

REFLEXÕES ACERCA DA PRODUTIVIDADE MORFOLÓGICA E DE SUA MEDIÇÃO: ESTUDO DE SUFIXOS NOMINALIZADORES DO PORTUGUÊS

Emanuel Souza de Quadros¹

manuquadros@gmail.com

RESUMO: Este trabalho apresenta uma discussão do conceito de produtividade morfológica e de alguns métodos de medição encontrados na literatura. Focamo-nos especificamente no índice de produtividade proposto por Baayen (1992) e o aplicamos a conjunto de sufixos nominalizadores do português, considerando um corpus de textos escritos retirados de blogs. Os resultados apontaram para o seguinte ordenamento de produtividade, do mais para o menos produtivo: X-ice, X-ismo, X-mento, X-ção, X-idade, X-ura e X-eza. O resultado inesperado nesse ordenamento foi a classificação de X-mento como mais produtivo que X-ção. A partir disso, discutimos algumas possíveis limitações desse método de medição e apresentamos perspectivas futuras de testagem.

PALAVRAS-CHAVE: produtividade; corpus; nominalizadores.

1. INTRODUÇÃO

Neste trabalho apresentamos algumas reflexões sobre a natureza da produtividade morfológica e sobre sua quantificação. Ilustramos algumas dessas questões com um levantamento do uso de sufixos nominalizadores em um *corpus* de língua escrita do português brasileiro.

Na seção 2, discutimos o conceito de produtividade. Na seção 3, tratamos do modo como a produtividade pode ser medida, tendo em vista a relação estreita que há entre o tipo de medição escolhido e o conceito de produtividade adotado. Em seguida, apresentamos nossos procedimentos metodológicos. Na seção 5, trazemos e discutimos os resultados obtidos, comparando-os com outros resultados presentes na literatura. Na seção 6, trazemos nossas considerações finais. E, por fim, na seção 7, fornecemos uma listagem dos dados coletados em nosso *corpus*.

¹ Universidade Federal do Rio Grande do Sul – UFRGS.

2. O CONCEITO DE PRODUTIVIDADE MORFOLÓGICA

A produtividade morfológica pode ser entendida nos termos de Schultink (1961), como “a possibilidade de os usuários de uma língua criarem, de forma não intencional, um número de formações incontável em princípio, por meio de um processo morfológico que serve como base para a correspondência entre forma e significado de palavras já conhecidas”.

Essa definição tem alguns aspectos que merecem ser considerados. Primeiramente, com a exigência de “não intencionalidade”, ela tenta capturar a ideia de que estamos interessados na capacidade intuitiva de os falantes de uma língua criarem palavras novas. Nesse sentido, exclui-se do domínio do estudo da produtividade formações criativas, muitas vezes de interesse estilístico, que podem utilizar recursos que escapam aos mecanismos gramaticais da língua.

Outro aspecto importante é a incontabilidade “em princípio” dessas formações. Padrões morfológicos produtivos são aqueles cujos produtos não se restringem a palavras já existentes na língua. Ainda que se possam listar todas as formas conhecidas que representam um padrão morfológico, p. ex., todas as formas nominais com o padrão X-ista, esse padrão ainda pode ser considerado produtivo, desde que seja possível encontrar uma forma nova que não tenha sido considerada por essa lista. Isto é, não há a possibilidade de listar exaustivamente todos os produtos de um padrão morfológico produtivo. Esses produtos só podem ser definidos intensionalmente, por meio de uma regra de formação de palavras, por exemplo.

Por fim, quando falamos em padrões morfológicos produtivos, falamos em mecanismos de formação de palavras que se utilizam de recursos morfológicos já existentes na língua. Isso é importante, porque a morfologia não é a única responsável pelo enriquecimento do léxico de uma língua. Palavras novas são incorporadas ao léxico por diversas vias, incluindo empréstimos, acronímia, truncamentos, etc. Nem todas essas vias representam padrões morfológicos regulares.

Padrões morfológicos diferenciam-se qualitativa e quantitativamente. Em termos qualitativos, padrões morfológicos podem apresentar diferentes contextos de ocorrência, com diferentes restrições linguísticas ou extralinguísticas. Em termos quantitativos, padrões morfológicos que possuem contextos de ocorrência semelhantes podem se diferenciar quanto a sua rentabilidade, gerando mais ou menos palavras novas.

