

LÉXICOS PARA ALÉM DO LEXICALISMO: NÃO TENDE FAZER ANÁLISE MORFOFONOLÓGICA NA PRIVACIDADE DE SUA GRAMÁTICA

LÉXICOS MÁS ALLÁ DEL LEXICALISMO: NO INTENTES REALIZAR ANÁLISIS MORFOFONOLÓGICOS EN LA INTIMIDAD DE TU GRAMÁTICA.

LEXICONS BEYOND LEXICALISM: DON'T TRY MORPHOPHONOLOGICAL ANALYSIS IN THE PRIVACY OF YOUR GRAMMAR

Luiz Carlos Schwindt
Universidade Federal do Rio Grande do Sul
schwindt@ufrgs.br
0000-0003-0533-589X

Resumo

Neste texto, aborda-se a polissemia do termo léxico nos estudos linguísticos, problematizando-se a distinção entre léxico e gramática na análise de fenômenos morfofonológicos. Argumenta-se que a compreensão do conhecimento morfofonológico passa pela descrição de pelo menos dois tipos de léxico: o *léxico em uso* (as formas do léxico permanente efetivamente empregadas pelos falantes) e o *léxico potencial* (palavras possíveis segundo os padrões da língua). A partir de estudos empíricos sobre harmonia vocálica metafônica, plurais de nomes fechados por ditongo nasal e marcação de gênero gramatical em português brasileiro, procura-se demonstrar que a melhor análise formal deve considerar o contraste entre esses *léxicos*. Discute-se, ainda, o lugar do léxico na arquitetura da gramática gerativa, em particular na Fonologia e Morfologia Lexical, na Teoria da Otimidade e na Morfologia Distribuída, defendendo-se a necessidade de computação gramatical para dar conta da interface morfologia–fonologia.

Palavras-chave: Léxico; morfofonologia; léxico potencial; léxico em uso; arquitetura da gramática.

Resúmen

En este texto se aborda la polissemia del término léxico en los estudios lingüísticos, problematizando la distinción entre léxico y gramática en el análisis de fenómenos morfofonológicos. Se argumenta que la comprensión del conocimiento morfofonológico pasa por la descripción de al menos dos tipos de

léxico: el léxico en uso (las formas del léxico permanente efectivamente empleadas por los hablantes) y el léxico potencial (palabras posibles según los patrones de la lengua). A partir de estudios empíricos sobre armonía vocálica metafónica, plurales de nombres terminados en diptongo nasal y la marcación de género gramatical en portugués brasileño, se procura demostrar que el mejor análisis formal debe considerar el contraste entre estos *léxicos*. Se discute, además, el lugar del léxico en la arquitectura de la gramática generativa, en particular en la Fonología y Morfología Léxica, en la Teoría de la Optimidad y en la Morfología Distribuida, defendiéndose la necesidad de computación gramatical para dar cuenta de la interfaz morfología–fonología.

Palabras clave: Léxico; morfofonología; léxico potencial; léxico en uso; arquitectura de la gramática.

Abstract

This text addresses the polysemy of the term lexicon in linguistic studies, problematizing the distinction between lexicon and grammar in the analysis of morphophonological phenomena. It is argued that understanding morphophonological knowledge requires describing at least two types of lexicon: the lexicon in use (forms from the permanent lexicon actually employed by speakers) and the potential lexicon (possible words according to language patterns). Based on empirical studies of metaphonic vowel harmony, plurals of nouns ending in nasal diphthongs, and grammatical gender marking in Brazilian Portuguese, it seeks to demonstrate that the best formal analysis must consider the contrast between these *lexicons*. The place of the lexicon in the architecture of generative grammar is also discussed, particularly in Lexical Phonology and Morphology, Optimality Theory, and Distributed Morphology, defending the necessity of grammatical computation to account for the morphology–phonology interface.

Keywords: Lexicon; morphophonology; potential lexicon; lexicon in use; architecture of grammar.

Recebido: 15/11/2025

Aceito: 3/2/2026

1. Contextualização

Neste texto abordo a polissemia do termo *léxico* em estudos linguísticos e problematizo a distinção entre léxico e gramática na análise de fenômenos morfofonológicos.¹ O título, que parafraseia parcialmente o título do artigo clássico de Marantz (1997), *No escape from syntax: don't try morphological*

¹ Este trabalho resume as principais ideias de duas conferências que ministrei, respectivamente, em dezembro de 2023, no XV Colóquio de Estudos Lexicais, na USP, e em junho de 2025, no ALFALito 2025: Morfologia, na UFMG.

analysis in the privacy of your own lexicon, contém uma falsa provocação. Falsa, porque minha abordagem não se funda em uma contradição importante ao que é mais caro à ideia do referido artigo, já que definitivamente o autor não sugere que a sintaxe seja o lugar da exponenciação fonológica de morfemas no modelo que adota, e de que é coproponente, a Morfologia Distribuída (Halle e Marantz, 1993). Divirjo, isso sim, quanto à abrangência da leitura do que se pode chamar de *léxico* em modelos formais e quanto à convicção de que morfofonologia resulta de computação gramatical.

O termo *léxico* (e termos dele derivados, como *lexical*, *lexicalizado* etc) vem sendo empregado para designar diferentes ideias em Linguística, como sugerem Harley e Noyer (2003).

- (i) Lexical(izado) = idiomatizado.
- (ii) Lexical(izado) = não construído pela sintaxe.
- (iii) Lexical(izado) = não sujeito a processos fonológicos que se aplicam sem exceções (processos lexicais vs. pós-lexicais, de acordo com o modelo da Fonologia e Morfologia Lexical). (Adaptado de Harley e Noyer, 2003: 467)

Essa polissemia, por decorrência, vem alimentando o debate sobre os limites do que se entende efetivamente por gramática, no sentido da pergunta clássica do gerativismo, qual seja, *o que sabemos quando sabemos uma língua?*

Entre as acepções de *léxico* sugeridas por Harley e Noyer (2003), os sentidos (i) e (iii) tocam mais diretamente à morfofonologia. Isso porque, quando assumimos que uma forma linguística é idiomatizada, negamos sua transparência, sua analisabilidade. Em outras palavras, a ideia de alomorfia fonologicamente condicionada cede espaço nesse caso à listagem de uma forma, a rigor memorizada. Também quando dizemos, no escopo de teorias como a Fonologia e Morfologia Lexical (FML), proposta por Kiparsky (1982) e Mohanan (1982), que um processo fonológico é lexical, ou que pertence a um determinado nível do léxico, estamos afirmando que está em plena interação com processos fonológicos, ao contrário dos processos pós-lexicais, que são puramente fonológicos ou fonéticos.

Assim, podemos nos perguntar sobre que mecanismos estão em jogo quando um falante nativo pratica os juízos que estão em (1), em que o asterisco sugere malformação.

- (1) sigo / *sego
- mãos / *mões
- a estudante / *a estudanta
- cafezinho / *cafeinho
- amamentação / *amamentamento
- solista / *solzista
- balonismo / *balóismo

Para os defensores da Hipótese Lexicalista, que a tradição atribui às ideias de Chomsky (1970), esses entre outros juízos sobre a boa formação de palavras são possíveis graças à *competência lexical*. Fato é que mesmo teorias que assumem ponto de vista menos específico do que esse, se pensadas no escopo do formalismo, parecem todas compartilhar o entendimento de que falantes conhecem as palavras de sua língua para além de memorizá-las.

Onde pretendo chegar mais precisamente aqui é na tese de que, no domínio do mapeamento entre unidades portadoras de significado e unidades sonoras, a distinção entre léxico e gramática é menos clara do que muitos supõem. E uma das principais razões para isso está no fato de que nosso conhecimento opera com pelo menos dois tipos de léxico, o *léxico em uso* e o *léxico potencial*.

