

A PRODUTIVIDADE DAS REGRAS DE FORMAÇÃO DE PALAVRAS

Maria Carlota ROSA
Universidade Federal do
Rio de Janeiro
e Alina VILLALVA
Faculdade de Letras de
Universidade de Lisboa

RESUMO

O conceito de produtividade das regras de formação de palavras tem sido apresentado e discutido por vários autores (cf. Aronoff 76, Basilio 80, Bauer 83, Scalise 84).

Independentemente da diversidade das teses defendidas, a produtividade das regras de formação de palavras tem sido estudada com base no conhecimento de falantes de uma determinada língua.

Neste trabalho pretendemos discutir a pertinência e a aplicabilidade deste conceito relativamente a sincronias passadas, às quais só é possível aceder através de textos escritos.

ABSTRACT

Productivity is a central point in morphological studies. As Aronoff (1976: 35) points out, "Though many things are possible in morphology, some are more possible than others".

Although this claim is not contested and some word-formation processes are felt as clearly productive, productivity is generally taken fairly intuitively, even in theoretical works.

The topic of this discussion is the extent to which this concept can be used. Synchronic and diachronic data are included in this paper.

0. O conceito de produtividade é recorrente nos estudos morfológicos e relaciona-se com a forma como são concebidos os mecanismos de formação de palavras. Mas este mesmo conceito, que tão frequentemente é utilizado como um argumento fundamental, é sempre tomado de forma um tanto intuitiva. Compreende-se, assim, a razão pela qual Suzanne Romaine (1983:181) afirma que o termo produtividade é usado "num sentido pré-teórico pelos gramáticos generativistas".¹

Pretendemos, aqui, apresentar alguns problemas suscitados pela utilização do conceito de produtividade que, em última análise, permitem questionar a pertinência da utilização do próprio conceito.

Assim, começaremos por apresentar as Regras de Formação de Palavras; em seguida discutiremos o conceito de produtividade destas regras; finalmente, faremos referência à noção de produtividade numa perspectiva diacrónica. Restringir-nos-emos a dados derivacionais.

1. As Regras de Formação de Palavras

É em Halle (1973) que se encontra a inovação formal para o estudo da formação de palavras, considerada por alguns

"O mecanismo que tem sido a força orientadora e unificadora da pesquisa no campo da morfologia, no âmbito do modelo generativo, nos últimos dez anos."
(Scalise, 1984:27)²

Esta inovação - as Regras de Formação de Palavras (RFPs) - consiste na explicitação de mecanismos que possibilitam gerar palavras passíveis de serem acumuladas no Dicionário.³ Tanto em Halle (1973) como em desenvolvimentos posteriores (como Aronoff (1976)), as RFPs devem especificar a categoria sintáctica da base a que determinado afixo pode ligar-se, a categoria sintáctica do elemento gerado, a estrutura interna e a informação semântica regular.

1.1 - As RFPs em Halle (1973)

Na proposta apresentada em Halle (1973), a distinção entre formas existentes, formas não existentes e formas agramaticais é feita por meio do traço [Inserção Lexical], uma vez que as RFPs, nessa proposta, não contêm condições ou restrições. Um Filtro, colocado entre a Componente de Regras de Formação de Palavras e o Dicionário, permitirá que tenham acesso a este apenas as formações existentes. O recurso ao filtro, no entanto, faz com que a questão da produtividade não se apresente, permanecendo inexplicados os motivos que tornam determinados processos mais produtivos que outros.⁴

1.2 - As RFPs em Aronoff (1976)

Aronoff retoma a noção de possível, mas não existente de Halle (1973), introduzindo, porém, conjuntos de condições para a operação das RFPs - condições positivas e condições negativas.

As condições positivas explicitam as classes de bases a que determinado afixo pode unir-se. Encontra-se em Aronoff (1976:63) uma exemplificação de condições positivas na série que constitui a parte (b) da regra de edição do prefixo inglês *un*, que reproduzimos em (1):

(1) Regra de un# negativo

a. $[X]_{Adj} \rightarrow [un\# [X]_{Adj}]_{Adj}$

interpretação semântica (grosso modo): $un\# X = not X$

b. Formas da base

1. $X_V \# en$ (en=participio passado)

2. $X_V \# ing$

3. $X_V \# able$

4. $X+Y$ (worthy)

5. $X+ly$ (seemly)

6. $X \# ful$ (mindful)

7. $X+al$ (conditional)

8. $X \# like$ (warlike)

As oito condições seguem parâmetros de produtividade, como se pode ver pelo comentário que segue a apresentação da regra:

"Certamente, cada uma delas [i.e., das condições positivas] terá algum índice de produtividade e coerência e si associado. A lista é dada, grosso modo, na ordem da produtividade..." (Aronoff 1976:63).

