

Princípios metodológicos para a detecção de neologismos da comunicação digital

DOI: <http://dx.doi.org/10.21165/el.v50i1.2961>

Ana Maria Ribeiro de Jesus¹

Resumo

O objetivo deste artigo é descrever alguns princípios metodológicos para a coleta e validação de neologismos da comunicação digital e, paralelamente, discutir os possíveis rumos a que o mundo digital conduz esses princípios. Com base nos estudos de Alves (2012), Cabré (2016), Auger (2010), Estopà (2015), Rey (1976), dentre outros, são apresentadas, inicialmente, as fases tradicionais do trabalho neológico e os principais critérios estabelecidos para a validação de neologismos. Abordamos, então, algumas limitações desses critérios, seguidas de propostas metodológicas que podem se somar a eles. Por fim, sintetizamos os passos metodológicos adotados na detecção das unidades lexicais neológicas de nossa pesquisa (dentre elas: *coronavirus*, *biscoitar*, *digital*, *intoxicação*, *lacrosfera* e *quarentener*), enfatizando os critérios que se mostraram funcionais em sua coleta e validação.

Palavras-chave: neologia; procedimentos metodológicos; comunicação digital.

¹ Universidade Federal do Espírito Santo (UFES), Vitória, Espírito Santo, Brasil; ana.m.jesus@ufes.br; <https://orcid.org/0000-0001-5479-5564>

Methodological principles for the detection of digital communication neologisms

Abstract

The objective of this article is to describe some methodological principles for the collection and validation of digital communication neologisms, and tangentially, to discuss the possible directions in which the digital world leads these principles. Based on studies by Alves (2012), Cabré (2016), Auger (2010), Estopà (2015), Rey (1976), among others, we present the traditional phases of neological work and the main criteria established for the validation of neologisms. Then, we address some limitations of those criteria, followed by methodological proposals that can work as additional steps. Finally, we synthesize the methodological steps adopted in the detection of the neological units of our research (among them: *coronavirus*, *biscoitar*, *figital*, *infoxicação*, *lacrosfera* and *quarentener*), emphasizing the criteria that proved to be functional in their collection and validation.

Keywords: neology; methodological procedures; digital communication.

Introdução

Os trabalhos de pesquisa em neologia envolvem, de modo geral, cinco grandes momentos após a escolha do subconjunto linguístico a ser analisado: a coleta do *corpus* textual que contém os candidatos a neologismos; o contraste desse *corpus* com outro maior, que contém os itens lexicais já estabilizados na língua; a detecção dos neologismos e sua validação a partir de critérios preestabelecidos; a descrição tipológica, estabelecendo-se uma postura teórica; e o monitoramento, que submete os neologismos a um período de “teste”, no intuito de observar seu desaparecimento ou sua lexicalização.

Essas fases podem variar de acordo com a área a que pertencem as unidades lexicais da pesquisa, o gênero textual que constitui o *corpus*, o enfoque da pesquisa e a instrumentação com que se trabalha, dentre outras motivações. Além disso, alguns pesquisadores contam com as novas tecnologias para o tratamento automatizado ou semiautomatizado em todas as fases, enquanto outros contam com esse recurso em apenas algumas fases, geralmente no momento do contraste entre os *corpora*.

De qualquer forma, é fundamental que todo esse arsenal metodológico acompanhe a notável difusão dos neologismos – efêmeros ou duradouros – na revolução digital dos últimos anos. A presença marcada de inúmeras comunidades linguísticas no mundo *on-line* enfatizou os processos de criação dessas unidades, representativas, em grande parte, da cultura global digital, das culturas de pequenos grupos e de um mundo em franca transição paradigmática. Nessa perspectiva, o objetivo deste artigo é descrever alguns princípios metodológicos para a coleta e validação de neologismos da comunicação

digital e, paralelamente, demonstrar como a ferramenta Google Trends pode ser utilizada em conjunto com os demais critérios já estabelecidos na área. Para tanto, inicialmente, são apresentadas as fases tradicionais do trabalho neológico e os principais critérios estabelecidos para a validação de neologismos. Abordamos, então, algumas limitações desses critérios, seguidas de propostas metodológicas que podem se somar a eles. Por fim, sintetizamos os passos metodológicos utilizados na coleta dos neologismos do *corpus* de análise de nossa pesquisa, enfatizando os critérios que se mostraram funcionais em sua coleta e validação.

As fases do trabalho neológico

Os primeiros estudos sistemáticos do fenômeno da neologia surgiram com os chamados Observatórios de neologismos, que têm a finalidade de analisar a criação lexical de um viés descritivo, observando a formação de palavras por meio dos mecanismos internos à língua ou por meio de elementos recebidos de idiomas estrangeiros. Os observatórios nasceram na década de 60, a partir da iniciativa pioneira do professor Bernard Quemada, que instaurou, na Universidade de Besançon, na França, o *Observatoire du français contemporain*, no intuito de analisar a neologia lexical a partir de um *corpus* jornalístico.

Nesses estudos iniciais, a detecção, a coleta e o tratamento dos neologismos eram feitos de forma totalmente artesanal. Atesta Cabré (2015) que os pesquisadores que iniciaram esse trabalho tinham a vasta tarefa de ler jornais inteiros diariamente, e levantar, ao longo da leitura, os candidatos a neologismos. Para tanto, contavam, em primeiro lugar, com a própria intuição linguística. Em segundo lugar, serviam-se de um conjunto dos principais dicionários de língua geral, selecionados como referência. Dessa forma, a intuição era confirmada por meio da comparação com as unidades registradas nas obras lexicográficas.

