

INOVAÇÃO LEXICAL E OBSOLESCÊNCIA LINGUÍSTICA NA ÁREA DE INFORMÁTICA: UM ESTUDO À LUZ DA LINGUÍSTICA DE *CORPUS*

Francisco Iací do NASCIMENTO

Sarah Virgínia Carvalho RIBEIRO

Antônio Luciano PONTES

Universidade Estadual do Ceará - UECE

Resumo: O léxico das línguas se renova constantemente por meio de processos vermiculares ou de empréstimos lexicais. Em algumas áreas como a informática a inovação tecnológica é intensa. Isso se reflete na língua com a criação de palavras novas para a denominação de novos objetos, conceitos, processos, etc. O objetivo deste trabalho é discutir como os neologismos da área de informática são introduzidos no léxico da língua portuguesa. Para tanto, apoiamos-nos, teoricamente, nos estudos sobre Neologia em português de Alves (2004, 2012) e Correia e Almeida (2012), e nos estudos sobre Linguística de *Corpus* de Sinclair (2005), Aluísio e Almeida (2006) e Mello e Sousa (2014). Definimos como objeto de estudo 15 (quinze) neologismos da área de informática que sofreram alteração morfossintática, catalogados por Assirati (1998). Para discutir o caráter neológico dessas 15 palavras, fizemos um estudo contrastivo no qual utilizamos a *Internet*, como *corpus* de estudo, e um conjunto de dicionários gerais, escolares e de aprendizagem como *corpus* de referência. Os resultados da análise mostraram que das 15 palavras estudadas 7 já foram incorporadas ao léxico da língua portuguesa, considerando o critério lexicográfico.

Palavras-chave: Neologismo. Linguística de *Corpus*. Informática. Léxico.

LEXICAL INNOVATION AND LINGUISTIC OBSOLESCENCE IN THE FIELD OF INFORMATICS: A STUDY UNDER THE LIGHT OF *CORPUS* LINGUISTICS

Abstract: The lexicon of languages is constantly renewed through vernacular or lexical loan processes. In some fields, such as computer science, technological innovation is intense and that is reflected on language through the creation of new words to name the new objects, concepts, processes etc. The purpose of this research is to discuss how neologisms from the informatics field are included in the Portuguese language lexicon. Therefore, we have based our theoretical background in the Portuguese studies on Neologism by Alves (2004, 2012), and Correia and Almeida (2012), and the *Corpus* Linguistics studies by Sinclair (2005), Aluísio and Almeida (2006) and Mello and Sousa (2014). We have defined, as our object of study, 15 (fifteen) neologisms from the informatics field which have undergone morphosyntactic changes, and which have been cataloged by Assirati (1998). In order to discuss the neological features of those 15 words, we have made a contrastive study using the *Internet* as our study *corpus* and a set of dictionaries

(general, school and learners) as our reference *corpus*. The analyses results showed that among the 15 words studied 7 have been incorporated into the lexicon of the Portuguese language taking into account the lexicographic criteria.

Keywords: Neologism. Corpus Linguistics. Informatics. Lexicon.

INNOVACIÓN LEXICAL Y OBSOLESCENCIA LINGÜÍSTICA EN EL ÁREA DE INFORMÁTICA: UN ESTUDIO A LA LUZ DE LA LINGÜÍSTICA DE CORPUS

Resumen: El léxico de las lenguas se renueva constantemente a través de procesos vernaculares o de préstamos lexicales. En algunas áreas como la informática la innovación tecnológica es intensa. Eso se refleja en la lengua con la creación de nuevas palabras para la denominación de nuevos objetos, conceptos, procesos, etc. El objetivo de este trabajo es discutir cómo los neologismos del área de informática son introducidos al léxico de la lengua portuguesa. Para ello, nos apoyamos teóricamente en los estudios sobre Neología en portugués de Alves (2004, 2012) y Correia y Almeida (2012) y en los estudios sobre Lingüística de Corpus de Sinclair (2005), Aluísio y Almeida (2006) y Mello y Sousa (2014). Definimos como objeto de estudio 15 (quince) neologismos del área de informática que sufrieron alteración morfosintáctica, catalogados por Assirati (1998). Para discutir el carácter neológico de esas 15 palabras, hicimos un estudio contrastivo en el que utilizamos Internet como corpus de estudio y un conjunto de diccionarios generales, escolares y de aprendizaje como corpus de referencia. Los resultados del análisis mostraron que de las 15 palabras estudiadas 07 fueron ya incorporadas al léxico de la lengua portuguesa, considerando el criterio lexicográfico.

Palabras clave: Neologismo. Lingüística de Corpus. Informática. Léxico.

INTRODUÇÃO

O final do século XX e início do século XXI são marcados por um grande desenvolvimento tecnológico. Equipamentos e tecnologias digitais invadiram nossa vida de forma extraordinária, mudando nossa relação com o mundo, deixando-nos cada vez mais conectados e dependentes das inovações tecnológicas trazidas por computadores, celulares, *tablets* etc. Hoje, estamos cercados de tecnologias por todos os lados – no trabalho, em casa e na rua – desde a hora que acordamos até a que vamos dormir. Esse uso intenso e constante de aparelhos tecnológicos se reflete na língua, especialmente, no léxico, uma vez que se tem a necessidade de criar novas palavras para denominar toda essa tecnologia, bem como os processos e ações que executamos com ela.

A influência da tecnologia no léxico das línguas é visível pelo uso intenso de neologismos. Isso ocorre porque o léxico é a parte mais aberta da língua e estabelece uma relação muito

estreita com a cultura, pois denomina as coisas da realidade extralinguística. Devido a essa abertura, o léxico se torna mais suscetível a acréscimos de novas formas por meio de processos vernaculares de criação lexical ou de empréstimos para atender às demandas linguísticas de uma determinada comunidade. Dessa forma, ao surgir um novo objeto, é preciso nomeá-lo. A velocidade da inovação tecnológica tem exigido constantemente a criação de palavras novas. Sendo assim, a informática se tornou uma das áreas em que a relação do léxico com a cultura fica bem mais evidente, uma vez que a cultura tecnológica é marcada pela velocidade na criação e consumo de novos equipamentos e tecnologias, pondo à mostra toda a dinamicidade e vitalidade do léxico das línguas na criação de novas palavras para nomear essas novas tecnologias.

No entanto, na área de informática, ao mesmo tempo em que palavras novas são criadas, outras deixam de ser usadas, acompanhando o ritmo de obsolescência tecnológica, dessa forma também teríamos uma espécie de obsolescência linguística, em que muitas palavras entram em desuso, tornam-se arcaicas, antes mesmo de entrar definitivamente para o léxico das línguas e serem registradas nos dicionários. Neste estudo, selecionamos 15 (quinze) neologismos da área de informática que sofreram algum tipo de adaptação, do ponto de vista sintático, catalogados por Assirati (1998). A escolha desses neologismos se deu pelo fato de já estarem em processo de integração ao léxico do Português, uma vez que são formados por um processo de derivação a partir de uma base estrangeira (*arjear, becapear, butar, debugar, dropar, escanear, getar, hobbista, inicializar, inputar, linkar, startar, updeitar, zipar*).

Há também uma carência de estudos em que se analisa e se discute o processo de renovação lexical, especialmente, a investigação dos neologismos que não vingaram e não entraram para o léxico da língua. Muitos estudos sobre neologia foram realizados na informática, catalogando os neologismos em uma determinada época, mas nenhum estudo posterior foi feito analisando a entrada dessas formas neológicas nos dicionários ou as formas que entraram em desuso antes de serem dicionarizadas.

Este estudo tem o propósito de discutir, à luz da Linguística de *Corpus*, como os neologismos são introduzidos ao léxico da língua. Para isso, baseamo-nos em pressupostos teóricos e metodológicos da Linguística de *Corpus*, principalmente, nos estudos de Sinclair

(2005), Aluísio e Almeida (2006) e Mello e Sousa (2014); e nas discussões teóricas e empíricas sobre Neologia feitas por Alves (2004, 2012) e por Correia e Almeida (2012).