Assim, a “possibilidade” de formação de palavras novas por meio de um padrão morfológico, mencionada na definição acima, pode ser entendida de duas formas. Pode-se

falar em possibilidade em termos dos contextos linguísticos em que esse padrão pode se manifestar. Por exemplo, os padrões X-idade e X-eza diferenciam-se porque somente bases não derivadas podem participar do segundo (*riqueza, pobreza, etc.*, mas *X-vel-eza, *X-al-eza); o primeiro aceita bases simples (*bondade, sinceridade*) e derivadas (*estabilidade, liberalidade*). Esses padrões diferenciam-se, portanto, porque um deles tem maior número de contextos possíveis em que pode se instanciar. Também se pode falar em diferentes possibilidades de formação no caso de padrões morfológicos que possuem condições gramaticais semelhantes, como X-ção e X-mento, mas que possuem diferentes probabilidades de instanciação. No caso desses dois afixos, observou-se em estudos anteriores que *-ção* ocorre em cerca de 60% de novas nominalizações, ao passo que *-mento* ocorre em cerca de 20% delas (Basílio, 2006). Assim, a possibilidade de que *-ção* forme uma palavra nova, de acordo com esses estudos, é maior do que a possibilidade de que *-mento* o faça.

A observação de que alguns processos tem mais possibilidades de formação do que outros, nesses dois sentidos, gera duas necessidades básicas: definir qualitativamente as possibilidades de formação de cada processo e estimar quantitativamente as possibilidades de uso de cada processo.

A maioria dos estudos em morfologia gerativa concentrou-se na investigação das restrições sobre o uso de padrões morfológicos. Essa investigação de natureza qualitativa gerou uma série de propostas de restrições, tanto específicas a determinados padrões morfológicos quanto gerais, como o bloqueio e a Hipótese da Base Unitária (Aronoff, 1976). A expectativa dos morfólogos dentro dessa tradição é de que a definição das restrições gerais e das restrições específicas sobre os padrões morfológicos possa descrever completamente as possibilidades de formação de palavras. “Assim, o grau de produtividade de uma regra de formação de palavras pode ser visto como inversamente proporcional à quantidade de restrições gramaticais sobre as regras de formação de palavras.” (Booij, 1977).

Outros morfólogos acreditam que mesmo uma definição exaustiva dessas restrições não é suficiente. Para eles, os processos morfológicos contam certa liberdade de variação em uma escala de produtividade. Em parte, essa questão é meramente terminológica. O uso da palavra produtividade é ambíguo. Trata-se da diferença estabelecida por Corbin (1987) entre disponibilidade e rentabilidade.

Mas, em parte, essa questão também tem importância teórica. Ainda que implicitamente, todos parecem reconhecer uma noção de rentabilidade ao lado de uma noção de disponibilidade. No entanto, pode-se entender a rentabilidade de um processo como um fator contingente, resultante da interação entre as restrições inerentes a um processo e suas

condições de uso real – como o fato de haver ou não necessidade comunicativa para o uso de uma formação nova. Ou pode-se entendê-la como algo de certa forma inerente ao processo, que faz parte da competência lexical dos falantes.

3. A QUANTIFICAÇÃO DA PRODUTIVIDADE

Independentemente de qual seja a origem das diferenças de uso entre os processos morfológicos, coloca-se a necessidade de quantificá-las. Uma medida sugerida por Aronoff (1976) é a razão entre o número de palavras produzidas por uma Regra de Formação de Palavras (RFP) e o número de palavras potencialmente produzidas por esta regra, formalizada por Baayen e Lieber (1991), como

$$I = \frac{V}{S}$$

em que I é o índice de produtividade, dado pela razão entre V, o número de palavras atestadas, e S, o número de palavras que poderiam ser formadas em princípio.