O avanço das novas tecnologias e sua aplicação nos Estudos do léxico abriu grandes horizontes para o trabalho de coleta, validação e descrição de neologismos, principalmente com os recursos provenientes da Linguística computacional e da Linguística de *corpus*. Entretanto, muito se deve aos primeiros observatórios de neologia que, desde o seu início, criaram as bases para a pesquisa da área e a metodologia que ainda hoje é adotada por grande parte dos pesquisadores, principalmente no que concerne ao chamado “*corpus* de exclusão”, sobre o qual comentaremos adiante.

A metodologia da pesquisa em neologia pode variar para atender os objetivos propostos em cada análise. É possível, também, que ela passe por ajustes no andamento da pesquisa, à medida em que novos problemas vão surgindo com as unidades coletadas, ou por causa de determinada situação não prevista anteriormente. Alguns passos, no entanto, são comuns à maior parte dos trabalhos. Observemos, então, os pontos em comum dos princípios adotados nas pesquisas de Adelstein (2015), Falk, Delphine e

Christophe (2018) e Alves (2012) para ter um panorama dos passos fundamentais de toda pesquisa neológica.

A metodologia apresentada por Adelstein (2015) é composta pelas três fases que, de acordo com a autora, são típicas da metodologia de trabalho dos principais observatórios de neologia da atualidade. Apresentando-as de forma sintetizada, são as seguintes:

1. Fase de detecção: Fase inicial que leva à reflexão acerca dos problemas em torno da teoria lexical, porque se trata do momento de decidir se uma forma lexical é nova.
2. Fase de contraste: Geralmente, na metodologia dos observatórios, o contraste com um *corpus* de dicionários é a maneira de confirmar o sentimento neológico.
3. Fase de tipologização: As unidades detectadas e validadas como neológicas são aplicadas a uma tipologia, elaborada de acordo com critérios teórico-metodológicos.

De modo semelhante, o estudo de Falk, Delphine e Christophe (2018) apresenta uma metodologia composta por três estágios, acompanhados pela especificação do tipo de processamento de cada um deles: manual ou automatizado. O objetivo dos autores é documentar os neologismos em seus respectivos contextos, a fim de que sejam disponibilizados para consulta *on-line*. As fases podem ser assim resumidas:

1. Fase de detecção: Extração automática de palavras novas, com base em um *corpus* de exclusão (processamento automático).
2. Fase de seleção: As palavras novas detectadas são validadas manualmente, ou seja, um especialista decide quais delas são, de fato, neológicas (processamento manual).
3. Fase de descrição: As palavras neológicas validadas são documentadas (processamento automático e manual).

Da metodologia corrente nos estudos de Alves (2012), por sua vez, podemos inferir cinco fases. Dentre elas, está incluída a fase de registro das unidades neológicas em uma base de dados, denominada *Base de neologismos do português brasileiro contemporâneo*, composta por textos do gênero jornalístico. De acordo com a autora, a base passou por diferentes metodologias, seguindo a evolução da Informática no Brasil. Alves inclui, também, um passo analítico em que se quantificam os processos de formação dos neologismos validados. As fases podem ser resumidas da seguinte forma:

1. Fase de detecção: Coleta semiautomática dos candidatos a neologismos, com auxílio de um extrator de neologismos.
2. Fase de contraste: Validação dos candidatos selecionados por meio do contraste com o *corpus* de exclusão, constituído por dicionários de língua, vocabulários ortográficos e por um conjunto de textos pré-determinados.
3. Fase de catalogação: Registro dos neologismos na base de dados, que contém referências a respeito da unidade lexical validada.
4. Fase de tipologização: Classificação do neologismo de acordo com categorias preestabelecidas.
5. Fase de análise: Quantificação da produtividade neológica no que concerne a processos vernaculares e empréstimos.

As três listagens apresentam estágios muito semelhantes, o que mostra, portanto, que determinados procedimentos são fundamentais a qualquer pesquisa em neologia. A análise das três propostas demonstra que a *coleta do corpus*, a *exclusão lexicográfica* e a *validação dos neologismos* são passos comuns a todas elas.

A observação das propostas revela também que, dependendo da ferramenta computacional que se tem à disposição, as partes “humanas” e automáticas podem ocorrer em diferentes momentos. Na proposta de Falk, por exemplo, a análise do especialista para decidir se uma forma é de fato nova ocorre após a parte automatizada, ou seja, após o contraste com o *corpus* de exclusão, em ordem inversa à primeira proposta. A fase 3 dos dois primeiros autores, bem como as fases 3, 4 e 5 de Alves estão ligadas aos objetivos finais de cada trabalho e, nesse caso, a tipologização das unidades lexicais neológicas também pode variar (cf. JESUS, 2020). A análise quantitativa, presente no trabalho de Alves, é um passo de grande importância para os estudos em neologia, uma vez que, na avaliação estatística da produtividade de cada processo de formação, é possível perceber como a linguagem está se movimentando e proceder a análises de cunho pragmático. Pode-se observar, por exemplo, o aumento ou a diminuição do número de neologismos por empréstimo e entender as motivações desse fenômeno.

Critérios tradicionais para a detecção de neologismos

Para que as fases do trabalho neológico se cumpram, é preciso que haja critérios preestabelecidos que determinem o caráter neológico das unidades lexicais. Sabemos que os neologismos, elementos resultantes dos processos de criação lexical, não têm existência por si mesmo; uma unidade neológica o é em relação a determinados parâmetros teóricos e conceituais.

No entanto, mesmo depois de décadas de observação sistemática do fenômeno da neologia nas línguas naturais, a delimitação do conceito de neologismo continua sendo um desafio e constitui uma questão fundamental aos pesquisadores da área. No título de um trabalho recente, Cabré (2016, tradução nossa²) expressa essa inquietação: “Por que é relativamente fácil detectar neologismos e tão complicado definir o que são”. Nesse sentido, muitos esforços têm sido feitos para se determinar critérios ideais para a definição das unidades neológicas. Trata-se de metodologias complementares, que são propostas à medida em que os estudos da linguagem avançam. Entretanto, a complexidade do conceito de neologia, reconhecida na fala de Cabré, mostra que os critérios ainda não são considerados totalmente satisfatórios e que, como consequência, estão sujeitos a questionamentos.

