

## Léxico aplicado à Saúde

**Gabriela Tieko Hirashima**  
Departamento de Computação  
Universidade Estadual de  
Londrina (UEL)  
Londrina, Brazil  
gabriela.hirashima@uel.br

**Pedro Hendling Hernandes**  
Departamento de Computação  
Universidade Estadual de  
Londrina (UEL)  
Londrina, Brazil  
deprao.pwhernandes@uel.br

**Cinthyán Renata  
Sachs C. de Barbosa**  
Departamento de Computação  
Universidade Estadual de  
Londrina (UEL)  
Londrina, Brazil  
cinthyán@uel.br

**Wagner Ferreira Lima**  
Departamento de Letras  
Vernáculas e Clássicas  
Universidade Estadual de  
Londrina (UEL)  
Londrina, Brazil  
wflima@uel.br

### ABSTRACT

The objective of this work was to show a simple, yet practical lexicon that helps in a big problem with Portuguese language: the incomprehension caused by the language itself. Due to its vast size and culture, words can have different meanings or even not exist in different regions of the country. This fact makes it difficult for doctors and patients to talk, which can lead to fatal errors. Thus, using as a basis a dictionary that addresses technical terms used with their popular variants and definitions in health, a lexicon was developed to reduce these difficulties found in the Brazilian Portuguese language.

### Keywords

Lexicon; Health; Hash Table.

### ACM Classification Keywords

Human-centered computing: Human Computer Interaction (HCI).

### RESUMO

O objetivo deste artigo foi mostrar um simples, porém prático programa que auxilie em um grande problema que a Língua Portuguesa possui: a incompreensão causada pela própria linguagem. Devido ao seu vasto tamanho e cultura, palavras podem possuir diferentes significados ou até mesmo não existir em diferentes regiões do país. Tal fato dificulta a conversa entre médicos e pacientes, o que pode levar a erros fatais. Dessa maneira, utilizando como base um dicionário que aborda termos técnicos utilizados com suas variantes populares e definições na saúde, um léxico foi desenvolvido para diminuir essas dificuldades encontradas na Língua Portuguesa do Brasil.

### Palavras-chave

Léxico; Saúde; Tabela Hash.

### INTRODUÇÃO

Para abordar o assunto da diversidade linguística de maneira coerente é preciso entender as variantes regionais que uma mesma palavra pode assumir. O problema de regiões diferentes do país possuir diversos sotaques, variantes linguísticas e costumes representam dificuldades para os nativos brasileiros e, ainda mais, para os estrangeiros.

Professores e médicos brasileiros participantes do programa Língua Portuguesa para o *Programa Mais Médicos para o Brasil*, relataram a existência de duas linguagens no Sistema Único de Saúde (SUS) [1]: a linguagem técnica, falada por profissionais da área de saúde e a linguagem geral, utilizada por pacientes ou familiares de pacientes. Este trabalho foi baseado nessas duas linguagens.

O desentendimento entre os grupos não ocorre somente entre estrangeiros e nativos. Até mesmo dentro do país, os dialetos se expressam de maneiras extraordinariamente diferentes e que podem variar por diversos fatores como propriedade intelectual, classe social, contato cultural, etc. Também há a existência de gírias próprias do corpo hospitalar, a qual também pode gerar confusão ao ser observada pelos de fora desse grupo [2]. Dessa maneira, não é surpreendente notar a existência do distanciamento entre o profissional de saúde e a população geral, durante atendimentos médicos.

Essa dificuldade, embora compreensível haja vista a diversidade encontrada no Brasil, deve ser superada pelo corpo hospitalar. Como citado por Sucupira [3], quando o foco é a doença e não o indivíduo, a não resolução do problema é frequente e causa insatisfação a todos os envolvidos. Quando consideramos o indivíduo, a forma com que ele se comunica é uma característica de grande efeito.

Outro ponto a ser considerado no período em que vivemos é a comunicação por meio de redes sociais. O impacto desse meio é muito relevante e abrangente, quando abordado de maneira efetiva. A linguagem é um fator de extrema relevância nesse caso e pode afastar ou aproximar o ambiente hospitalar e a população [4].

A Língua Portuguesa pode ser considerada morfológicamente rica porque é uma língua fusional. Tem mais variação nas formas de palavras do que o inglês, que é uma língua analítica, embora tenha menos variação do que uma linguagem aglutinante, tomando a análise morfológica um desafio [5].

A partir dessas afirmações, o léxico mostrado neste trabalho busca suprir a dificuldade em associação de palavras de cunho técnico e suas alternativas, oferecendo um método rápido e prático para o uso de profissionais e da população geral. O uso de uma interação simples entre o programa e o usuário é simplificado, tendo em vista o ganho de valor ao ser utilizado por todos, mesmo que não tenham conhecimento das linguagens de programação.

