

UMA ANÁLISE DO SUFIXO – URA COM BASE NA MORFOLOGIA CATEGORIAL

SUFFIX – URA ON CATEGORIAL MORPHOLOGY

Livy Maria Real Coelho¹

Resumo

Pretendemos estender neste artigo a possibilidade da análise morfológica através de conceitos e paradigmas categoriais. Para isso usaremos o aparato sintático da própria Gramática Categorial aplicado à morfologia. Para este teste, selecionamos o sufixo – *ura* de palavras como *assadura* e *pintura*. Pretendemos, então, discutir, a partir dos dados, o funcionamento desse morfema, suas regras de aplicação e, finalmente, como e o que a análise categorial nos permite abordar. Para o levantamento de dados, utilizamos o programa Listas (UNICAMP/IEL).

Palavras-chave: Morfologia; Gramática Categorial; Sufixo –ura; Análise Morfológica.

Abstract

The purpose of this study is to consider the possibilities of a categorial treatment to morphology. In order to do so, we consider as our object derivation, more specifically suffixation, in Brazilian Portuguese. We intend to demonstrate the relationship between the suffix and the word formed by it, through a compositional and categorial analysis. We will consider the suffix –ura (as in *assadura* and *pintura*) to show the application of this treatment on the level of morphology.

Key-words: Morphology; Categorial Grammar; -ura; Morphological Analysis.

1. Introdução

Considerando o tratamento proposto por Real (2006), que possibilita a análise morfológica através da Gramática Categorial, pretendemos aqui analisar o sufixo –ura do Português Brasileiro (doravante PB). Acreditamos que, se bem sucedida, a análise detalhada de um sufixo em especial dá ao novo modelo proposto mais validade.

Apesar de este não ser o foco deste trabalho, apresentaremos brevemente, na seção 2, a proposta da Morfologia Categorial. Na seção 3,

¹ PG/UFPR-CAPES. livyreal@gmail.com

veremos detalhadamente os dados e como é possível analisá-los. Discutiremos brevemente na seção 4, o que traz a visão categorial à morfologia.

2. Gramática Categorial

2.1. Gramática Categorial Canônica

Usada geralmente por semanticistas e sintaticistas, a Gramática Categorial (doravante GC) é uma ferramenta que representa claramente o paralelismo entre as relações sintáticas e semânticas das expressões. É baseada em um léxico altamente informativo e em apenas seis pares de regras que coordenam as possíveis interações dos itens lexicais.²

Na GC, não é necessário postular categorias arbitrariamente para os itens, já que existe a possibilidade de, a partir de algumas categorias básicas, chegar a todas as outras. O mecanismo de hierarquia de aplicação de regras é inexistente na GC, afinal, seu mecanismo lógico-formal nos possibilita chegar ao mesmo resultado final independentemente da ordem de aplicação das regras. Também não há na GC a distinção, por vezes problemática, entre *estrutura profunda* e *estrutura superficial*, já que isso não existe dentro da ferramenta. Vejamos um exemplo de aplicação do modelo categorial à análise sintática:

$$\left[\begin{array}{l} \text{EXP} \text{ Pedro corre.} \\ \text{CAT} \text{ S} \\ \text{REG} \text{ R1} \\ \text{CON} \left[\begin{array}{l} \text{EXP} \text{ Pedro} \\ \text{CAT} \text{ n} \\ \text{REG} \text{ lex} \end{array} \right] \left[\begin{array}{l} \text{EXP} \text{ corre} \\ \text{CAT} \text{ n/s} \\ \text{REG} \text{ lex} \end{array} \right] \end{array} \right]$$

² □ Abordaremos aqui apenas as regras ditas sintáticas dessa ferramenta. Com isso, não estamos desvalorizando o papel da semântica no processo de formação de palavras, mas sim considerando que, ao menos por ora, é suficiente demonstrar a aplicação da ferramenta escolhida abordando somente regras sintáticas.

Figura 1: Pedro corre.

Na análise acima usamos aqui a notação defendida por Pagani (2003), a *Estrutura de Características* (EC). “As características [à direita] são sempre representadas por símbolos atômicos, mas os valores [à esquerda] podem ser representados por um símbolo atômico ou por outra EC” (Pagani, 2003, p.394). *Exp* é a expressão, *cat* a categoria sintática, *reg* a regra de formação da expressão, *con* a constituição da entrada lexical. Pagani ainda propõe a característica *den* que seria a denotação formal da estrutura, porém como neste trabalho não entramos na semântica e na formalização das expressões, não utilizamos essa característica.

Na análise de *Pedro corre*. a *exp* é a própria expressão, em *reg R1* – que mostra que a expressão é formada pela aplicação da regra 1 da GC, a regra que promove a união de duas expressões (cf. Mcgee Wood, 1993), a *cat S* indica que a expressão é uma sentença e *com* traz as expressões que compõe a sentença: *Pedro* e *corre*. Em *con* duas novas estruturas de características são acionadas, o que explicita o processo todo de formação da sentença. *Pedro* é um nome (*N*) e está lexicalizado (*lex*) e *corre* (*N\|S*) é uma expressão que toma um *N* a esquerda (*Pedro*) e forma um *S* (*Pedro corre*.).

Assim, encontramos a categoria de *corre* sem que fosse necessário postular uma nova categoria que desse conta de verbos intransitivos. A extensão de tal análise às outras classes de palavras e expressões é facilmente viabilizável. Vejamos como encontramos a categoria para um verbo transitivo direto:

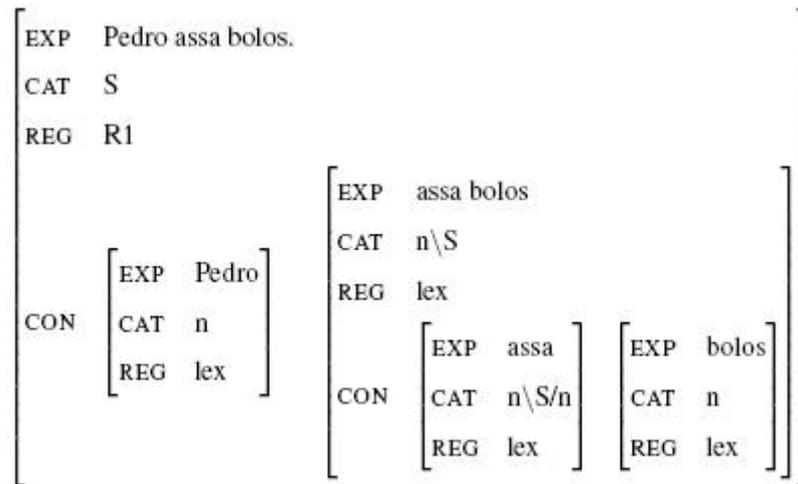


Figura 2: Pedro assa bolos.