São três, assim, os objetivos deste estudo:

- (i) discutir diferentes concepções de léxico em Linguística;
- (ii) analisar fenômenos que se classificam como morfofonológicos e seu comportamento no que se define como léxico em uso e léxico potencial;
- (iii) problematizar os limites do léxico numa arquitetura de gramática apta a lidar com fenômenos morfofonológicos.

Para dar conta desses objetivos, o texto que segue esta contextualização está organizado em três seções nucleares: na seção 2 abordo diferentes concepções de léxico e proponho meu entendimento sobre o que rotulo como léxico em uso e léxico potencial; na seção 3, a partir de uma proposta de definição para morfofonologia, recupero estudos que realizei sob a perspectiva dos dois tipos de léxico em foco; na seção 4 problematizo a relação de léxico e morfofonologia em modelos de base gerativa. Na seção 5, a título de palavras finais, apresento uma breve reflexão sobre a incontornável demanda por abstrações nos estudos envolvendo a interface morfologia–fonologia.

2. Léxicos

O estruturalismo saussuriano, mesmo sem apresentar uma definição alternativa explícita à tradição para o termo *léxico*, inova ao concebê-lo como mais do que uma lista de palavras (ou de signos, nos termos de sua teoria).

(...) Em outro ângulo, Saussure lida com a morfologia do léxico em termos do mecanismo de analogia, interpretada com a quarta proporcional. Diz ele, por exemplo, que 'carteiro' não vem de 'carta', mas de um modelo do tipo 'prisão::'prisioneiro'. Em referência a questões lexicais, diz Saussure que a analogia supõe um modelo e sua imitação regular. Uma forma analógica é feita à imagem de outra ou outras, segundo uma regra determinada (...); supõe a consciência e a compreensão da relação que une as formas entre si (SAUSSURE, 1916: 187). Assim, o mecanismo de analogia está na esfera da langue. Acrescenta ele, porém, que é necessário distinguir a compreensão da relação que une as formas geradoras entre si do resultado sugerido pela comparação, que corresponde à forma construída pelo falante, já que apenas esta última pertence à fala. Assim, conclui, a analogia nos ensina a separar a língua da fala, colocando esta como dependendo daquela. (Basilio, 2021: 161)

Ao abordar o mecanismo da analogia no âmbito da dicotomia língua/fala, Saussure (1916) antevê em certo sentido a ideia de formação de palavras, que se sedimentaria anos mais tarde com a morfologia gerativa. Para o autor, embora o léxico, enquanto lista de signos herdada, possa ser entendido como uma espécie de cristalização da fala, apresenta uma face sistemática, própria da língua, que se expressa pelo mecanismo da analogia. Garante-se, pela ideia de forma derivada como apenas relativamente arbitrária (em oposição à primitiva, que é completamente arbitrária), que signos sejam vistos relacionalmente no sistema também em função de sua estrutura interna, de suas partes

constitutivas (ex. [[metal]eiro]). Como afirma Basilio (2021), essa aparente contradição do *Cours*, contudo, salvaguarda que se achem no léxico tanto as formações previstas quanto as imprevistas (que tratamos como potenciais neste texto).

Não fixados em analogia, um conceito que depende de se assumirem relações paradigmáticas na descrição linguística, os estruturalistas norte-americanos aprofundaram como nenhuma outra corrente linguística a abordagem de constituição interna de palavras. Nesse sentido, léxico não é entendido exatamente como lista de *vocábulos*, mas como lista de *morfemas*.

Since every complex form is made up entirely of morphemes, a complete list of morphemes would account for all the phonetic forms of a language. The total stock of morphemes in a language is its lexicon. However, if we knew the lexicon of a language, and had a reasonably accurate knowledge of each sememe, we might still fail to understand the forms of this language. (Bloomfield, 1933: 162)

O estruturalismo norte-americano, também conhecido, nos termos de Hockett (1954), como abordagem *Item e Arranjo*, por privilegiar sistemas aglutinantes (em termos de tipologia morfológica), encontrou obstáculos para lidar com línguas em cuja morfologia não se observam apenas nem principalmente operações concatenativas. Processos como reduplicação (ex. malgaxe, *hafa* 'diferente' > *hafa-hafa* 'ligeiramente diferente'; sundanês, *guyon* 'gracejar' > *guguyon* 'gracejar repetidamente') ou infixação de vogais em raízes consonantais (ex. árabe egípcio, *ktb* 'raiz de escrever' > *kitab* 'livro', *katib* 'escritor'; hebraico, *gdl* 'raiz de amplitude, grandeza, crescimento' > *gadol* 'grande', *gadal* 'ele cresceu') desafiam o entendimento de léxico como lista de morfemas. Isso porque não é possível listar o reduplicante, já que sua forma depende da base de que é copiado, assim como não é possível prever fonotaticamente as vogais que serão inseridas em raízes de consoantes também idiossincráticas.

A consequência natural desse impasse estruturalista, que achará espaço na abordagem *Item e Processo*, é o entendimento de que à lista de itens (morfemas) presente no léxico se associam representações abstratas, como moldes (templates) sujeitos a operações (regras ou restrições) capazes de assegurar a boa formação das palavras. Esse entendimento se fundamenta na ideia de competência linguística, que nos primeiros modelos gerativos era noção bastante restrita à sintaxe. Com a hipótese lexicalista, como vimos, passa-se a conceber uma especialização dessa competência, a competência lexical, que inclui, nos termos de Basilio (1980):

a) o conhecimento de uma lista de entradas lexicais; b) o conhecimento da estrutura interna dos itens lexicais, assim como relações entre os vários itens; e c) o conhecimento subjacente à capacidade de formar entradas lexicais gramaticais novas (e, naturalmente, rejeitar as agramaticais). (Basilio, 1980: 11)

Essa ideia de competência lexical está na dependência da noção de produtividade, que, no caso das palavras, diferentemente da sintaxe, é de modo geral menos sistemática, se observada apenas da perspectiva dos itens existentes na língua. O problema, assim, de dar conta das palavras possíveis nas línguas está presente na formulação das regras de formação de palavras, sob diferentes modelagens teóricas, nas propostas de Halle (1973), Jackendoff (1975) e, de modo mais definitivo, de Aronoff (1976).

A questão, contudo, identificada por Basilio (1980), é que um único tipo de regra é incapaz de dar conta ao mesmo tempo da transparência e da produtividade de palavras, uma vez que identificamos estruturas que ainda assim não concorrem para a criação de novos itens na língua. Por exemplo, reconhecemos formas derivadas como *momentâneo*, *celeste*, *campestre*, *róseo* e *natalício*, mas não criamos palavras novas com os sufixos que as compõem (cf. Rocha, 1999: 42). Para lidar com isso, Basilio propõe que o léxico se constitui de regras de formação de palavras mas também do que chama de regras de análise estrutural.

O léxico gerativo — ideia à qual se filia essa proposta dos dois tipos de regras de Basilio — acaba por conter, nesse sentido, informações sobre palavras reais da língua e sobre palavras potenciais. É o que afirma Spencer (1991).

Then we would say that the linguistic lexicon is a list of actual words. Such a list is sometimes referred to as the permanent lexicon. We can contrast it with the (unbounded) list of potential words (which is often referred to as the conditional lexicon or potential lexicon. (Spencer, 1991: 48)

O autor problematiza a noção de léxico permanente diante do fato de que muitas das formas dicionarizadas não estão efetivamente em uso e assume o conceito de lexema como alternativa para restringir o léxico. Lexemas são entidades abstratas, que são vazias de substância fônica. Entretanto, parece que não formamos novas palavras apenas a partir de lexemas, mas de estruturas com constituição mórfica e fonológica (algo próximo da noção de forma de palavra).