Vejamos, inicialmente, os critérios tradicionais e seus pressupostos para, em seguida, pontuarmos algumas de suas limitações. São clássicos os três critérios estabelecidos por Rey (1976), propostos como parâmetros para se detectar e atestar as novas unidades lexicais. Seriam neologismos os elementos que atenderem ao seguinte:

- (a) Critério temporal: o surgimento da unidade lexical ocorre em período recente.
- (b) Critério psicolinguístico: a intuição dos falantes leva-os a perceber a unidade lexical como nova.
- (c) Critério lexicográfico: a unidade lexical não está registrada nos dicionários de língua geral.

Considerando os três critérios levantados por Rey, Cabré (1993) acrescenta, em seus estudos sobre a determinação do *status* de neologia das unidades candidatas, um quarto critério, que corresponde à observação do comportamento dos possíveis neologismos para atestar se eles são formalmente ou semanticamente instáveis. São estabelecidos, então, como parâmetros:

² No original: “Per què és relativament fàcil de detectar neologismes i tan complicat de definir què són”.

- (a) a diacronia: uma unidade é neológica se surgiu em um período recente.
- (b) a psicologia: uma unidade é neológica se desperta o sentimento neológico dos falantes.
- (c) a lexicografia: uma unidade é neológica se não está registrada nas obras de um *corpus* de exclusão.
- (d) a instabilidade sistemática: uma unidade é neológica se apresenta signos de instabilidade formal (morfológicos, gráficos, fonéticos) ou semântica.

Limitações dos critérios

Vários questionamentos já foram levantados no intuito de argumentar que os critérios temporal, psicológico e lexicográfico podem ser insuficientes para a determinação do caráter neológico de um item lexical. De fato, parece difícil estipular que estes seriam princípios aplicáveis, de forma generalizada, a todo tipo de neologismo. Muitas das expressões neológicas podem ter peculiaridades que escapam aos três critérios. Pode ocorrer, ainda, que os pressupostos de cada um dos critérios sejam “influenciados” pelos dados coletados na análise, ou pelo momento em que a pesquisa é conduzida, ou mesmo pelas obras selecionadas como *corpus* de exclusão, como veremos na sequência.

a. Limitações do critério temporal

O principal questionamento com relação ao primeiro critério, o temporal, está ligado ao caráter de redundância a ele atribuído. Se o “neologismo”, por definição, pressupõe uma unidade lexical nova, então a “novidade” é uma característica inerente a ele. Nesse sentido, a confirmação de que a palavra surgiu em um período recente não poderia ser um critério de detecção de neologismos, já que é propriamente o que o define (ESTOPÀ, 2015). Além disso, é difícil estabelecer com exatidão o momento de surgimento de uma palavra. O que é possível, no trabalho com *corpora* textuais, é estipular o período em que a unidade foi registrada por escrito pela primeira vez e, ainda assim, trata-se apenas de uma probabilidade.

Ao refletir a respeito dos princípios apresentados por Rey, Lavale-Ortiz (2019) explica que o próprio Rey já reconhecia o caráter relativo do critério temporal, ao afirmar que os neologismos deveriam ser estudados em um aspecto sincrônico, definido de acordo com os objetivos do pesquisador. Ainda assim, para a autora, esse critério não teria sentido do ponto de vista da Linguística Cognitiva, porque, mesmo que os trabalhos com neologismos possam ser delimitados temporalmente, a variação na linguagem gerada

pelos neologismos impõe um tratamento temporal que tem como base o momento presente do falante e os recursos textuais que se tem à disposição, que podem ser mais atuais ou mais antigos, dependendo dos propósitos do estudo.

Essa reflexão leva, também, ao questionamento a respeito do período de tempo que uma unidade lexical pode ser considerada e/ou percebida como neológica. Alves (2000) explica que o sentimento neológico começa a decrescer na medida em que o neologismo passa a fazer parte do discurso do seu criador e da comunidade sem marcas distintivas. Em outras palavras, a unidade lexical deve estar em uso durante determinado período, com certa constância, até que, após ter sido reconhecida como um neologismo, chegue ao que Alves denomina de *desneologização*. Alguns teóricos postulam, tanto na língua geral como nas línguas de especialidade, o período de dez anos como o tempo de permanência do sentimento neológico. Essa duração, entretanto, é considerada demasiadamente longa para a época contemporânea, em que a evolução digital e o desenvolvimento científico e tecnológico são extremamente rápidos. Por isso, a tendência dos pesquisadores atuais é restringir o tempo do sentimento neológico por um período de menos de cinco anos (PRUVOST; SABLAYROLLES, 2003).

b. Limitações do critério psicolinguístico

O critério psicolinguístico, ou psicológico, tem como pressuposto a competência intuitiva do falante ao deparar-se com uma palavra e considerá-la nova. Para Lavale-Ortiz (2019), do ponto de vista cognitivo, este critério constitui o único “método natural” para o estabelecimento de unidades neológicas, em comparação com a objetividade dos critérios temporal e lexicográfico:

O reconhecimento de um neologismo depende da visão subjetiva do falante, de sua perspectiva pessoal, da forma que constrói o mundo e da cultura em que se insere. [...] Embora possamos fazer uso de critérios objetivos em sua caracterização, estes não devem ser a base teórica para defini-lo e identificá-lo. (LVALE-ORTIZ, 2019, p. 215, tradução nossa³).