Agora a categoria de *assa* é N\S/N, ou seja, uma expressão que tomando um nome de cada lado forma uma sentença. É muito interessante notarmos as relações entre as categorias. N\S é a categoria de um verbo intransitivo tanto quanto é a categoria de um verbo transitivo unido a um nome à sua direita. Isto é dizer, as expressões têm o mesmo funcionamento, embora sejam compostas por itens diferentes. Tanto *corre* quanto *assa bolos* fazem o mesmo na sentença: formam-na unindo-se a um nome à esquerda.

2.2 Morfologia Categorical

Tais análises podem ser usadas também no nível morfo-lexical, como proposto em Real (2006), ao aplicar essas mesmas regras a expressões menores. Entendemos que um dos grandes acréscimos da GC ao estudo morfológico dá-se quanto à possibilidade de olhar para o léxico *composicionalmente*, o que para um estudo que busca regularidades lexicais é extremamente interessante. Temos, por exemplo, na Gramática Gerativa, um sufixo como *-ura* (de *altura*) entendido como de categoria nominal, pois, tem-se em vista, para determinar a categoria do morfema a categoria da palavra final.

Tal análise é consistente dentro do modelo, afinal a Gramática Gerativa parte sempre das estruturas complexas para se chegar nos elementos mínimos.

Com um modelo de análise como este não há diferença explícita, e.g., entre um sufixo que nominaliza verbos e um que nominaliza adjetivos. Ambos são sufixos de categoria sintática nominal, pode-se no máximo explicitar que tipo de categoria os sufixos nominalizam, mas isto é irrelevante dentro do modelo.

Dentro da GC, a visualização composicional das estruturas complexas, sejam elas sintagmas, frases ou palavras, é sempre relevante e clara, pois parte-se dos elementos mínimos para se constituir o complexo, parte-se dos itens lexicais. Assim, dentro desse modelo, entenderíamos um sufixo como *-ura* como uma partícula que faz de um adjetivo um nome e teríamos em sua representação formal uma categoria que explicitaria essa relação.

Vejamos a aplicação da GC no nível sintagmático. Em uma expressão como *bolo assado* teríamos que analisar *bolo*, *assado* e a relação entre eles. *Bolo* é um N, *assado* é uma palavra composta a partir do verbo *assar*, do qual já temos categoria, e do sufixo *-do*, formador de participípios/adjetivos. Para encontrarmos a categoria de *assado*, e conseqüentemente de *-do*, devemos ver sua função da expressão na sentença.

O uso canônico de *assado*³ pode ser visto em frases como “*Comi um bolo assado*” ou “*Bolo assado é uma delícia*”. Assim, *assado* funciona como adjetivo, isto é, se une a um nome para promover um nome complexo. Podemos ver claramente que *assado* não muda a categoria da expressão a que se une. Logo, precisamos para ele uma categoria como N\N, *assado*, então, precisa de um N para formar outro. Advérbios também trazem

³ Nesta breve explicação didática não entraremos no uso contextual da expressão que possibilita sentenças como “O assado estava delicioso.”

categorias semelhantes a essa, que não interferem na categoria da expressão a que se unem. Essa discussão, no entanto, não cabe nesta breve explicação.

Vejam, agora a aplicação da GC à morfologia. Tendo a categoria de *assado* (N\N) e a categoria de *assar* (N\S/N), a GC nos permite encontrar a categoria de *-do* através de uma conta matemática simples, similar às frações que aprendemos no primeiro grau. *Assado*, um N\N, é o resultado da multiplicação da categoria de *assado* N\S/N e da categoria de *-do*, que ainda não sabemos. Logo *-do* é definido através das categorias de assar e assado. É uma partícula que toma um N\S/N a esquerda e forma um N\N, assim é uma expressão da categoria (N\S/N)\(N\N). Apesar de extenso e impressionante, o que essa categoria diz é, a partir de um verbo transitivo, temos um adjetivo quando o *-do* é aliado a esse verbo. São todas essas noções muito intuitivas que podem ser entendidas mesmo sem um aparato lógico formal. Todavia, é possível – e desejável – uma aplicação formalizada da teoria, no entanto, aqui queremos antes explicitar como é intuitiva e simples essa análise. A estrutura de características de *assado*, então, seria algo como:

$$\left[\begin{array}{l} \text{EXP} \quad \text{assado} \\ \text{CAT} \quad n \backslash n \\ \text{REG} \quad R1 \\ \text{CON} \quad \left[\begin{array}{l} \text{EXP} \quad \text{assar} \\ \text{CAT} \quad n \backslash s / n \\ \text{REG} \quad R1 \end{array} \right] \left[\begin{array}{l} \text{EXP} \quad \text{do} \\ \text{CAT} \quad (n \backslash s / n) \backslash (n \backslash n) \\ \text{REG} \quad R1 \end{array} \right] \end{array} \right]$$

Figura 3: Assado

2.3 Sufixo –URA

Podemos assim encontrar categorias não só para palavras, sentenças e sintagmas, mas para qualquer item lexical que quisermos, morfemas, por exemplo. Neste trabalho elegemos o sufixo *-ura* para a demonstração do modelo. Encontraremos a categoria de *-ura* através desse

procedimento descrito acima e depois analisaremos outras palavras onde o mesmo sufixo aparece para verificar a aplicabilidade desta categoria.

Partimos, então, de *assadura*, por já termos a categoria de *assado*. *Assadura* é formada pela R1 (regra que une constituintes: *um assado e -ura*) e tem categoria N, pois também é uma expressão nominal.

