença infecciosa	saúde dos imigrantes
doença infetocontagiosa	saúde humana
doença isquémica	saúde infantil
doença parasitárias	saúde materna
doença profissional	saúde materno-infantil
doença respiratórias	saúde mental
doença transmissíveis	saúde oral
doença mental	saúde pública
doença mortal	saúde reprodutiva
doente	saúde sexual
doente evacuado	saúde sexual e reprodutiva
doente crónico	Cartão Europeu de Seguro de Doença
domínio da saúde	seguro de saúde
enfermagem	serviço de saúde
enfermeiro	Serviço Nacional de Saúde
Entidade Reguladora da Saúde	SIDA
entrada de doentes	sistema de saúde
equidade em saúde	sistema nacional de saúde
ERS	situação de saúde
estabelecimento de saúde	SNS
estabelecimento de saúde público	sofrer de doença
estabilizar	subsídio de doença
estado de saúde	subsistema de saúde
evacuação de doentes	taxa de mortalidade
evolução	taxa moderadora
evolutivo	técnico de saúde
exame	transporte de doentes
exame complementar	tratamento médico
experiências de saúde	trauma
extensão de saúde	unidade de saúde
farmacêutico	unidade de saúde familiar
fecundidade	utente
feminização	utilização dos serviços de saúde
fornecimento de medicamentos	vacina
fragilidade	vacinação
gabinete de/da saúde	vigilância da saúde
gravidez	

Appendix 2 - Items per subdomain

<p>1) Health rights and access to healthcare acesso a cuidados de saúde acesso à saúde acesso ao SNS atribuição beneficiar beneficiário comparticipação de medicamentos despesas de saúde direito à proteção da saúde direito à saúde direitos básicos de saúde direitos do doente/dos doentes acordo de saúde acordo internacionais de Saúde convenção mobilidade de doentes mobilidade internacional de doentes prestação por doença subsídio de doença proteção proteção da doença proteção da saúde protocolo de saúde Cartão Europeu de Seguro de Doença seguro de saúde subsistema de saúde utilização dos serviços de saúde vacinação</p>	<p>2) Portuguese health system organization ARS Direção Geral de Saúde Entidade Reguladora da Saúde ERS ISS Lei de Bases da Saúde Ministério da Saúde Organização Mundial de Saúde Plano Nacional de Saúde política de saúde Serviço Nacional de Saúde SNS sistema de saúde serviço de saúde sistema nacional de saúde</p>
<p>3) Portuguese health areas área de saúde saúde da criança saúde das populações saúde dos imigrantes saúde humana saúde infantil saúde materna saúde materno-infantil saúde mental saúde oral saúde pública saúde reprodutiva saúde sexual saúde sexual e reprodutiva</p>	<p>4) Healthcare providers/institutions centro de saúde estabelecimento de saúde estabelecimento de saúde público extensão de saúde gabinete de/da saúde hospital hospital público instituição de saúde Instituto Nacional de Saúde Dr. Ricardo Jorge maternidade prestador de saúde unidade de saúde unidade de saúde familiar admissão entrada de doentes taxa moderadora transporte de doentes</p>

<p>5) Healthcare professionals auxiliares de enfermagem enfermagem enfermeiro farmacêutico médico médico estrangeiro parteira pessoal de saúde profissional de saúde técnico de saúde</p>	<p>6) Health status/condition fecundidade feminização fragilidade gravidez mortal mortalidade mortalidade infantil morte muito boa saúde nado-vivo óbito boa saúde carência catar causa de doença causa de morte condição de saúde deficiência precoce prevalência da doença prevenção da doença preventivo privação problema de saúde problema de saúde prolongado risco agravado de saúde risco de saúde risco para a saúde saúde situação de saúde sofrer de doença</p>
<p>7) Diagnosis and treatment processes administração de medicamentos fornecimento de medicamentos análise aquisição de medicamentos assistência medicamentosa assistir doentes autoapreciação do estado de saúde consulta de saúde consulta de rotina cuidado de saúde cuidados básicos de saúde cuidados de saúde a menores cuidados de saúde continuados cuidados de saúde de emergência cuidados de saúde materno-infantil cuidados de saúde paliativos cuidados de saúde primários cuidados de saúde secundários cuidados de saúde urgentes estabilizar evacuação de doentes evolução evolutivo exame</p>	<p>8) Illnesses diabetes doença doença cardíaca doença cardiovascular doença cerebrovascular doença crônica doença de coração doença do aparelho circulatório doença do sistema nervoso doença grave doença infecciosa doença infetocontagiosa doença isquêmica doença parasitárias doença profissional doença respiratórias doença transmissíveis doença mental doença mortal SIDA trauma acidente</p>

<p>exame complementar instrumento médico internamento medicamento medicamento prescrito medicamentoso ministrar prescrição prescrição de medicamentos prescrito prestação de cuidados de saúde queixa de saúde rastreamento da doença receita tratamento médico vacina doente doente evacuado doente crónico utente</p>	
<p>9) Health related documentation atestados boletim de saúde boletim de vacinas declaração de doença</p>	<p>10) Other desigualdade em saúde determinante da/de saúde determinantes estruturais da saúde determinantes individuais da saúde determinantes sociais da/de saúde domínio da saúde equidade em saúde experiências de saúde incidência indicador de saúde Inquérito Europeu de Saúde Inquérito Nacional de Saúde medidas de saúde necessidade médica necessidade de saúde padrão de saúde razão de saúde taxa de mortalidade vigilância da saúde</p>