

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

PEDRO PERINI SURREAUX

A ATRIBUIÇÃO DE GÊNERO GRAMATICAL A ITENS NOVOS EM PORTUGUÊS
BRASILEIRO

PORTO ALEGRE

2020

PEDRO PERINI SURREAUX

A atribuição de gênero gramatical a itens novos em português brasileiro

Trabalho de Conclusão de Curso
apresentado como requisito parcial para
obtenção do título de Bacharel em Letras
pela Universidade Federal do Rio Grande
do Sul.

Orientador: Prof. Dr. Luiz Carlos Schwindt

PORTO ALEGRE

2020

AGRADECIMENTOS

Agradeço à Universidade Federal do Rio Grande do Sul pela formação gratuita e de qualidade que me proporcionou.

À minha família, meu pai, Luiz, minha mãe, Márcia, meu tio, Pedro, e meus irmãos, que me apoiam e me permitem seguir trilhando este caminho.

Aos meus amigos de longa data, Afonso, Leonardo, Ana Laura, Alessandra, Eduardo, Venancio, que estiveram comigo em todos os momentos e que serviram de exemplo e inspiração.

A Júlia, Rossana, Camila, Isabela, Beatriz e Pedro, meus amigos e colegas de pesquisa, com quem tanto aprendi, e que estiveram ao meu lado, dialogando e colaborando com suas experiências.

Aos professores e professoras do Instituto de Letras, Leandro, Valéria, Gabriel, Sérgio, Ingrid, que me inspiram e que para mim servem como exemplo de pesquisadores e profissionais da área da linguística.

Em especial, a meu orientador, Luiz, que acreditou em mim e me abriu tantas portas e oportunidades, também servindo como enorme exemplo e inspiração para que eu trilhasse este caminho.

Ao Grupo Círculo Linguístico da UFRGS, por possibilitar tanto o diálogo com colegas como o contato com diversos textos interessantes da área dos estudos linguísticos.

RESUMO

O presente Trabalho de Conclusão de Curso aborda o fenômeno do gênero gramatical e os processos com que se relaciona, visando a contribuir descritivamente para discussões sobre o tema. Toma-se como foco o sistema de gênero do português brasileiro e o processo de atribuição de gênero a substantivos novos na língua. Em português, todo substantivo em uso conta com um valor de gênero gramatical nuclear, feminino ou masculino, uma vez que determina a flexão de determinantes, como artigos e adjetivos, na concordância nominal. Embora só seja possível assegurar o gênero de um substantivo pela sintaxe, observam-se correlações entre gênero, traços semânticos e segmentos fonológicos no repertório da língua, associando-se, por exemplo, referentes do sexo feminino e a terminação átona /a/ ao feminino gramatical. Buscando-se o estudo do fenômeno da atribuição de gênero a novos itens na língua, objetiva-se realizar: revisão bibliográfica sobre o sistema de gênero gramatical do português e sobre o fenômeno da adaptação morfológica a sistemas de gênero; estudo do fenômeno a partir de empréstimos nominais do inglês e do árabe na língua; e experimentos com falantes de português brasileiro utilizando-se pseudosubstantivos para teste de critérios semânticos e formais possivelmente atuantes no processo. Entende-se ter papel no fenômeno o conhecimento internalizado dos falantes quanto a regras e padrões do sistema receptor. Assume-se também a atribuição de gênero por associação semântica, em que conceitos novos recebem o gênero de um substantivo análogo do português. Segundo os resultados, nos empréstimos do inglês, prevaleceram casos de analogia semântica, sendo terminações relevantes para o fenômeno restringidas pela fonotática do inglês; dentre os empréstimos do árabe, parte significativa dos femininos terminados em /a/ derivam palavras fechadas por vogais átonas análogas no árabe; pseudopalavras com terminações predominantemente femininas no léxico tiveram feminino atribuído na maioria dos casos; pseudopalavras que sugeriam certos tipos de analogia semântica com itens femininos foram em geral entendidas como femininas, enquanto que a atribuição do feminino a itens que supunham referentes do sexo feminino foi quase categórica. Os resultados sugerem o envolvimento de diferentes naturezas no fenômeno, dialogando com achados da literatura sobre outras línguas. Ainda, nos contextos observados, assume-se haver relativa sistematicidade na aplicação dos processos, para além dos casos condicionados pela semântica do referente, por vezes assumido como único critério sistemático.

Palavras-chave: Gênero gramatical. Atribuição de gênero. Empréstimos.

ABSTRACT

The present undergraduate thesis approaches the phenomenon of grammatical gender and the processes to whom it is related, aiming at contributing descriptively to discussions on the subject. In particular, this study focuses on the Brazilian Portuguese gender system and the process of morphological adaptation of novel nouns entering the language. In Portuguese, nouns in use contain a nuclear gender value, either feminine or masculine, as to determine the inflection of adjectives and articles, for instance, in nominal agreement. Although the gender of a given Portuguese noun may only be attested through syntax, correlations between gender, semantic features, and phonological units may be observed in the lexicon. Female referents and final unstressed vowel /a/, for example, are associated with the grammatical feminine. Seeking to study the phenomenon of gender assignment in novel nouns in Portuguese, this work aims at carrying out: a bibliographic review on the Portuguese gender system and on the phenomena of morphological adaptation of nouns entering grammatical gender systems; and the study of the phenomenon based on the cases of loanwords from English and Arabic in Portuguese, and on experiments with Brazilian Portuguese speakers, using pseudonouns to test the semantic and formal criteria potentially active in the process. It is assumed that speakers' internalized knowledge regarding linguistic patterns and rules of the receptor system plays a role in the process. Moreover, it is understood that gender assignment may occur by means of semantic association, where new concepts are assigned the gender values of associated Portuguese nouns. According to the results, semantic analogy prevailed as a criterion for gender assignment among English loanwords, given the phonotactical restraining of conditioning segments; most feminine loanwords from Arabic ending in /a/ derive from words in that language that ended in analogous unstressed vowels; pseudonouns with segments widely associated to feminine in the lexicon were assigned feminine gender in most cases; pseudonouns that suggested certain types of semantic analogy with feminine Portuguese nouns were generally understood as feminine, while the assignment of feminine to pseudonouns that presumed female referent stood out, being almost categorical. Results point to the involvement of different factors in the phenomenon, considering each context, in agreement with findings of the literature. Moreover, the process is assumed to apply with relative systematicity in cases other than those involving the semantics of the referent, sometimes seen as the only systematic criteria.

Keywords: Grammatical gender. Gender assignment. Loanwords.

SUMÁRIO

1 INTRODUÇÃO	5
2 O GÊNERO GRAMATICAL EM PORTUGUÊS	8
2.1 REVISÃO TEÓRICA	8
2.2 CARACTERIZAÇÃO DO SISTEMA DE GÊNERO EM PORTUGUÊS	12
3 O FENÔMENO DA ATRIBUIÇÃO DE GÊNERO A SUBSTANTIVOS NOVOS	14
3.1 REVISÃO TEÓRICA	14
3.2 ATRIBUIÇÃO DE GÊNERO A SUBSTANTIVOS NOVOS EM PORTUGUÊS	24
4 EMPRÉSTIMOS EM PORTUGUÊS	27
4.1 EMPRÉSTIMOS DO INGLÊS EM PORTUGUÊS BRASILEIRO	28
4.1.1 Metodologia	29
4.1.2 Resultados	30
4.1.3 Discussão	322
4.2 EMPRÉSTIMOS DO ÁRABE NO PORTUGUÊS	33
4.2.1 Metodologia	34
4.2.2 Resultados	34
4.2.3 Discussão	35
5 EXPERIMENTOS COM PSEUDOPALAVRAS	37
5.1 TESTE PARA CRITÉRIOS FORMAIS DE ATRIBUIÇÃO DE GÊNERO	37
5.1.1 Metodologia	37
5.1.2 Resultados	39
5.1.3 Discussão	4041
5.2 TESTE PARA CRITÉRIOS SEMÂNTICOS DE ATRIBUIÇÃO DE GÊNERO	422
5.2.1 Metodologia	42
5.2.2 Resultados	455
5.2.3 Discussão	48
6 CONSIDERAÇÕES FINAIS	52
REFERÊNCIAS	53

1 INTRODUÇÃO

O gênero gramatical integra a morfologia nominal do português, expressando-se em todos os nomes da língua. Dentre eles, esse traço manifesta-se de formas diferentes ao se contrastarem substantivos e demais itens nominais. Mesmo estando centrado no domínio da morfologia, o gênero em português só se torna observável através da sintaxe (ROCHA, 1998), no processo de concordância nominal. Nessa operação, determinantes – artigos, adjetivos e alguns pronomes – são flexionados para gênero e número de acordo com valores contidos no substantivo, que atua como núcleo do sintagma nominal. Assim, considerando-se seu uso, um dado substantivo deverá apresentar, em relação ao gênero, uma de suas duas expressões possíveis na língua, masculino ou feminino. Entende-se, então, a natureza morfossintática do gênero gramatical.

Línguas que possuem gênero gramatical, de modo geral, correlacionam expressões desse sistema a traços que integram outros níveis da gramática. De acordo com a classificação de Corbett (1991) para essas línguas, que considera a natureza e a expressividade dessas associações, o gênero do português poderia ser considerado um sistema híbrido, correlacionando-se em certo grau tanto à fonologia – em que se relaciona o segmento final *-a* e o gênero feminino –, quanto à semântica – dada a relação do feminino semântico ao feminino gramatical.

Neste trabalho, investigamos o sistema de gênero do português tendo em vista essas interações. Tomamos como foco o papel delas no fenômeno de atribuição de gênero na adaptação morfológica de itens novos que são implementados na língua, descrito no capítulo 3. Especificamente, buscamos melhor entender o que condiciona a escolha por um ou outro gênero quando substantivos estranhos são processados pelos falantes e sujeitos às estruturas da gramática do português. Por que, por exemplo, atribui-se masculino a *post*, mas feminino a *live*? Com o objetivo de clarificar questões como esta, analisamos alguns casos em que o fenômeno pode ser observado na língua: os empréstimos nominais de diferentes origens presentes no léxico do português, tema do capítulo 4, e os resultados de testes com pseudosubstantivos realizados com falantes de PB para fins do presente estudo, tema do capítulo 5.

A partir das análises e discussões sobre o sistema de gênero do português, esperamos que a investigação proposta neste trabalho possa contribuir para um melhor entendimento sobre essa categoria morfológica e os processos que ela integra enquanto fenômenos que dizem respeito à gramática das línguas do mundo.

2 O GÊNERO GRAMATICAL EM PORTUGUÊS

2.1 REVISÃO TEÓRICA

Neste capítulo, revisaremos alguns estudos que consideramos especialmente relevantes para a análise que empreendemos neste trabalho, por colaborarem para a descrição e explicação de sistemas de gênero gramatical e suas interações com outros componentes da gramática. Esses estudos configuram algumas das principais discussões sobre gênero em português, sendo eles Câmara Jr. (1970), Alcântara (2005), Armelin (2014), Schwindt (2011, 2018) e, em espanhol, Harris (1991).

Um dos primeiros autores a abordar o fenômeno do gênero gramatical em português foi Mattoso Câmara Jr., na obra *Estrutura da Língua Portuguesa* (1970). Partindo de uma perspectiva estruturalista de Item-e-Arranjo, o autor entende o sistema de gênero da língua como comportando uma forma marcada, o feminino, que se encontra em oposição ao masculino, a não marcada. O pressuposto baseia-se na observação dos segmentos finais dos itens nominais do português. As contrapartes femininas de pares de nomes sexuais, como *menino/menina* e *escritor/escritora* seriam exemplares de uma instância de flexão nominal, operada pela adição do morfema de gênero -a, semelhante àquela de número, onde o plural é instanciado pelo morfema -s e o singular pela ausência de marca. Segundo o autor, o caráter não marcado dos nomes masculinos evidencia-se em algumas observações dentro do recorte dos nomes sexuais: i) nos casos mais regulares, pares apresentam -a para feminino e -o para masculino; ii) é possível que o nome feminino termine em -a sem que a contraparte masculina traga elementos além da raiz, p. ex., *professor/professora*; iii) há casos em que há a ausência de marcas tanto no masculino como no feminino ex. (o) *estudante/(a) estudante*. A noção de marcação, neste caso, prevê que não existe a possibilidade de apenas a forma masculina de um par receber a marca de gênero: **a professor/o professoro, *a estudante/o estudanto*.

Harris (1991) analisa o sistema nominal do espanhol para explicar a exponenciação dos morfemas de gênero naquela língua. A relevância da análise para o presente estudo deve-se à relativa similaridade entre os sistemas nominais do português e do espanhol. A análise pressupõe a natureza modular das propriedades que subjazem os substantivos, adjetivos, verbos e pronomes. Segundo o autor, os

substantivos carregam diferentes traços linguísticos, que podem, por sua vez, estar relacionados às categorias de gênero (feminino, masculino, biforme ou comum de dois gêneros), às classes formais (identificadas pelos finais fonológicos -o, -a, -e) e às propriedades semânticas (animado, humano, sexuado, etc.).

Uma das características da abordagem de Harris é o aspecto modular atribuído aos sistemas nominais: um nome pode ser feminino ou masculino, terminar em uma variedade de segmentos fonológicos, e relacionar-se a referentes que apresentam sexo biológico ou gênero social masculino ou feminino, embora nenhuma dessas classes categoriais esteja necessariamente correlacionada a gênero.

Harris (1991) classifica os nomes da língua de acordo com a regularidade das correlações entre os diferentes traços a eles atribuídos; assim, aqueles pertencentes ao que rotula como *inner core* seriam os exemplos mais regulares, onde itens masculinos terminam em -o e femininos em -a, seguido por um número de casos com correlações gradativamente menos regulares, sendo os últimos casos os mais raros: aqueles em que nomes femininos terminam em -o e masculinos em -a. Essa classificação, que considera a recorrência das correlações, é análoga em quase todos os aspectos ao que acontece no português, como veremos em trabalhos que, até certo ponto, baseiam-se em Harris para analisar o sistema de gênero desta língua.

A perspectiva de Alcântara (2010) adota uma base modular semelhante à de Harris (1991) para a classificação dos nomes do português em relação às categorias de gênero e classe temática. Segundo a autora, os itens podem pertencer a uma de quatro classes formais: I - final em /o/; II - final em /a/; III - final em /e/; e IV, definida pela ausência de um elemento vocálico terminal, ou tema. Essas classes, assim como em Harris (1991), são potencialmente independentes em relação às duas classes de gênero dos nomes – feminino e masculino –, uma vez que podem ser observados na língua casos em todas as possíveis intersecções das classes formais com as classes de gênero, mesmo que em diferentes níveis, ex. *ponte/pente, casa/mapa, tribo/livro*.