Embora essa medida seja conceitualmente interessante, ela é impraticável. Tanto o número de palavras, V, quanto o número de palavras potenciais, S, só são facilmente mensuráveis no caso de processos não produtivos ou muito pouco produtivos. No caso de processos produtivos, V não pode ser computado, pois não há como enumerar as palavras formadas por uma RFP produtiva. Mesmo grandes *corpora* não conseguem representar todas as formações que já foram feitas. RFPs altamente produtivas tem justamente a propriedade de gerar um grande número de formações esporádicas, que podem não ser institucionalizadas. Há dificuldades também para se computar S, visto que o número de bases potenciais de uma RFP também pode ser infinitamente expandido, por meio de RFPs ou de outros mecanismos de enriquecimento lexical.

Outro problema desse índice de produtividade é o fato de que ele não parece nos dizer nada sobre a possibilidade de um padrão morfológico ser usado na criação de palavras novas. O índice I refere-se ao número de palavras já formadas, em relação ao número de palavras potenciais. Trata-se de uma medida do quanto esse padrão já foi utilizado até o momento da contagem, não da possibilidade de ele ser utilizado posteriormente.

Como não há um método de se medir essa possibilidade diretamente, podemos nos aproximar dela por métodos indiretos, baseados na identificação de correlatos da produtividade. Aronoff (1983) apresenta um correlato interessante entre produtividade e frequência de uso. O autor observa que a frequência média de ocorrência de produtos de

processos produtivos é menor que a de produtos de processos não produtivos.

Isso ocorre porque palavras que representam processos não produtivos precisam ser lexicalizadas. A lexicalização, conhecidamente, leva à aquisição de significados especializados pelas construções, tornando-as semanticamente mais complexas. Formas esporádicas, não lexicalizadas, formadas por processos produtivos, tendem a ter semântica mais simples, correspondendo diretamente ao significado composicional definido pelo padrão morfológico. Por hipótese, a complexidade semântica reflete-se na frequência de *tokens*, com o resultado de que palavras que exemplificam padrões pouco ou nada produtivos tendem a ser menos frequentes. Isso implica que podemos usar a frequência de *tokens* de um padrão morfológico como indicativo indireto da complexidade semântica das palavras que o exemplificam e, portanto, da produtividade desse padrão.

Partindo dessa observação, Baayen (1992) sugere um novo índice de produtividade, que considera a frequência de *tokens* dos processos morfológicos. O índice é dado pela seguinte fórmula:

$$\square = \frac{n_1}{N}$$

Em que n_1 é o número de palavras formadas pelo processo relevante que possuem frequência igual a 1 no *corpus* (*hapax legomena*), e N é o número total de ocorrências de palavras formadas por esse processo no *corpus*.

Neste trabalho, propomos uma aplicação desse índice de produtividade a alguns padrões de formação de nomes do português brasileiro, com o objetivo de testar os resultados frente a nossas intuições sobre a produtividade relativa dos padrões morfológicos e de compará-los com os resultados obtidos por outros estudos de produtividade da língua baseados em outros métodos.

