



**UNIVERSIDADE FEDERAL DA FRONTEIRA SUL  
CAMPUS CHAPECÓ  
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGUÍSTICOS (PPGEL)  
CURSO DE MESTRADO**

**IRDES MELLYNA BRANCO**

**ANGLICISMOS NO PORTUGUÊS BRASILEIRO: UM ESTUDO SOCIOLINGUÍSTICO  
EM *TWEETS* DO OESTE CATARINENSE**

Chapecó

2021

**IRDES MELLYNA BRANCO**

**ANGLICISMOS NO PORTUGUÊS BRASILEIRO: UM ESTUDO SOCIOLINGUÍSTICO**  
EM *TWEETS* DO OESTE CATARINENSE

Dissertação apresentada ao Programa de Pós-Graduação em Estudos Linguísticos – PPGEL – da Universidade Federal da Fronteira Sul – UFFS – como requisito para obtenção do título de Mestre em Estudos Linguísticos.

Orientadora: Prof.<sup>a</sup> Dr.<sup>a</sup> Cláudia Andrea Rost Snichelotto.

Chapecó

2021

**Bibliotecas da Universidade Federal da Fronteira Sul - UFFS**

Branco, Irdes Melyna

Anglicismos no português brasileiro: um estudo sociolinguístico em tweets do oeste catarinense / Irdes Melyna Branco. -- 2021.  
150 f. : il.

Orientadora: Prof.ª Dr.ª Cláudia Andrea Rost Snichelotto

Dissertação (Mestrado) - Universidade Federal da Fronteira Sul, Programa de Pós-Graduação em Estudos Linguísticos Chapecó, SC, 2021.

1. Anglicismo. 2. Ciberescrita. 3. Twitter. 4. Sociolinguística. 5. Empréstimo Linguístico. I. Snichelotto, Cláudia Andrea Rost, orient. II. Universidade Federal da Fronteira Sul. III. Título.

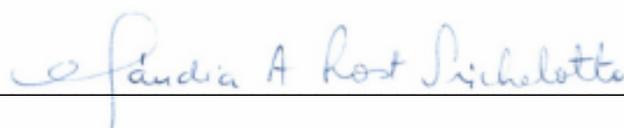
**IRDES MELLYNA BRANCO**

**ANGLICISMOS NO PORTUGUÊS BRASILEIRO: UM ESTUDO SOCIOLINGUÍSTICO  
EM *TWEETS* DO OESTE CATARINENSE**

Dissertação apresentada ao Programa de Pós-Graduação em Estudos Linguísticos – PPGEL – da Universidade Federal da Fronteira Sul – UFFS – como requisito para obtenção do título de Mestre em Estudos Linguísticos.

Este trabalho foi defendido e aprovado em 17 de dezembro de 2021.

**BANCA EXAMINADORA**



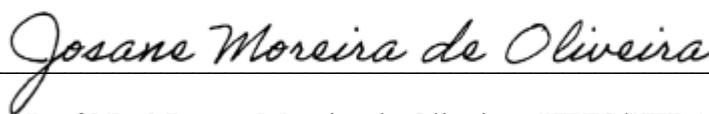
Prof.<sup>a</sup> Dr.<sup>a</sup> Cláudia Andrea Rost Snichelotto - UFFS

Orientadora e Presidente



Prof.<sup>a</sup> Dr.<sup>a</sup> Morgana Fabiola Cambrussi - UFFS

Avaliadora interna



Prof.<sup>a</sup> Dr.<sup>a</sup> Josane Moreira de Oliveira - UEFS/UFBA

Avaliadora externa

---

Prof.<sup>a</sup> Dr.<sup>a</sup> Raquel Gomes Chaves - UNESPAR

Avaliadora externa suplente

---

Prof.<sup>a</sup> Dr.<sup>a</sup> Athany Gutierrez - UFFS

Avaliadora interna suplente

À tia Dianete, que nos deixou este ano tão precocemente, mas que,  
enquanto pode, viveu em plenitude, dedico.

## AGRADECIMENTOS

Aos meus pais, Elenice e Wilamir, minha irmã, Sacha, e meus avós, Irdes e Jaime, por terem me dado o suporte e apoio necessários para continuar.

À professora Claudia Andrea Rost Snichelotto, minha orientadora, pela orientação, paciência e confiança, indispensáveis para a construção desta pesquisa, sempre com leituras atentas e palavras de conforto e motivação.

Às professoras, Morgana Fabiola Cambrussi, Josane Moreira de Oliveira, Athany Gutierrez e Raquel Gomes Chaves, membros das bancas de qualificação e defesa, agradeço pelas leituras dedicadas e pelos apontamentos essenciais para a construção deste trabalho.

À Universidade Federal da Fronteira Sul (UFFS), por tornar possível esta formação e oportunizar o crescimento da minha trajetória acadêmica e profissional e, principalmente, aos professores do Programa de Pós-Graduação em Estudos Linguísticos (PPGEL), por toda a troca de conhecimento e aprendizado, essenciais para a minha formação, em especial, à professora Aline Peixoto Gravina, por todos os ensinamentos repassados durante o estágio de docência, por trabalhar de forma séria e dedicada e ao mesmo tempo gentil.

Aos meus colegas de mestrado, pela troca de experiências, especialmente à Vanessa, por todo suporte e ajuda durante todo o percurso, à Cíntia e à Sara, pelos incentivos, ao Bruno, pelo suporte com o *software* R, ao Luan, pelas palavras de conforto, e a Elis, pela companhia durante as longas viagens até Chapecó.

À Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina (FAPESC), pelo auxílio financeiro a partir de 2020.

Às minhas amigas, Karine, Mariana, Rebeca e Samara, pelo incondicional incentivo.

À minha família carioca, por tudo o que me oportunizaram ao longo de minha vida, permitindo que eu conquistasse novos objetivos.

Aos meus padrinhos postigos, Lili e Célio, e à minha irmã de coração, Letícia, por toda a ajuda, os chás calmantes e as distrações necessárias em horas oportunas.

Enfim, agradeço a todos que, de alguma forma, contribuíram para que eu chegasse até aqui.

“Os idiomas são palco de mestiçagem e de interculturalidade.”

*John Robert Schmitz*

## RESUMO

O empréstimo linguístico é a incorporação de um vocábulo estrangeiro à língua receptora, já o estrangeirismo é um termo ou expressão alóctone que não está incorporado à língua que o recebe. O neologismo, por sua vez, é a criação de novas palavras, formadas ou derivadas de palavras já existentes, podendo ser da mesma língua ou não. A internet possibilita uma comunicação sem fronteiras e muito mais rápida que os meios de comunicação convencionais. O presente trabalho teve como objetivo descrever e analisar, sob a perspectiva Sociolinguística, a inserção de vocábulos oriundos da língua inglesa em *tweets* na ciberescrita em português. Levando em consideração as rápidas mudanças no meio cibernético e a velocidade com que os indivíduos trocam informações, este estudo sincrônico se faz necessário para que se possa identificar uma tendência de incremento de anglicismos no português brasileiro. Nossa amostra foi composta por 50 usuários do *Twitter*, estratificados em sexo, idade, ocupação e localização, selecionados por meio de uma busca avançada disponível na própria rede social. Foram coletados no total 2450 anglicismos de publicações da rede social *Twitter*, escritos por usuários localizados na região oeste catarinense, distribuídos em dois períodos de tempo: no primeiro, foram coletados *tweets* publicados entre março de 2019 e fevereiro de 2020, isto é, antes da declaração da pandemia da Covid-19; e, no segundo, considerando a necessidade de isolamento e distanciamento social, foram coletados *tweets* publicados entre março de 2020 e fevereiro de 2021. Nossa hipótese era que, devido à pandemia da Covid-19, haveria incremento no uso de anglicismos por consequência do aumento do trabalho e do estudo remotos e da maior necessidade de uso dos recursos tecnológicos virtuais no segundo período. A análise quantitativa dos dados, realizada por meio do software R (R CORE TEAM, 2020), revelou que, das sete variáveis independentes controladas, foram selecionadas como significativas apenas três: as variáveis linguísticas assimilação ou não dos termos estrangeiros ao sistema linguístico do português e classe morfológica dos anglicismos e a variável social faixa etária dos usuários do *Twitter* da amostra. Em síntese, os resultados revelaram que: houve um aumento no uso de anglicismos no segundo período analisado; os itens lexicais não assimilados e não-dicionarizados foram os mais recorrentes; os estrangeirismos são mais frequentes em relação aos empréstimos e aos neologismos na amostra nos dois períodos; o substantivo foi a classe morfológica com o maior número de ocorrências; os homens empregaram mais anglicismos que as mulheres nos dois períodos; os mais jovens usaram mais anglicismos que os mais velhos nos dois períodos; Chapecó é a microrregião com maior número de ocorrências; estudantes de ensino superior são os usuários que mais empregam anglicismos na amostra extraída de usuários do *Twitter*.

**Palavras-chave:** Empréstimo linguístico, estrangeirismo, anglicismo, ciberescrita, Twitter.

## ABSTRACT

The loanword is the incorporation of a foreign word into the receiving language, while foreignness is an allochthonous term or expression that is not incorporated into the receiving language. The neologism, in turn, is a language derived from new words or derived from existing words, which may or may not be the same. The internet makes communication possible without borders and much faster than the conventional means of communication. The present work aimed to describe and analyze, from a sociolinguistic perspective, the insertion of words from the English language in tweets in cyberwriting in Portuguese. Considering the fast changes in the cyber environment and the speed with which individuals exchange information, this synchronous study is necessary to identify (or not) an increasing trend in loanwords and anglicisms in Brazilian Portuguese. Our sample consisted of 50 Twitter users, selected through an advanced search available on the social network itself. A total of 2450 anglicisms of publications on the social network Twitter, written by users located in the western region of Santa Catarina, posted by users between March 2019 and February 2020, that is, before the declaration of the Covid-19 pandemic; and, in the second, considering the need for isolation and social distance, which made many changes in a professional and academic activities, were collected tweets published by users between March 2020 and February 2021. Our hypothesis was that, because of the Covid-19 pandemic, there was an increase in the use of anglicisms as a consequence of the increase in remote work and study and the use of virtual technological resources in the second course. From the quantitative analysis of our sample data, carried out using the R software (R CORE TEAM, 2020), it was found that of the seven independent variables controlled, only three were selected as significant: the linguistic variables assimilation or not of terms foreign to the Portuguese linguistic system and morphological class of anglicisms and the social variable age of the Twitter users in the sample. In summary, the results revealed that: there was an increase in the use of anglicisms in the second course analyzed; non-assimilated and non-dictionary lexical items were the most recurrent; foreign words are more frequent in relation to loans and neologisms in the sample in both time courses; the noun was the morphological class with the highest number of occurrences; men wrote more than women in both time courses; the younger ones used more anglicisms than the older ones in both time courses; Chapecó is the microregion with the highest number of occurrences; university students are the ones that most use anglicisms in the sample extracted from Twitter users.

**Keywords:** loanwords, foreignism, anglicism, cyberwriting, Twitter.

## LISTA DE ABREVIATURAS E SIGLAS

**CCJC** - Comissão de Constituição e Justiça e de Cidadania

**Cetic** - Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação

**CMC** - Comunicação mediada por computador

**CNPq** - Conselho Nacional de Desenvolvimento Científico e Tecnológico

**Covid-19** - *Coronavirus disease* (doença do coronavírus)

**FAPESC** - Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina

**IBGE** - Instituto Brasileiro de Geografia e Estatística

**INDL** - Inventário Nacional da Diversidade Linguística

**IPHAN** - Instituto do Patrimônio Histórico e Artístico Nacional

**LIBRAS** - Língua Brasileira de Sinais

**UFFS** - Universidade Federal da Fronteira Sul

**P1** - Primeiro período de análise (03/2019 a 03/2020)

**P2** - Segundo período de análise (03/2020 a 03/2021)

**PB** - Português Brasileiro

**PE** - Português Europeu

**PNAD** - Pesquisa Nacional por Amostra de Domicílios Contínua

**PPGEL** - Programa de Pós-Graduação em Estudos Linguísticos

**TICs** - Tecnologias da Informação e Comunicação

**VMPOSC** - Variação e Mudança no Português do Oeste de Santa Catarina

## LISTA DE FIGURAS

<b>Figura 1</b> - Captura da tela da página inicial de usuário do <i>Twitter</i> , em janeiro de 2021 .....	76
<b>Figura 2</b> - Mapa de Santa Catarina, com a mesorregião oeste destacada e dividida em suas cinco microrregiões .....	82
<b>Figura 3</b> - Busca avançada a partir de períodos no <i>Twitter</i> .....	86
<b>Figura 4</b> - Captura de tela com a categorização dos dados do usuário 1 .....	88
<b>Figura 5</b> - Usuário(a) do <i>Twitter</i> com informações públicas no perfil .....	115

## LISTA DE GRÁFICOS

<b>Gráfico 1</b> - Distribuição geral dos tweets com anglicismos por microrregiões do Oeste de Santa Catarina .....	87
<b>Gráfico 2</b> - Distribuição dos tweets com anglicismos separados por período .....	92
<b>Gráfico 3</b> - Anglicismos dicionarizados do primeiro período, separados por classe morfológica .....	96
<b>Gráfico 4</b> - Anglicismos não-dicionarizados do primeiro período, separados por classe morfológica .....	97
<b>Gráfico 5</b> - Anglicismos dicionarizados do segundo período, separados por classe morfológica.....	99
<b>Gráfico 6</b> - Anglicismos não-dicionarizados do segundo período, separados por classe morfológica .....	99
<b>Gráfico 7</b> - Anglicismos encontrados nos tweets do período P1 separados por classe morfológica .....	105
<b>Gráfico 8</b> - Anglicismos encontrados nos tweets do período P2 separados por classe morfológica .....	105
<b>Gráfico 9</b> - Anglicismos encontrados nos tweets do período P1 separados por sexo .....	108
<b>Gráfico 10</b> - Anglicismos encontrados nos <i>tweets</i> do período P2 separados por sexo .....	109
<b>Gráfico 11</b> - Anglicismos encontrados nos tweets do período P1 separados por faixa etária ..	112
<b>Gráfico 12</b> - Anglicismos encontrados nos tweets do período P2 separados por faixa etária ..	112
<b>Gráfico 13</b> - Anglicismos encontrados nos tweets do período P1 separados por microrregião .....	117
<b>Gráfico 14</b> - Anglicismos encontrados nos tweets do período P2 separados por microrregião .....	117
<b>Gráfico 15</b> - Anglicismos encontrados nos tweets do período P1 separados por ocupação ....	125

**Gráfico 16** - Anglicismos encontrados nos tweets do período P2 separados por ocupação ..... 124

## LISTA DE QUADROS

<b>Quadro 1</b> - Dados de Torrano (2010) .....	60
<b>Quadro 2</b> - Dados de Prado (2015) .....	63
<b>Quadro 3</b> - Dados de Castro (2017) .....	63
<b>Quadro 4</b> - Dados de Souza (2015) .....	64
<b>Quadro 5</b> - Resumo dos dados coletados pelos autores estudados .....	65
<b>Quadro 6</b> - Fatores sociais e linguísticos das pesquisas utilizadas neste trabalho .....	68
<b>Quadro 7</b> - Resumo das variáveis linguísticas e sociais controladas .....	86

## LISTA DE TABELAS

<b>Tabela 1</b> - Quem usa o Twitter? .....	84
<b>Tabela 2</b> - Frequência e proporção de anglicismos mais recorrentes nos <i>tweets</i> separados por período .....	93
<b>Tabela 3</b> - Frequência e proporção de anglicismos dicionarizados ou não, separados por P1 e P2 .....	96
<b>Tabela 4</b> - Anglicismos coletados nos <i>tweets</i> do P1 separados por classe morfológica e dicionarizados e não-dicionarizados. ....	98
<b>Tabela 5</b> - Anglicismos coletados nos <i>tweets</i> do P2 separados por classe morfológica e dicionarizados e não-dicionarizados .....	100
<b>Tabela 6</b> - Frequência e proporção de anglicismos dicionarizados separados por período .....	101
<b>Tabela 7</b> - Frequência e proporção de anglicismos não-dicionarizados separados por período .....	102
<b>Tabela 8</b> - Frequência e proporção de anglicismos assimilados ou não, separados por P1 e P2 .....	104
<b>Tabela 9</b> - Frequência e proporção de anglicismos assimilados ou não, separados por P1 e P2 .....	107
<b>Tabela 10</b> - Anglicismos mais recorrentes entre o público feminino, segundo P1 e P2 .....	113
<b>Tabela 11</b> - Anglicismos mais recorrentes entre o público masculino, segundo P1 e P2 .....	114
<b>Tabela 12</b> - Frequência e proporção de anglicismos assimilados ou não, separados por P1 e P2 .....	116
<b>Tabela 13</b> - Itens lexicais mais frequentes em <i>tweets</i> da faixa etária de 16 a 24 anos em P1 e P2 .....	118
<b>Tabela 14</b> - Itens lexicais mais frequentes em <i>tweets</i> da faixa etária de 25 a 34 anos em P1 e P2 .....	119

<b>Tabela 15</b> - Itens lexicais mais frequentes em tweets da faixa etária de 35 a 44 anos em P1 e P2 .....	120
<b>Tabela 16</b> - Itens lexicais mais frequentes em tweets da faixa etária de 45 a 54 anos em P1 e P2 .....	121
<b>Tabela 17</b> - Itens lexicais mais frequentes em <i>tweets</i> da microrregião de Chapecó em P1 e P2 .....	125
<b>Tabela 18</b> - Anglicismos coletados nos tweets do P1 separados por ocupação .....	128
<b>Tabela 19</b> - Anglicismos coletados nos tweets do P2 separados por ocupação .....	129

## SUMÁRIO

<b>1. INTRODUÇÃO</b>	<b>21</b>
<b>2. REFERENCIAL TEÓRICO</b>	<b>29</b>
2.1 A Sociolinguística Variacionista	29
2.2 A Metodologia da Pesquisa Variacionista	33
2.3 Os níveis linguísticos de variação	35
2.3.1 A variação lexical	37
2.4 A mudança sob a perspectiva da Sociolinguística variacionista	38
2.5 O ciberespaço e a ciberescrita	40
2.6 Léxico e lexicalização	43
2.6.1 Lexicalização por empréstimo	45
<b>3. REVISÃO DE LITERATURA</b>	<b>48</b>
3.1 O composto do léxico da Língua Portuguesa	48
3.2 Empréstimo linguístico, estrangeirismo ou neologismo?	50
3.2.1 Definição em dicionários	50
3.2.2 Definição em gramáticas descritivas	53
3.3 Anglicismo	54
3.3.1 Pesquisas sobre anglicismos no português do Brasil	55
<b>4. OBJETIVOS, QUESTÕES E HIPÓTESES</b>	<b>72</b>
4.1 Objetivo geral	72
4.2 Objetivos específicos	72
4.3 Questões e hipóteses	72
<b>5. METODOLOGIA</b>	<b>76</b>
5.1 A rede social Twitter	76
5.2 A sócio-história do oeste catarinense	80
5.3 A amostra	83
5.4 Os procedimentos de coleta de dados	86
5.5 Análise qualitativa dos dados	87
5.6 Restrições	89
5.7 Análise quantitativa dos dados: o Software R	90
<b>6. DESCRIÇÃO E ANÁLISE DOS DADOS</b>	<b>92</b>
6.1 Frequência geral	92
6.2 Variáveis linguísticas	95
6.2.1 Presença/ausência de anglicismos nos dicionários	95
6.2.3.1 Resultados e discussão	95
6.2.2 Anglicismos assimilados ou não ao português	102
6.2.3.1 Resultados e discussão	104
6.2.3 Classe gramatical	106

6.2.3.1 Resultados e discussão	107
6.3 Variáveis extralinguísticas	110
6.3.1 Sexo	110
6.3.1.2 Resultados e discussão	111
6.3.2 Faixa etária	115
6.3.2.2 Resultados e discussão	116
6.3.3 Localidade	122
6.3.3.2 Resultados e discussão	123
6.3.4 Ocupação	126
6.3.4.2 Resultados e discussão	126
<b>7. ORIGEM E POTENCIALIDADE SEMÂNTICO-PRAGMÁTICA DE ALGUNS ANGLICISMOS PRESENTES EM TWEETS</b>	<b>132</b>
<b>8. CONSIDERAÇÕES FINAIS</b>	<b>139</b>
<b>REFERÊNCIAS</b>	<b>143</b>

## 1. INTRODUÇÃO<sup>1</sup>

Durante interações comunicativas, a maioria dos falantes é capaz de diferenciar palavras da sua variedade de língua materna<sup>2</sup> daquelas oriundas de outros vernáculos; a exemplo disso, hoje temos, no português do Brasil, a ocorrência de vocábulos do inglês, como *on-line* e *like*. Estas palavras, a princípio, causariam estranhamento aos falantes do Português Brasileiro (doravante PB) entretanto hoje lhes são muito familiares devido ao uso cada vez mais frequente não só na fala como na escrita. A seguir, podemos ver duas ocorrências do uso destes vocábulos extraídas de *tweets* (ou tuítes)<sup>3</sup> da rede social *Twitter*<sup>4</sup>:

(1) “A gente percebe que se tornou adulto quando [está]tá pensando em trancar o cursinho pois [está]tá tendo aula *on-line*<sup>5</sup> e pagando MUITO mais que um cursinho *online* custa kkk” (publicado em 4 de maio de 2020, em Chapecó, grifos nossos);

(2) “[...] cada fav[favoritar]<sup>6</sup> é um fato por favor deem *like*<sup>7</sup> eu estou triste” (publicado em 16 de novembro de 2020, em Chapecó, grifo nosso).

Há aspectos importantes, nos dois *tweets* acima, que precisam ser considerados. O primeiro diz respeito ao gênero discursivo/tipo de texto e o segundo refere-se ao registro (formal ou informal) e ao suporte textual (virtual). Trata-se de uma rede social e microblog em que usuários comuns<sup>8</sup> conversam instantânea e espontaneamente entre si, reproduzindo um bate-papo informal. Utilizam uma ciberlinguagem que se caracteriza fortemente pela economia na escrita das palavras, sobretudo através de abreviações e redução de palavras, como “tá” e “fav” em (1) e (2). De acordo com Xavier (2005), nos momentos de intensas trocas verbais, abreviações e reduções em palavras e expressões são não só necessárias quanto esperadas, o que leva os usuários da internet a ter agilidade no manuseio do teclado. Essa economia, contudo, não se

<sup>1</sup> Esta pesquisa foi fomentada com bolsa de Mestrado do Programa FAPESC de Recursos Humanos em CTI - PPGEL (Edital nº 1081/GR/UFGS/2019).

<sup>2</sup> A língua materna caracteriza, geralmente, a origem e é usada, na maioria das vezes, no dia-a-dia. A língua materna, ou a primeira língua (L1) não é, necessariamente, a língua da mãe, nem a primeira língua que se aprende. Tampouco trata-se de apenas uma língua. Normalmente é a língua que aprendemos primeiro e em casa, através dos pais, e também é frequentemente a língua da comunidade. (SPINASSÉ, 2006)

<sup>3</sup> Os *tweets* (ou tuítes) são mensagens instantâneas de até 280 caracteres que são enviadas pela rede social *Twitter* via Internet.

<sup>4</sup> O *Twitter* é uma mescla de rede social e microblog para troca de mensagens entre usuários. Na seção 5.1, vamos detalhar mais essa fonte de onde extraímos nossos dados.

<sup>5</sup> remota com transmissão simultânea, tradução nossa.

<sup>6</sup> Daqui em diante, inserimos a escrita ortográfica das palavras entre colchetes para facilitar a compreensão do leitor.

<sup>7</sup> um curtir/gostar, tradução nossa.

<sup>8</sup> Há também usuários que representam instituições públicas e privadas e empregam uma linguagem mais formal e culta. Para mais informações, veja Freitas e Barth (2015).

restringe exclusivamente às abreviações, nem é a única característica da ciberlinguagem, mas que apresenta traços linguísticos bem marcados. (BARROS, 2006).

Outro aspecto a ser observado é que, por se tratar de um texto instantâneo, os termos *on-line* e *like* mantêm a escrita (e a pronúncia, no caso dos textos falados) original da língua-fonte, no caso o inglês. Há quem atribua o uso gratuito ou excessivo de estrangeirismo pretensioso e pedante e, por essa razão, a orientação é de que, no texto jornalístico, por exemplo, palavras dessa natureza sejam grafadas entre aspas e acompanhadas de explicação entre parênteses, por exemplo, “*spread*” (taxa de risco nos empréstimos internacionais). Também se recomenda que se traduza até mesmo palavras que se julguem muito conhecidas, pois nem todos os leitores saberão o significado dessas palavras. (MARTINS FILHO, 1997, p. 210). Na escrita formal da língua portuguesa, as palavras estrangeiras, segundo Luft (1991), que ainda não foram aportuguesadas devem guardar a sua grafia originária e, no texto vernáculo, devem ser escritas entre aspas ou grifadas.

Observamos que a rede social, nas suas regras de uso, permite grifos (em itálico e em negrito) no texto das publicações, mas isso deve ser feito fora da página ou do aplicativo do *Twitter* no dispositivo *iOS* ou *Android*, o que pode comprometer a agilidade dos usuários na postagem da mensagem. Além da pouca presteza na interação, nossa hipótese para a ausência de destaque dos termos estrangeiros se deve ao registro do termo *on-line* em dicionários do português (HOUAISS; VILLAR, 2008; FERREIRA, 2004):

[Inglês, de *on* (exprimindo a ideia de ‘posição em ou sobre; continuidade; funcionamento’), + *-line*, ‘linha’.]

1. *Advérbio*. Em conexão com, ou entre, sistemas de processamento e/ou transmissão de informação;
2. *Adjetivo [informática]*. Diz-se de equipamentos periféricos, arquivos, etc. que estão conectados a um sistema computacional, ou acessíveis para a utilização por este: *impressora on-line; banco de dados on-line*.
3. *Por extensão*. Diz-se do processo, ou atividade muito ágil e capaz de responder rapidamente a estímulos externos: *comunicação on-line*.
4. *Informática*. Nos sistemas distribuídos, diz-se da operação realizada em conexão com outros pontos do sistema, permitindo compartilhamento de informações e colaboração no processamento, como, p. ex., consultas e atualizações imediatas de bases de dados remotas. [Opõe-se à *offline*.](FERREIRA, 2004, p. 1440, grifos do autor)

Embora a palavra *on-line* esteja dicionarizada e devido a isso não apresente os destaques sugeridos acima, o termo *like* ainda não está registrado nos dicionários da língua portuguesa, ou seja, por se tratar de um texto espontâneo e instantâneo, no tuíte (2), o usuário não usou aspas, que estão acessíveis no teclado, seja do celular ou do computador, tampouco usou outro grifo.

Para cotejo com a língua inglesa, extraímos do Cambridge *Dictionary* (inglês-português) a definição de *like*<sup>9</sup>:

1. Gostar de; curtir (verbo). Exemplos: *I like to paint in my spare time*<sup>10</sup>; *He really likes her*<sup>11</sup>; *What do you like about him?*<sup>12</sup>
2. Como (preposição) - similar a alguém ou alguma coisa. Exemplo: *He looks like his father*<sup>13</sup>.
3. Como (conjunção) - informal. Exemplo: *Do it exactly like I told you to*<sup>14</sup>. (CAMBRIDGE, 2020, não paginado)

Outra hipótese para o não fornecimento do significado em português aos usuários do Twitter das palavras *on-line* e *like* de (1) e (2) se deve ao uso recorrente desses termos nas redes sociais. Observa-se que ambos estão adequadamente integrados ao português, ou seja, “um empréstimo verdadeiro (em oposição ao uso apenas esporádico de uma palavra não nativa) se integra na fonologia, morfologia e sintaxe da língua.” (BYBEE, 2020, p. 331).

Nas duas ocorrências, o termo *on-line* em (1), grafado com e sem hífen, é um nome (adjetivo) que atua, sintaticamente, como adjunto adnominal dos substantivos *aula* e *cursinho*. Em termos semânticos, *on-line* significa em tempo real e remoto. Em (2), o termo *like* é um nome (substantivo) que atua sintaticamente como objeto direto do verbo *dar*. Logo a expressão *dar um like*, em português, significa curtir o *tweet*, ou seja, gostar/apreciar a publicação (o tuíte) de um usuário da rede social. Na linguagem virtual, os usuários de redes sociais, como *Twitter*, *Facebook* e *Instagram*, empregam o termo *like* como uma ferramenta cuja função é a de indicar que as fotos, os *links*, as notícias postadas foram apreciadas (FREITAS; BARTH, 2015). Essas propriedades demonstram o potencial de adaptação dos empréstimos.

O estrangeirismo, segundo Garcez e Zilles (2001), consiste no emprego, na língua de uma comunidade, de palavras de outras línguas, mais especificamente, no caso do Brasil, o fenômeno ocorre com o uso de vocábulos e expressões estrangeiras no português. Porém, Bagno (2012) chama a atenção para a falta de adequação do termo empréstimo já que o ato de *emprestar* pressupõem a ação de *devolver*, aspecto que não ocorre no caso dos empréstimos linguísticos. Conforme Bybee (2020), a maioria das línguas toma palavras emprestadas de outras línguas, principalmente quando novos artefatos e conceitos são introduzidos a partir de outra cultura.

<sup>9</sup> Direito autoral da *Cambridge University Press* 2021.

<sup>10</sup> Eu gostaria de pintar em meu tempo livre (tradução nossa).

<sup>11</sup> Ela realmente gosta dela (tradução nossa).

<sup>12</sup> O que você gosta nele? (tradução nossa).

<sup>13</sup> Ele se parece com o pai dele (tradução nossa).

<sup>14</sup> Faça exatamente como eu disse (tradução nossa).

Ilari (2002) afirma que o acréscimo de palavras estrangeiras na língua portuguesa do Brasil já foi (e continua sendo!) visto como um problema por gramáticos, escritores e políticos sob o argumento - sem provas - de que as palavras estrangeiras corrompem o português do Brasil e, por isso, estas introduções deveriam ser combatidas. A ditadura Vargas (1930-1945) ficou marcada, dentre tantos aspectos políticos e econômicos, pelo impedimento do uso de empréstimos na língua portuguesa. De acordo com Zilles (2001, p. 146), nessa época, os imigrantes e seus descendentes foram proibidos de usar suas línguas de origem e tiveram até mesmo casas invadidas, livros queimados, escolas fechadas, além da proibição de dar aulas em outras línguas que não o português. Infelizmente essas restrições não pararam nesse período. A partir da Constituição de 1988, ainda se observa o esforço na promoção de projetos de leis e outras tentativas de abolir os estrangeirismos da língua portuguesa (ZILLES, 2001). A tentativa de eliminar os estrangeirismos mais marcante da história recente do país é o projeto de lei nº 1676/1999, do então deputado federal Aldo Rabelo, que foi apresentado à Assembleia Legislativa em 19 de abril de 2000 e tinha como proposta promover a “proteção, a defesa e o uso da língua portuguesa”, mas sobretudo coibir e multar o uso excessivo de estrangeirismos. Conforme Zilles (2001), tal projeto impunha a qualquer estrangeiro que vivesse há mais de um ano no Brasil o uso do português nos domínios públicos, dentre eles nos locais de trabalho. Vale ressaltar que a última ação legislativa sobre o respectivo projeto de lei de Rebelo, aconteceu em 13 de dezembro de 2007, quando a Comissão de Constituição e Justiça e de Cidadania (CCJC) aprovou por unanimidade tal proposta, no entanto a proposição ainda segue sujeita à apreciação do Plenário.

Há outros projetos de lei neste sentido tanto na esfera federal quanto nas esferas estaduais. Destacamos, a seguir, o projeto de lei nº 1.736-B/1996<sup>15</sup> do então deputado federal e atual presidente da República Federativa do Brasil, Jair Messias Bolsonaro, que tinha como intenção proibir o uso de vocábulos estrangeiros no PB. Vale destacar o seguinte trecho da justificativa do projeto:

Com efeito, cumpre adotar todas as precauções possíveis **contra o processo de desvalorização da língua portuguesa**. Trata-se no caso, de evitar a opção crescente por vocábulos de origem estrangeira para designar estabelecimentos comerciais e produtos em circulação no mercado, quando se sabe da existência de **termos congêneres, de formas até mais corretas e adequadas de expressão**, em nossa própria língua, quando se sabe, enfim, da indiscutível riqueza vocabular da língua portuguesa, que conta com mais de 400 mil

---

<sup>15</sup> O projeto de lei pode ser conferido na íntegra na página da Câmara Legislativa no link <[https://www.camara.leg.br/proposicoesWeb/prop\\_mostrarintegra;jsessionid=F0521145122C98F486D7302AE49288FD.proposicoesWebExterno1?codteor=1133712&filename=Dossie+-PL+1736/1996](https://www.camara.leg.br/proposicoesWeb/prop_mostrarintegra;jsessionid=F0521145122C98F486D7302AE49288FD.proposicoesWebExterno1?codteor=1133712&filename=Dossie+-PL+1736/1996)>.

palavras para representar com exatidão todas as espécies de objetos, fatos e situações. É urgente descolonizar a língua. (BOLSONARO, 1996, p. 2, grifos nossos).

O projeto de Bolsonaro foi felizmente rejeitado pela Comissão de Justiça e de Redação, em 15 de março de 2000, e arquivado, em 12 de setembro do mesmo ano. Contudo, vez ou outra a tentativa de eliminar vocábulos estrangeiros da língua portuguesa vira pauta nas câmaras e assembleias do país. No Rio Grande do Sul, por exemplo, o deputado estadual Raul Carrion propôs, por meio do projeto de lei nº 156/2009, a proibição do uso de estrangeirismos, obrigando a tradução dos vocábulos. A proposta, que previa “a obrigatoriedade da tradução de expressões ou palavras estrangeiras para língua portuguesa, sempre que houver em nosso idioma palavra ou expressão equivalente no âmbito do Estado do Rio Grande do Sul” (CARRION, 2009, p.1), foi aprovada pela Assembleia Legislativa gaúcha em 19 de abril de 2011, entretanto, Tarso Genro, governador à época, aprovou apenas o trecho do projeto que diz que órgãos públicos devem priorizar a língua portuguesa em documentos oficiais, sítios virtuais e materiais de propaganda. Tais propostas excessivas remetem a regimes autoritários, conforme Zilles (2001).

Todas essas tentativas são incongruentes e redundantes, haja vista que a Constituição garante a livre expressão a todos, incluindo os rituais religiosos em línguas africanas e as missas em latim (ZILLES, 2001) e também declara que o português é a língua oficial do país, de uso obrigatório no ensino e na comunicação oficial<sup>16</sup>. Nesse sentido, ao contrário do que muitos gramáticos, escritores e políticos imaginam, “[...] o léxico é um dos pontos em que mais claramente se percebe a intimidade das relações entre língua e cultura” (FARACO, 2005, p. 42). Portanto, “[...] a importação ininterrupta de palavras estrangeiras ao longo da história não modifica a estrutura gramatical da língua.” (FARACO; TEZZA, 1992, p. 35). Em resumo, segundo os autores,

As línguas são abertas no seu vocabulário, enriquecendo-o continuamente, e bastante fechadas em sua organização gramatical (o contato entre as línguas afeta essa organização só muito raramente). É claro que a organização gramatical também se modifica ao longo do tempo, mas antes por uma lógica interna da própria língua que por pressão de sistemas gramaticais estrangeiros. (FARACO; TEZZA, 1992, p. 35)

Nos dias atuais, muitas palavras emprestadas não são mais percebidas como estrangeiras e outras estão sendo inseridas aos poucos. Leitão (2006) localizou 1642 unidades lexicais

---

<sup>16</sup> A LIBRAS foi reconhecida como meio legal de comunicação e expressão a partir da Lei nº 10.436, de 24 de abril de 2002. Além da Libras, o Talian, Asurini do Trocará e Guarani Mbya também foram reconhecidas, em 2010, como as primeiras línguas de Referência Cultural Brasileira pelo IPHAN, fazendo parte do INDL, conforme dispõe o Decreto nº 7.387, de 9 de dezembro de 2010.

oriundas de outras línguas e recentemente incorporadas nas versões digitais de ambos os dicionários Aurélio (FERREIRA, 1999) e Houaiss (2001), dentre as quais 642 palavras e expressões provenientes da língua inglesa foram identificadas no Aurélio e 1072 foram registradas no Houaiss. Subtraindo-se a diferença entre os dois dicionários, temos 663 anglicismos a mais no segundo do que no primeiro. Dá-se o nome de anglicismo às palavras e às construções gramaticais (esse-ó-esse - *SOS* - “Save Our Souls” - “salvem nossas almas”; *Pecê - PC* - “personal computer” - “computador pessoal” etc. (ILARI, 2002, p. 23)) que o português recebeu do inglês. (ILARI, 2002). Em outras palavras, de acordo com o dicionário Aurélio (FERREIRA, 2004, p. 139), anglicismos são palavras ou locuções inglesas introduzidas em outras línguas e empregadas como se fossem destas.

O *Twitter* foi escolhido como *corpus* porque, além de ser um interessante banco de dados disponível na Internet, torna possível a pesquisa remota em meio à pandemia da Covid-19<sup>17</sup>, que fez com que muitas atividades profissionais e acadêmicas passassem para o formato virtual. Além da facilidade de acesso aos dados para pesquisa, os gêneros digitais têm, nos últimos anos, conquistado cada vez mais espaço nas mais diversas esferas sociais.

No caso do *Twitter*, é inegável a importância – inclusive política - que a rede social galgou nos últimos anos. O *Twitter* deixou de ser apenas uma rede social de interação, para ser uma plataforma de influências, pois marcas, políticos e celebridades utilizam o *site* com o intuito de interagir e influenciar seus interlocutores. (FREITAS; BARTH, 2015, p. 9)

Crystal (2005) afirma que a Internet exerce grande impacto sobre a linguagem principalmente porque a sociedade atual está cada vez mais conectada e mais presente nas redes sociais. Uma pesquisa realizada pela *Global Web Index* revelou que o Brasil é 2º em *ranking* de países que passam mais tempo em redes sociais, em média 225 minutos, enquanto que, globalmente, as pessoas gastaram quase 150 minutos por dia nas mídias sociais nos primeiros três meses de 2019 (BBC News Brasil, 2019). Por essa razão, o mundo virtual pode se constituir como um excelente espaço para análise da variação e propagação da mudança de uma determinada língua.

A partir de publicações de usuários na rede social *Twitter*, neste trabalho, sob a ótica da Sociolinguística Variacionista (WEINREICH; LABOV; HERZOG, 2006 [1968]; LABOV, 2008

---

<sup>17</sup> Segundo o Ministério da Saúde brasileiro, a Covid-19 é uma doença causada pelo coronavírus, denominado SARS-CoV-2, que apresenta um espectro clínico variando de infecções assintomáticas a quadros graves. A pandemia do novo coronavírus foi declarada no dia 11 de março de 2020, pelo diretor-geral da Organização Mundial da Saúde (OMS), Tedros Adhanom.

[1972]), descrevemos e analisamos os anglicismos empregados em uma amostra sincrônica da ciberescrita do português, mais especificamente em *tweets* de usuários da rede social *Twitter*.

Nosso recorte espacial é a mesorregião “oeste catarinense”, mais especificamente a microrregião de Chapecó. Esse recorte foi definido com o intuito de contribuir com a descrição do PB escrito, especialmente porque este estudo se insere em um projeto maior em curso denominado “Variação e Mudança no Português do Oeste de Santa Catarina” (ROST SNICHELOTTO, 2012)<sup>18</sup>, aprovado pelo Comitê de Ética em Pesquisa da UFFS (Processo 17011413.2.0000.5564), cujo objetivo é desenvolver pesquisas sobre fenômenos em variação e/ou mudança linguística do PB, em amostras sincrônica e diacrônica, na Mesorregião Grande Fronteira do Mercosul e adjacências.

A região oeste catarinense apresenta uma sócio-história repleta de diversidade étnica e cultural que nos faz entender a presença e a importância dos costumes, das tradições e das línguas para os habitantes da localidade. Segundo Radin, Benet e Milani (2003), os povos naturais - autóctones -, como os Guaranis e Kaingangos, no decorrer do tempo, uniram-se aos povos estrangeiros - alóctones -, como descendentes de imigrantes italianos, alemães e poloneses, que passaram a constituir a região.

O interesse por este estudo começou a partir das aulas dos componentes curriculares Diversidade e Mudança Linguística e Teorias Linguísticas do curso de mestrado. A escolha do recorte temporal para coleta do *corpus* de pesquisa surgiu em meio à pandemia do novo coronavírus (Sars-Cov-2), quando as pessoas começaram a passar mais tempo em frente às telas e, conseqüentemente, aumentaram a quantidade de tempo de acesso às redes sociais. O recorte temporal foi dividido em dois períodos - P1 (03/2019 a 02/2020) e P2 (03/2020 a 02/2021) - para efeito de comparação. A rede social *Twitter* foi escolhida porque acreditamos que os usuários monitoram menos a escrita nos *tweets* publicados. Por fim, do ponto de vista pessoal, a decisão pela coleta de tuítes da região oeste catarinense se deu porque, além de ser o local onde o *campus* Chapecó da UFFS está localizado, também é o local onde eu cresci e que chamo de lar.

Esta dissertação também se justifica em razão da possibilidade de se descrever e analisar o uso de novos vocábulos, advindos da língua inglesa, em uma variante do PB escrito e se insere na linha de pesquisa Diversidade e Mudança Linguística do PPGEL da UFFS, cujo objetivo é o

---

<sup>18</sup> Sob registro nº PES-2018-0879, o VMPOSC é desenvolvido pelo grupo de pesquisa Estudos Sociolinguísticos e Interfaces, certificado pela UFFS e pelo CNPq. Até este momento, já foram desenvolvidos nove trabalhos acadêmicos (um TCC da graduação em Letras Português e Espanhol e oito dissertações de mestrado em Estudos Linguísticos) sobre fenômenos linguísticos em variação e/ou mudança na região oeste catarinense.

desenvolvimento de pesquisas sobre fenômenos em variação ou mudança linguística do PB e das situações de bilinguismo ou de línguas em contato na Mesorregião Grande Fronteira do Mercosul e adjacências. Assim sendo, esta pesquisa também possibilita conhecimentos a respeito de usos cotidianos da língua a estudantes estrangeiros, a fim de que desenvolvam maior proficiência no uso do PB.

Organizamos esta dissertação em seis partes principais, assim definidas:

- **referencial teórico:** primeira parte da dissertação que trata dos pressupostos teórico-metodológicos da Teoria da Variação e da Mudança Linguística (WEINREICH; LABOV; HERZOG, 2006 [1968]); LABOV, 2008 [1972]). Esta seção está dividida em seis subseções com os seguintes conteúdos: a sociolinguística variacionista, a metodologia de pesquisa variacionista, os níveis linguísticos de variação, a internet no Brasil sob uma perspectiva sociolinguística, a abordagem baseada no uso e, por fim, léxico e lexicalização.

- **revisão de literatura:** segunda parte da dissertação em que, por meio de um levantamento bibliográfico em dicionários, gramáticas descritivas e em pesquisas linguísticas realizadas no PB, se trata da origem das palavras da língua portuguesa, da definição de empréstimos linguísticos, estrangeirismos e neologismos.

- **objetivos, questões e hipóteses:** terceira parte da dissertação que expõe os objetivos geral e específicos da dissertação. Após a apresentação dos objetivos, elenca-se as questões e suas hipóteses.

- **metodologia:** quarta parte da dissertação em que apresentamos os procedimentos metodológicos deste estudo. Buscamos, inicialmente, caracterizar a rede social *Twitter* de onde extraímos os dados que compõem o nosso corpus. Em seguida, descrevemos os procedimentos de coleta e análise dos dados, bem como especificamos os condicionantes linguísticos e extralinguísticos testados e especificamos algumas restrições que se fizeram necessárias.

- **descrição e análise dos dados:** quinta parte da dissertação em que tratamos dos fatores linguísticos e extralinguísticos que condicionam o uso de anglicismos em *tweets* escritos por usuários localizados no oeste catarinense.

- **origem e potencialidade semântico-pragmática de alguns anglicismos presentes em *tweets*:** sexta parte da dissertação em que confirmamos, preliminarmente, a origem, segundo os dicionários Aurélio e Houaiss, e a potencialidade semântico-pragmática de alguns anglicismos encontrados em *tweets* analisados previamente. Também exemplificamos seus usos a partir de *tweets* de usuários da rede social *Twitter* localizados na região oeste

catarinense. Caso o anglicismo não seja dicionarizado, lançamos mão dos dicionários Oxford e Cambridge, da língua inglesa.

- **considerações finais:** nesta última seção, responde-se às questões da subseção 4.2, sintetizam-se os principais resultados qualitativos e quantitativos. Também retomam-se alguns problemas conceituais discutidos no referencial teórico a partir da amostra coletada. Por fim, apontam-se as dificuldades encontradas durante a elaboração da pesquisa e sinalizam-se sugestões de pesquisas futuras a partir deste estudo.

