

Conjugador verbal do português brasileiro e análise morfofonológica dos radicais

Gustavo Lopez Estivalet¹

¹Laboratório de Processamento Linguístico (LAPROL) - Universidade Federal da Paraíba (UFPB), João Pessoa-PB, Brasil.
gustavoestivalet@hotmail.com

Abstract. *The main objective of this work was to develop a verbal conjugator of Portuguese with the generalities of regular verbs and specificities of irregular verbs. The secondary objectives were: i. outline the function of the developed instrument; ii. describe the general characteristics of verbal conjugation in Portuguese, and iii. analyze the morphophonological irregularities of verbal stems in Portuguese. The verbal conjugator was programmed in R language and allows the conjugation of regular and irregular verbs, as well as pseudoverbs and neologisms of Portuguese in three forms: simple, reflexive, and pronominal. This instrument is available for free on <https://lexicodoportugues.shinyapps.io/Conjugator/>.*

Resumo. *O objetivo principal deste trabalho foi desenvolver um conjugador verbal do português comportando as propriedades gerais dos verbos regulares e específicas dos verbos irregulares. Os objetivos secundários foram: i. delinear o funcionamento do instrumento desenvolvido; ii. descrever as características gerais da conjugação verbal do português e iii. analisar as irregularidades morfofonológicas dos radicais verbais do português. O conjugador verbal foi desenvolvido em linguagem de programação R e permite a conjugação de verbos regulares e irregulares, assim como pseudoverbos e neologismos do português em três formas: simples, reflexiva e pronominal. Este instrumento está disponível gratuitamente em <https://lexicodoportugues.shinyapps.io/Conjugator/>.*

1. Introdução

Nos últimos anos, uma intensa discussão em torno da arquitetura do funcionamento do processamento morfológico tem se dado, especialmente em relação às classes flexionais. Por um lado, a arquitetura de “item e processamento” prevê a representação de unidades mínimas no léxico mental e a computação morfológica para formação de formas flexionadas; por outro lado, a arquitetura de “palavra e paradigma” prevê a representação de formas flexionadas como palavras inteiras a partir das relações com os paradigmas flexionais estabelecidos e produtivos (Blevins, 2006). Neste sentido, grande parte da discussão tem se dado em torno da representação e do processamento das formas irregulares apresentando modificações morfofonológicas (e.g., impera[dor]/impera[triz], poder/p[O]de) e alomorfia (e.g., [ag]ir/[aj]o, [sab]er/[soub]e, ir/[v]ou); em suma, as formas irregulares são armazenadas no léxico mental como unidades inteiras ou possuem computações específicas para sua flexão?

As línguas latinas possuem morfologia rica e a conjugação verbal é uma categoria prototípica de produtividade flexional com grande número de formas em função de modo, tempo, pessoa e número (Villalva, 2007). Nesse sentido, a conjugação verbal de verbos específicos é frequentemente consultada para a verificação de formas irregulares assim como para apropriação de novas palavras, tais como estrangeirismos e neologismos. Para tanto, conjugadores verbais disponibilizados *online* na internet se tornaram uma ferramenta pertinente e frequentemente utilizada para verificação de formas verbais flexionadas, além disso, um algoritmo de conjugação e flexão verbal do português pode contribuir enormemente para o desenvolvimento e implementação do processamento de línguas naturais.

Sendo assim, o objetivo primário do presente trabalho foi desenvolver um conjugador verbal do português brasileiro (PB) para fins de pesquisa, consulta e aplicações computacionais. Os objetivos secundários foram: i. implementar um algoritmo de conjugação verbal do PB baseado na teoria linguística, ii. analisar as irregularidades morfofonológicas dos radicais verbais do PB e iii. oferecer uma ferramenta *online* e *offline* gratuita e aberta com aplicação em palavras e pseudopalavras do PB. Portanto, este trabalho justifica-se a partir da necessidade de descrição das características específicas da conjugação verbal do português e sua utilização no desenvolvimento de ferramentas de processamento de línguas naturais, assim como a exploração destes recursos em pesquisas em psicolinguística e linguística computacional. Ainda, esta ferramenta pode ser explorada como instrumento de consulta sobre a conjugação verbal do PB assim como recurso didático de ensino do PB (Estivalet & Meunier, 2015).