Isso fica mais evidente se pensamos em exemplos como o dos novos participios em português brasileiro (PB), que, na contramão do esperado, parecem eleger preferencialmente o padrão irregular, mas que, a rigor, podem estar se alinhando paradigmaticamente à 1ª pessoa do presente do indicativo, como sugere Chagas de Souza (2011).

- (2) lexema: FICAR
 eu *fico*_{Pres.Ind.1pes.}
 eu tinha *fico*_{Participio} em casa

Haspelmath (2024: 72) recupera a classificação de Harley e Noyer (2003) e amplia esse debate, sustentando que o termo léxico pode ser usado em quatro diferentes sentidos, correspondendo a quatro tipos de entidades lexicais: léxico como um componente da gramática que cria formas de palavra; léxico como um componente da gramática que armazena e cria lexemas; léxico como *inventorium*, isto é, o conjunto de elementos imprevisíveis de uma língua; léxico como *mentalicon*, isto é, o conjunto de elementos memorizados (o léxico mental).

Considerando-se que, à exceção de formas de palavras, unidades lexicais são em princípio entidades abstratas, a questão não trivial que se coloca é como se acessa o conhecimento internalizado sobre as palavras de uma língua. Pensando metodologicamente, sustento que uma boa análise em bases formais deve mirar as palavras potenciais a partir do mapeamento das propriedades das palavras reais da língua descrita, na perspectiva sugerida por Spencer, acima referida, que elaboro como segue.

This analysis starts from the premise that knowledge about word formation in a language must be seen, in addition to the perspective of the lexicon in use, from the perspective of the potential lexicon, that is, considering those words that satisfy all structural requirements to be formed, although they do not exist yet. (Schwindt, 2021: 144)

O melhor caminho para se chegar ao léxico em uso é o levantamento de dados em grandes *corpora*, considerando-se frequência dos itens na língua. O léxico potencial, por outro lado, pode ser alcançado por estudos de aquisição da fala e da escrita e de lapsos de fala (slips of tongue), por testes psicolinguísticos envolvendo jogos de palavras, línguas inventadas, logatomas, bem como por estudos sobre processamento lexical (efeito de priming, tempo de reação, rastreamento ocular etc).

O acesso aos léxicos em uso e potencial visa à descoberta do sistema e dos subsistemas. Essa é, *mutatis mutandis*, em perspectiva formal, a abordagem assumida pela Emergent Phonology and Morphology, que pode ser refinada em busca do aleatório que licencia ou é licenciado pelo regular em termos estocásticos, a perspectiva de MaxEnt, problematizada pela abordagem Deep Phonology e pela controvérsia Deep vs. Shallow Stochastic Phonology. Neste texto, por razões de foco e espaço, não pretendo conduzir o debate sob nenhuma dessas perspectivas em específico, ainda que julgue necessário registrar que, no sentido que justifica o olhar para esses léxicos em análises morfofonológicas, não se contradizem, em minha opinião.

3. Léxicos e morfofonologia

É mais ou menos consensual que, em algum nível de análise, morfologia interage com fonologia, e também que, combinados, seus objetos estão relacionados com algum entendimento de léxico. O termo morfofonologia, que tem sido empregado algumas vezes para nomear essa interação, pode se definir como (i) a área de estudos da Linguística que se ocupa de fenômenos que fazem referência a constituintes morfológicos e fonológicos e (ii) a interface, em abordagem gramatical, entre o módulo morfológico e o fonológico (Schwindt, 2024).

A relação entre morfofonologia e léxico passa por perguntas como as que seguem: como se mapeiam som e significado?; como se processam correspondências não biunívocas entre unidades morfológicas e unidades fonológicas (alomorfia)?; que conhecimentos estão em jogo quando julgamos uma palavra como possível ou impossível segundo o léxico de nossa língua?.

O exame de fenômenos morfofonológicos na perspectiva do léxico em uso contrastado ao léxico potencial tem contribuído particularmente para responder a última questão. A seguir exemplifico o papel desses léxicos em três recortes selecionados de alguns dos estudos sobre fenômenos morfofonológicos do PB que conduzi ou de cuja investigação participei.

3.1. Harmonia vocálica ou metafonia verbal

Postula-se na literatura da área que verbos de 2ª e 3ª conjugações estejam sujeitos a um processo de harmonia vocálica entre vogal temática e a última vogal da raiz (Harris, 1974; Wetzels, 1995; Schwindt e Wetzels, 2016), por

muitos compreendido como um processo metafônico. Em (3), as vogais das formas de base, aqui registradas em maiúsculas para indicar subespecificação de altura, assumem o mesmo valor de altura de suas respectivas vogais temáticas, que são apagadas quando ao tema verbal se associa a marca de 1ª pessoa do singular do presente do indicativo (e também das formas de subjuntivo).

- (3) f/E/r+i+r_{Inf.} → f/i/r+O_{Pres.Ind.1}
 b/E/b+e+r_{Inf.} → b/e/b+O_{Pres.Ind.1}

De acordo com estudo do léxico permanente, que relatamos em Schwindt e Quadros (2009), considerando 1.199 verbos listados no Dicionário Aurélio Eletrônico sujeitos a harmonia, 91,6% são harmonizados, com especial relevo para verbos de 2ª conjugação.²

A pergunta que se coloca a partir desse resultado é sobre a atividade do processo na sincronia da língua. Empreendemos, nessa direção, estudo experimental sobre o fenômeno, relatado em Schwindt (2007) e Schwindt e Quadros (2009). A pesquisa se caracterizou por três testes aplicados a informantes da Região Metropolitana de Porto Alegre envolvendo produtividade e transparência da harmonia vocálica verbal, brevemente descritos na sequência.

(i) Teste 1 (escrito), aplicado a 210 informantes, alunos de Ensino Médio, com a tarefa de escolher a forma flexionada para 12 pseudoverbos de 2ª e 3ª conjugações, seguindo padrões fonotáticos do PB, como no exemplo abaixo.

Geralmente eu não FÉRJO / FERJO / FIRJO à noite. (fergir)

(ii) Teste 2 (oral), aplicado a 35 informantes com escolaridade média ou superior, com a tarefa de ler frases como a que segue em voz alta, preenchendo suas lacunas.

FERGIR é algo bom de fazer entre um exercício e outro na academia. Quando eu _____, me sinto bem melhor. Por que tu não _____ também? Agora meu treinador quer que eu _____ todos os dias.

Além dos testes dedicados a apurar o grau de produtividade (no sentido de formação de novos vocábulos) da harmonia vocálica, elaboramos um terceiro instrumento, com vistas a verificar o grau de transparência desse fenômeno.

(iii) Teste 3 (escrito), aplicado a outros 60 participantes de escolaridade média, com a tarefa de escolher entre as opções oferecidas a forma de base do pseudoverbo flexionado, em frases como a que segue.

*Eu nem sempre SIRDO depois da academia.
 (a) SERDER (b) SERDIR*

² Nas investigações empíricas relatadas neste texto dados de dicionários oficiais foram classificados como *léxico permanente*, enquanto dados de grandes *corpora* de língua efetivamente em uso, mesmo que predominantemente de modalidade escrita, como *léxico em uso*. A distinção é crucial se está em jogo o grau de segurança na avaliação da sincronicidade desses dados. Ambas as fontes, porém, salvaguardados seus limites, podem figurar como ponto de contraste ao léxico potencial.

Os resultados dos testes 1 e 2 mostraram variabilidade relevante quanto à escolha da vogal da base, sem preferência clara por padrões harmonizados. O teste 3, por outro lado, revelou predominância de padrões harmonizados, com escolha de formas infinitivas contendo VT compatível em termos de altura com a vogal da raiz da forma flexionada. O estudo permitiu constatar, assim, com base nas amostras investigadas, que a harmonia vocálica metafônica é processo transparente na língua, apesar de não ser produtivo.