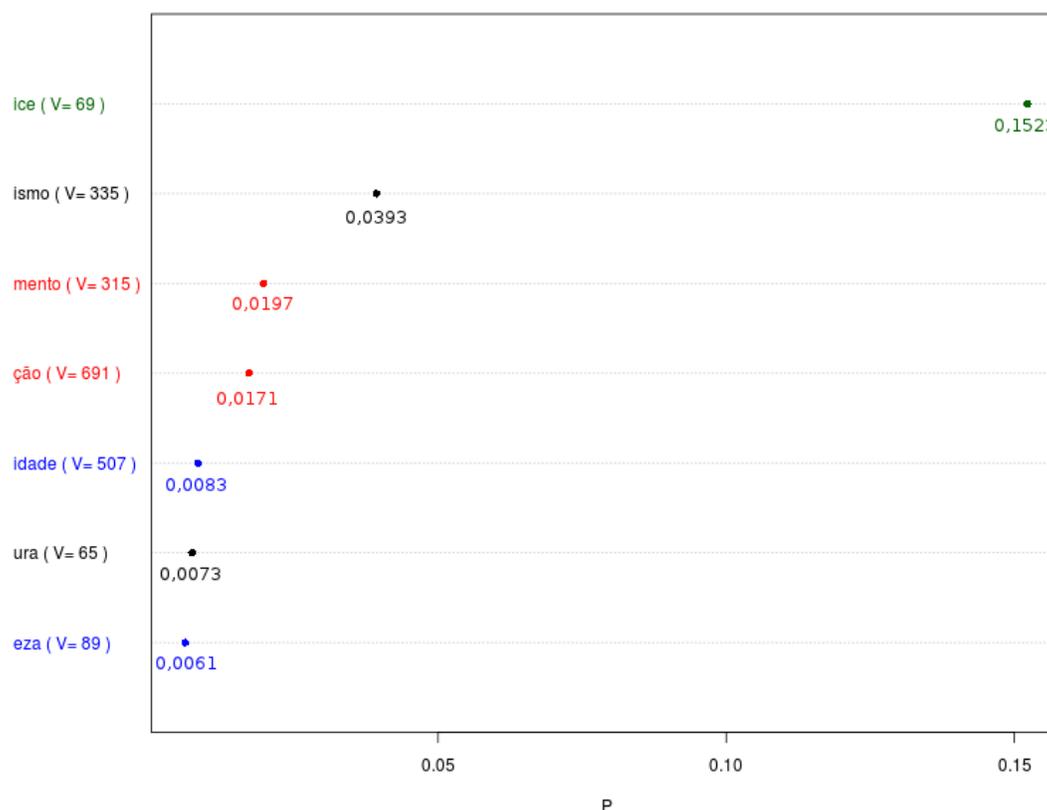
4. PROCEDIMENTOS METODOLÓGICOS

Nosso levantamento foi feito sobre um *corpus* de textos escritos de *blogs* iniciado por Quadros (2009). Atualmente, esse *corpus* apresenta 4.734.532 ocorrências de 134.301 palavras. O levantamento foi realizado com o auxílio de um programa computacional desenvolvido especificamente para a pesquisa e extração de dados deste *corpus*.

Os padrões morfológicos testados foram *X-ice*, *X-ismo*, *X-mento*, *X-ção*, *X-idade*, *X-ura* e *X-eza*.

5. RESULTADOS E DISCUSSÃO

Abaixo, vemos um quadro geral dos índices encontrados para cada um dos padrões considerados. Em cada linha do gráfico, encontra-se o afixo correspondente ao padrão morfológico e seu respectivo índice de produtividade.



O padrão X-ice mostra-se, com larga vantagem, o mais produtivo de acordo com esse índice. Esse resultado diferencia-se do de Sandmann (1988), em que esse padrão não apresentou uma frequência muito alta, contabilizando apenas três formações novas. Essa diferença nos resultados não é surpreendente, dadas as diferenças entre os *corpora* utilizados nesses estudos. Utilizamos aqui dados extraídos de textos de blogs, ao passo que Sandmann (1988) utiliza textos de jornais impressos. Em blogs, embora se trate de uma mídia escrita, a liberdade expressiva dos autores é muito maior do que em um jornal tradicional. Por esse motivo, podem ocorrer contextos como os seguintes, em que a ocorrência de X-ice é bastante motivada.

gando que não há limites para a **retardadice** humana neste sistema solar. Lemb

se flagra é mais um flagra de **pão-durice** do que de gambiarra, quem nos enviastei da exposição paulista, a **porralouquice** genial do artista estava bem disrica, tais como, aos 30 anos de **marmanjice**, ainda morar com os pais. (Mas

O resultado mais surpreendente neste estudo é o ordenamento entre os índices de produtividade de *X-mento* e *X-ção*. Esse resultado não se conforma aos obtidos em outros estudos que compararam esses dois padrões morfológicos. Em Sandmann (1988), a diferença entre os dois afixos também foi pequena, mas *-ção* apresentou 28 novas formações, ao passo que *-mento* apresentou 26.