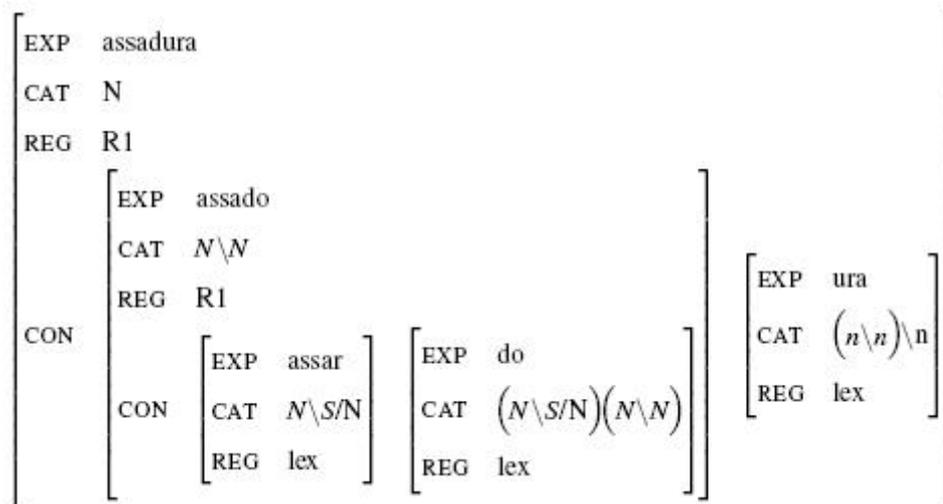


Figura 4: Assadura

Revisemos então o que foi dito até aqui para explicitar como chegamos na categoria de *-ura*. Em sua composição, *assadura* traz duas expressões *-ura* e *assado*, e esta por sua vez, traz mais duas *assar* e *-do*. *Assado*, como já vimos, é uma expressão que de um N atômico, forma um N complexo, *bolo assado* é tão nominal quanto *bolo*. Logo, *assado* é de categoria N\N, formada por *assar* e *-do*.

Assar é uma expressão que precisa de dois N, um de cada lado, para formar um S, como em *Pedro assa o bolo*. Logo, a expressão de *assar* é N\S/N *-do*, então, é um sufixo que de N\S/N forma um N\N. Agora temos o que

de mais interessante a GC traz para a morfologia, não temos que postular uma categoria para o sufixo, considerando um único aspecto. A partir da categoria da palavra selecionada e da palavra formada, encontramos a categoria do funtor (do sufixo) *-do* toma um N\S/N à esquerda formando N\N, logo é de categoria (N\S/N)\(N/N).

Agora, já com a categoria de *assado* e de *assadura* podemos encontrar facilmente a categoria de *-ura*, através do mesmo mecanismo toma um N/N (*assado*) e forma um N (*assadura*), portanto, é de categoria (N\N)\N.

$$\begin{bmatrix} \text{EXP} & \text{ura} \\ \text{CAT} & (n \setminus n) \setminus n \\ \text{REG} & \text{lex} \end{bmatrix}$$

Figura 5: -ura

Agora, com a categoria de *-ura* podemos testar se essa análise é possível para outras palavras formadas por esse sufixo. Testemos, então, com uma palavra que não seja formada por um particípio, como *alvura*, formada a partir de *alvo*.

$$\begin{bmatrix} \text{EXP} & \text{alvura} \\ \text{CAT} & \text{N} \\ \text{REG} & \text{R1} \\ \text{CON} & \begin{bmatrix} \text{EXP} & \text{alvo} \\ \text{CAT} & \text{N} \setminus \text{N} \\ \text{REG} & \text{lex} \end{bmatrix} \begin{bmatrix} \text{EXP} & \text{ura} \\ \text{CAT} & (n \setminus n) \setminus n \\ \text{REG} & \text{lex} \end{bmatrix} \end{bmatrix}$$

Figura 6: alvura

3. Dados

Aparentemente, a categoria que encontramos para *-ura* dá conta dos dados que temos. Analisaremos agora com mais cuidado os dados encontrados e suas possíveis análises, *i.e.*, suas bases, os processos morfológicos, regras em questão, etc.

Utilizamos o Programa Listas – IEL/UNICAMP para levantar o nosso corpus de trabalho. Esse programa permite a busca rápida de palavras com determinadas características fonéticas, morfológicas ou sintáticas em todo o dicionário Aurélio. Acreditamos assim trabalharmos com um *corpus* bastante considerável do PB. Ainda usamos algumas palavras que não estavam no corpus e que foram relevantes o suficiente para que nos déssemos conta de sua falta. São elas: *belezura, tessitura, tecedura, corredura, feitura, laqueadura e ranhura*. Embora não aparecessem no Aurélio, estas palavras estão dicionarizadas ou no Dicionário UNESP do Português Contemporâneo ou no Dicionário Etimológico Nova Fronteira ou no Houaiss. Entendemos que ao selecionar tais dicionários, com princípios de organização tão distintos, estamos contribuindo para a abrangência deste trabalho.

Por fim, listamos 210 palavras cuja grafia terminasse em *-ura*. Nessa listagem inicial observamos no *corpus* palavras que apesar de terminarem com a seqüência fonológica *-ura*, não são compostas do morfema com que estamos trabalhando. Nesta primeira revisão saíram 19 palavras, *aura, caradura, cenoura, cura, dura, lavoura, manicura, manjedoura, mesura, nunciatura*⁴, *pleura, púrpura, rasoura, salmoura, saracura, tanajura, vassoura, ventura, viatura*.

Algumas das palavras restantes não traziam uma base clara, como *punctura, vilegiatura* e *miniatura*, e por essa razão buscamos a etimologia dessas palavras, visando um tratamento mais preciso. Claro que nesse trabalho não buscamos um modelo diacrônico para tratar o léxico.

⁴ Tiramos *nunciatura* do *corpus* por não encontrarmos bases como *nunciado* ou *nunciato* no PB.

Entendemos, no entanto, que é o léxico o nível lingüístico onde a história da língua se faz mais presente e ignorar esse fato nos parece um tanto quanto reducionista. Assim, nos damos o direito de, nessa fase preliminar do trabalho, olhar para o percurso das palavras para poder delimitar com mais precisão exatamente onde é o sufixo *-ura* do PB que está atuando. Desse modo, algumas palavras aparentemente teriam explicação diacrônica. Assim, saíram da listagem mais 17 palavras: *aventura* (fr. *aventure*), *brochura* (fr. *brochure*), *clausura* (lat. *clausura*), *fervura* (lat. *fervura*), *figura* (fr. *figure*), *gravura* (fr. *gravure*), *leitura* (lat. *lectum*) *magistratura* (fr. *magistrature*)⁵, *musculatura* (fr. *musculature*), *nomenclatura* (lat. *nomenclatura*), *ossatura* (fr. *Ossature*), *pintura* (lat. *pictura*), *prefeitura* (lat. *praefectura*), *primogenitura* (lat. *primogenito*), *puntura* (lat. *punctura*), *ranhura* (fr. *ranhure*), *vilegiatura* (it. *villegiatura*).