O uso de um sistema específico como o proposto, com a associação de palavras e significados já verificados, permitiria ao usuário uma maior confiança em relação a outros métodos, como por exemplo, uma rápida busca em ferramentas de pesquisa. Embora simples, o uso de ferramentas generalizadas pode ser insatisfatório, devido ao grande volume de informação disponível na rede sem garantia de integridade, fator que pode confundir o usuário.

Com um sistema específico com base de dados confiável e estruturado, é possível fornecer agilidade e segurança ao usuário, fatores que são essenciais em diversas áreas, sobretudo, na Saúde.

Este trabalho está organizado como segue: a Seção 2 apresenta um referencial teórico e trabalhos relacionados na área de Processamento de Linguagem Natural na Saúde, a Seção 3 aborda a Coleta de Dados, a Seção 4 demonstra o funcionamento do Léxico, a Seção 5 explica a estrutura do código-fonte e a utilização da função hash desse léxico, e, finalmente, a Seção 6 apresenta as conclusões deste trabalho e propostas de trabalhos futuros.

## REFERENCIAL TEÓRICO

Há uma diversidade de terminologias clínicas com diferentes propósitos, por exemplo, as que representam o jargão médico e são denominadas “terminologias de interface”; as ontologias lidam com o conhecimento canônico, muitas vezes rotuladas de “terminologias de referência”, e as classificações, como CID-10, chamadas de “terminologias de agregação” [6].

Alguns trabalhos utilizam técnicas de Processamento de Linguagem Natural (PLN) e representação de conhecimento em ontologias [7].

Grande parte dos estudos realizados atualmente nessa área é, segundo Pereira e Rigo [7], baseada em construções

matemáticas para modelar os relacionamentos existentes nos textos tratados. As técnicas consistem, principalmente, na análise do número de incidência e da proximidade das palavras em uma mensagem.

Por meio de uma Gramática Livre de Contexto (GLC) e um algoritmo de conversão para Gramática Livre de Contexto Lexicalizada (GLCL), construções reais de uma interface na área da radiologia para consultas em Língua Portuguesa foi delineada por Barbosa [8] em cima de um caso de sarcoma de Ewing, onde incluía perguntas em vozes ativa e passiva, cláusulas relativas e interrogativas, suas combinações e pronomes (interrogativo, entre outros). Temos ainda utilizando a mesma gramática para um caso de câncer de pele que podemos encontrar em Pereira *et al.* [9].

Alguns trabalhos que utilizam de PLN na área médica foram encontrados. Desde os que abordam as anamneses do domínio da ginecologia [6] aos que tratam do desenvolvimento de sistemas de processamento de texto livre, capazes de extrair informação pertinente de um grande volume de textos médicos em Linguagem Natural (LN), como é o caso do trabalho de Ferreira, Teixeira e Cunha [10].

Um trabalho similar pode ser visto em Benicio [11] que faz o tratamento, cruzamento e padronização das informações, onde provavelmente será possível delinear um método terapêutico mais assertivo, a partir da detecção precoce do diagnóstico clínico.

Modelagem e desenvolvimento de um sistema de PLN (*chatbot*) em AIML foi desenvolvido por Souza *et al.* [12] para auxiliar na assistência de crianças de 0 a 2 anos de idade com fissuras labiopalatinas e anomalias craniofaciais, mediante a interação entre familiares e profissionais com um ambiente virtual de aprendizagem.

Porém, observando tais trabalhos nenhum deles têm um olhar mais específico em cima da terminologia da saúde. Assim, um Léxico da Saúde foi implementado.

## DESENVOLVIMENTO DO LÉXICO

É por meio dessa temática de coleta de dados que chegamos a um conceito muito conhecido na área de PLN: o *corpus*. Um conceito pare esse seria [13]: conjunto de textos escritos e registros orais em uma determinada língua e que serve como base de análise.

A principal fonte de dados utilizada para este trabalho é a de Silvestrow [1] em contribuições dos Estudos Terminológicos para os Profissionais da Saúde Básica do SUS. Ao longo de 10 anos, os termos técnicos e seus respectivos coloquiais foram estudados e correlacionados, a fim de documentar as variações léxicas e morfológicas dos termos técnicos da área de saúde.

A justificativa para a escolha do trabalho de Silverstrow [1] se justifica também pela citação de médicos estrangeiros do Programa *Mais Médicos para o Brasil* em Minas Gerais, além, por exemplo, de um professor de Medicina da UFMG

como colaboradores das definições dos termos médicos. Além deles, terminólogos e muitos professores universitários forneceram dados, correções e conhecimento adicional para a construção da base de dados, fazendo com que a fonte possua maior confiabilidade.