No estudo de Basilio (1996), sobre dados de língua falada do projeto NURC, foram encontradas 218 ocorrências de *-ção* e 73 de *-mento* – um uso três vezes maior do padrão *X-ção*. No estudo de Grodt (2009), sobre dados de fala do projeto VARSUL, foram encontradas 314 ocorrências de *-ção* e 195 de *-mento* – um uso 1,6 vezes maior do padrão *X-ção*. No entanto, esses resultados derivam de levantamentos do uso de formas regulares sufixadas por *ção* ou *mento*, independentemente de elas serem formações novas ou não. Em contraste, o índice utilizado em nosso estudo considera o número de *hapax legomena* que portam esses sufixos; por hipótese, isso nos dá uma medida da possibilidade de haver palavras novas com esses sufixos. Em outras palavras, podemos colocar a pergunta básica das medidas utilizadas nos estudos de Basilio (1996) e Grodt (2009) nos seguintes termos “Qual é a a frequência de uso desses padrões morfológicos?”; por outro lado, a pergunta básica do índice utilizado neste estudo é “Qual é taxa de crescimento do vocabulário desses padrões morfológicos; ou qual é a probabilidade de que um novo uso desse padrão morfológico revele uma palavra nova?” É esperado que as respostas a essas perguntas possam diferir um mesmo padrão morfológico.

Quando perguntamos sobre a frequência de uso dos padrões morfológicos em questão no nosso corpus, obtemos resultados que já não são tão distintos dos de Basilio (1996) e Grodt (2009). Encontramos 315 palavras sufixadas por *-mento* e 691 sufixadas por *-ção* – 2,2 vezes mais palavras em *-ção*. As palavras em *-mento* apresentaram 5472 ocorrências, e as em *-ção*, 10655 ocorrências – 1,95 vezes mais ocorrências de palavras em *-ção*. O fato é que a proporção de “palavras novas”, isto é, de *hapax legomena*, no universo de palavras formadas por *-mento* se mostrou maior do que no de palavras formadas por *-ção*. Não há informações sobre essas proporções nos estudos de Basilio (1996) e Grodt (2009) que nos permitam fazer uma comparação mais adequada entre esses estudos e o que apresentamos aqui.

Devemos considerar que nossos resultados não indicam necessariamente que há mais palavras novas sendo formadas com *-mento* do que em *-ção*. Como vimos, a medida de

produtividade que utilizamos é indireta e pode ser influenciada por diversos fatores. Ela depende da premissa de que *hapax legomena* costumam ser palavras novas. No entanto, outras palavras de baixa frequência, como arcaísmos e palavras formais, podem compor o conjunto de *hapax legomena*. Se o padrão *X-mento* apresentar uma distribuição maior desses tipos de palavras, ele pode apresentar um número maior de palavras de frequência 1, sem que isso revele formação de palavras novas.

Outra consideração importante é que o índice de Baayen (1992) é bastante sensível ao tamanho do *corpus* considerado. Ele é baseado na ideia de que palavras já existentes tendem a se repetir em uma amostra, ao passo que palavras novas têm natureza mais esporádica. Contudo, o *corpus* precisa ser suficientemente extenso para que as palavras existentes tenham chances de se repetir. Quanto maior for o *corpus*, maior será a proporção de palavras genuinamente novas no conjunto de *hapax legomena* e menor será a proporção de palavras já existentes que casualmente têm frequência baixa na variedade linguística considerada. Portanto, se o resultado observado entre *-mento* e *-ção* expressa um fato verdadeiro sobre produtividade relativa desses afixos, espera-se que *-mento* continue apresentando um índice maior do que *-ção* mesmo com um aumento do *corpus*.

Da mesma forma, espera-se que, com um aumento do *corpus*, a distância observada entre os índices de *X-idade* e *X-eza* também aumente. Sabemos que *X-eza* restringe-se a bases não derivadas. Como o vocabulário de bases não derivadas cresce de maneira lenta na língua, espera-se que haja poucas oportunidades de novas atestações desse padrão morfológico, ao passo que as atestações já feitas devem se repetir e deixar de ser *hapax legomena*. *X-idade*, por sua vez, apresenta distribuição menos restrita, podendo gerar palavras novas e, conseqüentemente, mais *hapax legomena*, a partir do conjunto sempre crescente de bases não derivadas.

Nota-se, então, que os resultados obtidos neste estudo carecem de confirmação, seja por meio do aumento progressivo do *corpus* utilizado, seja pelo emprego de métodos, ainda a serem investigados, que nos deem uma medida mais direta do número de palavras novas de cada padrão morfológico.