A partir desse modelo classificatório, Alcântara desenvolve uma análise com base na Morfologia Distribuída (HALLE; MARANTZ, 1993, 1994) para explicar a relação entre gênero e classe formal em português. Esse tratamento teórico coloca o português ao lado de outras línguas românicas, como o espanhol e o italiano, no que diz respeito à exigência de um sufixo temático em não verbos para cumprir uma condição de boa-formação da língua. A seleção de um dos três morfemas na

derivação dos nomes é atribuída a informações contidas nas raízes. Uma vez que o gênero masculino é o não marcado, ou *default*, cabe ao feminino uma informação específica, ou traço idiossincrático, a ser representado na raiz. Assim, de modo similar à regra de redundância apresentada por Harris (1991), a informação de *f*, feminino, nas raízes, seleciona a classe formal II na derivação, fazendo com que esta não precise estar explicitada, como é o caso de masculinos terminados em *-a*, p. ex., *cometa*.

Ao fazer isso, Alcântara estipula uma correlação inerente do gênero feminino com a terminação *-a*, prevista por regra, enquanto que a correlação dessa terminação com o masculino não é regular, sendo necessária uma informação que a especifique. Ainda, os casos raros em que femininos terminam em *-o* são explicados pela presença do traço de classe formal I, junto a *f*, na raiz, impedindo que o morfema II seja selecionado pela regra de redundância. Ou seja, é necessário que itens como *tribo* carreguem a informação específica quanto ao segmento terminal *-o*, de modo a impedir a aplicação da regra que relaciona itens femininos à vogal átona final *-a*.

Armelin (2014) distancia-se da perspectiva modular a respeito das classes nominais e vai em direção a um tratamento unificado, entendendo gênero e classe temática como expoentes de um mesmo nó sintático na derivação. Ao contrário de Alcântara (2010), essa perspectiva não considera valores internos à raiz, que permitiriam a projeção ou seleção de complementos. Assim, é o nó sintático de gênero que seleciona a raiz. A derivação, nos casos menos marcados, por *default*, exponenciaria os segmentos terminais *-o* e *-a* no caso de raízes selecionadas por nós de gênero contendo, respectivamente, os traços masculino e feminino. Formas que fogem do padrão, como *mapa* e *cometa*, de um lado, e *tribo* e *libido*, de outro, seriam explicadas por especificações nos nós sintáticos que selecionam as raízes desses itens e atribuem segmentos terminais irregulares na inserção fonológica.

Schwindt (2018) adota uma perspectiva otimalista realizacional para dar conta dos nomes em português. Este modelo de análise distancia-se ainda mais do tratamento dado por Harris e Alcântara, que consideram diferentes módulos para as palavras da língua. Para Schwindt (2011, 2018), os morfemas de gênero e classe temática são unidades independentes – por mais que fonologicamente idênticas em alguns casos –, exponenciando-se nos nomes com base em uma hierarquia realizacional em que diferentes traços competem, avaliados a partir de restrições

universais, no curso de sua superficialização. Embora a análise não esteja limitada a apenas estes casos, seu foco recai sobre os nomes temáticos (fechados por -o, -a e -e). À exceção do segmento final -e, que corresponde exclusivamente a um morfema de classe temática (CL3), cada final fonológico dos nomes temáticos (-a e -o) está potencialmente associado a dois morfemas, um de gênero e outro de classe formal (respectivamente: CL1 e FEM; e CL2 e MASC), sendo, portanto, unidades diferentes do ponto de vista morfológico.

Para melhor entendermos a abordagem realizacional proposta por Schwindt (2011, 2018), observemos a formalização da hierarquia em (1).

(CL3), /e/ >> (FEM), /a/ >> (CL2), /o/ >> (CL1), /a/ >> (MASC), /o/ (1)¹

Nela, CL3 aparece como o morfema mais marcado, o que está de acordo com os dados de produtividade levantados no estudo: os substantivos em português dividem-se de modo equilibrado entre nomes femininos e masculinos no montante de palavras fechadas por -e átono; além disso, nomes sexuais fechados por -e não superficializam FEM: a exemplo de (o) *estudante*/(a) *estudante*. A segunda posição, ocupada por FEM, justifica-se pela marcação de nomes femininos sexuais em -a, o que não acontece, à exceção de casos regulares (*aluno/aluna*), com nomes masculinos ex., *professor/professora*. A maior produtividade de itens “incomuns” masculinos – terminados em -a, como *mapa* e *sistema* – em relação a itens femininos desse tipo – terminados em -o, como *tribo* e *libido* –, sustentada pelos dados levantados em *corpora*, embasa o ordenamento dos morfemas de classe temática CL2 e CL1. Por fim, o morfema MASC, substanciado por /o/, é tratado com o morfema menos produtivo na língua, uma vez que casos de pares sexuais como *menino/menina* dizem respeito a um /o/ temático na forma masculina, e não a um morfema de gênero, sendo agramaticais casos em que MASC seria expresso fonologicamente: a *professora/o* **professoro*.

Dadas essas visões diversas sobre a organização do sistema de gênero em português, prosseguiremos, na próxima seção, para a caracterização do sistema em

¹ Adaptado de Schwindt (2011, p. 271).

linhas gerais, considerando seu caráter formal e semântico e chamando atenção para os aspectos que mais importam às nossas análises aqui apresentadas (seções 4 e 5).

2.2 CARACTERIZAÇÃO DO SISTEMA DE GÊNERO EM PORTUGUÊS

Para fundamentar sua análise apresentada acima, Schwindt (2018) traz uma descrição detalhada sobre a produtividade do sistema de gênero na língua, onde podemos observar a relação entre valores de gênero e demais marcas formais. Utilizaremos esses dados de produtividade para traçar aqui as linhas gerais sobre o sistema do português no que se refere a essas correlações internas à gramática. Conforme levantamento trazido pelo autor dos substantivos dicionarizados (17.049 *types* levantados do Dicionário Aurélio Eletrônico, à exceção de compostos, derivados de outros nomes e elementos de locuções), 51,9% são uniformes femininos, como *bola* e *ação*; 43,1% são uniformes masculinos, como *bolo* e *amor*; 2,6% são biformes, como *menino/menina*; e 2,3% são comuns de dois gêneros, como *dentista* e *atleta*. Já para os dados de uso falado (de transcrições de entrevistas do *corpus* do Projeto VARSUL; 4.800 *tokens* e 1.266 *types*), 49,4% dos substantivos são uniformes femininos; 41,1% são uniformes masculinos; 8% são biformes – diferença significativa em relação aos dados do dicionário – ; e 1,4% são comuns de dois gêneros.

Em relação à ocorrência desses tipos de gênero em associação com marcadores formais (terminações fonológicas, classes formais), 95,1% e 89,6% das palavras terminadas em *-a* são femininas e 99,9% e 100% das palavras terminadas em *-o* são masculinas para os dados do dicionário e do *corpus* de uso falado, respectivamente. Para as palavras terminadas em *-e*, masculino e feminino ocorrem em distribuição equilibrada, com 52,8% e 47,2%, respectivamente. Embora haja uma quase exclusividade do gênero masculino para as formas terminadas em *-o*, com a exceção de itens como *tribo* e *libido*, há um número significativo de itens masculinos terminados em *-a*, não sendo o sistema, assim, isomórfico em relação a essas marcas.

Ao contrário dos substantivos uniformes, o subconjunto de itens que se referem a entidades sexuadas exhibe variação quanto ao gênero a ser exponenciado. Itens comuns de dois gêneros são inerentemente variáveis, podendo ser usados tanto no masculino quanto no feminino. Ao contrário dos determinantes, entretanto, sua potencialidade está sujeita a fatores semânticos e pragmáticos, sendo fixado um valor

de gênero anteriormente à sintaxe. De modo semelhante, pares de substantivos biformes são intercambiáveis com base nos mesmos critérios. No levantamento de Schwindt (2018), vemos que, no uso falado, para o recorte dos sexuados, os substantivos biformes são mais expressivos do que os comuns de dois gêneros. Uma vez que a categoria dos biformes define-se pela oposição *-o/-a* para masculino/feminino, essa seria, segundo o autor, uma justificativa para a intuição dos falantes quanto à isomorfia dessa distribuição em relação aos substantivos da língua de modo geral, o que, na verdade, não acontece.

Ainda, em se tratando de substantivos com referentes sujeitos às categorizações semânticas de gênero social e sexo, há exceções quanto à codificação desses valores na morfologia. É o caso de nomes uniformes como *vítima*, *criança* e *membro*, por exemplo. Não se atestam itens que determinam seu gênero morfológico dessa forma fora do campo semântico da correspondência com sexo – como acontece em outras línguas, cujos sistemas de gênero expressam traços do referente, tais como animacidade, caráter humano, comestibilidade, etc. (CORBETT, 1991, p. 7-32).

Conforme Corbett (1991), sistemas de gênero gramatical, de modo geral, podem ser classificados enquanto semânticos, formais, ou ambos, em menor ou maior grau. O português pode ser considerado um sistema parcialmente formal, dada a marcação do gênero feminino com a vogal terminal *-a*. Na seção 5.1, que aborda o experimento com pseudopalavras para o teste de critérios formais, traremos ainda correlações de outros segmentos terminais com o gênero feminino, que fundamenta a elaboração das pseudopalavras utilizadas. Para o subconjunto de substantivos com referentes sexuados, o português aproxima-se dos sistemas de gênero semântico neste contínuo tipológico, estando associado com o feminino semântico.

3 O FENÔMENO DA ATRIBUIÇÃO DE GÊNERO A SUBSTANTIVOS NOVOS

Neste capítulo, buscaremos delimitar o fenômeno da atribuição de gênero gramatical na adaptação morfológica de substantivos novos. Para isso, em 3.1 serão abordados trabalhos da literatura que colaboram para sua caracterização enquanto fenômeno linguístico e para seu enquadramento no campo dos estudos linguísticos, sendo eles Carvalho (1986), Haspelmath (2009), Haugen (1950, 1956) e Poplack (2020). Além disso, serão abordados alguns trabalhos sobre o funcionamento da atribuição de gênero nos casos específicos em que se atesta: dos empréstimos nominais em línguas com sistema de gênero nominal. Tais trabalhos incluem Barkin (1980), Poplack, Pousada e Sankoff (1982), Corbett (1991), Kilarski e Krynicki (2005) e Thornton (2009). Em 3.2, prosseguimos para a localização do fenômeno e dos critérios nele envolvidos em relação à língua que tomamos como objeto (o português brasileiro), estabelecendo o ponto de partida para as análises apresentadas nas duas seções seguintes, 4 e 5.

Notamos que a designação de “palavra nova” é abrangente, se relacionando aos diferentes casos de neologismos, neste trabalho, aqueles que passam a integrar o léxico nominal do português. A caracterização desses produtos da ampliação lexical não é fácil de ser feita, e compete também aos estudos da terminologia e da lexicografia, para além de estudos como o aqui apresentado.

3.1 REVISÃO TEÓRICA

O contato linguístico é responsável por uma variedade de fenômenos interessantes aos estudos linguísticos. Processos decorrentes do contato entre línguas podem ser observados a nível fonológico, morfológico, sintático e semântico, estando normalmente associados à implementação de novos itens lexicais em uma língua a partir de outra (SANKOFF, 2001). Ao ser integrado ao léxico da língua receptora e estar sujeito a seus processos gramaticais, um empréstimo lexical necessariamente se adapta às estruturas do novo sistema (HAUGEN, 1956).

Conforme elencado em Carvalho (1986), o fenômeno dos empréstimos está entre os processos que possibilitam a ampliação lexical, ao lado da alteração de um item pré-existente; da seleção de uma variante em detrimento de outras em um

sistema estável; da criação sistemática de novos itens a partir de regras internas ao sistema; e da economia funcional, em que distinções pouco salientes são apagadas. Nota-se que o fenômeno aqui investigado diz respeito não só aos empréstimos nominais e sua acomodação ao sistema receptor, mas também ao caso análogo da ampliação lexical por meio de pseudopalavras em ambiente experimental, não considerado nas caracterizações da literatura, por seu condicionamento artificial.

Uma vez que parte de nossa investigação se dedica à análise dos empréstimos lexicais, fazem-se necessárias algumas considerações a respeito desse fenômeno, em vista de sua caracterização.

Haspelmath (2009) entende empréstimo lexical (*loanword*)² como uma palavra que passa a fazer parte de uma língua em algum ponto de sua história a partir de outra língua. Segundo o autor, as formas de origem dos empréstimos podem ser complexas ou mesmo frasais, mas perdem sua estrutura interna ao serem implementadas no novo sistema, onde são, em princípio, não analisáveis.

Ao entender os empréstimos lexicais como produto de processos de mudança linguística, Haspelmath (Ibid., p. 38) sugere que uma dada palavra em uma língua pode sempre ser um empréstimo em potencial, podendo ter sido implementada em estágios diacrônicos que escapam ao conhecimento científico. Valemo-nos dessa visão diacrônica sobre os empréstimos para atentar ao recorte a ser abordado em 4.2. Por mais remota sua implementação, e ainda que entendidas por parte dos falantes como parte do léxico nativo do português, o *status* das palavras de origem árabe enquanto empréstimos, somado ao pressuposto uniformitarista quanto à mudança linguística, justifica as análises e discussões aqui propostas quanto à atribuição de gênero a itens novos em português.

Haugen (1956) traz mais considerações interessantes para a tipologia dos empréstimos lexicais e para a caracterização de seus processos na gramática da língua receptora.

Qualquer similaridade entre [as línguas envolvidas] é importação, enquanto qualquer diferença entre elas é vista como substituição do material nativo. Substituição quer dizer que a imitação do modelo estrangeiro é menos que

² Há uma questão terminológica em relação aos termos *borrowing* e *loanword*, recorrentes na literatura, ambos correspondendo a *empréstimo*, em português. Enquanto *borrowing* se refere ao processo de empréstimo, não só de itens lexicais, mas de estruturas e padrões prosódicos, fonológicos, sintáticos etc., *loanword* designa especificamente o produto do processo de empréstimo lexical (HASPELMATH, 2009, p. 38-39).

perfeita, mas também quer dizer que passou a se tornar mais familiar àqueles que falam a língua nativa.” (HAUGEN, 1956, p. 388, tradução nossa³).