3.2. Plurais de nomes fechados por ditongo nasal

Em substantivos e adjetivos do português fechados por ditongo nasal, <ões> (aqui indicado ortograficamente por estratégia de simplificação), e não a marca padrão de plural -s, é a variante preferida.

- (4) a. ladrão / ladrões
b. mão / mãos / *mões
c. pão / pães / *pões

Em Schwindt *et al.* (2021), observamos, em levantamento no léxico em uso, considerando 9.245 dados do Corpus Brasileiro³, que o emprego dessas variantes minoritárias, <ãos> ou <ães>, prepondera em contextos de palavras monossílabas (ex. mãos, pães), confirmando o que observou Rizatto (2018), de paroxítonas (ex. órfãos, órgãos) e palavras formadas com sufixos gentílicos e assemelhados (ex. cristãos, catalães). No interior desses subléticos, essas alternantes apresentam também alta frequência.

Em Gaggiola e Schwindt (2025), apresentamos resultados de investigação no léxico potencial. Em experimento com pseudopalavras respondido por 116 participantes (1.392 respostas), examinamos o emprego das três alternantes de plural de palavras fechadas por <ão>, considerando possíveis condicionadores linguísticos. A seguir está um exemplo de frase-veículo empregada no teste, cuja escolha se definia após leitura e escuta das estruturas completadas por cada uma das opções.

Zicra era um país pacato no leste da Europa, até que seus habitantes, os ZICRÕES/ZICRÃOS/ZICRÃES, resolveram se rebelar.

Espelhando em boa medida o léxico em uso, o resultado da análise do léxico potencial apontou para o caráter preditivo das alternantes minoritárias, <ãos> ou <ães>, quando se tratava de palavras monossilábicas, de paroxítonas e de palavras afixadas por sufixos gentílicos. Uma possível explicação para a preferência por monossílabos pode ser a de proteção de sílabas iniciais, de modo que se preserve a vogal nuclear da forma de base, como apontaram Becker *et al.* (2012). Em relação às paroxítonas e aos gentílicos e assemelhados, a explicação pode estar na frequência desses padrões no léxico permanente (refletido no uso), que acaba por influenciar a escolha em formas novas.

Conforme Schwindt *et al.* (2021), paroxítonas estão em número muito reduzido no léxico em uso e apresentam preferência quase absoluta por <ãos>,

³ <https://www.linguateca.pt/acesso/corpus.php?corpus=CBRAS>

já gentílicos e assemelhados, menos escassos em ocorrências, distribuem-se de modo relativamente equilibrado entre <ãos> e <ões> e incluem uma percentagem não desprezível de casos de itens muito frequentes exibindo a alternante <ães>.

Aqui, a análise dos *léxicos* permitiu concluir, em relação à amostra estudada, que, padrões aparentemente menos comuns na língua podem ainda assim ser produtivos em contextos próprios, ou subléxicos. Interessantemente, neste caso, os dados do léxico em uso foram confirmados pelos resultados observados na análise do léxico potencial.

3.3. Gênero de substantivos fechados por -e

Também o estudo do uso do gênero gramatical em PB se beneficia de um olhar contrastivo entre léxicos.

A demanda descritiva existe, em primeiro lugar, porque os elementos fonológicos que fecham as palavras não correspondem categoricamente a gêneros gramaticais específicos, ou seja, há um aspecto idiossincrático na categoria gênero (que a diferencia de número, por exemplo) que é própria de fenômenos lexicais.

(4) Segmento final e gênero de substantivos do PB	Feminino	Masculino
Vogal /a/	casa	mapa
Vogal /o/	tribo	bolo
Vogal /e/	ponte	pente
Vogal tônica	avó	cipó
Ditongo oral	nau, lei	degrau, rei
Ditongo nasal	discussão, viagem	limão, homem
Consoante	variável, cor, luz	anzol, amor, giz

Uma segunda motivação da demanda descritiva da marcação de gênero frente aos diferentes léxicos se deve ao fato de que sua frequência de uso efetivo não corresponde plenamente à frequência de tipo observada no léxico dicionarizado. A seguir é apresentada uma síntese dos principais achados sobre produtividade de gênero tendo por base Schwindt (2018, 2020), entre outros estudos de que participei no âmbito do projeto que conduzimos em nosso grupo de pesquisa.

No que diz respeito à distribuição no léxico permanente, a partir da análise de 17.049 substantivos extraídos do Dicionário Aurélio Eletrônico, observamos que há mais nomes uniformes femininos do que masculinos em termos de tipos constantes do léxico permanente, em especial porque substantivos abstratos sufixados por -ção são categoricamente femininos (ex. adoração, rendição). Evidência de que se trata de particularidade do sufixo neste caso está no fato de que outras palavras fechadas por ditongo nasal podem ser tanto masculinas quanto femininas (ex. o arrastão, a gestão). Em relação ao padrão mais frequente do léxico, o das palavras fechadas por vogal átona em sílaba sem coda (ou seguida pelo marcador de plural -s), constatamos que praticamente todas as terminadas em -o são masculinas, a maioria das terminadas em -a é feminina, mas as terminadas em -e estão equilibradamente distribuídas entre os dois gêneros.

O gênero das palavras terminadas em -e também se sujeita a algumas restrições de natureza morfológica.⁴ Por exemplo, nomes fechados pela sequência fonológica /ite/ podem ser masculinos ou femininos (ex. o rebite, a elite), mas se forem derivados pelo sufixo -ite (com sentido de *inflamação*) serão sempre femininos (ex. a tendinite, a rinite, a flebite). Isso parece se aplicar semelhantemente a formas fechadas por -ose ou -dade (ex. neurose, maturidade). Também nomes fechados por -nte parecem ter comportamento regulado pela morfologia. Nesse caso, formas em que o sufixo está mais associado ao verbo que lhe serve de base, como resquício de particípio presente, são mais resistentes à entrada do marcador de feminino -a, como é próprio das categorias invariáveis. Isso pode levar ao juízo de que *presidenta do Brasil* é estrutura bem formada ao passo que **amanta das Letras* não é.

Em relação ao léxico em uso, examinamos inicialmente 4.800 dados extraídos de 24 entrevistas das três capitais que compõem a amostra-base do PROJETO VARSUL (Variação Linguística no Sul do Brasil).⁵ Constatamos, entre outros aspectos interessantes atinentes ao uso efetivo de marcas de gênero, que a distribuição dos segmentos finais se aproximou do observado no léxico permanente. Destacou-se, entretanto, o resultado para palavras terminadas em -e, que no uso privilegia o feminino, com um índice de 72,7%, contrastando com o equilíbrio entre os gêneros observado para os tipos dicionarizados. Ainda, segundo o Projeto ASPA (Avaliação Sonora do Português Atual), corpus de referência disponível online à época e adotado em nosso estudo, mais de 90% desses nomes são de alta frequência lexical.

Por fim, em relação ao léxico potencial, em Surreaux e Schwindt (2022), além de analisarmos 242 substantivos do inglês emprestados ao português a partir de corpora bibliográfico e observações empíricas, investigamos em média 1.200 dados de um teste escrito contendo pseudopalavras, divulgado em redes sociais, respondido por 89 falantes nativos de PB. O experimento continha 60 frases-veículo com lacunas a serem preenchidas com um determinante masculino ou feminino, como a exemplificada a seguir.

Já sabemos muito sobre A/O filate.