## 2. REFERENCIAL TEÓRICO

Ao nascermos, passamos a aprender de modo inconsciente uma determinada variedade de língua falada pela comunidade na qual o falante está inserido. A variedade pode ainda diferenciar em certos aspectos gramaticais e lexicais de outras variedades de uma mesma língua.

Muitas teorias e disciplinas<sup>19</sup> procuram descrever as variedades que formam uma mesma língua. Nessa perspectiva, este capítulo trata, inicialmente, dos pressupostos teórico-metodológicos da Teoria da Variação e da Mudança Linguística (WEINREICH; LABOV; HERZOG, 2006 [1968]); LABOV, 2008 [1972]). Na sequência, abordaremos em seis subseções os seguintes conteúdos:

- a sociolinguística variacionista, a metodologia de pesquisa variacionista, os níveis linguísticos de variação, a internet no Brasil sob uma perspectiva sociolinguística, a abordagem baseada no uso e, por fim, o léxico e a lexicalização.

Apresentamos, na seção a seguir, alguns dos principais pressupostos da Teoria da Variação e Mudança Linguística.

### 2.1 A Sociolinguística Variacionista

A Sociolinguística Variacionista - ou Teoria da Variação e Mudança Linguística - surge a partir da publicação do ensaio intitulado “Fundamentos empíricos para uma teoria da mudança linguística”, no qual Weinreich, Labov e Herzog (2006 [1968]) expõem um modelo teórico-metodológico para o estudo e a valorização dos fenômenos de variação e mudança

---

<sup>19</sup> Faraco (2005, p. 34) afirma que o estudo das variedades “[...] é realizado por diferentes disciplinas linguísticas: pela dialetologia (para as variedades geográficas), pela sociolinguística (para as variedades sociais e estilísticas), pela linguística histórica (para as variedades no tempo).”

linguística e da heterogeneidade inerente nos usos de uma língua. Os estudos de Labov (1972)<sup>20</sup> contribuíram para o reconhecimento de um modelo teórico-metodológico científico válido e confiável para se estudar sistematicamente a complexidade dos dados de fala e para se pesquisar a mudança linguística. Os resultados demonstraram que a variação linguística é inerente às línguas, regular e, enquanto tal, passível de análise linguística sistemática.

Nessa perspectiva, a língua é concebida como uma forma de comportamento social usada por indivíduos em um contexto também social, em que comunicam suas necessidades, emoções, ideias e visões de mundo uns aos outros. No entendimento de Labov (2008 [1972], p. 13),

[...] a base do conhecimento intersubjetivo na linguística tem de ser encontrada na fala - a língua tal como usada na vida diária por membros da ordem social, este veículo de comunicação com que as pessoas discutem com seus cônjuges, brincam com seus amigos e ludibriam seus inimigos.

O conceito de língua é bastante relevante para a linguística e apresenta vários sentidos que a ciência da linguagem propõe-se a diferenciar cautelosamente. Conforme Crystal (2000 [1985]), a língua em seu nível mais específico pode-se referir ao ato concreto de falar, em uma situação particular, a noção de *parole* ou desempenho saussuriana e chomskiana, respectivamente. O autor afirma ainda que o sistema linguístico que está por trás do uso individual da língua em um determinado tempo e lugar é identificado através do termo “ídioteleto”, seguidamente estendendo-se à análise sincrônica da língua (ou linguagem) de um indivíduo. Sob a perspectiva da sociolinguística, Crystal (2000 [1985], p. 159, grifos do autor) acrescenta:

Uma VARIÉDADE ou NÍVEL particular de fala/escrita também pode ser denominada ‘língua’ (ou ‘linguagem’), como em ‘língua(gem) científica’. Neste sentido, o termo está associado à limitação ESTILÍSTICA ou SOCIOLINGUÍSTICA de expressões como “‘língua de transição’ (cf. PIDGIN), o ensino de ‘línguas para fins específicos’, ou ‘instrumental’ (na LINGUÍSTICA APLICADA), etc.

Crystal (2000 [1985]) explica ainda que, em expressões como “primeira língua” ou “a língua inglesa”, o sentido é o sistema abstrato implícito à fala/escrita usadas coletivamente por uma comunidade, a noção de *langue*, ou o conhecimento deste sistema por um sujeito, a noção de competência. Nesta perspectiva, o termo idioma e língua são sinônimos e, para o autor, este último pode ser visto tanto em termos sincrônicos como diacrônicos.

<sup>20</sup> Os estudos clássicos de Labov são, de 1963, sobre a investigação da centralização da primeira vogal dos ditongos /ay/ e /aw/ em Martha's Vineyard; de 1966, sobre a estratificação do /r/ pós-vocálico nas lojas de departamento na cidade de Nova Iorque e, de 1972, sobre o apagamento da cópula no inglês falado por adolescentes negros do Harlem, em Nova York (LABOV, 2008 [1972]).

Esse sistema linguístico efetivamente usado em situações reais de interação em uma determinada comunidade é o objeto de estudo da Sociolinguística: “o objeto da descrição linguística é a gramática da comunidade de fala: o sistema de comunicação usado na interação social” (LABOV, 1982, p. 18 ). A comunidade de fala envolve a coletividade que usa concretamente a língua em um contexto histórico específico, ou seja, é composta por um grupo de indivíduos que compartilham traços linguísticos, que diferenciam o grupo de outros grupos de falantes.

Do latim *variātīō -ōnis*, a variação é um termo alternativo para variante e indica qualquer sistema de expressão linguística cujo uso seja dependente de variáveis de situação (CRYSTAL, 2000 [1985]). Na Sociolinguística, a variação é o processo pelo qual duas ou mais formas podem ocorrer no mesmo contexto com o mesmo significado e representam o mesmo estado de coisas (LABOV, 2008 [1972], p. 02). Portanto, segundo o autor (2008 [1972], p.313), “as variantes são idênticas em valor de verdade ou referencial, mas se opõem em sua significação social e/ou estilística”. Weinreich; Labov; Herzog (2006, p.105) conceituam variável linguística como “um elemento variável dentro do sistema controlado por uma regra”.

As regras variáveis estão intrinsecamente relacionadas aos fatores linguísticos, sociais e estilísticos, ou seja, a partir do contexto, os falantes alternam o uso de uma ou mais de variantes para referenciar a mesma coisa. Para verificar o vínculo entre as variáveis sociais e a variável linguística, os informantes das pesquisas de Labov (2008 [1972]), na ilha de Martha’s Vineyard e na cidade de Nova York, foram distribuídos segundo as suas características sociais prototípicas, como faixa etária, traços estilísticos, localidades, grupos étnicos, sexo, sotaque estrangeiro ou regional e ocupação.

É importante destacar que as palavras e construções, permeadas pelo fenômeno da variação, são repletas de significado social e, por essa razão, conforme Beline (2002), comunicam muito além de seu significado referencial/representacional que esses vocábulos e construções demandam. Assim, a variação dá indícios ao leitor ou ouvinte sobre a cultura na qual estamos inseridos em seus mais diversos aspectos, como a escolaridade, a idade, o contato com outras línguas, a profissão, o sexo<sup>21</sup> etc.

Portanto, conforme Coseriu (1987), a língua faz parte da constituição do indivíduo e de sua identidade e o identifica como pertencente a uma comunidade historicamente determinada ou como alguém que assume temporariamente a tradição idiomática de uma língua. Pelo poder que

<sup>21</sup> Para conhecimento da discussão sobre os termos sexo e gênero na Sociolinguística, ver Freitag (2015).

exerce na sociedade, a língua pode valorizá-lo, integrá-lo, discriminá-lo ou elevá-lo socialmente. Por meio da língua, acontecimentos sociais, culturais, políticos e históricos são expressos, o que a faz se modificar e se adaptar junto com a sociedade, moldando-se para representá-la. Labov (2008 [1972], p. 21) alerta ainda que

[...] não se pode entender o desenvolvimento de uma mudança linguística sem levar em conta a vida social da comunidade em que ela ocorre. Ou, dizendo de outro modo, as pressões sociais estão operando continuamente sobre a língua, não de algum ponto remoto do passado, mas como uma força social imanamente agindo no presente vivo.

Resumidamente, a sociolinguística estuda a língua em seu contexto real, considerando as relações entre os aspectos sociais e culturais e os sistemas linguísticos para a produção linguística, investigando a regularidade e as formas destes usos. Labov (2008 [1972], p. 19-20, grifos nossos) reforça ainda que a variação pode ser motivada por diversos processos:

Essas variações podem ser induzidas pelos processos de assimilação ou dissimilação, por analogia, **empréstimo**, fusão, contaminação, variação aleatória ou quaisquer outros processos em que o sistema linguístico interaja com as características fisiológicas ou psicológicas do indivíduo. A maioria dessas variações ocorre apenas uma vez e se extinguem tão rapidamente quanto surgem. No entanto, algumas são recorrentes e, numa segunda etapa, podem ser imitadas mais ou menos extensamente, e podem se difundir a ponto de **formas novas entrarem em contraste com as mais antigas** num amplo espectro de usos. Por fim, numa etapa posterior, uma ou outra das duas formas geralmente triunfa, e a regularidade é alcançada.

Labov (1996) investigou as mudanças por que as línguas passam no decorrer de suas histórias. De acordo com o autor, é preciso verificar os fatores históricos e sociais que influenciam as línguas, partindo do presente para o passado. Para isso, Labov (1996) criou um método que possibilita a verificação da variação e mudança linguística em tempo real e tempo aparente. Conforme Paiva e Duarte (2003), tempo aparente é uma metodologia mais rápida para analisar possíveis mudanças linguísticas. De modo geral, trata-se da comparação da fala de pessoas mais jovens com a fala de pessoas mais velhas, em que a diferença entre elas poderá mostrar mudanças. No entanto, Labov (1996) ressalta a necessidade da análise das diferenças dos grupos, o que pode ocasionar a influência a partir de fatores como classe social, escolaridade etc. Já o estudo em tempo real, segundo Romano (2012), se faz a partir da comparação de dois ou mais *corpora* coletados através de diferentes recortes sincrônicos, possibilitando ao pesquisador analisar documentos históricos que mostram a língua falada no passado.

Assim sendo, os falantes de uma certa comunidade reconhecem as características compartilhadas entre si, sendo possível que estes indivíduos possam distinguir outros falantes

que não compartilham as mesmas peculiaridades. Estas distinções correspondem à organização social e se espelham nos subgrupos de falantes e são também encontradas na fala de cada indivíduo, que pode variar de acordo com cada contexto.

A fim de observarmos o comportamento dos anglicismos na ciberescrita do português, tomamos como modelo teórico-metodológico a Teoria da Variação e Mudança Linguística.

## 2.2 A Metodologia da Pesquisa Variacionista

Na maioria das vezes, desenvolver pesquisas científicas não é uma tarefa fácil. O pesquisador encontra muitos percalços ao longo do desdobramento de um estudo. Entretanto, quando há métodos a serem seguidos, à medida que são bem utilizados, contribuem para uma pesquisa científica de qualidade.

Conforme Coelho *et al.* (2019), a noção de comunidade de fala se origina, como consequência, da tradição dessa concepção de Labov como norteadora da constituição de bancos de dados para a realização de pesquisas sociolinguísticas. Salomão (2011, p. 187) enumerou alguns grupos pioneiros que deram início às pesquisas em Sociolinguística Variacionista, desde a década de 1970 no Brasil<sup>22</sup>, cujo propósito é “investigar a linguagem relacionando-a a fatores sociais que distinguem diferentes comunidades de fala para a desconstrução da ideia de homogeneidade linguística.” Oliveira (2016) também destaca os primeiros projetos e grupos de pesquisas brasileiros - com registros a partir do final da década de 1960 - que tinham como objetivo a construção de bancos de dados para análises sociolinguísticas posteriores<sup>23</sup>. Hoje, existem vários grupos de pesquisa em todas as regiões do país.

---

<sup>22</sup> Os grupos pioneiros da sociolinguística variacionista, mencionados pela autora, são: o grupo do projeto Mobra Central, o grupo do projeto da Norma Urbana Oral Culta do Rio de Janeiro (Nurc) e do projeto Censo da Variação Linguística no Estado do Rio de Janeiro (Censo), tendo como coordenadores os professores Miriam Lemle, Celso Cunha e Anthony Naro, respectivamente, que citam ainda outros diversos grupos que em várias universidades brasileiras, seguem os pressupostos teórico-metodológicos da Sociolinguística, como o Programa de Estudos sobre o Uso da Língua (Peul), continuidade do Projeto Censo, o próprio Nurc – na Universidade Federal do Rio de Janeiro (UFRJ); o projeto de Variação Linguística da Região Sul do Brasil (Varsul) – na Universidade Federal de Santa Catarina (UFSC), na Universidade Federal do Rio Grande do Sul (UFRGS), na Universidade Federal do Paraná (UFPR) e na Pontifícia Universidade Católica do Rio Grande do Sul (PUC-RS). (SALOMÃO, 2011, p. 193).

<sup>23</sup> NURC – Projeto de Estudo da Norma Urbana Culta do Brasil (1969), Competências Básicas do Português (1977), PEUL – Programa de Estudos dos Usos da Língua (1980), Confluência Dialetal na Nova Capital Brasileira (1984), VARSUL – Variação Linguística Urbana da Região Sul (1989), VALPB – Variação Linguística na Paraíba (1994), Dialetos Sociais Cearenses (1996), LUAL – A Língua Usada em Alagoas (1997). (OLIVEIRA, 2016, p.483).

Os métodos tradicionais de coleta de dados de estudos linguísticos, como censos, questionários, excertos de peças teatrais e romances, testes psicológicos, relatórios etnográficos de normas comunitárias, por mais que perspicazes e produtivos, muitas vezes não permitem ao pesquisador se aproximar dos dados fundamentais da língua em uso, elementar para a sociolinguística (LABOV, 2008 [1972]). Por isso, Labov (2008 [1972]) concebeu um modelo de entrevista (entrevista sociolinguística) que fornecesse vários exemplos da fala espontânea, da fala emocionalmente carregada, da fala monitorada e do estilo de leitura. Mais recentemente, Labov (2001) desenvolveu outro modelo de análise chamado “árvore da decisão” em que segmenta a entrevista em oito contextos (resposta, língua, *soapbox*, residual, narrativa, grupo infância e tangente)<sup>24</sup> distribuídos pelos eixos de fala cuidada e de fala casual.

Definidos o método e os instrumentos de coleta de dados, Eckert (2012) distinguiu três ondas de investigação na prática das pesquisas sociolinguísticas, mas alerta a autora que não se trata de estágios que se sucedem ao longo da história.

A primeira começou com Labov (1966) e seu estudo de estratificação social do inglês na cidade de Nova Iorque. Segundo a autora, os resultados de Labov foram utilizados em inúmeros outros estudos durante as décadas de 1960 e 1970, não somente na América do Norte, mas também na Grã-Bretanha (WOLFRAM, 1969; TRUDGILL, 1974; MACAULAY, 1977), além dos estudos realizados no Panamá (CEDERGREN, 1973) e Irã (MODARESSI, 1978). De acordo com Eckert (2012, p. 88), estes estudos estabeleceram um padrão da estratificação socioeconômica da forma linguística, utilizando diferenciações regionais e étnicas:

Esses estudos estabeleceram um padrão regular de estratificação socioeconômica da forma linguística, com maior diferenciação regional e étnica na extremidade inferior da hierarquia socioeconômica, bem como maior uso de formas não padronizadas mais difundidas.<sup>25</sup>

A segunda onda, segundo Eckert (2012), aprofunda a atribuição dos fatores sociais envolvidos na variação da língua. O estudo de Labov na ilha de Martha’s Vineyard e o de Milroy (1980), inspirado na obra de Gumperz (1982), são exemplos de análise na perspectiva da segunda onda. Explicitamente contra a visão passiva da primeira onda, Milroy (1980) buscou forças positivas no uso da língua em Belfast pela classe trabalhadora:

<sup>24</sup> “A Árvore de Decisão organiza-se nos eixos da fala casual e da fala cuidada e os critérios são neles dispostos conforme uma escala de objetividade. Visa à segmentação e análise da fala na entrevista sociolinguística, para identificar especialmente os trechos de fala espontânea ou não monitorada produzidos pelo informante.” (BATTISTI *et al.*, 2021, p. 315).

<sup>25</sup> These studies established a regular pattern of socioeconomic stratification of linguistic form, with greater regional and ethnic differentiation at the lower end of the socioeconomic hierarchy as well as greater use of more widespread nonstandard forms (ECKERT, 2012, p. 88). (Tradução nossa).

Ela argumentou que redes densas múltiplas, típicas da classe trabalhadora, teriam um forte poder de imposição de normas locais, e ela procurou correlacionar os tipos de rede dos indivíduos com seu uso de variáveis vernáculas. O estudo mostrou tal correlação entre a variação, a densidade e a multiplicidade das redes sociais da classe trabalhadora feminina e foi seguido por estudos que mostraram uma relação entre o uso de variantes locais e também o envolvimento em redes locais etnicamente definidas (Edwards & Krakow 1985, Edwards 1991, Knack 1991). (ECKERT, 2012, p. 91)<sup>26</sup>.

Por último, a autora destaca que o principal movimento da terceira onda acontece a partir de uma visão da variação como um reflexo da complexidade das identidades sociais e práticas estilísticas dentro de uma comunidade. Eckert (2012) aponta como estudos de terceira onda pesquisas de Bucholtz e Hall (2005), Bucholtz (2010) e Irvine (2001). Camacho (2015) reitera que a terceira onda é mais recente na Sociolinguística e concentra seus esforços na variação como um recurso para a construção de significado social.

Após a coleta dos dados, Salomão (2011) relembra que a Sociolinguística variacionista trabalha com dados estatísticos que são tratados quantitativa e qualitativamente. A autora diz que os estudos sociolinguísticos baseiam-se em situações reais de linguagem com o falante real e que o sociolinguista busca coletar um número expressivo de dados a partir de amostras de fala/escrita de um determinado número de informantes, por meio de entrevista que leve o indivíduo a produzir um discurso espontâneo e informal. No caso desta pesquisa, em que os dados partiram de amostras de ciberescrita, o *Twitter* foi escolhido justamente por suas características de informalidade, bastante recorrente nas redes sociais virtuais.

### 2.3 Os níveis linguísticos de variação

A variação linguística decorre de fatores internos à língua, mas também é motivado pela interferência de fatores externos, de origem social - extralinguísticos. As línguas são essencialmente heterogêneas, visto que através delas damos conta das diversas situações sociais em que diariamente nos envolvemos. Conforme Castilho (2012, p.197, grifos nossos), as variações organizam a diversidade e heterogeneidade das línguas em níveis extralinguísticos:

---

<sup>26</sup> She argued that dense multiplex networks, typical of the working class, would have a strong local norm-enforcing power, and she sought to correlate individuals' network types with their use of vernacular variables. The study showed such a correlation between variation and the density and multiplexity of women's working-class social networks and was followed by studies showing a relation between the use of local variants and engagement in local ethnically defined networks as well (Edwards & Krakow 1985, Edwards 1991, Knack 1991). (ECKERT, 2012, p. 91). (Tradução nossa).

(1) variação\* geográfica, (2) variação sociocultural (3) variação individual, (4) variação de canal e (5) variação temática. Cada uma dessas variações, por sua vez, é organizada por um conjunto de variantes\*, ou seja, um conjunto de usos linguísticos considerados relevantes para a caracterização de uma variedade. Com isso, **entende-se por variação a manifestação concreta da língua**, e por variedade a soma idealizada das variações. Se fôssemos dispor esses conceitos numa hierarquia teríamos

variante > variação > variedade

Resumidamente, segundo Castilho (2019), compreende-se como (1) variação geográfica quando existe uma correlação entre a região de origem dos falantes e as marcas específicas que acompanham a produção linguística desses indivíduos. Por exemplo, brasileiros e portugueses não falam da mesma forma, como é o caso do dispositivo periférico chamado de “*mouse*” no Brasil e de “*rato*” em Portugal. Também cabe como exemplo de variação diatópica quando, numa mesma língua, a mesma palavra pode ser pronunciada de diferentes formas de acordo com o lugar, por exemplo, a forma como o -r é pronunciado no Rio de Janeiro é diferente da pronúncia do interior paulista. Já a (2) variação sociocultural se dá quando há uma ligação entre fatos linguísticos e o segmento social do qual o falante se origina. As variedades socioculturais são sistematizadas a partir de variáveis, por exemplo, (I) falante não escolarizado e (II) falante escolarizado. O controle de informantes não escolarizados e escolarizados é exemplo da variação diastrática - que tem relação com a situação socioeconômica e cultural do falante. Ainda conforme Castilho (2019), geralmente o falante menos escolarizado tende a não pronunciar o -r final dos verbos que marca o infinitivo, pronunciando algo como “comê” ao invés de “comer”, ao contrário dos falantes mais escolarizados, que tendem a pronunciar-lo. A (3) variação individual pode ser compreendida a partir de fatores como o registro, a idade e o sexo. Exemplo dessa variação é a forma como os adolescentes falam com mais gírias se comparados aos adultos. Um caso bastante atual é a gíria “*cringe*”, utilizada por adolescentes para referir-se a “algo vergonhoso”, termo até pouco tempo desconhecido pelos falantes mais velhos. No que se refere à (4) variação de canal, Castilho (2019) a conceitua como a comunicação linguística que ocorre em presença do interlocutor, quando falamos, ou na ausência, quando escrevemos, em outras palavras, a língua falada e a língua escrita. Por exemplo, podemos falar “Cê tá bem?”, mas escrevemos “Você está bem?”. Por fim, a (5) variação temática é caracterizada pelo modo como é tratado um assunto que está sendo desenvolvido, ou seja, essa variação distingue a linguagem do cidadão comum da linguagem de cientistas, políticos, clérigos e outros. Um exemplo desta variação é que o cidadão comum pode falar sobre um “pedido de liberdade” enquanto na linguagem jurídica o termo é “*habeas corpus*”.

Além da dimensão externa, a variação linguística pode ser investigada sob o ponto de vista da dimensão interna, isto é, a variação vista de dentro da língua. Conforme Coelho *et al.* (2019), a variação interna decorre dos seguintes níveis linguísticos: (I) variação lexical, que se dá a partir das palavras e suas variações em uma determinada língua (exemplos: abóbora e jerimum são vocábulos que nomeiam o mesmo alimento em diferentes regiões do Brasil); (II) variação fonológica, esta que ocorre quando existem diferentes pronúncias de um mesmo vocábulo (exemplos: o rotacismo é um caso de variante fonológica onde o falante troca a consoante -l pela consoante -r, como pode acontecer em “Framengo” no lugar de “Flamengo”); (III) variação morfofonológica, morfológica e morfossintática, estas ocorrem quando um morfema de um vocábulo é alterado (exemplos: “cantano” por “cantando”, eles “anda” por eles “andam”, uso dos pronomes “tu” e “você” para referência à segunda pessoa do singular etc.); (IV) variação sintática, que ocorre quando há diferentes construções para uma mesma oração (exemplos: “Eu o vi no mercado”/“Eu vi-o no mercado”; “Olha, eu não vou sair agora”/“Olha, eu não vou sair agora não”.); e (V) variação discursiva, esta que só é percebida dentro do discurso, pois os vocábulos isolados podem sugerir outros significados. Geralmente esta variação é sinalizada por marcadores discursivos, expressões de natureza adverbial e outros que são usados tanto na fala, quanto na escrita (exemplos: “quer dizer”, “digamos assim”, “aí”, “daí”, “então” etc.). Interessa-nos, em especial, o aprofundamento da variação lexical por tratarmos do uso de anglicismos na ciberescrita do PB.

### 2.3.1 A variação lexical

A variação lexical muitas vezes apresenta características mais explícitas que se tornam mais perceptíveis pelos falantes de uma língua, levando em consideração que ela se dá quando dois ou mais vocábulos se referem ao mesmo objeto, a exemplo de “aipim”, “mandioca”, ou ainda “macaxeira”, que nomeiam o mesmo alimento. Conforme Coelho *et al.* (2019), quando se aborda o tema da variação linguística, principalmente no meio escolar e acadêmico, os exemplos mais recorrentes dizem respeito ao nível do léxico, em outros termos, as palavras que compõem uma língua específica e, de modo geral, relacionadas à variação regional. Alguns exemplos dessa variação podem ser “abóbora”, “jerimum”; “pandorga”, “pipa”, “papagaio”. Percebe-se que, em todos os exemplos de variação lexical, há sempre duas ou mais variantes para designar a mesma variável. Vejamos outro exemplo, extraído de Görski e Rost (2008, p. 47-48):

[...] a fruta cítrica de cor alaranjada e sabor adocicado é conhecida em grande parte da região Sul como “bergamota” ou “vergamota”. Em São Paulo, é chamada de “tangerina”, no Rio de Janeiro de “laranja-cravo”, em Minas Gerais de “mexerica” e também em outros locais é conhecida como “laranja-mimososa” ou simplesmente “mimososa”.

Sobre os aspectos lexicais, fonético-fonológicos, morfológicos e sintáticos, Coelho *et al.* (2019, p. 25) afirmam que:

[...] na classificação dos dialetos em geral, os aspectos lexicais são menos sistematizáveis do que os fonético-fonológicos, morfológicos ou sintáticos, visto que esses últimos são regulados por condicionadores internos, além dos externos, enquanto os lexicais estão intimamente ligados a fatores extralinguísticos, de caráter cultural, sobretudo etnográficos e históricos.

Bagno (2012) reforça ainda que há um léxico característico, variável de região para região do Brasil, a exemplo de: “fruita” para “fruta”, “luita” para “luta”, “cunzinha” para “cozinha”, “dereito” para “direito” etc. Conforme o autor, a grande maioria destes vocábulos é resquício de fases anteriores da língua, podendo ser encontrados na literatura medieval e clássica.

Ainda segundo Coelho *et al.* (2019), as maiores contribuições para o estudo da variação lexical têm sido oferecidas pelos estudos geolinguísticos de diferentes regiões do Brasil. Esses estudos, desde os primeiros trabalhos da área, do filólogo Antenor Nascentes, na década de 1950, têm como principal finalidade a elaboração de um Atlas Linguístico do Brasil<sup>27</sup>.

## 2.4 A mudança sob a perspectiva da Sociolinguística variacionista

Castilho (2019) afirma que as línguas são também irremediavelmente voltadas para a mudança porque os seres humanos são dinâmicos e, por isso, as línguas necessitam adaptar-se às novas situações impostas.

Conforme Salomão (2011), uma das principais diferenças entre a sociolinguística e a linguística estruturalista é o objeto. A linguística estruturalista estuda apenas a substância material das palavras, não seu significado, enquanto a sociolinguística tem como objeto a fala, sendo fundamental a compreensão da variação e das mudanças linguísticas. Ainda sob esta perspectiva, Salomão (2011, p. 191) afirma que:

[...] a análise sociolinguística enfoca fundamentalmente o processo de interação fala/sociedade, justificando-se pela necessidade de compreender os fatores que

<sup>27</sup> O Projeto Atlas Linguístico do Brasil (Projeto ALiB) tem por meta a realização de um atlas geral do Brasil no que diz respeito à língua portuguesa. Disponível em: <<https://alib.ufba.br/>> Acesso em: 20 ago. 2021.

possam influenciar a operação de uma ou de outra variante, na busca de estabelecer uma sistematização ao processo de variação linguística.

Para Camacho (2015), a sociolinguística representou um dos principais esforços para romper com a dicotomia da linguística estruturalista ao considerar a heterogeneidade como princípio constitutivo da linguagem. Ainda de acordo com o autor, a Sociolinguística se estrutura a partir de fenômenos linguísticos coletados na comunidade, base para a análise concreta, diluindo dicotomias como língua/fala e competência/desempenho.

Guy e Zilles (2006, p. 42) afirmam que um dos aspectos mais importantes e relevantes das línguas humanas é a diversidade linguística:

Este é um ponto básico nas pesquisas e teorias sociolinguísticas e, em princípio, não precisamos de nenhuma pesquisa acadêmica formal para reparar na existência desta diversidade. Ela é evidente pela experiência de todo mundo; entretanto, em muitas sociedades, como é o caso da sociedade brasileira, a representação sociocultural da língua de certo modo oblitera essa percepção, fazendo crer que a língua de verdade não varia – ou, numa exacerbação idealizada, faz crer que a língua não deveria variar.

Conforme Weinreich; Labov; Herzog (1968) uma pesquisa empírica sobre as mudanças linguísticas em andamento pode ser subdividida em cinco problematizações distintas para as quais uma teoria da mudança deve fornecer respostas:

1. O problema da *transição* é encontrar o caminho pelo qual um estágio de uma mudança linguística evoluiu a partir de um estágio anterior. [...] Assim, são os aspectos do problema da transição, questões sobre a regularidade da mudança sonora, sobre a influência gramatical na mudança sonora sobre “cadeias que avançam” *versus* “cadeias que retrocedem”, sobre movimento constante *versus* alterações súbitas e descontínuas.
2. O problema do *encaixamento* é encontrar a matriz contínua de comportamento social e linguístico em que a mudança linguística é levada a cabo. O principal caminho para a solução está na descoberta das correlações entre esses elementos e o sistema não linguístico de comportamento social. As correlações se estabelecem por provas sólidas de variação concomitante, ou seja: mostrando-se que uma pequena mudança na variável independente é regularmente acompanhada por uma mudança da variável linguística numa direção previsível.
3. O problema da *avaliação* é encontrar os correlatos subjetivos (ou latentes) das mudanças objetivas (ou manifestadas) que foram observadas. A abordagem indireta deste problema correlaciona as atitudes e aspirações gerais dos informantes com seu comportamento linguístico. A abordagem mais direta é medir as reações subjetivas inconscientes dos informantes aos valores da própria variável linguística. (LABOV, 2008 [1972] p. 193).

Além dos três problemas acima, Weinreich, Labov e Herzog (1968) acrescentam outros dois, o dos fatores condicionantes (mudanças e condicionantes possíveis) e o da implementação (razões para mudanças ocorrerem em certa língua numa dada época).

A partir de nosso fenômeno de estudo, pode-se exemplificar os cinco problemas definidos por Labov (2008[1972]). No caso do problema de transição destacamos a palavra “*whatsapp*” que coexiste com “*whats*”, “*wpp*”, “*zap*” e outras, ambas referindo-se ao mesmo aplicativo, sendo a primeira nomenclatura original. Já quanto ao problema de encaixamento, destacamos a correlação entre o uso de anglicismos em *tweets* e o aumento do uso da rede social durante a pandemia. Levando em conta que o problema de avaliação se dá a partir do significado social da mudança, se é de prestígio ou não, conforme Salomão (2011), podemos exemplificá-lo também através do vocábulo “*whatsapp*” e suas variações “*zap*” ou “*zape*”. Em nossa pesquisa foram analisados os condicionantes faixa etária, sexo e ocupação. Nosso fenômeno se dá a partir das redes sociais virtuais e outras tecnologias, por isso, no que tange ao problema de implementação, podemos exemplificar o vocábulo “*internet*”, que já está dicionarizado e a palavra “*live*”, forma nova, que vem se destacando a partir da declaração da pandemia de Covid-19 e, por este motivo, não se trata de uma mudança completa, pois ainda está em curso.

Isto posto, traremos, na sequência, da internet no Brasil sob uma perspectiva sociolinguística, que é a base para a nossa coleta e análise de dados, além da construção deste trabalho como um todo.

## 2.5 O ciberespaço e a ciberescrita

Debord (1997) afirma que a cultura é a esfera geral do conhecimento e das representações da sociedade e a cibercultura é uma das partes deste todo que propicia uma comunicação sem fronteiras. Dentro do ciberespaço, a aceitação das redes sociais é um fenômeno global, capaz de formar uma ponte entre as culturas, um emaranhado de informações e um elo entre os indivíduos. As rápidas inovações tecnológicas caminham lado a lado com a ampliação lexical das línguas vivas. Lévy (1999, p. 92) define o ciberespaço como “o espaço de comunicação aberto pela interconexão mundial dos computadores e das memórias dos computadores”.

Para Ferreira (2009), já não existem mais controvérsias quanto à comunicação e à informação constituírem uma das bases fundamentais de todos os contatos sociais. Conforme o autor, desde o tempo das inscrições rupestres pré-históricas até o das redes computacionais virtuais, os indivíduos comunicam suas necessidades e anseios, formando, deste modo, uma rede relacional que sustenta os processos sociais. Neste sentido, percebe-se que a partir do século XXI a troca de informações passou a ser imediata e coletiva, com constante mobilização de ideias.

Conforme Briggs e Burke (2006, p. 303), foi entre setembro de 1993 e março de 1994 que uma rede, até então exclusiva de pesquisas acadêmicas, tornou-se a rede das redes aberta a todos:

Um dos principais usos da internet [...] era o envio de mensagens de correio eletrônico em linguagem “real”, sendo a maioria delas de pessoa para pessoa. Esse uso foi objeto de um número da *New Yorker*, “A idade digital”, em dezembro de 1999, que também continha um artigo intitulado “Carros inteligentes, tecnologia em movimento”. Havia imaginação, mas não fantasia, na abordagem adotada pelo artigo; não exatamente porque o escritor descrevia o correio eletrônico como o “retorno da palavra” depois de uma longa idade visual, mas porque o e-mail reacionário não voltava no tempo.

Lévy (1999) destaca o *e-mail* como uma das trocas de mensagens mais importantes do ciberespaço. Para Lévy (1999, p. 95), o correio eletrônico despertou grande interesse pela agilidade e facilidade da troca, envio e correção de mensagens sem a necessidade do papel:

De forma simétrica, não é mais necessário imprimir o texto para fazer com que chegue ao seu destinatário: pode ser enviado diretamente em sua forma digital inicial. [...] O correio eletrônico permite enviar, de uma só vez, uma mesma mensagem a uma lista (que pode ser longa) de correspondentes, bastando indicar essa lista. Assim, não é necessário fazer fotocópias do documento, nem digitar diversos números telefônicos, um após o outro.

Dando um salto, um aspecto marcante do início da década de 2000 é o surgimento das redes sociais, que possibilitam um grande fluxo de discurso social. Para uma utilização eficiente das redes, o usuário disponibiliza tempo, motivação e envolvimento nas diversas práticas sociais das suas áreas de interesse. A partir deste afunilamento de interesses, o ambiente virtual das redes sociais permite o estabelecimento de comunicação e, conseqüentemente, a disseminação de novos itens lexicais. Recuero (2009) afirma que o advento da CMC<sup>28</sup> amplificou a capacidade de conexão, tornando possível que redes fossem criadas e expressas no ciberespaço. Conforme a autora, essas redes sociais ganharam maior evidência em 2008, quando lançaram a campanha de Barack Obama para a presidência dos EUA e, no Brasil, com a catástrofe das fortes chuvas e deslizamentos em Santa Catarina, também em 2008. Foi durante esses dois eventos que “uma série de blogs, ferramentas de mensagens como o Twitter, mensageiros instantâneos e outros recursos foram utilizados para informar o resto do país a respeito dos acontecimentos.” (RECUERO, 2009, p. 16).

Para Rajagopalan (2013), meios como o *Facebook*, o *Instagram*, o *Whatsapp*, o *Twitter* e outros, fazem com que milhões, ou até bilhões, de pessoas se comuniquem todos os dias, a toda hora, a cada minuto, transitando pelas diversas formas de linguagem: da formal a informal, da

---

<sup>28</sup> Comunicação mediada por computador.

comum à especializada, das siglas ao léxico “facebookiano”, ao “internetês” padrão e outras tantas formas de iniciar uma comunicação. Os usuários das redes sociais precisam estar sempre muito atentos ao meio virtual, pois diariamente surgem novos “memes”, novas abreviações, novos estrangeirismos e, se a pessoa não estiver atualizada, a conversa poderá se tornar simplesmente incompreensível.

Conforme Carvalho (2018), “as redes sociais constituem um método de aprendizagem moderno e envolvente”. Já Germano e Taveira (2015) mostram que “os novos significados e significantes criados pelos usuários da rede social Facebook refletem não só o que está sendo tratado na mensagem ou publicação, mas também aspectos de sua cultura e identidade”. Neste mesmo sentido, Recuero (2009) define uma rede social como um conjunto de dois elementos: os atores, que são as pessoas, instituições ou grupos; e suas conexões, que são as interações ou laços sociais. De forma mais detalhada, a autora define o ator do ciberespaço:

Quando se trabalha com redes sociais na Internet, no entanto, os atores são constituídos de maneira um pouco diferenciada. Por causa do distanciamento entre os envolvidos na interação social, principal característica da comunicação mediada por computador, os atores não são imediatamente discerníveis. Assim, neste caso, trabalha-se com representações dos atores sociais, ou com construções identitárias do ciberespaço. Um ator, assim, pode ser representado por um *webblog*, por um *fotolog*, por um *twitter* ou mesmo por um perfil no Orkut. (RECUERO, 2006, p. 25)

Em 2004, no mesmo ano de criação e lançamento do Facebook, Marcuschi participou de uma conferência na Universidade de São Paulo (USP) em que dizia que “a internet é uma espécie de protótipo de novas formas de comportamento comunicativo; se bem aproveitada, ela pode tornar-se um meio eficaz de lidar com as práticas pluralistas sem sufocá-las” (MARCUSCHI, 2004, não paginado). Vale ressaltar ainda que, de acordo com Komesu (2009), uma estimativa divulgada pela imprensa, em agosto de 2002, apontava um número de 170.000 (cento e setenta mil) escreventes de *blogs*, levando em conta apenas os hospedados em *sites* brasileiros.

O último advento tecnológico que vem impactando a sociedade é a internet. Diante disso, Saliés e Sheperd (2013) afirmam que mais de mil línguas estão representadas no mundo virtual. As autoras expõem os resultados da análise estatística publicada pelo *site Internet World Stats* que mostra as dez línguas mais usadas até 2011: o inglês, o chinês, o espanhol, o japonês, o português, o alemão, o árabe, o francês, o russo e o coreano, nessa ordem, somando aproximadamente sete milhões de internautas dedicados cultural e socialmente na construção de sentido. Saliés e Shepherd (2013, p. 7) comentam ainda:

Na internet há também uma crescente expansão de mídias de toda natureza: as mais colaborativas, como as wikis, as mais dedicadas ao compartilhamento, como o Youtube, e as mais interacionais, como os Blogs, Twitter, Tumblr e Facebook. À medida que a internet expande essa infra-estrutura, atende a necessidades de comunicação, criando novas formas de colaboração, compartilhamento e interação.

Conforme Galli (2010), grande parte dos avanços tecnológicos está no processo evolutivo da comunicação, dirigindo-se para uma maior popularização da informação e, por conseguinte, do saber. Segundo a autora, a comunicação virtual introduz um conceito de descentralização da informação e do poder de comunicar. Todo computador e também os dispositivos móveis (*smartphones*, *tablets* e *smartwatches*), se conectados à internet, possuem a capacidade de transmitir palavras, imagens e sons. Esta nova mídia permite que qualquer pessoa possa construir um *site* na internet, sobre qualquer assunto:

O espaço cibernético tem se tornado um lugar essencial, um futuro próximo de comunicação completamente distinta da mídia clássica, pois como afirma Lévy (2000, p.13), “[...] todas as mensagens se tornam interativas, ganham uma plasticidade e têm uma possibilidade de metamorfose imediata”. (GALLI, 2010, p. 152).

Sob essa perspectiva, nota-se o potencial do ciberespaço como fonte de pesquisas linguísticas, principalmente considerando a ciberescrita. Há outros trabalhos sociolinguísticos em PB sobre o fenômeno da internet e/ou as redes sociais, que aqui destacamos: Freitag e Silva (2006), Soares e Peixoto (2010), Sampaio (2011), Roloff *et al.* (2015), Nascimento (2017), Borges (2016), Leite (2019), Salvador e Barros (2020).

À vista disso, tanto os empréstimos a partir do fenômeno das redes sociais, quanto quaisquer outros fenômenos linguísticos, precisam ser analisados e compreendidos perante a situação discursiva em que se encontram entrelaçados aos contextos linguísticos e extralinguísticos.

## 2.6 Léxico e lexicalização

As línguas têm formas de criar novas palavras a partir de recursos já existentes. Bybee (2020) descreve como as línguas adquirem novas palavras. De modo abrangente, léxico é sinônimo de vocabulário. Segundo Crystal (2000 [1985]), o léxico destaca-se principalmente dentro da gramática gerativa, na qual ele se refere a um compilado de todas as informações sobre as propriedades estruturais dos itens lexicais de uma determinada língua. Para Carvalho (1989), o

léxico de uma língua se compara a uma galáxia por viver em expansão permanente por incorporar as experiências pessoais e sociais da comunidade por meio de fatores linguísticos e extralinguísticos.

O termo lexical também pode apontar uma diferenciação da gramática, conforme Crystal (2000 [1985]). O autor expõe duas possibilidades para categorizar os vocábulos, as “palavras gramaticais” e as “palavras lexicais”: a primeira atribui-se às palavras que têm como única função pontuar as relações gramaticais como os vocábulos *de, e, para, a, como* e outros. Já as “palavras lexicais” são aquelas que possuem significação lexical, segundo Crystal (2000 [1985]), são os vocábulos que possuem conteúdo semântico. Bagno (2012) apresenta a clássica divisão dos estudos linguísticos entre léxico e gramática, que, segundo o autor, mostra que uma língua é composta por dois conjuntos: as palavras e as regras que determinam o funcionamento dessas palavras. Por outro lado, Bagno (2012) defende que, apesar da dualidade proposta, léxico e gramática não podem andar separados tendo em vista que, de acordo com o autor, há entre eles, uma relação intrínseca que torna impossível separá-los.

Bagno (2012) propõe a separação entre palavras e regras. No grupo das palavras situam-se o léxico e a morfologia e o grupo das regras é composto pela gramática e pela sintaxe. Dessa forma, o autor expõe os conceitos de “classes gramaticais abertas” e “classes gramaticais fechadas”, cujas primeiras seriam aquelas compostas por um repertório que se amplia diariamente com novas contribuições dos falantes, como os estrangeirismos, empréstimos, neologismos e outros. Por outro lado, as segundas, as fechadas seriam aquelas constituídas por um número finito de membros, ou seja, os pronomes, preposições, artigos etc. Ainda assim, Bagno (2012) indica que essa proposta de divisão de classes não se aplica a todas as línguas.

Castilho (2019) define a lexicalização como o processo de criação das palavras em que expressamos categorias e subcategorias cognitivas e seus traços semânticos, transformando impulsos mentais em ondas sonoras, em um mecanismo ainda obscuro. A lexicalização, ainda conforme Castilho (2019), percorre três caminhos: o da etimologia, quando um item da língua-fonte é integrado na língua-filha, a exemplo de palavra portuguesa *pêssego*, que tem origem na expressão latina *fructu persicu* (fruto da Pérsia); o caminho da neologia, que ocorre quando criamos uma palavra nova, não advinda da língua-fonte e organizada de acordo com as regras morfológicas da língua-alvo, a exemplo do verbo-ônibus<sup>29</sup> *coisar*, calcado no substantivo *coisa*; e, por fim, o caminho do empréstimo, que ocorre através do contato linguístico. A

---

<sup>29</sup> Verbo que se emprega como sinônimo de quase todos os outros.

lexicalização por empréstimo é especificamente o que nos interessa e será tratada na subseção seguinte.

### 2.6.1 Lexicalização por empréstimo

Bybee (2020) afirma que todas as línguas encontram formas de criar novos vocábulos a partir de recursos já existentes. Conforme a autora, à medida que novos conceitos e dispositivos vão surgindo em uma determinada cultura, os falantes de uma determinada língua buscam meios de criar novas palavras para que esses dispositivos e conceitos se insiram naquela determinada comunidade.

Entre os muitos fenômenos linguísticos que podem incorporar novos vocábulos a uma determinada língua, temos o empréstimo. Segundo Bagno (2012), qualquer língua viva no mundo recebe e receberá contribuições lexicais de diferentes línguas, os chamados empréstimos.

Do ponto de vista extralinguístico, Bagno (2012) afirma que fenômenos históricos, culturais, políticos, ideológicos e outros se combinam para que vocábulos de outras línguas, geralmente da língua dominante no contexto global, se instalem no léxico de outras línguas. Conforme o autor, é preciso analisar a questão dos estrangeirismos com imparcialidade, para que se possa constituir uma ideia criteriosa acerca do fenômeno linguístico. Atualmente, conforme Bagno (2012, p. 267, grifos do autor), as palavras estrangeiras derivam principalmente da língua inglesa e não afetam, conforme creem os leigos, a gramática da língua:

[...] os empréstimos são contribuições ao léxico da língua e não ao modo como a língua se estrutura morfossintaticamente nem fonologicamente. Ao contrário, quando um termo estrangeiro passa a ser empregado, ele é imediatamente adaptado aos hábitos fonéticos da língua de chegada, para mais adiante, se permanecer em uso, receber uma grafia coerente com o modo convencional de escrever o idioma.