A classificação dos empréstimos lexicais quanto à fidelidade às formas de origem se faz relevante por caracterizar os recortes de empréstimos apresentados no capítulo 4. Haugen (ibid.) divide os empréstimos lexicais em *loanwords*, em que tanto forma como significado são reproduzidos, como *tíquete*, *online*, etc.; *loanblends*, em que o empréstimo conserva significado e parte da forma – seria o caso da forma *goleiro* em português, em detrimento do empréstimo do equivalente *goalkeeper*–; *loan shifts*, onde apenas significado é importado – como *baixar* para o verbo do inglês *download*; e *loan translations*, em que é importado apenas significado, sendo a forma dada pela tradução literal da forma de origem, como os calques *cachorro-quente* e *palavra-chave* (para *hot dog* e *keyword*). Assim, são de nosso interesse os empréstimos lexicais caracterizados por Haugen como *loanwords*, uma vez que a integridade da forma de origem, vazia para o traço de gênero, garante o contexto para o fenômeno investigado.

Na adaptação de empréstimos, é possível que cada falante aplique um padrão individual ao novo item, propondo-o tacitamente através de seu uso contínuo à comunidade. É inapreensível a diversidade de fatores envolvidos em cada evento de produção e processamento dos empréstimos em seus estágios iniciais, ainda assim, itens lexicais emprestado gradualmente se estabilizam em novos sistemas linguísticos. Empréstimos lexicais são frequentemente classificados pelo seu grau de implementação ao sistema receptor. Comumente utiliza-se como critério o grau de adaptação fonológica ou mesmo grafêmica dos itens para diferenciá-los de casos de *code switching* ou de estrangeirismos, por exemplo.

Poplack (2020) aponta para um período de inconsistência dessas novas palavras na gramática, onde os empréstimos apresentam variações fonológicas e morfológicas, por exemplo, sendo mais expressivas que aquelas de produtos léxicos de outras naturezas; é nesse período que há variação de gênero em substantivos essencialmente uniformes. O modo como se classificam palavras em relação a esses critérios, no entanto, como nota a autora, depende do objetivo da análise a ser feita sobre os itens. Neste trabalho, como veremos na seção 4.1, dado o empréstimo em

³ “Any likeness between them is importation, while any difference between them is regarded as substitution. Substitution means that the imitation of the foreign model is less than perfect, but it also means that it has become more familiar to those who speak the native language”.

plena atividade de palavras do inglês no português brasileiro, consideramos como empréstimos, e não como casos de *code switching*, por exemplo, substantivos com pelo menos cem ocorrências em contextos diferentes em nosso *corpus* de referência, excetuando-se ocorrências em segmentos em inglês.

Poplack, Pousada e Sankoff (1982) é um dos primeiros estudos a abordar a questão da atribuição de gênero a empréstimos. Os autores utilizaram um corpus de 300 horas de entrevistas com falantes de espanhol porto-riquenho e francês de Montreal, de onde extraíram e compilaram todos os empréstimos do inglês detectados: um total de 872 ocorrências de 345 diferentes itens. Exemplos de segmentos contendo empréstimos incluem: “*Yo creo que mi hair está linda, beautiful*”. (“*Acho que minha hair está linda, beautiful*.”) (POPLACK; POUSADA; SANKOFF, 1982, p. 14) e “*Tes hot dogs sont pas prêtes, la steam, elle vient pas*”. (“*Seus hot dogs não estão prontos, o vapor não está vindo.*”) (Ibid., p. 14, grifo nosso).

Os autores consideram alguns critérios para a categorização dos itens. O primeiro critério, de “sexo fisiológico”⁴ (*physiological sex*), diferencia os substantivos entre aqueles que possuem um referente animado, que por sua vez pode diferenciar entidades masculinas e femininas de acordo com o gênero gramatical, e aqueles que não apresentam tal potencialidade. O critério de “tradução analógica” (*analogical translate*), por sua vez, identifica uma forma formal e semanticamente similar na língua receptora, p. ex., *establishment*, do inglês e *établissement*, do francês. Por último, considera-se a analogia com sufixos associados a um gênero na língua receptora, ou seja, se um empréstimo do inglês é fechado pela terminação *-tion*, é possível a analogia com o sufixo *-ción*, do espanhol, associado ao feminino.

Os achados do estudo apontam para a categoria semântica que distingue referentes de acordo com o “sexo fisiológico” (sexo biológico ou gênero social) como único critério significativamente regular para a atribuição de gênero. De modo geral, os critérios formais considerados no estudo variaram amplamente nas amostras analisadas, entre uma língua e outra e em cada uma delas tomadas separadamente.

Embora considerando a inexpressividade de critério fonológico para a atribuição de gênero, os autores chamaram atenção para o sufixo *-a* em espanhol – que, assim como no português, marca o gênero feminino – como apresentando

⁴ Neste trabalho, o termo “*physiological sex*” corresponde-se a seu equivalente próximo “sexo biológico”, e também à categoria de “gênero social”.

relativa regularidade, algo não verificável no francês. Os autores ainda ressaltam a relativa instabilidade dos empréstimos quanto ao gênero gramatical, excetuando-se os casos onde há um referente animado que suporte relação isomórfica entre sexo biológico/gênero social e gênero gramatical. Os autores notam que os itens passam por períodos em que mais de um gênero é adotado para cada empréstimo na comunidade de fala. “*Ouvimos le job, une interview em Paris, mas la job, un interview em Montreal*” (POPLACK; POUSADA; SANKOFF, 1982, p. 26).

A discussão apresentada pelos autores é pertinente à nossa investigação, primeiro, por reforçar a ideia da regularidade da atribuição de gênero gramatical condicionada pela correlação com traços semânticos de gênero e sexo; segundo, por pontuar os estágios de latência dos empréstimos nominais, quando ainda não há adaptação morfológica completa, produzindo os casos de variação de gênero em substantivos uniformes. Ainda que substantivos emprestados venham a se tornar estáveis no gênero masculino, interessam os critérios para atribuição de feminino em seus estágios iniciais. Ainda, em análises futuras, visamos à descrição e à análise de empréstimos recentes sujeitos a esse fenômeno e os fatores que condicionam a acomodação a um ou outro gênero.

Corbett (1991), na obra intitulada *Gender*, aborda a questão da atribuição de gênero a novos itens em seção dedicada à realidade psicolinguística do gênero gramatical. O autor propõe uma base explanatória para o fenômeno que serviu de referência para muitos estudos subsequentes – entre eles o de Thornton (2009), que se refere à obra como “a bíblia dos estudos de gênero” (Ibid., p. 34).

O autor classifica as línguas que possuem gênero gramatical de acordo com o componente que está mais associado à emergência do traço. Assim, algumas línguas baseiam-se mais em critérios fonológicos – como os segmentos terminais dos itens –, enquanto outras baseiam-se mais em informações do componente semântico – como, por exemplo, os traços de sexo ou gênero social para referentes humanos em português e espanhol. Podem, ainda, se basear em outros traços morfológicos dentro do sistema, como as classes flexionais de línguas como o latim e o russo, em que nomes são classificados de acordo com as formas do paradigma para caso, gênero e número.

Casos em que elementos fonológicos dos itens correlacionam-se com as classes de gênero do sistema são considerados pelo autor como exemplares do que

se classifica por "critérios fonológicos para atribuição de gênero". Corbett (1991) exemplifica esses casos mencionando o sistema altamente regular do hausa, língua afro-asiática falada na Nigéria. Nessa língua, palavras do gênero feminino infalivelmente terminam em *-aa*. Tal regularidade colabora para a atribuição do feminino a novos itens que adentram o sistema, quando estes terminam em segmentos percebidos como marca de gênero. Essa observação de Corbett sobre o hausa é relevante para nossas discussões sobre o papel da forma dos itens implementados no fenômeno da atribuição de gênero em português.

Os critérios semânticos para atribuição de gênero, segundo Corbett (1991), podem ser entendidos sob a generalização da associação conceitual. Esse mecanismo, presente em todas as línguas humanas e parte de nossa cognição, seria responsável por relacionar grupos de palavras sob uma mesma classe de gênero graças a suas propriedades semânticas, p. ex., o caso do português, mencionado acima, onde o gênero feminino é amplamente associado ao sexo feminino quando os referentes são animados. No entanto, o autor afirma que a atribuição de gênero a novos itens condicionada por critérios semânticos só se dá perante a ausência de fatores de natureza formal (fonológica ou morfológica), algo ainda disputado na literatura.

Neste trabalho, entendemos que, ao propor a generalização mencionada acima em relação ao condicionamento semântico no fenômeno da atribuição de gênero, Corbett deixa de considerar a complexidade dos diferentes critérios de base semântica potencialmente atuantes. Ainda, divergimos do posicionamento do autor quanto à sobreposição dos critérios de base formal sobre aqueles de natureza semântica. Nesse sentido, os resultados de nosso experimento com pseudopalavras corroboram com estudos que apontam maior influência dos fatores semânticos, em relação aos formais, no fenômeno de atribuição de gênero a itens novos.

Além das chamadas regras regulares (*normal rules*) para a atribuição de gênero, Corbett discute outros dois critérios já mencionados na literatura como importantes para o fenômeno. Segundo o autor, tais fatores devem ser considerados com cuidado, graças à sua natureza não sistemática: são critérios que se aplicam a casos isolados. É assim que o autor vê, por exemplo, o caso dos itens que adotam, por analogia semântica, o gênero de um item já presente na língua receptora. Portanto, quando Poplack, Pousada e Sankoff (1982) trazem à discussão o item

interview, emprestado do inglês para o francês, e explicam seu uso no feminino como sendo condicionado pela presença *a priori* de um item formal conceptualmente análogo no francês, para Corbett, não estariam em jogo regras sistemáticas.

Mais uma vez, vamos de encontro ao que supõe Corbett sobre o fenômeno. Ainda que, dado um par de línguas com sistemas de gênero análogos, exista a interferência do conhecimento bilíngue, como o exemplo mencionado, é possível observarem-se processos mais ou menos regulares que se opõem a essas analogias, como observamos em nossa análise dos empréstimos de origem árabe no português (4.2). Principalmente, a afirmação do autor quanto à não sistematicidade de processos de adaptação de novos itens ao sistema de gênero é prejudicial por negligenciar um objeto de análise que tem se provado frutífero para o debate científico.

O segundo caso não regular de Corbett (1991) é o que chama de “atribuição automática” (*automatic assignment*), que diz respeito às relações de marcação no interior de um sistema de gênero gramatical: quando o mecanismo de atribuição de gênero gramatical a um item novo não é capaz de computar regras formais que condicionem a escolha de um ou outro gênero, o gênero não marcado do sistema em questão seria atribuído ao item. Corbett discorda desse pressuposto, trazendo casos em que a explicação por critérios de natureza morfológica e fonológica substituem satisfatoriamente a atribuição por não marcação. Neste estudo, optamos por considerar a atribuição do não marcado como um fator independente, atuante na atribuição de gênero a itens novos em português. Contudo, entendemos o posicionamento de Corbett quanto a esse fator como compatível à visão aqui adotada. Ao assumirmos o feminino como gênero marcado da língua, são os casos em que ele é atribuído às novas palavras que devem ser explicados.

Barkin (1980) apresenta questões que interessam à nossa discussão, principalmente pelo enfoque que se distancia, de certo modo, dos traços gramaticais internos aos itens, revelando outras características importantes do fenômeno. A autora leva em consideração características dos empréstimos e fatores periféricos do fenômeno de atribuição que, até então, haviam sido pouco considerados na literatura: a competência bilíngue dos falantes, que utilizam os empréstimos, as fronteiras entre os dois sistemas por eles internalizados e os diferentes graus de adaptação percorridos pelos empréstimos, desde o momento em que passam a fazer parte do novo léxico até o momento em que estão plenamente inseridos no sistema.

Barkin (Ibid.) analisa os empréstimos do inglês em um *corpus* de língua falada composto de entrevistas com falantes de espanhol como primeira língua. Utilizando o sistema de classificação de Haugen (1953 apud BARKIN, 1980, p. 105), a autora classifica os itens levantados em i) inteiramente adaptados, ii) parcialmente adaptados e iii) não assimilados.

A importância desses fatores para o fenômeno se mostra, por exemplo, quando a autora aponta para a tendência dos falantes de não usar determinantes no mesmo sintagma de itens ainda não assimilados ao sistema lexical da língua em uso, ou seja, substantivos que não passaram por processos fonológicos adaptativos e são irreconhecíveis por monolíngues como palavras da língua em questão, como em casos de *code switching*.

Segundo Barkin, a tendência tem a ver com os níveis de consciência dos falantes em relação aos dois sistemas aos quais têm acesso. Evitar o uso de artigos e adjetivos com palavras totalmente alheias ao sistema, por exemplo, estaria relacionado à consciência dos falantes bilíngues quanto à ausência de gênero nominal no inglês.

Finalmente, a autora comenta a escassez de evidência para a atribuição de gênero condicionada por fatores formais – devida à baixa ocorrência de palavras terminadas em vogais átonas no inglês – e identifica as associações baseadas no componente semântico como aquelas que de modo mais regular condicionam a atribuição.

Kilarski e Krynicki (2005) analisam a atribuição de gênero a empréstimos do inglês no dinamarquês, no sueco e no norueguês. Os autores, a exemplo de Corbett (1991), dividem os critérios possivelmente atuantes na atribuição de gênero em dois grupos – formais e semânticos – executando uma análise minuciosa para determinar os critérios mais relevantes para os casos investigados. Essa análise dá-se através do modelo de função discriminante, que permite testar o grau de predizibilidade para atribuição de gênero em um dado item. Nesse modelo, são testados todos os critérios previamente alimentados e podem ser observados aqueles que se mostraram atuantes no emprego de um ou outro gênero.

O estudo busca conclusões em duas frentes: primeiramente, descobrir a importância relativa de diferentes critérios na atribuição de gênero dos itens vindos do inglês nas três línguas; em segundo lugar, verificar até que ponto as especificações

gramaticais desses itens determina, ou seja, é capaz de prever, seu gênero nas línguas receptoras. Para alcançar conclusões nessas duas direções, os autores estipularam um amplo conjunto de fatores variáveis que poderiam exercer um papel sobre a atribuição de gênero. Os 19 critérios escolhidos, em parte, referem-se a traços específicos das línguas germânicas do Norte, por exemplo, seus sistemas de gênero bipartites, para o sueco e o dinamarquês, e tripartite, para o norueguês, sem correlações com a semântica dos referentes sexuais, como é o caso do português e do espanhol.