As perguntas que guiaram esta pesquisa foram: i. Quais as regras de sufixação para flexão verbal do português? ii. Quais são as subclasses de verbos irregulares do português? iii. Como as variáveis grafotáticas e de n-grama influenciam a flexão verbal de formas irregulares do português? A hipótese nula é que a primeira classe de conjugação de verbos com infinitivos terminados em [-ar] é completamente regular e produtiva, enquanto as segunda e terceira classes de conjugação de verbos com infinitivos terminados em [-er] e [-ir], respectivamente, é irregular e improdutiva (Veríssimo & Clahsen, 2009). A hipótese de trabalho da presente pesquisa é que todas as classes de conjugação do português apresentam subclasses com diferentes graus de irregularidade e produtividade (Kilani-Schoch & Dressler, 2005) baseado nas estruturas grafotáticas e fonotáticas, assim como nas frequências de n-gramas do português (Albright, 2002).

Para realização da presente investigação, primeiramente foram pesquisados os mecanismos regulares e irregulares da flexão verbal do português e em seguida foi implementado um algoritmo computacional para flexão verbal do PB baseado na teoria linguística (Embick & Halle, 2005). Deu-se especial atenção para análise morfofonológica das irregularidades dos radicais com objetivo de definirem-se subclasses verbais com diferentes graus de irregularidades e produtividade, que por sua vez, possam também contemplar a irregularidades de neologismos, estrangeirismo e pseudopalavras do PB.

2. Conjugação verbal do português brasileiro

Em geral, os dicionários dos PB não apresentam as formas verbais flexionadas, mas somente as entradas lexicais nas formas infinitivas, já as gramáticas tradicionalmente apresentam as conjugações dos tempos verbais de verbos prototípicos. Atualmente, faz-se o uso de conjugadores verbais na internet quando se deseja consultar a forma flexionada específica de um verbo¹. Ainda, algumas ferramentas didáticas já foram desenvolvidas sobre a conjugação verbal, contudo continuam indisponíveis para o grande público, cobram taxas para sua utilização e/ou não apresentam a teoria utilizada para desenvolvimento (Vasilévski & Araújo, 2011). Sendo assim, a ferramenta e o algoritmo apresentados no presente trabalho têm como objetivo contornar estas limitações da descrição, implementação e utilização da conjugação verbal do PB.

2.1. Tempos verbais

A conjugação verbal do PB apresenta três modos, 12 tempos verbais, três pessoas do discurso e dois números de pessoas. O I. modo indicativo contém seis tempos verbais: 1. presente, 2. pretérito perfeito, 3. pretérito imperfeito, 4. pretérito-mais-que-perfeito, 5. futuro do presente e 6. futuro do pretérito; o II. modo subjuntivo contém três tempos verbais: 7. presente, 8. pretérito imperfeito e 9. futuro; o III. modo imperativo contém os tempos verbais: 10. afirmativo e 11. negativo. Além disso, a conjugação verbal do PB apresenta três formas nominais: i. infinitivo, ii. particípio passado e iii. gerúndio; o infinitivo contém o a. infinitivo impessoal e o b. infinitivo pessoal. As pessoas do discurso são 1ª, 2ª e 3ª nos números singular e plural, tipicamente representados pelos pronomes pessoais do caso reto: eu, tu, ele|ela, nós, vós, eles|elas, respectivamente. Além disso, destaca-se no PB a substituição do pronome de 2ª pessoa do singular “tu” pelo pronome “você” e do pronome de 1ª pessoa do plural “nós” pela expressão “a gente” com conjugação da 3ª pessoa do singular, assim como a substituição do pronome de 2ª pessoa do plural “vós” pelo pronome “vocês” com conjugação da 3ª pessoa do plural (Bassani & Lunguinho, 2011).