6. CONSIDERAÇÕES FINAIS

Fica clara neste trabalho, e na comparação deste com outros trabalhos presentes na literatura, a dificuldade de se identificar uma palavra nova de maneira objetiva. Se entendemos produtividade como a possibilidade de haver formação de palavras novas por

meio de recursos existentes na língua, é crucial que haja uma maneira de detectar esse tipo de formação para que se possa caracterizar a produtividade de um padrão morfológico.

Os resultados que obtivemos com a aplicação do índice proposto por Baayen (1992) apontam para uma altíssima produtividade do padrão morfológico *X-ice* ($P = 0,1523$), que é seguido por *X-ismo* ($P = 0,3993$), *X-mento* ($P = 0,0197$), *X-ção* ($P = 0,171$), *X-idade* ($P = 0,0083$), *X-ura* ($P = 0,0073$), *X-eza* ($P = 0,0061$). O ordenamento desses padrões, de forma geral, está de acordo com nossas intuições, o que indica que o índice considerado está próximo de uma medida de produtividade satisfatória. O único resultado que nos surpreendeu foi o ordenamento entre *X-mento* e *X-ção*, que deve ser investigado em estudos futuros.

7. LISTA DE HAPAX LEGOMENA

Seguem os *hapax legomena* encontrados em cada um dos padrões morfológicos.

X-ice

baitolice, birutice, branquelice, cachorrice, cacurice, cafonice, calhordice, charlatanices, cornice, corujices, credices, culinaries, danguices, esdruxulice, filosofices, frescurices, gabolices, gaiatice, gordice, manezice, marmanjice, mineirices, mongolice, mosconice, mulequice, pão durice, picaretice, porralouquice, putice, rameirice, ranzinzice, rapadurices, retardadice, sem-gracice, tagarelice, tontice, vigarice

X-ismo

achismo, alicismo, amarelismo, apoliticismo, aristocracismo, artificialismo, autodidatismo, barbarismo, bicameralismo, blueprintismo, botulismo, bovarismo, bruxismo, cabotinismo, calvinismo, cateterismo, charlatanismo, clubismo, coletivismo, coloquialismo, conceitualismo, construtivismo, corinthianismo, cosmopolitismo, crossfoxismo, cubismo, cutismo, denominacionalismo, derrotismo, desconstrutivismo, despotismo, determinismo, didatismo, diletantismo, dinossaurismo, dualismo, eclesiasticismo, egoismo, eleitoralismo, empirismo, empreguismo, entreguismo, ermitismo, escravismo, esnobismo, espiritismo, estilismo, exclusivismo, exotismo, fanboysismo, fashionismo, financismo, folclorismo, fordismo, gangsterismo, gatismo, gayzismo, getulismo, gramscismo, hacktivismo, hedonismos, hermafroditismo, hermetismo, humanitarismo, imobilismo, imoralismo, infantilismo, intimismo, laicismo, lobismo, luteranismo, magnetismo, malufismo, mercantilismo, militarismo, modelismo, municipalismo, mutismo, negrismo, onirismo, panfletarismo, parlamentarismo, passionalismo, paternalismo, patrimonialismo,

pentecostalismos, periodismo, peronismo, pieguismos, praticismo, preferencialismo, presidencialismo, priapismo, primitivismo, propagandismos, punhetismo, puritanismo, ranzinzismos, revanchismo, romancismo, sabujismo, satanismo, separatismo, sessentismo, socorrismo, sonambulismo, templarismo, tenentismo, teoricismo, tradicionalismo, transformismo, triunfalismo, tropicalismo, trotskismo, utilitarismo, vaginismo, vanguardismo, vegetarianismo, vigilantismo, vitimismo