Este artigo está organizado em quatro seções, além dessa introdução. Inicialmente, discutimos os postulados teórico-metodológicos da Neologia e da Linguística de *Corpus* para fundamentar teoricamente este trabalho. Depois, apresentamos a metodologia empregada para discutir e analisar o caráter neológico das palavras aqui estudadas. Em seguida, fizemos a análise dos dados desta pesquisa e apresentamos os resultados obtidos. E por último, tecemos algumas considerações finais.

1. NEOLOGIA E LINGUÍSTICA DE *CORPUS*

A pesquisa em Neologia, assim como em várias outras áreas de estudos da linguagem, tem se beneficiado também do aparato teórico-metodológico da Linguística de *Corpus*, seja no uso de parâmetros para compilação de *corpora*, seja no uso e desenvolvimento de ferramentas de extração automática ou semiautomática de neologismos. De acordo com Correia e Almeida (2012), para fazer um estudo em neologia tem-se, inicialmente, que construir um *corpus* de estudo ou *corpus* de extração. Depois, define-se um *corpus* de exclusão, geralmente constituído por dicionários gerais de língua. Em seguida, constrói-se uma ficha para o registro dos neologismos. Depois de organizados estes instrumentos, inicia-se o processo de extração de neologismos de forma manual ou semiautomática.

1.1. NEOLOGIA: PROCESSOS DE INOVAÇÃO VOCABULAR

O léxico de uma língua é dinâmico e se renova constantemente. Palavras deixam de ser utilizadas e se tornam arcaicas, obsoletas; outras podem surgir para atender a novas demandas e a necessidades linguísticas, os chamados neologismos. Esse processo de renovação lexical é denominado de Neologia. Já o elemento resultante desse processo – a palavra nova – é chamado de neologismo (ALVES, 2004). O termo Neologia se refere também ao “estudo dos neologismos que vão surgindo em uma língua” (CORREIA; ALMEIDA, 2012, p. 17).

A Neologia pode ser denominativa ou estilística. Denominativa quando resulta da necessidade de nomear novos objetos ou conceitos, que não existiam anteriormente. E a estilística ocorre quando se procura uma maior expressividade no discurso, a fim de se traduzir

ideias não originais de uma maneira nova, expressiva, criativa. Neste trabalho, nos deteremos à Neologia denominativa que pode resultar em uma novidade formal (significante novo) ou em uma novidade semântica (uma palavra já existente adquire uma acepção nova). Partindo dessa classificação, os neologismos podem ser definidos como sendo:

Uma unidade lexical cuja forma significante ou cuja relação significado-significante, caracterizada por um funcionamento efetivo num determinado modelo de comunicação, não se tinha realizado no estágio imediatamente anterior do código da língua (CORREIA; ALMEIDA, 2012, p.23).

Contudo, não é tarefa fácil identificar neologismos. De acordo com Correia e Almeida (2012), nos estudos de neologia se tem adotado alguns critérios para essa identificação, tais como o sentimento de novidade, o critério lexicográfico e a instabilidade formal.

O sentimento de novidade é um critério psicológico em que predomina a percepção do falante sobre a palavra tida como nova. Esse critério não é suficiente para sustentar um trabalho de natureza científica, uma vez que pode se deparar com uma palavra nova para ele, mas que já faz parte efetivamente do léxico da língua.

O critério lexicográfico é o mais utilizado pelos estudiosos, considerando como neológica uma palavra não registrada nos dicionários representativos do estado de uma língua em questão. Parte-se do princípio de que os dicionários gerais de língua são representativos do seu léxico em um dado momento. Geralmente, nas pesquisas sobre neologismos, esses dicionários compõem o *corpus* de exclusão, ou seja, se a palavra está registrada no dicionário não é considerada neologismo. Atualmente, há a possibilidade também de se usar como *corpus* de exclusão os *corpora* representativos de uma língua, para atestar se uma unidade é ou não neológica.

Já a instabilidade formal pode ser considerada também como um componente na identificação de unidades neológicas, isto é, caso uma palavra apresente sinais de instabilidade de natureza morfológica, fonética ou ortográfica, poderá ser considerada um neologismo. Contudo, é preciso ter cuidado com esse critério, sobretudo, quando se trata de estrangeirismos, uma vez que muitas palavras são grafadas da mesma forma que na língua estrangeira. Às vezes, contamos com duas formas de grafia para uma mesma palavra, uma aportuguesada e a outra tal qual a forma da língua estrangeira. Como, por exemplo, *xampu*, que os dicionários brasileiros

registram com essa grafia; já os portugueses, como *champô*; entretanto, nos meios de comunicação de forma geral se grafa *shampoo*. Vale salientar que, mesmo apresentando instabilidade formal, a palavra *xampu* não pode ser considerada uma forma neológica. Enfim, é preciso levar em conta esses critérios de forma combinada.

Por outro lado, vale destacar que a integração de um vocábulo novo ao idioma é feita pela comunidade linguística que decide pelo uso ou não do elemento neológico. Portanto, a frequência de uso de um neologismo também pode ser considerada um critério para determinar a sua incorporação ou não ao idioma. Atualmente, isso é perfeitamente possível de ser feito por meio dos *corpora* de língua geral já existentes e até mesmo se utilizando da *Internet* como *corpus*, no caso de usos mais informais.

A criação de novas palavras pode ser feita mediante processos vernaculares ou de empréstimos lexicais. Os processos vernaculares acontecem quando se recorre a regras próprias da língua para a construção de palavras novas; ou quando se atribui novos significados a palavras já existentes. Já os empréstimos ocorrem quando se importa palavras de outras línguas.

Desses processos, o empréstimo é o que se torna mais visível, mais saliente, uma vez que se introduz uma forma estranha ao sistema da língua que leva algum tempo para se adaptar a esse sistema, em alguns casos, mantendo-se como a forma original da língua estrangeira, os chamados estrangeirismos. Isso divide opiniões e gera polêmicas, principalmente, entre os puristas, por acharem que essas formas vão descaracterizar o idioma. Chegando-se inclusive a ser proposta uma lei para regulamentar o uso de estrangeirismos no Brasil.

Contudo, em tempos de globalização é muito difícil não ter contato com outras línguas, especialmente, a língua inglesa que se tornou a língua de comunicação internacional, das ciências, do turismo e dos negócios, devido ao predomínio econômico de países de fala anglo-saxônica.

1.2. NEOLOGISMOS POR EMPRÉSTIMO NO CAMPO DA INFORMÁTICA

No caso específico das áreas de tecnologia e de informática, a grande maioria das empresas desses setores utiliza a língua inglesa em seus produtos, gerando, muitas vezes, o uso excessivo de formas em inglês nas demais línguas. De acordo com Torrano (2010, p. 108), na área

de informática, os estrangeirismos são bem mais frequentes, representam 46% dos termos encontrados. Já na língua geral os empréstimos correspondem a 17% das criações neológicas (ALVES, 2012, p. 191). Como podemos constatar é uma diferença bastante significativa.

A grande quantidade de estrangeirismos na área de informática tem despertado o interesse de muitos pesquisadores, que se empenham em descrever como os empréstimos lexicais são usados. Assirati (1998) realizou um estudo sobre os neologismos no campo da informática com base em um *corpus* extraído de edições do caderno "Informática" de O Estado de S. Paulo e de entrevistas com técnicos. A estudiosa analisou os neologismos por empréstimo no vocabulário técnico-científico de informática, buscando descrever a sua integração ao sistema linguístico português, sob os pontos de vista fonético/fonológico, morfossintático e semântico. Desse estudo, fizemos um recorte dos neologismos que sofreram adaptação do ponto de vista morfossintático, com o intuito de analisarmos como essas formas neológicas que já passaram por algum tipo de adaptação morfossintática se integraram ao acervo lexical da língua portuguesa no Brasil. No quadro 1, feito com base nos dados da referida pesquisa, expomos o neologismo, o seu significado e a sua formação morfossintática.