O futuro do pretérito do indicativo também é conhecido como modo condicional. Além dos tempos simples, o português possui também as formas compostas construídas a partir dos verbos “ter/haver” conjugado nos diferentes tempos verbais seguido do particípio passado do verbo principal, que por sua vez podem ser regulares (e.g., am[ado], com[ido], dorm[ido]), irregulares (e.g., d[ito], ab[erto]) ou abundantes (e.g., entreg[ado]/entreg[ue], salv[ado]/sal[vo]).

2.2. Conjugações

O português apresenta tipicamente três classes definidas em relação à vogal temática, verbos com a forma infinitiva terminada em: [a]r, [e|o]r ou [i]r. A primeira classe em [-ar] é a classe mais regular e produtiva, possuindo maior número de verbos (mais de 4000 verbos). A segunda classe [-er|-or] apresenta irregularidade, número reduzido de verbos (cerca de 240 verbos) e verbos auxiliares e modais de grande utilização (e.g., ser, ter, haver, dever, fazer, querer, poder). A terceira classe em [-ir] apresenta verbos mais irregulares e um número mais reduzido (cerca de 160) (Villalva, 2007).

¹ Por exemplo: <https://www.conjugacao.com.br/>, <http://www.conjugador.com.br/>.

As formas verbais se apresentam em três conjugações: simples, reflexiva e pronominal. A conjugação simples apresenta apenas as formas flexionadas (e.g., eles lavaram); a conjugação reflexiva apresenta as formas flexionadas ligadas por hífen a clíticos expressados pelos pronomes reflexivos representados pelos pronomes pessoais do caso oblíquo átonos: me, te, se, nos, vos, se (e.g., eles lavaram-se); a conjugação pronominal apresenta as formas flexionadas, que por sua vez pode apresentar mudanças específicas, ligadas por hífen a clíticos expressados pelos pronomes objeto direto/indireto representados pelos pronomes pessoais do caso oblíquo átonos: me, te, o/lo/no/a/la/na/lhe, nos, vos, os/los/nos/as/las/nas/lhes (e.g., eles lavaram-no, lavá-las).

2.3. Variáveis morfofonológicas

Para o desenvolvimento do presente trabalho, observaram-se três variáveis em relação à forma dos radicais verbais: i. regularidade (regular, grau de irregularidade, supressão) (Embick & Halle, 2005), ii. morfofonologia (marca ortográfica da vogal temática, mudança de consoante ortográfica) e iii. distância ortográfica de Levenshtein (DOL) (número de inserções, exclusões e substituição entre formas) (Levenshtein, 1966).

Portanto, a partir dessas variáveis, propôs-se a classificação das formas irregulares em subclasses a partir do grau de irregularidade e da consistência grafotática entre os diferentes lexemas de um mesmo lema. Primeiramente os verbos foram classificados como i. regular, ii. irregular ou iii. supletivo; em seguida os verbos irregulares foram analisados a partir das formas alomórficas; enfim, avaliaram-se as irregularidades grafotáticas específicas das formas flexionadas (Albright, 2002).

3. Metodologia

3.1. Análise

Com o objetivo de se implementar um algoritmo computacional baseado diretamente na teoria linguística de “item e processamento”, mais especificamente de acordo ao modelo da morfologia distribuída, a flexão verbal do PB foi analisada a partir de dois tipos de morfemas: morfemas flexionais contendo traços morfossintáticos (sufixos) e morfemas lexicais de conteúdo semântico (radicais). Sendo assim, realizou-se uma análise computacional linguística i. das regras de aplicação dos sufixos flexionais (Bassani & Lunguinho, 2011) e ii. da morfofonologia e da alomorfia dos radicais lexicais verbais das subclasses dos verbos irregulares (Embick & Halle, 2005).