X-mento

acatamento, acobertamento, acomodamentos, acondicionamento, acotovelamento, afetamentos, afilamento, afogamentos, afrontamento, agendamento, agrupamento, aniquilamento, aprisionamento, arrendamento, asfaltamento, atilamento, atochamento, atordoamento, aviltamento, avistamentos, avivamento, balanceamento, barateamento, batimento, beneficiamento, branqueamento, bronzeamento, cadastramento, chamamento, chaveamento, confrontamento, congelamento, contentamento, corrimento, derramamentos, desarquivamento, descolamento, desentupimento, deslumbramento, despojamento, destelhamento, destrinchamento, dimensionamento, doutrinamento, embasamento, embotamento, emburrecimento, empalamento, empoderamento, enamoramento, encantamentos, encapsulamento, encarceramento, encascalhamento, enceramento, encolhimento, encorajamento, enfraquecimento, engessamentos, engorduramento, engrandecimento, enlaçamento, entrelaçamento, entupimento, envenenamento, escalonamento, escoamento, escoramentos, espancamentos, espelhamento, esquartejamento, falseamento, fardamento, favorecimento, florescimento, auto-fodimento, indiciamento, insistimento, instruimento, juramentos, ligamentos, melhoramentos, merecimento, nivelamento, patenteamento, presentimento, provimento, rankeamento, reaquecimento, recapeamento, recrudescimento, reequipamento, religamento, repareamento, repovoamento, resfriamento, ressecamento, ressurgimento, retardamento, rolamentos, sitiamento, sucateamento, televisionamento, tombamento, trackeamento, travamentos, versionamentos, zoneamento

X-ção

absorção, acomodações, adequação, afobação, agitação, alienação, alternância, amarração, ambientação, anexação, aniquilação, apelação, apreciação, audição, autolocalização, automação, autuações, averbação, babação, bajulação, barração, beijação, capitulações, carbonatação, castração, coação, comiseração, compressão, conclamação, concretização, confrontação, congelação, congratulações, congregações, constelações, constipação, contenção, contrações, contraposição, cremação, cristalização, crucificação,

declinação, decomposição, decretação, defecação, defecção, deflagração, deformações, degradação, degustação, deliberações, demolição, deposição, depuração, desaprovação, desatenção, desclassificação, desertificação, desfiliação, designação, desmotivação, desproteção, desvinculação, deterioração, diplomação, dissecação, dissertação, doutrinação, dramatizações, duplicação, ejaculação, elevação, elocubrações, emulação, encadernação, enjambração, entonações, entubação, equiparação, erradicação, escavações, estimagtização, exclamação, expropriação, facilitação, fechação, felicitação, figuração, forçação, formalização, fornicação, gestação, gongação, habitação, improvisação, incomodação, infecção, informatização, iniciação, injeção, insinuações, instauração, interceptação, interdição, intuição, irrigação, judiação, judicialização, lamentação, luxação, medicação, otávio-mesquitização, ministração, multiplicação, nomeação, objetivação, obtenção, ordenações, orkutização, pacificação, pagação, penalização, perdição, peregrinação, personificação, pichações, pirações, poluição, processualização, procuração, profissionalização, progressão, propiciação, protelação, quadrinização, racionalização, rasgação, reanimação, reapresentação, reavaliação, recomposição, reconfiguração, regeneração, regionalização, regulamentação, regularização, reimplementações, reinauguração, reinterpretações, reinvenção, relativização, remarcação, remuneração, renegociação, rentabilização, reposição, republicação, requalificação, restauração, retransmissão, retribuição, reutilização, revalidação, revogação, secreção, sedação, segmentação, segregação, sobrevalorização, sonorização, suavização, subordinação, terminações, torção, transliteração, transpiração, tributação, vacilação, validação, vedação, vinculação, viralização

X-idade

afabilidade, afetuosidade, aleatoriedade, amabilidade, amorosidade, aplicabilidade, assertividade, autossustentabilidade, baianidades, balneabilidade, benignidade, bipolaridade, bissexualidade, brasilidade, celestialidade, comercialidade, confidencialidade, conflitualidade, cosmetividade, descontinuidade, desproporcionalidade, dirigibilidade, disfuncionalidade, dispensabilidade, engenhosidade, escabrosidades, escalabilidade, escuridade, esterilidade, eternalidade, exiguidade, expressividade, fecundidade, ficcionalidade, frivolidade, furtividade, gauchidade, generalidades, geniosidades, genuinidade, graciosidade, habitualidade, heteronormatividade, hodiernidade, horizontalidade, ilegitimidade, imobilidade, impessoalidade, imponderabilidade, impopularidade, imputabilidade, inabilidade, inamovibilidade, infalibilidade, inflexibilidade, insalubridade, instantaneidade, intangibilidade, inteligibilidade, intransponibilidade, inumanidade, irreduzibilidade,