Podemos interrogar-nos, contudo, acerca dos critérios que levaram ao estabelecimento desse "algum índice de produtividade" quanto à aplicação do prefixo aos diferentes tipos de base e que

legitimam a hierarquia apresentada por Aronoff. A resposta parece ser a coerência semântica; no entanto, ainda na mesma página, Aronoff afirma que produtividade e coerência semântica "são, num certo sentido, a mesma coisa".

O outro conjunto de condições - as condições negativas - permite caracterizar as bases sobre as quais uma dada RFP não opera. Constatam, pois, a improdutividade de um afixo em relação a determinadas bases. Transcrevemos em (2), abaixo, um exemplo de condição negativa, tirado de Aronoff (1976:55):

(2) $X]_N - al]_A$

Condição: $X\# [Y]_{Vment}$

A afirmação contida em (2) é a de que o afixo inglês - al não forma novos adjectivos a partir de bases nominais em - ment derivadas de verbos.

Do conjunto das condições negativas, Aronoff destaca um tipo especial, a que dá o nome de bloqueio, i.e., a impossibilidade de que estejam "no léxico de uma única pessoa, ao mesmo tempo, duas palavras com o mesmo significado e o mesmo radical" (Aronoff, 1976:56). Por outras palavras, o bloqueio "é a não ocorrência de uma forma em virtude da simples existência de uma outra" (Id. et ib.: 43). O bloqueio, no entanto, só existe para palavras que vieram ocupar o lugar de outras que são de algum modo arbitrárias e, por essa razão, entraram no léxico (id. et ib.: 45). Processos produtivos (como a adição de *ness* a bases adjectivas com a forma *Xous*) nunca sofrem a acção do bloqueio, porque "não achamos qualquer razão para listar" as palavras assim formadas (Id. et ib.: 45).

2. A Produtividade

Os processos derivacionais são processos de formação de palavras cuja "irregularidade" é mais aparente que real.

A dificuldade sentida na captação das regularidades, ou seja, na formulação das RFPs, esteve na base da identificação de vários graus de produtividade. Mas os critérios em que essa distinção assentava nunca foram claramente enunciados.

Afirmações como a de Chomsky (1965:278) - "Os processos de derivação apresentam problemas muito maiores para qualquer tipo de gramática generativa (isto é, explícita) do que os sistemas de flexão. Isto resulta do facto de serem tipicamente esporádicos e apenas semi-produtivos"⁵ - que, de algum modo, estigmatizaram a morfologia no interior da gramática generativa, vulnerabilizando, conseqüentemente, o próprio modelo, estão longe de corresponder ao estado actual dos conhecimentos e às convicções teóricas dos linguistas que, no quadro desse mesmo modelo, trabalham sobre formação de palavras.

Em Aronoff (1976), este estado de coisas altera-se. Defende o autor que o conceito de produtividade só faz sentido quando aplicado a regras concorrentes, ou seja, a regras que seleccionam um mesmo conjunto de bases e geram um mesmo tipo morfológico de palavras.⁶ Pode entender-se, então, que está em causa, aqui, um conceito de produtividade relativa. Esta posição é também defendida por Suzanne Romaine (1983), que, no entanto, refere a necessidade de não considerar apenas condições morfológicas na definição de regras concorrentes.⁷

Mas, e de novo em Aronoff (1976), é possível encontrar uma referência a um outro tipo de produtividade. Neste caso, apenas uma regra é considerada, sendo a sua produtividade inversamente proporcional ao número de condições que lhe estão associadas.⁸ O que agora parece estar em jogo é um conceito de produtividade absoluta.

Esta distinção não basta, contudo, para definir produtividade. E essa definição acaba sempre por ser dada, mais ou menos explicitamente, por intermédio da noção de frequência.⁹

O número de palavras que uma regra permite formar e aquele que essa mesma regra efectivamente forma, ou melhor, formou até um dado momento, não são certamente coincidentes, pois, se o fossem, essa regra não permitiria gerar nenhuma outra palavra.