A partir das informações coletadas no referido documento, um banco de dados foi formado para a consulta, oferecendo alguns dos termos mais utilizados na área atendimento de saúde e armazenando os seguintes atributos citados: ‘Termo Pesquisado’, ‘Categoria’, ‘Variação(ões) Denominativa(s)’, ‘Variante(s) Popular(es)’ e ‘Definição’.

Nesse documento foram utilizadas 43 palavras das 84 disponíveis [1] para o dicionário de termos técnicos e 71 palavras populares alternativas a termos técnicos para o dicionário de palavras técnicas, totalizando 112 palavras disponíveis no trabalho.

A ideia deste trabalho é criar um *dataset* especializado para ser disponibilizado na web que pode ser usado na criação de um dicionário mais completo na área de saúde.

### DICIONÁRIO

O sistema possui o seguinte funcionamento: ao inicializar o programa, um menu será mostrado ao usuário com as funções de ‘Verificar os termos disponíveis no dicionário’, ‘Buscar palavra, seu significado e seu(s) termo(s) alternativo(s)’ e uma opção para ‘Sair’, como mostrado na Figura 1.

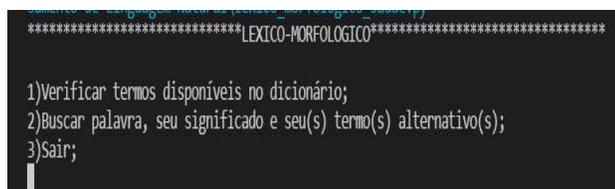


Figura 1. Menu – Estrutura de Controle do Programa  
[Fonte: Os autores]

Ao selecionar a opção 1, o programa mostra listas de termos médicos e populares disponíveis para a busca em tons amarelados, considerando o grupo de termos médicos e de termos populares. A cor amarela é utilizada para destacar o vocabulário nos símbolos terminais, os quais geralmente possuem fundo escuro. Para colorir foi utilizada a biblioteca Colorama disponível para Python, a qual é multiplataforma e roda para Linux, Windows e macOS. Nas Figuras 2 e 3 são mostradas como essas listas são apresentadas.

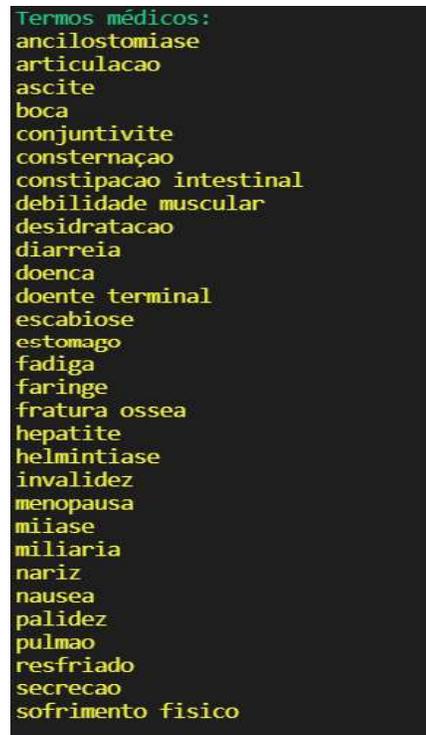


Figura 2. Lista de palavras do dicionário de termos técnicos  
[Fonte: Os autores]

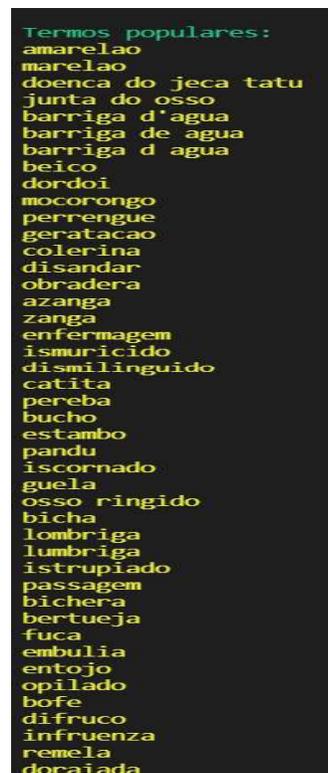


Figura 3. Lista de palavras do dicionário de termos técnicos  
[Fonte: Os autores]

Na Figura 4 a palavra ‘Articulação’ é pesquisada, sendo essa um termo médico.

```

1)Verificar termos disponíveis no dicionário;
2)Buscar palavra, seu significado e seu(s) termo(s) alternativo(s);
3)Sair;
2
Insira o termo a ser buscado:
Articulação

Termo Técnico
Termo : Articulação
Categoria : Termo médico/Anatomia
Variação(ões) Denominativa(s) : Junta, Juntas, Juntura
Variante(s) Popular(es) : Junta do Osso
Definição : Junção ou articulação óssea, joelho.