No que diz respeito aos estrangeirismos, observamos o padrão preferencial pelo masculino (ex. o job, o scanner), com uso de feminino orientado por analogia com palavras existentes no inventário do português, seja por equivalente tradutório (ex. a selfie = a foto) seja por critérios semânticos mais amplos (ex. a scooter = a moto), com algum espaço para equivalentes formais, quando a palavra estrangeira se apresenta fechada por <a> (ex. a creepypasta 'conteúdo de terror postado na internet'). Em relação ao teste com pseudopalavras, entre outros resultados, destacamos a preferência pelo masculino no caso de itens fechados por -e, excluídos os sufixos femininos, com índice de 72,7%, contrastando com o equilíbrio observado no léxico permanente

⁴ A palavra *tribo* é, efetivamente, o único caso de um substantivo genuíno de gênero feminino fechado por -o em uso em português brasileiro. Além de *tribo*, o dicionário registra *imago*, *virago*, de raro uso, e *libido*, este vocábulo variando entre feminino e masculino. Além desses casos, há deadjetivais como *soprano*, que cabem em geral numa paráfrase do tipo *ela é uma soprano = cantora do tipo soprano*, e truncamentos como *foto* e *moto*, femininos antes da redução.

⁵ <https://www.varsul.org.br>

e com a vantagem do feminino observada no léxico em uso, relatados em Schwindt (2018, 2020). Essa flutuação reforça o caráter lexical (mais do que gramatical) do gênero de palavras fechadas por -e, que toma o masculino como *default* potencial, restrito, porém, em contextos reais de comunicação.

Outros estudos em cuja análise se contemplam diferentes léxicos foram/vêm sendo desenvolvidos em nosso grupo de pesquisa. Entre eles, destacamos: investigação de fenômenos como afixação e palavra fonológica, competição entre os sufixos -ção e -mento, competição entre os sufixos -inho e -zinho, derivação de palavras com a estrutura base+I+Vsufixo, derivação de palavras com a estrutura base+n+Vsufixo, harmonização vocálica variável, redução da nasalidade em ditongos finais átonos; variação no emprego de róticos.

4. Léxico, morfofonologia e modelagem gramatical

A relação entre léxico, morfologia e fonologia recebeu diferentes tratamentos em se tratando de arquitetura de gramática em modelos de base gerativa. Em parte isso se deve ao próprio entendimento do termo léxico, como sugerido nas seções iniciais deste texto, que acaba inclusive por levar algumas teorias a se rotularem como lexicalistas ou não lexicalistas. A discussão é complexa, e não tenho pretensão de esgotá-la aqui, obviamente, mas apenas de contribuir para o delineamento mais preciso do problema.

Uma pergunta interessante quando se discute lugar da morfologia e do léxico (e da fonologia) na gramática é se formação de palavras está apartada de flexão, uma realidade para teorias fundadas na hipótese lexicalista, mas não para a MD, por exemplo. Isso porque as características consensualmente entendidas como lexicais, por assim dizer, são menos contestáveis na derivação, por exemplo, onde está em jogo mudança de significado nocional, do que na flexão, onde se manipulam apenas funções gramaticais a uma base em princípio nocionalmente estável. Daí muitas teorias se afiliarem à hipótese lexicalista fraca.

Stump (2001, cap. 1) classifica as teorias em dois tipos no que diz respeito à relação entre morfologia flexional e léxico: as lexicais e inferenciais. Nas teorias lexicais as associações entre afixos e propriedades morfossintáticas estariam listadas no léxico, em entradas lexicais, enquanto nas teorias inferenciais as relações formais sistemáticas entre a raiz de um lexema e a palavra plenamente flexionada são expressas por regras ou fórmulas (nesta última concepção, o verbo do inglês *likes* 'gosta', por exemplo, seria inferido da raiz *like* pela regra que associa -s a 3ª pessoa do singular). Para o autor, essa tipologia cruza com outra, a que separa as teorias em incrementais e realizacionais. Nas teorias incrementais, palavras são incrementadas por propriedades morfossintáticas somente quando a elas se associam os expoentes dessas propriedades. Nas teorias realizacionais, a associação do expoente é licenciada pela prévia associação da raiz com as propriedades morfossintáticas (ex. [*like*]_{Raiz}+ [*Pres. Indic. concord. 3ª pessoa*] ⇔ s). A seguir essas duas tipologias são apresentadas, com exemplos de modelos que, na visão de Stump, se enquadram em cada combinação.

- (5) Tipologia de teorias quanto à relação entre morfologia flexional e léxico (Stump, 2001, cap. 1)
- | | |
|----------------------------|--|
| Lexical incremental: | teoria sintática de formação de palavras (Lieber, 1992) |
| Lexical realizacional: | Morfologia Distribuída (Halle e Marantz, 1993) |
| Inferencial incremental: | Morfologia Articulada (Steele, 1995) |
| Inferencial realizacional: | teorias no modelo Palavra e Paradigma (Matthews, 1972; Zwicky, 1985; Anderson, 1992) |

O autor desenvolve diferentes argumentos, que escapam à abrangência deste artigo, para sustentar o último tipo de teorias, inferencial realizacional, como a proposta mais adequada para lidar com morfologia flexional.

Centrado em problematizar o lugar do léxico na arquitetura da gramática e sua relação com a morfofonologia, desejo discutir brevemente neste ponto do texto teorias classificadas como lexicais na tipologia de Stump. Nesse sentido, parece que foram de fato os modelos lexicais incrementais (os que motivaram a hipótese lexicalista e seus herdeiros) os de maior relevância na tradição dos estudos formais em morfologia. Ressalva se faz à Morfologia Distribuída, modelo não lexicalista de grande impacto nos últimos trinta anos, curiosamente classificado pelo autor como lexical realizacional, fato que aprecio mais adiante neste texto. O critério de máxima parcimônia (o menor número de sistemas gerativos associado a maior capacidade explicativa) tem sido sustentado como o principal critério para avaliação desses modelos.

Marco sujeito a controvérsia, atribui-se à publicação do artigo *Remarks on nominalization* (Chomsky, 1970) a proposição da hipótese lexicalista. Em palavras simples, trata-se da ideia de que existe um sistema computacional que forma palavras e que é independente, como módulo e quanto à natureza de suas operações, do sistema que forma frases.

Há aí uma importante redefinição de léxico, que, dito de modo amplo, até o modelo gerativo padrão não representava muito mais do que informação idiossincrática passível de ser processada pela sintaxe. Dessa ideia surgiram propostas do que se tornou conhecido como Morfologia Lexical (ou Morfologia Gerativa), muitas delas baseadas em variações da hipótese inicial, como a hipótese lexicalista forte, essa segundo a qual formação de palavras e flexão integram o assumido componente lexical.

A fonologia, por seu turno, com a publicação do *The Sound Pattern of English* (SPE), de Chomsky e Halle (1968), havia avançado mais do que a morfologia, uma vez que se assumia seu caráter gerativo. A melhor conciliação entre a fonologia e a morfologia gerativas acontece com a proposição da Fonologia e Morfologia Lexical (FML), proposta por Kiparsky (1982) e Mohanan (1982). Sua arquitetura, que, para os propósitos deste texto, sintetiza a versão mais convincente da hipótese lexicalista, é apresentada na Figura 1.