Foi assim que o francês *abat-jour* se transformou no abajur, e que o *billet*, o *bonnet*, o *maillot* e o *haut-bois* se transformaram em *bilhete*, *boné*, *maiô* e *oboé*. E é assim que muitos termos provenientes do inglês estão aos poucos se fonologizando-lexicalizando-morfossintaticizando no PB.

Bagno (2012) destaca ainda que qualquer enunciado feito em PB com palavras de origem estrangeira vai se constituir de acordo com as regras morfossintáticas da língua, ou seja, a ordem constituinte do enunciado, a concordância de gênero e número, as preposições e a flexão verbal serão em PB, apenas a palavra em si, como *notebook*, revelará a origem estrangeira, dificultando ao falante nativo compreender que se trata de um vocábulo emprestado.

Também Bybee (2020) destaca que um verdadeiro empréstimo - em oposição ao uso esporádico de uma palavra estrangeira - se integra na fonologia, morfologia e sintaxe da língua. A autora afirma que a adaptação fonológica geralmente substitui fonemas não nativos por nativos, trocando padrões fonotáticos<sup>30</sup>, além de regular padrões acentuais. Bybee (2020) aponta ainda para o fato de que, às vezes, os empréstimos impactam a língua receptora para além da expansão do léxico. Conforme a autora, podem entrar na língua receptora, através dos empréstimos, novos fonemas, novos padrões silábicos, novas distribuições de fonemas e até alguma morfologia derivacional. Entretanto, para que tais fenômenos ocorram, Bybee (2020) afirma que eles exigem um elevado grau de bilinguismo durante um grande período na comunidade.

Segundo Bybee (2020), o contato linguístico é uma das formas de expansão do léxico de uma língua. Os falantes atuais e, presumidamente, seus antepassados experienciam mais de uma língua em suas comunidades. De acordo com a autora, quando culturas entram em contato, ocorre a troca de objetos, alimentação e outros aspectos da sociedade que são acompanhados pelas palavras que nomeiam estes elementos.

Dos vários aspectos que constituem uma sociedade, a cultura alimentícia é uma das partes que escancara o contato linguístico. Conforme Bybee (2020), palavras comuns ao PB como *fondue*, *sushi* e *taco* têm a origem bastante clara para os falantes, porém o português vem importando palavras para denominar alimentos há muitos séculos. Segundo a autora, palavras como *laranja*, *batata* e *tomate* são tratadas como nativas, mas também são frutos de contato linguístico:

O primeiro registro de *laranja* em português é de 1377, e a palavra nos chegou do árabe *nārandja*, este do persa *nārang*, que por sua vez vem do sânscrito *nāraṅga*. A *batata* e o *tomate* são plantas originárias do continente americano: o termo *batata* entrou na língua no século XVI, pelo espanhol *patata*, atestado a partir de 1560, de uma língua do caribe do Haiti, para designar batata-doce. Por volta de 1590, o nome se estendeu para a batata branca comum, originária do Peru e que, por ironia da história, também é conhecida entre nós como *batata-inglesa*. (BYBEE, 2020, p. 329).

Os vocábulos acima mencionados são empréstimos de outras línguas completamente integrados ao PB. A autora reafirma ainda que os nomes que designam os alimentos viajam com os próprios alimentos, sendo fácil de perceber esse movimento. Ainda neste sentido, Bybee (2020) diz que o mesmo acontece com outros artefatos em que o progresso é cursado de modo

---

<sup>30</sup> Termo usado na FONOLOGIA para indicar as COMBINAÇÕES (ou “comportamento tático”) específicas de sons e FONEMAS de uma língua. (CRYSTAL, 2000 [1985]. p. 117).

semelhante, este pode ser o caso dos dispositivos e atividades que envolvem o mundo virtual e fazem parte do *corpus* desta pesquisa.

### 3. REVISÃO DE LITERATURA

Esta seção trata, primeiramente, da origem das palavras da língua portuguesa. Na sequência, diferenciamos, por meio de um levantamento bibliográfico em dicionários, gramáticas prescritivas e descritivas e em pesquisas linguísticas realizadas no PB, a definição de empréstimos linguísticos, estrangeirismos e neologismos.

#### 3.1 O composto do léxico da Língua Portuguesa

A evolução das línguas anda lado a lado dos fenômenos extralinguísticos, como a evolução histórica e cultural. Conforme Azeredo (2010), a língua portuguesa - que começou a ser escrita entre o final do século XII ou início do século XIII - possuía em seus primórdios um léxico que reunia cerca de 80% de palavras de origem latina e os outros 20% de palavras pré-romanas, germânicas e árabes. Esta origem, de acordo com Manzolillo (2014), se refere ao acervo lexical hereditário, isto é, o vocabulário que surgiu com a língua, que a ela forneceu o padrão fonético e morfológico.

A partir do momento em que a língua portuguesa passou a ser escrita, inúmeros fatores extralinguísticos contribuíram para que o idioma entrasse em contato com outras línguas. Vale ressaltar que a língua portuguesa é um idioma que exportou mais vocábulos do que importou, fato decorrente da história em que outras línguas foram predominantes, como o caso do francês e do inglês. Justamente por esses fatores extralinguísticos, assim reforça Tagliviani (1993) ao dizer que “os empréstimos servem admiravelmente para reconstruir a história cultural de uma nação e sua relação com os outros povos”.

Uma breve sistematização dos empréstimos recebidos pela língua portuguesa pode ser vista em Cunha (2003, p. 8):

Entre as línguas que contribuíram para o enriquecimento vocabular do português, já nas suas origens, salientam-se o francês e o provençal. Datam do século XIII os primeiros empréstimos a estes dois idiomas, devidos, particularmente, à linguagem dos trovadores.

A partir do último quartel do século XV, logo após as grandes viagens empreendidas pelos portugueses à África, à Ásia e à América, e a descoberta do caminho das Índias, foram introduzidos no português algumas centenas de vocábulos oriundos dos idiomas indígenas dessas regiões – africanismos, asiaticismos e americanismos.

Ainda conforme o autor, é somente no início do século XVI que a língua italiana passa a exercer forte influência em toda a Europa culta, incluindo Portugal. Segundo Cunha (2003), os empréstimos do italiano estão ligados principalmente ao campo das artes e ocorrem em grande número na língua portuguesa, sendo também comuns no castelhano e no francês. Já a partir da segunda metade do século XVI e durante todo o século XVII o italiano perde destaque, dando lugar ao castelhano. Cunha (2003) afirma ainda que, nesta época, muitos escritores portugueses eram bilíngues e dominavam também o espanhol, emprestando para o português muitos termos castelhanos. A partir da segunda metade do XVII até o final do século XIX é a França quem domina a cultura europeia:

De meados do século XVII, durante todo o século XVIII e todo o século XIX é a França que domina o panorama cultural da Europa. Da França partem as novidades que se irradiam por todo o mundo, acompanhadas naturalmente dos vocábulos franceses com que elas se nomeiam. É a época da invasão dos galicismos, tão repudiados pelos puristas de Portugal, do Brasil, da Espanha e dos países da América Espanhola. (CUNHA, 2003. p. 8)

É com o início da Revolução Industrial que o português e outras línguas passam a receber inúmeros anglicismos. Conforme Cunha (2003), o inglês é a linguagem internacional das ciências, da política, da administração e de outras áreas. Ainda segundo o autor, foi principalmente depois da Segunda Guerra Mundial que os Estados Unidos da América tiveram grande destaque e, por conseguinte, a língua inglesa também:

Os novos anglicismos, oriundos dos Estados Unidos, e, por isso mesmo, melhor denominados anglo-norte-americanismos, difundem-se em todas as línguas do mundo. Os anglo-norte-americanismos estão presentes na língua portuguesa e em todas as línguas de cultura, graças ao notável progresso material e cultural da grande nação do norte do Novo Continente. (CUNHA, 2003. p. 8-9)

Depois de vermos a trajetória da língua portuguesa, constatamos o quão presente está o empréstimo na história da língua portuguesa. Para Langacker (1972), “o empréstimo não é nunca uma necessidade linguística, visto ser sempre possível ampliar e modificar o uso das unidades lexicais existentes para fazer face às novas necessidades de comunicação”. Manzóllilo (2014) lembra que, nos dias atuais, com o desenvolvimento vertiginoso da ciência e da tecnologia, da moda e dos esportes, nota-se que muitas vezes o vocábulo estrangeiro é a única possibilidade viável para os povos que importam esses produtos do progresso, por conta da velocidade com que esses produtos surgem no mercado.

### 3.2 Empréstimo linguístico, estrangeirismo ou neologismo?

Não diferente dos demais períodos históricos, na atualidade - com toda a sua rapidez e velocidade de mudança nas mais diferentes esferas da sociedade - conforme Manzolillo (2014), a utilização de um vocábulo estrangeiro é, geralmente, a única forma de nomear o que é importado com tanta velocidade, tornando mais rara a troca de nomenclaturas estrangeiras. Todavia, é preciso entender como tais vocábulos estrangeiros são acolhidos por uma língua, neste caso o PB e os diferentes fenômenos envolvidos. Por este motivo, apresentamos, na sequência, as definições de três eventos que muitas vezes se confundem entre si: o empréstimo linguístico, o estrangeirismo e o neologismo.

#### 3.2.1 Definição em dicionários

Esta seção pretende apresentar a definição dos termos “empréstimo linguístico”, “estrangeirismo” e “neologismo”, a partir de Ferreira (2004), Houaiss e Villar (2008) e Aulete (2011).

O termo empréstimo linguístico é encontrado no dicionário Aurélio sob a seguinte definição:

[Do português arcaico *empréstido*] Substantivo masculino. 05. *Estudos da Linguagem*. Efeito causado pelo contato prolongado entre línguas diferentes, quer por influência de uma cultura sobre a outra, quer pelo aprendizado de uma língua por grupos sociais, e que se traduz em mudanças no sistema linguístico e/ou no léxico, que passa(m) a incluir elementos não vernáculos. (FERREIRA, 2004, p. 737, grifos do autor).

Ainda sobre o empréstimo linguístico, encontra-se a seguinte definição no dicionário Houaiss:

Substantivo masculino. 05. *Linguística*. incorporação ao léxico de uma língua de um termo pertencente a outra língua. [Dá-se por diferentes processos, tais como a reprodução do termo sem alteração de pronúncia e/ou grafia (*know-how*), ou com adaptação fonológica e ortográfica (*garçom, futebol*)]. (HOUAISS; VILLAR, 2008, p. 1128, grifos dos autores).

Neste mesmo sentido, o vocábulo empréstimo também é encontrado no dicionário Aulete:

Substantivo masculino. 04. *Linguística*. Inclusão de vocábulo de outra língua no vocabulário da língua vernácula (p. ex: *checkup, shopping, download* etc.) -

**Lexical.** Formação de palavra num léxico por adoção de palavra de léxico estrangeiro, que pode ser na forma da palavra ou expressão em seu formato e significado originais (como *ace*, *know-how*, *déjà-vu*, *blitzkrieg*), como adaptação fonética e gráfica, mantendo seu significado (como futebol, abajur), como uma tradução calcada no termo estrangeiro (como ‘salvar’ no sentido de ‘gravar’, de *save*) ou como deturpação de significado, por proximidade fonética com palavra estrangeira (como a atribuição do sentido ‘constatar’, ‘perceber’, ‘dar-se conta’ ao verbo *realizar*, por influência do inglês *realize*). (AULETE, 2011, p. 542).

Com relação à definição principal de “empréstimo linguístico”, os três autores apresentam o mesmo conceito de que é a incorporação de um vocábulo estrangeiro à língua receptora. É importante destacar que Aulete (2011) define o termo de forma mais detalhada, exemplificando mudanças fonéticas, gráficas e semânticas.

No que tange à semântica da palavra estrangeirismo, encontramos, no dicionário Aurélio, as seguintes definições para o vocábulo:

[De estrangeiro + ismo] Substantivo masculino. 01. *Estudos da linguagem.* Emprego de palavra, frase ou construção sintática estrangeira; peregrinismo. 02. Estrangeirice. (FERREIRA, 2004, p. 834).

Já no dicionário Houaiss, pode-se encontrar a seguinte definição:

Substantivo masculino. 01. Influência geral forte da cultura, dos costumes etc. de determinada nação sobre outra ou sobre uma parcela significativa dos indivíduos desta. 02. *Linguística.* Palavra ou expressão estrangeira usada num texto em vernáculo, tomada como tal e não incorporada ao léxico da língua receptora; peregrinismo, xenismo; empréstimo. 03. Mesmo que: estrangeirice (no sentido de 'afeição'). (HOUAISS; VILLAR, 2008, p. 1261, grifos dos autores).

A palavra estrangeirismo também é encontrada no dicionário Aulete com os significados a seguir:

Substantivo masculino. 01. *Linguística.* Emprego de frase, ou palavra, ou construção sintática estrangeira: *Era adepto do estrangeirismo, mais por afetação do que por necessidade.* 02. Ver *estrangeirice*. 03. Palavra, expressão ou frase de língua estrangeira usados em texto falado ou escrito em vernáculo. **Estrangeirice.** *Substantivo feminino.* 01. Aquilo que se faz ou que se diz sob a influência de culturas estrangeiras. 02. Simpatia exagerada pelas coisas estrangeiras; ESTRANGEIRISMO. (AULETE, 2011, p. 615, grifos do autor).

No que diz respeito ao termo “estrangeirismo” os três autores o definem como o uso de um termo ou expressão estrangeiro que não está incorporado à língua receptora.

Por fim, ao buscarmos a significação da palavra neologismo no principais dicionários do PB, encontramos a seguinte definição no Aurélio:

[De ne(o)- + -log(o)- + -ismo.] Substantivo masculino. 01. *Estudos da linguagem*. Palavra ou expressão nova numa língua, como, por exemplo, *dolarizar*, *dolarização*, no português. 02. *Por extensão*. Significado novo que uma palavra ou expressão de uma língua pode assumir. [Por exemplo: *zebra*, como “resultado inseperado”.] (FERREIRA, 2004, p. 1395).

No dicionário Houaiss, a palavra neologismo é encontrada sob as seguintes significações:

Substantivo masculino. *Linguística*. 01. Emprego de palavras novas, derivadas ou formadas de outras já existentes, na mesma língua ou não. 02. Atribuição de novos sentidos a palavras já existentes na língua. 03. Unidade léxica criada por esses processos. Coletivo: neologia. Antônimos: arcaísmo. Homônimos: neologismar<sup>31</sup>. (HOUAISS; VILLAR, 2008, p. 2009).

Ainda sobre a definição do vocábulo neologismo, podemos conceituá-lo do seguinte modo, conforme o dicionário Aulete:

Substantivo masculino. *Linguística*. 01. Uso de palavra ou expressão nova, geralmente com base em léxico, semântica, e sintaxe preexistentes, na mesma língua ou em outra. 02. *Por extensão*. Qualquer palavra ou expressão resultante desse processo. (AULETE, 2011, p. 966, grifos do autor).

O neologismo é, de acordo com os três dicionários consultados, a criação de novas palavras, formadas ou derivadas de palavras já existentes, podendo ser da mesma língua ou não. Ferreira (2004) define ainda como neologismo uma palavra ou expressão que assume novo significado.

De acordo com as definições encontradas sobre os três termos, pode-se identificar a ocorrência dos três fenômenos em nossa amostra de dados, como podemos ver nas ocorrências a seguir:

- Empréstimo linguístico:

(3) “[...] agora eu vou esperar ter *internet*<sup>32</sup> para ver como resolver o problema” (publicado em 28 de fevereiro de 2020, em Chapecó, grifos nossos);

- Estrangeirismo:

(4) “Tá todo mundo fazendo *live* e *stream*”<sup>33</sup> (publicado em 12 de maio de 2020, em Chapecó, grifos nossos),

- Neologismo:

<sup>31</sup> Verbo intransitivo. Produzir e/ou usar neologismos; neologizar. (HOUAISS; VILLAR, 2008, p. 2009)

<sup>32</sup> rede virtual, tradução nossa.

<sup>33</sup> ao vivo e transmissão de conteúdo, tradução nossa.

(5) “#FutebolFeminino #BRASIL #VARgonha<sup>34</sup> #FIFAWWC” (publicado em 13 de junho de 2020, em Chapecó, grifos nossos).

O vocábulo *internet* é considerado um empréstimo linguístico porque ele foi incorporado ao PB, e já está dicionarizado. A maioria dos falantes não o veem mais como estrangeiro. A palavra *live*, por outro lado, é considerada um estrangeirismo porque ainda não foi incorporada ao nosso idioma. Por último, *VARgonha* é considerado um neologismo por analogia, visto que é uma criação a partir das palavras *VAR* (*Video Assistant Referee*) e *vergonha*.

A gramática é, em termos globais, o modo como estão os princípios gerais da linguagem humana em uma determinada língua. Existem diversos tipos de gramáticas; entre elas, estão a descritiva, a teórica, a formal, a tradicional, a histórica e a gerativa. Interessa-nos a definição desses termos nas gramáticas descritivas, conforme tratamos a seguir.

### 3.2.2 Definição em gramáticas descritivas

Para compreender ainda mais a definição dos termos empréstimo, estrangeirismo e neologismo, buscamos tais conceitos a partir da perspectiva de Bagno (2012), Castilho (2019) e Moura Neves (2003).

Conforme Crystal (2000), a gramática descritiva é primeiramente a descrição de uma língua da forma como ela é encontrada em amostras da fala ou da escrita, em *corpus* sintetizados a partir dos falantes nativos. Ainda, segundo o autor, dependendo da base teórica, a gramática descritiva pode ainda fazer registros sobre a língua como um todo e, se essas afirmações forem explícitas e previrem a competência do falante, a gramática poderá ser “descritivamente adequada”. Por fim, Crystal (2000 [1985]) aponta ainda para o fato de que a abordagem mais antiga da gramática descritiva se opunha a alguns gramáticos prescritivos que tentavam estabelecer regras para o emprego social ou estilisticamente correto para o uso da língua.

Bagno (2012) mostra que o empréstimo linguístico se dá quando uma determinada língua recebe contribuições lexicais de outras línguas. Sob uma perspectiva extralinguística, o autor afirma que fenômenos culturais, políticos, históricos, econômicos e outros se juntam para que palavras de outras línguas, geralmente a língua dominante no cenário mundial, se instalem no léxico de outras.

---

<sup>34</sup> Neologismo por analogia de *VAR* (*Video Assistant Referee*) + *gonha* remetendo-se ao vocábulo *vergonha* para descrever o uso do árbitro assistente de vídeo (VAR) em jogos de futebol, tradução nossa.

Para Castilho (2019), a lexicalização por empréstimo acontece quando importamos palavras, sufixos e prefixos de línguas com que estivemos em contato direto ou indireto. Castilho (2019, p.114) afirma também que o contato direto se dá quando duas ou mais línguas ocuparam o mesmo território:

Historicamente isso ocorre nas seguintes possibilidades: (1) empréstimos tomados ao substrato linguístico\*, resultante do contato com povos que ocupavam anteriormente o território invadido, cuja cultura foi suplantada pelo invasor; (2) empréstimos tomados ao superstrato linguístico\*, resultante do contato com povos que invadiram um território sem suplantarem a cultura de seus ocupantes.

Sobre os estrangeirismos que ingressaram de forma indireta no PB, Castilho (2019, p.119) acrescenta que:

São considerados estrangeirismos as palavras francesas, espanholas, e norte-americanas que ingressaram no vocabulário do PB [português brasileiro] sem que tivesse ocorrido um contato direto com a cultura brasileira. Momentos de nacionalismo na cultura linguística brasileira levaram a campanhas contra os galicismos, os espanholismos e os anglicismos. Em tempos de globalização, dois movimentos antitéticos parecem esboçar-se: a incorporação de estrangeirismos *versus* sua repulsa, como forma de sustentar nossa identidade linguística.

Já no que concerne à lexicalização por neologia, Castilho (2019) afirma que este fenômeno ocorre quando criamos uma palavra nova, não herdada da língua-fonte, porém organizada de acordo com as regras morfológicas da língua receptora, como é o caso do verbo *coisar*, criado a partir do substantivo *coisa*.

Para Garcez e Zilles (2001), a discussão atual sobre os estrangeirismos se centraliza no uso de termos oriundos da língua inglesa, os anglicismos, que apresentamos na seção a seguir.

### 3.3 Anglicismo

Observando as diversas possibilidades de adições ao léxico, percebe-se que atualmente os anglicismos são os empréstimos mais recorrentes em diversas línguas dado o prestígio da língua inglesa, conforme apontam Correia e Almeida (2012, p. 69):

Atualmente, dado o predomínio da língua inglesa como a língua de comunicação internacional (científica, técnica, política), pelo predomínio geoestratégico de países de fala anglo-saxônica, a maioria das demais línguas são importadoras de palavras da língua inglesa.

Ferreira (2010, p. 46) define o termo anglicismo como uma palavra ou locução inglesa introduzida em outra língua e empregada como se fosse desta. Houaiss e Villar (2011, p. 58) definem o anglicismo como uma palavra ou expressão inglesa em outra língua. Garcez e Zilles (2001) apontam para o fato de que os anglicismos estão aumentando em diversas línguas, inclusive no PB, por causa de diversos fatores extralinguísticos, como o fato de a língua inglesa ser utilizada como língua franca para o contato internacional, além de as pesquisas e tecnologias serem registradas quase que em sua totalidade nessa língua.

Já Bechara (2009) conceitua anglicismo como um estrangeirismo e ressalta que a língua inglesa usa constantemente o grego e de tal forma que muitas vezes deturpa o emprego ou o significado original de determinadas palavras.

### 3.3.1 Pesquisas sobre anglicismos no português do Brasil

Localizamos, por meio de levantamento bibliográfico, algumas pesquisas que investigaram os anglicismos no PB sob diferentes perspectivas teóricas. Reunimos contribuições dos seguintes estudos: Carvalho (1989), Manzollilo (2000), Arraes (2006), Leitão (2006), Alves (2007), Torrano (2010), Correia e Almeida (2012), Pereira (2012), Ferreira e Gomes (2015), Souza (2015), Souza *et al.* (2015), Prado (2015), Ferreira (2016) e Castro (2017).

Apesar de os anglicismos parecerem algo mais recente, principalmente com a popularização do uso da internet e das redes sociais digitais, estes empréstimos já vêm sendo estudados há algum tempo. Carvalho (1989) afirma que o inglês funciona como uma língua franca praticamente em todo o mundo - uma língua que facilita o intercâmbio. Conforme a autora, a exportação de bens de consumo e tecnologias, além da expansão de empresas multinacionais, faz com que a língua inglesa exerça grande influência no léxico de vários povos. Carvalho (1989) informa que o acervo lexical da língua portuguesa no dicionário Aurélio gira em torno de 150.000 verbetes<sup>35</sup>, porém, segundo o dicionário Houaiss, com as terminologias especializadas, o acervo salta para 400.000 verbetes. Além disso, a autora afirma que esta nova “corrente estrangeira” geralmente é apresentada primeiro em seções especializadas dos grandes periódicos, antes de se integrar ao dicionário comum. Carvalho (1989) afirma que 95% dos vocábulos estrangeiros vêm da língua inglesa e, para comprovar tal afirmação, a autora criou um *micro-corpus* composto por 37 anglicismos retirados de reportagens e publicidades dos jornais

---

<sup>35</sup> “[...] o dicionário Houaiss, lançado em 2001, lista 400 mil palavras. Segundo Ieda Alves, o número de vocábulos que realmente existem é um pouco maior. “Se levarmos em consideração as palavras técnicas e científicas, devem existir cerca de 600 mil palavras na língua portuguesa”, estima.” (SATO, 2009, p. 1)

“Folha de São Paulo” e “Diário de Pernambuco”, dos quais escolheu três seções: economia, informática e eletrônica. A seguir, podemos ver um exemplo de anglicismos extraídos de cada seção, respectivamente:

O mais interessante *holding* de química do país, a Odebrecht. (*Folha de S. Paulo*, 27 maio 1987 in CARVALHO, 1989, p. 62).

Editor de texto e *open access* integrado, respectivamente. (*Diário de Pernambuco*, 06 jun. 1987 in CARVALHO, 1989, p. 64).

O sistema digital *Pulse-Code Modulation* (PCM) irá captar a qualidade do som... (*Folha de S. Paulo*, 02 jun. 1987 in CARVALHO, 1989, p. 65).

No quesito tecnologia, a autora afirma que, ao contrário da ciência, que independe de uma cultura particular, no sentido de universalidade, a tecnologia, por sua vez, é mensageira de uma forma de encarar o mundo, de uma cultura e acrescenta ainda:

Importar termos, fazer empréstimos linguísticos, é importar noções de um sistema de valores ou de pressuposições diferentes. Na ausência total ou parcial de equivalência linguística, conceitual ou funcional de uma língua e cultura para a outra, todo empréstimo é gerador de estruturas fônicas, morfológicas, sintáticas, semânticas e retóricas novas. (CARVALHO, 1989, p.74).

Para Manziolillo (2000), o empréstimo linguístico é o conceito que se refere às unidades lexicais que de alguma forma já estão integradas ao novo ambiente, a nova língua, ou seja, em última circunstância, são palavras que já estão dicionarizadas. No mesmo sentido, Carvalho (1989, p. 42 ) afirma que “o empréstimo tem sua origem no momento em que objetos, conceitos e situações nomeados em língua estrangeira transferem-se para outra cultura”. Já o fenômeno do estrangeirismo, no entendimento de Manziolillo (2000), diz respeito àquelas palavras francamente alienígenas, que ainda não foram incorporadas ao léxico da língua recebedora.

Manziolillo (2000, p. 21) ressalta que é comum, principalmente pela imprensa brasileira, o emprego estilístico de estrangeirismos e destaca alguns usados em jornais impressos: “winter sale - até 50% de desconto” (*O Globo*, 17/07/99); “O padre contou detalhes da love story dos noivos e explicou o ritual do sacramento” (*O Globo*, Segundo Caderno, 06/07/99, p. 3). O autor afirma ainda, no que diz respeito à morfossintaxe, que um bom sinal de integração ocorre quando a palavra estranha funciona como uma matéria-prima para outras formações lexicais, podendo ser realizadas por meio de processos de composição, derivação e a palavra-valise<sup>36</sup>. O autor apresenta os seguintes exemplos desses termos: *funkeiro*, *jazzista*, *bluesista*, *kartista* e

<sup>36</sup> Palavra-valise é um vocábulo que resulta da combinação da parte inicial de uma palavra com a parte final de outra, como em: motel (motor e hotel), namorido (namorado e marido), portunhol (português e espanhol), showmício (show e comício), blend. (Michaelis On-line, 2021)

*showmício*. O autor exemplifica ainda anglicismos que servem como base para novas unidades lexicais, a exemplo de *lobby* > lobista; *standard* > estandardização e *rock* > roqueiro. Manzolillo ressalta ainda que, “apesar de Ferreira (1999) já registrar as formas aportuguesadas roque e lóbi, indiscutivelmente, os lexemas estrangeiros rock e lobby [...] apresentam uma frequência de uso muito maior.” (MANZOLILLO, 2000, p. 23-24).

Leitão (2006) analisou anglicismos registrados nos dicionários Aurélio 3.0 (1999) e Houaiss 1.0 (2001), nos quais encontrou 652 e 1072 unidades lexicais, respectivamente, marcadas como recentemente incorporadas. Vale destacar que a autora afirma que não há o mesmo método lexicográfico que oriente a inserção de novos vocábulos advindos da língua inglesa nos dicionários Aurélio 3.0 (1999) e Houaiss 1.0 (2001):

[...] pode-se dizer que não há um procedimento lexicográfico uniforme que oriente ou que norteie a inserção de unidades lexicais de língua inglesa nos dicionários Aurélio 3.0 (1999) e Houaiss 1.0 (2001), as maiores obras lexicográficas do português. É importante notar que assim como não há critérios idênticos, não há ainda um método lexicográfico sobre o assunto. (LEITÃO, 2006, p. 135).

Baseando-se no fato de que os dicionários apresentam uma amostra atestada dos anglicismos importados para o PB, Arraes (2006) também utilizou os dicionários Aurélio e Houaiss como fonte de recolha de dados, ambos na versão *on-line*. Ao todo, foram coletados 2106 dados, dos quais 160 foram analisados. Os vocábulos analisados foram atribuídos a três classes de palavras, sendo elas: substantivo, adjetivo e verbo.

Procurando definir neologia por empréstimo, Alves (2007) foi em busca de anglicismos utilizados em matérias de jornais da década de 1980. Conforme a autora, o estrangeirismo é facilmente encontrado em vocabulários técnicos, como informática, economia, esportes, além da publicidade e do colunismo social. A título de exemplo, Alves (2007) cita as corridas automobilísticas como um esporte que emprega vários anglicismos no PB, como pode ser exemplificado a partir deste recorte de uma matéria da Folha: “Ayrton Senna foi pole-position pela 13ª. vez em uma temporada de 16 corridas. [...] Faltavam dois minutos para o final [...] quando Ayrton fez sua *flying lap* (volta rápida)” (F, 13-11-88: D-2, c. 1).

Segundo a autora, muitas vezes o emissor do estrangeirismo é consciente de que ele poderá não ser interpretado de modo adequado pelos receptores do texto. Por isso, conforme Alves (2007), muitas vezes o vocábulo estrangeiro é seguido de tradução, como em “[...] enquanto nos Estados Unidos o chamado *gray power* - o poder grisalho - assusta todos os

candidatos à presidência” (IE, 27-01-88: 28, c. 1). A autora relata ainda que, morfossintaticamente, a integração ao PB manifesta-se quando o estrangeirismo passa a formar derivados e compostos, como o caso de *new-jeca* (novo caipira). Alves (2007) destaca ainda que a maioria dos neologismos por empréstimo recebidos pelo PB são distribuídos principalmente entre a classe substantival e, raramente, entre as classes verbais e adjetivais.

No mesmo sentido, Pereira (2012) destaca que alguns empréstimos podem conservar sua grafia e/ou pronúncia originais e outros podem sofrer uma adaptação fonológica e ortográfica, como se pode notar nos exemplos *cheese* > x ou *show*, que manteve a forma original. Ainda sobre os anglicismos, Pereira (2012, p. 7, grifos do autor) destaca que os vocábulos oriundos da língua inglesa entraram em ascensão após a Segunda Guerra Mundial, principalmente após o advento da internet:

[...] com o avanço da informática e o [...] surgimento da internet, o fenômeno de empréstimos linguísticos da língua inglesa (anglicismos) passa de uso necessário de palavras que não encontram tradução no português brasileiro, como o caso de *hardware* e *software*, ao que, para muitos, passou a ser o uso desnecessário do estrangeirismo, como o caso de palavras *workshop*, *coffee break*, *sale*, *fast food*, etc.

Pereira (2012) ressalta ainda que muitos termos são inseridos em nosso léxico com adaptações morfossintáticas, como o caso de *logar*, que deriva do verbo inglês *to log in*, e o verbo deletar, de *to delete*. O autor exemplifica também as palavras *download* e *software* com os correspondentes baixar e programa, respectivamente, e a palavra *bullying*, que ainda não encontra uma palavra correspondente no PB.

Ferreira e Gomes (2015), por sua vez, analisam dados extraídos do programa televisivo *Manhattan Connection*, programa brasileiro transmitido em canal pago. Conforme os autores, a grande maioria dos empréstimos lexicais recebidos pelo português brasileiro são substantivos, sendo raro o empréstimo de adjetivos e verbos. Os autores destacam vocábulos como *bartender* (garçom de bar) e *crowdfunding*, que transitam nos discursos entre substantivo e locução adjetiva, *upgrade*, que originalmente é um verbo, mas, em português, exerce a função de substantivo e também palavras como *stress* > estresse, *football* > futebol e *delect* > delete, que sofreram mudanças morfológicas.

Ainda na busca pelo elencamento dos anglicismos, Ferreira (2016) analisou três versões digitais do Dicionário Aurélio, sendo elas: 3.0 (1999), 5.0 (2004) e 8.0 (2010). Durante a análise a autora encontrou 687 estrangeirismos de língua inglesa na versão 3.0 (1999), 689 na versão 5.0

(2004) e 888 na versão 8.0 (2010). Ferreira (2016) fez um levantamento de 2264 anglicismos nas três versões do Dicionário Eletrônico Aurélio.

A autora destaca sete áreas (rubricas) nas versões 3.0 (1999) e 5.0 (2004) do dicionário em que mais aparecem os anglicismos. No dicionário 3.0 (1999) a autora evidencia as seguintes áreas:

- 1) Cinema - 15 verbetes (Exemplos: *Black light*, *Flash* e *Storyboard*);
- 2) Economia - 25 verbetes (Exemplos: *Paper*, *Commodity* e *Blue chip*);
- 3) Esporte - 21 verbetes (Exemplos: *Body-board*, *Beach-soccer* e *Inning*);
- 4) Informática - 70 verbetes (Exemplos: *Assembler*, *Backbone* e *Bit*);
- 5) *Market* - 19 verbetes (Exemplos: *Display*, *Follow-up* e *Mock-up*);
- 6) Por extensão - 21 verbetes (Exemplos: *Briefing*, *Check-up* e *Courier*);
- 7) Televisão - 27 verbetes (Exemplos: *Broadcast*, *Dolly* e *Fade*).

Já na versão 5.0 (2004), a autora encontrou somente dois anglicismos a mais no total e destacou as seguintes áreas:

- 1) Cinema - 15 verbetes (Exemplos: *Close-up*, *Master* e *Prompter*);
- 2) Economia - 25 verbetes (Exemplos: *Dumping*, *Factoring* e *Float*);
- 3) Esporte - 24 verbetes (Exemplos: *Bodyboarding*, *Cross-country* e *Match*);
- 4) Informática - 72 verbetes (Exemplos: *Chat*, *Clipboard* e *Default*);
- 5) *Market* - 21 verbetes (Exemplos: *Endomarketing*, *Mailing list* e *Recall*);
- 6) Por extensão - 21 verbetes (Exemplos: *Cockpit*, *Flashback* e *Light*);
- 7) Televisão - 27 verbetes (Exemplos: *Network*, *Off-line* e *Replay*).

Por fim, na versão 8.0 (2010) do Dicionário Eletrônico Aurélio, Ferreira (2016) encontrou 888 verbetes oriundos da língua inglesa, evidenciando tanto o aumento considerável de verbetes quanto o aumento de rubricas. Destacaram-se as áreas (rubricas) a seguir:

- 1) Cinema - 16 verbetes (Exemplos: *Dolly-in*, *Set* e *Take*);
- 2) Economia - 30 verbetes (Exemplos: *Agribusiness*, *Crowding-out* e *Float*);
- 3) Esporte - 48 verbetes (Exemplos: *Hole-in-one*, *indoor* e *kickbox*);
- 4) Informática - 112 verbetes (Exemplos: *Drive*, *Enter* e *Freeware*);
- 5) Jornalismo - 15 verbetes (Exemplos: *Fanzine*, *Feature* e *Gossip*);
- 6) *Market* - 20 verbetes (Exemplos: *Merchandising*, *Prospect* e *Target*);

- 7) Música - 25 verbetes (Exemplos: *Anthem*, *Blues* e *Grunge*);
- 8) Por extensão - 26 verbetes (Exemplos: *Lockout*, *Nonsense* e *Offshore*);
- 9) Tecnologia têxtil - 15 verbetes (Exemplos: *Canvas*, *Oxford* e *Plush*);
- 10) Televisão - 26 verbetes (Exemplos: *Traveling*, *Videomaker* e *Zapping*).

Nota-se o grande aumento de verbetes oriundos da língua inglesa quando se comparam as versões 3.0 (1999) e 5.0 (2004) com a versão 8.0 (2010), principalmente no quesito informática e meios de comunicação. Ferreira (2016, p. 76) destaca:

Os dados evidenciam a ampliação lexical dessas áreas, fenômeno que deve ter ocorrido com maior rapidez em língua inglesa e, em razão da globalização dos meios de comunicação, essas unidades lexicais foram transpostas para a língua portuguesa, pela necessidade de se nomear novos referentes.

Correia e Almeida (2012) fazem distinção entre “estrangeirismo” e “empréstimo” e afirmam que, conforme a gramática tradicional portuguesa, “estrangeirismo” é uma unidade importada de outra língua que não passou por quaisquer adaptações à língua receptora, enquanto que “empréstimo” é um vocábulo estrangeiro que se adaptou ao sistema linguístico de acolhimento, isto é, dentro do cenário brasileiro, foi aportuguesado.

As autoras também destacam o predomínio do inglês como língua de comunicação internacional nos dias atuais, sobretudo nas áreas científica, política e técnica, resultado do predomínio geoestratégico de países de fala anglo-saxônica. Correia e Almeida (2012) lembram que a ascensão da língua inglesa é relativamente recente e passou a ocorrer no período após a Segunda Guerra Mundial. As autoras apresentam exemplos de anglicismos já dicionarizados pelo Aurélio e Houaiss como *escâner*, *escanear*, *xampu*, *estresse*, *sanduíche* e outros. Correia e Almeida (2012, p. 74) expõem ainda estudos mais recentes sobre a importação de palavras e mostram que, em resumo:

[...] toda e qualquer palavra que entra num sistema linguístico se adapta a esse sistema, ao menos fonologicamente, por isso que hoje em dia é discutível a manutenção dos conceitos de estrangeirismo e empréstimo. Porém, esses conceitos estão consagrados pela tradição [...]

Torrano (2010) procurou analisar neologismos empregados na área da informática, coletando um *corpus* da “Revista Info”, nas versões impressa e digitalizada dos anos 2007 e 2008. Ao todo foram analisadas 12 edições impressas e 12 edições digitalizadas, das quais foi possível retirar 769.435 ocorrências de neologismos. A partir desta extração a autora fez a divisão por número de ocorrências (frequência), reduzindo o *corpus* a 31.308 vocábulos. As unidades lexicais foram analisadas a partir dos seguintes critérios de exclusão, conforme Torrano

(2010): nomes de fabricantes de equipamentos, unidades de medida, vocábulos de reconhecido uso comum tanto em português quanto em inglês, registro em dicionário. Após a análise do *corpus*, a autora reconheceu 745 neologismos, cuja frequência média correspondia a até dez ocorrências por edição. Das 24 edições selecionadas para o *corpus*, somente 15 edições da revista foram analisadas, entre o período de janeiro de 2007 e novembro de 2008, devido a problemas técnicos.

Após revisão dos dados coletados, Torrano (2010) constatou que 46,31% do *corpus* de sua análise é constituído por estrangeirismos, ou seja, palavras que continuaram com sua forma original do inglês e que ainda não passaram por alterações segundo os padrões ortográficos e gramaticais do PB. A autora destaca também que a derivação (virtualização), as formações acronímicas (*laser*) e as siglas (*PDF*) são processos frequentes na formação de termos da área da Informática, sendo juntos responsáveis por 42,97% dos processos utilizados. Os processos de formação de neologismos do *corpus* de pesquisa de Torrano (2010, p. 45-46) foram separados da seguinte forma:

Quadro 1: Dados de Torrano (2010).

<b>Processo de formação</b>	<b>Quantidade</b>	<b>Percentual</b>
Estrangeirismo	345	46,31%
Derivação Prefixal	122	16,38%
Sigla	108	14,50%
Formação Acronímica	43	5,77%
Derivação Sufixal	33	4,43%
Composição por Subordinação	26	3,49%
Formação Sintagmática	25	3,36%
Neologismo Semântico <sup>37</sup>	20	2,68%
Decalque	12	1,61%
Composição por Coordenação	5	0,67%
Palavra-valise	3	0,40%

<sup>37</sup> “O neologismo semântico ou conceptual baseia-se na manutenção da forma com alteração do significado do termo, resultante de uma alteração no conjunto de semas referentes a uma unidade léxica.” (TORRANO, 2010, p. 19).

Analogia <sup>38</sup>	1	0,13%
Derivação Regressiva	1	0,13%
Estrangeirismo adaptado	1	0,13%
<b>Total</b>	745	100%

Fonte: Torrano (2010, p. 46-47)

De acordo com a autora, a numerosa presença do estrangeirismo, sobretudo do anglicismo, em sua pesquisa, mostra a influência da língua inglesa no campo da Informática. Segundo Torrano (2010), por se tratar de uma área que surgiu em um país de língua inglesa, os termos foram criados nessa língua, sendo dispensável - na maioria dos casos - a tradução. A autora afirma ainda que, por ter extraído o *corpus* de uma revista especializada em informática, foram encontrados muitos termos estrangeiros e estes não foram destacados entre aspas ou em itálico, outros marcam o plural sempre com -s no final. Torrano (2010, p. 89) mostra alguns exemplos:

Com um bom computador, um microfone, um software de edição e outro de compressão de áudio, qualquer amador produz <podcasts>. (Revista Info, jan. 2007 - destaque da autora).

Depois do sucesso de alguns serviços online, como o Gmail e o Del.icio.us, as <tags> (etiquetas) ficaram populares. No Vista, é possível adicionar <tags> e categorias aos arquivos multimídia, o que facilita a organização do HD e as buscas nele. (Revista Info, mar. 2007 - destaques da autora).

E esses recursos tendem a aumentar graças a arquitetura de <widgets> do Joost. (Revista Info, jun. 2007 - destaque da autora).

Torrano (2010) ressalta também que alguns estrangeirismos aparecem com grafia variante, como o caso de *touchscreen* ou *touch screen* e *touchpad* ou *touch pad*. No que tange aos estrangeirismos adaptados, somente o termo *blogue* apresentou alteração gráfica, cuja ortografia foi adaptada ao PB. A autora também mostra que a forma estrangeira pode conviver com seu decalque, tradução literal, como caso de placa-mãe, mas que estes estrangeirismos aparecem em menor quantidade.

Conforme Torrano (2010), as proporções entre os processos de formação se mantiveram estáveis durante as 15 edições da Revista Info e, apesar de o estrangeirismo ter sido o processo mais recorrente, a soma dos processos vernáculos permaneceu maior, o que afasta a ideia de uma invasão dos estrangeirismos ou ainda uma ameaça ao PB.

<sup>38</sup> “Palavra de difícil classificação, o termo *conservapedia* (= conservar + pedia) formou-se por analogia à forma *Wikipédia*”. (TORRANO, 2010, p. 107).

Souza *et al.* (2015) demonstram a partir dos estrangeirismos o poder da língua inglesa nas mais diversas esferas da sociedade brasileira. Assim como os outros estudiosos, anteriormente mencionados, estes autores reafirmam a relevância nos dias atuais da língua inglesa, não só no Brasil, mas em vários outros países, sobretudo devido a fatores sociais, econômicos e também graças aos avanços tecnológicos e científicos. Para exemplificar a relevância do inglês no português, Souza *et al.* (2015, p. 37-38, grifos dos autores) exemplificam alguns termos facilmente encontrados em manuais de instruções, como:

Power ou Play ao invés de Liga, *Stop* no lugar de Pare ou Desligue, *Open/Close* para abrir e fechar uma gaveta de *Compact Disk* (CD), *Stand By* para colocar o aparelho em modo de espera, *Tunning*, etc. O mesmo acontece com os *video games*, todos com marca nesse idioma: *Mega Joy*, *Play Station*, *Dynavision*, *Gun Boy* entre outros. Alguns de nossos alimentos hoje também são atingidos trazendo marcas como *Qualy*, *Waffles*, *Chips*, *Cream Cracker*, *Claybom* etc. Quando uma pessoa vai a uma lanchonete, para fazer ou comprar um lanche, pede um *hot-dog*, um *hamburger*, *cheeseburger*, acompanhado de uma *Diet Coke*.

É importante ressaltar também que, para os autores, hoje o conhecimento do inglês pode oferecer melhores salários, além do melhor uso e manuseio de recursos acessíveis no mercado atual. Conforme Souza *et al.* (2015), o inglês, por questões de dominação histórica e econômica, tornou-se uma língua universal. O Brasil sofre essa influência principalmente dos Estados Unidos, presentes no cenário mundial de tal forma que, de acordo com os autores, os países periféricos não estão mais sendo dominados somente pela força armada, mas pela matéria-prima do império cultural, a língua.

Partindo do pressuposto de que a língua inglesa está presente nas mais diversas áreas da sociedade, Prado (2015) construiu seu *corpus* de pesquisa com nomes de estabelecimentos comerciais do interior do estado de São Paulo, para representar o PB e também nomes de comércios da capital portuguesa, Lisboa, representando o Português Europeu (PE). Os dados do PB foram coletados a partir do *site* Guia Mais ([www.guiamais.com.br](http://www.guiamais.com.br)), do qual a autora optou por fazer um recorte e coletar dados somente do interior de São Paulo. Ao todo foram encontrados 7271 comércios dos quais 862 apresentaram algum elemento da língua inglesa. Para a coleta de dados do PE, Prado (2015) utilizou os seguintes *sites*: Info Empresas ([www.infoempresas.com.pt](http://www.infoempresas.com.pt)), Restaurantes Lisboa ([www.restauranteslisboa.net](http://www.restauranteslisboa.net)), Hotéis ([www.hoteis.pt](http://www.hoteis.pt)) e Páginas Amarelas ([www.pai.pt](http://www.pai.pt)). A autora encontrou 9.093 estabelecimentos comerciais, dos quais 446 apresentaram elementos do inglês. A seguir, exemplificamos o Quadro 2, que representa os nomes comerciais com elementos em inglês no português do Brasil:

Quadro 2: Dados de Prado (2015).