Nessa perspectiva, a autora atribui maior importância ao critério psicolinguístico em detrimento dos outros, porque o relaciona ao processo espontâneo e regular que explica a maneira como os falantes usam a linguagem enquanto capacidade cognitiva. Sablayrolles (2010), por sua vez, entende o “sentimento neológico” como critério problemático, justamente porque se relaciona à subjetividade do julgamento dos falantes:

3 No original: “El reconocimiento de un neologismo depende de la visión subjetiva del hablante, de su perspectiva personal, de la forma en la que construye el mundo y de la cultura en la que se inserta. [...] Aunque podamos hacer uso de criterios objetivos en su caracterización, estos no deben ser la base teórica para definirlo e identificarlo”.

A percepção de novidade varia de acordo com os indivíduos e os objetos analisados. O mesmo objeto, ainda percebido como novo ou recente por um, já é considerado velho e obsoleto por outro. Não é diferente com as inovações lexicais que são os neologismos: os julgamentos de falantes nativos podem divergir sobre eles [...]. Vários fatores responsáveis pelas variações são identificados, alguns em relação aos indivíduos (idade, cultura, experiência no domínio...), outros aos objetos (alguns tipos de inovação são mais perceptíveis que outros...), outros, por fim, aos objetivos que se pretende com a observação da neologia. (SABLAYROLLES, 2010, p. 12, tradução nossa⁴).

Em outra perspectiva, Cabré (2016) mostra que a intuição pode levar a uma contradição com o critério lexicográfico, uma vez que algumas unidades lexicais podem não ser percebidas como novas pela intuição do falante e, ao mesmo tempo, não estar registradas nos dicionários de língua geral. O que está por trás desse impasse, segundo a autora, é o fato incontestável de que o sentimento neológico depende de circunstâncias geográficas, sociais e temporais a que estão submetidos os falantes, ainda que todos eles sejam, por definição, competentes em sua própria língua.

c. Limitações do critério lexicográfico

O critério mais usado nas pesquisas em neologia é o lexicográfico, que compreende o uso de um *corpus de exclusão*: contrasta-se o candidato a neologismo com palavras já presentes em obras dicionarísticas. Em outras palavras, a ausência da unidade lexical em obras lexicográficas, geralmente dicionários de língua geral, atribui-lhe a qualidade de neologismo. Trata-se de um critério de seleção objetiva que parte do pressuposto que os dicionários servem de parâmetro para a determinação do caráter neológico ou não-neológico da unidade lexical. Como vimos, esse critério tem sido usado desde a criação, no início dos anos 60, do primeiro observatório de neologia e pelos outros observatórios que posteriormente surgiram.

Apesar de ser amplamente utilizado para a caracterização dos termos neológicos, esse critério apresenta limitações principalmente com relação à detecção de neologismos semânticos, uma vez que estes surgem quando um novo sentido é atribuído a uma forma já estabelecida na língua, ou seja, já registrada nas obras lexicográficas. Esse processo torna

4 No original: "La perception de la nouveauté est variable selon les individus et les objets examinés. Le même objet sera encore perçu neuf ou récent par l'un et déjà ancien et vieilli pour un autre. Il n'en va pas autrement des innovations lexicales que sont les néologismes : les jugements des locuteurs natifs peuvent diverger à leur sujet [...]. Plusieurs facteurs responsables des variations sont identifiables qui tiennent les uns aux personnes (âge, culture, expérience dans le domaine...), d'autres aux objets (certains types d'innovation sont plus perceptibles que d'autres...), d'autres enfin aux objectifs poursuivis dans la veille néologique".

problemática principalmente a extração automática de neologismos, já que a máquina reconhece, no contraste com o *corpus* de exclusão, apenas a morfossintaxe de palavras e expressões. A dificuldade de detecção ocorre principalmente em unidades simples, como “gatilho”, mas também em combinatórias, como “velho normal”, contextualizadas, respectivamente, nos exemplos 1 e 2:

1. “Fui postar um áudio do meu pai super fofo me dando bom dia e a menina mandou eu apagar porque era <gatilho>”, publicou a internauta que, sem saber, dava início a uma longa corrente de conteúdos e perfis relacionados ao acontecido⁵.
2. Quando o <velho normal> pode ser, de novo, o novo normal: Como a pandemia mudou ou não sua maneira de encarar seus investimentos?⁶

Além disso, a validação de uma unidade neológica pressupondo sua não dicionarização irá depender da quantidade de dicionários selecionados como *corpus* de exclusão e das características de cada um deles quanto à apresentação da macro e microestrutura. Esse tipo de pressuposto atribui, aliás, certa autoridade à obra dicionarística, tornando-a confiável enquanto representante das palavras e sentidos existentes em uma língua. Entretanto, a ausência de unidades lexicais da nomenclatura de um dicionário pode se dar por conta de princípios preestabelecidos na elaboração da obra: o registro linguístico selecionado, a frequência estipulada para que as palavras fossem coletadas no *corpus*, além de questões ideológicas, editoriais etc. A esse respeito, ensina Cabré (2016, p. 130, tradução nossa⁷):

Uma breve reflexão sobre o que significa “estar em um dicionário” faz-nos perceber a relativa arbitrariedade dos dicionários: cada um deles coleta as unidades que correspondem às suas fontes de referência ou a determinadas condições de recepção. Não estar em um dicionário pode significar que uma unidade é nova, mas também que, apesar de não o ser, não se considerou pertinente ou adequado coletá-la..

5 Disponível em: <https://www.anf.org.br/o-gatilho-que-ganhou-forte-significado-nas-redes-sociais/>

6 Disponível em: <https://valorinveste.globo.com/blogs/fernando-torres/post/2020/06/quando-o-velho-normal-pode-ser-de-novo-o-novo-normal.ghtml>

7 No original: “Una breu reflexió sobre què significa ‘haver estat recollit en un diccionari’ ens fa adonar de la relativa arbitrietat dels diccionaris: cada diccionari recull les unitats que corresponen a les seves fonts de referència o a unes determinades condicions de recepció. No ser en un diccionari pot voler dir que és una unitat nova, però també que, tot i no ser-ho, no s’ha considerat pertinent o adequat recollir-la”.