Quadro 1 - Neologismos da informática que sofreram adaptação morfossintática

NEOLOGISMO	CONCEITO	FORMAÇÃO
<i>arjear</i>	Compactar usando o programa ARJ.	Verificamos a base nominal inglesa <i>arj</i> + vogal de ligação /e/ + sufixo verbal português -ar, o que constitui uma formação <i>híbrida</i> .
<i>becapear</i>	Fazer uma cópia de um arquivo, ou dados ou disco.	Termo formado pelo verbo original <i>back up</i> , adaptado gráfica e fonologicamente <i>becap</i> (i) + vogal de ligação e + sufixo verbal -ar.
<i>butar</i>	Dar início ao programa.	Termo formado pela base inglesa <i>boot</i> , adaptada <i>but</i> + sufixo verbal -ar. Constitui um <i>hibridismo</i> .
<i>debugar</i>	Retirar os defeitos de um programa; depurar.	Termo formado pela base inglesa <i>debug</i> + sufixo verbal -ar.
<i>deletar</i>	Apagar, remover um caractere, ou um texto.	Este termo formou-se a partir do verbo inglês <i>delete</i> - em que observamos a supressão da vogal <i>e</i> final - + sufixo verbal -ar.
<i>dropar</i>	Excluir procedimentos, tabelas, etc.	<i>Hibridismo</i> formado pela forma verbal <i>drop</i> + sufixo verbal -ar.
<i>escanear</i>	Passar uma imagem ou texto para o computador.	<i>Hibridismo</i> formado pela base inglesa <i>scan</i> + vogal de ligação e + sufixo verbal -ar.
<i>getar</i>	Entrar no sistema.	<i>Hibridismo</i> formado pela base inglesa <i>get</i> + sufixo verbal -ar.
<i>hobbista</i>	Designa aqueles que	<i>Hibridismo</i> formado pela base inglesa <i>hobby</i> +

	têm um <i>hobby</i> que é praticado através do computador.	sufixo nominal <i>-ista</i> , em que observamos a supressão da consoante <i>y</i> .
<i>inicializar</i>	Dar partida; colocar a máquina em funcionamento.	Hibridismo formado pelo verbo inglês <i>initialize</i> , em que verificamos a adaptação gráfica e fonológica de <i>inicial</i> para <i>inicial</i> + sufixo verbal <i>-izar</i> .
<i>inputar</i>	Inserir dados no computador	Hibridismo formado pelo verbo inglês <i>input</i> + sufixo verbal <i>-ar</i> . O termo original <i>input</i> , ao formar o verbo <i>inputar</i> , concorre com o homônimo <i>imputar</i> , que tem um sentido bastante diverso.
<i>linkar</i>	Ligar.	Hibridismo formado pelo verbo inglês <i>link</i> + sufixo verbal <i>-ar</i> . Da forma verbal <i>linkar</i> , surgiram formas derivadas como <i>linkado</i> , <i>linkagem</i> , etc.
<i>startar</i>	Iniciar o computador.	Hibridismo formado pela forma inglesa <i>start</i> + sufixo verbal <i>-ar</i> .
<i>updeitar</i>	Atualizar a informação por meio da alteração ou adição de dados específicos em um arquivo principal.	Hibridismo formado pela base inglesa adaptada <i>updeit</i> (i) + sufixo verbal <i>-ar</i> . Esta formação também poderia ser evitada, utilizando-se a forma <i>atualizar</i> .
<i>zipar</i>	Compactar arquivos.	Hibridismo formado a partir do nome <i>PKZIP</i> , que é um programa de compactação. Sua formação é bastante inusitada: a forma final <i>ZIP</i> < <i>PKZIP</i> + sufixo verbal <i>-ar</i> .

Fonte: elaborado pelos autores a partir do estudo de Assirati (1998)

Alguns dos neologismos listados no quadro, mesmo já tendo sofrido algum tipo de adaptação morfossintática, ainda são estranhos ao nosso sistema ortográfico, tais como *inputar*, *linkar*, *startar*, *updeitar*. Em nossa seção de análise (seção 3), veremos que essa instabilidade ortográfica ainda permanece, principalmente, nas formas que não foram dicionarizadas. A seguir, discutiremos o conceito de Linguística de *Corpus*, como a pesquisa em linguística pode se beneficiar dos métodos e ferramentas dessa disciplina e como podemos utilizar a *Internet* como *corpus*.

1.3. A LINGUÍSTICA DE CORPUS

No início deste artigo, mencionamos a influência da tecnologia no léxico das línguas. No entanto, a própria Linguística de *Corpus* (LC doravante) foi influenciada pelo avanço nos recursos computacionais aplicados ao estudo do léxico e está vivenciando um momento de crescente e contínua expansão em sua aplicação. Vale salientar que “a utilização de *corpus* sempre foi um recurso empregado em pesquisas linguísticas” (ALUÍSIO; ALMEIDA, 2006, p. 156), sobretudo, nos

estudos lexicográficos e lexicológicos. O que vemos recentemente é uma mudança na concepção do que é considerado um *corpus*.

1.3.1. Conceito de corpus

Existem diferentes definições para o que denominamos de *corpus* para a LC. Apresentamos, como forma de exemplificar, apenas duas delas. Para Sanchez, (1995, p. 8-9) é

Um conjunto de dados linguísticos (pertencentes ao uso oral ou escrito da língua, ou a ambos), sistematizados segundo determinados critérios, suficientemente extensos em amplitude e profundidade, de maneira que sejam representativos da totalidade do uso linguístico ou de algum de seus âmbitos, dispostos de tal modo que possam ser processados por computador, com a finalidade de propiciar resultados vários e úteis a descrição e análise (Tradução dos autores)¹.

Apresentamos, a seguir, a definição de Sinclair (2005, p. 16), por ser ele considerado um dos maiores linguistas de *corpus*, sendo o pioneiro na compilação de um dicionário (COBUILD) a partir de um *corpus* computadorizado:

Um *corpus* é uma coleção de extratos de textos linguísticos em formato eletrônico, selecionados de acordo com critérios externos para representar, o melhor possível, uma língua ou variedade linguística como fonte de dados para pesquisa linguística (Tradução dos autores)².

Nas definições mais atuais de *corpus* encontradas na literatura, devemos procurar por quatro características fundamentais. Segundo McEnery e Wilson (1996 *apud* ALUÍSIO; ALMEIDA, 2006, p. 157), são elas: 1- amostragem e representatividade; 2- tamanho finito; 3- formato eletrônico; e 4- referência padrão. Percebemos que estas características são contempladas, em sua maioria, nas definições apresentadas acima.

¹ Un corpus lingüístico es un conjunto de datos lingüísticos (pertenecientes al uso oral o escrito de la lengua, o a ambos) sistematizados según determinados criterios, suficientemente extensos em amplitud y profundidad de manera que sean representativos del total del uso linguístico o de alguno de sus ámbitos, y dispuestos de tal modo que pueden ser procesados mediante ordenador com el fin de obtener resultados varios y útiles para la descripción y el análisis.

² A corpus is a collection of pieces of language text in electronic form, selected according to external criteria to represent, as far as possible, a language or language variety as a source of data for linguistic research.

De acordo com Mello e Souza (2014), a LC surgiu da necessidade que estudiosos da língua sentiram de se apoiar em usos reais para fazerem generalizações ou esboçarem teorias a respeito do funcionamento linguístico. Por intermédio de diferentes *softwares* e ferramentas é possível fazer uso de uma abordagem empirista, que tem como ponto central a noção de linguagem enquanto sistema probabilístico, sendo possível evidenciar e quantificar padrões, ou seja, a correlação entre traços linguísticos e contextos situacionais de uso da linguagem.