```

**Figura 4. Busca pela palavra ‘Articulação’**  
[Fonte: Os autores]

Nas Figuras 5 e 6, exemplos de buscas realizadas de maneira alternativa são exibidos. Pode-se observar que a simplificação da palavra de busca resulta em uma maior flexibilização e compreensão por parte do programa.

Assim, uma palavra é definida da seguinte maneira:

```

articulacao = {
    "Termo" : "Articulação",
    "Categoria" : "Termo médico/Anatomia",
    "Variação(ões) Denominativa(s)" : "Junta, Junta, Juntas,
Juntura",
    "Variante(s) Popular(es)" : "Junta do Osso",
    "Definição" : "Junção ou articulação óssea, joelho."
}

```

Ainda seguindo nessa mesma linha, pode-se observar que a variante popular da palavra ‘Articulação’ é Junta do Osso. Uma pesquisa dessa variante é exemplificada na Figura 7.

```

1)Verificar termos disponíveis no dicionário;
2)Buscar palavra, seu significado e seu(s) termo(s) alternativo(s);
3)Sair;
2
Insira o termo a ser buscado:
articulacao

Termo Técnico
Termo : Articulação
Categoria : Termo médico/Anatomia
Variação(ões) Denominativa(s) : Junta, Juntas, Juntura
Variante(s) Popular(es) : Junta do Osso
Definição : Junção ou articulação óssea, joelho.

```

**Figura 5. Busca pela palavra ‘Articulacao’**  
[Fonte: Os autores]

```

1)Verificar termos disponíveis no dicionário;
2)Buscar palavra, seu significado e seu(s) termo(s) alternativo(s);
3)Sair;
2
Insira o termo a ser buscado:
ArticUlação

Termo Técnico
Termo : Articulação
Categoria : Termo médico/Anatomia
Variação(ões) Denominativa(s) : Junta, Juntas, Juntura
Variante(s) Popular(es) : Junta do Osso
Definição : Junção ou articulação óssea, joelho.

```

**Figura 6. Busca pela palavra ‘ArticUlação’**  
[Fonte: Os autores]

```

1)Verificar termos disponíveis no dicionário;
2)Buscar palavra, seu significado e seu(s) termo(s) alternativo(s);
3)Sair;
2
Insira o termo a ser buscado:
junta do osso

Termo Popular
Termo : Junta do Osso
Categoria : Termo Popular Informal/Anatomia
Termo(s) técnico(s) : Articulação
Variação(ões) Denominativa(s) : Junta, Juntas, Juntura
Variante(s) Popular(es) : -
Definição : Junção ou articulação óssea, joelho.

```

**Figura 7. Busca pela palavra ‘Junta do Osso’**  
[Fonte: Os autores]

Por fim, a opção Sair de número 3, apenas finaliza a execução do programa.

### TABELAS HASH

Tabelas hash têm demonstrado gerar bons resultados para a Interface em Linguagem Natural para Banco de Dados (ILNBDs) e apresentam como características dispersão dos dados, velocidade de respostas e simplicidade de programação [15].

A estrutura do código é separada entre definição dos elementos, dicionário com base em *hash table* [16] e a estrutura de controle de fluxo do programa. Também conta com uma função para manipular a entrada do usuário e converter para o formato aceito pelo dicionário criado.

O uso do dicionário nesse caso se justifica pelo conjunto de informações de cada termo que contém variações denominativas, variantes populares e definição. Além disso, a praticidade e eficiência da função *hashing* da própria linguagem Python se tornam grandes aliadas.

Na referida linguagem, a função *hashing* recebe um elemento e retorna a chave como um valor inteiro. No caso de strings, os elementos diferentes possuirão chaves (índices) devido ao valor aleatório que a função proveniente do Python insere nos índices, durante o processo de hashing [16].