Outra questão a ser considerada é que um tempo relativamente longo pode passar desde a criação de um item lexical e seu registro em um dicionário. Se, durante esse tempo, o item estiver em uso frequente por grande parte dos falantes, ele não será mais percebido como neológico, ainda que não esteja registrado em obras lexicográficas.

Algumas tendências metodológicas

Como já enfatizamos, as propostas metodológicas que continuamente surgem acompanham os avanços dos estudos da linguagem e não têm caráter excludente, mas complementar. Assim, apesar das limitações do filtro lexicográfico, o contraste dos candidatos a neologismos com um *corpus* de exclusão continua sendo fundamental na pesquisa neológica (cf. ALVES, 2012). Por isso, no intuito de tornar esse critério mais funcional, os estudos atuais propõem que o *corpus* seja ampliado, de forma a incluir não apenas fontes lexicográficas, mas também fontes textuais de vários tipos, a depender do subconjunto linguístico a que pertencem os candidatos a neologismos: livros, artigos, textos jornalísticos etc. Essa ampliação tornou-se possível, principalmente, após o desenvolvimento da Linguística de *Corpus* e a grande evolução dos *corpora* eletrônicos, que passaram a ser compilados e armazenados em bancos de dados, constando de milhares a milhões de palavras e constituindo um filtro textual mais amplo. Além disso, o uso da internet como filtro, ou seja, o recurso à frequência de *sites* de busca também tem se somado aos filtros lexicográfico e textual. A ferramenta *on-line* Google Trends, sobre a qual falaremos no próximo item, é mais um interessante recurso para se quantificar a frequência de uso, por meio da quantificação do interesse dos usuários pela busca dos termos ao longo do tempo.

Ainda como complemento aos três principais parâmetros já descritos, alguns autores sugerem outras pistas que podem levar à comprovação do caráter neológico de um item lexical. Auger (2010) propõe que se observe a forma de apresentação da palavra no texto: uma unidade pode ser um neologismo se apresentar marcas distintas (itálico, negrito etc.), se for seguida por uma reformulação frástica, geralmente definicional, o que indica que foi percebida como nova pelo autor do texto, ou se for acompanhada por um sinônimo, demonstrando a tentativa de esclarecimento do sentido da nova unidade.

Outros filtros são apontados por Estopà (2015), para quem a percepção do caráter neológico é gradual, uma vez que não é a mesma para diferentes falantes; ou seja, trata-se de uma percepção altamente individual e que está relacionada a fatores como conhecimento de mundo, idade, experiência ou acervo lexical do falante. Ainda assim, a autora acredita que, na pesquisa em neologia, é possível encontrar parâmetros que possam demonstrar essa gradação. Dentre eles, estão a frequência de uso, o âmbito de uso (se é empregado de forma particular ou entre determinado grupo) e o canal de uso (meio em que a unidade lexical é empregada: redes sociais, meios de comunicação etc.).

A esse respeito, lembra Cabré (2016) que o próprio qualificador “novo” é gradual: é possível que uma palavra seja mais nova ou menos nova, e questões como tempo e condição enunciativa tornam o conceito de “novo” ainda mais relativo. A definição de neologismo a que chega a autora demonstra a concepção de uma unidade marcada pelo tempo e que varia em relação ao grupo de falantes, ao lugar em que eles se encontram e ao contexto-situacional: “um neologismo é uma unidade lexical nova em relação a um dado momento, aos conhecimentos dos interlocutores e ao seu uso em uma dada situação e a partir de uma dada perspectiva enunciativa” (CABRÉ, 2016, p. 130, tradução nossa⁸).

Nesse sentido, como bem esclarece Alves (2000), a criação neológica não pode ser dissociada dos componentes frásico e textual, porque a grande maioria dos neologismos é criada no âmbito das sentenças e dos textos em que estas se inserem, ou seja, em um contexto pragmático. O fenômeno da neologia, bem como o seu produto, o neologismo, vai além de expressões ou estruturas. Assim, se o material que os pesquisadores têm em mãos é textual, é de fundamental importância considerar o contexto em que os textos do *corpus* são lidos e interpretados, de forma que os candidatos a neologismos estejam em conformidade com a situação comunicacional em que ocorrem.

Os termos da cultura digital e a detecção de neologismos

Os neologismos que irão ilustrar essa discussão são provenientes de gêneros textuais representativos da cultura digital, os quais datam de 2016 a 2020, dentre eles, textos midiáticos, *posts* de redes sociais e memes digitais. Após a triagem dos textos que fariam parte do *corpus*, estes foram inseridos na ferramenta computacional *AntConc*⁹, que gerou uma lista de palavras-chave em contraste com um *corpus* de referência. Essa lista compôs os pré-candidatos a neologismos que, em seguida, passaram por uma seleção e validação. Nos *posts* de redes sociais e nos memes, a produtividade da inovação lexical é particularmente notável, porque são gêneros textuais abertos à manifestação do grande público digital. Qualquer pessoa munida de um dispositivo com conexão à internet pode ter voz na “cultura participatória”. A difusão dos neologismos fica por conta de outro fenômeno fundamental da cultura digital contemporânea: o compartilhamento. Por meio dele, a audiência digital expressa seus interesses culturais, políticos e ideológicos.

Demonstraremos, sucintamente, como se comportou cada critério metodológico no que concerne à detecção e validação de treze neologismos do *corpus* de nossa pesquisa. Dentre os critérios, incluímos a ferramenta Google Trends¹⁰ (doravante Trends), que mostra o

8 No original: “Un neologisme és una unitat lèxica nova en relació amb un moment donat, amb els coneixements dels interlocutors i amb l’ús en una situació i des d’una perspectiva enunciativa determinades”.