Desta constatação resultou a distinção entre palavras possíveis e palavras existentes. E desta distinção resulta um novo

problema, que consiste em saber o que são palavras possíveis e o que são palavras existentes.

Admitindo que palavras possíveis são todas aquelas que uma dada regra permite gerar, desloquemos a discussão para o lado das palavras existentes, onde, pelo menos, dois critérios de delimitação se confrontam: por um lado, o da contagem das palavras dicionarizadas e, por outro, o dos juízos dos falantes nativos.

O primeiro critério, ou seja, a contagem das palavras dicionarizadas, não parece ser plenamente aceitável. Sem esquecer a diferença entre um dicionário, registo escrito das palavras efectuado por falantes, e o léxico, componente da gramática, deve considerar-se o facto de as palavras formadas por regras não apresentarem idiosincrasias que exigem a sua dicionarização: em princípio, num dicionário do português, não são registadas formas derivadas onde intervêm sufixos diminutivos, nem formas flexionadas, a menos que algum processo de deriva semântica tenha interferido. Por outro lado, o registo num dicionário é um processo mais lento do que o de formação de palavras, o que, inevitavelmente, dá origem a um desfasamento.

Quanto ao segundo critério, ou seja, a consideração dos juízos dos falantes nativos, talvez se possa pensar no que resultaria de um inquérito sobre a existência de palavras apresentadas em(3), retiradas do corpus do Português Fundamental:

- (3)
1. analfabética
 2. coitadinhita
 3. ciclista
 4. para-brisas

ou sobre a aceitabilidade de hipóteses alternativas do tipo das que surgem em (4):

- (4)
1. ordenação/ordenamento
 2. congelação/congelamento
 3. variação/variabilidade

Assim, e a menos que se proceda por amostragem, não é possível identificar a produtividade de uma regra com base no número de palavras por ela formadas.

No entanto, a produtividade não deixa de ser algo de bastante intuitivo. Por exemplo: pode-se dizer que, em português, a formação de advérbios de modo em *-mente* é sentida, sem dificuldades, como produtiva. Constatações como essa, porém, não implicam que a produtividade deva ser considerada um "conceito" ou um "instrumento teórico" em morfologia generativa, mas sugerem a sua aceitação como um índice (Bauer 1983:88).¹⁰

Assim, maior produtividade absoluta indicia uma regra com menor número de condições e, inversamente, menor produtividade absoluta é indicadora de uma regra fortemente condicionada.¹¹

Por último, e admitindo que a existência de regras concorrentes é o resultado de uma situação de conflito que caracteriza o processo de evolução de um dado sistema lingüístico, a produtividade relativa pode ser sentida como um índice do sentido dessa mesma evolução.¹²

3. A produtividade em Estudos Diacrónicos.

Vimos que produtividade, em sua "utilização pré-teórica", pode apresentar problemas no estudo de sincronias presentes. Tentaremos demonstrar que, ao serem focalizados estágios de uma língua pertencentes ao passado, também nos estaremos debruçando com dificuldades. Para isso utilizaremos um exemplo retirado do português. Vamos propor a seguinte questão: "De que modo avaliar a veracidade de uma assertiva, como a que se encontra em Machado (1977), que nos apresenta *-udo* como um sufixo que "teve grande vitalidade no português arcaico"?

Ao lado da terminação *-udo* para a derivação de adjectivos denominais, havia, já desde o latim vulgar, a formação de participios passados fracos da segunda conjugação em *-ūtu-*, de que proveio a terminação *-udo*¹³ do participio passado de segunda conjugação em

português: sabudo, perdudo, evudo, vinudo e venudo (a par de vlido), conoçudo (a par de conocido). Por volta do século XVI (Williams, 1938:188-9), já a segunda conjugação portuguesa substituiu a terminação de participio passado em -udo por -ido.

No entanto, muitos "participios passaram a adjectivos: por exemplo...., escondudo 'escondido, secreto', espendudo 'espalhado', teudo 'tido, obrigado'" (Huber, 1933:213). Não há necessidade de se comprovar a produtividade da adição da -udo na formação de participios. Trata-se de um processo flexional e, enquanto flexão, produtivo por natureza.

Vamos restringir-nos a -udo formador de adjectivos denominais em casos como sanhudo e siudo.