Os dicionários são efetivos nesse caso, já que possuem velocidade, além da efetividade. Ambas são características importantes para o desempenho do programa. Para a visualização da rapidez introduzida ao funcionamento, vamos exemplificar a busca pelo primeiro elemento inserido no dicionário técnico, “ancilostomíase”, mostrando o tempo necessário para completar a operação (Figura 8).

```

*****LEXICO-MORFOLOGICO*****
1)Verificar termos disponíveis no dicionário;
2)Buscar palavra, seu significado e seu(s) termo(s) alternativo(s);
3)Sair;
2
Insira o termo a ser buscado:
ancilostomíase

Termo Técnico
Termo : Ancilostomíase
Categoria : Termo médico/Doença
Variação(ões) Denominativa(s) : Ancilostomose
Variante(s) Popular(es) : Amarelão, Marelão, Doença do Jeca Tatu
Definição : Infecção intestinal causada por nematódeos cujos sintomas são palidez, tom amarelado da pele e anemia.

Tempo de busca: 7.00139993896484 ms

```

**Figura 8. Busca por palavra ‘Ancilostomíase’**  
[Fonte: Os autores]

As colisões utilizando esse método (dicionário) são raras, porém não inexistentes. Como citado, a função hashing nativa da linguagem insere um valor aleatório ao utilizar a função *hash()*, de maneira a tornar os valores mais distintos e também para garantir a segurança do processo, o que dificulta o acontecimento de colisões quando tratamos de um vocabulário com um elemento identificador único. É importante observar, porém, que adicionar um termo com um identificador igual a um outro citado, substitui as informações anteriores.

Pode-se observar na Figura 8 que o tempo total para buscar e retornar o primeiro elemento inserido é de 7.001 ms. No dicionário criado para este trabalho, evitamos o referido problema ao estender a definição dos termos disponíveis quando havia qualquer colisão [14] (Figura 9).

Importante lembrar que embora não tenha sido projetado para operar exclusivamente com esse léxico, o dicionário desenvolvido na linguagem de programação Python tem a função de selecionar os verbetes já ordenados por aquele software. Isso também é observado em outros léxicos como o Indicalex [17].

```

amarelao = {
    "Termo" : "Amarelão",
    "Termo(s) técnico(s)" : "Ancilostomíase ou Hepatite",
    "Categoria [Ancilostomíase]" : "Termo Popular Informal/Doença",
    "Variação(ões) Denominativa(s) [Ancilostomíase]" : "Ancilostomose",
    "Variante(s) Popular(es) [Ancilostomíase]" : "Marelão, Doença do Jeca Tatu",
    "Definição [Ancilostomíase]" : "Infecção intestinal causada por nematódeos cujos sintomas são palidez, tom amarelado da pele e anemia.\n",
    "Categoria [Hepatite]" : "Termo Popular Informal/Doença",
    "Variação(ões) Denominativa(s) [Hepatite]" : "-",
    "Variante(s) Popular(es) [Hepatite]" : "Amarelão",
    "Definição [Hepatite]" : "Infecção hepática causada por agentes infecciosos ou tóxicos, Hepatite."
}

```

**Figura 9. Definição da palavra ‘Amarelão’**  
[Fonte: Os autores]

O último elemento inserido é o verbo “vomitar” que pode ser visualizado na Figura 10.

```

Termo Técnico
Termo : Vomitar
Categoria : Termo médico/Sintoma
Variação(ões) Denominativa(s) : Êmese
Variante(s) Popular(es) : Lançá
Definição : Expelir golfo, vômito ou secreção similar.

Tempo de busca: 3.0002593994140625 ms

```

**Figura 10. Busca por palavra ‘Vomitar’**  
[Fonte: Os autores]

É possível notar que o tempo levado para encontrar tal termo é de 3.000 ms, nesse caso. Com o dicionário popular, o tempo para encontrar o primeiro e o segundo termos inseridos (Amarelão e Lançá) é ilustrado na Figura 11. Repetindo essa mesma pesquisa em uma amostra de dez ocorrências, o tempo médio levado para encontrar a palavra ‘Vomitar’ foi de 2.291 ms com um desvio padrão de amostra de 0.494.

```

Termo Popular
Termo : Amarelão
Termo(s) técnico(s) : Ancilostomíase ou Hepatite

Categoria [Ancilostomíase] : Termo Popular Informal/Doença
Variação(ões) Denominativa(s) [Ancilostomíase] : Ancilostomose
Variante(s) Popular(es) [Ancilostomíase] : Marelão, Doença do Jeca Tatu
Definição [Ancilostomíase] : Infecção intestinal causada por nematódeos cujos sintomas são palidez, tom amarelado da pele e anemia.

Categoria [Hepatite] : Termo Popular Informal/Doença
Variação(ões) Denominativa(s) [Hepatite] : -
Variante(s) Popular(es) [Hepatite] : Amarelão
Definição [Hepatite] : Infecção hepática causada por agentes infecciosos ou tóxicos, Hepatite.

Tempo de busca: 4.9991607666015625 ms

```

```

Termo Popular
Termo : Lançá
Categoria : Termo Popular Informal/Sintoma
Termo Técnico : Vomitar
Variação(ões) Denominativa(s) : Êmese
Variante(s) Popular(es) : -
Definição : Expelir, pela boca, com esforço e em golfadas, o alimento ingerido.

Tempo de busca: 1.0027885437011719 ms

```

**Figura 11. Busca por palavras ‘amarelao’ e ‘lanca’**  
[Fonte: Os autores]

A partir desses testes, é perceptível que não há variações de tempo grandes o suficiente para que haja problema. Com exceção do termo ‘Amarelão’ que requer mais dados exibidos, a variação de tempo é pequena e não causa incômodo ao usuário, um dos objetivos de auxílio do léxico.