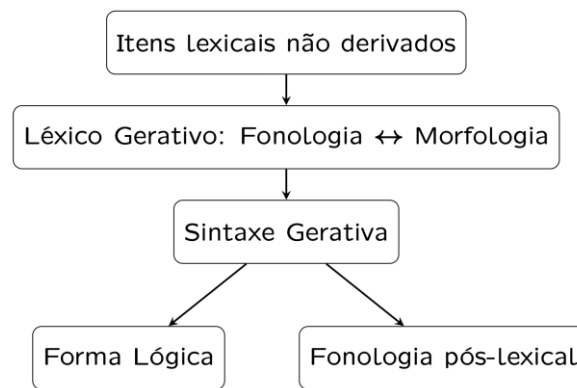


Figura 1: Arquitetura da Fonologia e Morfologia Lexical

O diferencial da FML em relação à Morfologia Lexical padrão está no fato de que o componente gerativo da fonologia, pós-sintático no SPE, é agora combinado pré-sintaticamente à morfologia, num grande componente chamado *léxico*, que se alimenta dos itens não derivados (o *léxico lista*, próprio do que entendemos por dicionário mental). Se, porém, essa teoria tem a vantagem de promover a interação plena entre regras fonológicas e morfológicas, enquanto modelo unidirecional não resolve a dificuldade de acesso de regras sintáticas às operações morfológicas e fonológicas. Além disso, apenas operações fonológicas irrestritas lexicalmente se processam sobre o output da sintaxe, as regras rotuladas como pós-lexicais.

Onze anos mais tarde surgem outros dois modelos de grande impacto, que seleciono para contrastar à LPM: a Teoria da Otimidade (TO), proposta por McCarthy e Prince (1993) e Prince e Smolensky (1993), e a Morfologia Distribuída (MD), proposta por Halle e Marantz (1993). Trata-se de duas teorias de base gerativa, mas muito distintas em suas concepções sobre computação linguística. Entretanto, sua diferença mais básica, eu diria, está no fato de que a TO é orientada fonologicamente, ao passo que a MD é orientada sintaticamente. Quero dizer com isso que, apesar de sua pretensão ser cobrir plenamente o que se entende por gramática em linguística gerativa (pode-se fazer alguma sintaxe e morfologia com TO; pode-se fazer alguma fonologia e morfologia com MD), fica claro que são pensadas da ótica de uma e de outra subárea.

Léxico na TO é sinônimo de input. Isso porque, na visão desse modelo não incidem restrições sobre o léxico, princípio conhecido como Riqueza da Base. Não há, além disso, um componente lexical, morfológico, sintático ou semântico com operações em série intra ou intermodulares. Na proposta mais básica da teoria, gramática consiste em um ranking de restrições universais violáveis que avalia em paralelo a boa formação de candidatos a output, por suas marcas e por sua maior ou menor fidelidade ao input. Na Figura 2 está a arquitetura da TO.

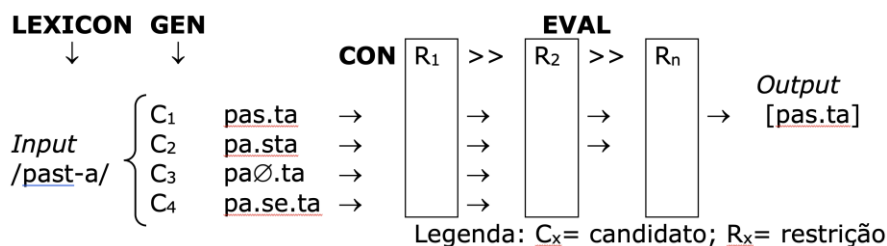


Figura 2: Arquitetura da Teoria da Otimidade
(Adaptada de Kager, 1999: 22 por Schwindt, 2010 (2005: 262))

Considerando-se a tipologia de Stump (2001), pode-se dizer que o modelo básico da TO é lexical incremental, uma vez que morfemas estão associados a traços ou segmentos fonológicos no input e, em princípio, essa associação não se sujeita a avaliação do ranking de restrições (a menos que esteja em jogo algum tipo de mudança estrutural). Morfologia na TO está fundamentalmente contemplada no input e na definição de restrições, muitas vezes identificada como domínio de fidelidade, marcação ou alinhamento generalizado (conforme Schwindt, 2010).

Ainda que a adoção de restrições universais violáveis possa representar uma vantagem sobre o emprego de múltiplos expedientes menos ou mais categóricos (princípios, parâmetros, regras, filtros etc), a falta de derivação na TO tem custo alto ao modelo. Nesse sentido, parte importante do desenvolvimento da teoria foi direcionada para dar conta de interações opacas (não aparentes ou não verdadeiras na superfície). Também o debate sobre o papel de restrições sobre o léxico tem alimentado subteorias no escopo da TO. Destaco nesse particular dois tópicos aos quais vimos nos dedicando em nosso grupo de pesquisa: o estudo de restrições sobre estrutura de morfema, com um papel especial sobre a fonotaxe, e o estudo de subléticos.

Na MD não há léxico no sentido tradicional do termo, enquanto lista de palavras ou lexemas, nem mesmo de morfemas no sentido de amálgama entre som e significado. A classificação de Stump (2001), todavia, sugere tratar-se de uma teoria lexical realizacional. É lexical, interpretando a ideia do autor, por contar com listas, mais estritamente porque associa, na lista 1, afixos a propriedades morfossintáticas (e mesmo semânticas, a depender de sua concepção; cf. Resende, 2023) e, na lista 3, os nós terminais sintáticos a significados nocionais e informações extralinguísticas. É, por outro lado, um modelo realizacional, porque os elementos manipulados na computação sintática são desprovidos de substância fônica; abstratos, só são exponenciados fonologicamente pelas associações promovidas na inserção de vocabulário, alimentada pela lista 2. Na Figura 3 está a arquitetura da MD.

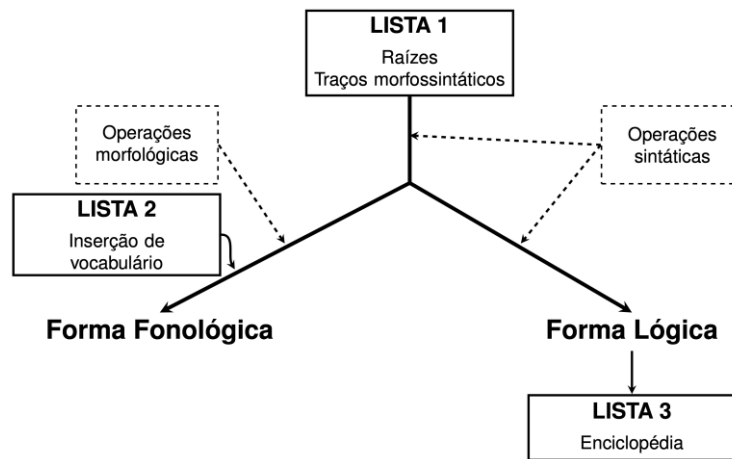


Figura 3: Arquitetura da Morfologia Distribuída
(Adaptada de Bassani e Minussi, 2015: 140)

Na visão de Marantz (1997), a MD não está em contradição com a proposta de Chomsky (1970). Para o autor, há uma leitura equivocada do texto, no sentido de se privilegiarem as propriedades idiossincráticas das nominalizações, enquanto o foco do artigo são as propriedades sintáticas e semânticas sistemáticas dessas formações e as razões por que tais propriedades não permitiriam derivá-las de sentenças transformacionalmente. Chomsky não propõe, na visão de Marantz, nenhuma regra lexical especial ou adiciona à gramática operações que teriam lugar num suposto componente lexical.