<b>Categoria</b>	<b>Nº de nomes coletados</b>	<b>Porcentagem</b>
Academias desportivas	84	10%
Automóveis - peças e serviços	67	8%
Cabeleireiros e institutos de beleza	207	24%
Informática - equipamentos e assistência	85	10%
Hotéis e motéis	64	7%
Lavanderias	26	3%
Móveis	34	4%
Padarias e confeitarias	17	2%
Restaurantes e bares	104	12%
Roupas	174	20%
<b>Total</b>	<b>862</b>	<b>100%</b>

Fonte: Prado (2015, p. 102).

De modo mais abrangente, Castro (2017) pesquisa os processos e usos de neologismos na internet. A autora fez a análise da formação morfológica das inovações lexicais. Dos dez neologismos estudados por Castro (2017), destacamos dois que surgiram a partir de anglicismos: tuiteiro e shippo. Vejamos, na sequência, como a autora apresentou os dados:

Quadro 3: Dados de Castro (2017).

<b>Neologismo:</b> tuiteiro
<b>Contexto:</b> Matéria do jornal Estadão intitulada “O ‘tuiteiro em chefe’”
<b>Significado:</b> Usuário do twitter. Pessoa que usa demasiadamente o twitter.
<b>Estrutura morfológica:</b> tuit + eiro: tuiteiro <div style="text-align: center;">                 <b>nome sufixo</b> </div>
<b>Categoria morfosintática:</b> adjetivo
<b>Tipo de unidade:</b> derivação sufixal
<b>Fonte:</b> www. internacional.estadao.com.br
<b>Domínio de referência:</b> interpessoal

<b>Neologismo:</b> Shippo
<b>Contexto:</b> Página do facebook intitulada “Shippo”
<b>Significado:</b> Ato de torcer para que duas pessoas ou personagens formem um par romântico.
<b>Estrutura morfológica:</b> shippar - shippo (1ª pessoa do singular) <i>Relationship - ship - shippar</i>
<b>Categoria morfossintática:</b> verbo
<b>Tipo de unidade:</b> derivação regressiva
<b>Fonte:</b> www.facebook.com/estoushippingando
<b>Domínio de referência:</b> interpessoal

Fonte: Castro (2017, p. 14-15).

Conforme Castro (2017), o dado “tuiteiro” forma-se da palavra “twitter” mais o sufixo “eiro”, sendo uma unidade formada a partir de derivação sufixal, processo que une dois vocábulos, formando uma nova palavra. Já o dado “shippo” foi criado através de derivação regressiva e, de acordo com a autora, o verbo “shippar” pode ser conjugado em todas as pessoas pronominais. “Shippo” é um neologismo que deriva da palavra inglesa “relationship” e é usado somente em contextos informais e está no domínio interpessoal, ainda segundo Castro (2017).

Souza (2015) retirou uma lista de neologismos para análise de seu estudo do *Projeto AC/DC: corpo Corpus Brasileiro*, que se trata de uma coletânea de aproximadamente um bilhão de palavras do PB, resultado de projeto coordenado por Tony Berber Sardinha, com financiamento da Fundação de Amparo à Pesquisa de São Paulo (Fapesp). Os dados do *corpus* foram filtrados do *Projeto AC/DC* pela ordem de ocorrência de lexicalização, sendo separados em mínima/ baixa frequência (limite de 01 a 10 ocorrências, com 100% das amostras coletadas), baixa (limite de 25 e 100 ocorrências, com 10% das amostras coletadas), média (101 a 999 ocorrências, com 5% das amostras coletadas) e alta (com mais de 1000 ocorrências, com 1% das amostras coletadas). Dos dados coletados e analisados pelo autor, destacamos os que têm origem na língua inglesa:

Quadro 4: Dados de Souza (2015).

<b>Amostra</b>	<b>Quantidade de ocorrências</b>	<b>Sentidos identificados a partir de certos contextos de uso</b>
(subs.) internauta (subs.) internautas	<b>2.063</b> casos / <b>1.875</b> casos	1. (Substantivo): pessoa ou usuário da internet.

(v./subs./adj.) fake (v./subs./adj.) fakes	<b>187 casos / 26 casos</b>	1. (Substantivo): imitação. 2. (Adjetivo): falso. 3. (Verbo): fingir.
(subs.) cracker	<b>120 casos</b>	1. (Substantivo): pessoa ou usuário que faz invasões em <i>sites</i> ; posta e criptografa arquivos.
(v.) plugar	<b>37 casos</b>	1. (Verbo): conectar; ligar (sentido físico). 2. (Verbo): conectar; interligar (sentido metafórico).
(subs.) interneteiro	<b>10 casos</b>	1. (Substantivo): usuário viciado na internet.
(subs.) analfabyte (subs.) analfabytes	<b>01 caso / 01 caso</b>	1. (Substantivo): pessoa que desconhece as ferramentas da internet e a própria informática.

Fonte: Souza (2015, p. 109).

Souza (2015) destaca que o centro da língua é o “Dispositivo Conversacional” e é ele que comanda os sistemas semântico, discursivo, gramatical e lexical. O autor afirma também que estes sistemas são independentes e dispõem de características próprias que não interferem nas outras. Ainda segundo o estudo de Souza (2015), os itens lexicalizados que são selecionados pelos falantes se juntam com outros elementos linguísticos do léxico, formando um todo na comunicação. A partir desta perspectiva, o autor aponta para o fato de, independentemente de os falantes utilizarem empréstimos linguísticos, estrangeirismos ou neologismos, o “Dispositivo Conversacional” atuará na interação verbal.

Embora a maioria dos autores escreva sobre os empréstimos linguísticos de forma geral, percebe-se, a partir da leitura de seus estudos, que o empréstimo linguístico mais frequente advém da língua inglesa, os anglicismos. Por isso, nesta seção, limitamo-nos a arrolar somente os empréstimos advindos da língua inglesa presentes nos estudos anteriormente mencionados. Na sequência podemos ver o total de dados coletados por cada autor estudado e a fonte de onde as informações foram extraídas:

Quadro 5: Resumo dos dados coletados pelos autores estudados.

<b>AUTOR (A)</b>	<b>Total de dados coletados</b>	<b>FONTE DOS DADOS</b>
------------------	---------------------------------	------------------------

Carvalho (1989)	37	Folha de São Paulo e Diário de Pernambuco.
Manzollilo (2000)	20	Periódicos da década de 1990.
Arraes (2006)	2.106	Dicionário Aurélio Eletrônico Século XXI e Dicionário Eletrônico Houaiss da língua portuguesa.
Leitão (2006)	1.724	Dicionário Eletrônico Aurélio 1.0 (1999) e Dicionário Eletrônico Houaiss 1.0 (2001).
Alves (2007)	33	Periódicos da década de 1980.
Torrano (2010)	769.435	12 edições da Revista Info Imprensa (2008) e um CD-Rom com 24 edições digitalizadas (2007 e 2008).
Correia e Almeida (2012)	25	Dicionários Aurélio e Houaiss.
Pereira (2012)	17	Estabelecimentos comerciais e dicionários.
Ferreira e Gomes (2015)	11	Programa televisivo <i>Manhattan Connection</i> , exibido pela Globo News.
Souza (2015)	6.583	Coletânea de palavras do Projeto AC/DC: corpo <i>Corpus</i> Brasileiro .
Souza <i>et al.</i> (2015)	46	Manuais de aparelhos eletrônicos, marcas de jogos, marcas alimentícias, nomes de estabelecimentos comerciais e músicas.

Prado (2015)	1.308	Nomes de estabelecimentos comerciais do interior de São Paulo (PB) e nomes de estabelecimentos comerciais na capital portuguesa, Lisboa (PE).
Ferreira (2016)	2.264	Dicionário Eletrônico Aurélio 1.0 (1999), 3.0 (2004) e 8.0 (2010).
Castro (2017)	10	<i>Facebook</i> , jornais digitais e <i>blogs</i> .

Fonte: elaborado pela autora (2021)

Acima, vemos muitos autores que de alguma forma estudaram sobre os anglicismos. A maioria utilizou como fonte para suas pesquisas os dicionários Aurélio e Houaiss ou jornais e revistas, mas também vimos alguns autores que foram em busca de outras fontes como Prado (2015) que foi em busca de anglicismos em nomes de comércios brasileiros e portugueses; Castro (2017) que pesquisou na rede social *Facebook*, em jornais digitais e *blogs*; Souza (2015) que utilizou o banco de dados já constituído Projeto AC/DC e Souza *et al.* (2015) que basearam sua pesquisa em manuais de aparelhos eletrônicos, marcas de jogos, marcas alimentícias, nomes de estabelecimentos comerciais e músicas. Nem todos pesquisaram sob a ótica da sociolinguística como Carvalho (1989) e Manzolillo (2000). Destacamos Leitão (2006), que fez um estudo lexicográfico, embasando-se em autores com Biderman (1978), Carvalho (1989) e Alves (1990); Arraes (2006), que concentrou-se na análise da estrutura morfolexical dos anglicismos, embasou seu estudo em modelos anteriores como Halle (1973), Jackendoff (1975), Aronoff (1976) e Basílio (1980); Prado (2015) focou seus estudos nas adaptações de nível fonético e fonológico, além das estruturas morfossintáticas, a partir de estudos como os de Paiva (1991), Garces e Zilles (2004) e Trask (2004); Ferreira (2016), fez um estudo lexicográfico verificando o acréscimo de estrangeirismos de base inglesa em três versões do dicionário Aurélio, para isto, utilizou estudos como base teórica Sapir (1961) e o estudos de Biderman (1987), Vilela (1994), Faraco (2006) e outros; por fim, salientamos o estudo de Torrano (2010) que em seu estudo lexicográfico analisou os processos de criação de neologismos na área da informática, para tanto, utilizou os estudos de autores como Guilbert (1975), Barbosa (1981), Câmara Jr. (1985) e Alves (1990) .

Na sequência, vemos o Quadro 6, com os fatores linguísticos e sociais das pesquisas sobre anglicismos utilizadas neste trabalho:

Quadro 6: Fatores sociais e linguísticos das pesquisas utilizadas neste trabalho.

<b>AUTOR (A)</b>	<b>fatores LINGUÍSTICOS</b>	<b>fatores SOCIAIS</b>
Carvalho (1989)	fonológicos (adaptação dos sons), morfológicos (adaptação gráfica, decalque, incorporação de empréstimos na forma original e derivação) e semânticos (modificação dos significados dos vocábulos).	Data e local das publicações dos periódicos.
Manzollillo (2000)	morfológicos (processos de composição, derivação e palavra-valise) e semânticos (monossemia e polissemia).	Data e local das publicações dos periódicos.
Arraes (2006)	morfológicos (processos de derivação e pós-derivação), interpretação semântica e morfossintaxe.	Ano da edição dos dicionários e formato (impresso ou digital)..
Leitão (2006)	morfológicos (truncação, palavra-valise, reduplicação, derivação e neologismo)	Ano da edição dos dicionários e formato (impresso ou digital)..
Alves (2007)	morfológicos (neologismos formados por derivação prefixal e sufixal, truncação, palavra-valise, reduplicação e derivação regressiva), sintáticos (mudança	Data e local em que os periódicos foram publicados.

	de função) e fonológicos (neologismos fonológicos).	
Torrano (2010)	morfológicos (composição) e neologismos fonológicos, sintáticos e semânticos.	Ano das publicações das revistas e formatos (impresso ou digital).
Correia e Almeida (2012)	morfológicos e morfossintáticos (composição, derivação, recomposição, amálgamas, truncação, abreviação vocabular, siglas e acrônimos).	Data da edição dos dicionários e formatos (impresso ou digital).
Pereira (2012)	morfológicos, morfossintáticos, fonológicos e semânticos (adaptação ortográfica e mudança da pronúncia, troca de significado).	Datas das edições dos dicionários e local dos comércios.
Ferreira e Gomes (2015)	morfológicos (classes morfológicas dos empréstimos analisados, mudança de função, alterações ortográficas) e estilísticos (metáfora).	Público-alvo do programa, tipo de transmissão do canal (TV fechada/paga), assuntos abordados e data dos programas nos episódios analisados.
Souza (2015)	semânticos (sentido e significado), morfológicos e morfossintáticos (derivação prefixal e sufixal, composição coordenativa e subordinativa, acrônimos, siglas e aglutinações).	Data das coletas dos vocábulos e formato do corpus (digital).

Souza <i>et al.</i> (2015)	morfológicos (incorporação de estrangeirismos ao léxico do PB).	Data e local em que foram coletados os dados.
Prado (2015)	morfológicos (onomástica comercial - estudo dos nomes de estabelecimentos comerciais).	Local dos estabelecimentos comerciais e data em que os dados foram coletados.
Ferreira (2016)	morfológicos (incorporação de anglicismos ao léxico do PB).	Data de edição dos dicionários e formato (eletrônica).
Castro (2017)	morfológicos, morfossintáticos e semânticos (inovações lexicais, amálgama, significado, derivação sufixal, derivação, derivação regressiva, tipo de unidade, temática e domínio de referência).	Blogs, <i>sites</i> de notícias e redes sociais em que os dados foram obtidos e as datas da coleta.

Fonte: elaborado pela autora (2021).

Após esta breve diferenciação entre os conceitos sobre empréstimo linguístico, estrangeirismo e neologismo, pode-se perceber que estes fenômenos linguísticos ao mesmo tempo em que se diferenciam em certos aspectos, se entrecruzam dentro da língua.

## 4. OBJETIVOS, QUESTÕES E HIPÓTESES

### 4.1 Objetivo geral

- Descrever e analisar os anglicismos empregados em uma amostra sincrônica de *tweets* da rede social *Twitter*, escritos por usuários localizados na região oeste catarinense.

### 4.2 Objetivos específicos

- Verificar a atuação de variáveis linguísticas e extralinguísticas no uso dos anglicismos presentes na amostra de *tweets*; e
- Identificar propriedades formais e funcionais caracterizadoras dos anglicismos presentes na amostra de *tweets*.

### 4.3 Questões e hipóteses

#### Questão 1

Quais variáveis linguísticas e extralinguísticas atuam no uso dos anglicismos presentes na amostra de *tweets*?

#### Hipótese 1

Labov (2008 [1972]) indica que certas variáveis de natureza linguística e extralinguística podem condicionar e ser responsáveis pela implementação de uma variante e que outras, ao contrário, não demonstram qualquer efetividade na aplicação de uma regra variável.

Dividimos os condicionantes linguísticos que podem atuar no uso dos anglicismos em dois tipos: formais e funcionais. Em termos formais, a maior parte dos anglicismos são nomes (substantivos e adjetivos) ou verbos, que, por consequência, assimilam os recursos do sistema morfológico da língua - por exemplo, a ordem constituinte do enunciado, a concordância de gênero e número. Segundo Bagno (2012), podem desempenhar diferentes funções sintáticas na sentença. Em termos funcionais, muitos desses vocábulos originários do inglês como verbos, por

exemplo, mantêm ou mudam de categoria gramatical, alteram ou podem estar sendo usados com sentido diverso da língua de origem.

Dentre os condicionantes extralinguísticos, os fatores sexo, idade, localidade e ocupação podem favorecer ou desfavorecer o uso dos anglicismos. O sexo é um fator extralinguístico que pode atuar no uso da língua, a depender do fenômeno linguístico investigado. Segundo Labov (2008 [1972], p. 353), “as mulheres adotam as formas inovadoras mais rapidamente que os homens”. Por essa razão, nossa hipótese é que, apesar de os dados estatísticos levantados pelo *Twitter* (2018) revelarem que a maioria dos tuiteiros brasileiros é do perfil masculino (53%), os anglicismos assimilados seriam mais utilizados pelas mulheres nos tuítes. Outro fator a ser controlado é a faixa etária. Faraco (2005, p. 186) lembra que “[...] a predominância de uma variante entre os mais jovens e sua pouca ocorrência entre os mais velhos pode estar indicando uma mudança em progresso, isto é, que uma das variantes está sendo abandonada em favor de outra.” Por essa razão, nossa hipótese é que os usuários da faixa etária mais jovem da rede social utilizem mais anglicismos não assimilados dos que os usuários mais velhos. Em relação à localidade, não temos uma hipótese razoável para o comportamento dos usuários quanto às microrregiões e às ocupações em relação aos anglicismos. Quanto à escolaridade e à ocupação, acreditamos que os usuários com maior grau de escolaridade e com formação em nível superior tendem a empregar mais anglicismos do que os aqueles com menor escolaridade.

## **Questão 2**

Quais anglicismos emergiram em uma amostra sincrônica de *tweets* da rede social *Twitter*?

## **Hipótese 2**

É sabido que as línguas se enriquecem pelo mútuo contato. Portanto, com o confinamento decretado pelo governador do estado de Santa Catarina a partir da declaração da pandemia da Covid-19 (Sars-Cov-2), os catarinenses passaram a usar mais as redes sociais e os meios eletrônicos, se fazendo necessário o uso de ferramentas virtuais que, de modo geral, são nomeadas em língua inglesa. Diante da perspectiva do inglês como *língua franca*<sup>39</sup> em múltiplos contextos, seja nos nomes das lojas, nas conversas cotidianas presenciais ou virtuais, no ramo da

---

<sup>39</sup> Expressão usada na SOCIOLINGUÍSTICA e muitas vezes no vocabulário cotidiano, indicando uma LÍNGUA auxiliar usada para permitir que grupos de pessoas que falam línguas nativas diferentes possam estabelecer uma comunicação de rotina. O inglês é a língua franca mais famosa do mundo. (CRYSTAL, 2000 [1985]. p. 160. Grifos do autor).

moda e no ramo jornalístico, além da crescente dominação da cultura americana na língua portuguesa, conforme Guerra (2005, p. 435), pretendíamos encontrar, nos *tweets* dos usuários localizados no oeste catarinense, os anglicismos relacionados ao campo das Tecnologias da Informação e Comunicação (TICs), por exemplo, *bug*, *chat*, *like*, *meet*, *stalk*, *troll* e suas derivações como trollar, bugar, stalkear, entre outros. Acreditamos que novos anglicismos sejam incorporados à amostra principalmente porque os espaços de trabalho e de estudo tradicionais foram obrigados a migrar para o campo virtual. Também acreditamos, com base em Bybee (2020), que, além de objetos novos na cultura, alguns empréstimos não têm motivação linguística de necessidade, tendo em vista que as palavras são importadas mesmo quando a língua receptora já tem uma palavra para o objeto ou o conceito. Portanto, nossa hipótese é de que os anglicismos sejam preferidos aos termos vernáculos disponíveis.

### **Questão 3**

Quais as propriedades formais e funcionais dos anglicismos presentes na amostra de *tweets* do oeste catarinense?

### **Hipótese 3**

Na descrição e definição de fenômenos linguísticos, critérios formais (centrados na forma morfológica ou sintática dos elementos linguísticos), semânticos (centrados na significação) ou funcionais (centrados na associação entre formas e significados) são empregados (GÖRSKI; ROST, 2008, p. 60).

Em termos formais, alguns anglicismos, como *on-line*, *link* e *site*, já estão registrados nos dicionários Aurélio e Houaiss da língua portuguesa. Esses anglicismos dicionarizados encontram-se registrados na forma original (sem alteração gráfica), em forma de decalque (tradução literal de um empréstimo para morfemas equivalentes da língua receptora) ou aportuguesados (com adaptação fonológica e de grafia). Outros anglicismos, como *like*, *troll* e *meet*, porque não estão registrados em dicionários da língua portuguesa, são também pouco conhecidos e, por essa razão, em termos funcionais, necessitam de tradução e/ou explicação. Os estrangeirismos lexicais podem ser divididos em dois grupos: os que se assimilam de tal maneira à língua que os recebe, que só são identificados como empréstimos pelas pessoas que lhes conhecem a história, e os que facilmente mostram não ser assimilados e se apresentam na vestimenta estrangeira (BECHARA, 2009, p. 500). Bybee (2020) destaca que um verdadeiro empréstimo - em oposição ao uso esporádico de uma palavra estrangeira - se integra na

fonologia, morfologia e sintaxe da língua. No PB, Bagno (2012) argumenta que qualquer enunciado originado por palavras estrangeiras vai se constituir de acordo com as regras morfossintáticas da língua, o que torna impossível a compreensão por parte de um falante nativo da língua da qual o vocábulo foi emprestado.

Em razão de nossa amostra se constituir de *tweets* de usuários de uma rede social da Internet, nossa hipótese é de que os anglicismos não assimilados ao português serão mais recorrentes entre os usuários da amostra do que os assimilados. Bybee (2020, p. 330) afirma que “É bastante fácil ver que as palavras que designam alimentos viajam junto com os alimentos. Um roteiro semelhante vale para outros objetos cujo progresso mundo afora pode ser traçado de maneira parecida”.

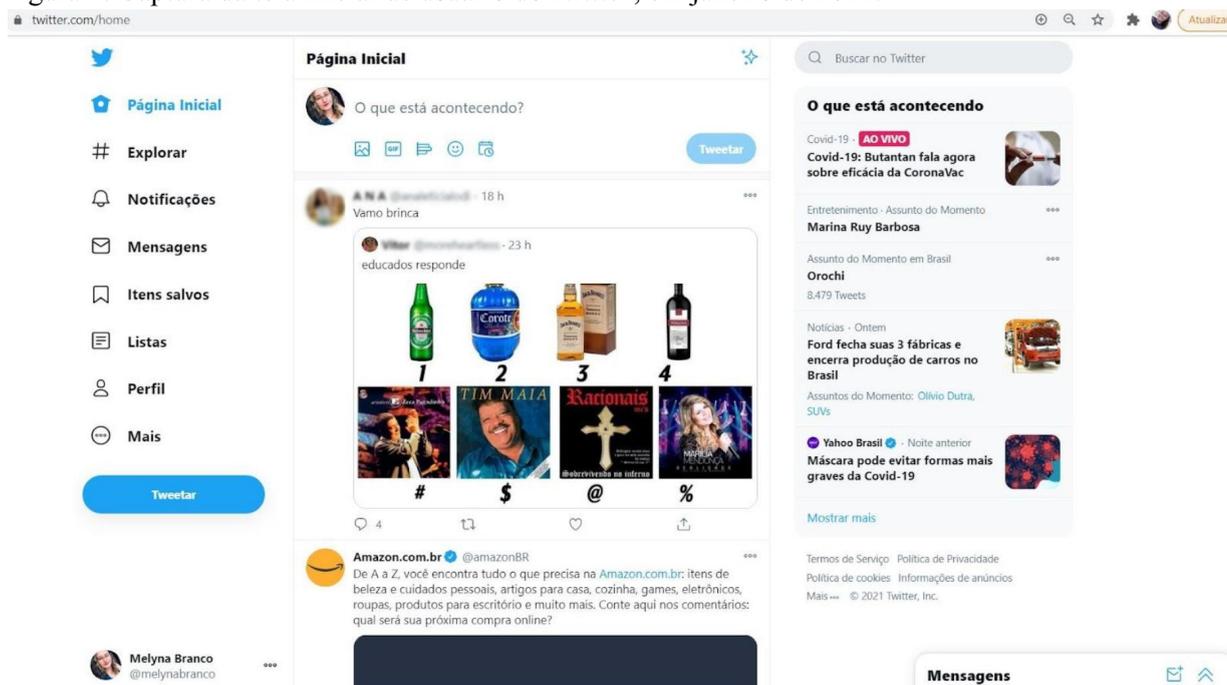
## 5. METODOLOGIA

Nesta seção, apresentamos os procedimentos metodológicos deste estudo. Descrevemos cada uma das etapas da pesquisa. A análise qualitativa partiu, inicialmente, da caracterização da rede social *Twitter* de onde extraímos os dados que compõem a nossa amostra. Procuramos ainda evidenciar, resumidamente, a sócio-história da mesorregião oeste de Santa Catarina que selecionamos como a área territorial de onde selecionamos os *tweets*. Em seguida, descrevemos a amostra e os procedimentos de coleta dos dados e especificamos as variáveis linguísticas e extralinguísticas testadas e as restrições identificadas. A análise quantitativa dos dados foi efetuada por meio do Programa R e amparada nos pressupostos teóricos da Sociolinguística Variacionista e também nos autores que pesquisaram os anglicismos no PB, conforme levantamento efetuado na subseção 3.3.1.

### 5.1 A rede social *Twitter*

Com o advento da internet e sua popularização, as redes sociais vêm se destacando no cotidiano da sociedade. O *Twitter* foi criado por Jack Dorsey em 2006 nos EUA, obtendo rápido crescimento no mundo e no Brasil, conforme Recuero e Zago (2009). De acordo com Barton e Lee (2015), publicar mensagens curtas regularmente, ou microblogar - *twittar* - transformou-se gradualmente em uma prática social comum. É importante ressaltar que “em abril de 2010, mais de 100 milhões de contas de usuários haviam sido criadas no Twitter (*The Economic Times*, 2010), atraindo pessoas de todo o mundo.” (BARTON; LEE, 2015, p. 75). Na Figura 1, a seguir, vemos a página inicial da rede social na Internet:

Figura 1: Captura da tela inicial de usuário do *Twitter*, em janeiro de 2021.



Fonte: *Twitter* (2021, não paginado)

Nesta rede social, os usuários devem responder basicamente à pergunta “O que está acontecendo?”. O *Twitter* permite também que o usuário crie um perfil público para que ele possa interagir com outras pessoas através das mensagens publicadas, além de mostrar sua rede de contatos. No *Twitter*, cada usuário pode determinar quem deseja seguir (a categoria *following* indica aqueles usuários de quem receberá as atualizações) e também poderá selecionar por quais usuários deseja ser seguido (os *followers*, em inglês, os seguidores, em português, são aqueles que vão receber as suas atualizações). Por definição, as atualizações são públicas, mas os usuários podem configurar suas contas para que elas se tornem de acesso privado, uma vez que é preciso aprovar ou rejeitar os seguidores (RECUERO; ZAGO, 2010).

Os propósitos comunicativos são muito diversificados no *Twitter* e variam conforme o perfil do usuário da rede social. “Assim, um usuário comum pode tratar de uma gama ilimitada de assuntos. Porém, um usuário que representa uma entidade ou um usuário que representa uma empresa possui uma limitação temática.” (FREITAS; BARTH, 2015, p. 23). Barton e Lee (2015, p. 75) mostram um importante estudo de Honeycutt e Herring (2009) de uma época em que *Twitter* era em grande parte em inglês:

Seu estudo coletou amostras de tweets postados em quatro períodos de tempo de um dia, cobrindo os principais fusos horários. Apesar de mais de treze línguas (incluindo japonês e espanhol) terem sido registradas nesses quatro períodos, o inglês ainda era a língua dominante em todos eles (variando de mais de 35% a 68%).

Os autores destacam ainda que, apesar da importância de o estudo de Honeycutt e Herring (2009) ter mostrado dados importantes sobre as línguas usadas nas publicações no *Twitter*, não é incluída a identidade de quem publica os microtextos, fato fundamental para a sociolinguística.

A palavra *Twitter* já está dicionarizada em língua portuguesa, sendo categorizada nos dicionários Houaiss e Michaelis como um substantivo masculino e é definida como uma rede social do tipo *microblog* projetada para o envio e a recepção em tempo real pela internet de mensagens instantâneas, os *tweets*, com até 280 caracteres. O dicionário Houaiss *On-line* descreve ainda a etimologia da palavra inglesa *Twitter*, que, em português, significa a ação de emitir gorjeios, o canto de um pássaro, ou ainda o ato de piar.

Freitas e Barth (2015) caracterizam o “*Twitter* como um suporte material – mesmo que virtual – que fixa os textos, os *tweets*”. Porém a rede social não fixa qualquer texto, tendo em vista que os enunciados fazem parte de toda uma configuração específica como a situação comunicativa. “Assim, mesmo que exista uma variação muito grande do estilo e dos temas dos *tweets*, podemos considerá-los, sim, como um gênero híbrido e não textos de diferentes gêneros sendo vinculados pelo mesmo suporte.” (FREITAS; BARTH, 2015, p. 24).

Conforme Silva e Valls (2012), o *Twitter* possui uma linguagem própria, com termos e expressões característicos, além de também serem comuns contrações e abreviações que só fazem sentido neste ambiente. A seguir, listamos alguns termos característicos do ambiente da rede social e conceituados pelo próprio *Twitter* (2020):

- *Tweet*: são mensagens publicadas no *Twitter* que contêm texto, fotos, um GIF<sup>40</sup> e/ou um vídeo. Os *tweets* aparecem para o remetente em sua *timeline*<sup>41</sup> e em sua página do perfil. Já para os destinatários, os *tweets* aparecem na *timeline* de qualquer usuário que siga o remetente.

Freitas e Barth (2015) observaram que o perfil dos usuários determina as diferenças em relação ao conteúdo temático e ao estilo dos *tweets*. Porém, mesmo que seja observada variação no estilo e nos temas dos *tweets*, estes se constituem como um gênero híbrido e não textos de diferentes gêneros sendo vinculados pelo mesmo suporte.

---

<sup>40</sup> *Graphics Interchange Format* ou formato de intercâmbio de gráficos é um formato de imagem muito usado na Internet e que foi lançado em 1987 pela CompuServe, para disponibilizar um formato de imagem com cores. (BRITO, 2012. não paginado).

<sup>41</sup> Linha do tempo, tradução nossa.

- *Retweet*: é uma nova postagem de um *tweet*. Através do recurso *retweet*, o usuário pode compartilhar conteúdo rapidamente com seus seguidores. É possível retuitar os próprios *tweets* ou as publicações de outra pessoa. Alguns usuários usam o termo “RT” no início de um *tweet* para indicar que estão publicando novamente o conteúdo de outra pessoa.
- *Timeline*: é a linha do tempo ou a página inicial do usuário que exhibe o fluxo de *tweets* de contas que o usuário segue. Há também a possibilidade de ver conteúdos sugeridos pela própria rede social.
- *Hashtag*: é a escrita de uma palavra antecedida pelo símbolo “#”. Através desta função, o usuário pode indexar palavras-chave ou tópicos no *Twitter*. Esta função foi criada no *Twitter* e permite que as pessoas sigam facilmente os tópicos de seu interesse.
- Arroba: é o símbolo “@” que antecede o nome do usuário, também conhecido como identificador e é exclusivo da conta de cada usuário e aparece na URL<sup>42</sup> do perfil. O nome de usuário é utilizado para acessar o *Twitter* e também para procurar o perfil de outro usuário.
- *Trending topics*: são os assuntos do momento e podem ser encontrados através da aba explorar. Os *trending topics* são determinados por um algoritmo e, por padrão, são personalizados com base no que o usuário segue, nos interesses e na localização do tuiteiro. Esse algoritmo tem a função de identificar os tópicos mais populares da atualidade, possibilitando ao usuário que ele descubra os assuntos que estão em discussão no *Twitter* em tempo real.

Conforme os termos de serviço do *Twitter* (2020), somente pessoas com 13 anos ou mais podem ter uma conta de usuário desta rede social. Os termos alertam ainda que o usuário é responsável pelo uso dos serviços do *Twitter* e por todo o conteúdo fornecido, incluindo a conformidade com leis, regras e regulamentos aplicáveis e sugere ao usuário que publique apenas conteúdos com os quais ele se sinta confortável em compartilhar com outras pessoas. Os termos de serviço da rede reiteram ainda que o *Twitter* não aprova, apoia, declara nem garante a integridade, veracidade, exatidão ou confiabilidade de qualquer conteúdo ou comunicado publicado pelos serviços, tampouco endossa opiniões expressas pelos usuários da rede. Conforme o acordo do *Twitter* (2020), o usuário deve compreender que, ao utilizar o *Twitter*,

---

<sup>42</sup> *Uniform resource locator* ou localizador-padrão de recursos é a forma padronizada de representação de diferentes documentos, mídia e serviços de rede na internet, capaz de fornecer a cada documento um endereço único. (Dicionário Houaiss *On-line*).

pode estar exposto a conteúdo ofensivo, prejudicial, impreciso, inadequado ou, em alguns casos, a publicações identificadas de modo indevido ou fraudulentas.

Por fim, a rede social reitera que todo o conteúdo é de responsabilidade exclusiva da pessoa que o criar e que o *Twitter* se reserva ao direito de remover o conteúdo que viole o acordo do usuário, incluindo, por exemplo, violações de direitos autorais ou de marcas comerciais ou outra apropriação indébita de propriedade intelectual, falsidade ideológica, conduta ilegal ou assédio. Aqui é importante mencionar a disseminação das *fake news* no *Twitter* e os casos em que famosos, políticos e outras pessoas públicas têm seus *tweets* apagados por violarem as regras da rede. O caso que mais chamou atenção foi o do ex-presidente dos EUA, Donald Trump, que teve seu perfil excluído permanentemente em janeiro de 2021 pelo próprio *Twitter* por incitar atos violentos.

## 5.2 A sócio-história do oeste catarinense

A região oeste catarinense, segundo Radin, Benet e Milani (2003), foi incorporada ao estado de Santa Catarina pelo acordo de limites com o Paraná, em 1916, com aproximadamente 28.000 km<sup>2</sup>.

Em sua sócio-história, a região se constitui a partir de uma diversidade étnica e cultural que nos faz entender os costumes, as tradições e as línguas dos habitantes inseridos nestas localidades. Os povos naturais - autóctones -, como os Guaranis e *Kaingangs*, no decorrer do tempo, uniram-se aos povos estrangeiros - alóctones -, como descendentes de imigrantes italianos, alemães e poloneses, que passaram a constituir a região, de acordo com Bavaresco (2005), a partir da década de 1920, depois de o tropeirismo ter aberto os primeiros caminhos e pontos de parada e também após a resolução dos limites com o Paraná, com o fim da Guerra do Contestado. É importante ressaltar também que “bem antes da colonização, algumas famílias de caboclos já habitavam a região.” (BAVARESCO, 2005, p. 16). Essas minorias étnicas utilizam em diversos contextos uma ou mais variedades linguísticas que não correspondem a uma variedade do PB. Um exemplo destas situações eram as poucas escolas da região Oeste no início do século XX e, segundo Bavaresco (2005), os professores ministravam as aulas em alemão ou italiano, mais uma forma de excluir os caboclos e índios da vida social. Vale lembrar que foi somente em 1937 que o governo federal desenvolveu uma campanha de nacionalização quando começaram a surgir escolas mantidas pelo estado.

De acordo com Argenta (2014), a região que hoje chamamos de Oeste catarinense era conhecida como ‘Xaçepó’. Ainda de acordo com o autor, as terras eram a maioria em regiões montanhosas e isoladas. Já Onghero (2014) destaca que a atuação das companhias colonizadoras na região teve início a partir da definição dos limites entre o Paraná e Santa Catarina em 1916 e a criação do município de Chapecó em 1917. O autor afirma ainda que as terras desta região permaneceram abandonadas até meados de 1940.

Conforme Radin, Benedet e Milani (2003), a participação dos italianos na colonização do oeste foi bastante expressiva, o que gerou conflitos étnicos, pois, na região, predominavam as famílias de luso-brasileiros e caboclos. Os descendentes de italianos migraram, em sua maioria, das Colônias Velhas do estado vizinho, Rio Grande do Sul, em busca de melhores condições de vida e também pelo fato de que o solo sul-rio-grandense já estava em aparente esgotamento. Vale lembrar ainda que os migrantes colonizadores do Oeste catarinense estavam entre os elementos desejados para a ocupação dos espaços vazios do território brasileiro. Deste modo, os colonos que compravam as terras na região Oeste catarinense queriam que elas estivessem limpas, isto é, sem moradores. Bavaresco (2005, p. 115) rememora esta história:

Quando as concessões de terras aumentaram na região, os posseiros eram expulsos das terras pelas colonizadoras, pois, para poderem comercializar os lotes, faziam o que chamavam de “limpeza da área”. Nesses casos, o caboclo era visto como intruso. Para isso, valiam-se de todas as artimanhas possíveis.

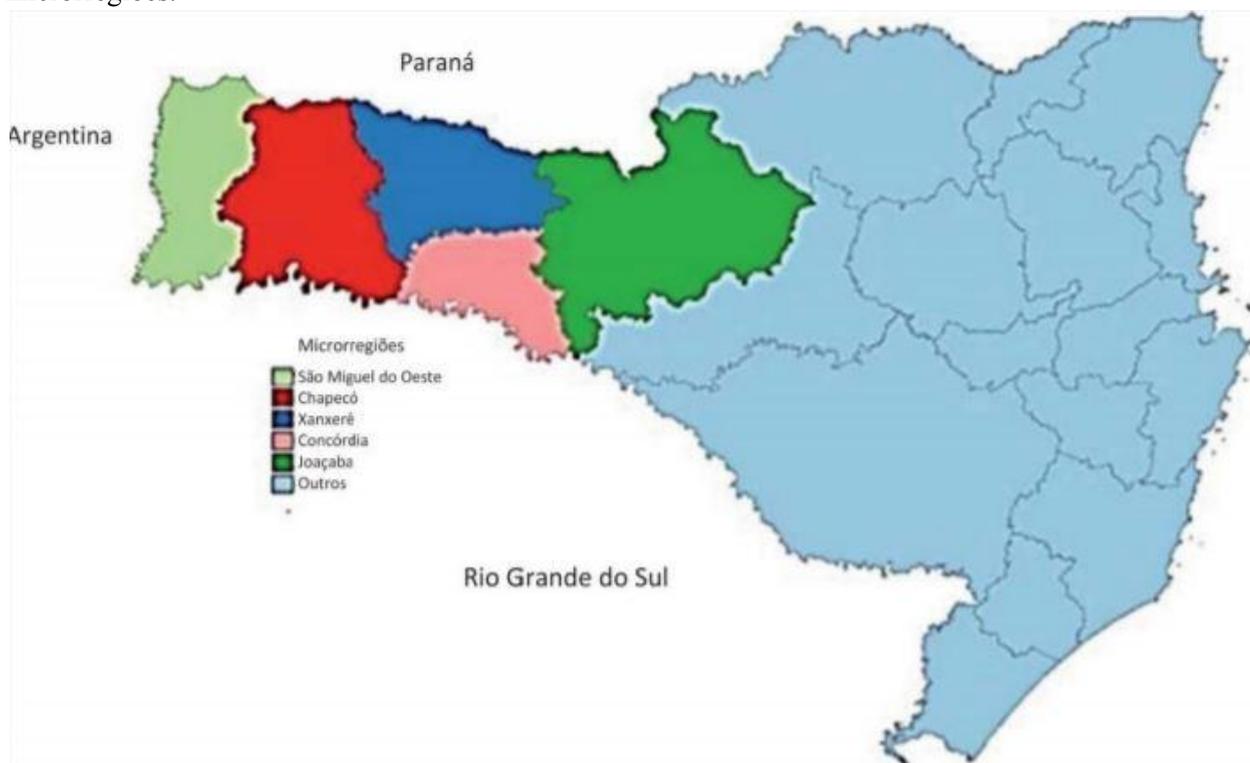
Ainda conforme Bavaresco (2005), não havia interesse dos colonizadores quanto à integração dos caboclos nos aspectos culturais e educacionais. De acordo com o autor, além da segregação econômica, havia um enorme isolamento cultural, pois, quando havia escola, as aulas eram ministradas em italiano ou alemão. A prática era permitida pelo governo do estado catarinense nas zonas de colonização, onde se podia dar aulas em qualquer idioma.

Acerca da imigração dos alemães para o Brasil, primeiramente, de acordo com Werlang (1999), deu-se devido ao atraso em que se encontrava a Alemanha em relação à Revolução Industrial, entretanto, em uma segunda etapa, foi devido ao excesso populacional e à crise econômica naquele país. Esse período coincidiu com o estímulo brasileiro à imigração, com a intenção de substituir a mão de obra escrava. No Sul do Brasil, os colonos alemães se instalaram, primeiramente, no Rio Grande do Sul. Werlang (1999) sustenta que, entre 1830 e 1844, o governo alemão proibiu a emigração para o Brasil e, após a liberação, a procura pelo país foi bastante considerável. Ainda de acordo com o autor, a principal causa da migração para Santa Catarina foi a falta de terras e também a infertilidade do solo gaúcho.

Saltando dos primórdios para a atualidade, é importante ressaltar que a região do Oeste catarinense vem recebendo um número elevado de imigrantes haitianos e venezuelanos. Porém, de acordo com Fornara (2016), como a imigração desse povo é recente e ainda em processo de locomoibilidade, são ainda poucas as pesquisas linguísticas<sup>43</sup> sobre a influência destes novos idiomas na variedade do PB falada na região do Oeste de Santa Catarina.

Na mesorregião do Oeste estão localizados 118 municípios (FISHER; LÜCKMANN, 2020, p. 31), os quais são divididos em cinco microrregiões: Chapecó (37 municípios), Concórdia (15 municípios), Joaçaba (27 municípios), São Miguel do Oeste (22 municípios) e Xanxerê (17 municípios). Estes municípios representam 40% das cidades catarinenses. Ainda conforme Fisher e Lückmann (2020), a população da mesorregião foi estimada, em 2019, em 1.403.281 habitantes, os quais se dividem da seguinte forma: 449.161 na microrregião de Chapecó, 352.143 na microrregião de Joaçaba, 181.753 na microrregião de São Miguel do Oeste, 163.708 na microrregião de Xanxerê e 147.463 na microrregião de Concórdia. A Figura 2 apresenta as cinco microrregiões da mesorregião Oeste:

Figura 2: Mapa de Santa Catarina, com a mesorregião oeste destacada e dividida em suas cinco microrregiões.



Fonte: Fisher e Lückmann (2020, p. 33).

<sup>43</sup> Pesquisas sobre imigrantes da atualidade na região Oeste: Costa (2016); Soares (2016); Pinheiro, (2018); Ribeiro (2018); Borsati (2019); Mattos (2021); Zamaro (2021).

A microrregião de Chapecó é formada por 37 municípios, são eles: Águas de Chapecó, Águas Frias, Bom Jesus do Oeste, Caibi, Campo Erê, Caxambu do Sul, Chapecó, Cordilheira Alta, Coronel Freitas, Cunha Porã, Cunhataí, Flor do Sertão, Formosa do Sul, Guatambu, Iraceminha, Irati, Jardinópolis, Maravilha, Modelo, Nova Erechim, Nova Itaberaba, Novo Horizonte, Palmitos, Pinhalzinho, Planalto Alegre, Quilombo, Saltinho, Santa Terezinha do Progresso, Santiago do Sul, São Bernardino, São Carlos, São Lourenço do Oeste, Saudades, Serra Alta, Sul Brasil, Tigrinhos e União do Oeste. Já a microrregião de Concórdia é constituída por 15 municípios: Alto Bela Vista, Arabutã, Arvoredo, Concórdia, Ipira, Ipumirim, Irani, Itá, Lindóia do Sul, Paial, Peritiba, Piratuba, Presidente Castello Branco, Seara e Xavantina. A microrregião de Joaçaba, por sua vez, é constituída por 27 municípios: Água Doce, Arroio Trinta, Caçador, Calmon, Capinzal, Catanduvas, Erval Velho, Fraiburgo, Herval d'Oeste, Ibiam, Ibicaré, Iomerê, Jaborá, Joaçaba, Lacerdópolis, Lebon Régis, Luzerna, Macieira, Matos Costa, Ouro, Pinheiro Preto, Rio das Antas, Salto Veloso, Tangará, Treze Tílias, Vargem Bonita e Videira. Localizada no extremo Oeste, a microrregião de São Miguel do Oeste é composta por 22 municípios: Anchieta, Bandeirante, Barra Bonita, Belmonte, Descanso, Dionísio Cerqueira, Guaraciaba, Guarujá do Sul, Iporã do Oeste, Itapiranga, Mondai, Palma Sola, Paraíso, Princesa, Riqueza, Romelândia, Santa Helena, São João do Oeste, São José do Cedro, São Miguel da Boa Vista, São Miguel do Oeste e Tunápolis. Por fim, a microrregião de Xanxerê é formada por 17 municípios: Abelardo Luz, Bom Jesus, Coronel Martins, Entre Rios, Faxinal dos Guedes, Galvão, Ipuacu, Jupiá, Lajeado Grande, Marema, Ouro Verde, Passos Maia, Ponte Serrada, São Domingos, Vargeão, Xanxerê e Xaxim.