irreversibilidade, lascividade, legibilidade, magnanimidade, manobrabilidade, massividade, materialidade, mediandade, melosidade, mendacidade, minoridade, mortandade, multifuncionalidade, natividade, necessidade, nordestinidade, oficialidade, opressividade, orfandade, paranormalidade, partenidade, pateticidade, pecaminosidade, perenidade, permeabilidade, pessoalidade, polaridade, pontualidade, positividade, prematuridade, prioridade, proatividade, prolixidade, puerilidade, radicalidade, reflexividade, regionalidade, resolutividade, responsividade, ridicularidade, rugosidades, sacralidade, sensitividade, sensorialidades, sentimentalidade, severidade, sincronicidade, socialidade, subsidiariedade, teatralidade, tecnicidades, tecnicidade, tensidade, territorialidade, tortuosidade, transitoriedade, ventosidade, verticalidade, vivacidade

X-ura

apertura, assaduras, atadura, brancura, brandura, cavalgaduras, chatura, crucifatura, fervura, finura, formosura, gramatura, levedura, nervura, pisaduras, planura, prematuras, proposituras, singradura, ternuras

X-eza

afositeza, alteza, aspereza, bambeza, crudeza, estreiteza, indelicadezas, justeza, pequenenezas, torpezas

Referências bibliográficas

1. ARONOFF, Mark. *Word Formation in Generative Grammar*. Cambridge, Mass.: MIT Press, 1976.
2. ARONOFF, Mark. Potential words, actual words, productivity and frequency. In: *Proceedings of the 13th International Congress of Linguistics*, 1983. p. 163-171.
3. BAAYEN, R Harald. On frequency, transparency, and productivity. In: Booij, G. E. e van Marle, J. (eds), *Yearbook of Morphology 1992*, Dordrecht: Kluwer Academic Publishers, 1992. p. 181-208.
4. BAAYEN, R Harald; LIEBER, Rochelle. Productivity and English derivation: a corpus-based study. *Linguistics*, v. 29, p. 801-844, 1991.
5. BASILIO, Margarida. Formação e uso da nominalização deverbal sufixal no português falado. In: Castilho, A. T. e Basilio, M. (orgs), *Gramática do Português Falado*, v. 4., Campinas: Editora da UNICAMP, 1996. p. 23-33.
6. BASILIO, Margarida. *Formação e classes de palavras no português do Brasil*. São Paulo: Contexto, 2006.

7. BOOIJ, Geert. *Dutch morphology: a study of word formation in Generative Grammar*. Dordrecht: Foris, 1977.
8. CORBIN, Danielle. *Morphologie dérivationelle et structuration du lexique*. Tübingen: Niemeyer, 1987.
9. GRODT, Aline. *Um estudo sobre produtividade derivacional no português falado no sul do Brasil*. 92 f. Dissertação de Mestrado - Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2009.
10. QUADROS, Emanuel S. *A estrutura e o uso da parassíntese no português*. 2009. 54 f. Monografia (Trabalho de Conclusão do Curso de Letras) - Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2009.
11. SANDMANN, Antônio J. *Formação de Palavras no Português Brasileiro Contemporâneo*. Curitiba: Scientia et Labor, 1988.

ABSTRACT: This paper discusses the notion of morphological productivity and some methods for measuring productivity found in the literature. We focus on the index of productivity proposed by Baayen (1992) and apply to a sample of nominalizing suffixes in Portuguese, given a *corpus* of written language extracted from blogs. Our results show the following ranking of the suffixes in terms of productivity, from the most to the least productive: X-ice, X-ismo, X-mento, X-ção, X-idade, X-ura e X-eza. The surprising result in this ranking was the positioning of X-mento as more productive than X-ção. From this results, we discuss some possible limitations of this method of measuring productivity and point to future ways of testing.

KEYWORDS: productivity; corpus; nominalizers.