Além de critérios morfológicos dessa natureza, outros critérios morfológicos considerados foram o *status* derivacional dos itens, o sufixo e as ocorrências no plural e no singular. Já os critérios fonológicos incluíram polissilabidade, quantidade do último fonema, quantidade do penúltimo fonema, número de consoantes finais, presença ou não de parada glotal final (para o dinamarquês), entre outros. Finalmente, os critérios semânticos considerados identificavam os nomes como animados ou inanimados, concretos ou abstratos, pessoais ou não pessoais, e também de acordo com a presença de um equivalente na língua receptora, como, por exemplo, em português, o uso de *hot dog* em vista da palavra equivalente *cachorro-quente*.

Os resultados quantitativos revelaram que, em dinamarquês e sueco, o gênero comum (um dos dois gêneros desses sistemas) está super-representado no subconjunto dos empréstimos, ou seja, havia relativamente mais nomes de gênero comum dentre os empréstimos do que no repertório da língua como um todo. O modelo de análise discriminante aplicado corroborou o quadro, tendo selecionado o tipo de gênero como uma das variáveis relevantes para o fenômeno. Em relação à predizibilidade na atribuição de gênero gramatical a um dado item – um dos objetivos na utilização do modelo citado – uma regularidade meramente moderada foi verificada nos sistemas gramaticais analisados: o modelo tendo selecionado, no melhor cenário, 71.9% dos casos do dinamarquês.

O trabalho de Thornton (2009) sobre atribuição de gênero a empréstimos traz algumas das bases que utilizaremos para a análise dos empréstimos do inglês no português. Adotamos o modelo da autora para essa análise pelo seu poder descritivo. Os diferentes critérios semânticos por ela propostos dão conta da diversidade de casos observados no recorte de empréstimos do inglês.

A abordagem da autora se dá, inicialmente, sobre as bases descritivas de Corbett (1991), operando uma divisão entre critérios de natureza formal, de um lado, e semântica, de outro, para a atribuição de gênero a itens novos nas línguas. Analisando o comportamento de topônimos, neologismos e empréstimos do inglês no italiano, umas das principais contribuições da autora é o aprofundamento dos critérios de natureza semântica envolvidos na atribuição de gênero gramatical.

Thornton (2009) considera a analogia semântica como um conjunto de critérios que podem condicionar a atribuição ativa de gênero gramatical. Ao estudar certos grupos de palavras, a autora estipulou critérios semânticos menores potencialmente ativos no fenômeno: *gênero do hiperônimo*, *equivalente* e *associação*. No primeiro caso, quando um empréstimo entra em uma categoria conceitual maior já presente na língua receptora, é possível que adquira o gênero do substantivo genérico que rotula aquela categoria; assim, *uma yamaha* designa uma motocicleta (feminino), enquanto que *um yamaha* designa um instrumento musical (um teclado, masculino). Já a associação de conceitos equivalentes diferencia-se por, neste caso, a língua receptora contar com uma palavra que designa estritamente o mesmo elemento nomeado pelo empréstimo; enquanto que, naquele caso, frente à ausência do conceito no léxico mental dos falantes, um conceito associado é acionado e seu gênero projetado no empréstimo. Assim, a palavra *call*, usada no feminino em português, tem seu gênero atribuído através do acionamento de *chamada*, um equivalente tradutório; por outro lado, *ecobag*, um conceito novo, é usado no feminino diante do acionamento de *bolsa*, um conceito associado já existente no léxico mental dos falantes de português.

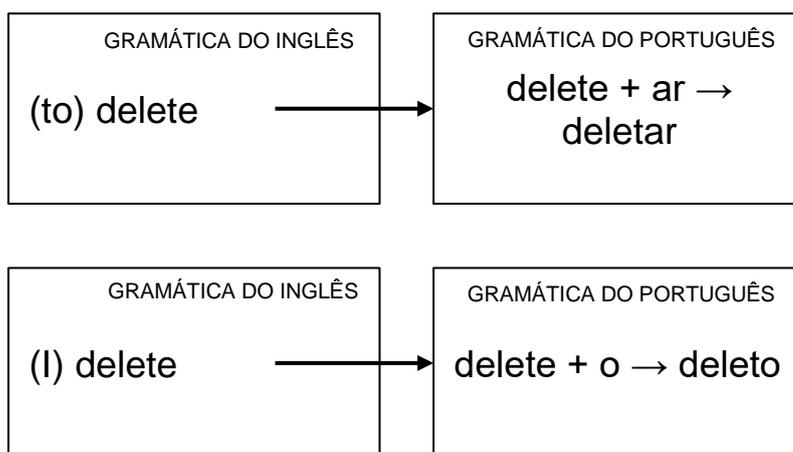
Por fim, Clegg (2010) fundamenta um de nossos experimentos com pseudopalavras que serão abordados no capítulo 5, especificamente o que visa ao teste dos critérios formais na atribuição de gênero. Seu trabalho se baseia em levantamentos feitos previamente na literatura sobre sufixos estatisticamente relacionados com o gênero feminino, presentes em dicionários de língua espanhola. O autor elabora um experimento onde os sufixos são testados através das respostas dos falantes às questões, que revela se há ou não o processamento das palavras com tais sufixos no gênero feminino. Por exemplo, os sufixos *-dad* e *-umbre*, presentes em grande parte em palavras femininas no léxico dicionarizado, se mostraram

condicionantes para a atribuição de feminino nas pseudopalavras *alaridad* e *colidumbre*, criadas para o estudo.

3.2 ATRIBUIÇÃO DE GÊNERO A SUBSTANTIVOS NOVOS EM PORTUGUÊS

Atribuir características morfológicas ausentes em sua língua de origem é um processo de adaptação que leva à estabilidade dos itens dentro do sistema. A acomodação dos empréstimos ao sistema morfológico da língua receptora pode ser observada, por exemplo, em empréstimos verbais do inglês no português brasileiro, por exemplo. Para que possam funcionar na gramática da língua receptora, os verbos recebem os morfemas abertamente expressos do paradigma verbal do português, conforme demonstra a Figura 1:

Figura 1 – Adaptação morfológica do verbo ‘deletar’ em português



Fonte: elaborado pelo autor.

A forma verbal *delete* é implementada em português brasileiro, podendo ser expresso por *deletar*, com a afixação de *-ar*, mas também tomando a forma de primeira nominativa de primeira pessoa singular *-o* e todas as demais formas correspondentes ao paradigma flexional dos verbos na língua receptora. O fenômeno da adaptação lexical dos empréstimos interessa a investigações de cunho fonológico dado o pressuposto de que regras inerentes ao sistema receptor aplicam-se no processo. Da mesma forma, o fenômeno faz-se interessante para a morfologia por permitir que se observem não só os produtos da aplicação dessas regras, mas também as condições

contextuais que determinam o surgimento de dadas formas em detrimento de outras, como é o caso dos empréstimos a que se atribui o gênero feminino.

Como vimos, há diferentes propostas quanto aos fatores possivelmente atuantes na atribuição de gênero a itens que passam a integrar as línguas investigadas pelos trabalhos abordados. Para fundamentar sua análise sobre empréstimos do inglês em variedades do francês e do espanhol, Poplack, Pousada e Sankoff (1982) reúne pressupostos recorrentes na literatura sobre o tema, conforme apresentados no Quadro 1.

Quadro 1 – Critérios para atribuição de gênero recorrentes na literatura

- | |
|--|
| <ul style="list-style-type: none"> a. Sexo do referente animado b. Forma fonológica c. Equivalente semântico na língua receptora d. Item homófono na língua receptora e. Afixo homófono na língua receptora f. Forma grafêmica g. Conformação ao gênero não marcado |
|--|

Fonte: adaptado de Poplack, Pousada e Sankoff (1982, p. 5-6⁵, tradução nossa⁶).

Com base nos critérios elencados em (1) acima, busca-se estipular uma relação de critérios adaptada às análises propostas neste trabalho, apresentadas e discutidas nas próximas seções (4 e 5).

Primeiramente, se faz relevante o critério exposto em (a), que, conforme apresentado, faz parte do sistema do português e é observado, por exemplo, na morfologia dos substantivos comuns de dois gêneros. Em relação aos empréstimos em português brasileiro, tomemos como exemplo o caso do item do inglês *popstar*, a que se atribui feminino gramatical dado um referente feminino, passando a integrar o subconjunto dos comuns de dois gêneros.

Entendemos que os critérios em b., d., e. e f. correspondem, de modo geral, à atribuição de gênero gramatical por índices de natureza formal, não se fazendo relevantes as diferenças entre cada caso, expressas pela subcategorização desses critérios. Portanto, como veremos, não se distinguem a atribuição de gênero a itens

⁵ As referências, conforme apresentadas pelos autores, dão-se da seguinte forma: pressupostos em (a-e) atribuem-se a Haden, Joliat (1940); Reed (1942); Weinreich (1953); Haugen (1969); e Ibrahim (1973). Aquele em (f), a Arndt (1970). Aquele em (g), a Haugen (1969) e Correa-Zoli (1973).

⁶ Originais: (a), *physiological sex of animate referent*; (b), *phonological shape*; (c), *host language semantic equivalent*; (d), *host language homophone*; (e), *host suffix*; (f), *graphemic image*; e (e), *unmarked gender of the host language*.

introduzidos em sua forma grafêmica – possivelmente como ocorre com parte dos empréstimos do inglês, dada sua difusão por suportes escritos dos ambientes virtuais – daquela dos empréstimos por uso falado, como os derivados de línguas indígenas ou africanas, em contato com o português brasileiro desde a colonização – por exemplo, *pitanga* e *berimbau*.

Por outro lado, ao adotar-se o critério em c., que sugere associação com um equivalente na língua receptora, propõe-se sua subcategorização, considerando-se o já apresentado modelo de Thornton (2009). Assim, entende-se a *analogia com equivalente* sugerida em c. como uma das possibilidades de condicionamento por *analogia semântica* com o léxico da língua receptora, ao lado da *analogia por hiperônimo* e *analogia por conceito associado* (Ibid., p. 194).

Por fim, a perspectiva que adotamos quanto à gramática do português implica a existência do critério referido em g., sendo a não marcação do masculino um pressuposto teórico fundamental em perspectivas como a de Câmara Jr. (1970), discutidas no capítulo 2.

Feitas as considerações sobre os critérios de atribuição de gênero no Quadro 1, dispõem-se no Quadro 2 aqueles que serão utilizados na análise dos casos propostos neste trabalho:

Quadro 2 – Critérios para atribuição de gênero considerados

1. Traço semântico do referente
2. Forma do item
3. Equivalente semântico na língua receptora
4. Hiperônimo na língua receptora
5. Conceito associado na língua receptora
6. Conformação ao gênero não marcado

Fonte: elaborado pelo autor.

Propõe-se ainda o agrupamento desses fatores de acordo com a natureza dos processos e com sua interação com a gramática, conforme demonstra o Quadro 3.

Quadro 3 – Critérios para atribuição de gênero agrupados

1. Semântica do referente
 - i. Sexo feminino ou gênero social feminino
2. Analogia semântica com léxico receptor
 - i. Equivalente semântico na língua receptora
 - ii. Hiperônimo na língua receptora
 - iii. Conceito associado na língua receptora
3. Forma do item
4. Conformação ao gênero não marcado

Fonte: elaborado pelo autor.

Os fatores elencados acima dizem respeito a interações diversas às quais se sujeita o traço de gênero na gramática do português, e são importantes para uma análise que considera os critérios para a atribuição de gênero ativa, processada no momento da entrada de novos itens.

4 EMPRÉSTIMOS EM PORTUGUÊS

Considerando-se o que afirma Corbett (1991) sobre o fenômeno da atribuição de gênero gramatical a novos itens e a possibilidade de sua observação no fenômeno dos empréstimos, apresentamos neste capítulo a análise de dois casos em que um grande número de palavras estrangeiras foram adaptadas ao português. Primeiramente, na seção 4.1, abordamos o caso dos empréstimos do inglês para o português brasileiro, que rende vários substantivos novos cujo processo de empréstimo está ainda em atividade na língua. Em 4.2, analisamos o recorte dos itens de origem árabe, que entraram nas variedades românicas da Península Ibérica principalmente entre os séculos VIII e XV AD, variedades que viriam a originar, entre outros, o português. Como veremos, esses dois recortes são interessantes por incluírem itens que divergem, não só em seu período de entrada na língua, mas também em sua estrutura, o que nos possibilita a observação de fatores diversos possivelmente atuantes na atribuição de gênero.

Ainda, a exemplo dos critérios formais que assumimos presentes na adaptação morfológicas dos itens árabes, atestam-se alguns exemplos de substantivos vindos de outras línguas que parecem ter sido empregados sistematicamente no feminino devido a fatores formais, uma vez que já eram fechadas por fonemas ou grafemas analisáveis como *-a* nas formas de origem. Levando em conta a diversidade das famílias linguísticas das línguas de origem, exemplos incluem os empréstimos femininos *jabuticaba*, *sauna* e *catana*, respectivamente do tupi-guarani /i'apoti'kaba/, do finlandês *sauna*, e do japonês *かたな*, /kĩ'tʰanə/.

4.1 EMPRÉSTIMOS DO INGLÊS EM PORTUGUÊS BRASILEIRO

O inglês compartilha as raízes indo-europeias do português e de outras línguas que, hoje em dia, contam com sistemas de gênero nominal. Essa língua, no entanto, passou por processo de empobrecimento morfológico em relação a essa categoria nominal em etapas anteriores da língua⁷. Atualmente, o gênero gramatical em inglês não se apresenta como sistema morfossintático, embora ainda se atestem,

⁷ Já no inglês antigo, a língua não contava mais com as expressões de gênero nos substantivos, que antes distinguíam feminino, masculino e neutro, e que perduram, por exemplo, no alemão moderno.

residualmente, no sistema pronominal da língua, a expressão de gênero semântico pelas formas pronominais pessoais de terceira pessoa *he* e *she*, *ele* e *ela*.

Para a investigação proposta neste trabalho, analisaremos o conjunto de substantivos que passam a fazer parte do léxico do português brasileiro a partir do inglês. Segundo classificação de Haugen (1956), são casos de *loanword*, em que se conservam significado e forma de origem, descartando-se a atribuição guiada por formas já familiares ao sistema. Ainda, a ausência de expressão de gênero nos substantivos do inglês é relevante por excluir a possibilidade de atribuição por analogia com o gênero dos itens na língua doadora, a que estão sujeitos empréstimos do francês e do espanhol, por exemplo.