1.3.2. Pesquisa com corpus

A cada dia nos deparamos com novas possibilidades para o uso de *corpora* em estudos linguísticos, desde a simples verificação da frequência de uso de uma palavra ou expressão, até a aplicação na análise de textos de literatura. A lexicologia e a lexicografia, o ensino de línguas e a tradução são algumas das áreas que têm se beneficiado ampla e favoravelmente da LC.

Mello e Souza (2014) discutem a compilação de *corpora*, o desenvolvimento de ferramentas para análise de *corpora*, a descrição da linguagem e a exploração do uso de descrições baseadas em *corpora* para várias aplicações como as principais áreas da LC.

As ferramentas computacionais geralmente utilizadas na LC mencionadas pelas mesmas autoras são: Programas para listar palavras (frequenciadores); Concordanciadores; Etiquetadores; e Ferramentas de Engenharia Textual.

Dependendo dos objetivos da pesquisa a ser conduzida, as ferramentas a serem utilizadas são selecionadas, bem como o tipo e o tamanho do *corpus*. Comentaremos a seguir um dos *corpora* que tem sido usado com frequência cada vez maior, o *corpus web*.

1.3.3 A Internet como corpus

Existem diversos *corpora* disponíveis *online* em diferentes línguas, e o pesquisador deve selecionar aquele que mais se adéqua a sua pesquisa. Quando nenhum dos *corpora* que estão disponíveis pode ser utilizado, cabe a ele construir o seu próprio *corpus* de estudo.

Rios (2010, p. 4) defende que a *web* deve ser utilizada como *corpus*, principalmente para os estudos linguísticos com idiomatismos. A autora escreve: “Diversos autores mencionam a *web*

como uma fonte válida para a obtenção de dados que atestem o uso real da língua, apesar das dificuldades provenientes do fato de não se tratar exatamente de uma base de dados linguística.” Fletcher (2005 *apud* RIOS 2010) lista oito vantagens em se considerar os dados provenientes da *web* para a pesquisa em Linguística: atualidade, espontaneidade, completude, escopo, diversidade linguística, custo, conveniência e representatividade.

Há duas formas de se trabalhar a *web* no que se refere à utilização dos dados nela contidos para fins de pesquisa linguística: a *web* para um *corpus* (*web for corpus-WfC*) e a *web* como um *corpus* (*web as corpus-WaC*). Nesta se considera a *web* como um *corpus* em si, enquanto que naquela ela é vista como uma fonte em potencial de coleta de textos para a criação de um *corpus* com características definidas pelo pesquisador. Na próxima seção, abordaremos a metodologia empregada para esta pesquisa.

2. METODOLOGIA

A descrição do léxico de uma língua exige procedimentos metodológicos de caráter mais qualitativo, contudo, faz-se necessário também quantificar ocorrências para que se tenha uma noção de frequência de uso das formas lexicais em um determinado *corpus*.

Neste trabalho, inicialmente, definimos os neologismos a serem estudados (ver Quadro 01) para em seguida consultarmos os *corpora* do português do Brasil, a fim de verificar suas ocorrências. Tivemos, contudo, muita dificuldade de encontrar um *corpus* acessível, tanto que recorremos ao *Corpus* do Português³, no qual, infelizmente, não foi possível encontrar um número significativo de ocorrências. Diante desse impasse, resolvemos usar a *Internet* como *corpus* de estudo, além dos principais dicionários gerais de língua, dos escolares e dos de aprendizagem como *corpus* de referência.

Escolhemos como “motor” de busca, o buscador comercial do *Google*. Inserimos cada um dos neologismos entre aspas para garantir que as ocorrências, pelo menos, em termos formais, fossem apenas da palavra em questão. Depois disso, filtramos as buscas por país (Brasil), idioma (páginas em Português) e período (de 01 de janeiro de 1998 a 31 de julho de 2015). Com relação ao período, resolvemos escalonar em intervalos de quatro: período I – de 01 de janeiro

³ Disponível em: <http://www.corpusdoportugues.org/xp.asp>. Acesso em: 19/06/2017.

de 1998 a 31 de dezembro de 2002; período II – de 01 de janeiro de 2003 a 31 de dezembro de 2006; Período III – de 01 de janeiro de 2007 a 31 de dezembro de 2010; e período IV – de 01 de janeiro de 2011 a 31 de julho de 2015. Esse escalonamento objetiva uma análise mais acurada das ocorrências e uma visão da tendência da quantidade de usos no período.

Buscamos os neologismos no *Google* e anotamos, a cada busca realizada, a quantidade de ocorrências por período (ver Tabela 1, na seção 3). Depois disso, buscamos cada um dos neologismos no *corpus* de referência, composto por:

- **Dicionários Gerais** – Ferreira (2010), Aulete (2015), Michaelis (2012) e Houaiss (2009);
- **Dicionários Escolares** (tipo 4) – Bechara (2011), Houaiss (2011), Borba (2011), Aulete (2011);
- **Dicionários Gerais Monolíngues *online***: um de inglês britânico - Cambridge (2015) e um de inglês americano – Collins (2015);
- **Dicionários de Aprendizagem Bilíngues** (inglês e português) – Minidicionário Bilíngue Prático (2008) e Oxford Escolar (2009).

Depois de catalogar os dados das buscas no Google e nos dicionários, iniciamos a análise e a discussão dos referidos dados com base em dois critérios: a quantidade de ocorrências dos neologismos no *corpus* de estudo (*Internet*) e o registro ou não deles no *corpus* de referência (dicionários). A seguir apresentamos essa análise, buscando compreender como os neologismos em estudo são usados efetivamente na *Internet* e como eles são registrados nos dicionários consultados.

3. ANÁLISE DOS DADOS E RESULTADOS

Depois de buscar e sistematizar as ocorrências de cada forma neológica, iniciamos o processo de análise. Na Tabela 1, temos a quantidade de ocorrências por período de quatro anos. Vale salientar que as buscas feitas no *Google* retornaram apenas ocorrências de forma. Para ocorrências de sentido, é preciso investigar cada ocorrência e mapear os sentidos que estão sendo empregados. Dessa forma, as quantidades da Tabela 1 se referem apenas às ocorrências dos neologismos da forma como foram escritos, exatamente da maneira que estão escritos na Tabela 1.

Tabela 1- Quantidade de ocorrências pesquisadas no motor de busca Google

NEOLOGISMO	PERÍODO				TOTAL
	P I (1998/2002)	P II (2003/2006)	P III (2007/2010)	P IV (2011/2015)	
deletar	353.000	294.000	724.000	1.650.000	2.810.000
linkar	848	3.890	31.600	78.600	114.000
inicializar	2.330	6.630	28.700	58.800	94.600
escanear	625	2.230	15.500	49.600	67.100
dropar	218	1.100	15.700	32.900	41.900
startar	167	870	2.750	6.220	9.850
debugar	153	580	2.430	5.530	8.520
zipar	148	447	2.380	4.420	7.250
hobbista	71	116	690	2.500	3.300
butar	57	118	718	2.350	2.940
inputar	03	38	175	356	569
getar	03	02	113	321	461
becapear	07	08	33	57	105
updeitar	0	02	03	03	08
arjear	0	02	01	01	04

Fonte: elaborada pelos autores

Percebe-se, mediante uma leitura panorâmica da tabela, que em geral as quantidades de ocorrências crescem a cada período. Isso talvez se deva em virtude do crescimento do número de usuários da *Internet* nos últimos anos, porém esse aumento constante também pode indicar que essas formas neológicas estão sendo aceitas pelos falantes que a cada período as utilizam mais, fazendo com que elas entrem para o léxico da língua e sejam registradas nos dicionários.

Quando passamos a analisar as quantidades de ocorrência por palavras, constatamos que todas as palavras foram encontradas no período pesquisado (1998 a 2015). Entretanto, há palavras com número de ocorrências baixíssimo. Por exemplo, *arjear* (4) e *updeitar* (8), indicando que essas formas neológicas não vingaram, uma vez que foram substituídas por outras: *arjear* foi substituída por *zipar*; *updeitar*, pela forma com o verbo suporte – fazer *update* – e pela forma vernácula ‘atualizar’ que adquiriu uma nova acepção – substituir um sistema trocando por uma versão mais recente.