Para a palavra ‘Amarelão’, em uma amostra de dez ocorrências de busca, o tempo médio foi de 4.29 ms com um desvio padrão de amostra de 0.952. Já para a palavra ‘Lançá’, também em uma amostra de dez ocorrências, o tempo médio de busca foi de 2.658 ms com desvio padrão de amostra de 0.471.

No caso de busca por um termo não existente no vocabulário, temos um tempo de busca em cerca de 1 ms, como exemplificado na Figura 12. Em uma amostra de dez ocorrências buscando por ‘Palavra não existente’, o tempo médio de busca foi de 0.994 ms com desvio padrão de amostra de 0.002.

```

1)Verificar termos disponíveis no dicionário;
2)Buscar palavra, seu significado e seu(s) termo(s) alternativo(s);
3)Sair;
2
Insira o termo a ser buscado:
Palavra não existente

Palavra inserida não está disponível no dicionário

Tempo de busca: 0.9992122650146484 ms

```

**Figura 12. Busca por palavras ‘Nome Não Existente’**  
[Fonte: Os autores]

Cada elemento do dicionário técnico possui os seguintes atributos: Sua chave-geral (o termo técnico simplificado), responsável por identificar o elemento, e os valores, ‘Termo’, ‘Categoria’, ‘Variação(ões) Denominativa(s)’, ‘Variante(s) Popular(es)’ e ‘Definição’. Já elementos do dicionário de termos populares possuem: ‘Termo’, ‘Categoria’, ‘Termo Pesquisado’, ‘Variação(ões) Denominativa(s)’, ‘Variante(s) Popular(es)’ e ‘Definição’.

Para demonstrar visualmente, novamente será utilizado o termo “Articulação” para exemplificar sua definição na Figura 13.

```

13 articulacao = {
14     "Termo Pesquisado" : "Articulação",
15     "Categoria" : "Termo médico/Anatomia",
16     "Variação(ões) Denominativa(s)" : "Junta, Junta, Juntas, Juntura",
17     "Variante(s) Popular(es)" : "Junta do Osso",
18     "Definição" : "Junção ou articulação óssea, joelho."
19 }

```

**Figura 13. Definição da palavra ‘Articulação’**  
[Fonte: Os autores]

Dessa maneira, os elementos são criados já dentro do programa. Para a criação dos dicionários, os elementos e as chaves são associados da seguinte maneira: a chave entre aspas, separada por dois pontos e o elemento definido no escopo.

No código do Léxico aplicado à Saúde temos duas estruturas, para os termos técnicos e populares. Um é chamado de ‘dicionario\_tecnico’ e possui o formato apresentado na Figura 14, e assim como o nome supõe, relaciona os termos utilizados pelos profissionais de saúde aos outros alternativos.

```

dicionario_tecnico = {
    "ancilostomiase" : ancilostomiase,
    "articulacao" : articulacao,
    "ascite" : ascite,
    "boca" : boca,
    "conjuntivite" : conjuntivite,
    "consternacao" : consternacao,
    "constipacao intestinal" : constipacao_intestinal,
    "debilidade muscular" : debilidade_muscular,
    "desidratacao" : desidratacao,
    "diarreia" : diarreia,
    "doenca" : doenca,
    "doente terminal" : doente_terminal,
    "escabiose" : escabiose,
    "estomago" : estomago,
    "fadiga" : fadiga,
    "faringe" : faringe,
    "fratura ossea" : fratura_ossea,
    "hepatite" : hepatite,
    "helmintiase" : helmintiase,
    "invalidez" : invalidez,
    "menopausa" : menopausa,
    "miase" : miase,
    "miliaria" : miliaria,
    "nariz" : nariz,
    "nausea" : nausea,
    "palidez" : palidez,
    "pulmao" : pulmao,
    "resfriado" : resfriado,
    "secrecao" : secrecao,
    "sofrimento fisico" : sofrimento_fisico,
}