A implementação ativa desses itens no português brasileiro relaciona-se ao processo da globalização e ao fato de o inglês ser falado como língua franca nos meios de comunicação virtuais. A adoção de novos conceitos nesses contextos se substancia, em maior parte, por substantivos emprestados, classificados como empréstimos culturais (HASPELMATH, 2009, p. 49). Em menor grau, alguns dos empréstimos são implementados para denotar conceitos já conhecidos, com equivalentes em português, os quais configurariam casos de *core borrowing*, motivados por prestígio ou estilo.

4.1.1 Metodologia

Nosso *corpus* foi composto, inicialmente, pelos *corpora* de Cano e Prado (2006) e Orsi (2015), que reúnem empréstimos do inglês em português brasileiro referentes ao domínio da informática e da moda, respectivamente. Dada a contínua implementação desses itens à língua, passamos a compor o *corpus* com itens levantados por observações empíricas, verificando-se seu uso efetivo na língua por suas ocorrências no Corpus do Português: Web/Dialects, que conta com mais de um bilhão de palavras retiradas de páginas da *web* e é atualizado mensalmente.

O *corpus* totalizou 242 substantivos, dos quais: 183 são uniformes empregados predominantemente no masculino, como *notebook* e *hamburger*; 52 são uniformes femininos, como *webcam* e *picape*; e 17 são substantivos comuns de dois gêneros, como *crush* e *influencer*. Dada a elaboração do *corpus*, partiu-se para a codificação dos itens de acordo com as variáveis: *segmento terminal* – terminadas em *-a*, *-e*

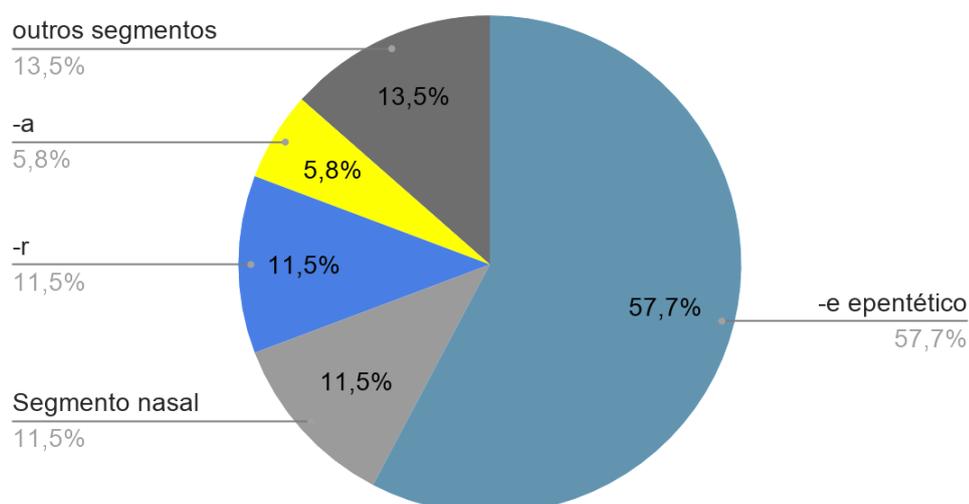
epentético, *-er*, *-l*, *-s*, *-ks*, etc.; e *tipo de gênero* – feminino, masculino ou comum de dois gêneros. Dados nossos pressupostos sobre a marcação do feminino na língua, centramos nossa análise nas 52 palavras com uso significativo no feminino, as quais foram codificadas também de acordo com a variável *critério para atribuição de gênero feminino* – processo detalhado abaixo. Por fim, os dados foram processados estatisticamente utilizando o programa IBM SPSS Statistics.

Foi aplicado o modelo de Thornton (2009), apresentado acima, para classificar os empréstimos femininos de acordo com analogia semântica com itens femininos já presentes no inventário do português enquanto critério para a atribuição de gênero. Assim, poderiam atuar no processo a analogia com um conceito associado (*associate*), com um hiperônimo (*hyperonym*), ou um conceito equivalente (*equivalent*). Tomando como exemplos itens de nosso *corpus*, *ecobag* estaria relacionado aos conceitos associados *bolsa* ou *sacola*; *legging* seria o hipônimo de *calça* (calça do tipo *legging*); e *homepage* seria um equivalente tradutório de *página inicial*. Para o caso dos critérios formais, determinou-se a terminação átona final *-a* – ou grafema *<a>* – como possível condicionante para atribuição de feminino.

4.1.2 Resultados

Conforme demonstra a Figura 2, em relação à variável *segmento terminal*, os empréstimos femininos dividiram-se entre fechados por: *-e* epentético (58%, como *trademark* e *picape*), consoantes nasais (11,5%, como *touchscreen* e *webcam*), *-r* (11,5%, itens como *newsletter* e *happy hour*), *-a* (6%, como *creepypasta* e *mídia*) e *outros segmentos* (13%, como *jukebox*, *skill*, *fake news*). Notamos ainda que 47% dos dezessete itens comuns de dois gêneros apresentavam a terminação *-er*, do sufixo agentivo do inglês (como *hater*, *gamer*, *youtuber*). Esses itens, quando utilizados no feminino, seguem o padrão do funcionamento dos itens comuns de dois gêneros em português, fazendo referência ao sexo do referente, como em ‘*a youtuber*’, designando referente feminino.

Figura 2 – Segmento final de empréstimos femininos do inglês



Fonte: elaborado pelo autor.

Quanto aos critérios possivelmente atuantes na atribuição de gênero aos empréstimos, os 52 itens femininos distribuíram-se entre aqueles condicionados por analogia com: conceito associado (54%), equivalente tradutório (25%), hiperônimo (7%); e por presença do segmento final *a* na forma de origem (5%, três casos: *creepypasta*, *mídia* e *sinuca*) (Tabela 1). Não foi possível relacionar três dos itens aos critérios de atribuição considerados: *gangue*, *van* e *performance*. Nota-se que são itens que foram implementados no português em um período anterior àquele da maioria dos outros itens, o que faz com que os conceitos e palavras possivelmente associados a eles no momento da atribuição, e a própria associação, sejam mais difíceis de serem resgatados.

Tabela 1 – Critério de atribuição de gênero a itens femininos do inglês

Critérios	Exemplos
-a → fem. (3)	<i>mídia</i> , <i>creepypasta</i> , <i>sinuca</i>
Item associado no inventário (31)	<i>selfie</i> 'foto', <i>playlist</i> 'lista'
Equivalente no inventário (10)	<i>night</i> 'noite', <i>timeline</i> 'linha do tempo'
Hiperônimo no inventário (5)	<i>legging</i> 'calça', <i>scooter</i> 'motocicleta'
Nenhum critério (3)	<i>gangue</i> , <i>van</i> , <i>performance</i>

Fonte: elaborado pelo autor.

4.1.3 Discussão

Devido à fonotática do inglês, predominaram no recorte a atribuição de gênero por critérios semânticos, a partir da analogia com itens já presentes no inventário. Ou seja, a pouca ocorrência na língua de itens terminados em *-a* ou quaisquer outros segmentos que possam ser associados ao feminino em português faz com que a forma desses empréstimos não seja relevante para a escolha na maioria dos casos.

A prevalência do condicionamento por analogia com conceito associado dentre os critérios semânticos relaciona-se com a mencionada adoção de novos conceitos no contexto de globalização. Novas tecnologias e novas práticas, por exemplo, chegam aos falantes sem que se possa encontrar prontamente em seus repertórios palavras que as denotem. Os chamados empréstimos culturais representam uma quase necessidade de implementação de empréstimos. Por outro lado, *core borrowings* configuram a adoção espontânea de empréstimos em vista de equivalentes na língua receptora (HASPELMATH, 2009). Envolvendo fatores de natureza estilística e o valor atribuído à língua de origem, embora pouco numerosos na amostra, são bastante estáveis em português brasileiro, por exemplo, *night* e *bike*.

Nos três itens para os quais se propõe condicionamento por critério formal, nota-se a natureza diversa dos processos que resultaram na presença do segmento vocálico final nos três casos observados. A palavra *mídia*, por exemplo, entra no inglês através do latim, como *media*, plural de *medium* ('meio'); embora estranha ao padrão fonotático do inglês, manteve-se a vogal átona final. *Creepypasta* deriva do portmanteau *copypasta* (que, por sua vez, não figurou como empréstimo em português), usado em comunidades da internet para se referir a contos que são difundidos ao serem copiados (*copy*) e colados (*pasted*) sucessivamente⁸. A terminação *-pasta*, nesse caso, está relacionada ao registro na grafia do processo de apagamento do segmento final *-d* no inglês e do abaixamento da última vogal. É semelhante o caso de *sinuca*, com origem na forma do inglês *snooker*, possivelmente associado ao feminino em português frente a sua pronúncia com apagamento do *-r* final e abaixamento da vogal precedente: /'snu.ka/.

⁸ <https://time.com/2818192/creepypasta-copypasta-slender-man/>. Acessado em 24/10/2020.

4.2 EMPRÉSTIMOS DO ÁRABE NO PORTUGUÊS

Os empréstimos do árabe representam um recorte numeroso de itens com origem compartilhada em português, tendo entrado na língua em sua maioria entre os séculos VIII e XV AD. Foram quase oito séculos de contato entre falantes de, entre outros, o árabe andaluz e falantes de variedades do latim vulgar então presentes na Península Ibérica que viriam a originar as línguas ibero-românicas modernas, como o catalão, o português e o espanhol.

Segundo Corriente (2008), o léxico nominal das línguas românicas ibéricas modernas constitui-se de: uma minoria de palavras das línguas ibéricas de substrato, pré-romanas; uma porção de palavras de etimologia germânica, oriundas da breve presença de grupos germânicos, ao rastro da queda do Império Romano; uma maioria de palavras latinas, derivadas do latim vulgar de onde tem origem; e uma quantidade considerável de palavras do árabe. Analisaremos estas, por representarem um dos mais numerosos recortes de empréstimos com uma origem comum, mas também por gozarmos de uma variedade de fontes lexicográficas, que apontam não só para a origem destes empréstimos no português, como para a forma de étimos dos itens nas variedades do árabe então em vigor.

De acordo com a classificação de Haugen (1956), a maior parte do léxico árabe implementado é representada por empréstimos culturais, denotando conceitos até então ausentes na visão de mundo dos povos nativos da Península. São itens que representam novos bens de consumo, como *laranja*, *azeitona* e *romã*; tecnologias do domínio têxtil, como *algodão*, *alfinete* e *almofada*; e elementos comerciais, como *armazém*, *açougue*, *aduana*.

Em português, o critério potencialmente atuante seria aquele relacionado à morfofonologia dos nomes do português, já que muitas das palavras em árabe terminam em segmentos que poderiam ser percebidos como -a átono final, em formas árabes como as formas de étimo de *alfafa* e *alfazema*: <faʃfaʃa> e <al-ḵuzāmā>⁹. Dado o distanciamento temporal em relação ao período da implementação da maioria desses itens, entendemos ser praticamente inapreensível o papel de critérios de natureza semântica para a atribuição de gênero a esses itens, como se observaram

⁹ Conforme informação etimológica disponível no Dicionário Priberam da Língua Portuguesa, onde <a> e <â> representam vogais átonas finais curta e longa, respectivamente.

no recorte dos itens do inglês. Uma análise com base no modelo de Thornton (2009), exigiria uma consideração do ambiente conceitual dos falantes que participaram do contato linguístico na época em que ele ocorreu.

4.2.1 Metodologia

Para a análise dos fatores possivelmente atuantes na atribuição de gênero a empréstimos do árabe, levantamos os 159 substantivos do dicionário Priberam da Língua Portuguesa contendo a notação etimológica *ár*. A escolha da obra deu-se, para além de sua reconhecida qualidade, pela possibilidade de acesso ao recorte proposto e pela riqueza de disposição de informações etimológicas, como a forma dos étimos, importante para nossa análise.

Os itens levantados foram codificadas de acordo com as variáveis: *tipo de gênero* (feminino, masculino ou comum de dois gêneros), *segmento terminal* (-a, -o, -e, ditongos ou consoantes) e *grafema terminal do étimo*, conforme notação da fonte consultada (o que inclui formas como <a>, <â>, <m>, <h> e clusters consonantais). De forma análoga ao recorte realizado na etapa anterior com os itens do inglês, e pelas mesmas razões, tomamos como objeto central os 75 substantivos femininos uniformes identificados, desta vez, de acordo com sua classificação de gênero nas notações morfológicas trazidas pelo dicionário utilizado.

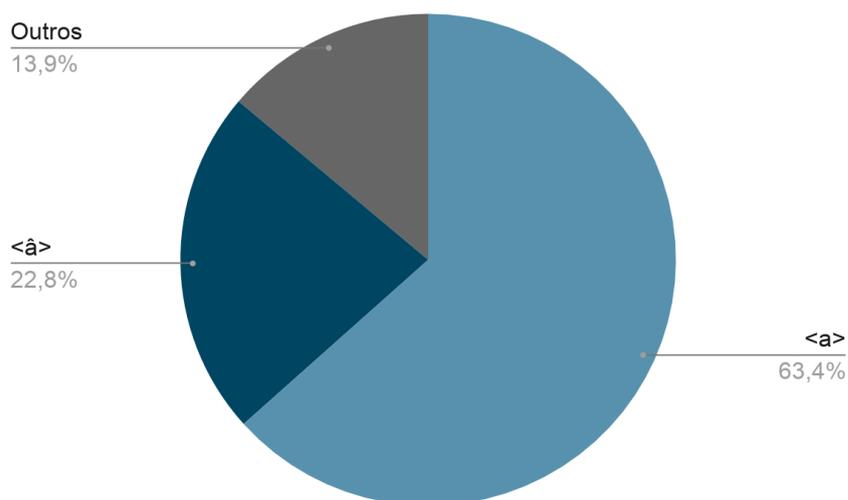
Uma vez compilado e codificado o *corpus*, os itens reunidos foram analisados estatisticamente através do programa IBM SPSS Statistics, em vista da comparação entre as ocorrências do gênero feminino, as terminações dos empréstimos em português e suas terminações nas formas de étimo. Assim, considerando o referido critério formal para atribuição de gênero, buscou-se identificar palavras que continham um segmento final análogo a -a em árabe e que foram empregadas no feminino em português.

4.2.2 Resultados

Conforme mostra a Figura 3, dos 75 itens femininos, 73 (97%) terminam na vogal átona final -a em português, um em -e (*alface*) e um em ditongo nasal (*monção*). Desses itens fechados por -a, 65 (87%) possuem étimo fechado por um grafema que

pode ser generalizado como a marca de feminino na língua: <a> e <â>, segundo a notação da fonte consultada – são itens como em *argola*, derivado de <al-gulla>; e *almofada*, derivado da forma <al-mukhaddâ>. O restante dessas palavras apresentou étimos terminados em <m>, <q>, <bd>, entre outros.