Em relação às operações de sufixação para flexão verbal, partiu-se da 1ª classe de verbos terminados em [-ar] produtiva e procedeu-se com a formulação de regras de ajuste para as 2ª [-er] e 3ª [-ir] classes (e.g., IND.PRET.IMP.3S: [va]_{PRET.IMP} → [ia]_{PRET.IMP}/_[e*i*]_{Th}, ama[va] → com[ia][dorm[ia]). Em relação aos radicais verbais, realizou-se primeiramente uma análise morfofonológica ortográfica das formas flexionadas (lexemas) que possuíam divergências em relação às formas infinitivas (lema) no caractere imediatamente anterior a vogal temática (e.g., [caç]ar~[cac]ei, [ag]ir~[aj]o, [passe]ar~[passei]o); em seguida, analisaram-se os radicais alomórficos através da comparação da consistência grafotática dos demais caracteres finais dos radicais verbais (e.g., [sab]er~[se]i~[soub]e~[saib]a, [v]ir~[venh]o~[vinh]a). Todas as análises foram realizadas a partir do dicionário UNITEX/DELAf (Muniz, 2004).

3.2. Algoritmo

O algoritmo do conjugador verbal do português foi implementado no programa R (R Core Team, 2014), sendo exploradas principalmente as funções de base `paste` e `gsub`. A primeira permite a concatenação de dois ou mais conjuntos de caracteres e foi utilizada para formação de formas verbais flexionadas através da operação de junção entre o radical (formado pela raiz e vogal temática) e os sufixos flexionais. A segunda permite a substituição de caracteres específicos e foi utilizada para as operações de ajuste ortográfico (morfofonológico). As expressões regulares (REGEX) foram especialmente exploradas para manipulação de caracteres e para o desenvolvimento de funções específicas (Fitzgerald, 2012), explorando-se especialmente os conjuntos de caracteres, limitadores de palavra e operadores de repetição (e.g., $\wedge.[aei]r\$$ = infinitivo, formas finalizadas pelas vogais temáticas seguidas de morfema infinitivo).

Para análise dos radicais alomórficos, explorou-se a DOL entre as formas dos radicais dos lexemas e lemas. A DOL é definida a partir do cálculo do número de operações de inserções, exclusões e substituições necessárias para comparação de duas formas (e.g., `sab/saib`: 1 inserção, `sab/soub`: 1 substituição + 1 inserção = 2, `sab/se`: 1 exclusão + 1 substituição = 2) (Levenshtein, 1966).

3.3. Processamento

Os seguintes procedimentos para formação de formas verbais flexionadas foram implementados: i. para os verbos regulares, procede-se de forma linear a junção do radical com os sufixos flexionais, seguida das regras de ajuste; ii. para os verbos irregulares, uma terceira etapa realiza a análise da irregularidade e produtividade do radical e ajusta-os de acordo aos radicais alomórficos correspondentes (Albright, 2002). As operações alomórficas das subclasses dos verbos irregulares são realizadas a partir da análise da estrutura grafotática dos radicais.

Para as conjugações reflexiva e pronominal, são acrescentados os pronomes clíticos na posição de ênclise correspondentes a seleção do usuário: tipo de conjugação (simples, reflexiva, pronominal); conjugação pronominal (`mas/sing`, `mas/plur`, `fem/sing`, `fem/plur`). Enfim, são realizadas operações de ajustes (e.g., `amar/amá-lo`). Além disso, o conjugado possui a opção de apresentação dos rótulos das formas flexionadas (e.g., `lavas_C1.IND.PRE.2S`, classe: 1ª conjugação [-ar], modo: indicativo, tempo verbal: presente, pessoa: 2ª, número: singular).