```

**Figura 14. Dicionário Técnico** [Fonte: Os autores]

```

dicionario_popular = {
    "amarelao" : amarelao,
    "marelao" : marelao,
    "doenca do jeca tatu": doenca_do_jeca_tatu,
    "junta do osso" : junta_do_osso,
    "barriga d'agua" : barriga_agua,
    "barriga de agua" : barriga_agua,
    "barriga d'agua" : barriga_agua,
    "beico" : beico,
    "dordoi" : dordoi,
    "mocorongo" : mocorongo,
    "perrengue" : perrengue,
    "geratacao" : geratacao,
    "colerina" : colerina,
    "disandar" : disandar,
    "obradera" : obradera,
    "azanga" : azanga,
    "zanga" : zanga,
    "enfermagem" : enfermagem,
    "ismuricido" : ismuricido,
    "dismilinguido" : dismilinguido,
    "catita" : catita,
    "pereba" : pereba,
    "bucho" : estomago,
    "estambo" : estomago,
    "pandu" : estomago,
    "iscornado" : fadiga,
    "guela" : faringe,
}

```

Figura 15. Dicionário Popular [Fonte: Os autores]

Para o controle de fluxo de programa, uma condição de *while* associada a condições de *if* são aninhadas para garantir o funcionamento do código, o qual o encerramento é obtido apenas quando a condição de saída é passada no console (Condição de número 3 da Figura 1).

Todos os testes foram realizados em uma máquina com o sistema operacional Windows 10, processador AMD Ryzen 7 e memória de 16GB DDR4.

### CONCLUSÕES E TRABALHOS FUTUROS

A Língua Portuguesa é rica ao ser vista por vários ângulos, e mais ainda, quando analisada morfológicamente. De maneira orgânica, as palavras possuem diversas flexões e variantes, o que por si, já a torna mais complexa do que muitas outras línguas. Considerando essa dificuldade natural que a própria linguagem traz, a variação regional de sotaques e costumes, elevam o nível de acessibilidade de forma exponencial.

Essa complexidade se importa para a comunicação na área da saúde, a qual já é observada há muito tempo. O desentendimento causado por profissionais da área de saúde por utilizarem somente termos técnicos e o dos pacientes por utilizarem termos populares ou regionais desconhecidos pelos profissionais se torna uma barreira perigosa durante diagnósticos, consultas ou procedimentos corriqueiros.

A partir dessa dor observada no ambiente de saúde, um programa como o apresentado neste trabalho, pode ser de extremo auxílio, se desenvolvido com mais profundidade. Assim, possivelmente um sítio para disponibilização desses recursos possibilitaria o desenvolvimento por parte dos grupos interessados em ações de avaliação conjunta em áreas específicas, como a Medicina.

Embora instrumentos de pesquisa possam auxiliar na busca de tais termos desconhecidos, uma ferramenta especializada tendo como base um trabalho extenso e completo sobre o

léxico aplicado à saúde, se sobressairia de maneira positiva e confiável. Um possível uso, por exemplo, seria um analisador de sentenças completas. Aliando os significados dos termos citados nos dicionários com a análise de sentença natural, seria possível realizar uma “tradução” de uma frase técnica para uma de mais fácil compreensão, conciliando também, a velocidade para encontrar resultados. Muitas vezes isso não é levado em conta em alguns sistemas que são utilizados no âmbito da saúde.

Muitas pesquisas [18] mostram como os pacientes fazem uma correlação com determinada doença, ainda que de maneira pouco elaborada, o que não significa que as orientações recebidas anteriormente foram aprendidas. Assim, este trabalho poderia ter como trabalhos futuros o uso de perguntas (já no enfoque de Análise Sintática do Processamento de Linguagem Natural e não como este trabalho visando um analisador Léxico-morfológico) feitas pelos pacientes acerca de determinadas doenças ou diagnósticos recebidos.

### REFERÊNCIAS

1. Tatiana M. M. Silvestrow, 2017. *Contribuições dos Estudos Terminológicos para os Profissionais da Saúde Básica do SUS*. Tese de Doutorado. Faculdade de Letras da Universidade Federal de Minas Gerais, Belo Horizonte.
2. Ana C. D. Cavalcanti e Maria J. Coelho. 2007. A linguagem como ferramenta do cuidado do enfermeiro em cirurgia cardíaca em *Escola Anna Nery Revista de Enfermagem*, 11, 2 (Jun., 2007), 220-226. <https://doi.org/10.1590/S1414-81452007000200006>
3. Ana C. Sucupira. 2007. A importância da relação médico-paciente e das habilidades de comunicação na formação do profissional de saúde. *Interface – Comunicação, Saúde e Educação*. 11, 23 (Dez., 2007). <https://doi.org/10.1590/S1414-32832007000300016>.
4. Libânio F. V. F. Pinto. 2013. *Léxico e Dicionário. Literacia e comunicação em saúde: eficácia das estratégias de comunicação digital em saúde: doenças respiratórias crônicas*. Dissertação de Mestrado. Faculdade de Medicina da Universidade de Lisboa, Lisboa.
5. Lucas B. Galhardi, Cinthyan R. S. C. de Barbosa, Joao Coelho Neto e Jacques D. Brancher. 2018. Analisador Léxico-Morfológico de Redações de Estudantes no Estilo do ENEM. In *Anais da XXIII Conferência Internacional sobre Informática na Educação (TISE'18). Nuevas Ideas en Informática Educativa, Volumen 14*, Santiago de Chile, 509-513.
6. Amanda D. Souza e Eduardo R. Felipe. 2021. Processamento de Linguagem Natural aplicado à anamneses do domínio da ginecologia. *Fronteiras da Representação do Conhecimento*, 1, 2 (Dez, 2021).
7. Felipe R. Pereira e Sandro J. Rigo. 2013. Utilização de Processamento de Linguagem Natural e Ontologias na