Figura 3 – Distribuição de terminações dos étimos de femininas do árabe terminadas em /a/



Fonte: elaborado pelo autor.

4.2.3 Discussão

A coocorrência observada entre grafemas do étimo com o segmento -a dos itens importados do árabe para o português sugere a presença de um critério formal de atribuição de gênero também na diacronia. Tal critério determina que o feminino seja atribuído a itens analisáveis como contendo a marca de feminino em português, representado por *fem.* → -a.

No entanto, por maior o distanciamento genético entre as línguas envolvidas no fenômeno, há marcantes semelhanças morfofonológicas entre os sistemas nominais das duas línguas. O árabe, assim como o hebraico e outras línguas semíticas, vale-se de um sistema de gênero relativamente semelhante ao das línguas românicas, em que seus substantivos se subdividem entre femininos e masculinos. Mais ainda, atenta-se à semelhança fonológica de um grupo de substantivos do árabe com subsistemas léxicos da maioria das línguas ibero-românicas, incluindo-se o português. Nomes femininos, em muitas variedades do árabe, são fechados por um

segmento vocálico átono *-a*, derivada da terminação *-at*, ainda atestável em algumas variedades e emergente enquanto fenômeno fonossintático no restante delas (CORRIENTE, 2008). Esta coincidência entre as gramáticas das línguas observadas traz alguns obstáculos para nossa análise. Assim como em línguas aproximadas geneticamente, é possível que o gênero dos empréstimos tenha sido deliberadamente transferido por falantes bilíngues ou com algum conhecimento do árabe da então variedade românica que originaria o português.

5 EXPERIMENTOS COM PSEUDOPALAVRAS

Para analisar com maior profundidade o fenômeno da atribuição de gênero, foram elaborados e aplicados dois experimentos, que visam testar os critérios de natureza semântica e formal que podem agir na atribuição de gênero a um novo item que entra no sistema do português. Ao passo que a análise dos *corpora* de empréstimos nominais que se implementam naturalmente ao português nos fornecem dados numerosos e sujeitos a toda a real complexidade do uso, experimentos como os que apresentaremos abaixo permitem que sejam manipulados a forma, o conteúdo, e os contextos de atribuição dos itens, aceitando as limitações desses contextos artificiais em detrimento de algumas vantagens do ponto de vista analítico.

5.1 TESTE PARA CRITÉRIOS FORMAIS DE ATRIBUIÇÃO DE GÊNERO

Com o objetivo de testar os critérios fonológicos potencialmente atuantes no fenômeno de atribuição de gênero a substantivos novos em português, elaboramos dois experimentos com pseudopalavras. Nesta seção, apresentamos o primeiro deles, que trata dos critérios formais de atribuição de gênero, aqueles em que o falante escolhe o gênero do item novo com base em segmentos fonológicos, grafêmicos, ou padrões morfológicos da forma processada. Em português, isso quer dizer, por exemplo, atribuir o feminino ao entender-se a terminação de uma palavra nova como *-a*, a marca de gênero da língua, correlacionada estatisticamente ao feminino. Pela segunda razão, também é possível que a percepção de outros segmentos em uma palavra nova também oriente a escolha pelo feminino. Sendo assim, foram testadas no experimento uma série de terminações coincidentes com o gênero feminino e que poderiam estar a ele associadas.

5.1.1 Metodologia

Como vimos no caso dos empréstimos do árabe, a correlação entre segmentos fonológicos e expressões de gênero parece estar ativa no processamento de novos itens em português, o que se reflete no gênero atribuído a itens novos. Para este experimento, tomou-se como fundamento o *corpus* do Dicionário Eletrônico Aurélio e

do Projeto Varsul para o levantamento de associações formais com o gênero feminino para além de *-a*. Há uma série de outros segmentos finais recorrentes na língua que estão estatisticamente relacionados ao gênero feminino, dentre eles, os sufixos *-(ç)ão*, *-agem*, *-ade*, *-ite*, *-ose*, e o segmento *-e*. Aplicamos às pseudopalavras elaboradas estes segmentos finais a fim de testar se a associação dessas formas com o feminino no léxico estaria também ativa no fenômeno de atribuição de gênero a novas palavras pelo falante.

O experimento contava com 60 questões, 40 das quais eram questões distratoras, elaboradas com base na fonotática do português, mas terminando em segmentos que não ocorrem significativamente em palavras femininas, como os finais *-o*, *-l*, *-r* e *-s*. As questões eram dispostas em ordem aleatória para cada sujeito que iniciava o teste, a fim de neutralizar a influência do cansaço e da automatização nos resultados. Os participantes foram expostos a frases com as palavras-teste, cada uma das quais trazia uma pseudopalavra, precedida por uma lacuna que deveria ser preenchida com um dos dois artigos definidos do português, tornando explícito o gênero que cada sujeito atribuiu aos substantivos ao adaptá-los à morfologia da língua. Por exemplo, o experimento contou com a frase “Já sabemos muito sobre _ gutação”. Assumimos que os participantes preencheriam a lacuna com *<o>*, se processassem a palavra como substantivo masculino, ou *<a>*, se a entendessem como feminina. O questionário oferecia as duas alternativas na forma, das quais apenas uma poderia ser escolhida. O acesso às questões seguintes só era permitido uma vez respondidas as questões anteriores; desse modo, não eram possíveis questões com duas ou nenhuma das alternativas escolhidas.

O único requisito para participar do experimento era que os sujeitos fossem falantes nativos de português. O questionário iniciava com questões relativas às variáveis sociais *idade*, *sexo*, *escolaridade*, *cursos de graduação* (opcional/se aplicável) e *línguas faladas além do português* (opcional/se aplicável). O experimento foi elaborado e veiculado por meio da plataforma Google Forms e difundido através de redes sociais. Os dados correspondentes aos resultados foram convertidos em planilhas no Google Sheets e posteriormente processados através do programa SPSS Statistics.

5.1.2 Resultados

O teste foi encerrado contando com 89 respondentes. Quanto às variáveis sociais, 67% identificam-se com o gênero feminino, 32% com o masculino e 1%, um respondente, com outro gênero; em relação à idade, 80% tinham entre quinze e vinte e cinco anos, 15% entre vinte e seis e quarenta anos, e 5% tinham mais de quarenta anos; quanto à escolaridade, 82% possuíam Ensino Médio completo, 17% Ensino Superior completo e 1%, um participante, tinha Ensino Fundamental completo ou incompleto.

Quanto às variáveis sociais que podiam ou não ser aplicáveis aos respondentes, tentamos isolar estudantes de Letras ao amálgama dos respondentes que estavam cursando ou que eram formados em diversos outros cursos de graduação, por acreditarmos que possíveis conhecimentos técnicos e metalinguísticos dos integrantes daquele grupo pudessem ter influência sobre os resultados do teste. Nessa variável, 12% dos respondentes possuíam graduação ou estavam se graduando em Letras, e o restante, 88%, em outros cursos.

Em relação ao domínio de línguas para além do português, 39% dos participantes responderam que falavam também o inglês, 26% que falavam inglês e outra língua, 17% que dominavam uma língua adicional que não o inglês, 8% declararam falar três línguas além do português, e 10% responderam que não falavam línguas adicionais ou forneceram respostas incompletas à questão.

De modo geral, as terminações testadas mostraram-se condicionantes para a atribuição de feminino às pseudopalavras. Na Tabela 2, estão expostas as frequências com que o feminino foi atribuído, ou seja, com que o artigo definido feminino <a> foi escolhido como resposta a cada uma das questões relativas às vinte pseudopalavras testadas.

Tabela 2 – Atribuição de feminino às palavras-teste

Pseudopalavra	Atribuição de feminino	Segmento testado
jamência	97,7%	-a
nantarra	92%	-a
passina	95,5%	-a
ralufata	88,5%	-a

delitagem	100%	-agem
volinagem	96,6%	-agem
ravidagem	98,9%	-agem
celisão	93,2%	-ão
gutação	95,5%	-ão
arredição	97,7%	-ão
estrização	98,9%	-ão
bameridade	95,4%	-dade
lenvidade	97,7%	-dade
sagueidade	97,7%	-dade
troje	15,7%	-e
filate	19,3%	-e
dalome	21,6%	-e
debasite	69%	-ite
zemalite	75%	-ite
renutite	86,4%	-ite

Fonte: elaborada pelo autor.

As pseudopalavras representativas de cada segmento testado não apresentaram diferenças estatísticas relevantes dentro de cada grupo. Na Tabela 3, observam-se os dados referentes à frequência da atribuição do feminino às palavras com cada segmento, tomadas conjuntamente.

Tabela 3 – Atribuição de feminino às terminações testadas

Segmento	Atribuição de f.
-agem	98,50%
-dade	96,90%
-ão	96,32%
-a	93,42%
-ite	76,80%
-e	18,86%

Fonte: elaborada pelo autor.

As variáveis sociais controladas não pareceram interferir no gênero atribuído às palavras testadas, excluindo-se a possibilidade de diferentes níveis de

escolaridade, gênero, idade ou fluência em outras línguas desempenharem um papel na representação das formas linguísticas ativadas na gramática mental dos sujeitos durante o teste.

5.1.3 Discussão

Considerando-se os resultados do instrumento e os dados levantados nos *corpora*, que embasaram a elaboração das pseudopalavras, é possível assumir que o tratamento desses novos itens em relação ao gênero é baseado nas correlações existentes no inventário da língua receptora.

A atribuição de gênero gramatical a itens novos, conforme exposto anteriormente, pode ser condicionada por pelo menos dois conjuntos principais de variáveis: as de natureza semântica e as de natureza formal (CORBETT, 1991). Os fatores formais foram testados no experimento exposto acima, tendo se mostrado atuantes na escolha de gênero feminino para as pseudopalavras do teste que terminavam em *-agem*, *-dade*, *-ão*, *-a* e *-ite*, devido a associação que reflete correlações presentes no léxico entre o feminino e esses segmentos terminais. Notavelmente, o alto grau de consistência com que os informantes atribuíram o gênero feminino às palavras com terminação *-agem* revela uma forte associação desse segmento final com o gênero feminino em português. Uma das pseudopalavras com esse sufixo, *delitagem*, foi a única a ter o feminino a ela atribuído 100% das vezes pelos 89 respondentes.

A Tabela 3 mostra haver um ordenamento para os graus de associação entre os segmentos e o gênero feminino. Os três segmentos que mais foram entendidos como femininos também foram os itens mais coincidentes com o gênero feminino nos *corpora* consultados, com *-dade* sendo exclusivo ao feminino nesses dados, e *-agem* encontrando exceções em palavras comuns de dois gêneros, como *personagem*, que pode ser empregada tanto no feminino quanto no masculino: *uma personagem/um personagem*. Já para a terminação *-ão*, há ocorrência em algumas palavras masculinas uniformes e em palavras biformes, como *chão* e *charlatão*, respectivamente, o que pode estar por trás da menor associação desse segmento com o feminino, conforme sua posição inferior na classificação da Tabela 3. A menor quantidade de atribuição de feminino atestada no instrumento para as pseudopalavras

contendo *-ite*, por outro lado, pode estar relacionada à identidade desse sufixo com o segmento terminal de algumas palavras masculinas frequentes, como *convite* e *limite*.

Já às palavras com final em *-e*, o masculino foi atribuído em mais de 80% dos casos. Os dados dos *corpora* consultados mostram para uma divisão equilibrada entre as palavras femininas e masculinas com esse segmento, com um pouco mais de palavras femininas no dicionário: 49%, em relação às 45,5% masculinas. Esses dados, no entanto, não levam em consideração a presença dessa vogal final em sufixos. Quando excluímos do recorte as palavras terminadas em finais como *-dade*, *-ite* e *-ete* são excluídas da figura, chegamos a 28% das palavras femininas nos dados do dicionário; ao aplicar o mesmo tratamento analítico aos dados do *corpus* de base de uso, o percentual cai de 73% para 57%. Assim, a maior incidência da terminação *-e* em palavras masculinas quando essas não estão associadas a sufixos pode explicar o baixo grau de associação entre as palavras testadas que terminam nesse segmento e o gênero feminino.

Por fim, não se observou correspondência relevante entre as escolhas dos participantes para o gênero dos itens no teste e os dados das variáveis sociais analisadas.

5.2 TESTE PARA CRITÉRIOS SEMÂNTICOS DE ATRIBUIÇÃO DE GÊNERO

A fim de testar o papel dos critérios semânticos de atribuição de gênero para além do escopo dos empréstimos, formulamos um segundo experimento psicolinguístico, que simula episódios de atribuição de gênero a substantivos inéditos ao falante, induzindo pistas de natureza semântica potencialmente atuantes, conforme proposto por Thornton (2009) e atestado no recorte de empréstimos do inglês em PB.

Neste experimento, as pistas foram oferecidas nas questões através de definições, estipuladas para cada pseudopalavra e pensadas a fim de simularem os critérios de atribuição que devem estar atuantes no evento de tomada de conhecimento do falante a respeito tanto da nova palavra quanto de seu referente suposto. Esses critérios dizem respeito à associação do feminino gramatical com a propriedade semântica [feminino] – para mulheres ou fêmeas de animais –, interna à gramática da língua. Ainda, conforme proposto por Thornton (2009), consideramos três tipos de associação por analogia com itens já presentes no inventário do falante.

São casos como o da atribuição do feminino a *playlist*, por associação ao item já conhecido *lista*.

5.2.1 Metodologia

O experimento consistia em 24 questões, cada uma contendo uma das pseudopalavras formuladas. Dezesseis delas são questões-teste, com itens que terminam em segmentos não licenciados em português. As oito restantes são questões distratoras, com terminações regulares na língua. As questões trazem as pseudopalavras seguidas de breves definições dos conceitos estipulados.

As questões-teste dividem-se igualmente entre aquelas com itens de definições que condicionam a atribuição do gênero feminino (1) e de definições que condicionam a atribuição do masculino (2).

(1) Horb: um tipo de cadeira de balanço.

(2) Melateb: o mais novo dentre os irmãos do rei.

A fim de neutralizar interferências da forma fonológica no processo de atribuição de gênero, excluímos terminações grafêmicas nativas do português, uma vez que segmentos como *-a*, *-e*, *-l*, *-r*, *-s*, *-ão*, etc., podem estar, de alguma forma, associadas a valores de gênero. Para isso, optamos por aproximar suas formas àquelas de empréstimos do inglês, como *hightech* e *Facebook*, que, conforme exposto, não sofrem, de modo geral, influência da fonologia no fenômeno de atribuição de gênero – sendo, em geral, pronunciados com uma vogal epentética final [ɪ] em português brasileiro: [haj'tɛ.kɪ] e [fej.sɪ'bu.kɪ].