As palavras com maior número de ocorrências são *deletar*, *linkar*, *inicializar*, *escanear* e *dropar*. Isso indica que essas palavras estão introduzidas ao léxico, contudo, é preciso verificar se já foram dicionarizadas (critério lexicográfico). Por fim, as unidades neológicas com menor

número de ocorrências são *becapear*, *getar*, *inputar*, *butar*, *hobbista*, *debugar* e *startar*, além de *arjear* e *updeitar*, supracitadas.

Depois do levantamento das ocorrências no *corpus* de estudo, buscamos nos dicionários gerais de língua os neologismos em estudo. No Quadro 2, temos os resultados dessas buscas, o símbolo (√) indica que a palavra foi encontrada no(s) dicionário(s), enquanto o símbolo (-) indica que não.

Quadro 2 - Registro dos neologismos nos dicionários gerais de língua portuguesa.

NEOLOGISMO	DICIONÁRIO GERAL			
	Ferreira	Aulete	Michaelis	Houaiss
Arjear	-	-	-	-
Becapear	-	-	-	-
Butar	-	-	-	-
Debugar	-	-	-	√
Deletar	√	√	√	√
Dropar	√	√	-	-
Escanear	√	√	√	√
Getar	-	-	-	-
Hobbista	-	-	-	-
Inicializar	√	√	√	√
Inputar	-	-	-	-
Linkar	-	-	-	-
Startar	-	-	-	√
Updeitar	-	-	-	-
Zipar	√	√	√	√

Fonte: Elaborado pelos autores

No quadro, constatamos que as palavras *arjear*, *becapear*, *butar*, *getar*, *hobbista*, *linkar* e *updeitar* não estão registradas em nenhum dos dicionários consultados. Isso indica que elas continuam sendo formas neológicas e talvez sejam usadas apenas dentro da área de informática.

Já as palavras *debugar*, e *startar* (*estartar*) estão registradas em apenas um dicionário, o Houaiss (2009). A palavra *dropar* foi encontrada em dois dicionários, Ferreira (2010) e Aulete (2015), porém não apresenta o sentido descrito por Assirati (1998), qual seja “excluir procedimentos, tabelas, etc.”. Está registrada apenas com o sentido de descer ou deslizar, referindo a manobras feitas por esquetistas e surfistas. Por fim, as formas *deletar*, *escanear*, *inicializar* e *zipar* estão registradas nos quatro dicionários consultados.

Levando em conta o critério lexicográfico podemos afirmar que *debugar*, *startar*, *deletar*, *escanear*, *inicializar* e *zipar* deixaram de ser neologismos e já fazem parte do léxico geral da língua portuguesa. O critério lexicográfico está em consonância com a quantidade de ocorrências nas buscas do Google: *deletar* - 2.810.000; *inicializar* - 94.600; *escanear* - 67.100; *startar* - 9.850; *debugar* - 8.520; *zipar* - 7.250. A única exceção é *linkar* que teve uma alta quantidade de ocorrências, 114.000 (cento e quatorze mil ocorrências), mas não está registrada ainda em nenhum dos dicionários, mesmo em sua forma aportuguesada (*linicar*). Isso pode ser um indício de que os dicionários brasileiros não utilizam o critério da frequência para incorporar novas palavras às suas nomenclaturas.

Investigamos também como os neologismos em estudo estão registrados nos dicionários escolares de língua portuguesa e nos dicionários de aprendizagem de língua inglesa, uma vez que esses dicionários se destinam a um público jovem que faz uso intenso de tecnologia no seu cotidiano. No Quadro 3, temos os resultados da pesquisa dos registros, ou não, das formas neológicas nos quatro dicionários tipo 4 do MEC, que se destinam a alunos do ensino médio.

Quadro 3- Registro dos neologismos nos dicionários escolares tipo 4

NEOLOGISMO	DICIONÁRIO ESCOLAR (TIPO 4)			
	Bechara	Houaiss	Borba	Aulete
Arjeaar	-	-	-	-
Becapeaar	-	√	-	-
Butaar	-	-	-	-
Debugaar	-	-	-	-
Deletaar	√	√	√	√
Dropaar	√	-	-	√
Escanear	√	√	√	√
Getaar	-	-	-	-
Hobbista	-	-	-	-
Inicializar	√	√	√	√
Inputaar	-	-	-	-
Linkaar	-	-	-	-
Startaar	-	-	-	-
Updeitaaar	-	-	-	-
Zipaar	√	√	√	√

Fonte: elaborado pelos autores

Da mesma forma que os dicionários gerais, os dicionários escolares tipo 4 registram os neologismos *deletar*, *escanear*, *inicializar* e *zipar*. Registram também *dropar* apenas com o sentido de descer ou deslizar, referindo-se a manobras de esquiteistas e surfistas.

Diferentemente dos dicionários gerais, os escolares do tipo 4 não registram *debugar* e *startar* e somente o Houaiss (2011) registra *becapear*. Não estão registradas as palavras *arjear*, *butar*, *getar*, *hobbista*, *inputar*, *linkar* e *updeitar*.

De forma geral, com exceção de “arj”, todas as palavras foram encontradas também nos dicionários de inglês consultados. Apesar de em português haver a ocorrência do neologismo *arjear*, em inglês, se trata tão somente de uma extensão de programa de computadores (.arj) e não uma palavra de verdade. Os quatro dicionários de inglês consultados registram os verbos *delete*, *backup*, *scan*, *input* e *update* e seus usos relacionados à informática.

Os dois dicionários *online* consultados e o Dicionário Prático Bilíngue (2008) registram os verbos *boot* e *debug*. Contudo, o dicionário Oxford (2009) não registra essa palavra como verbo, somente como substantivo e sem relação com computadores.

Os verbos *drop*, *get*, e *start* estão registrados em todos os dicionários, mas sem qualquer menção ao uso específico na informática. O mesmo fato ocorre com o substantivo *hobby*.

Apesar de todos os dicionários de inglês consultados registrarem o verbete *link*, apenas o dicionário Oxford (2009) faz uma menção específica ao uso desse verbo ou substantivo na informática. Somente os dicionários de inglês *online* registram o verbo *initialize* e seu uso relacionado à informática.

Apenas o dicionário *online* Cambridge (2015) registra o verbo *zip* com o significado de “comprimir ou reduzir arquivos de computador”. Os outros dicionários, apesar de registrarem esse verbo, tão somente o fazem com o significado de “fechar com zíper”.

No decorrer da investigação, percebemos que há uma instabilidade no registro escrito dos neologismos em estudo. Fazemos, a seguir, uma descrição mais detalhada de cada um desses neologismos para descrever suas várias formas de registro, bem como compreender os sentidos que são efetivamente empregados pelos usuários.

Arjear é o “ato de compactar um arquivo usando o programa ARJ”. Como podemos constatar na tabela 2, há poucas ocorrências dessa palavra (04). Nas buscas no *Google*, encontramos também duas ocorrências de *desarjear*. Talvez isso ocorra porque esse programa

era utilizado para compactar arquivos do primeiro sistema operacional MS-DOS, que caiu em desuso depois do lançamento dos sistemas da família *Windows*. Dessa forma, a obsolescência tecnológica faz com que a palavra tenha pouco uso, restringindo a contextos específicos e fazendo com que ela não entre para o léxico geral da língua. Essa palavra não está registrada em nenhum dos dicionários consultados. Possivelmente, foi um termo muito utilizado pelos usuários do sistema MS-DOS, porém, hoje, tornou-se arcaico ou antiquado, acompanhando a obsolescência do referido sistema. Foi um neologismo que não vingou. As poucas ocorrências estão em textos que contam a história da informática.

Nenhum dos dicionários em inglês consultados, sejam os *online* monolíngues ou os escolares bilíngues, registra ARJ. Fato que não nos surpreende, uma vez que se trata de uma extensão de programa de computadores, ou seja, não se refere a uma palavra propriamente dita.