Ainda retiramos palavras que de alguma forma já haviam sido contabilizadas, e.g., as formas compostas da *palavra cultura*: *agricultura*, *apicultura*, *avicultura*, *cacaicultura*, *cafeicultura*, *agricultura*, *escultura*, *floricultura*, *horticultura*, *psicultura*, *pomicultura*, *sericultura*, *silvicultura*, *suinocultura*, *triticultura*, *vinicultura*, *viticultura*, e ainda, *compostura*, *conjuntura*⁶, *descompostura*, *desembocadura*, *desenvoltura*, *desventura*, *impostura*, *fotogravura*, *incultura*, *infra-estrutura*, *padre-cura*, *pirogravura*, *porventura*, *reabertura*, *sobrecostura*, *xilogravura* e *zincogravura*.

Retiramos ainda as palavras que são formadas a partir do morfema retroativo, chamado também pela literatura de *regressivo*, como *procura*, que surge de *procurar*. São essas: *captura*, *censura*, *costura*, *estrutura*, *jura*, *manufatura*, *procura*, *rasura*, *satura*, *usura*, *comissura*. Tiramos também palavras que precisariam de formas verbais postuladas para que pudessem ser formadas através de um processo sincrônico: *quadratura*, *ranhadura*.

⁵ Para *magistratura*, o dicionário Nova Fronteira aponta a etimologia francesa adotada acima, o Houaiss, contudo, assume o radical latino *magistrat-*,

⁶ Entendemos que ao analisar *juntura*, analisamos o mesmo processo que ocorre em *conjuntura*, assim como em *postura* e *compostura*, etc.

Listamos abaixo o *corpus* total com que trabalhamos, entre parênteses registramos o que assumimos como argumento, *i.e.* base, tomado pelo sufixo na formação de cada uma dessas palavras e à direita listamos que significado a palavra formada parece assumir. Utilizamos: PR(opriedade), R(esultado), I(nstrumento) e P(rocesso).

ABERTURA	(ABERTO)	R/P
ABOTOADURA	(ABOTOADO)	P/I
ABREVIATURA	(ABREVIADO)	P/R/I
ALTURA	(ALTO)	PR
AMARGURA	(AMARGO)	PR
AMASSADURA	(AMASSADO)	P/R
ANDADURA	(ANDANDO)	P
APERTURA	(APERT-)	R/P
ARMADURA	(ARMADO)	I
ARQUITETURA	(ARQUITET-)	P/R
ARRANHADURA	(ARRANHADO)	R
ASSADURA	(ASSADO)	P/R
ASSINATURA	(ASSINADO)	P/R/I
ATADURA	(ATADO)	I/P/R
BENZEDURA	(BENZ-)	P
BRANCURA	(BRANCO)	PR
BRANDURA	(BRANDO)	PR
CANDURA	(CÂNDIDO)	PR
CATADURA	(CATADO)	P/R
CAVALGADURA	(CAVALGADO)	P/R/I
CERCADURA	(CERCADO)	P
CERZIDURA	(CERZIDO)	P/R
CHANFRATURA	(CHANFRADO)	P
COBERTURA	(COBERTO)	P/R/I
CORDURA	(CORDO)	PR
CORREDURA	(CORRE-)	P
COSEDURA	(COSE-)	P/R
CRIATURA	(CRIADO)	R
CULTURA	(CULTO)	PR
CURVATURA	(CURVADO)	R
DIABRURA	(DIABRIL)	PR
DIREITURA	(DIREITO)	PR
DITADURA	(DITADO)	P/R
DOÇURA	(DOCE)	PR
DOIRADURA	(DOIRADO)	P
DOURADORA	(DOURADO)	P
EMBOCADURA	(EMBOCADO)	P/R
ENVOLTURA	(ENVOLTO)	P/R
ENVERGADURA	(ENVERGADO)	P/PR
ESCRITURA	(ESCRITO)	P/R
ESFOLADURA	(ESFOLADO)	P/R
ESPESSURA	(ESPESSO)	P/R
ESTATURA	(ESTADO)	P/R
FARTURA	(FARTO)	PR

FECHADURA	(FECHADO)	I/P/R
FEITURA	(FEITO)	P
FRITURA	(FRITO)	P/R
FERRADURA	(FERRADO)	I/P/R
FERVURA	(FERV-)	P
FINURA	(FINO)	PR
FORMATURA	(FORMADO)	P/R
FORMOSURA	(FORMOSO)	PR
GORDURA	(GORDO)	PR
GOSTOSURA	(GOSTOSO)	PR
GROSSURA	(GROSSO)	PR
FRESCURA	(FRESCO)	PR
INVESTIDURA	(INVESTIDO)	P/R
JUDICATURA	(JUDICADO)	P
JUNTURA	(JUNTO)	P/R
LAQUEADURA	(LAQUEADO)	P/R
LARGURA	(LARGO)	PR
LAVADURA	(LAVADO)	P
LAVRATURA	(LAVRADO)	P/R
LEGISLATURA	(LEGISLADO)	P/R
LESURA	(LESO)	PR
LICENCIATURA	(LICENCIADO)	P/I/R
LIGADURA	(LIGADO)	P/R/I
LONJURA	(LONGE)	PR
LOUCURA	(LOUCO)	PR
MACHUCADURA	(MACHUCADO)	P/R
MORDEDURA	(MORDE-)	P/R
NEGRURA	(NEGRO)	PR
PARTITURA	(PARTIDO)	I
PISADURA	(PISADO)	P/R
POSTURA	(POSTO)	P/PR
QUEIMADURA	(QUEIMADO)	R/P
QUENTURA	(QUENTE)	PR
RACHADURA	(RACHADO)	R
RAPADURA	(RAPADO)	R/P
ROÇADURA	(ROÇADO)	P
ROEDURA	(ROED-)	P/R
RONCADURA	(RONCADO)	P/R
ROTURA	(ROTO)	PR
SECURA	(SECO)	PR/R
SEMEADURA	(SEMEADO)	P/R
SEPULTURA	(SEPULTO)	I/R
SOLDADURA	(SOLDADO)	P/R
SOLTURA	(SOLTO)	PR
TECEDURA	(TECE-)	P/R
TECITURA	(TECIDO)	P/R
TEMPERATURA	(TEMPERADO)	P/PR
TERNURA	(TERNO)	PR
TEXTURA	(TEXTIL)	PR
TINTURA	(TINTO)	P/R/I
TONSURA	(TONSO)	PR
TONTURA	(TONTO)	PR
TRAVESSURA	(TRAVESSO)	PR
TREMURA	(TREM-)	P/R
UNTURA	(UNTO)	P/R