9 Disponível em: <https://www.laurenceanthony.net/software/antconc/>

10 Disponível em: <https://trends.google.com.br/trends/?geo=BR>

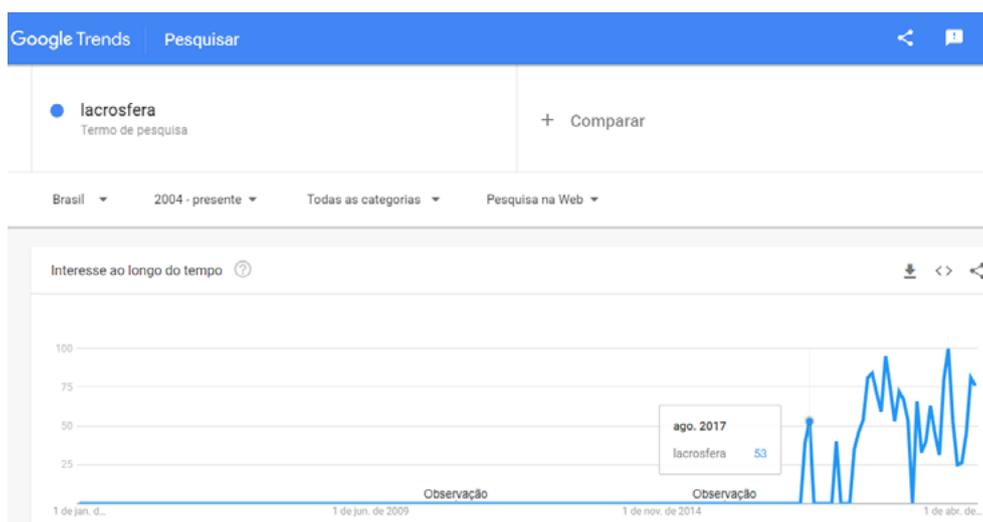
aumento ou declínio da popularidade da busca por termos em períodos específicos. Para tanto, a ferramenta apresenta gráficos com a frequência em que determinado termo ou expressão é procurada, bem como a frequência com que determinado assunto está em alta dentre as publicações da rede. Além disso, ela possibilita a seleção de filtros por região (países) ou sub-região (estados) e por períodos de tempo, com início possível a partir de 2004. Para nossa pesquisa, o filtro da região foi determinado como “Brasil” e o período foi determinado como “2004 – presente”, para que houvesse melhor visualização dos picos do gráfico nos anos atuais caso a unidade pesquisada fosse neológica. Para exemplificar os padrões do Trends, serão utilizados os neologismos *lacrosfera* e *cancelamento*.

Lacrosfera é um neologismo formado pelo processo de cruzamento vocabular, também conhecido como palavra-valise. Trata-se de um tipo de redução em que duas unidades lexicais são privadas de seus elementos para formarem um novo: um perdendo sua parte final e o outro, a parte inicial (ALVES, 2000). No termo em questão, os dois elementos são os substantivos “lacre” – que, no contexto, tem um sentido também neológico, proveniente do novo uso do verbo “lacrar”: encerrar a discussão, arrasar – e “esfera”, com o sentido de “mundo esférico, fechado”. O termo é usado geralmente em contexto político, para ironizar pontos de vista progressistas. Um dos contextos de ocorrência de *lacrosfera* é representado pelo exemplo 3:

3. A pesquisa CNI/Ibope, divulgada hoje, confirma, no entanto, que a <lacrosfera> continua falhando miseravelmente em sua avaliação de Jair Bolsonaro.

A busca dessa unidade lexical na ferramenta Trends, com as configurações supracitadas de região e de período, retorna o seguinte gráfico:

Figura 1. Busca do termo “lacrosfera” na ferramenta Google Trends



Fonte: <https://trends.google.com.br/trends/explore?date=all&geo=BR&q=lacrosfera>

Como se trata de um neologismo formal, ou seja, a combinação das palavras que o constituem resulta em uma forma inédita, é possível observar uma linha reta antes de agosto de 2017, indicando que a quantidade de buscas pelo termo equivale a zero (e, conseqüentemente, a alta probabilidade de a palavra não existir nesses anos anteriores). Em 2017, ocorre o primeiro aumento repentino da busca pelo termo, indicado pelo primeiro pico no gráfico. O percurso do zero ao pico é um grande indicativo de detecção de uma unidade lexical neológica. O último e maior pico do gráfico ocorreu em março de 2020, e o gráfico termina em alta, o que indica que o item lexical continua em uso.

A unidade lexical *cancelamento*, já antiga na língua, tornou-se semanticamente neológica ao receber o sentido de “desprezo por alguém que cometeu um erro – de acordo com o julgamento dos ‘canceladores’”. O emprego do neologismo ocorre principalmente em memes e viralizou em *posts* de redes sociais. A figura 2, abaixo, contextualiza o uso no Twitter, e os exemplos 4 e 5, na sequência, apresentam o uso em textos midiáticos:

Figura 2. Neologismo semântico “cancelamento”