Foram escolhidos os grafemas finais ⟨*t*, *d*, *p*, *b*, *k*, *g*, *v*, *f*⟩, que representam, respectivamente, as consoantes /*t*, *d*, *p*, *b*, *k*, *g*, *v*, *f*/. Esses grafemas foram divididos aleatoriamente entre as palavras-teste, porém, de modo com que se repetissem, cada uma, duas vezes no inventário de pseudopalavras. Já aos itens das questões distratoras, foram atribuídos, também aleatoriamente, os segmentos finais ⟨*o*, *a*, *l*, *r*, *s*⟩, referentes a finais fonológicos licenciados no português.

O texto de cada questão, ou seja, a definição dos conceitos criados para o experimento, que induziria um ou outro gênero para cada item, foi formulado considerando-se os critérios semânticos mencionados. Os critérios de analogia com o inventário propostos por Thornton (2009) foram representados por dez das

dezesseis questões-teste. As seis restantes traziam definições que apontavam para entidades sexuadas, a fim de testar o critério que atribui o gênero gramatical feminino a palavras que designam mulheres e fêmeas de animais.

O modo específico como distribuimos esses critérios dentre as questões – mais questões de analogia com itens do inventário do que de correlação com traços do referente – deve-se ao fato de os critérios de analogia semântica terem sido pensados nos termos de Thornton (2009). A autora subdivide os critérios semânticos nas três categorias menores, *analogia com conceito associado*, *analogia com hiperônimo* e *analogia com equivalente*. A fim de que os três tipos de critérios analógicos fossem representados de maneira satisfatória, dedicamos um maior número de questões a essa categoria.

As pseudopalavras que simulam os critérios de atribuição por sexo/gênero são definidas com referentes animados, marcados para um desses traços semânticos, ou seja, fêmeas de animais ou humanos do sexo/gênero social feminino, nas questões femininas, como em (3) a.; e referentes com o traço semântico masculino para atribuição do gênero gramatical masculino, como em (3) b..

(3) Exemplos de pseudopalavras com referente sexuado

a. *Sark: a fêmea do rinoceronte.*

b. *Klag: o líder espiritual de certas tribos nômades do Polo Sul.*

A distribuição escolhida em relação aos critérios a serem testados nas pseudopalavras, relegando menos questões ao critério de analogia com hiperônimo, deve-se ao fato de vermos a atribuição por analogia por conceito associado e analogia com hiperônimo como potencialmente idênticas no processo não artificial de atribuição de gênero a novos itens, como na atribuição de gênero a empréstimos. Mesmo indo parcialmente de encontro à visão de Thornton (2009) nesse aspecto, estipulamos questões em que tentamos simular a associação por hiperônimo, a fim de testar o modelo no que diz respeito ao índice de atribuição em ambiente artificial.

As definições das questões que testam o critério de *analogia por conceito associado* utilizam como definições sintagmas cujos núcleos apresentam o gênero sugerido, como vemos em (4) a. para masculino e (4) b. para feminino. Já as questões relativas à *analogia por hiperônimo* trazem definições iniciando em “tipo de X”, em que X é o hiperônimo da pseudopalavra. Em (5) a., o hiperônimo na definição é “cadeira” cujo gênero feminino assumimos que seria associado a *horb*. Finalmente, para

simularmos a *atribuição por equivalente*, formulamos definições que remetessem a um item específico a ser associado às pseudopalavras, como em (6) a.-b.

(4) Exemplos de pseudopalavras com analogia por conceito associado

a. *Beltav*: sapato tradicional do Nepal.

b. *Zarf*: ponte levadiça que dá acesso a castelos ilhados.

(5) Exemplos de pseudopalavras com analogia por hiperônimo

a. *Horb*: tipo de cadeira de balanço.

b. *Waltröst*: tipo de chapéu usado por trabalhadores rurais da Escócia.

(6) Exemplos de pseudopalavras com atribuição por equivalente

a. *Lusp*: objeto usado para varrer o chão.

b. *Gurd*: livro sagrado do Islamismo.

O experimento foi incorporado na plataforma Google Forms e difundido *online*, por contatos pessoais e em publicações em redes sociais. Antes de chegar nas questões descritas acima, que compõem o experimento linguístico em si, os participantes, presumidamente falantes nativos de português, deveriam responder questões sobre sua estratificação social, como gênero, idade e grau de escolaridade, e, quando aplicáveis, sobre as línguas faladas além do português, os cursos de Ensino Superior e a experiência no exterior (onde e por quanto tempo viveu).

5.2.2 Resultados

Em relação às variáveis sociais, os 152 respondentes dividem-se do seguinte modo: 70% se identificaram com o gênero feminino, 29% com o masculino e 1%, dois indivíduos, com outro gênero; 53% têm idades entre quinze e vinte e cinco anos, 35% têm de vinte e seis a quarenta anos, e 23% têm mais de quarenta anos. Quanto ao nível de escolaridade, 49% dos participantes possuem Ensino Médio completo; 48% possuem Ensino Superior completo, e 3% dividem-se entre aqueles com Ensino Fundamental completo ou incompleto. No que diz respeito a seus cursos de graduação, os dados aqui entendidos como potencialmente relevantes, conforme justificado em 5.1.1, mostraram que 18% dos indivíduos eram graduados ou estavam se graduando em Letras, e, 82%, em outros cursos.

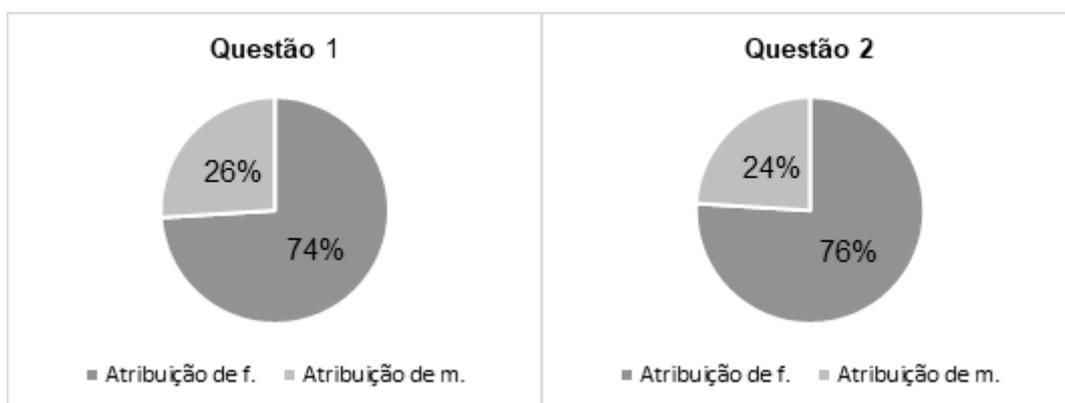
Em relação às línguas adicionais faladas pelos participantes, 27%, não fala outras línguas além do português, 33% falam português e inglês, 23% falam também

uma terceira língua, 6% falam duas línguas além de português e do inglês, e os outros 10% falam, além do português, outras línguas excluindo-se o inglês, como espanhol e árabe, e espanhol e alemão.

Finalmente, na questão que solicitava informações sobre a experiência dos respondentes vivida no exterior, também operamos alguns amálgamas para melhor analisar os grupos. Com base nas respostas, além de considerarmos a quantidade de tempo passada nos países estrangeiros (mais, ou menos, que um ano), separamos os países entre anglófonos (Inglaterra, Estados Unidos, Irlanda), hispanófonos (Espanha, Argentina, Chile) e “germanófonos” (Dinamarca e Alemanha). Consideramos importantes essas distinções pela semelhança gramatical do espanhol com o português, o que poderia impactar os resultados diferentemente do conhecimento de indivíduos que viveram em países onde se falam línguas germânicas. Para este grupo, ainda, a divisão entre países que falam inglês e países que falam alemão ou dinamarquês foi considerada, já que, primeiro, nosso experimento foi formulado com base nos padrões fonotáticos e no *corpus* de empréstimos do inglês, e, segundo, essas outras línguas germânicas possuem também um sistema de gênero nominal, diferentemente do inglês, não restrito aos pronomes. Assim, para além dos 82% que nunca viveram fora do Brasil: 2% viveram mais de um ano e 3% menos de um ano em país anglófono; 3% viveram mais de um ano e 1% viveu menos de um ano em país hispanófono; 3% viveram mais e 1% menos de um ano em países classificados como germanófonos; 1%, dois respondentes viveram em Portugal por menos de um ano; 1% viveu em diversos países; e 3% dos respondentes deram respostas incompletas, respondendo ou apenas o lugar, ou apenas a duração do período vivido no exterior.

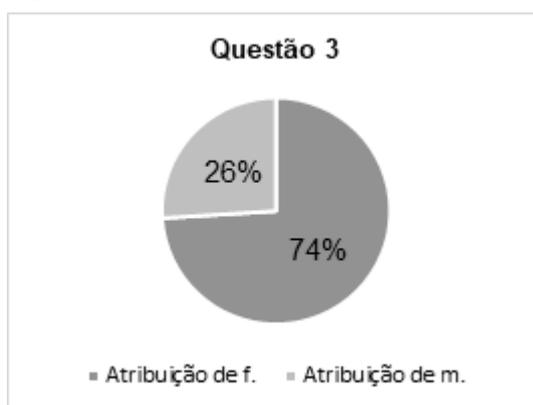
Quanto às respostas às questões que consideram a analogia com itens já presentes no inventário, aquelas que induziam analogia com um conceito associado tiveram 74% e 76% de respostas femininas (Figura 4) – por questão de praticidade, chamaremos as respostas que cumpriram nossa expectativa em relação ao gênero atribuído de “acertos”. A questão que induzia analogia com um hiperônimo feminino obteve 74% de acertos (Figura 5). As questões com analogia por conceito equivalente tiveram 74% e 76% de respostas femininas (Figura 6). As três questões que definiam entidades animadas femininas tiveram, duas delas, 90% de acertos e, uma, 97% (Figura 7).

Figura 4 – Questões com item associado feminino



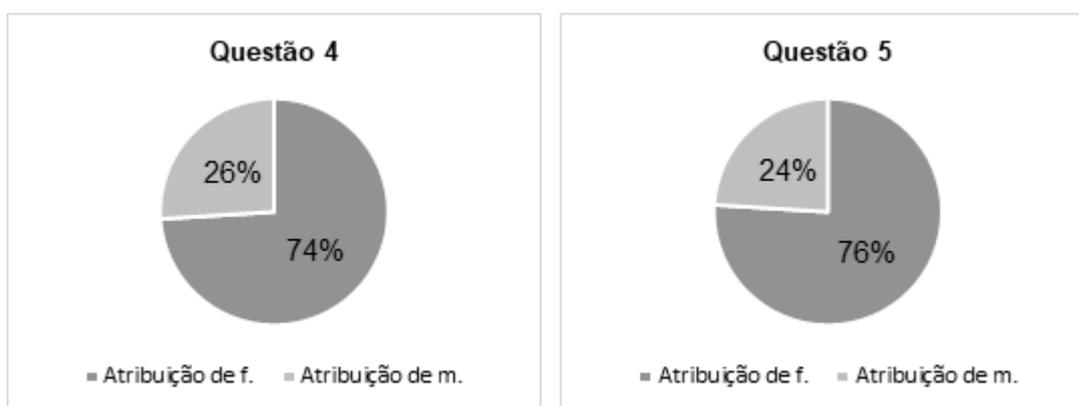
Fonte: elaborado pelo autor.

Figura 5 – Questão com hiperônimo feminino



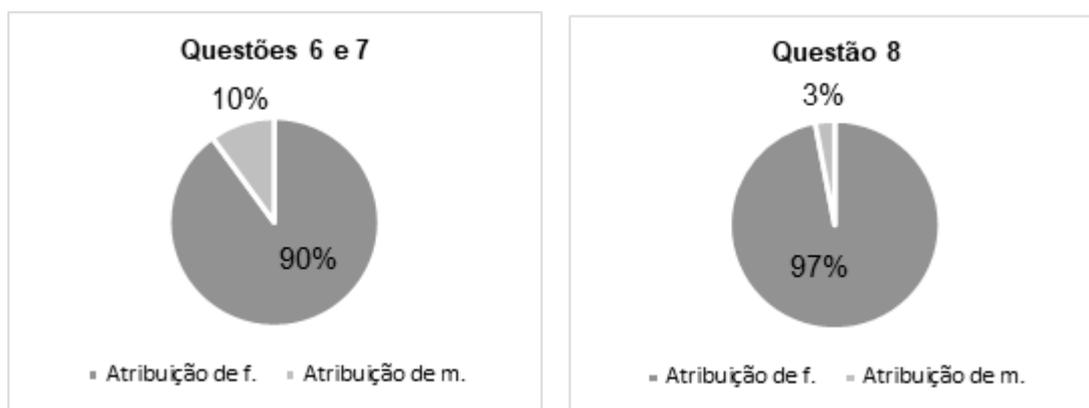
Fonte: elaborado pelo autor.

Figura 6 – Questões com item equivalente feminino



Fonte: elaborado pelo autor.

Figura 7 – Questões com referentes semanticamente femininos



Fonte: elaborado pelo autor.

Das questões que induziam o gênero masculino, as questões por conceito associado apresentaram 48% e 83% de respostas com pseudopalavras masculinas. A questão com analogia com hiperônimo teve 85,5% de acertos. As duas questões que induziam a resposta apontando para um conceito equivalente tiveram 84% e 78% das respostas masculinas. E, finalmente, as três questões que descreviam seres animados no gênero masculino tiveram 90%, 85% e 91% de respostas masculinas.

5.2.3 Discussão

Concentramos nossa discussão nas questões desenhadas para induzir o gênero feminino através de um dos quatro critérios semânticos tomados como base no estudo. Isso se deve ao fato de enxergarmos o feminino como o gênero marcado da língua. No que diz respeito ao fenômeno de atribuição de gênero gramatical a itens inéditos a um dado falante, apenas a atribuição do gênero feminino precisa ser explicada, já que, quando o masculino é atribuído, embora os mesmos critérios possam ter atuado, há o critério que atribui esse gênero, não marcado, na ausência de demais índices, ou da pouca saliência desses para a marcação.