4. Resultados e discussão

4.1. Conjugador verbal do português brasileiro

Como resultado, o algoritmo aberto e livre implementado no programa R (R Core Team, 2014) e o aplicativo disponibilizado gratuitamente na internet permitem a conjugação completa de todos os verbos do português, além disso, esse instrumento permite a conjugação verbal de neologismos e estrangeirismo, assim como de pseudoverbos que obedeçam as regras grafotáticas do PB (Keuleers & Brysbaert, 2010). Observa-se que estas características são exatamente as principais limitações de outros conjugadores verbais do português (Vasilévski & Araújo, 2011).

Na Figura 1 abaixo, são apresentados exemplos do funcionamento do aplicativo online: a esquerda a conjugação simples em [-ar] com exibição dos rótulos, no centro a conjugação reflexiva em [-er] e a direita a conjugação pronominal feminina/plural em [-ir]. Na parte superior, a caixa de texto permite a entrada de qualquer verbo na forma infinitiva (i.e., REGEX: $\wedge.[*[\text{aeoi}]\text{r}\$]$). Abaixo, seleciona-se o tipo de conjugação, no caso de conjugação pronominal, pode-se selecionar o gênero e número da conjugação. Enfim, a opção TAG apresenta os rótulos de cada forma flexionada. Portanto, as formas geradas pelo algoritmo podem facilmente ser utilizadas por outros programas de processamento de línguas naturais explorando os rótulos contendo os traços morfossintáticos das formas flexionadas (Estivalet & Margotti, 2014).

The figure shows three screenshots of a web application for Portuguese verb conjugation. Each screenshot displays the input verb, conjugation type, gender/number, and a table of conjugated forms with their corresponding tags.

Left Screenshot (Verbo: lavar):

pres	perpast	imppast
lavo_C1.IND.PRE.1S	lavei_C1.IND.PRP.1S	lavava_C1.IND.IMP.1S
lavas_C1.IND.PRE.2S	lavaste_C1.IND.PRP.2S	lavavas_C1.IND.IMP.2S
lava_C1.IND.PRE.3S	lavou_C1.IND.PRP.3S	lavava_C1.IND.IMP.3S
lavamos_C1.IND.PRE.1P	lavamos_C1.IND.PRP.1P	lavávamos_C1.IND.IMP.1P
lavais_C1.IND.PRE.2P	lavastes_C1.IND.PRP.2P	laváveis_C1.IND.IMP.2P
lavam_C1.IND.PRE.3P	lavaram_C1.IND.PRP.3P	lavavam_C1.IND.IMP.3P

Center Screenshot (Verbo: comer):

pres	perpast	imppast	pluspast	presfut
como-me	comi-me	comia-me	comera-me	comer-me-ei
comes-te	comeste-te	comias-te	comeras-te	comer-te-ás
come-se	comeu-se	comia-se	comera-se	comer-se-á
comemo-nos	comemo-nos	comiamo-nos	coméramo-nos	comer-nos-emos
comeis-vos	comestes-vos	comeis-vos	coméreis-vos	comer-vos-eis
comem-se	comeram-se	comiam-se	comeram-se	comer-se-ão

Right Screenshot (Verbo: abrir):

pres	perpast	imppast	pluspast	presfut
abro-as	abri-as	abria-as	abriria-as	abrir-las-ei
abre-las	abriste-las	abria-las	abriria-las	abrir-las-ás
abre-as	abriu-as	abria-as	abriria-as	abrir-las-á
abrimo-las	abrimo-las	abrimo-las	abrimo-las	abrir-las-emos
abris-las	abristeis-las	abris-las	abris-las	abrir-las-eis
abrem-nas	abriram-nas	abriram-nas	abriram-nas	abrir-las-ão

Figura 1. Exemplos de utilização do aplicativo online.