Para comprovar-se a veracidade da afirmação de Machado, podemos elaborar uma lista de formas que efectivamente existirem, através de consulta a dicionários.

Machado (1977) fornece-nos as seguintes palavras em -udo¹⁴ com registos de entrada nos séculos XII e XIII:¹⁵

- (5)
1. barbudo - 1132
 2. sesudo - séc. XII
 3. siudo - séc. XIII
 4. cornudo - 1209
 5. cabeçudo - 1220
 6. sanhudo - séc. XIII

Se os verbetes fornecidos pelo próprio Machado são um reflexo, ainda que não exaustivo, da história do léxico português, a entrada de 6 elementos em -udo - ou de 5, no caso de contarmos sesudo e siudo como uma única palavra - ao longo de dois séculos é indicio de que -udo era parte de uma RFP com algum grau de produtividade.

Analisando o corpus que Machado (1977) nos fornece, encontramos nomes referentes a partes do corpo e ao que, grosso modo, poderíamos considerar "características do comportamento humano", tais como sanha e siso, como bases para essas formações. As entradas em séculos posteriores - e a leitura de Machado (1977) sugere um crescendo, com os séculos XVIII e XIX detendo o maior número de entradas em -udo (cf. nota 13) - parece confirmar essa hipótese. Podemos, contudo, diante de um corpus tão diminuto, tirar conclusões com um mínimo de segurança? Em outras palavras: seriam quaisquer nomes de partes do corpo, acrescidos do sufixo -udo, palavras possíveis para os falantes da época, embora não tenham vindo a concretizar-se como palavras existentes?

O recurso a dados de séculos posteriores, ainda que possível, deve ser feito com cautela, uma vez que podem ter sido levadas a efeito alterações quanto ao tipo de base privilegiada. Parece-nos, também, que não há como estabelecer, por exemplo, se alguma dessas formas apresentava, na época, alguma idiossincrasia. Se é difícil estabelecer a medida em que -udo era produtivo, também é difícil estabelecer que regras (se as havia) concorriam com a adição de -udo.

Não há, evidentemente, como proceder ao levantamento de dados por meio de inquéritos junto a falantes. Temos acesso apenas a textos escritos, que nos proporcionam poucos dados. E supomos que esses dados, se pudessem ser confrontados com as respostas de falantes a questionários, fornecer-nos-iam respostas nem sempre coincidentes com aquelas que os textos trouxeram até nós, tal como sucede na época actual.

NOTAS

1. Em nenhum outro quadro teórico a produtividade é explicitada; a restrição de Romaine quanto aos generativistas deve-se ao facto de, neste modelo, ser concedida atenção fundamental aos processos de formação de palavras.
2. Esta opinião não é compartilhada por todos os linguistas. Jensen (1987: 206), e.g., considera demasiado o entusiasmo de Scallise pelas RFPs.
3. Em Halle (1973), a determinação do Dicionário será feita pela lista de Morfemas, pelas RFPs e pelo Filtro. Para Aronoff (1976:43), será o lugar para "todas e apenas aquelas palavras que são excepcionais, i.e., arbitrárias em pelo menos um dos seus vários traços..."
4. Cf. Lieber (1983:256) para críticas quanto ao recurso a filtros como solução para problemas.
5. Na tradução portuguesa de Chomsky (1965), a expressão *semi-produtivos* é objecto de uma nota explicativa, onde se pode ler "Isto é, grande parte dos elementos da língua não se submete de modo regular a estes processos" (p. 278).
6. Aronoff (1976:37): "Our method will be to compare in some detail two WRFs which we know to differ in productivity... we therefore must take two rules which operate on the same base and have outputs of the same lexical category and subcategorization."
7. Romaine (1983:182): "Aronoff draws the conclusion that there is

no way to say that one word formation rule is more productive than another without taking into account the morphology of the base.

I do not dispute this point; in fact, I can add some support to it..."

Idem: 188: "there are other limitations on the productivity of bases which are phonological."

Idem: 192: "Another important factor in productivity is meaning. I use the term in its broadest possible sense to refer not just to referential meaning, or semantics, but also to pragmatic and syntactic aspects of meaning."