Becapear é “fazer uma cópia de um arquivo, ou dados ou disco.” Essa palavra apresenta duas grafias: *becapear* (105 ocorrências) e *backupear* (1.070 ocorrências). Essa forma parece não ser a preferida dos usuários, que acabam utilizando mais a forma com verbo suporte, “fazer backup” que teve 75.300 ocorrências nas buscas no *Google*. Essa expressão também apresenta instabilidade gráfica, sendo registrada como “fazer back up” (4.090 ocorrências) e “fazer becape” (117). Comparando a quantidade de ocorrências das duas formas (*becapear* e *fazer backup*), percebemos que a segunda é a preferida dos usuários e pela quantidade de ocorrências talvez já fosse possível estar dicionarizada. O que não ocorre em nenhum dos dicionários consultados. Apenas o dicionário Houaiss (2011) registra a forma *becapear* que tem uma baixa frequência de uso. Vale salientar que tanto os dicionários gerais quanto os escolares consultados registram *backup* na escrita estrangeira ou *becape* aportuguesada, mas nenhum registra a forma “fazer backup”. Ambos os dicionários monolíngues trazem uma entrada em que especificam ser esse um termo da computação.

Os dois dicionários escolares registram também o termo, especificando seu uso na informática e ambos traduzem o verbo como “fazer *backup*” o que corrobora a baixa frequência de uso em português do neologismo *becapear*. O dicionário Oxford (2009) traz inclusive uma nota de referência remetendo o usuário para a entrada computador, em que há uma figura desse objeto e um microglossário das palavras mais comuns relacionadas a ele.

Butar é o ato de dar início a um programa de computador. Nas buscas que fizemos utilizando a grafia “butar”, a maioria dos casos se referia a variação da escrita do verbo “botar”. Então, fizemos outras buscas utilizando a forma “bootar” que retornou 11.700 ocorrências. Fizemos também uma busca com a forma “dar boot” que resultou em 21.400, sendo, portanto, mais usada do que a forma “bootar”. Nenhum dos dicionários consultados registra a forma “butar”, porém alguns registram a forma “boot”. Aulete (2011, 2015) registra como subentrada a forma “dar boot”. Portanto, a forma “butar”, por não está registrada nos dicionários, ainda continua sendo uma forma neológica, mesmo o sentido já tendo sido registrado de outra forma (*dar boot*).

O verbete *boot* e a variante *boot up* são registrados com o sentido relacionado ao computador no dicionário Cambridge (2015). O Collins (2005) registra os diferentes usos dessa palavra que pode ser usada tanto como substantivo como verbo em inglês. Dentre os usos, dois são registrados como específicos da informática, ambos como verbos. O Dicionário Bilingue Prático (2008) registra o uso do verbo como inicializar computadores, programas etc. Contudo, o dicionário Oxford (2009) somente registra o uso como substantivos.

Debugar é o ato de “retirar os defeitos de um programa (depurar).” Nas buscas que realizamos na *Internet*, obtivemos 8.520 ocorrências dessa palavra, que está registrada apenas no dicionário Houaiss (2009). Pesquisamos também a forma neológica *bugar* que teve 34.330 ocorrências e as formas flexionadas *bugou* e *bugado* que tiveram 34.100 e 107.000 ocorrências, respectivamente. Isso demonstra que o verbo *bugar* tem um uso mais frequente do que *debugar*. Pesquisamos ainda a expressão “tirar o bug”, que, com apenas 1.490 ocorrências, não se mostrou muito produtiva. Enfim, como já foi registrada em um dicionário *debugar* não é mais uma forma neológica, contudo, devido sua alta frequência, a forma *bugar* também já deveria estar registrada nos dicionários.

Ambos os dicionários *online* (inglês) registram o verbo *debug* e seus usos relacionados à informática, como também outros usos. Dentre os dicionários escolares, somente o Dicionário Bilingue Prático (2008) registra esse uso.

Deletar é a palavra com o maior número de ocorrências (2.810.000) e com registros em todos os dicionários. Isso demonstra que essa palavra está plenamente integrada ao léxico da língua portuguesa, chegando inclusive a ser empregada em sentido figurado e fora do contexto

da informática. Por exemplo: “vou te deletar do meu coração”. Portanto, não é mais uma forma neológica. Por sua vez, os quatro dicionários de inglês pesquisados registram o verbo *delete* e seu uso relacionado à informática.

Na área de informática, **dropar** significa “excluir procedimentos, tabelas etc.” Já no esporte, significa o surfista ou esquetista descer de uma onda ou rampa. Nas pesquisas realizadas na *Internet*, *dropar* retornou 41.900 ocorrências, no entanto, analisando as primeiras páginas de resultados retornadas pelo *Google*, não encontramos nenhuma ocorrência no sentido da área de informática. Ela está registrada apenas com o sentido da área do esporte nos dicionários Ferreira (2011), Aulete (2011) e Bechara (2011). Enfim, como sentido de excluir procedimentos não está registrado, *dropar* continua sendo uma forma neológica no campo da informática. Vale salientar que a pesquisa de neologismos com LC precisa ser feita também com análises qualitativas, uma vez que as ferramentas computacionais só conseguem rastrear neologismos de forma, não conseguindo detectar os neologismos semânticos. Isso também se aplica quando usamos a *Internet* como *corpus* de estudo.

Nenhum dos quatro dicionários de inglês registra um uso específico do verbo *drop* na informática. Somente o Oxford (2009) registra o significado “excluir”, mas não necessariamente relacionado à informática.

Escanear significa passar uma imagem ou texto para o computador por meio de um escâner. Nas buscas no *Google*, encontramos 67.100 ocorrências. Buscamos também a forma *scanear* que resultou em 16.400. Pesquisamos também as formas flexionadas *escaneou* (3.070), *escaneie* (37.100) e *escaneado* (49.300). Se considerarmos as várias grafias e as formas flexionadas, teremos mais de 100.000 mil ocorrências. Todos os dicionários consultados registram a palavra. Além do sentido da área de informática, os dicionários Houaiss (2009) e Aulete (2015) registram também um sentido da medicina: 1. varrer (algo) com um escâner. Ex.: “escanear o baço de um paciente” (HOUAISS, 2009); 2. Med. Obter informações minuciosas sobre (órgão do corpo) varrendo-o com um escâner: *escanear* o fígado/o útero. (AULETE, 2015). Enfim, *escanear* também já entrou para o léxico geral da língua portuguesa, ou seja, não é mais um neologismo.

Todos os dicionários de inglês consultados registram o verbo *scan* e seus diversos usos, incluindo o mais específico da informática.

Getar significa entrar no sistema. Trata-se de uma palavra que teve poucas ocorrências nas buscas, com apenas 461 ocorrências no período analisado. O baixo índice de ocorrências talvez seja porque outras palavras foram adotadas para o sentido de “entrar no sistema” ou “iniciar o sistema”. É um uso restrito ao contexto específico da informática, não estando registrado em nenhum dos dicionários consultados. Portanto, essa palavra ainda pode ser considerada como uma forma neológica.

O verbo *get* é bastante usado em inglês, e todos os dicionários consultados trazem vários significados e usos dele. Contudo, nenhum é registrado como específico para a informática ou computação. Existem outros verbos que são usados para denominar a mesma ação, tais como *login* ou *sign in*. Provavelmente, esses termos são mais utilizados porque são específicos da área da informática, enquanto *get in(to)* pode ser utilizado em vários outros contextos.

Na área de informática, **hobbista** designa aqueles que têm um *hobby* que é praticado por meio do computador. Nas buscas na Internet, encontramos 3.300 ocorrências para esta palavra, entretanto, da mesma forma que *dropar*, quando analisamos os resultados das primeiras páginas constatamos que o uso de *hobbista* está relacionado àquele indivíduo que gosta de se divertir praticando uma determinada atividade. Dessa forma, não podemos afirmar que as 3.300 ocorrências são da acepção que a palavra tem dentro da informática. Para fazer isso é preciso analisar ocorrência por ocorrência em cada *site* listado pela busca do *Google*, o que foge ao foco deste trabalho. Nos dicionários de língua portuguesa consultados, nenhum registra a palavra *hobbista*, porém registram o estrangeirismo *hobby*. Portanto, trata-se de uma forma que ainda não entrou para o léxico da língua, continuando com uma forma neológica na área de informática.