4.2. Análise do corpus verbal

A análise do dicionário UNITEK/DELA (Muniz, 2004) apresentou o total de lemas e lexemas e o total de formas simples de lemas e lexemas conforme a Tabela 1 abaixo. No primeiro par de colunas, apresenta-se o número total de lemas (infinitivos) e lexemas (formas flexionadas) encontrados no dicionário. No segundo par de colunas, apresenta-se o número de lemas e lexemas somente das conjugações simples, isto é, sem as formas reflexivas e pronominais. O terceiro par de colunas apresenta o número de lemas e lexemas das formas irregulares, ou seja, formas flexionadas que apresentaram DOL > 0. O quarto par de colunas apresenta o número de formas supletivas, considerada com substituição da primeira letra e/ou DOL > 2. Destaca-se que a DOL entre cada par lexema/lema foi calculada a partir das formas ortográficas após a decomposição do morfema de infinitivo [-r] e do morfema de vogal temática [-a|e|o|i-] do lema em relação ao lexema com mesmo número de letras (e.g., fal[es]/fal[ar] = 0, cac[es]/caç[ar] = 1, durm[a]/dorm[ir] = 1, soub[e]/sab[er] = 2).

Sendo assim, pode-se calcular o coeficiente de família morfológica flexional, definido pela relação entre o número de lexemas em função do número de lemas (i.e., CFM = lexemas/lemas) (Baayen, 2008). Observa-se um CFM bastante alto para todas formas encontradas no dicionário ($10.430.109/14.278 = 730$), esse coeficiente alto se deve ao grande número de formas flexionadas (lexemas) por forma infinitiva (lema), pois o corpus completo apresenta todas formas flexionadas de todos tipos de conjugação (i.e., simples, reflexiva e pronominais). Diferentemente, o CFM das formas simples está mais próximo do valor esperado ($925.646/14.278 = 64$), pois: 6 pessoas/número X 12 tempos verbais = 72, considerando-se as formas homógrafas (Villalva, 2007). Em seguida, o CFM das formas irregulares ($25.299/2.109 = 12$) indica que apenas 12 formas

flexionadas de cada infinitivo são irregulares, ou seja, em torno de 18,75% das formas totais da conjugação completa. Enfim, nos verbos supletivos, o CFM ($301/10 = 30$) indica que cerca de 47% das formas são supletivas, ou seja, possui modificação da primeira letra ou modificações em mais de 2 letras do radical (Albright, 2002).

Tabela 1. Número de lemas (Lm) e lexemas (Lx) do dicionário UNITEX/DELAFA.

Classe	Lemas	Lexemas	Lm Sim.	Lx Sim.	Lm Irr.	Lx Irr.	Lm Supl.	Lx Supl.
1ª [-ar]	12.718	9.394.572	12.718	824.594	1.352	16.175	1	43
2ª [-er] ²	839	605.587	839	53.958	471	6.316	6	193
3ª [-ir]	721	429.950	721	47.094	281	2.808	3	65
Total	14.278	10.430.109	14.278	925.646	2.104	25.299	10	301

Destaca-se que de todas as formas analisadas, 14,73% dos lemas apresentam 2,73% de lexemas irregulares; dentre os irregulares, apenas 0,47% dos lemas apresentam 1,19% de lexemas supletivos. Portanto, mais de 85% de todas as formas verbais flexionadas do PB são produzidas a partir de lemas regulares através de regras simples de junção entre o radical regular e os sufixos flexionais (Rocha, 1999). Todavia, destaca-se que muitos verbos frequentes do PB estão entre os menos de 15% de verbos irregulares, sendo assim, cabe realizar-se uma análise das irregularidades dos radicais a partir de uma perspectiva gradual, para enfim estabelecerem-se subclasses.