URDIDURA	(URDIDO)	P/R
VERDURA	(VERDE)	PR
VESTIDURA	(VESTIDO)	P

Já com essa nova lista composta de 104 palavras, procuramos qual seria a categoria sintática da base da palavra final. Inicialmente pensamos que o *-ura* tomava palavras de categorias verbais e nominais, como *fritar>fritura* e *gordo>gordura*. Analisamos então que tipo de relação tinha a palavra final com a palavra base. Quando a palavra base era um nome a relação era a de *propriedade*: *gordura* é a propriedade do *gordo*. Já quando as bases eram verbais, as relações poderiam ser de *resultado*, *processo* ou de *instrumento*, como em *arranhar>arranhadura*, *varrer>varredura* e *armar>armadura*, respectivamente. Haveria, ainda, palavras que traziam mais de um desses sentidos, como *assinatura* e *assadura*.

Além de listar todas as possíveis bases, analisamos os verbos que serviriam de palavras bases quanto a seus comportamentos sintático e semântico, como tentativa de definir o que exatamente traz os diferentes significados para a palavra final, *i.e.*, porque e como bases verbais davam resultados semânticos diferentes.

Analisamos a transitividade do verbo e as propriedades aspectuais. Apesar de a grande maioria dos verbos ser transitiva e a maioria das palavras finais conter, entre outras, a idéia de processo, não conseguimos destacar nenhuma característica das bases que parecesse ser relevante para as palavras finais. Muitas palavras como *fritura* aparecem com idéia de *processo* e idéia de *resultado*. Analisamos isto como um fenômeno polissêmico que ocorre depois da formação da palavra final: um único morfema resulta em uma única palavra, mas esta assume diferentes significados a depender do contexto em que aparece, conforme já discutido.

Claro que outras soluções poderiam ser pensadas, como assumir que temos vários *-ura* formando, a partir da mesma base, palavras diferentes. Não adotamos essa solução por entendermos que os sentidos assumidos

pelas palavras finais são muitos próximos e facilmente relacionáveis não sendo necessário criar um novo morfema para dar conta de todas as acepções que as palavras finais podem tomar. Pensamos também que criar um novo morfema a cada sentido que encontrássemos para uma palavra final incharia o sistema mais que dobrando o número de partículas a serem memorizadas. Parece-nos que os possíveis resultados do *-ura* estão intrinsecamente relacionados com a distribuição de outras partículas do léxico. Reparamos durante a classificação que o *-ura* quando não assume determinados sentidos, não o faz por bloqueio paradigmático (cf. Assis Rocha, 1998, p141), *i.e.*, há já um outro morfema fazendo as vezes desse sentido.

Já com a lista pronta e esses testes feitos, percebemos que uma generalização maior poderia ser feita, se assumíssemos que as palavras bases seriam na verdade as formas participiais dos verbos: *andadura* viria de *andado* e não de *andar*, *feitura* de *feito*, *fritura* de *frito*, *catadura* de *catado*. Essa hipótese pareceu-nos melhor por dois motivos: i. unir duas categorias que a gramática tradicional trata por coisas distintas, participios e adjetivos, mas que podem ser analisadas sob um mesmo rótulo. ii. dar conta do porque vogais temáticas permanecem em algumas palavras finais e em outras não, *e.g.*, em *andadura* a vogal permanece, diferentemente de *fritura*, que tem a vogal temática perdida, afinal *andadura* viria de uma base já com a presença da vogal temática (*andado*), enquanto *fritura*, por não vir de *fritado*, não guardaria essa mesma vogal.

Para não abandonarmos nossa antiga teoria⁷, tentamos dar conta desse fato no âmbito fônico procurando ambientes que proporcionariam a permanência ou não da vogal temática. No entanto parecia não haver um condicionante para a permanência ou não da vogal temática e conseqüente

⁷ Como o defendido por Lakatos (apud Borges Neto, 2004, p.142): inserimos, para corroborar a nossa teoria, explicações secundárias que abrangem também os possíveis contra-exemplos ao que postulamos, a fim de proteger “o núcleo” da hipótese.

inserção da consoante que poderia ser *d/t*. Temos para esse sufixo três formas encontradas: *-ura* (como em *fritura*), *-dura* (como em *andadura*) e *-tura* (como em *assinatura*). Postulamos então que é um único sufixo, afinal tem sempre o mesmo comportamento, mas que conta com três formas fônicas, formas essas extremamente parecidas. Aqui, caímos no mesmo problema que estruturalistas e gerativas caem ao postular formas subjacentes, no entanto acreditamos que temos motivos para considerarmos o *-ura* uma única entrada lexical: seus sentidos são radicalmente próximos, sua realização é muito similar. Enquanto Hoeksema (1985) postula que uma entrada lexical tem que ter identidade fonológica, aqui acreditamos ser essa identidade algo flexível.

Não encontramos condicionantes para a seleção de *d* ou *t*. Nota-se, por exemplo, que em *abreviatura* e *abotoadura*, temos o mesmo contexto fonológico, fonético e morfológico. Diante de tal fato, assumimos como arbitrária a escolha entre *t* e *d*. As grandes semelhanças entre esses sons e sua distribuição complementar em outros casos do PB (como em *quantidade*, *quantitativo*, etc), bem como a dificuldade de diferenciar esses sons durante a aquisição de escrita e fala, suportam nossa análise. No entanto, reconhecemos que parte do nosso intuito nesse trabalho, buscar a regularização do léxico, já não poderá compreender todos os fenômenos com os quais deparamos, pois, como os cientistas que recriminamos, estamos jogando no *saco do léxico* alguns dos problemas que encontramos. Podemos, ao menos, caso adotarmos a segunda hipótese levantada, dar conta do porque essa consoante, ora realizada como *t* ora como *d*, não aparece em muitos casos. Assumindo que as palavras vêm do participípio, poderemos dar conta, através dos participípios reduzidos, como *frito*, *leso* e *posto*, da falta da vogal.