- Análise Qualitativa de Frases Curtas. *Renote*. 11, 3 (Dez., 2013). <http://dx.doi.org/10.22456/1679-1916.44431>.
8. Cinthyan R. S. C. de Barbosa. 1998. *Gramática para Consultas Radiológicas em Língua Portuguesa*. Dissertação de Mestrado. Departamento de Informática da Universidade Federal do Rio Grande do Sul, Porto Alegre.
  9. Yago H. Pereira, Cinthyan R. S. C. de Barbosa, Arthur A. Artoni, José L. V. M. Mioni e Jacques D. Brancher. 2020. Interface em Linguagem Natural para uma Sublinguagem de Câncer de Pele. In *Anais do XI Computer on the Beach*, Balneário Camburiú, 469-474. <https://doi.org/10.14210/cotb.v11n1.p469-474>.
  10. Liliana Ferreira, António Teixeira e João P. Cunha. 2008. Linguateca e Processamento de Linguagem Natural na área de Saúde: alguns comentários e sugestões. In *Actas do Encontro Linguateca: 10 anos*. 43-48. Perspectiva sobre a Linguateca.
  11. Diego H. P. Benicio. 2020. *Aplicação de Mineração de Texto e Processamento de Linguagem Natural em Prontuários Eletrônicos de Pacientes para Extração de Transformação de Texto em Dado Estruturado*. Dissertação de Mestrado. Programa de Pós-Graduação em Tecnologia da Informação da Universidade Federal do Rio Grande do Norte, Natal.
  12. Lyana C. e Souza, H. Silva, E. G. Silva, Márcia C. A. F. Moraes e Luciana P. Maximino. 2015. TEO: um processador de linguagem natural para otimização da assistência de crianças com fissuras labiopalatinas e anomalias craniofaciais. In *Anais do IV Simpósio Internacional de Fissuras Orofaciais e Anomalias Relacionadas. Hospital de Reabilitação de Anomalias Craniofaciais da Universidade Estadual de São Paulo*. Bauru.
  13. Rodrigo Macedo. 2018. *Introdução a Processamento de Linguagem Natural*. Recuperado Setembro 26, 2022 de <https://medium.com/ensina-ai/introdu%C3%A7%C3%A3o-a-processamento-de-linguagem-natural-174936c096b>
  14. Python Standard Library. *Built in Functions. Documentation*. Retrieved September 26, 2022 from <https://docs.python.org/3.7/library/functions.html#hash>
  15. Moreno, F. C., Barbosa, C. R. S. C. e Manfio, E. R. 2021. Tabelas Hash para um Léxico Digital. *Revista de Informática Teórica e Aplicada (RITA)*, 28, 2 (Ago., 2021) 26-38. <https://doi.org/10.22456/2175-2745.107128>
  16. Davide Matromatted. *Python Hashtable: Understanding dictionaries. The Python Corner*. Retrieved September 26, 2022 from <https://thepythoncorner.com/posts/2020-08-21-hash-tables-understanding-dictionaries/>
  17. Fabiane C. Altino, Celciane A. Vasconcelos, Cinthyan R. S. C. de Barbosa, Edio R. Manfio, Fábio C. Moreno, Maria R. M. Gobbo e Marcelo O. Sousa. 2019. Indicalex: áreas do saber a serviço do estudo da língua. In *Anais do XLVI Seminário Integrado de Software e Hardware (SEMISH'19)*, Belém, 1-11. <https://doi.org/10.5753/semish.2019.6562>
  18. Magda C. Q. Dell'Acqua, Janete Pessuto, Silvia C. M. Bocchi e Renata C. P. M. dos Anjos. 1997. Comunicação da equipe multiprofissional e indivíduos portadores de hipertensão arterial. *Revista Latino-Americana de Enfermagem*, 5, 3 (Jul., 1997), 43-48. <https://doi.org/10.1590/S010412902014000400019>.