4.3. Subclasses verbais

A análise das subclasses irregulares que apresentam DOL = 1 com modificações na última vogal ou consoante de ligação final do radical apontou 18 subclasses no PB. Observa-se o que a 1ª conjugação regular em [-ar] é a classe que apresenta menos subclasses (3), mas produtivas com verbos terminados em “-car” com modificações entre “c/qu” e terminados em “-çar” com modificações entre “ç/c” (e.g., educar/eduque, caçar/cace). Esse mesmo padrão produtivo pode ser encontrado de forma invertida na 2ª conjugação em [-er] “-cer” “c/ç” (e.g., escurecer/escureço). Outras subclasses semiprodutivas são as terminações em “-ger” com modificações entre “g/j” da 2ª conjugação e inversamente em “-gir” com modificações entre “g/j” da 3ª conjugação em [-ir] (e.g., proteger/protejo, emergir/emerjo), conforme apresentada na Tabela 2. Portanto, essas modificações na última letra do radical permitem estabelecer as seguintes regras de alomorfia ortográfica simples para os casos apontados: $c/_ar \rightarrow qu/_e|i$, $\varsigma/_ar \rightarrow c/_e|i$, $c/_er|ir \rightarrow \varsigma/_a|o$, $g/_er|ir \rightarrow j/_a|o$, $gu/_er|ir \rightarrow g/_a|o$ (Estivalet & Margotti, 2014).

Em seguida, as subclasses que apresentaram produtividade (mais de 10 lemas) com DOL = 1 e modificações na penúltima letra do radical (alomorfia de vogal) são: -azer (45), -elir (11), -erir (33) e -izar (14) (e.g., fazer/fizer, impelir/impilo, ferir/firo, enraizar/enraízo). Somente a subclasse “-izar” pertence a 1ª conjugação e, por sua vez, apresenta simplesmente uma modificação de acento “i/i” na última vogal do radical.

² Desses, as seguintes formas eram em [-or]: Lemas = 39, Lexemas = 31.746, Lexemas Simples = 2.574, Lemas Supletivos = 1, Lexemas Supletivos = 63.

Tabela 2. Frequências das subclasses verbais que apresentam irregularidades na última letra do lema (infinitivo) em seus lexemas (formas flexionadas).

	[-ar]	[-er]	[-ir]	Total
Vogal	-iar (10)	-oer (8), -uer (8)	-uir (28)	54
Consoante		-ber (5), -cer (381),		
	-car (954),	-der (4), -ger (21),	-cir (2), -dir (13),	
	-çar (366)	-rer (6), -ver (2), -zer (79)	-gir (87), -rir (3), -vir (4)	1927
Total	1330	514	137	1981

Finalmente, a análise das subclasses restante com DOL > 1 apresentou 57 subclasses (1ª conjugação [-ar] = 5, 2ª conjugação [-er] = 30, 3ª conjugação [-ir] = 22), com padrões produtivos (mais que 5 lemas) em: -aber (8), -aver (10), -azer (7) e -guir (14). Portanto, a partir do funcionamento dessas subclasses verbais, as regras de alomorfia do radical podem ser estabelecidas para conjugação de verbos irregulares (Embick & Halle, 2005), assim como para adequação de conjugação de neologismos, estrangeirismos e pseudoverbos do PB (Keuleers & Brysbaert, 2010).

Enfim, os verbos irregulares podem ser classificados em i. supletivos, ii. ajuste da última vogal ou consoante de ligação do radical, iii. alomorfia da última vogal do radical e/ou iv. alomorfia do radical. A análise desenvolvida e algoritmo implementado no presente instrumento apresenta relações diretas com a teoria linguística e estruturas hierárquicas das classes e subclasses verbais (Kilani-Schoch & Dressler, 2005). As subclasses verbais produtivas do PB analisadas aqui são apresentadas na Figura 2.

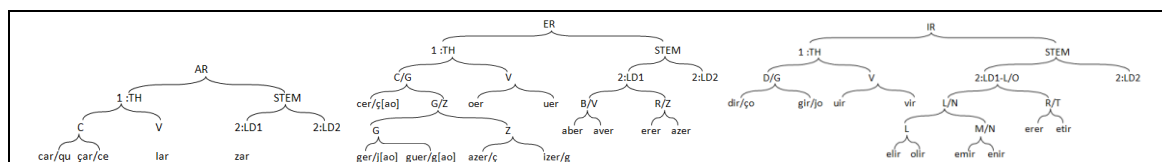


Figura 2. Subclasses verbais produtivas do PB.