Tal hipótese traz algumas soluções e regularidade ao sistema, mas como não poderia deixar de ser, nos gera problemas nos seguintes casos: i. onde um único verbo tem um participípio reduzido e um normal: como *envolver* > *envolvido* / *envolto*, gerando *envoltura*, apenas do participípio reduzido, e ainda: *fritura* (de *frito* e não de *fritado*), *juntura* (de *junto* e não de *juntado*, forma

também encontrada), *secura* (de *seco* e não de *secado*), *soltura* (de *solto* e não de *soltado*), *untura* (de *unto*, e não de *untado*); ii. nos casos onde a palavra final parece vir de uma possível, porém não realizada, forma de particípio reduzido, como *pintura*, *cintura*, *fervura* e *tintura*, ou ainda, de possíveis verbos como *ratar* (*ratadura*) e *escravar* (*escravatura*); iii. no caso de palavras que mesmo com o particípio existindo insistem em vir do radical infinitivo, como *arquitetura* (e não *arquitetad-ura*), *candidatura* (e não *candidadura*), *varredura*, *corredura*, *benzedura*, *mordedura*, *cozedura*, *roedura*, *tremura*.

Quanto ao problema ii., optamos simplesmente por excluir tais palavras do trabalho. Embora seja possível através de formas subjacentes, pertencentes ao já citado léxico possível, abranger tais dados, optamos por trabalhar apenas com o léxico realizado, o que exclui verbos não realizados das entradas lexicais passíveis de servirem como elemento à formação de novas palavras.

Notamos, depois de muitos testes, que se entendêssemos que o *-ura* adota sempre particípios, e que em existindo particípios reduzidos, o sufixo os preferiria, daríamos conta de um bom número de dados, como os do problema i. Percebemos também que as palavras do problema iii., poderiam ser divididas sob dois rótulos: **a.** palavras com um vasto número de sílabas seguidas cuja consoante é uma oclusiva alveolar (surda ou sonora) **b.** palavras que tem por base verbos de segunda conjugação.

Quanto a **a**, podemos admitir que o falante por economia lingüística opta por contrair uma das oclusivas; embora isso vá contra Halle (apud Borges Neto, 2004, p.140), voltaremos mais tarde a essa questão, já que argumentos de economia lingüística, em geral, são discutíveis. Quanto a **b**, buscamos, no latim, a etimologia dessas palavras, nos radicais de *infectum* e particípio, e não encontramos nada que condicionasse diacronicamente o uso da vogal baixa em detrimento da alta. Por exemplo: *morder* em latim é *mordeo* (primeira pessoa do singular do presente do indicativo), *mordere* (infinitivo), *morsum*

(particípio), seu radical de *infectum* é *mord(e)*⁸- e o de particípio, onde esperávamos encontrar uma solução, é *mors*-. Ao contrário de *pungo*, *pungere*, **punctum** e *pingo*, *pingere*, **pictum** - que trazem já no particípio latino a forma tomada pelo sufixo em português, explicando, assim, as formas *punctura* e *pintura* - a forma latina de *morder* não explica a presença do e em *mordedura*.

Poderíamos, então, propor uma regra de *ablaut* entre *e/i* para verbos de segunda conjugação, até porque para falantes de PB, essa alternância é ainda mais recorrente do que a alternância por nós assumida anteriormente entre *t/d*. No entanto, se assumíssemos que a palavra vem do radical do infinitivo e não do particípio, como era nossa primeira hipótese, esses casos ficariam explicados. O problema aqui não é assumir o *ablaut* ou não, mas sim ignorar que a forma em e é na verdade a forma de infinitivo.

Tais verbos nos fizeram oscilar entre as duas hipóteses, embora a regularização da hipótese tardia seja extremamente interessante para um estudo lexical. No entanto, nos questionamos sobre os dados que, não sendo contemplados, parecem invalidar, ou, ao menos, enfraquecer nossa análise.

Entendemos que para se abster da contemplação de um dado, *i.e.*, para podermos retirá-lo do corpus e dizer que esse dado, apesar de não abrangido pela nossa hipótese, não a invalida, existem duas razões: uma *quantitativa* e uma *qualitativa*. “*Dentro de um programa de investigação, uma teoria será, normalmente, eliminada, por uma teoria melhor. Uma teoria é melhor que outra se apresentar conteúdo excedente ou se tiver (prometer) maior poder heurístico (p. ex., uma teoria sintática X é melhor que uma teoria sintática Y se explicar mais fatos que Y ou se permitir mais facilmente, digamos, uma semântica.*” (Borges, 2004, p.144) Essas duas razões de Borges para a escolha de uma teoria se aproximam bastante com as que explanaremos, embora nossa discussão aqui seja sobre a relevância de dados.

⁸ e é o que a Gramática Tradicional chama de vogal temática, entretanto, nessa conjugação, a vogal está presente em todas as formas da palavra, podendo ser analisada como parte do radical.

Como ele, postulamos o fator numérico e o fator heurístico negativo de Lakatos, “*princípios metodológicos com duas funções: proteger o núcleo de refutações, i.e., a heurística negativa induz a modificações nas hipóteses auxiliares e não no núcleo) e impedir tentativas de explicações de tipos radicalmente diferentes (p. ex., o uso de explicações mentalistas no programa de instigação do behaviorismo clássico).*” (apud Borges, 2004, p.142).

A primeira se dá quando poucos dados não são contemplados em comparação ao número dos que são. Por exemplo, temos em português, apenas quatro palavras formadas claramente a partir de substantivos e o sufixo *-ura*, são elas *musculatura*, *nervura*, *ossatura* e *dentadura*. Todas são formadas por palavras que remetem a partes do corpo e ainda, na história dessas palavras, notamos que *musculatura* e *ossatura* vêm do francês *musculature* e *ossature*, e até que elas entraram no português no mesmo ano, 1873.⁹ Entendemos que apenas esses quatro dados gerados a partir de substantivos não são suficientes para postularmos ainda mais uma regra para o funcionamento do sufixo com o qual trabalhamos. Assim, por uma razão *quantitativa*, podemos assumir que essas palavras já não fazem parte do nosso objeto de estudo, estando assim, incapazes de invalidar nossa hipótese.