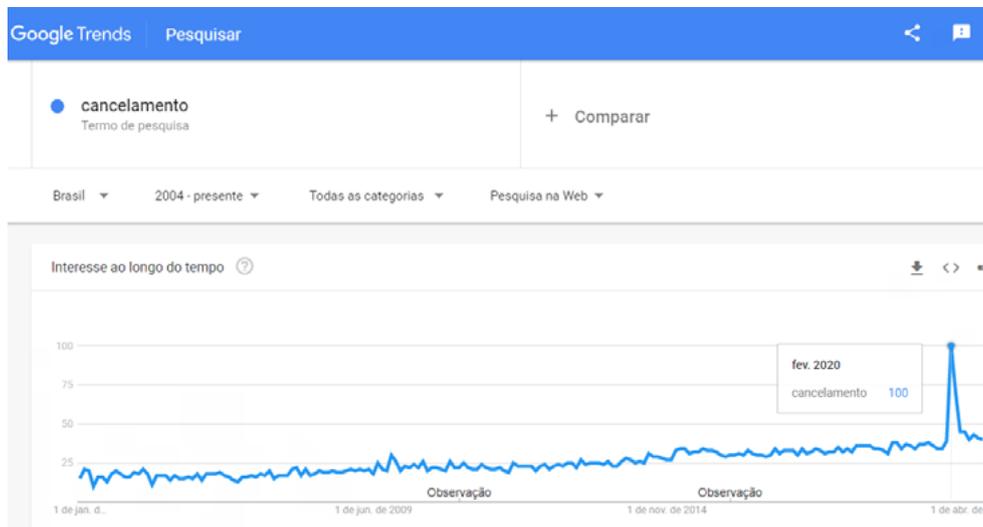
Fonte: <https://twitter.com/anacronices/status/1201110568540803072>

4. Pense em uma celebridade que disse algo controverso ou se envolveu em alguma polêmica – ela provavelmente foi sentenciada ao <“cancelamento”> pelo tribunal das redes sociais.
5. O <“cancelamento”> é um ataque à reputação que ameaça o emprego e os meios de subsistência atuais e futuros do cancelado¹¹.

Na busca por esse neologismo semântico na ferramenta Trends, é possível observar a frequência de uso do termo nos anos anteriores ao pico do gráfico, diferentemente do neologismo anterior:

¹¹ Disponível em: <https://www.bbc.com/portuguese/geral-53537542>

Figura 3. Busca do neologismo “cancelamento” na ferramenta Google Trends



Fonte: <https://trends.google.com.br/trends/explore?date=all&geo=BR&q=cancelamento>

Como se trata de um neologismo cujo processo de formação foi semântico, e não formal, a linha do gráfico apresenta movimentações desde o período inicial. Em fevereiro de 2020, quando a unidade lexical começou a ser usada com o novo significado nas redes sociais e nas mídias, houve um aumento repentino significativo no gráfico. Em outras palavras: a ferramenta poderá indicar que estamos diante de um neologismo semântico quando exibe uma alta de uso, pelo público digital, de uma unidade lexical já antiga.

Além da frequência de uso, os aumentos repentinos nos gráficos do Trends são interessantes para se observar a imprevisibilidade das unidades lexicais, um dos critérios adicionais propostos por Auger (2010), ligado à “observação de frequência irregular”. Com esses recursos, a ferramenta oferece um ótimo auxílio na validação das unidades neológicas de nossa pesquisa, somando-se aos critérios tradicionais e às marcas textuais sugeridas por Auger.

No quadro a seguir, listamos, em ordem alfabética na primeira coluna, uma pequena amostra dos neologismos coletados no *corpus* de análise da pesquisa. Para cada um deles, estão assinalados os critérios que se mostraram funcionais em sua coleta e validação. Constam, na primeira linha horizontal, os critérios tradicionais propostos por Rey (lexicográfico e psicolinguístico). O critério temporal não é apresentado, porque, como vimos, ele pressupõe que uma unidade é neológica se surgiu em um período recente; como todas surgiram recentemente – e esse é o próprio conceito de neologismo –, todas as células da coluna correspondente ao parâmetro temporal estariam assinaladas. Por isso, consideramos que sua apresentação não seria necessária. Em seguida, sob o rótulo de “marcas textuais”, estão incluídas todas as pistas tipográficas e semânticas sugeridas por Auger (2010): aspas, itálico, reformulação frástica, e/ou definicional, apresentação de

um sinônimo etc. Por fim, a coluna “frequência de uso” mostra a busca do neologismo no Trends e é assinalada se o gráfico respondeu com picos nos últimos cinco anos:

Quadro 1. Produtividade dos critérios de detecção de neologismos

Critérios Neologismos	Lexicográfico	Psicolinguístico	Marcas textuais	Frequência de uso (Google Trends)
anti-vacina	x			x
cancelamento		x	x	x
corongavírus	x	x		x
biscoitar		x		x
finalmente os refrescos				x
figital	x	x	x	
gatilho		x		x
infoxicação	x	x		x
jantar		x		x
lacrosfera	x	x		x
lockdown	x	x	x	x
novo normal		x	x	x
quarentener	x	x		x

Fonte: Elaboração própria

O quadro indica a importância de se utilizarem critérios variados na identificação de neologismos: na amostra apresentada, apenas um dos itens lexicais cumpriu os quatro critérios preestabelecidos (e também o critério temporal, omissos no quadro). Trata-se de *lockdown*, que cumpre: o critério lexicográfico, por não estar registrado no *corpus* de exclusão; o critério psicolinguístico, por ser reconhecido como novo pela intuição linguística; o critério das marcas textuais, porque é grafado, geralmente, em itálico, entre aspas ou acompanhado de uma pequena explicação; e o critério do Trends, cujo gráfico apresenta frequência altíssima em maio de 2020.

Os neologismos semânticos *cancelamento*, *biscoitar*, *gatilho* e *jantar* escapariam ao critério de exclusão lexicográfica e, em uma extração automática, não seriam selecionados. Além disso, o sentimento neológico do público digital jovem os faz reconhecer como neológicos, mas, provavelmente, outros grupos de falantes não os reconheceriam. Nem todos eles trazem marcas textuais, mas todos apresentam picos no gráfico gerado pelo Trends.