### 5.3 A amostra

A amostra foi extraída da rede social *Twitter* a partir de um conjunto de *tweets* publicados por 50 usuários comuns da rede, estratificados em sexo, faixa etária, ocupação e localidade (que indicaram a localização na mesorregião oeste de Santa Catarina). Entendemos como usuários comuns as pessoas que preenchem o perfil de estratificação definido por nós, portanto não foram considerados perfis de empresas, instituições públicas e privadas, clubes esportivos e outras corporações. As seguintes variáveis foram coletadas:

- a) faixa etária: a data de nascimento ou idade pode ser visualizada no perfil do usuário. Os usuários do *Twitter* têm, em geral, entre 16 e 54 anos, segundo *Global Web Index* (2017) e *ComScore* (2018).

- b) sexo: o sexo deve ser especificado na configuração da conta do usuário e a informação não é exibida publicamente. O usuário pode escolher entre sexo masculino ou feminino ou inserir em “outro” seu gênero. O sexo dos usuários será identificado pelo nome do usuário.
- c) ocupação: a ocupação é colocada de forma voluntária na descrição de cada perfil. Quando não encontramos a ocupação na descrição, buscamos pela informação, de forma manual, nos *tweets* do usuário.
- d) localidade: há duas possibilidades de encontrar a localidade do usuário: na descrição do perfil do usuário ou nas informações disponíveis no rodapé de cada *tweet*.

Na tabela a seguir, podemos visualizar os resultados de um levantamento do perfil do usuário do *Twitter* feito em 2018:

Tabela 1: Quem usa o *Twitter*?

<b>Sexo</b>	<b>Porcentagem</b>
Homens	57%
Mulheres	43%
<b>Faixa etária</b>	<b>Porcentagem</b>
16-24 anos	32%
25-34 anos	29%
35-44 anos	19%
45-54 anos	11%

Fonte: *Global Web Index*, 2017; *ComScore*, 2018.

A amostra final foi composta por 50 usuários do *Twitter*, dividida entre 29 homens e 21 mulheres, respeitando a porcentagem de quem usa o *Twitter*, segundo a *Global Web Index* (2017) e *ComScore* (2018), 57% e 43%, respectivamente. Além de separar por sexo, os usuários também

foram extratificados por faixa etária. De acordo com a Global Web Index (2017), quatro faixas etárias são as mais recorrentes entre os tuiteiros: 16 a 24 anos, que corresponde a 32% dos usuários, equivalente a 10 homens e 7 mulheres da amostra; 25 a 34 anos, equivalente a 29% dos brasileiros que utilizam o *Twitter* e 8 homens e 6 mulheres da amostra; 35 a 44 anos, representando 19% dos tuiteiros e 7 homens e 5 mulheres da coleta; e, por fim, 45 a 54 anos, correspondendo a 11% dos usuários da rede e a 4 homens e 3 mulheres da amostra.

De acordo com a última estimativa populacional do Instituto Brasileiro de Geografia e Estatística (IBGE), divulgada em 1º de junho de 2021, há, no Brasil, 213.317.639 (duzentos e treze milhões, trezentos e dezessete mil e seiscentos e trinta e nove) habitantes e, segundo a pesquisa desenvolvida pela Global Web Index (2017) e ComScore (2018), o país possui em torno de 8,6 milhões de usuários ativos no Twitter, correspondendo a 4,1% da população brasileira.

Nossa amostra também foi constituída por dois períodos de tempo: no primeiro, foram coletados *tweets* publicados pelos usuários entre março de 2019 e fevereiro de 2020 e, no segundo, considerando a necessidade de isolamento e distanciamento social em razão da pandemia da Covid-19, que fez com que muitas atividades profissionais e acadêmicas passassem para o formato remoto, foram coletados *tweets* publicados pelos usuários entre março de 2020 e fevereiro de 2021. Nossa hipótese é que, devido à pandemia da Covid-19, tenha havido incremento no uso de anglicismos por consequência do aumento do trabalho e estudo remoto e do uso dos recursos tecnológicos virtuais no período.

Outros critérios adotados para a seleção dos *tweets* foram:

- e) *tweets* públicos: somente foram considerados os *tweets* públicos feitos por usuários comuns localizados na região oeste catarinense, portanto foram descartados os *retweets* feitos pelos usuários selecionados, ou seja, compartilhamento de mensagens de terceiros.

Nosso interesse versa, neste estudo, sobre o uso de estrangeirismos, mais especificamente os anglicismos do campo das TICs, por exemplo, *bug*, *chat*, *like*, *meet*, *stalk*, *troll* e suas derivações como *trollar*, *buguei*, *stalkear* e outros. Inicialmente, os *tweets* foram coletados de forma manual, diretamente do *Twitter*. Posteriormente, contamos com a ajuda de um programador para agilizar o processo e, por fim, encontramos o *site vicinitas.io*, que possibilita o *download* automático de 3.200 *tweets* por usuário, de forma gratuita, no formato de planilha do Excel. Estas planilhas foram utilizadas para que fosse possível montar o *corpus* desta pesquisa, protegendo a identidade de cada usuário da rede social.

## 5.4 Os procedimentos de coleta de dados

Os dados foram coletados de publicações feitas - os *tweets* - no contexto informal da ciberescrita do *Twitter*, realizadas por usuários da rede que se localizam nas microrregiões de Chapecó, Concórdia, Joaçaba e Xanxerê, no oeste catarinense.

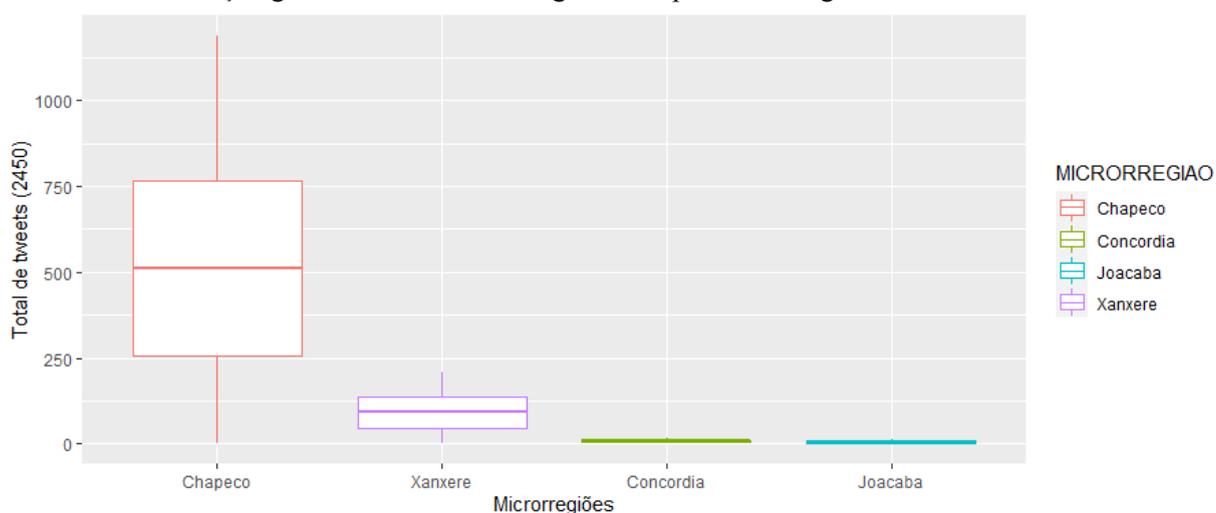
Primeiramente, selecionamos os informantes que preenchem os requisitos especificados na seção anterior. Na sequência, extraímos todos os *tweets* públicos desses usuários nos dois períodos estimados (março de 2019 a fevereiro de 2020 e março de 2020 a fevereiro de 2021). Estas informações foram coletadas a partir da ferramenta de “busca avançada” disponibilizada dentro da própria rede social. Esta busca está dividida em cinco categorias e dentro delas há subcategorias. As categorias principais estão divididas em “palavras”, “contas”, “filtros”, “engajamento” e “datas”. Para esta pesquisa foi utilizado o filtro data, que possibilita escolher a data de início e data final que interessa à busca. A localização foi filtrada a partir do comando “near:”, que delimita o local (cidade, estado ou país) em que os *tweets* foram publicados. Vejamos um exemplo na imagem a seguir:

Figura 3: Busca avançada a partir de períodos no *Twitter*.

A imagem mostra a interface de busca avançada do Twitter, especificamente a seção 'Datas'. O título 'Datas' está em negrito. Abaixo dele, há duas seções: 'De' e 'Para'. Cada seção contém três campos de seleção: 'Mês', 'Dia' e 'Ano'. No campo 'De', 'Mês' é 'Setembro', 'Dia' é '1' e 'Ano' é '2019'. No campo 'Para', 'Mês' é 'Setembro', 'Dia' é '1' e 'Ano' é '2020'. Todos os campos possuem uma seta para baixo indicando que são menus suspenso.

Fonte: *Twitter* (2021, não paginado)

Foram coletados ao todo 32.312 *tweets*, distribuídos em 14.721 (45,56%) relativos ao período de março de 2019 a fevereiro de 2020 e 17.591 (54,44%) concernentes ao período de março de 2020 a fevereiro de 2021. Após a coleta, todos os *tweets* públicos desses usuários foram lidos, mas selecionamos apenas os que continham anglicismos na sua escrita. No Gráfico 1 da sequência, podemos ver a distribuição da frequência de *tweets* com anglicismos, separados pelas microrregiões:

Gráfico 1: Distribuição geral dos *tweets* com anglicismos por microrregiões do Oeste de Santa Catarina.

Fonte: elaborado pela autora (2021).

Os resultados do Gráfico 1 revelam que, considerando a região oeste catarinense, foram levantados dados de usuários das localidades de Chapecó, Concórdia, Joaçaba e Xanxerê. No total, foram coletados 2450 *tweets*, dentre os quais a maior parte ( $2037/83 = 14\%$ ) foi extraída de Chapecó, que se constitui a cidade com maior população na microrregião, seguida das cidades de Xanxerê ( $368/15 = 2\%$ ), Concórdia ( $29/1 = 18\%$ ) e Joaçaba ( $16/0 = 65\%$ ).

### 5.5 Análise qualitativa dos dados

Após a coleta, cada item lexical foi transcrito e categorizado segundo as variáveis independentes (variáveis previsoras) linguísticas e extralinguísticas (ou sociais) controladas. O Quadro 7 resume a descrição das variáveis e de seus respectivos fatores:

Quadro 7: Resumo das variáveis linguísticas e sociais controladas.

Variáveis linguísticas e sociais	Descrição
1. Faixa Etária	Distribuída em quatro grupos: 16 a 24 anos, 25 a 34 anos, 35 a 44 anos e 45 a 54 anos.
2. Sexo	Dividido em masculino e feminino.
3. Ocupação	Distribuída a partir da descrição de cada usuário em seu perfil público no <i>Twitter</i> . As ocupações encontradas foram: administrador/a, advogado/a, agrônomo/a, atleta, barbeiro/a, comunicador/a, dentista, <i>designer</i> gráfico, empresário/a, engenheiro/a controle e automação, estudante

	ensino médio, estudante ensino superior, fisioterapeuta, funcionário/a público/a, gastrônomo/a, motorista, pastor/a, professor/a, programador/a, psicólogo/a, tecnólogo/a em processos gerenciais, vendedor/a e zootecnista. Três usuários não informaram a ocupação.
4. Microrregião	Distribuída em quatro microrregiões do Oeste: Chapecó, Concórdia, Joaçaba e Xanxerê.
5. Período publicação dos tweets	Os tweets foram separados em dois períodos: março de 2019 a fevereiro de 2020 (P1) e março de 2020 a fevereiro de 2021 (P2).
6. Item lexical	Todos os anglicismos foram isolados do contexto e categorizados.
7. Classe morfológica	Os itens lexicais foram categorizados como: adjetivo, advérbio, conjunção, interjeição, marcador discursivo, preposição, substantivo e verbo.
8. Classificação dos anglicismos	Os anglicismos foram separados com o objetivo de saber se eles foram utilizados em sua forma original ou se sofreram algum processo de adequação/ajuste à língua portuguesa.
9. Dicionarização	Todas as palavras foram pesquisadas em quatro dicionários da língua portuguesa: Aurélio, Bechara, Houaiss e Moura Neves, sendo posteriormente categorizadas em dicionarizadas e não-dicionarizadas.
10. Contexto	Um trecho mais amplo de cada tweet foi preservado para a compreensão e a classificação de cada anglicismo.

Fonte: elaborado pela autora, (2021).

Na Figura 4, apresentamos como as variáveis foram distribuídas e os dados codificados na planilha em extensão .csv (*comma separated values*) para análise no *software* R:

Figura 4: Captura de tela com a categorização dos dados do usuário 1.

PARTICIPANTE	FAIXA.ETARIA	SEXO.GENERO	OCUPACAO	MICRORREGIAO	TWEETS.TOTAL	TWEETS.AN.MEDIANA	FREQ.ANG.19.20
USER1.1M	16-24	masculino	programador/mestrando	Chapeco	747	463	31
USER1.1M	16-24	masculino	programador/mestrando	Chapeco	747	463	31
USER1.1M	16-24	masculino	programador/mestrando	Chapeco	747	463	31
USER1.1M	16-24	masculino	programador/mestrando	Chapeco	747	463	31
FREQ.ANG.20.21	DATA.TWEET	ITEM.LEXICAL	CLASSE.MORFOLOGICA	ORIGINAL/APORTUGUESADO	DICIONARIZADO	CONTEXTO	
114	D1	boxes	substantivo	original	sim	essa nova entrada dos boxes	
114	D1	grid	substantivo	original	sim	Os pilotos no final do treino pare	
114	D1	widjet	substantivo	original	nao	SpotifyAjuda pq foi removido o v	
114	D1	podcast	substantivo	original	nao	parabéns pelo último podcast sc	

Fonte: elaborada pela autora, (2021).

Esses dados foram categorizados e analisados qualitativamente, amparados nos pressupostos teóricos da Sociolinguística Variacionista e também no levantamento bibliográfico sobre anglicismos no PB, conforme a seção 2.

## 5.6 Restrições

Os 50 usuários geraram 32.312 (trinta e dois mil trezentos e doze) *tweets* (com e sem anglicismos) e, por uma questão de delimitação da pesquisa, optamos por analisar apenas a mediana<sup>44</sup> de cada faixa etária dividida pelo sexo, descritas na Tabela 1. Através da mediana foram analisados 25.104 (vinte e cinco mil cento e quatro) *tweets*, nos quais foram encontradas 2.450 (duas mil quatrocentos e cinquenta) ocorrências de anglicismos. Para que fosse possível a análise dos dados no período determinado, fizemos a mediana de cada faixa etária e sexo, para posterior análise. As medianas encontradas foram: 16 a 24 anos: 463 (M), 1119 (F); 25 a 34 anos: 285 (M), 567 (F); 35 a 44 anos: 288 (M), 484 (F); 45 a 54 anos: 1463 (M), 116 (F).

Na análise qualitativa dos *tweets*, foram excluídos todos os *tweets* identificados como “frases” e também expressões cristalizadas em inglês, por entendermos que se tratava de contexto mais amplo de uso da língua inglesa ou até de uma situação de bilinguismo. Assim, optamos por excluir 129 *tweets* no total. Abaixo, estão exemplificadas algumas das ocorrências excluídas da análise quantitativa:

(6) “[...] minha orientadora me passou um artigo "muito bom" que ela estava usando para escrever um artigo, segunda página desse artigo: "*as proposed in*"<sup>45</sup> [minha orientadora]".” (publicado em 13 de julho de 2020, em Chapecó, grifo nosso).

(7) “Ai você está fazendo umas lições no Duolingo e se depara com uma dessas: "*My horse paints beautiful pictures.*"<sup>46</sup>” (publicado em 26 de outubro de 2019, em Chapecó, grifo nosso).

(8) “Senhor trás (sic) as alianças *because I do*”<sup>47</sup> (publicado em 24 de julho de 2019, em Chapecó, grifo nosso).

<sup>44</sup> Mediana é “[...] o valor que se obtém quando todos os dados são colocados em ordem de tamanho e você escolhe aquele que divide essa distribuição ao meio.” (GRIES, 2019, p. 111).

<sup>45</sup> como proposto em, tradução nossa.

<sup>46</sup> Meu cavalo pinta lindas imagens, tradução nossa.

<sup>47</sup> porque eu aceito, tradução nossa.

## 5.7 Análise quantitativa dos dados: o *Software R*

Em nossa análise quantitativa dos dados, foram utilizados os *softwares* livres R e RStudio (R CORE TEAM, 2020)<sup>48</sup>. Trata-se de uma linguagem de programação para gráficos e cálculos estatísticos, a partir dos estudos e tutoriais de Oushiro (2014), que disserta sobre o processamento de dados com o RStudio para análises sociolinguísticas.

Esta perspectiva permite a análise quantitativa comparativa acerca do aumento no uso dos anglicismos durante o período da pandemia do novo coronavírus, a Covid-19.

Ritter, They e Konzen (2019) explicam que o R é um tipo de linguagem e um *software* computacional e gráfico. Conforme os autores, o R é um *software* de código aberto, livre e gratuito, isto quer dizer que os usuários podem executar o programa como quiserem, além de poderem adaptá-lo às necessidades, haja vista que se tem livre acesso aos códigos-fonte.

Levando em consideração que o R é uma linguagem de programação, Oushiro (2014) relembra que o programa permite ao usuário customizar inúmeras tarefas que ele julgue necessário executar e, deste modo, tenha maior controle sobre os dados obtidos. Por outro lado, o fato de o R ser um *software* livre impossibilita ao usuário clicar em botões com funções pré-definidas e limitadas, já que geralmente é o usuário que define as funções que deseja executar. Conforme a autora, uma sequência de linhas no *software* R é chamada de *script* ou código e, além disso, segundo Oushiro (2014), tais funções podem parecer complicadas no início, mas, conforme o usuário se familiariza com o ambiente do R, se tornará menos complexo. A autora afirma ainda que geralmente é necessário criar um *script* somente uma vez, já que o usuário pode salvar o código e reutilizá-lo quantas vezes forem necessárias, modificando apenas pequenas partes para adaptá-lo às novas demandas.

Para que a análise de nossos dados no R fosse possível, primeiro fizemos uma planilha no *software* do *Google* com 15 variáveis. Na sequência salvamos este arquivo no formato “valores separados por vírgula” (.csv) e o importamos para o R. Para a análise, rodamos os pacotes *tidyverse*<sup>49</sup> e *magrittr* para utilizar o operador *pipe* `%>%`<sup>50</sup>. Para as rodadas de dados e para criação dos gráficos, os principais *scripts* utilizados foram:

<sup>48</sup> Nestes sites <https://cran.r-project.org/> e <https://www.rstudio.com/products/rstudio/download/>, podem ser feitos *downloads* gratuitos do R e RStudio.

<sup>49</sup> “O *{tidyverse}* é um pacote guarda-chuva que consolida uma série de ferramentas que fazem parte do ciclo da ciência de dados. Fazem parte do *{tidyverse}* os pacotes *{ggplot2}*, *{dplyr}*, *{tidyr}*, *{purrr}*, *{readr}*, entre muitos outros”. (DAMIANI *et al.*, 2015, não paginado)

<sup>50</sup> “A ideia do operador `%>%` (*pipe*) é bem simples: usar o valor resultante da expressão do lado esquerdo como primeiro argumento da função do lado direito”. (DAMIANI *et al.*, 2015, não paginado)

- Para ver os dados do arquivo:

```
summary(nome do arquivo)
```

```
str(nome do arquivo)
```

- Para ver a frequência de determinada variável:

```
novo.nome <- nome variável %>%
```

```
count(nome coluna analisada) %>%
```

```
mutate(media = prop.table(n)) %>%
```

```
print(nome do arquivo)
```

- Para fazer um gráfico *boxplot* (diagrama de caixa):

```
ggplot(nome do arquivo, aes(x=nome variável, y= nome variável))+
```

```
geom_boxplot(aes(color= nome) +
```

```
labs(x = "nome", y = "título desejado") +
```

```
scale_x_discrete(limits = c("nome1", "nome2", "nome3"))
```

- Para testar a significância das variáveis

```
novo.nome <- glm (nome variável 1 ~ nome variável 2, data =
```

```
nome do arquivo, family = binomial)
```

```
summary (nome do arquivo)
```

É importante ressaltar que qualquer erro de digitação resulta em erro na hora de rodar os *scripts*, por isso é preciso ter muita atenção na hora de rodar os dados e se atentar para caracteres simples como uma letra maiúscula ou minúscula que pode alterar o código. Para que conseguíssemos analisar os nossos dados no R, foi preciso muitas horas de estudo sobre o *software* e muita tentativa e erro na hora de rodar os *scripts* até que obtivéssemos nossos resultados.

## 6. DESCRIÇÃO E ANÁLISE DOS DADOS

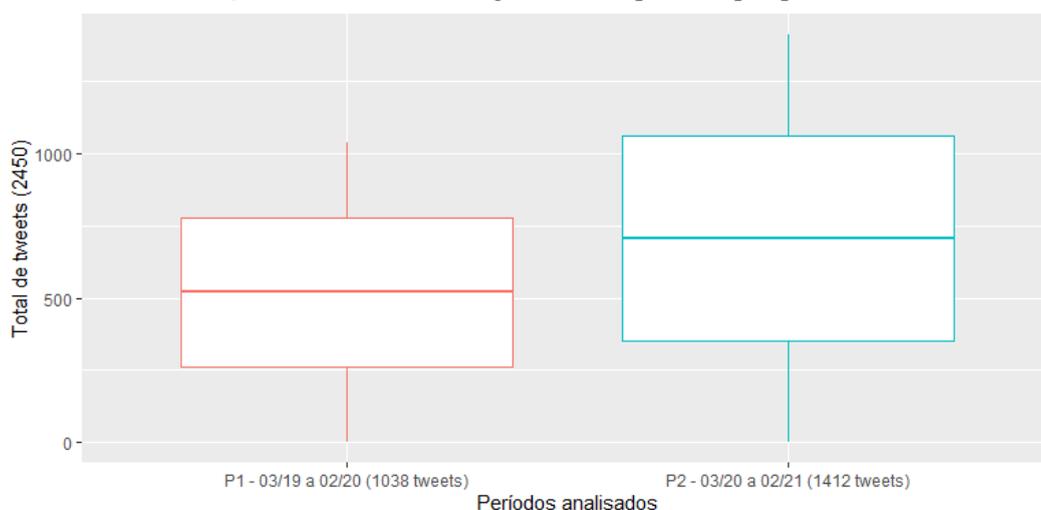
Nesta seção, realizamos a descrição e a análise dos anglicismos empregados em uma amostra sincrônica de *tweets* da rede social *Twitter* escritos por usuários localizados na região oeste catarinense. Iniciamos com a apresentação da frequência geral dos dados e, na sequência, tratamos dos resultados das variáveis linguísticas e extralinguísticas que condicionam o uso de anglicismos em *tweets* escritos pelos usuários da amostra.

### 6.1 Frequência geral

Vimos que a frequência de dados totais de *tweets* coletados foi de 32.312, dentre os quais 14.721 (45,56%) relativos ao primeiro período (março de 2019 a fevereiro de 2020) e 17.591 (54,44%) concernentes ao segundo período (março de 2020 a fevereiro de 2021). Esse incremento na frequência de ocorrência confirmou nossa hipótese inicial de que naturalmente, em decorrência da pandemia da COVID-19, haveria aumento na frequência de ocorrência dos *tweets* entre os dois períodos investigados. Porém, sabemos que o aumento da frequência é decorrente principalmente da transformação de espaço de trabalho e estudo, pois muitas pessoas tiveram de passar mais tempo diante das plataformas digitais.

Também observamos se o aumento na frequência de ocorrência de *tweets* conduziu (ou não) à entrada de novos anglicismos na ciberescrita do português empregado nas mensagens. A seguir, apresentamos o total de anglicismos extraídos dos *tweets* separados por período:

Gráfico 2: Distribuição dos *tweets* com anglicismos separados por período



Fonte: elaborado pela autora (2021).

Conforme se observa no Gráfico 2, no período de coleta dos dados, do total de 2450 anglicismos coletados, a maior frequência ( $1412/2450 = 58\%$ ) ocorreu no segundo período, ao passo que a menor frequência ( $1038/2450 = 42\%$ ) adveio do primeiro período. Esse resultado corrobora nossa hipótese de que, em decorrência do isolamento e distanciamento social devido à pandemia da COVID-19, houve aumento dos domicílios com acesso à rede de 83%, segundo o Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação (Cetic). De acordo com a pesquisa TIC Domicílios 2020 (Edição COVID-19 - Metodologia Adaptada), realizada pelo Comitê Gestor da Internet do Brasil (CGI.br), o Brasil possui 152 milhões de usuários de Internet, o que corresponde a 81% da população do país com dez anos ou mais.

O estudo do CGI.br também observou que o crescimento da proporção de domicílios com acesso à internet se deu em todos os segmentos analisados: nas áreas urbanas e rurais, em todas as regiões, em todas as faixas de renda familiar e estratos sociais. Os domicílios das classes C (91%) e D e E (64%) apresentaram mais diferenças em comparação a 2019 (80% e 50%, respectivamente) e as diferenças regionais diminuíram:

O principal tipo de conexão domiciliar foi a banda larga fixa (68%), com especial aumento das conexões por cabo ou fibra óptica, em consonância com o que revelou a última edição da TIC Provedores. O estudo também mostrou que houve um aumento na presença de computador (*desktop*, portátil ou *tablet*) nos domicílios (passou de 39% em 2019 para 45% em 2020), revertendo uma tendência de declínio que vinha se desenhando nos últimos anos. (Cetic, 2021. não paginado).

Em nossa amostra, identificamos ao todo 437 diferentes anglicismos nos *tweets* nos dois períodos. Em vista disso, nosso próximo interesse era verificar, dentre os itens lexicais empregados na amostra, quais os anglicismos mais recorrentes nos *tweets*. Na Tabela 2, optamos por apresentar os resultados dos dez itens lexicais mais frequentes distribuídos por período:

Tabela 2: Frequência e proporção de anglicismos mais recorrentes nos *tweets* separados por período.

1º período - 03/19 a 02/20			2º período - 03/20 a 02/21		
Item lexical	Apl./Total	%	Item lexical	Apl./Total	%
<b><i>Postar</i></b>	37/1038	3,55	<b><i>Live</i></b>	123/1412	8,70
<b><i>On-line</i></b>	36/1038	3,46	<b><i>On-line</i></b>	58/1412	4,10

<i>Day</i>	33/1038	3,17	<i>Postar</i>	53/1412	3,74
<i>Internet</i>	32/1038	3,07	<i>Internet</i>	42/1412	2,97
<i>Story</i>	30/1038	2,88	<i>Fake</i>	34/1412	2,40
<i>Insta</i>	28/1038	2,69	<i>Insta</i>	33/1412	2,33
<i>Whats</i>	27/1038	2,59	<i>Show</i>	32/1412	2,26
<i>Fake</i>	24/1038	2,31	<i>Top</i>	32/1412	2,26
<i>Top</i>	21/1038	2,02	<i>Story</i>	30/1412	2,12
<i>Show</i>	20/1038	1,92	<i>Site</i>	24/1412	1,70
<b>Total geral</b>	277/1038	26,61	<b>Total geral</b>	447/1412	31,61

Fonte: elaborada pela autora (2012).

Podemos observar, na Tabela 2, que a maioria dos anglicismos tem relação com o mundo da cibercultura - assim como na pesquisa de Torrano (2010) - com exceção das palavras *day*, *top* e *show*. Nota-se também que, em ambos os períodos, a maioria dos itens lexicais com maior ocorrência não está dicionarizado (*postar*, *whats*, *fake*, *story* e *live*). É importante ressaltar ainda que *day* e *whats* aparecem apenas no primeiro período, enquanto *live* e *site* foram empregados como os vocábulos com maior ocorrência apenas no segundo período.

Na próxima seção, passamos à verificação das variáveis linguísticas (presença/ausência de anglicismos nos dicionários contemporâneos da língua portuguesa, assimilação ou não dos termos estrangeiros ao sistema linguístico do português e classe morfológica dos anglicismos) e extralinguísticas (faixa etária, sexo, ocupação e localidade) que atuam no uso dos anglicismos presentes nos *tweets*. Das sete variáveis independentes controladas, o programa estatístico selecionou como significativas apenas três: as variáveis linguísticas assimilação ou não dos termos estrangeiros ao sistema linguístico do português e classe morfológica dos anglicismos e a variável social faixa etária dos usuários do *Twitter* da amostra analisada. Mas, na sequência, optamos por apresentar todos os resultados quantitativos de cada variável controlada.

## 6.2 Variáveis linguísticas

As três variáveis linguísticas consideradas a fim de observar o comportamento dos itens lexicais mais recorrentes na amostra foram a presença/ausência de anglicismos nos dicionários contemporâneos da língua portuguesa, a assimilação ou não dos termos estrangeiros ao sistema linguístico do português e a classe morfológica dos anglicismos coletados nos dois períodos.

### 6.2.1 Presença/ausência de anglicismos nos dicionários

Para caracterização dos anglicismos dicionarizados e não-dicionarizados, tomamos como base dois dicionários monolíngues da língua portuguesa (HOUAISS; VILLAR, 2008; FERREIRA, 2004) representativos do léxico mais recente da variedade brasileira. Também foi consultado o material produzido por Neves (2003), que apresenta um estudo exaustivo dos usos vigentes no português contemporâneo, além de também ser consultado Bechara (2009).

Leitão (2006) encontrou 642 palavras e expressões provenientes da língua inglesa no dicionário Aurélio (1999) e 1072 itens no Houaiss (2001), ambos em formato digital. Apenas 409 eram comuns aos dois dicionários, ou seja, há, no dicionário Houaiss, 663 anglicismos registrados a mais que o anotado pelo Aurélio.

Na sequência, apresentamos algumas ocorrências de *tweets* com anglicismos cujo registro está e não está, respectivamente, dicionarizado em português:

(9) “[...] consegui melhorar a *internet*<sup>51</sup> em 12x a velocidade” (publicado em 13 de abril de 2020, em Chapecó - grifos nossos).

(10) “[...] dupla curvatura apresentado por mim em uma *Live*<sup>52</sup> com o mestre kkkk” (publicado em 12 de fevereiro de 2021, em Catanduvas - grifos nossos).

#### 6.2.3.1 Resultados e discussão

Dentre os anglicismos coletados na amostra, observamos a frequência de 1566/2450 (64%) ocorrências de anglicismos não-dicionarizados e 884/2450 (36%) de dicionarizados. Essa frequência mais alta para os anglicismos não-dicionarizados se mantém mesmo que consideremos os dados separados por período. Observem-se os resultados na Tabela 3, a seguir:

<sup>51</sup> rede (virtual), tradução nossa.

<sup>52</sup> transmissão ao vivo, tradução nossa.

Tabela 3: Frequência e proporção de anglicismos dicionarizados ou não, separados por P1 e P2.

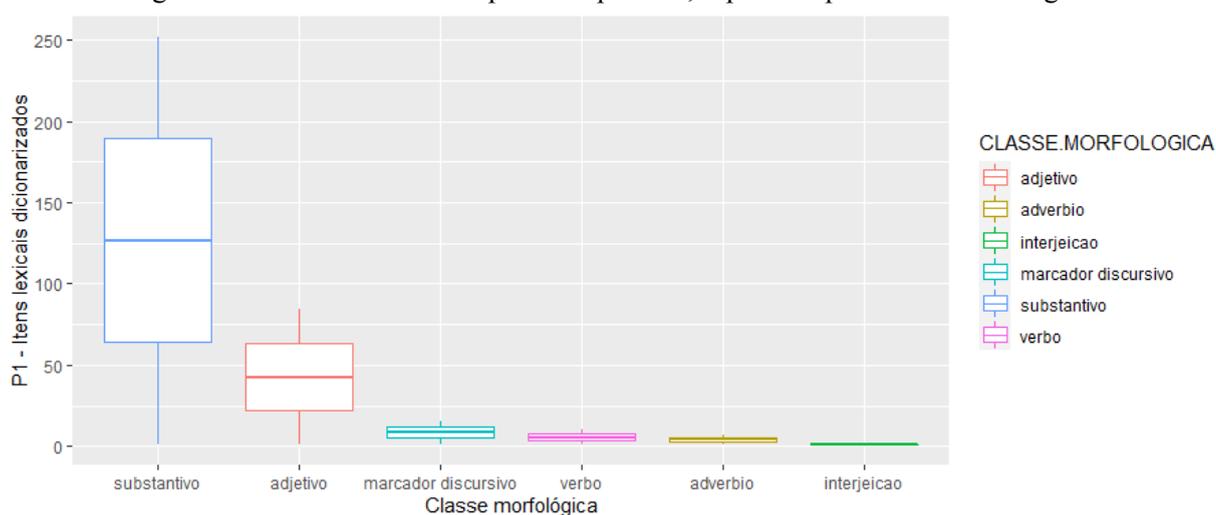
Período	Anglicismo não-dicionarizado		Anglicismo dicionarizado		Total geral
	Apl./Total	%	Apl./Total	%	
<b>P1</b>	670	64,5%	368	35,5%	1038
<b>P2</b>	896	63,5%	516	36,5%	1412
<b>Total parcial</b>	1566	64%	884	36%	2450
p-value < 2.2e-16					

Fonte: elaborada pela autora, (2021).

Os resultados da Tabela 3 revelam que, assim como no estudo de Torrano (2010), os anglicismos não-dicionarizados têm frequência um pouco mais elevada de ocorrência (P1 - 64,5% e P2 - 63,5%) do que os dicionarizados (P1- 35,5% e P2 - 36,5%) nos dois períodos em que os *tweets* foram coletados.

Na sequência, vemos, no Gráfico 3, os anglicismos dicionarizados coletados no P1, separados por classe morfológica:

Gráfico 3: Anglicismos dicionarizados do primeiro período, separados por classe morfológica.



Fonte: elaborado pela autora, 2021.

O gráfico 3 mostra que, no primeiro período, os itens lexicais dicionarizados são representados principalmente por substantivos (252/368 ocorrências = 68%), seguidos dos adjetivos com menos ocorrências (84/368 dados = 23%). Na sequência, veremos um exemplo de cada classe morfológica:

(11) Substantivo: “#F1noSporTV essa nova entrada dos **boxes** ficou muito melhor” (publicado em 21 de junho de 2019, em Chapecó - grifos nossos).

(12) Adjetivo: “Hoje, foiii **topppppppzeraaa**” (publicado em 6 de outubro de 2019, em Ponte Serrada - grifos nossos).

(13) Marcador discursivo: “**Ok**, estou selecionado no mestrado” (publicado em 8 de novembro, em Chapecó - grifos nossos).

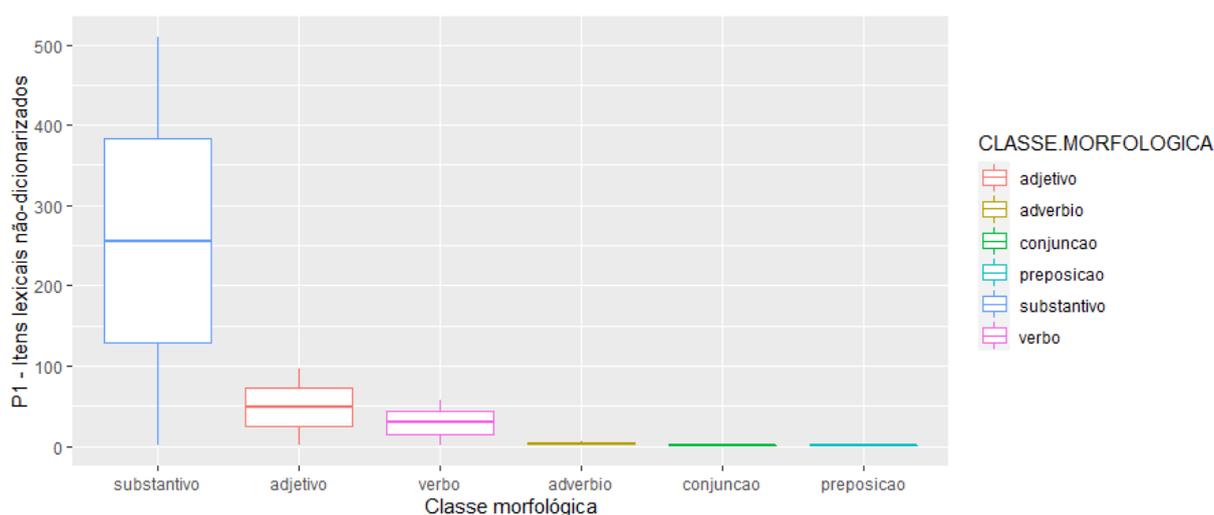
(14) Verbo: “[...] me **stalkeiem** no spotify e vão ver o que é diversidade” (publicado em 7 de dezembro, em Chapecó - grifos nossos).

(15) Advérbio: “[...] quase 2020 e eu novamente mais uma vez **again** chorando com san junipero #BlackMirror” (publicado em 22 de dezembro de 2019, em Chapecó - grifos nossos).

(16) Interjeição: “Oque (sic) dói mais: ( ) Parto ( ) Tiro na Testa ( X ) Escutar o seu **Bye Bye**” (publicado em 19 de março de 2019, em Águas Frias - grifos nossos).

A seguir, vemos, no Gráfico 4, os anglicismos não-dicionarizados coletados no P1, separados por classe morfológica:

Gráfico 4: Anglicismos não-dicionarizados do primeiro período, separados por classe morfológica.



Fonte: elaborado pela autora (2021).

Igualmente, no Gráfico 4, os itens lexicais não-dicionarizados do P1 também são representados em maior número pelos substantivos (509/670 ocorrências = 76%), novamente seguidos pelos adjetivos (97/670 ocorrências = 14%).

Na Tabela 4, reunimos os anglicismos coletados no primeiro período por classe morfológica e presença/ausência nos dicionários da língua portuguesa:

Tabela 4: Anglicismos coletados nos *tweets* do P1 separados por classe morfológica e dicionarizados e não-dicionarizados.

<b>Anglicismos dicionarizados</b>	<b>Apl./Total</b>	<b>%</b>	<b>Anglicismos não-dicionarizados</b>	<b>Apl./Total</b>	<b>%</b>
Substantivo	252/1038	24,28	Substantivo	509/1038	49,03
Adjetivo	84/1038	8,09	Adjetivo	97/1038	9,34
Marcador discursivo	14/1038	1,35	Verbo	56/1038	5,39
Verbo	10/1038	0,96	Advérbio	6/1038	0,58
Advérbio	7/1038	0,67	Conjunção	1/1038	0,09
Interjeição	1/1038	0,09	Preposição	1/1038	0,09
<b>Total</b>	<b>368/1038</b>	<b>35,44</b>	<b>Total</b>	<b>670/1038</b>	<b>64,52</b>

Fonte: elaborada pela autora, (2021).

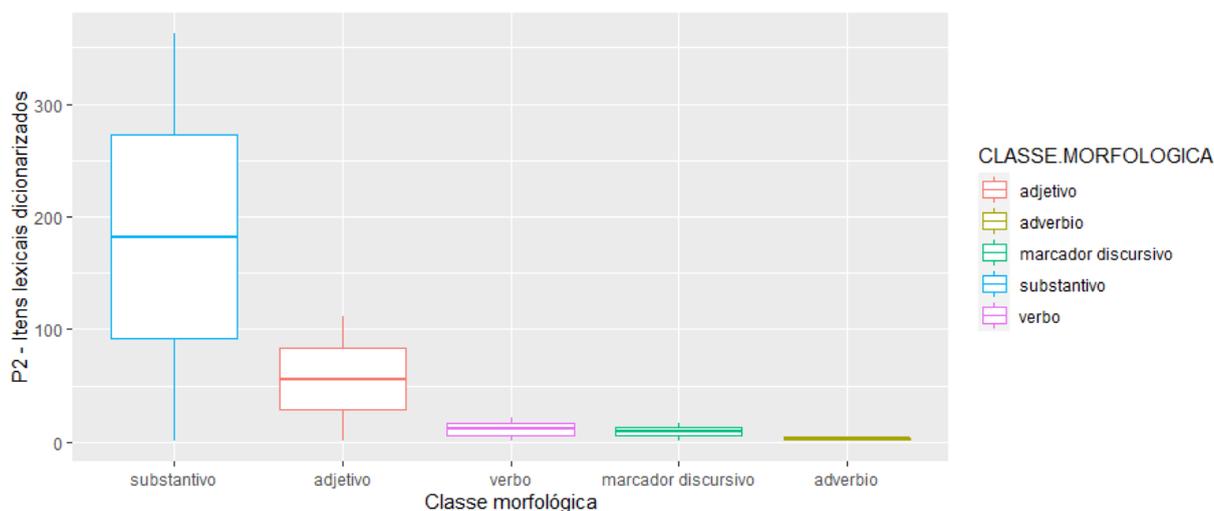
Na Tabela 4, vemos que, no primeiro período, os anglicismos da categoria substantivos são a maior ocorrência de dados, independentemente de serem eles dicionarizados (252) ou não-dicionarizados (509). Ao se tratar dos itens dicionarizados, além dos substantivos, na sequência, os anglicismos mais recorrentes são da categoria adjetivos (84), marcadores discursivos (14), verbos (10), advérbios (7) e localizamos somente um registro de interjeição. Já os vocábulos não-dicionarizados também são ocorrências de adjetivos (97), verbos (56), advérbios (6), além de um registro de conjunção e um de preposição. Vejamos, na sequência, dois exemplos:

(17) Dicionarizado: “Terminamos de jantar e faltou luz, depois de horas sem luz, cancelaram o *show*.” (publicado em 22 de março de 2019, em São Carlos - grifos nossos).

(18) Não-dicionarizado: “#BolsonaroLiberaRinhadeGalo é a melhor *tag* do ano” (publicado em 8 de março de 2019, em Águas Frias - grifos nossos).

Vejamos, a seguir, os resultados para os itens lexicais de origem inglesa dicionarizados do segundo período segundo a classe morfológica no Gráfico 5.

Gráfico 5: Anglicismos dicionarizados do segundo período, separados segundo a classe morfológica.

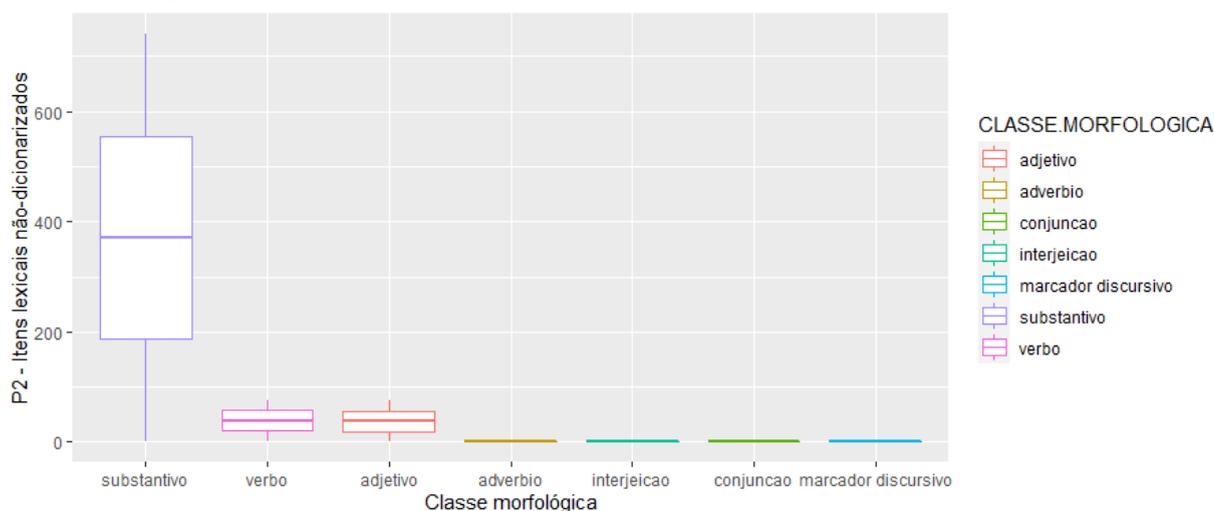


Fonte: elaborado pela autora (2021).

Podemos observar, no Gráfico 5, que os itens lexicais dicionarizados do segundo período analisado aparecem em maior número como substantivos (363/516 ocorrências = 70%), seguidos de frequência relativamente menor pelos adjetivos (111/516 ocorrências = 22%).

Vejamos, a seguir, os resultados para os itens lexicais de origem inglesa não-dicionarizados do segundo período, separados por classe morfológica no Gráfico 6:

Gráfico 6: Anglicismos não-dicionarizados do segundo período, separados por classe morfológica.



Fonte: elaborado pela autora (2021).