A segunda razão, a *qualitativa*, de certa forma também se aplica aos dados acima comentados. Isso porque havia em todos os dados que retiramos do corpus, um *único princípio* que parecia agir sobre os processos realizados: o fato de todos representarem partes do corpo. Quando parece haver um só fator para as palavras se portarem de maneira irregular, podemos entender que tal fator está aquém das intenções --- ou se formos sinceros, das possibilidades --- de explicações do nosso estudo. Isto é dizer: existem formas regulares e vamos tratá-las, as que se portam de forma irregular são regidas por um fator X

⁹ Coincidentemente, ou não, as outras duas palavras, *nervura* e *dentadura*, também entraram no mesmo ano no português, 1844 (cf. Dicionário Etimológico Nova Fronteira).

que não é abordado. Poderíamos, por exemplo, dizer que tais palavras, por virem via empréstimo, não entrariam em nosso estudo. Isso até é muito coerente dentro deste trabalho, já que inicialmente retiramos do corpus palavras que vinham para o português via empréstimo, *e.g.*, *villegiatura*. É, no entanto, importante ressaltar tal fator (o da razão *qualitativa*) e que ele seja explicável *lingüísticamente*. Por exemplo, não poderíamos retirar os dados com ablaut e nem os com alternância *d/t* assumindo ou uma forma ou outra porque estas distinções não são *suficientes* para tal, ou seja, porque existem teorias - que nesse trabalho servem de teorias secundárias - que sustentam no sistema lingüístico, essa alternância.

Caímos, então, no conceito de *suficiente* para ser retirado de um *corpus* ou não. Entendemos que para o intuito deste trabalho dados ligados ao sufixo *-ura* apenas por razões diacrônicas (*e.g. vilegiatura*), ou dados que parecem ter sofrido especialização (*moldura*¹⁰) não precisam ser considerados.

Em lingüística - na ciência, em geral - é comumente visto que para tentar salvar um modelo, tomam-se saídas *ad hoc*: jogam-se problemas para outros níveis, que se diz não pretender estudar; postulam-se como reais estruturas que não necessariamente estão claras, nem para o leigo, nem para o lingüista¹¹, etc. Imaginemos agora a grande quantidade de soluções desse tipo a que teremos que chegar neste trabalho, em especial por tratarmos do léxico - possivelmente o nível mais irregular de todos, onde ficam evidentes empréstimos entre línguas, suas evoluções, muitas das possíveis alternâncias

¹⁰ *Moldura* sai do trabalho tanto por não vir de moldado --- a forma seria **moldadura* --- quanto por aparentemente vir de um substantivo *molde*. Veremos a frente que a hipótese de que esta palavra venha de **moldo* um possível particípio reduzido de *moldar* não é tão absurda.

¹¹ Exemplos disso são a postulação do morfema zero estruturalista ou ainda do arquifonema nasal de Mattoso Camara (1972, p.49)

fonológicas¹², etc. É então para fugirmos dessas soluções que estabelecemos esses dois critérios que tiram palavras do corpus. Preferimos antes não abranger todos os dados - explicitando-os, obviamente - mas ter uma hipótese consistente para todos os que conseguirmos contemplar e ainda, sem postular um sem número de regras que dão conta de um fator pouco recorrente na língua.

Toda essa discussão foi trazida pelos casos de palavras finais com a alternância *e/i* em relação aos participípios. O que há de qualitativamente igual em todos esses casos é o fato de que todas as palavras surgem de verbos de segunda conjugação. Este parece ser o único fator que une sob um único rótulo todas as palavras que não são abordadas pela explicação encontrada. No entanto, este fator, a conjugação, não é *suficiente* para que retiremos dados do nosso *corpus*. Não é possível uma explicação que exclua, sem um ótimo motivo, dados de uma pesquisa por estes serem de segunda conjugação¹³; poderíamos claro, assumir que verbos de segunda conjugação sofrem *ablaut*. Isso pareceu-nos uma melhor solução: estamos usando uma ferramenta já aceita pelos lingüistas (a alternância *e/i*) e ainda conseguimos traçar um padrão de regularidade onde ela atuaria, traço que não conseguimos encontrar na relação *d/t*.

Antes, porém, de assumirmos o *ablaut*, procuramos todas as palavras finais que vieram de segunda conjugação, afinal antes estávamos olhando do problema para a origem da palavra, e agora, observando se o mesmo problema se verifica em todas as palavras com essa origem

¹² Aqui não assumimos que todas as alternâncias fonológicas possam ser vistas no nível morfológico pensando em casos como a alternância *s/z* em [casaz.amarelas] e [casas.pretas] que se dá no nível sintático. No entanto há muitas explicitadas no nível morfo-lexical.

¹³ Claro que se estivéssemos fazendo um trabalho de nível fonético/fonológico poderíamos fazê-lo, em razão da diferença das vogais temáticas e as possíveis implicações disso no sistema.

poderemos, talvez, listar alguma regularidade. Encontramos então *feitura* e *escritura*, dos participios reduzidos *feito* e *escrito*. Ora, já havíamos percebido que em verbos detentores de mais de um participio, era sempre o participio reduzido a base selecionada, o que mostra uma certa hierarquização na hora da escolha da base. Olhando então para estes dados, é visível que verbos de segunda conjugação que possuem participios reduzidos (como *escrever* e *fazer*) funcionam como todas as outras formas da língua.

Então, observamos os verbos de outras conjugações: por que não haveria ablaut entre *e/i* em verbos de terceira conjugação, por exemplo? Porque nunca encontramos algo como **partetura* de *partido*, se encontramos *varredura* de *varrido*? Mais uma vez buscamos as respostas nas diferentes formas que os verbos assumem a depender de sua conjugação: *partir*>*partido*>*partitura*, *assar*>*assado*>*assadura*, mas não *correr*>*corrido*>**corridura*. Ao irmos além do radical de participio, e notarmos também o radical de infinitivo, parece que encontramos uma possível solução para o problema: na seleção de bases o *-ura* tomaria sempre a forma participial, quando esta existisse em sua forma reduzida, como *frito*. No entanto, em não existindo tal forma, seria selecionado como base o radical do infinitivo: como em *benzedura* e *varredura*. Tal hipótese, por mais abstrata que pareça, abrange todos os nossos dados deverbais, inclusive os verbos de primeiras e terceiras conjugações e os casos de *arquitetura* e *candidatura*.