O que mais se destacou em relação aos resultados foram as respostas às questões que denotavam referente sexuado. Podemos ver na Figura 4, em comparação às Figuras 1-3, a significativa inclinação dos sujeitos à atribuição de feminino para itens associados a entidades animadas femininas. A força dos critérios de atribuição de gênero semântico baseado em associações internas ao sistema linguístico – o que em português pode ser representado por *f.* - feminino, *m.* - masculino

– vai ao encontro das ideias de Corbett (1991) no que concerne critérios semânticos para as línguas em geral (Ibid., p. 14). Esse pressuposto está também atestado nas conclusões da análise de Poplack, Pousada e Sankoff (1982) sobre os empréstimos do inglês no francês e no espanhol, segundo as quais o único critério significativamente relevante para a atribuição de gênero gramatical foi o denominado “sexo fisiológico” (*physiological sex*)¹⁰. É o que se observa também no estudo de Barkin (1980), que, ao opor critérios formais a semânticos, como os que relacionam feminino gramatical a feminino semântico, aponta para uma maior regularidade no condicionamento deste tipo.

Enquanto nas três questões com pseudopalavras que descreviam animais ou humanos do sexo/gênero feminino tivemos 90%, 90% e 96% de atribuições de feminino, nas três questões desse tipo que induziam o gênero masculino, tivemos 91%, 90% e 48% de atribuição do masculino. Entendemos que o último valor, visivelmente contrastante, seja fruto de uma falha no *design* da questão, que acabou por gerar ambiguidade com uma palavra feminina não sexuada em português (*guarda*, comum de dois que significa ‘pessoa que tem como função guardar alguém ou alguma coisa’, e *guarda*, substantivo feminino abstrato, que significa ‘o ato de guardar ou de ter sob guarda’). Ainda sobre as questões de gênero feminino, pudemos observar uma quantidade maior de acertos (97%) para a alternativa que designava uma entidade humana (Figura 7) se comparada às duas questões que designavam fêmeas de animais, Questões 6 e 7 (ambas com 90%). É possível que haja maior margem para a não marcação do gênero semântico na morfologia quando se trata de fêmeas de animais e palavras masculinas, ou ainda de machos e palavras do gênero feminino. Isso se explica se considerarmos a gramática do português, onde não são raros os casos de epicenos, como “rinoceronte fêmea” ou “girafa macho”. Por outro lado, são raros itens como “criança” e “testemunha”, que não marcam o sexo do referente na morfologia. Embora não tenhamos controlado esse traço semântico nas questões, a observação merece uma análise criteriosa, e pretendemos considerá-la em experimentos e estudos futuros.

As quatro questões de analogia com conceitos previamente conhecidos pelos sujeitos, que descreviam os objetos através de palavras com os valores de gênero que esperávamos que fossem escolhidos, tiveram 74,3% e 76,3% de acertos nas

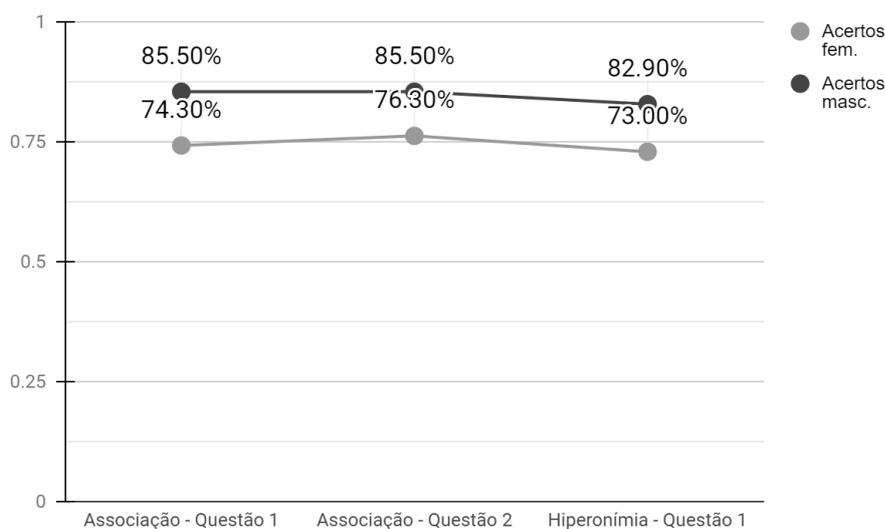
¹⁰ Conforme explicação do termo em nota de rodapé, na seção 3.1.

questões que induziam o feminino, e 85,5% e 85,5%, nas relativas ao masculino. Nota-se que, por mais que a tendência seja forte de ambos os lados, a atribuição do masculino foi mais recorrente em itens desse tipo. Quanto às questões que definiam um hiperônimo para a pseudopalavra ("*tipo de cadeira[...]*", "*tipo de sapato[...]*"), tivemos 73% de acertos para o feminino, e 82,9% de masculino, acompanhando a tendência das questões por associação. Para as questões de *associação com conceito equivalente*, houve volume bastante baixo de acertos (Figura 6), o que se deve a uma falha no *design* adotado. Parece-nos que a descrição, mesmo tendo sido escrita da forma mais direta possível – por exemplo, uma vassoura como "*objeto usado para varrer o chão*" – não foi suficiente para ativar este item na memória do falante-sujeito. Em experimento futuro, talvez uma alternativa mais eficiente seria utilizarmos imagens ou sons para tal.

Considerando que as questões de *equivalente* apresentaram problemas de *design*, tomemos as duas questões de *associação* e a questão de *hiperonímia*, que somam três questões para cada valor de gênero, como representativas dos critérios de analogia com itens do inventário.

Ainda que, de modo geral, as seis questões tenham atingido um índice alto de acertos em relação ao gênero esperado, é possível observar na Figura 8 uma maior segurança dos falantes quanto à atribuição do masculino para os três casos. Essa tendência pode estar relacionada com a maior robustez do gênero masculino, que seria atribuído tanto nos casos que itens masculinos análogos determinaram a atribuição, quanto para o restante dos casos, onde não houve índices para atribuição e o masculino foi atribuído por não marcação.

Figura 8 – Comparação de acertos masc./fem. em questões de analogia com inventário



Fonte: elaborado pelo autor.

Finalmente, as variáveis sociais, de modo geral, não mostraram ter influência nas escolhas dos sujeitos durante a atribuição de gênero aos itens. Ainda assim, nas questões que forçaram a atribuição com itens associados já presentes no inventário, sujeitos mais jovens apresentaram menor tendência a atribuir o gênero esperado às pseudopalavras. O mesmo é verdade para sujeitos com inglês como língua adicional. É possível que falantes nesses recortes estejam menos propensos a atribuir ao novo item o gênero feminino de um item associado já conhecido.

Essa tendência pode estar relacionada a um maior conhecimento da gramática do inglês, onde não há marcação de gênero para substantivos, sendo replicado esse comportamento de não marcação mesmo quando um item feminino análogo é ativado. Assim, falantes com menos conhecimento, menos tempo de contato, ou contato desde uma menor idade com o inglês teriam menor propensão a adaptar itens nominais estranhos à morfologia do português, marcando-os para a categoria gênero.

É interessante notar que, em relação aos itens com referente sexuado, a alta tendência de marcação de gênero para as palavras que denotavam entidades femininas se distribuiu uniformemente dentre os grupos sociais. Além de, como afirmado na literatura, se tratar de critérios mais robustos para a atribuição de gênero, consideremos novamente o conhecimento sobre a gramática do inglês, que pode ter influenciado as questões de analogia com o inventário. Nesta língua, há marcação de gênero pronominal categórica quando referentes sexuados são recuperados por

pronomes, o que justificaria este comportamento mesmo estando as duas gramáticas atuantes na atribuição. Ou seja, limitada ao sistema pronominal, a gramática do inglês marca morfologicamente o gênero semântico de sexuadas – através dos pronomes pessoais de terceira pessoa *he* e *she* e seus correlatos –, com base de sexo biológico ou gênero social do referente. O mesmo não é verdade para entidades não sexuadas, que seriam designadas pelo pronome neutro de terceira pessoa *it* e seus correlatos.

6 CONSIDERAÇÕES FINAIS

Neste trabalho abordamos o gênero gramatical em português e seu envolvimento no processo de adaptação morfológica de novos itens na língua. Conforme sugerido por autores como Corbett (1991) e Poplack, Pousada e Sankoff (1982), a multiplicidade de fatores linguísticos envolvidos no processo, somado à consideração de fatores extralinguísticos envolvidos na adoção e difusão de cada item em cada evento de uso, tornam o detalhamento do processo de difícil apreensão. No entanto, ao analisarmos casos em que processos naturais geram contexto para o fenômeno, bem como casos em que tal contexto foi gerado artificialmente – o caso dos empréstimos e dos experimentos com pseudopalavras, respectivamente – foi possível constatar a atuação de critérios de natureza formal e semântica, relacionados com a produtividade de gênero da língua e a distribuição das correlações relacionadas no léxico. Assim, tanto a vogal átona final *-a*, atrelada ao feminino em nível morfofonológico, quanto segmentos como *-agem* e *-dade*, correlacionados estatisticamente ao feminino no léxico, parecem relevantes na atribuição de gênero dada sua presença em itens novos a serem processados. Por outro lado, notamos também que a ativação da analogia semântica com itens já presentes nos repertórios lexicais dos falantes pode projetar o gênero destes itens às novas palavras, conforme descrito pelos critérios de atribuição de Thornton (2009).

Futuras investigações que venham a contribuir para o entendimento desses processos incluem sua observação em outros recortes de empréstimos no léxico do português – como nos itens do tupi-guarani e do grego –, novos experimentos linguísticos que testem os critérios formais e semânticos simultaneamente ou em um maior número de contextos, bem como a análise do fenômeno na aquisição linguística de línguas de gênero, tanto por crianças quanto por aprendizes do português como língua adicional, por exemplo.

REFERÊNCIAS

- ALCÂNTARA, C. As classes formais do português brasileiro. **Letras de Hoje**, Porto Alegre, v. 45, n. 1, p. 5–15, mai. 2010.
- ARMELIN, P. Classifying nominals in Brazilian Portuguese: a unified account for gender and inflectional class. In: VESELOVSKÁ, L.; JANEBOVÁ, M. (orgs.). **Complex visibles out there**: Proceedings of the Olomouc Linguistics Colloquium 2014: Language Use and Linguistic Structure. Olomouc: Palacký University, 2014.
- AUDRING, J. Gender as a complex feature. In: PACIARONI, T.; THORNTON, A. M.; LOPORCARO, M. (eds.). **Language Sciences**, v. 43, p. 5–17, 2014.
- BARKIN, F. The role of loanword assimilation in gender assignment. **Bilingual Review / La Revista Bilíngüe**, v. 7, n. 2, p. 105–112, 1980.
- CÂMARA JR., J. M. **Estrutura da língua portuguesa**. 42. ed. Petrópolis, RJ: Vozes, 1970.
- CANO, W. M.; PRADO, D. de F. Os estrangeirismos da área da informática no Aurélio XXI. **Revista Alfa**, São Paulo, v. 50, n. 2, p. 265–275, 2006.
- CARVALHO, N. M. **Empréstimos lingüísticos**. São Paulo: Editora Ática, 1986.
- COMRIE, B. Grammatical Gender Systems: A Linguist's Assessment. **Journal of Psycholinguistic Research**, v. 28, p. 457–466, 1999.
- CORBETT, G. **Gender**. Cambridge: Cambridge University Press, 1991.
- CORRIENTE, F. **Dictionary of Arabic and Allied Loanwords**. Boston: Brill, 2008.
- HALLE, M; MARANTZ, A. Distributed Morphology and the Pieces of Inflection. In: HALE, K. L.; KEYSER, S. J. (eds.) **The View from Building 20**: Essays in Linguistics in Honor of Sylvain Bromberger. Cambridge: MIT Press, 1993. p. 111–176.
- _____. Some Key Features of Distributed Morphology. **Papers on Phonology and Morphology**, Cambridge, v. 21, p. 275–288, 1994.
- HARRIS, J. W. The exponence of gender in Spanish. **Linguistic Inquiry**, v. 22, n. 1, p. 27–62, 1991.
- HASPELMATH, M. Lexical borrowing: Concepts and issues. In: HASPELMATH, M.; TADMOR, U (Eds.). **Loanwords in the world's languages: a comparative handbook**. Berlin: Mouton de Gruyter, 2009. p. 35–54.
- HAUGEN, E. The analysis of linguistic borrowing. **Language**, v. 26, n. 2, p. 210–231, 1950.

_____. (Ed.). **Bilingualism in the Americas: a bibliography and research guide**. Tuscaloosa: American Dialect Society, 1956.

KILARSKI, G.; KRYNICKI, G. Not arbitrary, not regular: the magic of gender assignment. In: DELBECQUE, N.; AUWERA, J.; GEERAERTS, D. (Eds.). **Perspectives on Variation**. Berlin: Mouton de Gruyter, 2005. p. 235–250.

ORSI, V. A presença de empréstimos da língua inglesa na revista brasileira Glamour. In: COLÓQUIO DE MODA - 8ª EDIÇÃO INTERNACIONAL, 11., 2015, Curitiba. **Anais eletrônicos** [...]. Curitiba: 2015. Disponível em: <http://www.coloquiomoda.com.br/anais/Coloquio%20de%20Moda%20-%202015/COMUNICACAO-ORAL/CO-EIXO4-COMUNICACAO/CO-4-A-PRESENCA-DE-EMPRESTIMOS-DA-LINGUA-INGLESA-NA.pdf>. Acesso em: 17 jun. 2018

POPLACK, S. A variationist perspective on language contact. In: ADAMOU, E.; MATRAS, Y. (Eds.) **Routledge Handbook of Language Contact**. New York/London: Routledge, 2020. p. 46–62.

POPLACK, S.; POUSADA, A.; SANKOFF, D. Competing influences on gender assignment: variable process, stable outcome. **Lingua**, Amsterdam, v. 57, p. 1–28, 1982.

PRIBERAM. Dicionário Priberam da Língua Portuguesa, 2020. Disponível em: <https://dicionario.priberam.org/>. Acesso em: 8 out. 2020.

ROCHA, L. C. **Estruturas morfológicas do português**. Belo Horizonte: Ed. da UFMG, 1998.

SCHWINDT, L. C. Zeros na morfologia nominal portuguesa à luz da Optimal Interleaving Theory. **Revista virtual de estudos da linguagem – ReVEL** [Edição especial n. 5], 2011.

_____. Exponência de gênero e classe temática em português brasileiro. **Revista DELTA**, São Paulo, v. 34, n. 2, p. 745–768, 2018.

THORNTON, A. M. Constraining gender assignment rules. **Language Sciences**, v. 31, n. 1, p. 14–32, 2009.