Da mesma forma, o Gráfico 6 mostra que os substantivos (739/896 ocorrências = 82%) apresentam maior frequência de ocorrência dentre os anglicismos não-dicionarizados do segundo período. Entretanto, ao contrário do que se viu nos resultados expressos pelos gráficos 3 e 4, do

P1, e pelo Gráfico 5, do P2, no Gráfico 6, são os verbos que ocupam a segunda maior ocorrência (76/896 ocorrências = 8,5%), praticamente empatando com os adjetivos (75/896 ocorrências = 8,4%).

Na Tabela 5, reunimos os anglicismos coletados no segundo período por classe morfológica e presença/ausência nos dicionários da língua portuguesa:

Tabela 5: Anglicismos coletados nos tweets do P2 separados por classe morfológica e dicionarizados e não-dicionarizados.

<b>Anglicismos dicionarizados</b>	<b>Apl./Total</b>	<b>%</b>	<b>Anglicismos não-dicionarizados</b>	<b>Apl./Total</b>	<b>%</b>
Substantivo	363/1412	25,70	Substantivo	739/1412	52,33
Adjetivo	111/1412	7,86	Verbo	76/1412	5,38
Verbo	22/1412	1,55	Adjetivo	75/1412	5,31
Marcador discursivo	15/1412	1,06	Advérbio	2/1412	0,14
Advérbio	5/1412	0,35	Interjeição	2/1412	0,14
<b>Total</b>	516/1412	36,52	Marcador discursivo	1/1412	0,07
			Conjunção	1/1412	0,07
			<b>Total</b>	672/1412	63,44

Fonte: elaborada pela autora, (2021).

Na Tabela 5, observamos que a ocorrência dos substantivos é bem maior do que as outras classes morfológicas, sendo 25,70% de ocorrências para itens dicionarizados e 52,33% para itens não-dicionarizados. De maneira bastante ínfima aparecem duas ocorrências de interjeição (*What*) nos itens não-dicionarizados e uma ocorrência de conjunção (*but*) também não-dicionarizada. Não há ocorrências de preposições dicionarizadas ou não no P2, não havendo também interjeições e conjunções dicionarizadas. Confira a seguir algumas ocorrências encontradas:

(19) “Ele : "sim, vou pensar ainda se vou comprar", “*Whaaaat*” (publicado em 7 de maio de 2020, em São Carlos - grifos nossos.)

(20) “Indo no aeroporto *but sadly* não é pra viajar”, (publicado em 16 de julho de 2020, em Chapecó - grifos nossos)

Assim como nos estudos de Torrano (2010), constatamos em nossa pesquisa que tanto em P1 quanto em P2, os anglicismos não-dicionarizados são os mais recorrentes na amostra. No primeiro período os vocábulos não-dicionarizados representam 64,52% do total dos dados coletados e, no segundo período, os itens lexicais que não foram encontrados em dicionários representam 63,44% dos anglicismos.

Na próxima etapa, voltamos nossa análise para identificar os itens lexicais de origem inglesa mais empregados nos *tweets* a fim de observar o registro desses itens pelos dicionários da língua portuguesa. Na Tabela 4, a seguir, apresentamos os resultados relativos à frequência de ocorrência dos anglicismos dicionarizados mais empregados separados por período:

Tabela 6: Frequência e proporção de anglicismos dicionarizados separados por período.

P1			P2		
Anglicismo dicionarizado	Apl./Total	%	Anglicismo dicionarizado	Apl./Total	%
<i>On-line</i>	36/368	9,78	<i>On-line</i>	58/516	11,24
<i>Internet</i>	32/368	8,69	<i>Internet</i>	42/516	8,14
<i>Top</i>	22/368	5,97	<i>Show</i>	32/516	6,20
<i>Site</i>	25/368	6,79	<i>Top</i>	32/516	6,20
<i>Show</i>	20/368	5,43	<i>Site</i>	24/516	4,65
<b>Total geral</b>	110/368	36,66	<b>Total geral</b>	188/516	36,43

Fonte: elaborada pela autora (2021).

A Tabela 6 mostra que os cinco itens lexicais dicionarizados com maior ocorrência são os mesmos (*on-line*, *internet*, *top*, *site* e *show*) nos dois períodos, mas não na mesma ordem. Nota-se também que três dos cinco itens, assim como na pesquisa de Pereira (2012), têm relação com as TICs: *on-line*, *internet* e *site*.

Na Tabela 7, a seguir, apresentamos os resultados relativos à frequência de ocorrências dos anglicismos não-dicionarizados mais empregados separados por período:

Tabela 7: Frequência e proporção de anglicismos não-dicionarizados separados por período.

P1			P2		
Anglicismo não-dicionarizado	Apl./Total	%	Anglicismo não-dicionarizado	Apl./Total	%
<i>Postar</i>	37/670	5,52	<i>Live</i>	123/896	13,72
<i>Fake</i>	33/670	4,92	<i>Postar</i>	58/896	6,47
<i>Day</i>	33/670	4,92	<i>Fake</i>	34/896	3,79
<i>Story</i>	30/670	4,47	<i>Insta</i>	33/896	3,68
<i>Insta</i>	28/670	4,17	<i>Story</i>	30/896	3,34
<i>Whats</i>	22/670	3,28	<i>Whats</i>	22/896	2,45
<b>Total geral</b>	183/670	27,28	<b>Total geral</b>	300/896	33,45

Fonte: elaborada pela autora, (2021).

Pode-se inferir dos resultados da Tabela 7 quais os anglicismos não-dicionarizados com mais ocorrências nos dois períodos analisados. Tanto no P1 quanto no P2, os itens lexicais que ainda não são encontrados em dicionários de PB são seis substantivos (*fake*, *day*, *story*, *insta*, *whats* e *live*) - como na pesquisa de Souza (2015), em que a maior frequência de ocorrência é de substantivos - e apenas um verbo (*postar* - no sentido de publicar textos e/ou material audiovisual em redes sociais virtuais) entre os vocábulos não-dicionarizados nas cinco maiores ocorrências em ambos os períodos.

### 6.2.2 Anglicismos assimilados ou não ao português

Consideramos que os itens lexicais de origem inglesa são assimilados, isto é, são empréstimos ou neologismos, ou não assimilados, ou seja, são estrangeirismos. De outro modo, desejamos observar se os anglicismos se adaptaram - ou não - ao entrarem na língua receptora, no caso, o português.

Luft (1991, p. 347-354) lista uma série de anglicismos correntes empregados no português, os quais já estão aportuguesados ou ainda não foram assimilados. Na primeira categoria, considera-se que alguns nomes estrangeiros se adaptaram ao sistema fonológico do

português e têm, por exemplo, o seu plural consoante as normas vigentes. São exemplos de neologismo os adjetivos estressado/estressados e blogueiro/blogueiros. Neste caso, por exemplo, os adjetivos estão flexionados em número, sem pretender coincidir com as regras do plural da língua originária. Outro exemplo de anglicismo assimilado, ou empréstimo, é o caso das palavras *sportman/sportmen*, que se regulam pelas normas da língua estrangeira, o que, em geral, é o procedimento recomendado pelas gramáticas normativas (por exemplo, BECHARA, 2009, p. 105-106) e pelos textos científicos (a exemplo de CARVALHO, 1989).

No PB, Bagno (2012) destaca que qualquer enunciado originado por palavras estrangeiras vai se constituir de acordo com as regras morfossintáticas da língua. Bybee (2020) destaca que um verdadeiro empréstimo - em oposição ao uso esporádico de uma palavra estrangeira - se integra na fonologia, morfologia e sintaxe da língua.

A seguir, exemplificamos algumas ocorrências de anglicismos assimilados e não assimilados, nesta ordem, empregados em *tweets* da nossa amostra:

(21) “[...] agora eu vou esperar ter **internet** para ver como resolver o problema”, (publicado em 28 de fevereiro de 2020, em Chapecó, grifo nosso).

(22) “Meninas que passam creme no corpo todos os dias e mesmo assim tão sempre com a pele ressecada me **add**<sup>53</sup>”, (publicado em 15 de julho de 2020, em Chapecó, grifo nosso).

(23) “Eu to chorandoooooo com essa **trolagem** do Instagram MKKKKK”, (publicado em 22 de maio de 2020, em Chapecó, grifo nosso).

Em 21 vemos o vocábulo *internet*, que é considerado um empréstimo linguístico porque ele já está incorporado aos dicionários do PB. Na sequência, no exemplo 22, o vocábulo *add* (adicionar), representando um anglicismo não assimilado, isto é, um estrangeirismo. Por último, temos um exemplo de neologismo através do *tweet* que contém a palavra *trolagem*, que se ajustou às regras morfossintáticas do português ao reunir um verbo em inglês e um sufixo em português.

Nossa hipótese foi a de que encontraríamos a maior parte dos itens lexicais relacionados às TICs ainda não assimilados nos *tweets*, isto é, ainda como estrangeirismos, em razão de a democratização dessa tecnologia ser relativamente recente no país e principalmente em decorrência da pandemia da Covid-19, que obrigou muitos usuários a acessá-la.

---

<sup>53</sup> adicionem, tradução nossa.

### 6.2.3.1 Resultados e discussão

Esta variável linguística foi selecionada como significativa pelo programa estatístico. Vejamos os resultados gerais na Tabela 8 a seguir:

Tabela 8: Frequência e proporção de anglicismos assimilados ou não, separados por P1 e P2.

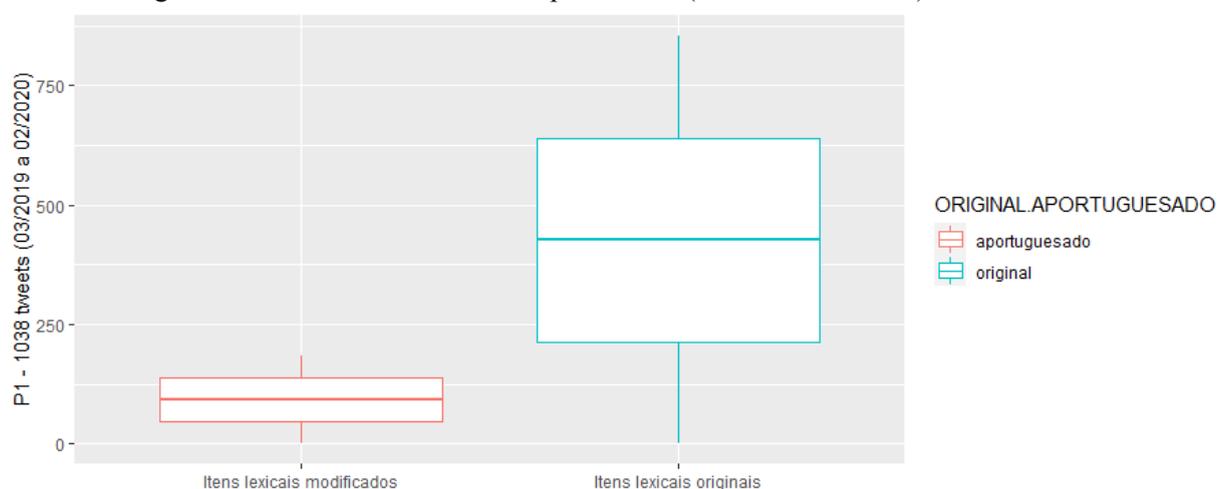
Período	P1		P2		Total geral
	Apl./Total	%	Apl./Total	%	
<b>Anglicismo assimilado</b>	183/404	45	221/404	55	404
<b>Anglicismo não-assimilado</b>	855/2046	42	1191/2046	58	2046
<b>Total parcial</b>	1038/2450	42	1412/2450	58	2450
p-value < 4.84e-11					

Fonte: elaborada pela autora, (2021).

Os resultados da Tabela 8 revelam que os anglicismos assimilados e não assimilados são mais frequentes no segundo período da amostra, conforme nossa expectativa. Os itens assimilados ao PB são alguns pontos percentuais superior em P1 em relação a P2 e, em contrapartida, os não-assimilados apresentaram também alguns pontos percentuais superior em P2 em relação a P1. Essa pouca diferença percentual no emprego dos anglicismos nos dois períodos de análise pode ser explicada, porque “Os empréstimos lexicais durante muito tempo sofreram as críticas dos puristas, mas hoje vão sendo aceitos com mais facilidade, exceto aqueles comprovadamente desnecessários e sem muita repercussão em outros idiomas de cultura do mundo.” (BECHARA, 2009, p. 500),

Vejamos, a seguir, os resultados mais detalhados dos itens lexicais de origem inglesa assimilados ou não português do Brasil coletados no primeiro período, conforme Gráfico 7.

Gráfico 7: Anglicismos assimilados ou não no período P1 (03/2019 a 02/2020).

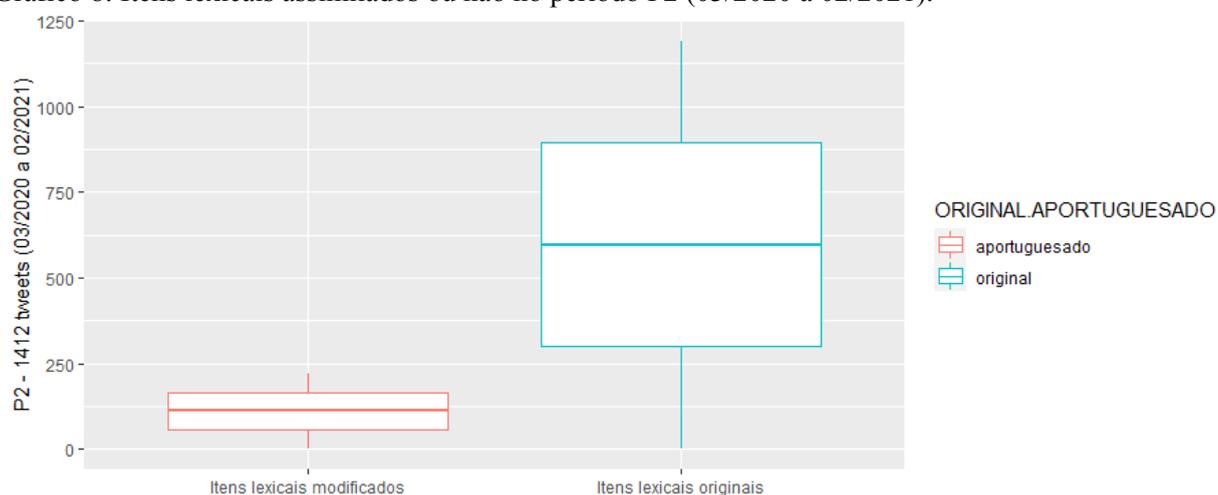


Fonte: elaborado pela autora, (2021).

No Gráfico 7, percebemos que, no primeiro período analisado, os itens lexicais originais (855/1038 ocorrências = 82%) aparecem em um número bem maior do que os neologismos, isto é, quando o item apresenta aportuguesamento (183/1038 ocorrências = 18%).

Vejamos, a seguir, os resultados para os itens lexicais de origem inglesa assimilados ou não à língua portuguesa coletados no segundo período, conforme o Gráfico 8.

Gráfico 8: Itens lexicais assimilados ou não no período P2 (03/2020 a 02/2021).



Fonte: elaborado pela autora, (2021).

Igualmente, no segundo período analisado, os itens lexicais aparecem em maior número no formato original (1191/1412 ocorrências = 84%) e em frequência menor os neologismos (221/1412 ocorrências = 16%).

Constatamos que, nos dois períodos, os itens lexicais não assimilados foram os mais recorrentes, com uma diferença bastante acentuada. Em P1, 82% dos anglicismos são não assimilados e, em P2, 84% dos itens lexicais coletados. Esses resultados confirmam nossa hipótese de que encontraríamos a maior parte dos itens lexicais relacionados às TICs ainda não assimilados nos *tweets*, isto é, ainda como estrangeirismos.

Por fim, cabe comentar que ainda que os gráficos 7 e 9 indicam que a proporção de neologismos é baixa em relação aos empréstimos e estrangeirismos. Resta-nos saber, dentre os anglicismos originais, quais os itens da amostra que se enquadraram entre empréstimos (item lexical dicionarizado) ou estrangeirismos (item lexical não-dicionarizado). Para isso, voltamos aos resultados das tabelas 4 e 5 e vimos que, no conjunto de itens lexicais dicionarizados, por período de amostragem, 5 anglicismos são empréstimos (*on-line, internet, top, site, show*). Esses 5 itens somam juntos 298 ocorrências (12%) totais dos dois períodos de análise. Por outro lado, 6 anglicismos são estrangeirismos (*fake, day, story, insta, whats e live*) porque não estão incorporados aos dicionários do PB. Esses 6 itens somam juntos 483 ocorrências (20%) totais dos dois períodos de análise. Logo, os estrangeirismos parecem corresponder à maior parte dos dados coletados na amostra de *tweets* nos dois períodos se comparados aos neologismos e aos empréstimos.

### 6.2.3 Classe gramatical

A maior parte dos anglicismos presentes no português do Brasil é da categoria dos nomes (substantivos e adjetivos) ou verbos, que, por consequência, assimilam os recursos do sistema morfológico da língua e podem desempenhar diferentes funções sintáticas na sentença. Poplack *et al.* (1998 apud BYBEE, 2020, p. 331) alerta que “onde o empréstimo acontece, os substantivos são os mais frequentes, talvez por causa do seu alto grau de conteúdo lexical e do seu menor grau de integração no discurso do que classes como verbos e adjetivos.” Também a adaptação de verbos emprestados de uma língua flexional é tarefa difícil, embora a maioria das línguas tenha recurso para fazer isso, porém “[...] é bom lembrar que o empréstimo de verbos é muito menos frequente que o de substantivos”. (BYBEE, 2020, p. 334).

Na sequência, apresentamos algumas ocorrências de *tweets* com anglicismos usados como adjetivo, substantivo e verbo, respectivamente, em diferentes funções sintáticas:

(23) “As vezes eu até me assusto com o quão *stalker*<sup>54</sup> eu sou”, (publicado em 30 de março de 2020, em Arvoredo, grifo nosso).

(24) “Se tu é a atendente do *chat*<sup>55</sup> da Sul América , parabéns tu me causou ódio mais uma vez”, (publicado em 23 de agosto de 2019, em Xaxim, grifo nosso).

(25) “*Trolei*<sup>56</sup> a Erika e a tia com o filtro do cachorro, não consigo parar de rir”, (publicado em 9 de dezembro de 2019, em Capinzal, grifo nosso).

Nossa hipótese, baseada em Poplack *et al.* (1998 *apud* BYBEE, 2020, p. 331), era que a categoria substantivo fosse mais recorrente que as demais, devido ao mundo globalizado em que vivemos, onde os contatos de nações e de culturas são propiciados com muita facilidade e rapidez por meio das TICs. Segundo Bybee (2020, p. 330), “É bastante fácil ver que as palavras que designam alimentos viajam junto com os alimentos. Um roteiro semelhante vale para outros objetos cujo progresso mundo afora pode ser traçado de maneira parecida”. Por essa razão, acreditamos, principalmente em decorrência da pandemia da Covid-19, que muitos dos substantivos empregados pelos usuários do Twitter serão derivados das TICs.

### 6.2.3.1 Resultados e discussão

Esta variável linguística foi selecionada como significativa pelo programa estatístico. Vejamos os resultados gerais na Tabela 9 a seguir:

Tabela 9: Frequência e proporção de anglicismos assimilados ou não, separados por P1 e P2.

Período	P1		P2		Total geral
	Apl./Total	%	Apl./Total	%	
<b>Substantivo</b>	761/1863	41	1102/1863	59	1863
<b>Adjetivo</b>	181/367	49	186/367	51	367
<b>Verbo</b>	66/164	40	98/164	60	164
<b>Marcador discursivo</b>	14/30	47	16/30	53	30

<sup>54</sup> perseguidor, tradução nossa.

<sup>55</sup> bate-papo virtual, tradução nossa.

<sup>56</sup> Zoei, tradução nossa.

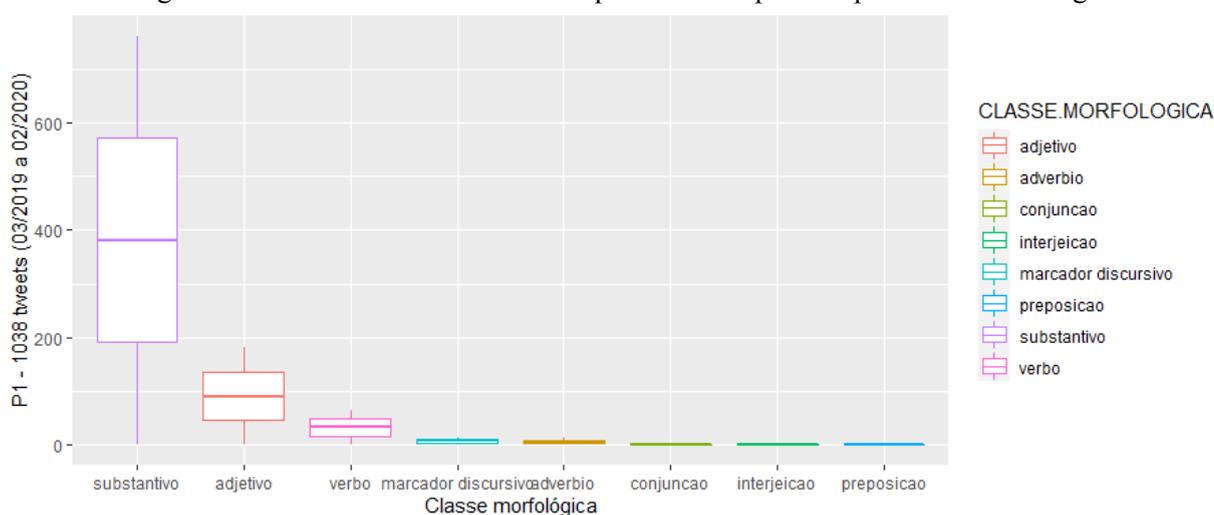
<b>Advérbio</b>	13/20	65	7/20	35	20
<b>Conjunção</b>	1/2	50	1/2	50	2
<b>Interjeição</b>	1/3	33	2/3	67	3
<b>Total parcial</b>	1038/2450	42	1412/2450	58	2450
					p-value < 1.46e-12

Fonte: elaborada pela autora, (2021).

Os resultados da Tabela 9 revelam que os anglicismos da categoria são substantivo mais recorrentes nos dois períodos da amostra, seguidos das categorias morfológicas verbo e adjetivo, conforme nossa expectativa. Esse resultado confirma nossa hipótese, baseada em Poplack *et al.* (1998 *apud* BYBEE, 2020, p. 331), de que a categoria substantivo fosse mais recorrente que as demais. À exceção dos advérbios, todos os itens derivados de diferentes classes gramaticais tiveram incremento do uso no segundo período em relação ao primeiro período.

Na sequência, vemos a análise dos dados de anglicismos separadas por períodos:

Gráfico 9: Anglicismos encontrados nos *tweets* do período P1 separados por classe morfológica.



Fonte: elaborado pela autora (2021).

Observamos, no Gráfico 9, que, no P1, os substantivos (761/1038 ocorrências = 73%) aparecem com maior destaque, seguidos pelos adjetivos (181/1038 ocorrências = 17%), verbos (66 ocorrências = 7%) e, com menor frequência, totalizando 30/1038 ocorrências = 3%, os marcadores discursivos (14 ocorrências), os advérbios (13 ocorrências), as conjunções (1/1038

ocorrência), as interjeições (1/1038 ocorrência) e as preposições (1/1038 ocorrência). A seguir, listamos alguns exemplos da nossa amostra:

(26) “[...] **ok** agr eu tenho certeza q só minha mãe me entende”, (publicado em 14 de outubro de 2019, em Chapecó - grifos nossos).

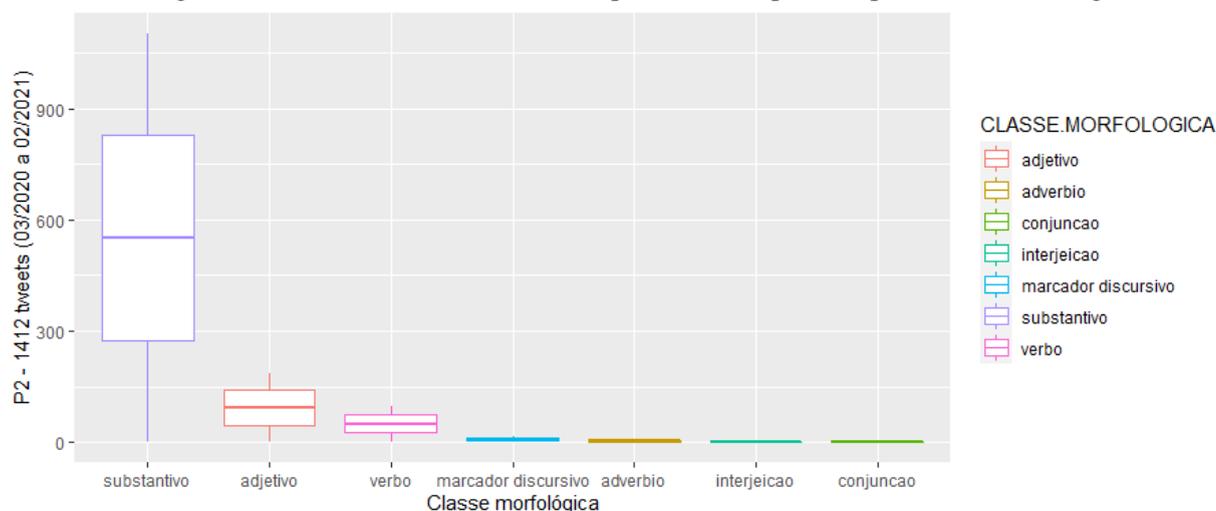
(27) “[...] quer brigar cmg **ok** não me importo, só me deixa em paz”, (publicado em 10 de maio de 2019, em Vargeão - grifos nossos).

(28) “[...] *playlist* de lavar o carro foi um *this is* ppa, pra tomar banho já um instrumentalzera **and chill**” (publicado em 7 de dezembro de 2019, em Xaxim - grifos nossos).

(29) “Oque (sic) dói mais: ( ) Parto ( ) Tiro na Testa ( X ) Escutar o seu **Bye Bye** depois do *I Love You*”, (Publicado em 19 de março de 2019, em Águas Frias - grifos nossos).

(30) “[...] frança fodendo com o brasil no feminino **since** 2017”, (publicado em 23 de junho de 2019, em Chapecó - grifos nossos).

Gráfico 10: Anglicismos encontrados nos *tweets* do período P2 separados por classe morfológica.



Fonte: elaborado pela autora (2021).

De modo semelhante, os substantivos (1102/1412 ocorrências = 78%) também representam a maioria dos itens lexicais coletados no segundo período, seguidos pelos adjetivos (186/1412 ocorrências = 13%), os verbos (98/1412 ocorrências = 7%) e, de modo menos frequente, totalizando 26/1412 ocorrências = 2%, os marcadores discursivos (16/1412 ocorrências), os advérbios (7/1412 ocorrências), as interjeições (2/1412 ocorrências) e as conjunções (1/1412 ocorrência).

Nesta subseção, identificamos os anglicismos da classe dos substantivos como mais recorrentes em ambos os períodos, representando 73,31% dos dados coletados no P1, e 78,05% dos itens lexicais analisados no P2. Também observamos que os adjetivos (P1 17,43% e P2 13,18%) e os verbos (P1 6,35% e P2 6,94%) são as outras duas classes morfológicas mais frequentes na amostra de dados coletados. Câmara Jr. (2008) afirma que artigos, preposições, conjunções etc. pertencem às "classes gramaticais fechadas", com um número restrito de membros. Tal afirmação pode ser confirmada a partir das classes morfológicas que aparecem com menor frequência em nossa amostra: os marcadores discursivos (P1 1,34% e P2 1,13%), os advérbios (P1 1,25% e P2 0,49%), as conjunções (P1 0,09% e P2 0,07%), as interjeições (P1 0,09% e P2 0,14%) e as preposições (P1 0,09% e P2 0,0%), que somados representam menos que 3% dos dados coletados em cada período.

### 6.3 Variáveis extralinguísticas

As quatro variáveis extralinguísticas consideradas a fim de observar o comportamento dos itens lexicais mais recorrentes na amostra foram o sexo, a idade, a localidade e a ocupação dos usuários da nossa amostra.

#### 6.3.1 Sexo

Dentre as variáveis extralinguísticas, o sexo tende a atuar no uso da língua. Segundo alguns estudos de Labov (2008 [1972], p. 353), “as mulheres adotam as formas inovadoras mais rapidamente que os homens”, quando estas são variantes prestigiadas. Por outro lado, os homens tendem a liderar a mudança quando as variantes são socialmente desprestigiadas, ao passo que as mulheres, nesse mesmo contexto, assumem uma atitude conservadora.

Na sequência, apresentamos quatro ocorrências de *tweets*, dois publicados por mulheres e dois publicados por homens, respectivamente, em que se observa o uso dos anglicismos:

(31) “[...] os *bugs*<sup>57</sup> do meu celular aiaiaii”, (publicado em 9 de fevereiro de 2020, em Xanxerê, grifo nosso).

---

<sup>57</sup> erros, tradução nossa.

(32) “[...] tenho mais vergonha de falar no *meet*<sup>58</sup> que na sala (não faz sentido)”, (publicado em 18 de agosto de 2020, em Videira, grifo nosso).

(33) “Eu escrevo super fomal [formal] nos *emails*<sup>59</sup> pras profs[professoras] e nos trabalhos ai chega nas ead (sic) eu escrevo no *chat*<sup>60</sup> "to [estou] comeno [comendo] gelo””, (publicado em 15 de junho de 2020, em Joaçaba, grifos nossos).

(34) “1 *like*<sup>61</sup> vou ir chorar no mesmo lugar de sempre”, (publicado em 13 de agosto de 2019, em Concórdia, grifo nosso).

Apesar de os dados estatísticos levantados pelo *Twitter* (2018) revelarem que a maioria dos tuiteiros brasileiros é do sexo masculino (57%), em 2019, havia mais mulheres do que homens no Brasil, conforme a PNAD (Pesquisa Nacional por Amostra de Domicílios Contínua), registro mais recente dos cidadãos residentes no país, segundo o sexo. De acordo com o IBGE (2019), a população brasileira é composta por 51,8% de mulheres e 48,2% de homens. Portanto, acreditamos que os anglicismos seriam mais recorrentes na ciberescrita dos homens por serem a maioria dos usuários desta rede social.

### 6.3.1.2 Resultados e discussão

Os gráficos 11 e 12 que representam o comportamento dos 50 usuários com relação ao uso dos anglicismos separados por sexo nos dois períodos analisados:

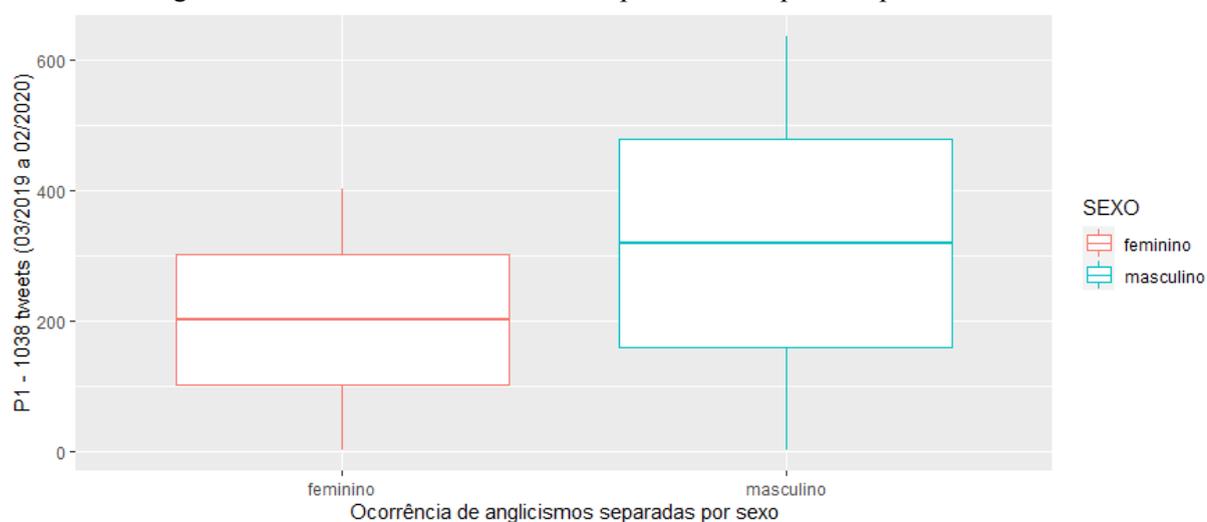
---

<sup>58</sup> encontro, tradução nossa. Neste caso, o usuário está se referindo à plataforma *Google Meet*.

<sup>59</sup> correio eletrônico, tradução nossa.

<sup>60</sup> bate-papo virtual, tradução nossa.

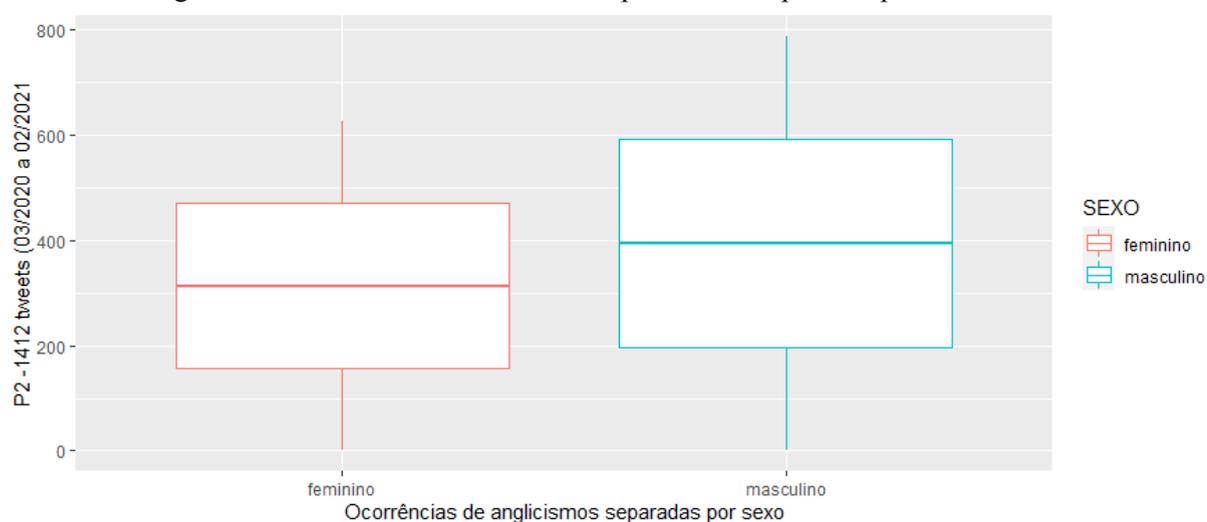
<sup>61</sup> curtida, tradução nossa.

Gráfico 11: Anglicismos identificados nos *tweets* do período P1 separados por sexo.

Fonte: elaborado pela autora, (2021).

No Gráfico 11, observamos que, de modo geral, os anglicismos coletados no P1 ocorrem em maior número em *tweets* de usuários do sexo masculino (636/1038 ocorrências = 61%), seguido pelas ocorrências de usuários do sexo feminino (402/1038 = 39%). Este gráfico corrobora com os dados apresentados na Tabela 1, que mostra que a maioria dos perfis no *Twitter* é de homens (57% dos usuários), e, portanto, com nossa hipótese de que os anglicismos seriam mais recorrentes na ciberescrita dos homens por serem a maioria dos usuários desta rede social.

Na sequência, apresentamos o Gráfico 12, com os resultados do P2, separados por sexo:

Gráfico 12: Anglicismos identificados nos *tweets* do período P2 separados por sexo.

Fonte: elaborado pela autora, (2021).

Igualmente no P2, de modo geral, o maior número de dados coletados é representado por anglicismos publicados por homens, com 786/1412 ocorrências (55,66%), em relação às

mulheres, com 626/1412 ocorrências (44,33%), confirmando os resultados da Tabela 1 e a nossa hipótese.

A seguir, temos as tabelas 7 e 8, que mostram os dez anglicismos mais recorrentes em *tweets* separados por sexo:

Tabela 10: Anglicismos mais recorrentes entre o público feminino, segundo P1 e P2.

<b>Feminino</b>					
<b>P1</b>			<b>P2</b>		
<b>Item lexical</b>	<b>Apl./Total</b>	<b>%</b>	<b>Item lexical</b>	<b>Apl./Total</b>	<b>%</b>
<i>Fake</i>	20/402	4,97	<i>Live</i>	48/626	7,66
<i>Insta</i>	17/402	4,22	<i>Impeachment</i>	30/626	4,79
<i>News</i>	16/402	3,98	<i>Postar</i>	28/626	4,47
<i>Postar</i>	16/402	3,98	<i>Insta</i>	27/626	4,31
<i>Ok</i>	13/402	3,23	<i>Fake</i>	25/626	3,99
<i>Show</i>	12/402	2,98	<i>Internet</i>	22/626	3,51
<i>Boy</i>	11/402	2,73	<i>Story</i>	22/626	3,51
<i>Whats/wpp/zap</i>	11/402	2,73	<i>On-line/on</i>	20/626	3,19
<i>Internet</i>	10/402	2,48	<i>BBB</i>	19/626	3,03
<i>Tweet/tt</i>	10/402	2,48	<i>Top</i>	19/626	3,03
<b>Total</b>	136/402	33,78	<b>Total</b>	260/626	41,49

Fonte: elaborada pela autora, (2021).

Ao olharmos para os dados das informantes femininas da amostra, percebemos que, no primeiro período, de modo geral, há maior frequência de anglicismos relacionados ao meio virtual. De modo específico, os termos mais empregados pelas mulheres são: *Fake* (4,97%), *Insta* (4,22%), *News* (3,98%), *Postar* (3,98%), *Whats/wpp/zap* (2,73%), *Internet* (2,48%) e *Tweet/tt* (2,48%). Dos 402 itens coletados em P1, 136 correspondem aos dez vocábulos com

mais ocorrências, equivalentes a 33,78% do total da amostra. Já no segundo período, de modo geral, sete das dez palavras mais recorrentes apresentam alguma relação com as TICs. De modo específico, os anglicismos mais frequentes são: *Live* (7,66%), *Postar* (4,47%), *Insta* (4,31%), *Fake* (3,99%), *Internet* (3,51%), *Story* (3,51%) e *On-line/on* (3,03%). Há 260/626 ocorrências dos dez itens mais recorrentes em P2, correspondendo a 41,49% dos dados.

Na sequência, podemos observar as dez palavras com maior incidência dos dois períodos analisados, em *tweets* do público masculino:

Tabela 11: Anglicismos mais recorrentes entre o público masculino, segundo P1 e P2.

Masculino					
P1			P2		
Item lexical	Apl./Total	%	Item lexical	Apl./Total	%
<i>On-line/on</i>	39/636	6,13	<i>Live</i>	75/786	9,54
<i>Day</i>	31/636	4,87	<i>On-line/on</i>	38/786	4,83
<i>Every</i>	29/636	4,55	<i>Life</i>	32/786	4,07
<i>Story</i>	25/636	3,93	<i>Motivation</i>	25/786	3,18
<i>Site</i>	23/636	3,61	<i>Postar</i>	25/786	3,18
<i>Internet</i>	22/636	3,45	<i>Internet</i>	20/786	2,54
<i>Postar</i>	21/636	3,30	<i>Mtb/mountain bike</i>	20/786	2,54
<i>VAR</i>	20/636	3,14	<i>Show</i>	19/786	2,41
<i>Whats/wpp/zap</i>	16/636	2,51	<i>Site</i>	19/786	2,41
<i>E-book</i>	14/636	2,20	<i>Design</i>	17/786	2,16
<b>Total</b>	217/636	37,69	<b>Total</b>	290/786	36,86

Fonte: elaborada pela autora, (2021).

Ao olharmos para os dados dos informantes masculinos da amostra, percebemos que também eles recorrem aos vocábulos relacionados com meio virtual frequentemente. No primeiro período, de modo específico, há oito itens que se destacam: *On-line/on* (6,13%), *Story* (3,93%), *Site* (3,61%), *Internet* (3,45%), *Postar* (3,30%), *VAR* (3,14%), *Whats/wpp/zap* (2,51%) e *E-book* (2,20%). Há 217 registros dos dez itens mais recorrentes em P1, equivalente a 37,69% dos 636 vocábulos do mesmo período. No que diz respeito ao segundo período, de modo específico, observamos entre as dez palavras mais recorrentes que seis têm relação com as TICs:

Live (9,54%), *On-line/on* (4,83%), *Postar* (3,18%), *Internet* (2,54%), *Site* (2,41%) e *Design* (2,16%). Os dez vocábulos mais recorrentes contabilizam 290/786 registros, correspondendo a 36,86% do total de dados de P2.

Em síntese, os resultados dos dois períodos analisados mostraram que os anglicismos seriam mais recorrentes na ciberescrita dos homens por serem a maioria dos usuários desta rede social.

### 6.3.2 Faixa etária

A idade dos usuários, segundo Labov (1972 [2008]), pode sinalizar “[...] a predominância de uma variante entre os mais jovens e sua pouca ocorrência entre os mais velhos pode estar indicando uma mudança em progresso, isto é, que uma das variantes está sendo abandonada em favor de outra” (*apud* Faraco 2005, p. 186).

Pesquisa realizada pelo *Global Web Index* (2017) e *Comscore* (2018), divulgada pelo *Twitter* (2018), apontou que a maioria dos usuários tem entre 16 e 54 anos, distribuídos nas seguintes faixas etárias: entre 16 e 24 anos (32%), entre 25 e 34 anos (29%), entre 25 e 44 anos (19%) e entre 45 e 54 anos (11%). De acordo com os termos de uso do *Twitter* (2020), para criar uma conta na rede social, o usuário deve ter, no mínimo, 13 anos, sendo obrigatório informar a data de nascimento, cabendo ao tuiteiro torná-la pública ou não aos demais usuários. Na Figura 5 podemos ver um exemplo de perfil aberto:

Figura 5: Usuário(a) do *Twitter* com informações públicas no perfil.



Fonte: *Twitter* (2021, não paginado)

Nossa hipótese era que os usuários mais jovens empregassem mais os anglicismos do que os mais velhos, que tenderiam ao uso menos frequente desses itens lexicais por força do menor tempo de contato com as redes sociais.

Caracterizamos os usuários da nossa amostra a partir de quatro faixas etárias: 16 a 24 anos, 25 a 34 anos, 35 a 44 anos e 45 a 54 anos.

### 6.3.2.2 Resultados e discussão

Esta variável extralinguística foi selecionada como significativa pelo programa estatístico. Os resultados da análise mostram que nossa hipótese se confirma em parte no que tange à maior frequência de uso dos anglicismos pelas duas faixas etárias mais jovens. vejamos os resultados gerais na Tabela 12 da sequência:

Tabela 12: Frequência e proporção de anglicismos assimilados ou não, separados por P1 e P2.

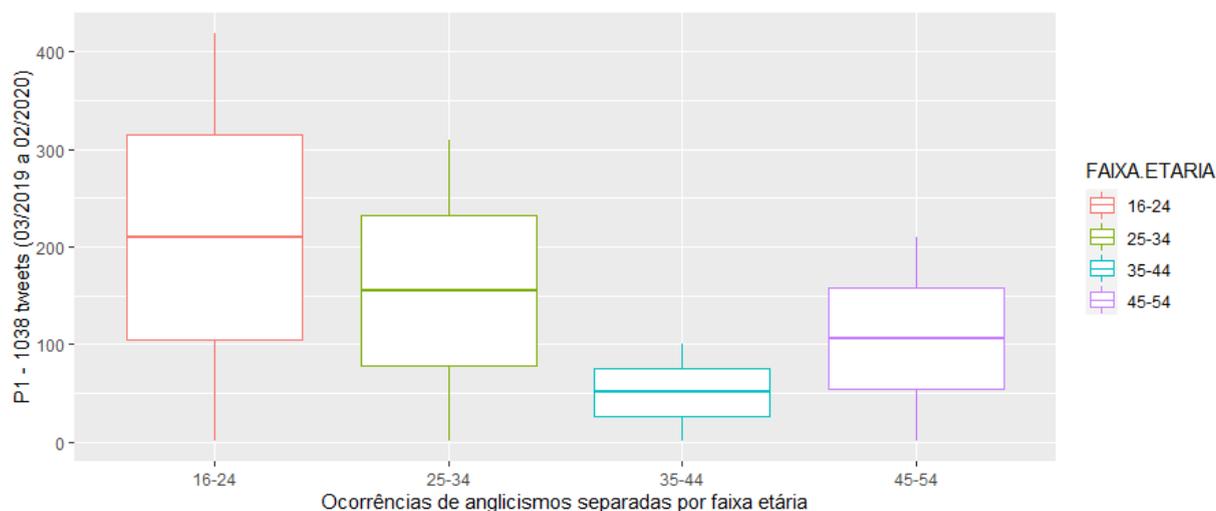
Período	P1		P2		Total geral
	Apl./Total	%	Apl./Total	%	
16 a 24 anos	418/921	45	503/921	55	921
25 a 34 anos	310/823	38	513/823	62	823
35 a 44 anos	101/324	31	223/324	69	324
45 a 54 anos	209/382	55	173/382	45	382
<b>Total parcial</b>	1038/2450	42	1412/2450	58	2450
p-value < 2.10e-09					

Fonte: elaborada pela autora, (2021).