What Chomsky really discussed in "Remarks" is summarized by this quote, Chomsky (1970: 17): "We might extend the base rules to accommodate the derived nominal directly (I will refer to this as the 'lexicalist position'), thus simplifying the transformational component; or alternatively, we might simplify the base structures, excluding these forms [the nominalizations], and derive them by some extension of the transformational apparatus (the 'transformational position')." Note that the crucial issue here is about extending the base rules (i.e., allowing N's to take complements) rather than adding operations to a place called "the lexicon." (Marantz, 1997: 214-215)

Sobre o lugar da fonologia na gramática da MD, Marantz (1997: 205-206) também apresenta ponto de vista interessante, ainda que contestável em dada medida, como aponto em seguida. O autor afirma que, se constituintes prosódicos, em particular palavras fonológicas, não coincidem com itens lexicais, mas dependem de mapeamento com essas unidades, faz sentido que sejam construídas pós-sintaticamente. Contribui para essa tese o fato de que as regras fonológicas, além de poderem se aplicar a unidades menores do que o item lexical, podem se aplicar muitas vezes a unidades maiores do que ele. Finaliza esse debate sugerindo que, para ser sustentado, o léxico computacional dos lexicalistas depende de evidência de que a unidade palavra corresponde a algum domínio especial relevante para a sintaxe e para o componente interpretativo, isto é, para significados especiais e/ou para correspondências especiais de som/significado. O problema da reflexão de Marantz, me parece, reside tanto em sua visão terminologicamente estrita de léxico (intrateórica, a rigor), quanto na redução da interação da morfologia com a fonologia. O fundamento da FML é o de que a fonologia opera *durante* a construção da palavra gramatical, ou seja, relaciona-se com peças menores do que o átomo sintático. Concebê-la pós-sintaticamente tem como consequência operar em *loop* com o componente que

processa morfemas ou tratá-la toda como reparo, negligenciando essa interação, problema enfrentado em trabalhos como o de Embick e Shwayder (2018).

A MD vem sendo, como a TO, aprimorada a partir de muitos de seus pontos mais sensíveis. Destaco dois que têm me interessado mais diretamente. O primeiro deles diz respeito à natureza das raízes, sobretudo no que se refere ao debate sobre a relevância de estar previsto já na lista 1 seu significado, de caráter nocional, e mesmo seu conteúdo fonológico. O segundo tem a ver com o tratamento das regras e processos fonológicos sistemáticos num modelo de inserção tardia. Os dois pontos têm alimentado muitas propostas de aperfeiçoamento da teoria. Em particular o segundo provocou algumas propostas de combinação entre TO e MD, em que o ranqueamento de restrições pode se processar no lugar da inserção de vocabulário (Trommer, 2001; Haugen 2011; Wallace, 2013). Essas implementações são todas sujeitas a controverso debate, em especial porque pesam sobre a vantagem da MD de operar com um único sistema computacional, o sintático, sugerindo se refazer seu cálculo de parcimônia.

Para fechar esta seção, considero importante apenas recuperar uma fragilidade da natureza dos modelos formais que afeta diretamente a ideia de exploração de léxicos que defendi até aqui. O fato é que nenhum desses modelos foi pensado para lidar com uma distinção entre língua institucionalizada, uso e potencialidade. Parte-se da crença mais geral de que o léxico potencial é o mais preditivo entre os três, e que deve explicar o uso e as formas sedimentadas nos dicionários. Trata-se, em plano mais geral, do mesmo problema que se identifica quando está em jogo variação, pois, mesmo a TO, que gera inúmeros candidatos a outputs, orienta-se, em seu modelo padrão, para um só output ótimo. Também nesse campo as propostas de aprimoramento são muitas. Omitindo detalhes por questões de foco, afirmo que é a inclusão de índices quantitativos nos modelos (frequências ou probabilidades que integram o conhecimento internalizado do falante) que têm assegurado maior progresso nessa reflexão.⁶

5. Palavras finais

Neste artigo tratei das diferentes concepções de léxico frente à descrição e análise de fenômenos morfofonológicos em modelos de base formal. O objetivo, além de aprofundamento epistemológico sobre o conceito de léxico, foi sustentar a tese de que a melhor análise passa por empiria: é a que considera o léxico em uso contrastado ao léxico potencial, e mesmo ao léxico permanente ou institucionalizado.

⁶ Como aponta, com pertinência, um dos pareceristas, o incremento dos modelos formais com recursos que contemplam o uso, como, por exemplo, a inclusão de índices quantitativos em estruturas ou regras/restrições responsáveis por processar essas estruturas, redimensiona o entendimento de competência linguística preconizado pelos primeiros modelos gerativos. Não o invalida, porém, em minha percepção, se ainda se sustenta a ideia de que falantes operam com regras/restrições e de que as processam hipotético-dedutivamente, isto é, partem de uma gramática de fundo universal, mas, a cada confronto com os dados, redimensionam suas hipóteses (seus modelos, em outras palavras). Permitindo-me algum reducionismo, eu diria que a diferença está no fato de que a própria hipótese é quantificada neste caso, e não mera etapa estática em resposta à observação da frequência.

Esse modo de ver, que é hoje compartilhado por muitos, no âmbito formal ainda desafia em certa medida a hipótese de autossuficiência da abordagem abstrata. De outra parte, associar modelagem gramatical ao exame empírico desafia igualmente abordagens concretistas, fixadas tão somente no uso.

A pergunta que se coloca hoje à análise linguística não é diferente da pergunta inicial dos modelos gerativos, mas é desafiada pelos avanços dos estudos experimentais e dos recursos computacionais: em que medida o que captamos no contínuo da comunicação corresponde às representações que subjazem esse dado? e para que importa essa informação subjacente?

A ideia de um léxico potencial que dialoga com um léxico em uso (e que ganha a feição institucionalizada como léxico permanente) e sua relação com processos morfofonológicos sustenta em boa medida uma abordagem de uso sem o reducionismo do tipo *what you see is what you get*.

A extensão da distância entre representações concretas e representações abstratas, porém, vem sendo redimensionada a cada novo estudo que se aventura a questionar pressupostos teórico-descritivos do ponto de vista de seu processamento e computação. Reentender memória (que, afinal, sustenta o argumento de listas menos ou mais ricas) faz reentender parcimônia, e reentender parcimônia força repensar os modelos. Em outras palavras, ao menos no que diz respeito à morfofonologia, parece que somos convidados a deixar o interior das salas do léxico ou da gramática para nos aproximarmos de suas janelas, porque é nelas que parecem residir as principais perguntas e muito provavelmente é lá que residem as melhores respostas.

Agradecimento

Agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico pelo apoio financeiro ao projeto que abarca este estudo, em forma de Bolsa de Produtividade em Pesquisa, processo 309576/2022-9. Agradeço também ao Pedro Eugênio Gaggiola, aos pareceristas anônimos e aos editores pelas sugestões para aprimorar o texto.

Referências bibliográficas

- Anderson, Stephen R. 1992. *A-Morphous Morphology*. Cambridge: Cambridge University Press.
- Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.
- Basilio, Margarida. 1980. *Estruturas Lexicais do Português: Uma Abordagem Gerativa*. Petrópolis: Vozes.
- Basilio, Margarida. 2021. Saussure e as estruturas lexicais. *Diadorim*, Rio de Janeiro, vol. 23, n. 2, p. 159-166.
- Bassani, Indaiá de Santana e Rafael Dias Minussi. 2015. Contra a seleção de argumentos pelas raízes: nominalizações e verbos complexos. *Revista Virtual de Estudos da Linguagem*, v. 13: 139-173.