Em todos os dicionários de inglês consultados, encontramos a palavra *hobby*. Contudo, acreditamos que o neologismo *hobbista* venha do substantivo *hobbyist*, que designa quem pratica ou tem um *hobby*. Cabe ressaltar que somente os dois dicionários *online* registram essa palavra (*hobbyist*). O *Cambridge* (2015) traz ‘*a computer hobbyist*’ como exemplo de uso para esta entrada. É possível também encontrar artigos e notícias da área de informática na *Internet*, onde o termo é usado para diferenciar o programador profissional daquele não licenciado para exercer tal função, sendo possível perceber certa conotação negativa em seu uso nesse caso.

Inicializar é colocar o computador para funcionar, iniciar, ligar. É uma palavra que tem um número de ocorrência elevado (94.600 para inicializar, 4.900 para inicializou, 20.400 para inicialize e 26.600 para inicializado), somando mais de 150.000 mil ocorrências. Todos os dicionários consultados já registram a palavra “inicializar”, portanto, já entrou para o léxico geral da língua portuguesa.

Somente os dicionários de inglês *online* registram esse verbo e seu uso relacionado à informática. Os dicionários escolares bilíngues registram apenas o verbo *initiate* que deu origem a *initialize*.

Na área de informática, **inputar** significa inserir dados no computador. Nas buscas na Internet, obtivemos 569 ocorrências no período analisado, um número que pode ser considerado baixo, indicando que é uma palavra pouco usada. Nenhum dos dicionários consultados registra essa palavra, nem mesmo como uma acepção da palavra *imputar* ou como homônimo. Porém, registram o estrangeirismo *input*. Enfim, podemos considerar que *inputar* continua sendo um neologismo, não tendo entrado ainda para o léxico geral da língua portuguesa. Esse verbo está registrado em todos os dicionários de inglês consultados com referência ao uso específico na informática.

Linkar significa ligar um *link* ou *hiperlink*. Nas buscas feitas na Internet, foi uma das palavras com maior número de ocorrências (114.000). Buscamos também com a grafia aportuguesada *lincar*, que resultou em 3.890 ocorrências. Já a busca com as formas flexionadas *linkou*, *linkei*, *linkado* retornou 11.900, 14.500 e 62.300 ocorrências, respectivamente. Somando as ocorrências de todas as formas, constata-se que *linkar* é uma palavra muito usada, mas que ainda não foi registrada em nenhum dos dicionários consultados. Apenas o estrangeirismo *link* está registrado em todas as obras consultadas. O caso de *linkar* é muito interessante, porque mesmo tendo uma frequência de uso grande, ainda não foi registrada nos dicionários, que trazem palavras menos frequentes como *becapear* e *debugar*. Isso pode ser um indício de que os dicionários de língua portuguesa não usam a frequência de uso baseada em *corpus* para incorporar novas palavras às suas nomenclaturas. Enfim, mesmo *linkar* tendo um número grande de ocorrências, pelo critério lexicográfico, ainda é considerado um neologismo.

Apesar de todos os dicionários de inglês consultados registrarem o verbete *link*, apenas o dicionário Oxford (2009) faz uma menção específica ao uso desse verbo ou substantivo na

informática. Os outros não fazem esse registro, talvez porque o significado nessa área é o mesmo do uso comum (elo, ligação, unir etc.).

Startar significa iniciar o computador. Obtivemos 9.850 ocorrências dessa palavra nas buscas da *Internet*. É uma palavra que ainda sofre flutuação na grafia, sendo usadas as formas *startar* e *estartar* (591). Apenas o dicionário Houaiss (2009) a registra com a forma aportuguesada. Portanto, é uma palavra que já está incorporada ao léxico geral da língua.

Essa palavra está registrada em todos os dicionários de inglês consultados, mas sem qualquer menção específica ao uso na área da informática. Acreditamos que não seja necessário porque a mesma não adquire um significado específico nesse contexto.

Updeitar é o ato de atualizar a informação por meio da alteração ou adição de dados específicos em um arquivo principal. É uma palavra que apresenta uma flutuação gráfica grande, sendo encontradas as formas *updatar* e *updatear*. As buscas na *Internet* retornaram poucas ocorrências: *updeitar* – 8; *upadatar* – 102; e *updatear* – 48. Isso demonstra que é uma palavra pouco utilizada. Buscamos também a expressão “fazer *update*” que resultou em 2.950 ocorrências, indicando certa preferência pela forma com o verbo suporte. Há também preferência pela forma vernácula atualizar com mais de 100.000 (cem mil) ocorrências. Essa nova acepção da palavra atualizar já está dicionarizada, Ferreira (2010) e Aulete (2015) já a registram. Contudo, nenhum dos dicionários consultados registra *updeitar* nem o estrangeirismo *update*. Portanto, podemos considerar que essa palavra ainda é um neologismo.

O verbo *update* está registrado em todos os dicionários de inglês consultados sendo traduzido como “atualizar”, o que corrobora a baixa frequência do neologismo *updeitar*.

Zipar significa compactar arquivos. As buscas que realizamos na *Internet* resultaram em 7.250 ocorrências. Buscamos também pelo particípio, “zipado”, que resultou em 35.600 ocorrências. Já nas consultas aos dicionários, verificamos que todos registram esta palavra. Destarte, não é mais neologismo, estando incorporada ao léxico geral da língua portuguesa.

Dos dicionários de inglês consultados, apenas o Cambridge (2015) registra o verbo *zip* com o significado de “comprimir ou reduzir arquivos de computador”. Todos os outros registram

esse verbo somente com o significado de “fechar com zíper”. No entanto, em ambos os dicionários *online* e no Oxford (2010), encontramos registrada a forma *unzip* (descompactar).

CONSIDERAÇÕES FINAIS

A criação de novas palavras e a sua incorporação ao léxico revelam a dinamicidade das línguas e os vínculos estreitos que mantêm com a cultura, a sociedade e o momento histórico em que ela é usada. Hoje, a rapidez com que novas tecnologias são criadas e postas em consumo faz com que as línguas tomem de empréstimo palavras da língua onde aquelas tecnologias foram criadas e denominadas inicialmente. Por outro lado, o inglês, como língua das ciências e dos negócios, tem sido a língua preferida para denominar a maioria dos novos produtos tecnológicos. Isso faz com que seja, hoje, a língua que mais empresta palavras às demais, em especial, à língua portuguesa, sobretudo na área de informática. Contudo, esses empréstimos, mesmo quando mantêm sua forma estrangeira, moldam-se ao sistema linguístico da língua receptora. Podem ser adotados por um tempo, mas sem a certeza de entrarem definitivamente para o léxico da referida língua. Vale salientar que a rapidez com que tecnologias são substituídas também pode interferir nesse processo.