Os resultados da Tabela 12 revelam que os anglicismos foram mais recorrentes em três das quatro faixas etárias no segundo período da amostra, à exceção do grupo dos mais velhos que, em termos de percentual, empregaram mais itens lexicais do inglês no primeiro período. Esse resultado também confirma nossa hipótese de que os usuários mais jovens empregassem mais os anglicismos nos dois períodos em relação aos mais velhos, que tenderiam ao menor uso dos estrangeirismos por força do menor tempo de contato com as redes sociais.

Vejam, na sequência, os gráficos 13 e 14, que separam os dados coletados em P1 e P2 por faixa etária:

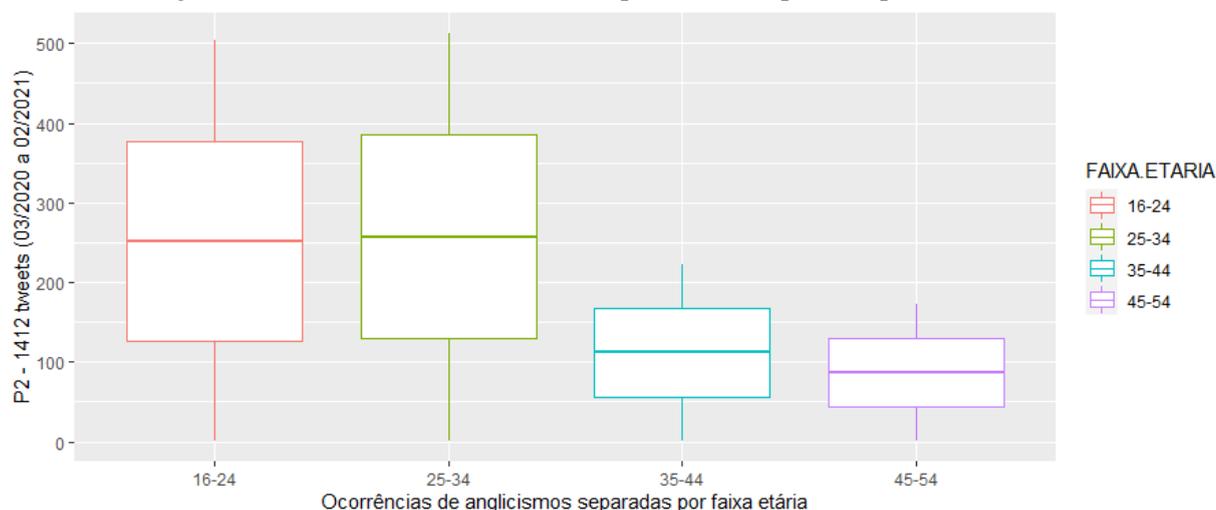
Gráfico 13: Anglicismos identificados nos *tweets* do período P1 separados por faixa etária.



Fonte: elaborado pela autora, (2021).

No gráfico 13, podemos observar as ocorrências de anglicismos coletados no primeiro período separadas pelas quatro faixas etárias analisadas. Apesar de os dados da Tabela 1 informarem que a maior porcentagem de uso do Twitter se dá da menor para a maior faixa etária, nos nossos dados, obtivemos uma pequena diferença. De modo específico, a faixa etária de 16 a 24 anos representa 418/1038 dados (40,27%) dos anglicismos obtidos no P1, já o segundo intervalo etário, 25 a 34 anos, totaliza 310/1038 ocorrências (29,86%) dos dados totais do P1. A terceira faixa etária, 35 a 44 anos, registra 101/1038 dados (9,73%) dos itens lexicais. Por fim, o quarto intervalo de idades, 45 a 54 anos, representa 209/1038 dados, (20,13%) dos dados totais do primeiro período analisado.

Gráfico 14: Anglicismos identificados nos *tweets* do período P2 separados por faixa etária.



Fonte: elaborado pela autora, (2021).

No Gráfico 14, notamos que houve uma mudança de comportamento quando comparamos P1 e P2. Desta vez, de modo geral, o número de ocorrências das duas primeiras faixas etárias é bastante similar. De modo específico, os usuários entre 16 e 24 anos apresentam 503/1412 dados (35,62%) do total dos itens coletados no segundo período. Já a segunda faixa etária, de 25 a 34 anos, registra 513/1412 ocorrências (36,33%) totais. A terceira faixa etária, de 35 a 44 anos, representa 223/1412 itens lexicais (15,79%). Por fim, o último intervalo de idades, de 45 a 54 anos, registra 173/1412 dados (12,25%). A seguir, discriminamos, nas tabelas 9, 10, 11 e 12, os itens lexicais mais frequentes em cada uma das quatro faixas etárias analisadas:

Tabela 13: Itens lexicais mais frequentes em *tweets* da faixa etária de 16 a 24 anos em P1 e P2.

16 a 24 anos					
P1			P2		
Item	Apl./Total	%	Item	Apl./Total	%
<i>Insta</i>	22/418	5,26	<i>Live</i>	44/503	8,74
<i>Dm/direct</i>	17/418	4,06	<i>Postar</i>	25/503	4,97
<i>Ok</i>	16/418	3,82	<i>Internet</i>	24/503	4,77
<i>Postar</i>	16/418	3,82	<i>On-line/on</i>	22/503	4,37
<i>Top</i>	14/418	3,34	<i>Insta</i>	20/503	3,97
<i>Whats/wpp/zap</i>	13/418	3,11	<i>Top</i>	16/503	3,18
<i>Internet</i>	12/418	2,87	<i>Story</i>	15/503	2,98
<i>Show</i>	11/418	2,63	<i>Tt(Twitter)</i>	14/503	2,78
<i>Snap</i>	11/418	2,63	<i>Ok</i>	13/503	2,58
<i>Dog</i>	9/418	2,15	<i>Show</i>	12/503	2,38
<b>Total</b>	141/418	33,69	<b>Total</b>	205/503	40,72

Fonte: elaborada pela autora, (2021).

A Tabela 13 revela, de modo específico, os dez itens lexicais mais recorrentes no primeiro período equivalem a 33,69% do total dos dados (418/1038) dos anglicismos

empregados pelos usuários da primeira faixa etária. Destacamos os que, de alguma forma, se relacionam com a cibercultura: *Insta* (5,26%), *Dm/direct* (4,06%), *Postar* (3,82%), *Whats/wpp/zap* (3,11%), *Internet* (2,87%) e *Snap* (2,63%). Já no segundo período, de modo específico, as dez palavras mais frequentes correspondem a 40,72% da amostra total (503/1412) dos anglicismos empregados pelos usuários da primeira faixa etária.. Destacamos, neste recorte analisado, os seguintes itens: *Live* (8,74%), *Postar* (4,97%), *Internet* (4,77%), *On-line/on* (4,37%), *Insta* (3,97%), *Story* (2,98%) e *Tt* (2,78%).

Vejamos a seguir as palavras mais recorrentes na segunda faixa etária analisada:

Tabela 14: Itens lexicais mais frequentes em *tweets* da faixa etária de 25 a 34 anos em P1 e P2.

25 a 34 anos					
P1			P2		
Item	Apl./Total	%	Item	Apl./Total	%
<i>Day</i>	30/310	9,67	<i>Life</i>	30/513	5,84
<i>Every</i>	29/310	9,35	<i>Motivation</i>	25/513	4,87
<i>Story</i>	19/310	6,12	<i>Postar</i>	20/513	3,89
<i>Postar</i>	14/310	4,51	<i>BBB</i>	17/513	3,31
<i>Baby</i>	7/310	2,25	<i>Design</i>	17/513	3,31
<i>Girl</i>	7/310	2,25	<i>Live</i>	17/513	3,31
<i>Whats/wpp/zap</i>	7/310	2,25	<i>On-line/on</i>	13/513	2,53
<i>E-mail</i>	6/310	1,93	<i>Insta</i>	11/513	2,14
<i>Top</i>	6/310	1,93	<i>Motion</i>	10/513	1,94
<i>Tweet/tt</i>	6/310	1,93	<i>Fake</i>	10/513	1,94
<b>Total</b>	131/310	42,19	<b>Total</b>	170/513	33,08

Fonte: elaborada pela autora, (2021).

Assim como na primeira faixa etária analisada, de modo geral, os anglicismos publicados por usuários entre 25 a 34 anos apresentam termos relacionados às TICs entre os mais recorrentes. Destacam-se, de modo específico, no primeiro período: *Story* (6,12%), *Postar* (4,51%), *Whats/wpp/zap* (2,25%), *E-mail* (1,93%) e *Tweet/tt* (1,93%). Os dez vocábulos mais recorrentes em P1 equivalem a 42,19% do total da amostra (310/1038) dos anglicismos empregados pelos usuários da segunda faixa etária. No segundo período, de modo específico, sobressaem-se os seguintes itens: *Postar* (3,89%), *Design* (3,31%), *Live* (3,31%), *On-line/on* (2,53%), *Insta* (2,14%) e *Fake* (1,94%). Há 170 registros das palavras com maior frequência em P2, correspondendo a 33,08% do total (513/1412) dos anglicismos empregados pelos usuários da segunda faixa etária.

Em seguida, apresentamos a Tabela 15, com os vocábulos de origem inglesa mais recorrentes empregados pela terceira faixa etária analisada:

Tabela 15: Itens lexicais mais frequentes em *tweets* da faixa etária de 35 a 44 anos em P1 e P2.

35 a 44 anos					
P1			P2		
Item	Apl./Total	%	Item	Apl./Total	%
<i>Fake</i>	16/101	15,84	<i>Live</i>	30/223	13,45
<i>News</i>	15/101	14,85	<i>Impeachment</i>	18/223	8,07
<i>VAR</i>	11/101	10,89	<i>News</i>	12/223	5,38
<i>Link/hiperlink</i>	7/101	6,93	<i>App</i>	11/223	4,93
<i>Internet</i>	6/101	5,94	<i>Fake</i>	11/223	4,93
<i>Postagem/post</i>	4/101	3,96	<i>Link/hiperlink</i>	9/223	4,03
<i>Postar</i>	3/101	2,97	<i>Postar</i>	8/223	3,58
<i>Show</i>	3/101	2,97	<i>Download</i>	7/223	3,13
<i>Day</i>	2/101	1,98	<i>E-book</i>	6/223	2,69
<i>Golden</i>	2/101	1,98	<i>E-mail</i>	6/223	2,69

<b>Total</b>	69/101	68,31	<b>Total</b>	118/223	52,88
--------------	--------	-------	--------------	---------	-------

Fonte: elaborada pela autora, (2021).

Assim como nas faixas etárias anteriores, de modo geral, os usuários do Twitter entre 35 e 44 anos também apresentam uma tendência de uso de anglicismos relacionados à cibercultura. Ressaltamos, de modo específico, no primeiro período, os seguintes vocábulos: *Fake* (15,84%), *News* (14,85%), *VAR* (10,89%), *Link/Hiperlink* (6,93%), *Internet* (5,94%), *Postagem/post* (3,96%) e *Postar* (2,97%). As dez palavras mais recorrentes em P1 correspondem a 68,31% do total de itens empregados pelos usuários da terceira faixa etária. Destacamos, de modo específico, no segundo período, os seguintes anglicismos: *Live* (13,45%), *News* (5,38%), *App* (4,93%), *Fake* (4,93%), *Link/Hiperlink* (4,03%), *Postar*, (3,58%), *Download* (3,13%), *E-book* (2,69%) e *E-mail* (2,69%). A dezena de palavras com mais frequência corresponde a 52,88% do total da amostra (223) dos anglicismos empregados pelos usuários da terceira faixa etária.

Vejam, a seguir, as palavras de origem inglesa mais utilizadas pela quarta e última faixa etária analisada:

Tabela 16: Itens lexicais mais frequentes em *tweets* da faixa etária de 45 a 54 anos em P1 e P2.

45 a 54 anos					
P1			P2		
Item	Apl./Total	%	Item	Apl./Total	%
<i>On-line/on</i>	36/209	17,22	<i>Live</i>	32/173	18,49
<i>Site</i>	16/209	7,65	<i>Mtb/mountain bike</i>	20/173	11,56
<i>Mtb/mountain bike</i>	14/209	6,69	<i>On-line/on</i>	20/173	11,56
<i>E-book</i>	13/209	6,22	<i>Site</i>	16/173	9,24
<i>VAR</i>	11/209	5,26	<i>Impeachment</i>	12/173	6,93
<i>Internet</i>	10/209	4,78	<i>Show</i>	10/173	5,78
<i>Bike</i>	6/209	2,87	<i>Top</i>	8/173	4,62
<i>Live</i>	6/209	2,87	<i>Fake</i>	5/173	2,89

<i>Baby</i>	5/209	2,39	<i>VAR</i>	5/173	2,89
<i>Boy</i>	5/209	2,39	<i>Internet</i>	4/173	2,31
<b>Total</b>	122/209	58,34	<b>Total</b>	132/173	76,27

Fonte: elaborada pela autora, (2021).

Igualmente, de modo geral, os usuários entre 45 e 54 anos recorrem com frequência a anglicismos relacionados às TICs. No primeiro período, de modo específico, destacam-se os seguintes itens: *On-line/on* (17,22%), *Site* (7,65%), *E-book* (6,22%), *VAR* (5,26%), *Internet* (4,78%) e *Live* (2,87%). Foram coletados 209 itens empregados pelos usuários desta faixa etária em P1, cujos dez vocábulos mais recorrentes equivalem a 58,34% do total dos anglicismos empregados pelos usuários da quarta faixa etária. No segundo período, de modo específico, evidenciam-se as seguintes palavras: *Live* (18,49%), *On-line/on* (11,56%), *Site* (9,24%), *Fake* (2,89%), *VAR* (2,89%) e *Internet* (2,31%). Foram localizadas 173 ocorrências de anglicismos produzidos por esta faixa etária, dentre as quais 132 estão entre os dez itens mais recorrentes, que equivalem a 96,27% do total dos anglicismos empregados pelos usuários da quarta faixa etária.

Vimos nesta subseção, que, em ambos os períodos de análise, a faixa etária com o maior número de ocorrências de anglicismos foi de 16 a 24 anos, corroborando a pesquisa realizada pelo *Global Web Index* (2017) e *Comscore* (2018), que apontou que esta faixa etária é a que mais utiliza o *Twitter*, seguida das faixas etárias: 25 e 34 anos, 25 e 44 anos e 45 e 54 anos. Apenas no primeiro período houve uma pequena divergência em que a última faixa etária (45 a 54 anos) teve mais registros de anglicismos que a faixa anterior (35 a 44 anos).

### 6.3.3 Localidade

Com o intuito de contribuir com a descrição do PB, a amostra deste trabalho é um recorte de *tweets* das microrregiões de Chapecó, Concórdia, Joaçaba e Xanxerê, localizadas no oeste de Santa Catarina. Na sequência, transcrevemos dois *tweets* que contêm anglicismos e que foram publicados em municípios da região oeste catarinense:

(35) “Dizem que a *bad*<sup>62</sup> é bem normal”, (publicado em 03 de agosto de 2020, em Ponte Serrada, grifo nosso).

<sup>62</sup> tristeza, tradução nossa.

(36) “Que gracinha os dois dando *like*<sup>63</sup> juntos”, (publicado em 09 de janeiro de 2020, em Caibi, grifo nosso).

A microrregião de Chapecó possui 37 municípios e sua população foi estimada, em 2019, em 449.161 habitantes. Não temos, contudo, uma hipótese razoável para o comportamento dos usuários dessa microrregião em relação aos anglicismos já que a rede social é aberta a postagens de usuários do mundo todo.

Durante a coleta de dados nos deparamos com a dificuldade de encontrar os 50 usuários do Twitter, com perfis abertos, dentro da microrregião de Chapecó e, por este motivo, ampliamos nossa coleta de dados para as demais microrregiões do Oeste, também reunindo dados das microrregiões de Concórdia, Joaçaba e Xanxerê. Nenhum dado foi coletado na microrregião de São Miguel do Oeste. Para confirmar que de fato os *tweets* foram escritos por indivíduos que moram/moraram na região Oeste de Santa Catarina, separamos algumas publicações com termos e expressões característicos da fala da região:

(37) “Já "**potchou**" sua bolacha maria hj ?” (publicado em 20 de abril de 2019, em Vargeão, grifo nosso).

(38) “E nós aqui na clínica, apostamos 5 **pila** cada um em um placar do jogo do grêmio x flamengo..” (publicado em 2 de outubro de 2019, em Seara, grifo nosso).

(39) “[...] me tirem uma dúvida **piazada** vontade de ficar ouvindo os áudios antigos da alemoa é sintoma de que? (publicado em 20 de agosto de 2020, em Águas Frias, grifo nosso).

(40) “[...] **crem** ma tu é carente né guria faz assim arrume uma casa pra tu **passa uma bassora** arrume um **lote pra tu carpi** vai **passar umas maquinada de veneno** vai se ocupá” (publicado em 18 de novembro de 2020, em Águas Frias, grifos nossos).

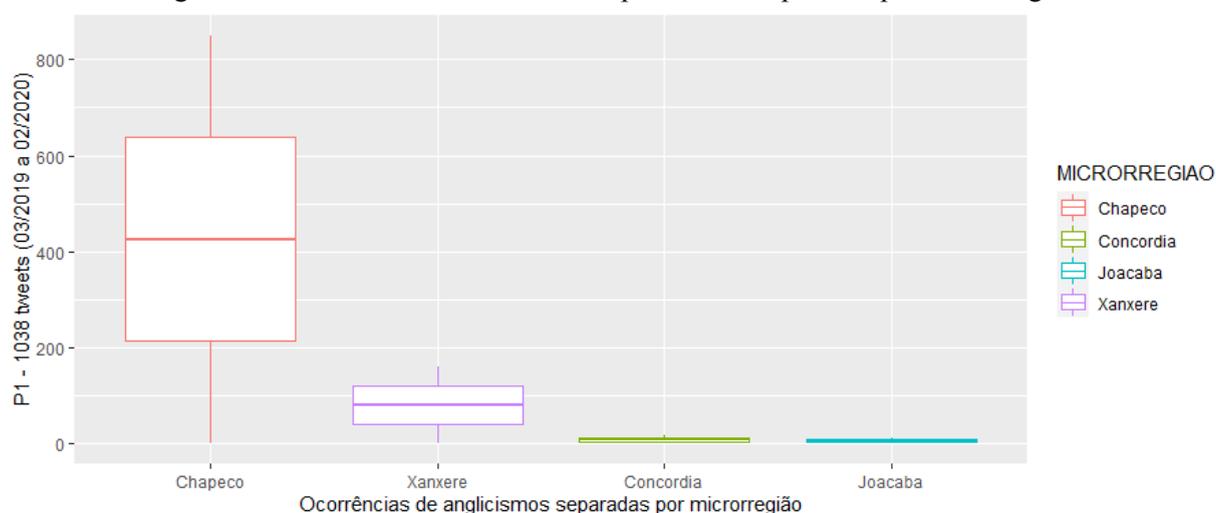
(41) “Me **kuaiando** com o Xamã” (17 de fevereiro de 2021, em Vargeão, grifo nosso).

### 6.3.3.2 Resultados e discussão

Vejamos os gráficos 15 e 16 que representam os anglicismos presentes nos tweets coletados em P1 e P2, separados por microrregião:

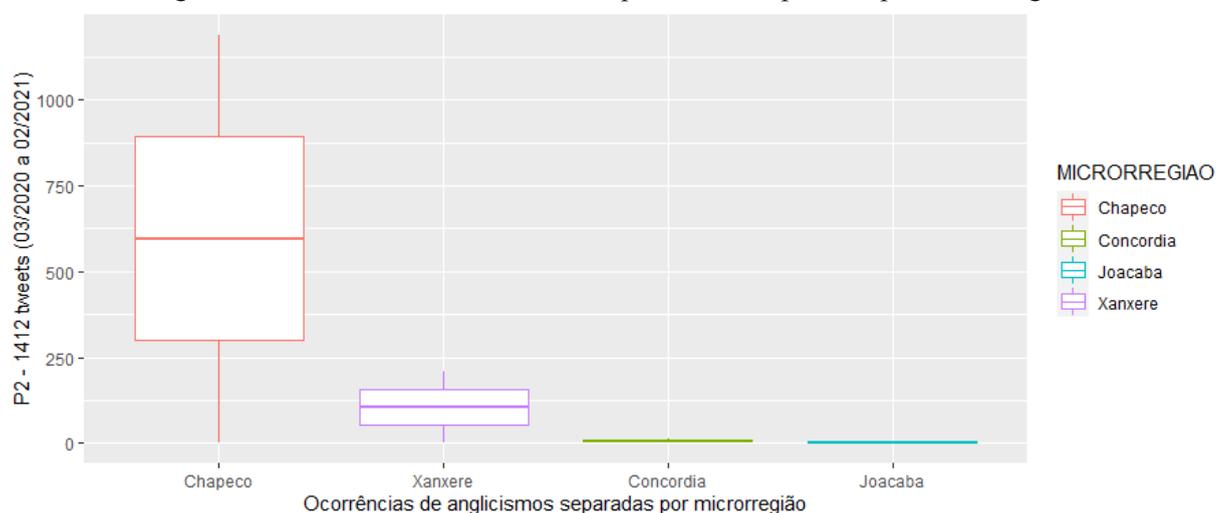
---

<sup>63</sup> curtida, tradução nossa.

Gráfico 15: Anglicismos identificados nos *tweets* do período P1 separados por microrregião.

Fonte: elaborada pela autora (2021).

No Gráfico 15, podemos observar a distribuição dos anglicismos - coletados no P1 - separados por microrregião. De modo específico, Chapecó é o local com maior destaque, registrando 848 ocorrências (81,69%) de anglicismos, depois temos a microrregião de Xanxerê, com 161 ocorrências (15,51%), em menor número temos Concórdia com 17 registros (1,63%) e Joaçaba, microrregião com menor representatividade, registrando apenas 12 ocorrências (1,15%).

Gráfico 16: Anglicismos identificados nos *tweets* do período P2 separados por microrregião.

Fonte: elaborado pela autora (2021).

Da mesma maneira, o Gráfico 16 representa os anglicismos separados por microrregião coletados no segundo período. Os dados de P2 ficaram na mesma sequência das localidades do que em P1. De modo específico, Chapecó registrou 1189/1412 (84,20%) ocorrências de anglicismos, Xanxerê representa 207/1412 registros (14,66%) e, novamente em números menos

expressivos, Concórdia teve 12/1412 dados coletados (0,84%) no P2 e, por fim, Joaçaba registrou 4/1412 ocorrências (0,28%).

Vejamos, a seguir, a Tabela 13 que apresenta os itens lexicais mais recorrentes na microrregião de Chapecó nos dois períodos estudados:

Tabela 17: Itens lexicais mais frequentes em *tweets* da microrregião de Chapecó em P1 e P2.

<b>Microrregião de Chapecó</b>					
<b>P1</b>			<b>P2</b>		
<b>Item</b>	Apl./Total	%	<b>Item</b>	Apl./Total	%
<i>On-line/On</i>	36/848	4,24	<i>Live</i>	101/1189	8,49
<i>Day</i>	33/848	3,89	<i>On-line/On</i>	47/1189	3,95
<i>Every</i>	29/848	3,41	<i>Postar</i>	41/1189	3,44
<i>Story</i>	29/848	3,41	<i>Internet</i>	38/1189	3,19
<i>VAR</i>	29/848	3,41	<i>Fake</i>	34/1189	2,85
<i>Internet</i>	27/848	3,18	<i>Impeachment</i>	34/1189	2,85
<i>Postar</i>	25/848	2,94	<i>Show</i>	31/1189	2,60
<i>Site</i>	24/848	2,83	<i>Top</i>	31/1189	2,60
<i>Fake</i>	23/848	2,71	<i>BBB</i>	30/1189	2,52
<i>Whats/wpp/zap</i>	19/848	2,24	<i>Life</i>	30/1189	2,52
<b>Total</b>	274/848	32,26	<b>Total</b>	417/1189	35,01

Fonte: elaborada pela autora, (2021).

A partir da Tabela 17, de modo geral, é possível notar que a maioria dos vocábulos com mais ocorrências têm relação com as TICs. No primeiro período, de modo específico, oito palavras têm relação com o mundo virtual: *On-line/on* (4,24%), *Story* (3,41%), *VAR* (3,41%), *Internet* (3,18%), *Postar* (2,94), *Site* (2,83%), *Fake* (2,71%) e *Whats/wpp/zap* (2,24%). Os dez itens mais recorrentes de P1 são equivalentes a 32,26% do total de dados (848) da microrregião no mesmo período. Já no segundo período analisado, de modo específico, há cinco vocábulos

relacionados às TICs: *Live* (8,49%), *On-line/on* (3,95%), *Postar* (3,44%), *Internet* (3,19%) e *Fake* (2,85%). Em P2 foram coletados 1189 itens na microrregião; destes 417 estão entre as palavras mais recorrentes, correspondendo a 35,01% do total.

Vimos nesta subseção que a microrregião de Chapecó é o local com o maior número de emprego de anglicismos nos *tweets*, registrando 848/1038 ocorrências (81,69%) no primeiro período e 1189/1412 (84,20%) no segundo períodos, seguida pelas microrregiões de Xanxerê, com 161/1038 ocorrências (15,51%) em P1 e 207/1412 (14,66%) em P2; Concórdia com 17/1038 registros (1,63%) em P1 e 12/1412 (0,84%) em P2 e, por último, Joaçaba com 12/1038 ocorrências (1,15%) em P1 e 4/1412 (0,28%) em P2.

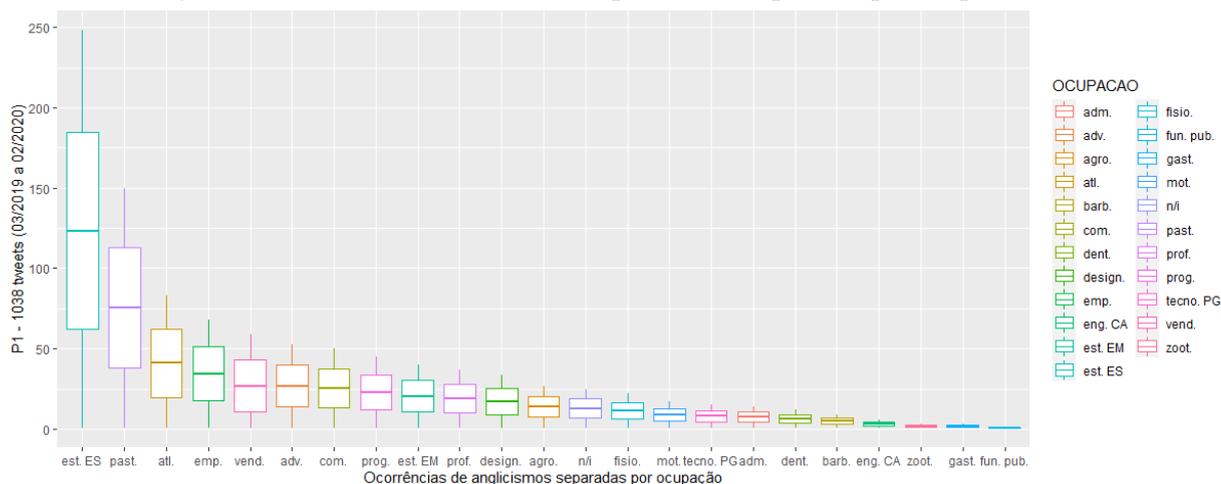
### 6.3.4 Ocupação

A variável “profissão e ocupação” foi incluída na tentativa de relacionar a incidência (ou não) do uso de anglicismos e os meios nos quais os falantes estão envolvidos. Conforme Labov (1972 [2008]), a ocupação é de extrema relevância para a definição de classe social sob o ponto de vista sociolinguístico.

Para encontrar a profissão e/ou ocupação dos falantes que compõem o *corpus*, se fez necessário um levantamento manual de tais informações a partir das contas dos usuários no *Twitter*. Não sendo possível encontrar tais dados na rede social analisada, recorreremos aos perfis dos usuários no *Instagram* e/ou *Facebook*.

#### 6.3.4.2 Resultados e discussão

Nesta subseção podemos observar a distribuição dos dados de P1 e P2 a partir da ocupação dos usuários. Vejamos na sequência o Gráfico 17:

Gráfico 17: Anglicismos encontrados nos *tweets* do período P1 separados por ocupação.

Fonte: elaborado pela autora, (2021).

De modo geral, nota-se, no Gráfico 17, que 22 ocupações foram encontradas entre os 50 usuários analisados em P1, além de 3 usuários que não informaram a ocupação. De modo específico, as cinco ocupações com maior número de registros de anglicismos em P1 são: estudantes de Ensino Superior, com 245 registros (23,60%), pastor(a), com 150 ocorrências (14,45%), atleta, representando 85 dados (8,19%), empresário(a), com 68 ocorrências (6,55%) e vendedor(a), com 65 registros (6,26%).

Tabela 18: Anglicismos coletados nos *tweets* do P1 separados por ocupação.

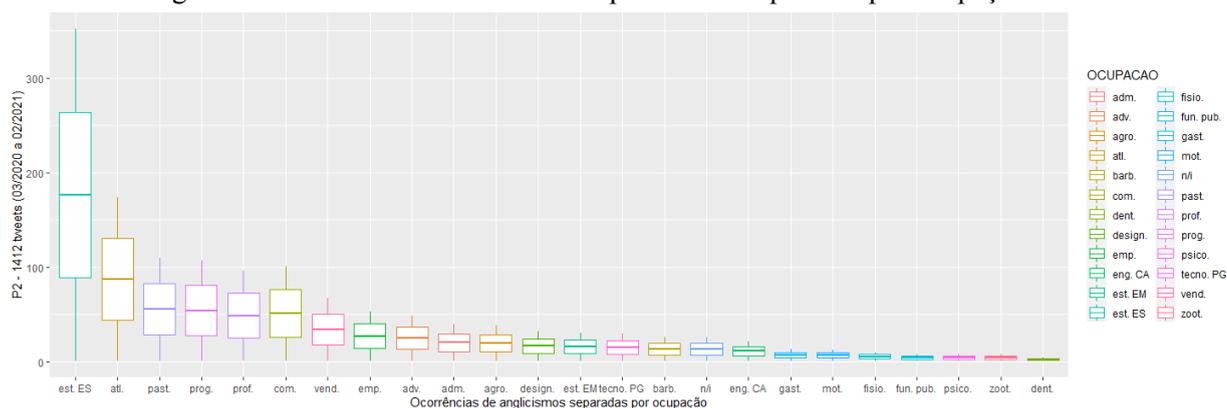
Ocupação	Apl./Total	%
Estudante ES (11 usuários)	245/1038	23,60
Pastor (1 usuário)	150/1038	14,45
Atleta (2 usuários)	85/1038	8,19
Empresário (1 usuário)	68/1038	6,55
Vendedor(a) (2 usuários)	65/1038	6,26
Advogado(a) (3 usuários)	53/1038	5,11
Estudante EM (2 usuários)	52/1038	5,01
Comunicador(a) (5 usuários)	50/1038	4,82
Programador(a) (2 usuários)	45/1038	4,33
Professor(a) (3 usuários)	37/1038	3,56
Designer (2 usuários)	34/1038	3,27
Agrônomo(a) (2 usuários)	27/1038	2,60
Não informado (3 usuários)	25/1038	2,41
Fisioterapeuta (1 usuário)	22/1038	2,12

Motorista (1 usuário)	17/1038	1,64
Tecnólogo(a) de Processos Gerenciais (1 usuário)	15/1038	1,45
Administrador (1 usuário)	14/1038	1,35
Dentista (1 usuário)	12/1038	1,16
Barbeiro (1 usuário)	9/1038	0,87
Engenheiro/a Controle e Automação (1 usuário)	6/1038	0,57
Zootecnista (1 usuário)	3/1038	0,29
Gastrônomo (1 usuário)	3/1038	0,29
Servidor público (1 usuário)	1/1038	0,10
<b>Total</b>	<b>1038/1038</b>	<b>100</b>

Fonte: elaborada pela autora, (2021).

Na Tabela 18, observamos detalhadamente as 22 ocupações encontradas em P1, sendo que as cinco com mais ocorrências (estudante ES, pastor, atleta, empresário e vendedor(a)) representam juntas 59,05% dos dados totais de anglicismos coletados no primeiro período. As outras 17 ocupações representam 38,54% dos itens lexicais de P1 e, por último, os usuários que não informaram (três pessoas) a ocupação representam 2,41% das ocorrências de anglicismos no primeiro período.

Gráfico 18: Anglicismos encontrados nos *tweets* do período P2 separados por ocupação.



Fonte: elaborado pela autora, (2021).

De modo geral, vemos, no Gráfico 18, que foram registradas 23 ocupações<sup>64</sup> no segundo período, com exceção dos que não informaram a ocupação (três usuários). De modo específico, as cinco ocupações com maior emprego de anglicismos são: estudante de Ensino Superior, com 352 registros (24,92%), atleta, com 174 ocorrências (12,32%), pastor, representando 110 dados

<sup>64</sup> Psicólogo/a não teve registros de anglicismos no primeiro período analisado.

(7,79%), programador(a), com 107 ocorrências (7,58%) e, por fim, comunicador(a), com 101 registros (7,15%).

Tabela 19: Anglicismos coletados nos *tweets* do P2 separados por ocupação.

Ocupação	Apl./Total	%
Estudante IES (10 usuários)	352/1412	24,92
Atleta (2 usuários)	174/1412	12,32
Pastor (1 usuário)	110/1412	7,79
Programador(a) (2 usuários)	107/1412	7,58
Comunicador(a) (4 usuários)	101/1412	7,15
Professor(a) (4 usuários)	96/1412	6,80
Vendedor(a) (2 usuários)	67/1412	4,74
Empresário ( 1 usuário)	53/1412	3,75
Advogado(a) (3 usuários)	48/1412	3,40
Administrador (1 usuário)	39/1412	2,76
Agrônomo(a) (2 usuários)	38/1412	2,69
Designer (1 usuário)	32/1412	2,26
Tecnólogo de Processos Gerenciais (1 usuário)	29/1412	2,05
Barbeiro (1 usuário)	26/1412	1,84
Não informado (3 usuários)	26/1412	1,84
Engenheiro Controle e Automação (1 usuário)	21/1412	1,49
Estudante EM (2 usuários)	30/1412	2,12
Gastrônomo (1 usuário)	13/1412	0,92
Motorista (1 usuário)	12/1412	0,85
Fisioterapeuta (1 fisioterapeuta)	10/1412	0,70
Zootecnista (1 usuário)	8/1412	0,56
Psicólogo (1 usuário)	8/1412	0,56
Funcionário público (1 usuário)	8/1412	0,56
Dentista (1 usuário)	4/1412	0,28
<b>Total</b>	<b>1412/1412</b>	<b>100</b>

Fonte: elaborada pela autora, (2021).

De modo específico, podemos observar, na Tabela 19, todas as 23 profissões registradas no segundo período analisado. As cinco ocupações (estudante ES, atleta, pastor, programador(a) e professor(a)) retratam 59,76% dos dados de anglicismos coletados no P2. As outras 18 ocupações encontradas somam 38,40% e os indivíduos (três pessoas) que não informaram a ocupação representam apenas 1,84% dos dados.

No que se refere às ocupações, é importante salientar que há entre os 50 usuários: 11 estudantes de Ensino Superior, 5 comunicadores, 4 professores, 3 advogados, 2 agrônomos, 2 atletas, 2 designers, 2 estudantes de Ensino Médio, 2 programadores, 2 vendedores, além de 3 usuários que não conseguimos obter esta informação. As demais ocupações são representadas por um usuário cada.

Nesta subseção, identificamos que os homens empregaram mais anglicismos mais que as mulheres no primeiro e no segundo período. Sobre a relação entre a idade dos usuários e o uso de anglicismos, identificamos que a faixa etária de 16 a 24 anos representa 418/1038 dados (40,27%) dos anglicismos obtidos no P1, seguida pelas faixas etárias de 25 a 34 anos totaliza 310/1038 ocorrências (29,86%); 35 a 44 anos registra 101/1038 dados (9,73%) e, por último, 45 a 54 anos representa 209/1038 dados (20,13%) dos dados totais do primeiro período analisado. Já em P2 o uso de anglicismos sofre uma pequena mudança, sendo que, novamente, a faixa etária com maior número de ocorrências é a de 16 a 24 anos com 503/1412 (35,62%), seguida pelas faixas etárias: 25 a 34 anos, que registra 51/1412 ocorrências, (36,33%); 35 a 44 anos, representando 223/1412 itens lexicais (15,79%), e, por fim, 45 a 54 anos, com o registro de 173/1412 dados (12,25%).

Com relação à localidade, veremos na sequência, os dados obtidos em P1 e P2, respectivamente. A microrregião de Chapecó é o local com maior destaque, registrando 848/1038 ocorrências (81,69%) no primeiro período e 1189/1412 (84,20%) em P2, seguida pelas microrregiões de Xanxerê, com 161/1038 ocorrências (15,51%) em P1 e 207/1412 (14,66%) em P2; Concórdia com 17/1038 registros (1,63%) em P1 e 12/1412 (0,84%) em P2 e, por último, Joaçaba com 12/1038 ocorrências (1,15%) em P1 e 4/1412 (0,28%) em P2. Por fim, no que se refere às 22 ocupações com ocorrências no primeiro período, destacaram-se: estudantes de Ensino Superior, com 245 registros (23,60%); pastor, com 150 ocorrências (14,45%); atleta, representando 85 dados (8,19%); empresário, com 68 ocorrências (6,55%) e, por último, vendedor(a), com 65 registros (6,26%). Já no segundo período, as 23 ocupações com maior registro de anglicismos foram: estudante de Ensino Superior, com 352 registros (24,92%); atleta,

com 174 ocorrências (12,32%); pastor, representando 110 dados (7,79%); programador(a), com 107 ocorrências (7,58%) e, por fim, comunicador(a), com 101 registros (7,15%).

## 7. ORIGEM E POTENCIALIDADE SEMÂNTICO-PRAGMÁTICA DE ALGUNS ANGLICISMOS PRESENTES EM *TWEETS*

Nesta seção tratamos da origem, segundo os dicionários Aurélio e Houaiss, e da potencialidade semântico-pragmática de alguns dos anglicismos encontrados em *tweets* da amostra analisada. Também exemplificamos seus usos a partir de *tweets* de usuários da rede social *Twitter* localizados na região oeste catarinense.

Em termos funcionais, muitos desses vocábulos originários do inglês como verbos, por exemplo, podem manter ou mudar de categoria gramatical, bem como podem estar sendo usados com o mesmo sentido ou com significado diverso da língua-fonte. Bechara (2009, p. 501) chama a atenção “para o fato de que o inglês vai constantemente ao grego e ao latim buscar-lhes os tesouros, mas os usa com tal liberdade, que muitas vezes deturpa o emprego ou o significado original.”. Em razão disso, investigamos a origem, segundo os dicionários Aurélio e Houaiss, e a potencialidade semântico-pragmática dos anglicismos encontrados nos *tweets* analisados. Os anglicismos que não estão dicionarizados, em PB, foram analisados a partir dos dicionários Oxford e Cambridge da língua inglesa.

A seguir, apresentamos dois exemplos de anglicismos em *tweets* que mantêm o sentido e outros dois que mudam o sentido no uso em português, respectivamente:

(42) “Maaano, como que vcs[vocês] conseguem mais amei do que *like*<sup>65</sup> nas foto do face ??????”, (publicado em 9 de janeiro de 2021, em Chapecó, grifo nosso).

(43) “Quería ouvir o melhor *podcast* mas aparentemente a *internet* tá de *troll*<sup>66</sup> cmg [comigo], (publicado em 27 de maio de 2020, em Xanxerê, grifo nosso).

(44) “[...] vai parar de me *stalkear*<sup>67</sup> e me chamar qndoo [quando]?”, (publicado em 24 de agosto de 2019, em Herval d’Oeste, grifo nosso).

(45) “Eu não consigo entender nada nessas aulas no *meet*<sup>68</sup>”, (publicado em 22 de maio de 2020, em Palmitos, grifo nosso).

Bybee (2020) destaca que, além de objetos novos na cultura, alguns empréstimos não têm motivação linguística de necessidade, tendo em vista que as palavras são importadas mesmo

<sup>65</sup> curtida, tradução nossa.

<sup>66</sup> sacanagem, tradução nossa.

<sup>67</sup> espionar, tradução nossa.

<sup>68</sup> encontro, tradução nossa. Neste caso, o usuário está se referindo à plataforma *Google Meet*.

quando a língua receptora já tem uma palavra para o objeto ou o conceito. São exemplos desse tipo os empréstimos como *business* (negócio), *smart* (esperto), *game* (jogo) ou *party* (festa). A seguir, apresentamos duas ocorrências de *tweets* contendo anglicismos que não têm motivação linguística de necessidade:

(46) “Tinha um *site* pra assistir os **games**<sup>69</sup> da NBA agora deu pau, alguém ai com alguns links pra salvar?”, (publicado em 12 de janeiro de 2021, em Xaxim, grifo nosso).

(47) “Dps[depois] da pandemia **pool party**<sup>70</sup> aqui em casaaaaa p[a]ra todos os meus seguidoreess qm[quem] n[não] vir eu vo[vou] fica[ficar] bolado”, (publicado em 10 de janeiro de 2021, em Chapecó, grifo nosso).

A partir dos 2450 *tweets* da amostra, observamos que os itens lexicais coletados, em termos funcionais, se mantêm na classe gramatical de origem, sem alterações significativas ao passarem a ser utilizadas na ciberescrita em PB. Também notamos que nossa hipótese de que os vocábulos relacionados às TICs, em geral, mantiveram o significado da língua-fonte. Por último, verificamos que muitos tuiteiros preferiram utilizar no lugar de termos equivalentes em PB, como no caso da utilização da palavra “*selfie*” ao invés de “fotografia” ou “retrato”. Na sequência, extraímos da nossa amostra dois exemplos de substantivos e dois exemplos de adjetivos, respectivamente, relacionados às TICs direta ou indiretamente:

(48) “É real cara eu não sei mais fazer **selfie** sozinho.”, (publicado em 7 de maio de 2020, em São Carlos - grifos nossos).

(49) “[...] queria que existisse o **download** de leituras para o cérebro”, (publicado em 28 de setembro, em Chapecó - grifos nossos).

(50) “[...] juro q li a primeira parte q estava “**cropada**””, (publicado em 31 de agosto de 2020, em Chapecó - grifos nossos).

(51) “**Tiltado** as 8h da manhã já”, (publicado em 1 de agosto de 2020, em Vargeão - grifos nossos).

Identificamos, nos *tweets*, alguns verbos - e suas derivações, como *to bug* (9 ocorrências), *to chat* (2 ocorrências), *to like* (9 ocorrências), *to meet* (3 ocorrências), *to stalk* (12 ocorrências) e

<sup>69</sup> jogos, tradução nossa.

<sup>70</sup> festa na piscina, tradução nossa.

*to troll* (6 ocorrências). É importante destacar que a partícula ‘*to*’ aparece apenas nos verbos em sua forma infinitiva, mas, para buscá-los nos dicionários é preciso removê-la.

*To bug*: no dicionário Oxford, o verbo “*to bug*” significa “gravar ou ouvir (uma conversa) usando um microfone escondido<sup>71</sup>” (tradução nossa). O termo também é encontrado como substantivo “*bug*” e significa inseto, por outro lado, no dicionário Cambridge, o termo também é encontrado como um substantivo com o conceito de um “erro ou problema na forma como algo funciona<sup>72</sup>”. Em português, o vocábulo é encontrado como um substantivo masculino nos dicionários Aurélio e Houaiss, derivado da informática. No dicionário Aurélio, o conceito é de “erro de programação ou defeito de execução de um programa (geralmente causado por inconsistência no seu código ou incompatibilidade com outros programas, que estejam simultaneamente em execução)” (FERREIRA, 2004, p. 336). Já o dicionário Houaiss informa que a origem do vocábulo é de 1622, mas desconhecida, e significa “defeito, falha ou erro na lógica ou no código de um programa que provoca seu mau funcionamento” (HOUAISS; VILLAR, 2008, p. 526). O Houaiss define esse termo também como “inseto tipo percevejo, piolho”; por extensão pejorativa “defeito ou imperfeição inesperada”; ou ainda por metonímia<sup>73</sup> “o que suga, mantém-se escondido ou disfarçado para causar dano ou obter algo por meio ilícito”. Vejamos a ocorrência a seguir:

(52) “[...] meu *insta*[Instagram] é todo errado, fui mandar um *stories buguei*<sup>74</sup> tudo” (publicado em 17 de novembro de 2019, em Chapecó, grifo nosso)

Observa-se que o termo “bug” possui significados divergentes quando os dicionários o conceituam como verbo ou como substantivo. O significado empregado pela informática para o PB é o do substantivo que significa erro ou defeito, geralmente, de um *software*. Porém, pode-se visualizar, no exemplo extraído do *Twitter*, que o usuário da rede social usou o termo como verbo, aportuguesando a conjugação do vocábulo (buguei). No contexto da publicação, compreende-se que o usuário quis dizer que ele travou (deu erro) no aplicativo do *Instagram*.