8. Aronoff (1976: 36): "A WFR whose general productivity is high will have few morphological restrictions on the class of bases to which it attaches."
9. Embora seja claramente rejeitada: Scalise (1984: 157): "the productivity of a WRF cannot be identified simply on the basis of the frequency with which it applies. It is also necessary to take into consideration the morphological restrictions on the words that can be taken as the base of a given WFR". Também Aronoff (1976: 36).
10. Bauer (1983: 88): "Thus there is an inverse relationship between productivity/lexicalization such that the most productive patterns are not lexicalized processes, and fully lexicalized processes are not productive. This leads to the view... that productivity is not an either/or phenomenon, but presents a cline."
11. Cf. nota 8.
12. Clark & Berman (1984).
13. Para Lausberg (1966: 390), as formas em -udo de participio passado, seriam, na Península Ibérica, "galicismos literários", não a continuação de formas latinas. Segundo ele, "...essa for-

mación es desconocida en sard., esp. y port."

14. Não incluímos palavras para as quais não era consignada qualquer indicação temporal.

15. São os seguintes os denominais registados em Machado (1977):

1. barbaçudo - séc. XVI
2. barbudo - 1132
3. botocudo - séc. XVIII
4. cabeçudo - 1220
5. carnudo - 1813
6. carrencudo - séc. XVI
7. cornudo - 1209
8. dentudo - séc. XVI
9. espadaúdo - séc. XVI
10. façudo - séc. XVII
11. felpudo - séc. XVI
12. facinho - séc. XVI
13. folhudo - séc. XVII
14. lanzudo - 1890
15. lombudo - séc. XVII
16. mamalhudo - 1873
17. mamudo - 1813
18. membrudo - séc. XVI
19. nembrudo - séc. XIV
20. narigudo - 1813
21. nervudo - 1813
22. olhudo - 1813
23. pentafaçudo - 1813
24. papudo - séc. XVIII
25. pegudo - 1890
26. peludo - séc. XVIII
27. pencudo - 1890
28. penudo - séc. XVI
29. quartaludo - séc. XVII

30. rabudo -	séc. XVIII
31. rechonchudo -	séc. XVIII
32. sanhudo -	séc. XIII
33. sesudo -	séc. XII
34. sisudo -	séc. XIII
35. taludo -	séc. XVI
36. testudo -	1813
37. tetudo -	1813

BIBLIOGRAFIA

- Aronoff, Mark (1976). *Word formation in generative grammar*. Cambridge, Massachusetts: MIT.
- Bauer, L. (1979). Patterns of productivity in new formations denoting persons using the suffix -er in modern English, *Chiers de Lexicologie*, 25. 2: 26-31.
- Bauer, L. (1983). *English word-formation*. Cambridge: Cambridge University Press
- Chomsky, Noam (1965). *Aspectos da teoria da sintaxe*. Trad., int., notas e apêndice de José António Meireles e Eduardo Paiva Raposo. Coimbra: Arménio Amado 1978.
- Clark, Eve V. & Berman, Ruth A (1984). Structure and use in the acquisition of word formation. *Language*, 60.3: 542-590.

-
- Halle, Morris (1973). Prolegomena to a theory of word formation. *Linguistic Inquiry*, 4. 1: 3-16.
- Huber, Joseph (1933). *Gramática do português antigo*. Lisboa: Fundação Calouste Gulbenkian, Lisboa, 1986.
- Jensen, John (1987). Review (Scalise, 1984). *Journal of Linguistics*, 23. 1: 206-209.
- Lausberg, Heinrich (1962). *Lingüística Românica. Segunda parte: Morfologia*. Madrid Gredos. 1966.
- Lieber, Rochelle (1983). Argument linking and compounds in English, *Linguistic Inquiry*, 14. 2: 251-285.
- Machado, José Pedro (1977). *Dicionário etimológico da língua portuguesa com a mais antiga documentação escrita e conhecida de muitos dos vocábulos estudados*. Lisboa: Horizonte. 5 vols.
- Matthews, P. H. (1974). *Morfologia. Introducción a la teoría de la estructura de la palabra*. Trad. y adaptado por Rafael Moury Casas. Madrid: Paraninfo. 1980.
- Romaine, Suzanne (1983). On the productivity of word formation rules and limits of variability in the lexicon. *Australian Journal of Linguistics*, 3: 177-200.
- Scalise, Sergio (1984). *Generative morphology*. Dordrecht: Foris.
- Williams, Edwin (1938). *Do latim ao português. Fonologia e morfologia históricas de língua portuguesa*. Rio de Janeiro: Tempo Brasileiro. 1975.