Agora então temos formadas duas hipóteses: i. o *-ura* toma bases nominais ou verbais; ii. o *-ura* toma bases adjetivas, participais (e quando não as encontra toma bases verbais que teriam formado os adjetivos). Embora a primeira hipótese contemple 136 dados e a segunda 124 (isso porque a primeira não abrange os dados *belezura*, *canelura*, *cintura*, *escravatura*, *levedura*, *moldura*, *musculatura*, *nervura*, *ossatura* e *moldura*¹⁴), a segunda

¹⁴ Algumas dessas palavras, como *moldura*, *cintura*, *escravatura*, *musculatura*, *nervura*, *ossatura* seriam explicadas mesmo pela segunda hipótese, se assumíssemos entradas lexicais subjacentes, porém não o fazemos, como já discutido.

hipótese se presta a uma maior regularização nos dados, mas não abrange o mesmo número de dados. A primeira é mais abrangente, mas temos o *-ura* tomando por base mais de uma categoria sintática, além de não termos nessa hipótese a distinção adjetivo/substantivo, a qual prezaríamos bastante, por pretendermos uma análise categorial dos dados, onde tal distinção está na base da teoria e é demonstrável formalmente. Na primeira análise também teríamos que assumir um ablaut ou uma vogal [a] presente em *ossatura* e *escravatura*. Neste caso, assumir tal mudança nos parece um tanto mais complicado, pois dessa vez trata-se do surgimento de um som onde ele não era necessário, enquanto nos outros casos tratava-se de uma alternância recorrente na língua.

Antes de elegermos por fim uma ou outra, pensemos mais uma vez na nossa categorização de palavras finais quanto ao seu significado. Tínhamos separado-as em processos (*cavalgadura*), instrumento (*abotoadura*), resultado (*criatura*) e propriedade (*brancura*). Analisemos para essas distinções para localizar algum tipo de regularidade em relação às palavras base. Encontramos algum nível de regularidade na hipótese ii.: bases vindas de adjetivos puros resultam em propriedade e bases participiais resultam em qualquer uma das outras categorias. Assim, nos pareceu interessante a relação traçada entre semântica e sintaxe.

Quanto à hipótese i., uma regularização só é traçável se dissermos que bases verbais resultam em *instrumento*, *resultado* ou *processo* e que bases nominais, além de resultarem em *propriedade*, resultam também em substantivos de massa, (como *musculatura*, *ossatura*), ou ainda em novos substantivos, porém sem uma característica semântica definida (como *canela>canelura* e *molde> moldura*). Se assumíssemos tal hipótese, teríamos ainda novas categorias semânticas para as palavras finais que ainda não tínhamos notado, e estas seriam, ainda, confusas e nada delimitadas.

Assim, tanto pela traçável relação semântica base-palavra final, quanto pela regularidade que a hipótese ii. traz, optamos neste trabalho por

trabalhar com ela. Temos em vista, sim, todas as suas falhas, mas para o nosso propósito --- a regularização a partir do menor número de regras e postulações possíveis --- esta nos parece melhor por ora. Queremos ressaltar que trabalhar com a hipótese ii. Não significa que a primeira seja falsa, pois “o falseamento e o abandono de teorias são processos independentes. O verdadeiro teste de uma teoria, então, está na capacidade de prever fatos novos. Se o faz, refutações e anomalias podem ser ignoradas.” (Borges Neto, 2004, p.144). Se pensarmos nessa capacidade de prever fatos novos, citada por Borges, ainda temos mais um ponto a favor da teoria ii., alguns substantivos formados pelo *-ura* a partir de um radical de infinitivo, começam a aparecer, já em textos escritos, formados pelos participios: como *varredura* e *varridura*¹⁵. Entretanto, independente do quanto pensemos nas teorias que encontramos e do quanto gostemos delas (ou não), provavelmente existem ainda outras hipóteses que poderiam ser levantadas e das quais nem nos aproximamos.

4. Conclusão

A partir do levantamento de dados e das representações que fizemos, evidenciamos que a análise categorial é uma ferramenta que também pode ser aplicada à morfologia. Questões teóricas, como diferentes níveis de

¹⁵ Buscamos, no google.com, a ocorrência de alguns pares de palavras (claro que esta ferramenta não é absolutamente confiável e que muitas entradas são contadas mais de uma vez, mas acreditamos que a simples ocorrência da forma não canônica em textos escrito já demonstra uma possível mudança no processo de seleção do *-ura*). Encontramos: 704.000 para *varredura* e 8 para *varridura*; 377.700 para *mordedura* e 88 para *mordidura*, 160 palavras para *roedura* e 1 para *roidura*. Tais palavras parecem indicar que o *-ura*, numa tentativa de regularidade, começa a tomar sempre participios, independente de estes serem mínimos ou não. Encontramos, também, na fala oral, expressões como *benzidura*, e não *benzedura*, como o dicionarizado. É possível que essas formas populares também corroborem a hipótese de que, hoje, o *-ura* toma especialmente participios como argumentos.

estrutura e a dificuldade de rotular morfemas, são resolvidas ao se trabalhar com essa hipótese e os resultados são abrangentes e elegantes.

Acreditamos, porém, que o maior ganho deste trabalho é a exposição clara dos processos de formação de palavra, bem como das funções exercidas por cada um dos elementos que participam desse processo. Esse trabalho visou antes uma aplicação do modelo com base empírica do que a sua defesa, para uma discussão mais abrangente das possibilidades da aplicação da GC neste nível, vide Real (2006).

Referências Bibliográficas

- ASSIS ROCHA, Luiz Carlos. *Estruturas morfológicas do português*. Editora UFMG: Belo Horizonte, 1998.
- BORBA, Francisco S.(org.). *Dicionário UNESP do Português Contemporâneo*. Editora UNESP: São Paulo, 2004.
- BORGES NETO, José. *Ensaio de filosofia da lingüística*. Parábola: São Paulo, 2004.
- CUNHA, Antonio Geraldo da. *Dicionário Etimológico Nova Fronteira da Língua Portuguesa*. Nova Fronteira: Rio de Janeiro, 1982.
- Dicionário de Latim-Português*. Porto: Porto Editora, 2001.
- Dicionário eletrônico HOUAISS da Língua Portuguesa*, versão 1.0.5a.
- HOEKSEMA Jack. *Categorical Morphology*. Garland Publishing: Nova Iorque e Londres, 1985
- MATTOSO CAMARA JR, Joaquim. *Estrutura da língua portuguesa*. Vozes: Petrópolis, 1972.
- MCGEE WOOD, Mary. *Categorical Grammar*. Routledge, 1993.
- PAGANI, Luis Arthur. Gramática Categorical através de estrutura de características. in *Revista Letras* (60), Curitiba: Editora UFPR, 2003. p. 385-410.

REAL, Livy. Sufixação: uma nova proposta de análise in *Anais do VII Ce/Sul*, 2006. Pelotas: UCPEL, 2006. p. 221..227