Realizamos esta pesquisa com o propósito de analisar e discutir como neologismos formados por processos vernaculares, a partir de bases estrangeiras, nesse caso do inglês, foram incorporados ao léxico da língua portuguesa do Brasil. Para isso, levantamos as ocorrências de uso dessas palavras na *Internet*, no período de 1998 a 2015, subsequente à descrição das mesmas como neologismos. Os resultados das análises mostram que:

- As palavras *becapear*, *debugar*, *deletar*, *escanear*, *inicializar*, *estartar* e *zipar* deixaram de ser neologismos, passando a fazer parte do léxico geral da língua portuguesa;
- As palavras *arjeaar*, *butar*, *dropar*, *getar*, *hobbista*, *inputar*, *linkar* e *updeitar* continuam sendo neologismos, uma vez que ainda não estão registradas em nenhum dos dicionários de língua portuguesa consultados;
- *Arjeaar* pode ser considerada uma palavra obsoleta, pois tem apenas 4 ocorrências. Em inglês, a forma ARJ não foi registrada em nenhum dicionário, possivelmente seja também um neologismo;

- *Linkar* mesmo tendo um número expressivo de ocorrências não está ainda registrada nos dicionários. Isso pode indicar que os dicionários de língua portuguesa não utilizam a frequência de uso para atualizar sua nomenclatura e incorporar novas palavras aos dicionários. Dado o grande número de ocorrências e o uso constante e crescente dessa palavra, sugerimos a sua incorporação ao léxico geral da língua portuguesa registrado nos dicionários;
- Em alguns casos, parece que há uma preferência pela formação com o verbo suporte: *dar boot*, *fazer becape* (*back up*). Talvez isso se dê porque as formas *butar* e *becapear* soam estranhas aos nossos ouvidos. No caso de *butar*, há o verbo botar em português, pronunciado como [butar], talvez a preferência pela forma com o verbo suporte seja para evitar homonímia com o verbo já existente;
- As palavras *arjear* e *zipar* são sinônimas e disputam a preferência pelo uso, sendo que *zipar* acabou sendo incorporada ao léxico, uma vez que está associada a uma tecnologia que ainda é usada, o que não ocorre com *arjear*;
- As palavras *getar*, *inicializar* e *startar* também são sinônimas, sendo *inicializar* a preferida pelos usuários.

Por fim, vale salientar que a incorporação de novas palavras ao léxico registrado deve ser feita pela frequência de uso, pois é a comunidade linguística que decide no uso cotidiano as melhores formas de denominar a realidade.

REFERÊNCIAS

ALUISIO, Sandra Maria e ALMEIDA, Gladis Maria Barcelos. O que é e como se constrói um corpus? Lições aprendidas na compilação de vários *corpora* para pesquisa linguística. **Caleidoscópio**. Vol. 4, n.3, p. 156-178, set/dez 2006. Unisinos.

ALVES, Ieda Maria. **Neologismo**. Criação lexical. São Paulo: editora Ática, 2004, 2ªed.

_____. Uma metodologia para descrição dos neologismos. In: GONÇALVES, A. V. e GÓIS, M. L. S. (org.). **Ciências da Linguagem**: o fazer científico? Campinas – SP: Mercado das Letras, 2012. Vol. 1. pag. 179 a 195.

ASSIRATI, Elaine. Therezinha. Neologismos por empréstimo na informática. **Alfa**, São Paulo, 42(n.esp.): 121-145, 1998.

CORREIA, Margarita e ALMEIDA, Gladis Maria de Barcellos. **Neologia em português**. São Paulo: Parábola Editorial, 2012.

MELLO, Heliana Ribeiro de; SOUZA, Renato Rocha. A linguagem da ciência: prospecção de dados baseados em corpora. **Texto Livre: Linguagem e Tecnologia**, [S.l.], v. 7, n. 1, p. 158-166, Jul. 2014. ISSN 1983-3652. Disponível em: <<https://goo.gl/YtFZRq>>. Acesso em: 01 Set. 2015.

RIOS, Tatiana Helena Carvalho. **A Lingüística de Corpus para a descrição de idiomatismos**. Disponível em: <encurtador.com.br/gyBIL>. Acesso em: 01 set. 2015.

SANCHEZ, Aquilino. **Corpus Lingüístico del español contemporáneo: fundamentos, metodología y aplicaciones**. Madrid: SGEL, 1995.

SINCLAIR, John. "Corpus and Text - Basic Principles". In: **Developing Linguistic Corpora: a Guide to Good Practice**, ed. M. Wynne. Oxford: Oxbow Books, 2005. 1-16. Disponível em: <<https://goo.gl/untqC5>> [Acesso em: 19 de jun de 2017.](#)

TORRANO, Sandra Delneri Petean. **Produtividade e Criatividade do Léxico: os neologismos na área da informática**. 123 f. Dissertação (Mestrado). Programa de Pós-Graduação em Filologia e Língua Portuguesa. Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2010.

Dicionários consultados em Língua portuguesa:

Dicionários gerais

AULETE, Caldas. **Aulete Digital o dicionário da língua portuguesa**. Disponível em: <http://www.aulete.com.br/>. Acesso em: 27/08/2015.

FERREIRA, Aurélio Buarque de Holanda. **Dicionário Eletrônico Aurélio – Versão 5.8.3**. Curitiba: Editora Positivo, 2010.

HOUAISS, Antônio. **Houaiss Eletrônico - versão 3.0**. Rio de Janeiro: Editora Objetiva, 2009.

WEISZFLOG, Walter (Editor). **Michaelis Dicionário de Português Online**. Editora Melhoramentos, 2012. Disponível em: <http://zip.net/bqrVsF>. Acesso em: 27/08/2015.

Dicionários escolares tipo 4

AULETE, Caldas. **Novíssimo Aulete dicionário contemporâneo de língua portuguesa** [organizador Paulo Geiger]. Rio de Janeiro: Lexikon, 2011.

BORBA, Francisco S. **Dicionário Unesp do português contemporâneo**. Curitiba: Piá, 2011.

HOUAISS, Antônio. **Dicionário Houaiss Conciso**. [organizador Mauro de Salles Villar] São Paulo: Moderno, 2011.

BECHARA, Evanildo. **Dicionário da língua portuguesa Evanildo Bechara**. Rio de Janeiro: Editora Nova Fronteira, 2011.

Dicionários consultados em Língua inglesa:

Dicionários gerais online

Cambridge Dictionaries online. Disponível em: <<https://goo.gl/FqBn9n>> Acesso em: 01/09/2015.

Collins online. Disponível em: <<https://goo.gl/nBUH3T>> Acesso em: 01/09/2015.

Dicionários escolares bilíngues

Minidicionário Bilíngue Prático: inglês-português, português-inglês. 2ª Ed. Curitiba. Ed. Positivo, 2008.

Dicionário Oxford escolar para estudantes brasileiros de inglês: português-inglês e inglês-português. Edição atualizada. Oxford University Press. 2009.

Francisco Iacó do NASCIMENTO

Doutorando e Mestre em Linguística Aplicada pela Universidade Estadual do Ceará com interesse na área de Lexicografia (2013). Possui Especialização em Língua Portuguesa: Leitura e Produção de textos, pela Faculdade Vale do Jaguaribe (2006), e Especialização em Gestão e Avaliação da Educação Pública, pela Universidade Federal de Juiz de Fora (2011), graduação em Letras pela Universidade Estadual do Ceará (1998). Atualmente é professor efetivo da EEM José Francisco de Moura, pertencente a rede de ensino do Estado do Ceará. Membro do Grupo de Pesquisa Lexicografia, Terminologia e Ensino - LETENS/UECE.

Sarah Virgínia Carvalho RIBEIRO

Doutoranda em Linguística Aplicada na Universidade Estadual do Ceará - UECE, instituição onde concluiu seu mestrado em Letras - Língua Inglesa (1997) e graduação (1987) em Letras - Português e Língua Estrangeira. Coordenou o Programa Mulheres Mil no Campus Fortaleza do Instituto Federal de Educação, Ciência e Tecnologia do Ceará onde é docente desde 1993



Antônio Luciano PONTES

É Mestre em Língua Portuguesa (UFPB). Possui doutorado em Lingüística (UNESP), sob a orientação da Professora Doutora Ieda Maria Alves (USP). Membro do grupo de pesquisa cadastrado no CNPq Lexicografia, Terminologia e Ensino (LETENS). Professor Titular no curso de Letras da UERN, onde está ligado ao Programa de Pós-graduação em Letras (PPGL) e Mestrado Profissionalizante de Letras (PROFLETRAS). Está vinculado ao Programa de Pós-graduação de Linguística Aplicada (PosLA) da Universidade Estadual do Ceará.

Recebido em 13 julho 2019 - Aceito em 23 agosto 2019