Neste ponto, é importante enfatizar o que já mencionamos anteriormente: a ferramenta indica não o uso do termo em questão, mas o aumento ou declínio da popularidade da busca pelo termo em períodos específicos. Nesse sentido, devem-se reconhecer as limitações desse processo, já que, nem sempre, o aumento significativo das buscas por um termo ou a frequência com que determinado assunto está em alta faz destes um neologismo semântico. Por isso, novamente, há que se considerar os outros parâmetros para validação. Além disso, não se trata, no caso do gráfico, de uma análise diacrônica dos neologismos, mas de uma ênfase na movimentação formal e semântica por que passam essas unidades e a possibilidade de contraste sincrônico e dinâmico com os outros elementos da língua, considerando-se a criação lexical por parte dos falantes e o uso situado das unidades neológicas.

Considerações finais

Já há algum tempo, a prática tradicional de coleta de novas palavras apenas pela leitura de textos foi substituída pela coleta automatizada. Após a automatização, houve propostas de alteração dos critérios de reconhecimento de neologismos e a união de procedimentos como os *corpora* de exclusão e as mais avançadas ferramentas tecnológicas para extração de neologismos. A evolução dos modelos teóricos da pesquisa linguística e neológica espelha não só a dinamicidade das unidades lexicais de uma língua, mas também o lugar cativo da Inteligência Artificial:

Impulsionada pelo progresso tecnológico, a oferta de novos meios de comunicação digital renova-se, de fato, a um ritmo desenfreado, contribuindo, assim, para a proliferação da atividade neológica. Paralelamente, a crescente interação entre ciências da linguagem e ciência da computação oferece aos linguistas meios cada vez mais sofisticados de analisar os fenômenos da linguagem. (ALTMANOVA; ZOLLO, 2019, p. 19, tradução nossa¹²).

12 No original: "Portée par le progrès technologique, l'offre des nouveaux supports de communication numérique se renouvelle en effet à un rythme effréné, contribuant ainsi au foisonnement de l'activité néologique. Parallèlement, l'interaction croissante entre les sciences du langage et l'informatique offre aux linguistes des moyens de plus en plus sophistiqués pour analyser les phénomènes langagiers".

Os critérios tradicionais para o estabelecimento da neologia continuam sendo importantes, funcionais e frequentemente usados nas principais pesquisas da área. No entanto, se o conjunto de critérios estipulados para o trabalho com determinado *corpus* for mais amplo, será possível a detecção de neologismos que escapariam da exclusão lexicográfica ou da intuição neológica do pesquisador.

Ainda que determinados critérios metodológicos não sejam aplicáveis a todo tipo de unidade neológica, a adaptação deles ao subconjunto linguístico analisado pode ser realizada de forma criteriosa e, conseqüentemente, potencializar os resultados, principalmente no que concerne à detecção e à validação dos neologismos. A quantificação dos resultados com o auxílio de ferramentas que apresentam a frequência de uso em períodos atuais pode ser feita para fins de observação da língua em movimento em decorrência dos processos de renovação lexical.

REFERÊNCIAS

ADELSTEIN, A. Metodología de trabajo neológico y tipologías: aspectos de neología semántica. In: ALVES, I. M.; PEREIRA, E. S. (org.). *Neologia das línguas românicas*. São Paulo: Humanitas, 2015.

ALVES, I. M. Uma metodologia para a descrição de neologismos. In: GONÇALVES, A. V.; GÓIS, M. L. S. (org.). *O fazer científico*. Campinas: Mercado das Letras, 2012.

ALTMANOVA, J.; ZOLLO, S. D. Présentation. *Neologica*, Paris, n. 13, p. 19-22, 2019.

AUGER, P. Néologicité et extraction néologique automatisée. In: CABRÉ, M. T. *et al.* (ed.). *Actes del I Congrés Internacional de Neologia de les Llengües Romàniques*. Barcelona: IULA; Documenta Universitaria, 2010.

CABRÉ, M. T. Per què és relativament fàcil de detectar neologismes i tan complicat de definir què són: breu apunt epistemològic. In: OBSERVATORI DE NEOLOGIA (ed.). *Mots d'avui, mots de demà*. Barcelona: Institut de Lingüística Aplicada, Universitat Pompeu Fabra, 2016. p. 127-132.

CABRÉ, M. T. La neologia: un nou camp a la cerca de la seva consolidació científica. *Caplletra*, n. 59, p. 125-136, 2015.

CABRÉ, M. T. *La terminología: teoría, metodología, aplicaciones*. Barcelona: Editorial Antártida; Empúries, 1993.

ESTOPÀ, R. Sobre neologismos y neologicidad: reflexiones teoricas con repercusiones metodologicas. *In: ALVES, I. M.; PEREIRA, E. S. (org.). Neologia das línguas românicas*. São Paulo: Humanitas, 2015.

FALK, I.; DELPHINE, B.; CHRISTOPHE, G. The Logoscope: a Semi-Automatic Tool for Detecting and Documenting French New Words. *ArXiv*, v. 1, out. 2018.

GUILBERT, L. Théorie du néologisme. *Cahiers de l'Association Internationale des Etudes Françaises*, v. 25, p. 9-29, 1972.

JESUS, A. M. R. Tipologias dos neologismos: breve percurso histórico. *GTLex*, Uberlândia, v. 4, n. 1, p. 54-67, 2020.

LAVALE-ORTIZ, R. M. Bases para la fundamentación teórica de la neología y el neologismo: la memoria, la atención y la categorización. *CLAC*, n. 80, p. 201-226, 2019.

PRUVOST, J.; SABLAYROLLES, J. F. *Les néologismes*. Paris: Presses Universitaires de France, 2003.

REY, A. Néologisme: un pseudo-concept? *Cahiers de Lexicologie, Revue Internationale de Lexicologie et de Lexicographie*, v. 28, p. 3-17, 1976.

SABLAYROLLES, J. F. Neologismes, classes d'objets et extraction automatique. *In: CABRÉ, M. T. et al. (ed.). Actes del I Congrés Internacional de Neologia de les Llengües Romàniques*. Barcelona: IULA; Documenta Universitaria, 2010.