- Becker, Michael, Andrew Nevins e Jonathan Levine. 2012. Asymmetries in Generalizing Alternations to and from Initial Syllables. *Language* 88, no. 2: 231-268.
- Bloomfield, Leonard. 1933. *Language*. New York: Henry Holt.
- Chagas de Souza, Paulo. 2011. Uma Discussão sobre a Flexão de Participípios no Português Brasileiro. *Revista Virtual de Estudos da Linguagem* 9, no. 16: 55-82.
- Chomsky, Noam. 1970. Remarks on Nominalization. In *Readings in English Transformational Grammar*, edited by Roderick A. Jacobs and Peter S. Rosenbaum, 184-221. Waltham, MA: Ginn.
- Chomsky, Noam e Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper & Row.
- Embick, David and Kobey Shwayder. 2018. Deriving morphophonological (mis)applications. In *From sounds to structures: beyond the Veil of Maya*, edited by Roberto Petrosino, Pietro Cerrone, and Harry Van Der Hulst. Mouton: De Gruyter: 193-248.
- Gaggiola, Pedro Eugênio e Luiz Carlos Schwindt. 2025. Alternantes minoritárias no plural de nomes terminados em <ão> no português brasileiro: contextos favorecedores. *Revista do GEL*, v. 21, n. 2, p. 79-101.
- Halle, Morris. 1973. Prolegomena to a Theory of Word Formation. *Linguistic Inquiry* 4, no. 1: 3-16.
- Halle, Morris e Alec Marantz. 1993. Distributed Morphology and the Pieces of Inflection. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, edited by Kenneth Hale and Samuel Jay Keyser, 111-176. Cambridge, MA: MIT Press.
- Harley, Heidi e Rolf Noyer. 2003. Distributed Morphology. In *The Second Glot International State-of-the-Article Book*, edited by Lisa Cheng and Rint Sybesma, 463-496. Berlin: Mouton de Gruyter.
- Harris, James W. 1974. Evidence from Portuguese for the 'Elsewhere Condition' in Phonology. *Linguistic Inquiry* 5, no. 1: 61-80.
- Haspelmath, Martin. 2024. Four Types of Lexicon: A New Typology. *Word Structure* 17, no. 1: 68-92.
- Haugen, Jason D. 2011. Reduplication in Distributed Morphology. *Morphology* 21: 505-548.
- Hockett, Charles F. 1954. Two Models of Grammatical Description. *Word* 10, no. 2-3: 210-234.
- Jackendoff, Ray. 1975. Morphological and Semantic Regularities in the Lexicon. *Language* 51, no. 3: 639-671.
- Kager, René. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.
- Kiparsky, Paul. 1982. Lexical Morphology and Phonology. In *Linguistics in the Morning Calm*, edited by In-Seok Yang, 3-91. Seoul: Hanshin.
- Lieber, Rochelle. 1992. *Deconstructing Morphology: Word Formation in Syntactic Theory*. Chicago: University of Chicago Press.
- Marantz, Alec. 1997. No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon. In *University of Pennsylvania Working Papers in Linguistics*, vol. 4, 2, edited by Alexis Dimitriadis, Laura Siegel et al. Philadelphia: University of Pennsylvania: 201-225

- Matthews, Peter Hugoe. 1972. *Inflectional Morphology: a Theoretical Study Based on Aspects of Latin Verb Conjugation*, Cambridge University Press.
- McCarthy, John J. e Alan S. Prince. 1993. *Prosodic Morphology I: Constraint Interaction and Satisfaction*. Technical Report RuCCS-TR-3. New Brunswick, NJ: Rutgers University Center for Cognitive Science.
- Mohanan, K. P. 1982. *Lexical Phonology*. PhD diss., Massachusetts Institute of Technology.
- Prince, Alan e Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. Technical Report RuCCS-TR-2. New Brunswick, NJ: Rutgers University Center for Cognitive Science.
- Resende, Maurício. 2023. As raízes da semântica e a semântica das raízes: notas sobre "conteúdo semântico" em Morfologia. *Caderno de Squibs: temas em estudos formais da linguagem*, v. 9, n. 2, p. 18-42.
- Rocha, Luiz Carlos de Assis. 1999. *Estruturas Morfológicas do Português*. Belo Horizonte: UFMG.
- Saussure, Ferdinand de. 1916. *Cours de Linguistique Générale*. Edited by Charles Bally and Albert Sechehaye. Paris: Payot.
- Schwindt, Luiz Carlos. 2007. Paradigmatic Correspondences in the Brazilian Portuguese Verbal Vowel System. *Acta Linguistica Hungarica*, v. 54, 35-50.
- Schwindt, Luiz Carlos. 2010. Morfologia. In: Leda Bisol e Luiz Carlos Schwindt (org.) *Teoria da Otimidade: Fonologia*. Campinas: Pontes. p. 207-227.
- Schwindt, Luiz Carlos. 2018. Exponência de gênero e classe temática em português brasileiro. *Revista DELTA*, v. 34, n. 2, p. 745-768.
- Schwindt, Luiz Carlos. 2020. Predizibilidade da marcação de gênero em substantivos no português brasileiro. In: Danniel Carvalho e Dorothy Brito. (org.). *Gênero e língua(gem): formas e usos*. 1ed., Salvador / Bahia: Editora da UFBA, v. 1, p. 279-294.
- Schwindt, Luiz Carlos. 2021. Underlying representation of [w]-final words in Brazilian Portuguese: evidence from morphological derivation. *Acta Linguistica Academica* 68, 1-2, p. 139-157, Budapest: Akadémiai Kiadó.
- Schwindt, Luiz Carlos. 2024. Current issues in the Portuguese phonology-morphology interface. In: André Zampaulo (ed.) *The Routledge Handbook of Portuguese Phonology*.
- Schwindt, Luiz Carlos, Pedro Eugênio Gaggiola e Isabela Prisco Petry. 2021. Frequência e distribuição de plurais irregulares no corpus brasileiro. *Revista de Estudos da Linguagem*. 29, n. 2, 1289-1324.
- Schwindt, Luiz Carlos e Emanuel Souza de Quadros. 2009. A harmonia vocálica verbal no léxico dicionarizado do PB. *Letrônica*, v. 2, p. 58-70, Porto Alegre: PUCRS.
- Schwindt, Luiz Carlos e Leo Wetzels. 2016. The morphology and phonology of inflection. In: João Costa, Sergio Menuzzi e Leo Wetzels. *Handbook of Portuguese Linguistics*. Wiley-Blackwell, p. 188-209.
- Spencer, Andrew. 1991. *Morphological Theory: An Introduction to Word Structure in Generative Grammar*. Oxford: Blackwell.
- Steele, Susan. 1995. Towards a Theory of Morphological Information. *Language* 71, no. 2: 260-309.
- Stump, Gregory T. 2001. *Inflectional Morphology: A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.

- Surreaux, Pedro Perini e Luiz Carlos Schwindt. 2022. Marcação de gênero gramatical em formações novas em português brasileiro. *Estudos Linguísticos e Literários*, Salvador, n. 72, p. 390–414.
- Trommer, Jochen. 2001. Distributed Optimality. PhD diss., Universität Potsdam.
- Wallace, Rebekah. 2013. Distributed Morphology, Roots, and Morphophonological Conditioning in the Verbal System of Ulwa. PhD diss., University of Edinburgh.
- Wetzels, W. Leo. 1995. Mid Vowel Alternations in the Brazilian Portuguese Verb. *Phonology* 12, no. 3: 281-304.
- Zwicky, Arnold M. 1985. How to Describe Inflection. In *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society*, edited by Mary Niepokuj et al., 372-386. Berkeley: Berkeley Linguistics Society.

Nota

Este artigo foi integralmente idealizado e elaborado por seu autor.

Disponibilidade dos dados

Este artigo foi integralmente idealizado e elaborado por seu autor. Os dados utilizados são também de pesquisas por ele desenvolvidas, isoladamente ou em colaborações. Essas pesquisas estão fielmente listadas nas referências, e todas as informações a elas relacionadas podem ser acessadas em suas fontes originais, espelhadas no website pessoal do autor: <https://sites.google.com/site/luizschwindt>.

Nota de Aceitação

Este texto foi aceito para publicação pelo único Diretor-Editor da revista, Adolfo Elizaincín, que agiu de acordo com o disposto na "Declaração de comportamento ético" da revista *Lingüística* (https://www.mundoalfal.org/sites/default/files/revista/Declaracao_comp_etico.pdf), primeiro parágrafo do capítulo "Obrigações do Diretor-Editor". O Diretor-Editor, os revisores e os autores devem aderir explicitamente a esta declaração.