5. Considerações finais

Como resultado, o instrumento desenvolvido permite a conjugação completa de todos os verbos do português brasileiro. Além disso, ele permite a conjugação de pseudoverbos que obedeçam às estruturas e regras grafotáticas do português. O instrumento possui como opções as conjugações: simples, reflexiva e pronominal (masc/fem, sing/plur); assim como a apresentação dos rótulos das formas flexionadas. O conjugador verbal do PB está disponível gratuitamente: <https://lexicodoportugues.shinyapps.io/Conjugator/>; o algoritmo em: http://www.lexicodoportugues.com/downloads/lexporbr_conjugator.R/.

A análise das irregularidades morfofonológicas e alomórficas dos radicais verbais do PB permitiu uma melhor compreensão dos fenômenos relacionados a hierarquia e processos alomórficos e ortográficos dos radicais verbais do PB. Espera-se que a análise linguística do sistema verbal, assim como o instrumento aqui desenvolvido, sejam úteis para trabalhos e investigações futuras em torno da descrição e funcionamento do sistema verbal do PB.

Referências

- Albright, A. (2002). Islands of Reliability for Regular Morphology: Evidence from Italian. *Language*, 78(4), 684–709.
- Baayen, R. H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics*. Cambridge: Cambridge University Press.
- Bassani, I. de S., & Lunguinho, M. V. (2011). Revisitando a flexão verbal do português à luz da Morfologia Distribuída: um estudo do presente, pretérito imperfeito e pretérito perfeito do indicativo. *ReVEL*, 199–227.
- Blevins, J. P. (2006). Word-based morphology. *Journal of Linguistics*, 42(03), 531.
- Embick, D., & Halle, M. (2005). On the status of stems in morphological theory. In Geerts, T. & Jacobs, H. (Eds.), *Proceedings of Going Romance 2003*. Amsterdam: John Benjamins, 59–88.
- Estivalet, G. L., & Margotti, F. W. (2014). Diálogos entre a flexão verbal do Português e do Francês (Dialogues entre la flexion verbale du Portugais et du Français). *Estudos da Lingua(gem)*, 12(2), 31–49.
- Estivalet, G. L., & Meunier, F. (2015). The Brazilian Portuguese Lexicon: An Instrument for Psycholinguistic Research. *PLOS ONE*, 10(12), e0144016.
- Fitzgerald, M. (2012). *Introdução às expressões regulares*. São Paulo, Brasil: Novatec Editora.
- Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, 42(3), 627–633.
- Kilani-Schoch, M., & Dressler, W. U. (2005). *Morphologie naturelle et flexion du verbe français* (Vol. 488). Tübingen: Gunter Narr Verlag Tübingen.
- Levenshtein, V. I. (1966). Binary Codes Capable of Correcting Deletions, Insertions, and reversals. *Soviet Physics*, 10(8), 707–710.
- Muniz, M. C. M. (2004). *A construção de recursos linguístico-computacionais para o português do Brasil: o projeto de Unitex-PB*. Universidade de São Paulo.
- R Core Team. (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna/Austria.
- Rocha, L. C. de A. (1999). *Estruturas Morfológicas do Português*. Belo Horizonte, MG: Editora UFMG.
- Vasilévski, V., & Araújo, M. J. (2011). Tratamento dos sufixos modo-temporais na depreensão automática da morfologia dos verbos do português. *LinguaMATICA*, 3(2), 107–118.
- Veríssimo, J., & Clahsen, H. (2009). Morphological priming by itself: A study of Portuguese conjugations. *Cognition*, 112(1), 187–194.
- Villalva, A. (2007). *Morfologia do Português*. Lisboa, PT: Universidade Aberta.