I. *To chat*: na versão digital do dicionário Oxford, o verbo “*to chat*” é definido como a ação de “conversar amigavelmente de um jeito informal ou trocar mensagens *on-line* em tempo

<sup>71</sup> *To bug*: [verb] record or listen to (a conversation) using a concealed microphone; [noun] a small insect. (Dicionário Oxford aplicativo para iOS)

<sup>72</sup> *Bug*: [noun] a mistake or problem in the way something works. (Dicionário Cambridge On-line)

<sup>73</sup> Metonímia:[substantivo feminino]: figura de retórica que consiste no uso de uma palavra fora do seu contexto semântico normal, por ter ela significação com alguma relação objetiva, de contiguidade, material ou conceitual, com o conteúdo ou o referente ocasionalmente pensado. (HOUAISS; VILLAR, 2008, p. 1911)

<sup>74</sup> travei, tradução nossa.

real com um ou mais usuários simultâneos de uma rede de computadores<sup>75</sup> (tradução nossa). No dicionário Cambridge também encontram-se conceitos neste mesmo sentido de que o termo “to chat” é “falar com alguém de um jeito amigável ou comunicar-se com alguém na internet enviando mensagens que você pode ler e responder imediatamente<sup>76</sup>” (tradução nossa). Já em português, “*chat*” é encontrado como um substantivo masculino com o significado de “forma de comunicação através de rede de computadores, similar a uma conversação, na qual se trocam, em tempo real, mensagens escritas; bate-papo *on-line*, bate-papo virtual, papo *on-line*, papo virtual” (FERREIRA, 2004, p. 452). O dicionário Houaiss conceitua “*chat*” como um “forma de comunicação a distância, utilizando computadores ligados à internet, na qual o que se digita no teclado de um deles aparece em tempo real no vídeo de todos os participantes do bate-papo; etimologia inglesa (1530) ‘falar, emitir sons semelhantes ao da fala, de modo rápido, indistinto ou pouco articulado; conversar de modo informal’, de origem expressiva imitativa” (HOUAISS; VILLAR, 2008, p. 694), classificando-o também como substantivo masculino. A seguir exemplificamos o uso do termo no *Twitter*:

(53) “A maturidade no *chat*<sup>77</sup> da aula *online*: não tem KKKK” (publicado em 24 de março de 2020, em Chapecó, grifo nosso).

Apesar de o termo “*chat*” ser classificado como verbo em inglês e como substantivo em português, os conceitos encontrados são similares, ambos relacionados com conversas informais e/ou virtuais simultâneas. No exemplo extraído do *Twitter*, o usuário usa como um substantivo e, pelo contexto do *tweet*, pressupõe-se que ele está falando sobre o bate-papo de uma aula remota.

II. *To like*: o dicionário Oxford descreve a expressão verbal “*to like*” com definição de achar [algo] “agradável, divertido ou satisfatório<sup>78</sup>” (tradução nossa); o dicionário apresenta ainda o conceito a partir do contexto das redes sociais, que indica a “aprovação ou apoio (a alguém ou alguma coisa) através de um ícone ou *link* específicos<sup>79</sup>” (tradução nossa). No dicionário Cambridge, “*to like*” aparece como “curtir ou aprovar algo ou alguém<sup>80</sup>” ou ainda para “mostrar que você acha que algo é bom em um *site* de rede social clicando em um símbolo ou na

<sup>75</sup> *To chat* [verb.]: talk in a friendly and informal way; exchange messages online in real time with one or more simultaneous user of a computer network. (Dicionário Oxford aplicativo para iOS)

<sup>76</sup> *To chat* [verb]: to talk with someone in a friendly way; to communicate with someone on the internet by sending messages that you can read and reply to immediately. (Dicionário Cambridge *On-line*)

<sup>77</sup> bate-papo virtual, tradução nossa.

<sup>78</sup> *To like*: [verb] find agreeable, enjoyable, or satisfactory. (Dicionário Oxford aplicativo para iOS)

<sup>79</sup> *To like*: [in the context of social media] indicate one's approval of or support for (someone or something) by means of a particular icon or link. (Dicionário Oxford aplicativo para iOS)

<sup>80</sup> *To like*: [verb] to enjoy or approve of something or someone. (Dicionário Cambridge *On-line*)

palavra ‘*like*’<sup>81</sup> (tradução nossa). O termo “*to like*” não é encontrado nos dicionários Aurélio e Houaiss da língua portuguesa. Na sequência pode-se verificar um exemplo extraído do *Twitter*:

(54) “[...] se inscrevam e deem aquele *like*<sup>82</sup> p[a]ra fortalecer a cena” (publicado em 20 de maio de 2020, em Chapecó, grifo nosso).

No que se refere ao termo “*to like*” vemos que, apesar de não estar dicionarizado em língua portuguesa, é utilizado pelo usuário do *Twitter* sem que a tradução seja necessária. No *tweet* em questão o vocábulo “*like*” é usado no sentido de curtir/gostar.

III. *To meet*: no dicionário Oxford o termo “*to meet*” aparece como um verbo com o significado de “encontrar ou por acaso estar na presença ou companhia de alguém; encontrar ou experimentar uma situação ou atitude específicas<sup>83</sup>” (tradução nossa). Ainda sobre o conceito de “*to meet*”, o termo é encontrado no dicionário Cambridge sob a definição de “ver ou falar com alguém pela primeira vez; sair junto com alguém intencionalmente; marcar um encontro com uma pessoa ou grupo de pessoas para falar com eles sobre algo<sup>84</sup>” (tradução nossa). O vocábulo “*to meet*” não foi encontrado nos dicionários de língua portuguesa, Aurélio e Houaiss. Na sequência, pode-se observar uma amostra da palavra “*to meet*”, encontrada neste estudo:

(55) “Casa limpa, roupa lavada, incenso no sal grosso, sentado para finalizar o último trabalho de interiores do semestre, tomando meu café, no *meet*<sup>85</sup> com o amigo que eu gosto. Paz” (publicado em 3 de dezembro de 2020, em Chapecó, grifo nosso).

Quanto ao vocábulo “*to meet*”, vemos que ambos os dicionários da língua inglesa apresentam conceitos semelhantes e, apesar de o vocábulo não estar dicionarizado em língua portuguesa, podemos ver no exemplo que o termo é utilizado sem a necessidade de tradução, mas como um substantivo no sentido de encontro/reunião. Não fica claro se no *tweet* o usuário se refere a um encontro presencial ou virtual.

---

<sup>81</sup> *To like*: to show that you think something is good on a social networking website by clicking on a symbol or the word ‘like’. (Dicionário Cambridge On-line)

<sup>82</sup> curtir, tradução nossa.

<sup>83</sup> *To meet*: [verb] arrange or happen to come into the presence or company (someone); encounter or experience (a particular situation or attitude). (Dicionário Oxford Aplicativo para iOS)

<sup>84</sup> *To meet*: [verb] to see and talk to someone for the first time; to come together with someone intentionally; to arrange to see a person or group of people in order to speak with them about something. (Dicionário Cambridge On-line)

<sup>85</sup> encontro, tradução nossa.

IV. *To stalk*: O termo “*to stalk*” é encontrado no dicionário Oxford sob a denominação “perseguir ou se aproximar furtivamente<sup>86</sup>” (tradução nossa). Já no dicionário Cambridge o verbo “*to stalk*” tem o significado de “seguir um animal ou pessoa o mais próximo possível, sem ser visto ou ouvido, geralmente para capturá-los ou matá-los; seguir ilegalmente e observar alguém durante um período de tempo<sup>87</sup>” (tradução nossa). O vocábulo “*to stalk*” não está inserido nos dicionários de língua portuguesa Houaiss e Aurélio. Na sequência, podemos ver um *tweet* em que o termo é empregado:

(56) “Fui **stalkear**<sup>88</sup> e vi o que não queria kkkkkkk de novo” (publicado em 19 de novembro de 2019, em Chapecó, grifo nosso).

No exemplo do uso do “*to stalk*” retirado do *Twitter* percebe-se que, independentemente de o termo não estar dicionarizado na língua portuguesa, ele foi usado pelo usuário sem necessidade de tradução. Nota-se também que o vocábulo foi aportuguesado, foram acrescentadas à palavra as letras “e”, “a” e “r” formando o neologismo “**stalkear**”. Nesse caso, o termo foi utilizado no sentido de observar/monitorar algo ou alguém.

V. *To troll*: No dicionário Cambridge o vocábulo “*to troll*” significa “deixar uma mensagem insultuosa ou ofensiva na internet para incomodar alguém, ou para chamar a atenção ou causar problemas; fazer ou dizer intencionalmente algo irritante ou ofensivo para aborrecer alguém<sup>89</sup>” (tradução nossa). O dicionário Oxford conceitua o termo “*to troll*” como “fazer uma ofensa deliberada ou provocativa em uma publicação on-line com o objetivo de chatear alguém ou provocar uma resposta raivosa<sup>90</sup>” (tradução nossa). O verbo “*to troll*” não é encontrado nos dicionários Aurélio e Houaiss da língua portuguesa. Adiante podemos observar um exemplo do uso do termo “*to troll*” em PB:

(57) “[...] fui tentar **trollar**<sup>91</sup> o Lucas mas o praga é esperto kkkk” (publicado em 1º de abril de 2019, em Cordilheira Alta, grifo nosso).

<sup>86</sup> *To stalk*: [verb] pursue or approach stealthily. (Dicionário Oxford Aplicativo para iOS)

<sup>87</sup> *To stalk*: [verb] to follow an animal or person as closely as possible without being seen or heard, usually in order to catch or kill them; to illegally follow and watch someone over a period of time. (Dicionário Cambridge On-line)

<sup>88</sup> bisbilhotar, tradução nossa.

<sup>89</sup> *To troll*: [verb] (informal) to leave an insulting or offensive message on the internet in order to upset someone, or to get attention or cause trouble; to intentionally do or say something annoying or offensive in order to upset someone. (Dicionário Cambridge On-line)

<sup>90</sup> *To troll*: [verb] make a deliberately offensive or provocative online post with the aim of upsetting someone or eliciting an angry response. (Dicionário Oxford Aplicativo para iOS)

<sup>91</sup> zoar, tradução nossa.

No *tweet* com o exemplo do verbo “*to troll*” podemos ver que o usuário da rede social aportuguesou o termo adicionando as letras “a” e “r”, remetendo aos verbos no infinitivo em português. Não houve necessidade de tradução da palavra e ela foi usada no sentido de enganar/ludibriar.

Vimos nesta seção que alguns anglicismos, principalmente os relacionados às TICs, já estão inseridos de forma bastante natural na ciberescrita dos tuiteiros. Dentre os exemplos utilizados nesta parte, apenas um usuário utilizou aspas para identificar o anglicismo aportuguesado: “*cropada*”, indicando que ele sabe que é uma palavra ainda não-dicionarizada em PB. Observamos também, a partir da nossa amostra que os vocábulos coletados, em termos funcionais, se mantêm na classe gramatical de origem, sem alterações significativas ao passarem a ser utilizadas na ciberescrita do PB, mostrando uma possível preferência de uso destes anglicismos pelos usuários do *Twitter*.

## 8. CONSIDERAÇÕES FINAIS

Esta pesquisa teve como objetivo descrever e analisar os anglicismos empregados em uma amostra sincrônica de *tweets* da rede social *Twitter* escritos em PB por usuários localizados na região oeste catarinense, tendo como *corpus* as publicações de 50 usuários da rede, separados em dois períodos.

A primeira questão levantada pretendeu identificar quais anglicismos emergiram do *corpus* da pesquisa, especialmente durante o período anterior e posterior à declaração da pandemia de Covid-19. Nossa hipótese é que os anglicismos relacionados ao campo das Tecnologias da Informação e Comunicação (TICs) seriam os mais recorrentes, o que foi confirmado durante a análise dos dois períodos coletados. Para este estudo, foram coletados 32.302 *tweets*, dos quais 14.721 (45,56%) são relativos ao primeiro período (março de 2019 a fevereiro de 2020) e 17.591 (54,44%) referentes ao segundo período (março de 2020 a fevereiro de 2021), confirmando a hipótese do aumento do uso das redes sociais durante a pandemia da Sars-Cov-2. Do total de anglicismos coletados, 2450 ocorrências foram selecionadas a partir da mediana de cada faixa etária, separadas por sexo e, posteriormente analisados. Em nosso *corpus*, identificamos ao todo 437 diferentes itens lexicais nos *tweets* nos dois períodos. No primeiro período foram coletados 1038 *tweets* (42% da amostra), que apresentaram com maior frequência os seguintes vocábulos: *postar* (3,55%), *on-line* (3,46%), *day* (3,17%), *internet* (3,07%), *story* (2,88%), *insta* (2,69%), *whats* (2,59%), *fake* (2,31%), *top* (2,02%) e *show* (1,92%), que juntos correspondem a 27% da amostra de P1. Já no segundo período foram coletados 1412 *tweets* (58% da amostra), que exibiram como mais recorrentes os seguintes itens lexicais: *live* (8,70%), *on-line* (4,10%), *postar* (3,74%), *internet* (2,97%), *fake* (2,40%), *insta* (2,33%), *show* (2,26%), *top* (2,26%), *story* (2,12%) e *site* (1,70%), correspondendo a 32% da amostra total de P2.

Já a segunda questão aspirava identificar propriedades formais e funcionais caracterizadoras dos anglicismos presentes nos *tweets*. Como hipótese, apresentava-se o fato de que os anglicismos não dicionarizados no português seriam mais recorrentes que os dicionarizados, o que também se confirmou durante o decorrer desta pesquisa.

Como terceiro questionamento, procuramos saber quais variáveis linguísticas e extralinguísticas atuaram no uso dos anglicismos presentes nos *tweets* durante os períodos analisados. Das sete variáveis independentes controladas, o programa estatístico selecionou como significativas apenas três: as variáveis linguísticas assimilação ou não dos termos estrangeiros ao sistema linguístico do português e classe morfológica dos anglicismos e a

variável social faixa etária dos usuários do Twitter da amostra analisada. Dentre os condicionantes linguísticos, destacamos a classe morfológica, pressupondo de que os substantivos teriam maior recorrência, seguidos dos adjetivos e verbos, o que se confirmou durante a análise. Os resultados da análise quantitativa mostraram que a classe morfológica mais recorrente foi a dos substantivos, seguida pelos adjetivos em ambos os períodos. Identificamos também que os vocábulos não dicionarizados tiveram maior frequência tanto em P1, quanto em P2. Constatamos ainda que, nos dois períodos, os itens lexicais não assimilados foram os mais recorrentes, o que confirmou nossa hipótese de que encontraríamos a maior parte dos itens lexicais relacionados às TICs ainda não assimilados nos *tweets*, isto é, ainda como estrangeirismos. Vimos também que a proporção de neologismos é baixa em relação aos empréstimos e estrangeirismos da amostra. Dentre os condicionantes extralinguísticos, controlamos o sexo, a faixa etária, a localidade e a ocupação. Os resultados da análise quantitativa mostraram que homens apresentaram maior ocorrência de anglicismos em relação às mulheres, o que confirmou nossa expectativa. Sobre a faixa etária, levantamos a hipótese de que os usuários mais jovens da rede social utilizariam mais anglicismos dos que os usuários mais velhos, resultado que se confirmou por haver mais ocorrência de anglicismos nas primeiras duas faixas etárias (16-24 anos e 25-34 anos). Em relação à localidade e à ocupação, não tínhamos uma hipótese razoável para o comportamento dos tuiteiros a partir destes condicionantes, porém destacamos aqui que foi mais fácil o acesso aos dados dos informantes da microrregião de Chapecó e, por isso, há um número maior de *tweets* deste local na amostra. Já sobre a ocupação, frisamos que a ocupação com maior ocorrência de *tweets* é de estudantes de ensino superior. Por fim, na quarta questão, ansiamos saber em que medida a entrada de anglicismos novos na ciberescrita durante o período da pandemia de Covid-19 permitiria também avistar inovações no PB. Acreditamos na hipótese de que novos anglicismos seriam incorporados ao PB escrito, tendo em vista a pandemia que obrigou as pessoas a migrarem ao mundo virtual para realizarem atividades do cotidiano. Também pressupomos que os anglicismos seriam preferidos no lugar de palavras equivalentes no PB. Ambas as hipóteses, sobre o quarto questionamento, foram confirmadas a partir da análise do *corpus* coletado.

Conforme Labov (2008 [1972]), uma pesquisa empírica sobre a mudança linguística em andamento pode ser subdividida em cinco problematizações distintas - transição, encaixamento, avaliação, fatores condicionantes e implementação - para as quais uma teoria da mudança deve fornecer respostas. No caso do problema de transição destacamos a palavra “*whatsapp*”, que coexiste com as versões “*whats*”, “*wpp*”, “*zap*” e outras, todas referindo-se ao mesmo aplicativo,

sendo a primeira a nomenclatura original. Levando em conta que o problema de avaliação se dá a partir do significado social da mudança, se é de prestígio ou não, conforme Salomão (2011), podemos exemplificá-lo também através do vocábulo “*whatsapp*” e suas variações. Já quanto ao problema do encaixamento, destacamos a correlação entre o uso de anglicismos em *tweets* e a faixa etária dos usuários, sendo mais recorrentes o uso entre os mais jovens, como confirmamos em nossa amostra. Em nossa pesquisa utilizamos os fatores sociais condicionantes como a faixa etária, o sexo e a ocupação. E, por fim, sobre o problema de implementação, exemplificamos com o vocábulo “*internet*”, que já está dicionarizado, e a palavra “*live*” (a mais recorrente em P2), forma nova, que vem se destacando a partir da declaração da pandemia de Covid-19 e, por este motivo, não se trata de uma mudança completada, ainda está em curso.

Das dificuldades encontradas durante o percurso desta pesquisa, destacamos a coleta de dados dentro do *Twitter*. Apesar de existir a ferramenta de busca avançada na rede, há aspectos que dificultam a seleção da amostra, como: encontrar perfis abertos e com uma frequência mínima de publicações; encontrar a ocupação dos usuários também é um fator dificultante porque nem todos colocam sua profissão na descrição do perfil, sendo necessário em alguns casos recorrer a outras redes sociais como o *Instagram* e *Facebook*. A busca pelos anglicismos também demanda bastante tempo por ser um processo que exige a leitura de cada *tweet*, sem automatização desta parte do processo. Por ser um processo que demanda muito tempo, foi preciso optar por apenas 50 usuários do *Twitter*, o que gera uma margem de erro de 13,85%, para que a confiabilidade fosse de 95%, seria necessário coletar *tweets* de 354 usuários, gerando, em média, 228.698 *tweets*, o que seria impossível de analisar em tempo hábil.

Destacamos a importância de um estudo sobre a ocorrência de anglicismos pós pandemia para saber se houve um aumento deste fenômeno também na fala do Oeste catarinense. Salientamos também que outros fenômenos podem ser estudados a partir do *corpus* coletado para esta pesquisa, como o caso do apagamento do R dos verbos no infinitivo, que observamos com certa frequência nos *tweets* da amostra.

Observamos também a possibilidade de analisar outras amostras de escrita e fala para verificar se estes novos anglicismos, que passaram a ser utilizados durante a pandemia e como os falantes e suas comunidades vão se comportar com relação a estes novos itens lexicais ao longo do tempo.

Por fim, destacamos a importância deste levantamento de dados para futuras atualizações dos dicionários de PB pelos lexicógrafos, além de também ser relevante para o ensino do português como língua materna e estrangeira.

## REFERÊNCIAS

ALMEIDA, Sandy Tavares de; AZEVEDO, Ana Claudia Oliveira; PEREIRA, Marcia Helena de Melo. **Variação linguística em um perfil institucional do Twitter: (in)adequação à situação sociocomunicativa?** XIII Colóquio Nacional - VI Colóquio Internacional do Museu Pedagógico. Salvador: Universidade Estadual do Sudoeste da Bahia (UESB), 2019.

ALVES, Ieda Maria. **Neologismo: criação lexical**. 3. ed. São Paulo: Ática, 2007.

AMARAL, Luís I. C. **A Concordância Verbal de Segunda Pessoa do Singular em Pelotas e suas Implicações Linguísticas e Sociais**. 181 f. Tese (Doutorado em Letras) – Programa de Pós-Graduação em Letras, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, 2003.

ARAÚJO, Glauco. **Assembleia Legislativa do RS aprova lei que proíbe uso de 'estrangeirismo'**. In: G1, São Paulo: 2011. Disponível em: <http://g1.globo.com/brasil/noticia/2011/04/assembleia-legislativa-do-rs-aprova-lei-que-proibe-us-o-de-estrangeirismo.html>. Acesso em: 31 jan. 2021.

ARGENTA, Denise (org.). **50 Anos depois: inventário e salvaguarda do patrimônio imaterial de Nova Erechim**. Pinhalzinho, SC: Museu Histórico de Pinhalzinho, 2014.

ARRAES, Flávia Cristina Cruz Lamberti. **Empréstimos linguísticos do inglês, com formativos latinos, adotados pelo português do Brasil**. Brasília: Departamento de Linguística, Línguas Clássicas e Vernáculas (LIV), Universidade de Brasília (UNB), 2006.

AULETE, Francisco Júlio de Caldas. **Novíssimo dicionário contemporâneo da língua portuguesa**. [Organizador: GEIGER, Paulo]. Rio de Janeiro: Lexicon, 2011.

BARROS, Cláudia Graziano Paes de. **Letramento Digital**. Considerações sobre a leitura e a escrita na internet. v. 12, n. 1. Cuiabá: Revista Polifonia, Universidade Federal de Mato Grosso (UFMT), 2006. p. 133-156.

BAGNO, Marcos. **Gramática Pedagógica do Português Brasileiro**. São Paulo: Parábola Editorial, 2012.

BATTISTI, Elisa. SANTOS, Bruna Silva dos. GUTERRES, Rodrigo Lerner. TEIXEIRA, Thomaz Torres. CUNHA, Victória Goulart. Alternância de estilo em entrevistas sociolinguísticas: uma aplicação dos critérios da árvore de decisão laboviana. **Revista do GEL**, v. 18, n. 3, São Paulo, 2021, p. 311-334.

BAVARESCO, Paulo Ricardo. **Ciclos econômicos regionais: modernização e empobrecimento no Extremo Oeste Catarinense**. Chapecó: Argos, 2005.

BBC News Brasil. **Brasil é 2º em ranking de países que passam mais tempo em redes sociais**. Época Negócios, Tecnologia, set. 2019. Disponível em: <https://epocanegocios.globo.com/Tecnologia/noticia/2019/09/brasil-e-2-em-ranking-de-paises-que-passam-mais-tempo-em-redes-sociais.html>. Acesso em: 13 jan. 2021.

- BECHARA, Evanildo. **Moderna gramática portuguesa**. 37. ed. rev., ampl. e atual. conforme o novo Acordo Ortográfico. Rio de Janeiro: Nova Fronteira, 2009.
- BELINE, Ronald. A variação linguística. In: FIORIN, José Luiz (org.). **Introdução à linguística: I. Objetos teóricos**. São Paulo: Contexto, 2002. p. 122-140.
- BOLSONARO, Jair Messias. **Projeto de lei nº 1.736-B/1996**. Brasília: 1996. Disponível em: [https://www.camara.leg.br/proposicoesWeb/prop\\_mostrarintegra;jsessionid=F0521145122C98F486D7302AE49288FD.proposicoesWebExterno1?codteor=1133712&filename=Dossie+-PL+1736/1996](https://www.camara.leg.br/proposicoesWeb/prop_mostrarintegra;jsessionid=F0521145122C98F486D7302AE49288FD.proposicoesWebExterno1?codteor=1133712&filename=Dossie+-PL+1736/1996). Acesso em: 31 jan. 2021.
- BRIGGS, Asa. BURKE, Peter. **Uma história social da mídia: De Gutenberg à internet**. Tradução: DIAS, Maria Carmelita Pádua. 2. ed. Rio de Janeiro: Zahar, 2006. p. 266-305.
- BYBEE, Joan. **Mudança linguística**. Tradução Marcos Bagno. Petrópolis, RJ: Editora Vozes, 2020.
- CAMACHO, Roberto Gomes. Uma breve retrospectiva da pesquisa sociolinguística. In: PARREIRA, Maria Cristina *et. al.* **Pesquisas em linguística no século XXI: perspectivas e desafios teóricos-metodológicos**. Araraquara, SP: Cultura Acadêmica Editora, 2015.
- CÂMARA dos deputados. **PL 1736/1996**. Disponível em: <https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=25269>. Acesso em: 31 jan. 2021.
- CÂMARA dos deputados. **PL 1676/1999**. Disponível em: <https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=17069>. Acesso em: 31 jan. 2021.
- CÂMARA JR., Joaquim Mattoso. **Estrutura da língua portuguesa**. Petrópolis: Vozes, 2008.
- CARRION, Raul. **Projeto de lei nº 156/2009**. Porto Alegre: 2009. Disponível em: <http://proweb.procergs.com.br/Diario/DA20090817-01-100000/EX20090817-01-100000-PL-156-2009.pdf>. Acesso em: 31 jan. 2021.
- CARVALHO, Nelly. **Empréstimos linguísticos**. São Paulo: Editora Ática S. A., 1989.
- CASTILHO, Ataliba T. de. **Nova gramática do português brasileiro**. 5. reimpressão - São Paulo: Editora Contexto, 2019.
- CASTRO, Janaína Souto de. **Os processos de formação e uso dos neologismos na internet e seu reflexo na revisão de texto**. 22 f. Monografia (Especialização em Revisão de Texto) – Instituto CEUB de Pesquisa e Desenvolvimento, Centro Universitário de Brasília, Brasília, 2017.
- COELHO, Izete Lehmkuhl *et al.* **Para conhecer sociolinguística**. 2. reimpressão - São Paulo: Editora Contexto, 2019.
- CORREIA, Margarita; ALMEIDA, Gladis M<sup>a</sup> de Barcellos. **Neologia em português**. São Paulo: Parábola Editorial, 2012.
- CRESCE o uso de Internet durante a pandemia e número de usuários no Brasil chega a 152 milhões, é o que aponta pesquisa do Cetic.br. **Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação**. São Paulo, 18 ago. 2021. Disponível em:

<https://www.cetic.br/pt/noticia/cresce-o-uso-de-internet-durante-a-pandemia-e-numero-de-usuarios-no-brasil-chega-a-152-milhoes-e-o-que-aponta-pesquisa-do-cetic-br/>. Acesso em: 1 out. 2021.

CUNHA, Antônio Geraldo da. **Dicionário etimológico da língua portuguesa**. 4. ed. Rio de Janeiro: Lexicon, 2010.

CUNHA, Antônio Geraldo da. **Os estrangeirismos da língua portuguesa: vocabulário histórico-etimológico**. São Paulo: Humanitas FFLCH/USP, 2003.

CUNHA, Celso. CINTRA, Lindley. **Nova gramática do português contemporâneo**. Rio de Janeiro: Lexicon, 2017.

COSERIU, E. **O homem e sua linguagem**. Tradução Carlos Alberto da Fonseca e Mário Ferreira. 2.ed. Rio de Janeiro: Editora Presença, 1987.

CRYSTAL, David. **Dicionário de linguística e fonética**. Tradução e adaptação Maria Carmelita Pádua Dias. Rio de Janeiro: Jorge Zahar Editor, 2000 [2. ed. A dictionary of linguistics and phonetics. Oxford, England: Basil Blackwell, 1985].

DAMIANI, Athos. MILZ, Beatriz. LENTE, Caio. FALBEL, Daniel. CORREA, Fernando. TRECENTI, Julio. LUDUVICE, Nicole. AMORIM, William. **Ciência de dados em R**. São Paulo: Instituto de Matemática e Estatística da Universidade de São Paulo (IME-USP), 2015. Disponível em: <<https://livro.curso-r.com/>> Acesso em: 10 dez. 2021.

DEBORD, Guy. **A sociedade do espetáculo**. Tradução Estela Abreu dos Santos. Rio de Janeiro: Contraponto. 13ª ed., 2013. p. 119-135.

DEFINIÇÃO do Valor-P. **Data Science**. 2020. Disponível em: <https://datascience.eu/pt/matematica-e-estatistica/definicao-do-valor-p/>. Acesso em: 10 dez. 2021.

ECKERT, Penelope. **Three waves of variations study: the emergence of meaning in the study of sociolinguistic variation**. California: Stanford University, 2012.

FARACO, Carlos Alberto (org.). **Estrangeirismos: guerras em torno da língua**. São Paulo: Parábola Editorial, 2001.

FARACO, Carlos Alberto. **Linguística histórica: uma introdução ao estudo da história das línguas**. São Paulo: Parábola Editorial, 2005.

FARACO, Carlos A.; TEZZA, Cristóvão. **Prática de textos para estudantes universitários**. 9. ed. Petrópolis, RJ: Vozes, 1992.

FERREIRA, Aurélio Buarque de Holanda. **Novo dicionário Aurélio da língua portuguesa**. 3. ed. Curitiba: Positivo, 2004.

FERREIRA, Aurélio Buarque de Holanda. **Mini Aurélio: o dicionário da língua portuguesa**. 8. ed. Curitiba: Positivo, 2010.

FERREIRA, Delson. **Manual de sociologia: dos clássicos à sociedade da informação**. 2. ed. São Paulo: Atlas, 2009. p. 165-181.

- FERREIRA, Giselle Vasconcelos dos Santos. GOMES, Nataniel dos Santos. *World English em Manhattan Connection: uma Análise dos Anglicismos na Linguagem Televisiva*. Círculo
- FERREIRA, Maira de Oliveira. **Estrangeirismos da Língua Inglesa: um estudo em três versões do dicionário Aurélio eletrônico**. Três Lagoas: Universidade Federal de Mato Grosso do Sul (UFMS), 2016.
- Fluminense de Estudos Filológicos e Linguísticos. **Revista Philologus**, ano 21, nº. 63 - Supl.: Anais da X CNLF. Rio de Janeiro: CiFEFil, set./dez. 2015.
- FISHER, Augusto; LÜCKMANN, Luiz Carlos. **A mesorregião do Oeste catarinense: análise comparativa de indicadores demográficos, econômicos e sociais de Santa Catarina**. Joaçaba, SC: Editora Unoesc, 2020.
- FORNARA, Ana Elizabeth. **Estratégias de fomento à manutenção de línguas minoritárias no Oeste catarinense**. 30 f. TCC (Graduação em Letras Português e Espanhol - Licenciatura) - Universidade Federal da Fronteira Sul (UFFS), Chapecó, 2016.
- FREITAG, Raquel Meister Ko.; "(Re)Discutindo Sexo/Gênero na Sociolinguística", p. 17-74 . In: Freitag, Raquel Meister Ko.; Severo, Cristine Gorski (Org). **Mulheres, Linguagem e Poder - Estudos de Gênero na Sociolinguística Brasileira**. São Paulo: Blucher, 2015. ISBN: 978-85-8039-121-3, DOI 10.5151/9788580391213-0001
- FREITAS, Ernani Cesar; BARTH, Pedro Afonso. Gênero ou suporte? O entrelaçamento de gêneros no Twitter. **Revista (Con) Textos Linguísticos (UFES)**, v. 9, p. 8-26, 2015. Disponível em: <https://periodicos.ufes.br/contextoslinguisticos/article/view/8888>. Acesso em: 8 jan. 2021.
- GALLI, Fernanda Correa Silveira. Linguagem da Internet: um meio de comunicação global. In: MARCHUSCHI, Luiz Antônio. XAVIER Antonio Carlos. **Hipertexto e gêneros digitais: Novas formas de construção de sentido..** São Paulo: Editora Cortez, 2009. p. 147-164.6
- GAÚCHA Zero Hora. **Tarso Genro decide vetar parcialmente projeto sobre estrangeirismos**. Porto Alegre: 2011. Disponível em: <https://gauchazh.clicrbs.com.br/amp/tarso-genro-decide-vetar-parcialmente-projeto-sobre-estran-geirismos-3315117.html>. Acesso em: 31 jan. 2021.
- GÖRSKI, Edair M.; ROST SNICHELOTTO, Cláudia A. **Introdução aos estudos gramaticais**. Florianópolis: LLV/CCE/UFSC, 2008 v. 1. 89 p.
- GRIES, Stefan Thomas. **Estatística com R para linguística**. Organização e tradução Heliana R. Mello *et al.* Belo Horizonte: FALE/UFMG, 2019.
- GUERRA, Rogério F. Língua Franca e estrangeirismos. **Revista de Ciências Humanas**, v. 46, nº. 2, p. 435 - 451, out. 2012. Florianópolis: Universidade Federal de Santa Catarina (UFSC), 2012.
- HASSELSTRON, Munick Maria. **Línguas de imigração em contato com o português no Oeste catarinense: crenças e atitudes linguísticas**. 149 f. Dissertação (Mestrado em Estudos Linguísticos) - Programa de Pós-graduação em Estudos Linguísticos, Universidade Federal da Fronteira Sul (UFFS), Chapecó, 2018.
- HOUAISS, Antônio; VILLAR, Mauro de Salles. Grande dicionário Houaiss da língua portuguesa. Rio de Janeiro: Objetiva, 2008 [2001].

ILARI, Rodolfo. **Introdução ao estudo do léxico** - brincando com as palavras. São Paulo: Contexto, 2002.

INSTITUTO do Patrimônio Histórico e Artístico Nacional (Iphan). **Três línguas são reconhecidas pelo Iphan como Referência Cultural Brasileira**. Brasília - DF, nov. 2014. Disponível em: <http://portal.iphan.gov.br/noticias/detalhes/114>. Acesso em: 27 jun. 2021.

KOMESU, Fabiana Cristina. *Blogs e as práticas de escrita sobre si na internet*. In: MARCHUSCHI, Luiz Antônio. XAVIER Antonio Carlos. **Hipertexto e gêneros digitais**: Novas formas de construção de sentido.. São Paulo: Editora Cortez, 2009. p. 135- 146.

LABOV, William. **Padrões Sociolinguísticos**. Tradução Marcos Bagno, Maria Marta Pereira Sherre, Carolina Rodrigues Cardoso. São Paulo: Parábola Editorial 2008 [1ª ed.: Sociolinguistic patterns. Philadelphia: University of Pennsylvania Press, 1972].

LABOV, William. Where does the linguistic variable stop? A response to Beatriz Lavandera. In: **Working Paper in Sociolinguistics**, Texas, 1978.

LABOV, William. Building on empirical foundations. In: LEHMANN, Winfred Philip; MALKIEL, Yakov. **Perspectives on historical linguistics**. Amsterdam; Philadelphia: John Benjamins, 1982, p. 17-92.

LÉVY, Pierre. **Cibercultura**. Tradução: COSTA, Carlos Irineu da. São Paulo: Editora 34, 1999.

LEITÃO, Isabela Custódio. **Anglicismos no português do Brasil**: um estudo lexicográfico Aurélio Houaiss. 2006. 140 f. Dissertação (Mestrado em Linguística e Língua Portuguesa) - Universidade Estadual Paulista, Faculdade de Ciências e Letras de Araraquara, 2006.

LOPES, Jezebel Batista. **Variação, percepções e atitudes linguísticas dos chapecoenses frente à referência à segunda pessoa do singular**. 332 f. Dissertação (Mestrado em Estudos Linguísticos) - Programa de Pós-graduação em Estudos Linguísticos, Universidade Federal da Fronteira Sul (UFFS), Chapecó, 2017.

LUFT, Celso P. **Novo manual de português, gramática, ortografia oficial, redação, literatura, textos e testes**. 17.ed. São Paulo: Globo, 1991.

MANZOLILLO, Vito César de Oliveira. **Ainda em torno da dicotomia empréstimo/estrangeirismo**. III Letras em Foco - Semana de Letras (FFP/UERJ). Rio de Janeiro, 2000.

MANZOLILLO, Vito César de Oliveira. **Empréstimo Linguístico**: o que é, como e porque se faz. Cadernos do CNLF, Vol. XVIII, Nº 03 – Minicursos e Oficinas. Rio de Janeiro: CiFEFiL, 2014.

MARCUSCHI, Luiz Antônio; XAVIER, Antônio Carlos. **Hipertexto e Gêneros Digitais**. Rio de Janeiro: Lucerna, 2004.

MARTINS FILHO, Eduardo Lopes. **Manual de Redação e Estilo**. 3. ed. revista ampliada — São Paulo: O Estado de S. Paulo, 1997.

OLIVEIRA, Josane Moreira de. A sociolinguística laboviana: festejando o cinquentenário e planejando o futuro. **Cadernos de Estudos Linguístico**, Campinas, set./dez. 2016, p. 481-501.

ORGANIZAÇÃO Mundial da Saúde declara pandemia de coronavírus. **Agência Brasil**, Brasília, 11 mar. 2020. Disponível em: <https://agenciabrasil.ebc.com.br/geral/noticia/2020-03/organizacao-mundial-da-saude-declara-pandemia-de-coronavirus>. Acesso em: 2 dez. 2020.

OUSHIRO, Livia. **Tratamento de dados com o R para análises sociolinguísticas**, p.134-177. In Raquel Meister Ko. Freitag (Organizadora). Metodologia de Coleta e Manipulação de Dados em Sociolinguística, São Paulo: Editora Edgard Blücher, 2014.

PEREIRA, Aguinaldo. Reações causadas pela Anglofonia no Cenário Político Brasileiro. **Web-Revista Sociodialeto**, Vol. 2. Nº. 2. Campo Grande: Universidade Estadual de Mato Grosso do Sul (UEMS), nov. 2012.

PRADO, Natália Cristine. **O uso do inglês em contexto comercial no Brasil e em Portugal: questões linguísticas e culturais**. São Paulo: Editora UNESP; São Paulo: Cultura Acadêmica, 2015.

QUANTIDADE de homens e mulheres. **Instituto Brasileiro de Geografia e Estatística (IBGE)**, Rio de Janeiro, 18 mar. 2020. Disponível em: <https://educa.ibge.gov.br/jovens/conheca-o-brasil/populacao/18320-quantidade-de-homens-e-mulheres.html>. Acesso em: 1 nov. 2021.

RADIN, José Carlos. BENEDET, José Higino. MILANI, Maria Luiza. (Orgs.). **Facetas da colonização italiana: Planalto e Oeste Catarinense**. Joaçaba: UNOESC, 2003.

RAJOGOPALAN, Kanavillil. Como o internetês desafia a linguística. In: SALIÉS, Tânia G. SHEPHERD, Tania G. **Linguística da internet**. São Paulo: Editora Contexto, 2013. p. 37-54.

RECUERO, Raquel. **Redes sociais na internet**. Porto Alegre: Editora Meridional, 2009.

RECUERO, Raquel. ZAGO, Gabriela. **Em busca das “redes que importam”**: redes sociais e capital social no *Twitter*. *Líbero*, São Paulo, v. 12, n. 24, p. 81-94, dez. de 2009.

RECUERO, Raquel. ZAGO, Gabriela. **“RT, por favor”**: Considerações sobre a difusão de informações no Twitter. *Revista Fronteiras – Estudos midiáticos*. Unisinos. maio/ago. 2010.

RITTER, Matias do Nascimento. THEY, Ng Haig. KONZEN, Enéas. **Introdução ao software estatístico R**. Imbé, RS: UFRGS/CECLIMAR, jun. 2019.

ROST SNICHELOTTO, Cláudia Andrea. **Varição e Mudança no Português do Oeste de Santa Catarina**, 2012. Plano de Trabalho Chamada Pública FAPESC n. 04/2012 Universal.

SALOMÃO, Ana Cristina Biondo. Variação e Mudança Linguística: Panorama e Perspectivas da Sociolinguística Variacionista no Brasil. **Fórum Lingüístico**, Florianópolis, v. 8, n.2, p.187-207, jul./dez. 2011. Disponível em: <https://periodicos.ufsc.br/index.php/forum/article/view/1984-8412.2011v8n2p187/21673>. Acesso em: 13 jan. 2021.

SATO, Paula. **É possível calcular quantas palavras surgem por dia na Língua Portuguesa?** **Nova Escola**, São Paulo, maio 2009. Disponível em: <https://novaescola.org.br/conteudo/2539/e-possivel-calcular-quantas-palavras-surgem-por-dia-na-lingua-portuguesa>. Acesso em: 27 jun. 2021.

SILVA, Heloisa Spolador; VALLS, Valéria Martin. **Retenção de conhecimento na Internet: o papel do Twitter**. CRB-8 Digital, São Paulo, v. 1, n. 5, p. 124-147, jan. 2012.

SILVA, Jaqueline Alves da. **Estrangeirismos no Facebook**. 2015. 18 f. TCC (Licenciatura em Letras Português) Instituto de Letras, Departamento de Linguística, Português e Línguas Clássicas, Universidade de Brasília (UNB), Buritis-MG, 2015.

SOBRE a doença: o que é a Covid-19. **Ministério da Saúde**. Brasília, 13 mar. de 2020. Disponível em: <https://coronavirus.saude.gov.br/sobre-a-doenca#o-que-e-covid>. Acesso em: 2 dez. 2020.

SOUZA, Adilio Junior de. **Lexicalização e neologismo: análise funcional em corpus digital**. 155 f. Dissertação (Mestrado em Linguística), Centro de Ciências Humanas, Letras e Artes, Programa de Pós-graduação em Linguística, Universidade Federal da Paraíba (UFPB), João Pessoa, 2015.

SOUZA, Sonia Maria da Fonseca. BORGES, Cláudia Gualberto. SILVA, Jônatas Guimarães. RABELO, Priscila Banca. **Estrangeirismos: Vestígios de Poder da Língua Inglesa**. **Revista Eletrônica Saberes Múltiplos**, vol. 01, Nº. 02, Nova Iguaçú: Universidade de Iguaçú (UNIG), 2015. p. 32-46.

SPINASSÉ, Karen Pupp. Os conceitos Língua Materna, Segunda Língua e Língua Estrangeira e os falantes de línguas alóctones minoritárias no Sul do Brasil. **Revista Contingentia**, Vol. 1, nov. Porto Alegre, 2006. p. 01-10

SHEPHERD, G. Tania; SALIÉS, G. Tânia (Orgs.). **Linguística da Internet**. São Paulo: Editora Contexto, 2013.

TERMOS de serviço do *Twitter*. **Twitter**. 2020. Disponível em: <https://twitter.com/pt/tos> Acesso em: 26 dez. 2020.

TWITTER suspende perfil de Donald Trump permanentemente. **R7**. São Paulo: jan. 2021. Disponível em: <https://noticias.r7.com/internacional/twitter-suspende-perfil-de-donald-trump-permanentemente-08012021#/foto/1>. Acesso em: dez. 2021.

**TWITTER Marketing BR. Perfil do Twitter para profissionais de marketing e comunicação - Insights e novidades para inspirar o mercado brasileiro**. São Paulo. Disponível em: <https://twitter.com/TwitterMktgBR>. Acesso em: 8 jan. 2021.

TORRANO, Sandra Delneri Petean. **Produtividade e criatividade do léxico: os neologismos na área da informática**. São Paulo: Departamento de Letras Clássicas e Vernáculas - Universidade de São Paulo (USP), 2010.

WEINREICH, Uriel; LABOV, William; HERZOG, Marvin. **Fundamentos empíricos para uma teoria da mudança linguística**. Tradução Marcos Bagno. São Paulo: Parábola Editorial, 2006. [1ª ed. Empirical Foundations for a Theory of Language Change *in*: Directions for Historical Linguistics - A Symposium. Austin: University of Texas Press, 1968, p. 95-199].

WERLANG, Alceu Antonio. Colonização Ítalo-brasileira, Teuto-brasileira e Teuto-russa no Oeste de Santa Catarina – A atuação da CIA Territorial Sul Brasil. **Cadernos do CEOM**, v. 13, nº 11, Chapecó: Grifos, 1999. p. 11-53.

XAVIER, Antonio Carlos. Reflexões em torno da escrita nos novos gêneros digitais da internet. **Revista Investigações - Linguística e Teoria Literária**, v. 18, n.º. 2. Pernambuco: Universidade Federal de Pernambuco (UFPE), 2005.

ZILLES, A. M. S.; GARCEZ, P. M. Estrangeirismos: desejos e ameaças. In: FARACO, Carlos A.. (Org.). **Estrangeirismos: guerras em torno da língua**. São Paulo: Parábola, 2001. p. 15-36.