

Diminutivo em Português Brasileiro: sufixação ou infixação?^{1*}

(Diminutives in Brazilian Portuguese: suffixation or infixation?)

Maximiliano Guimarães¹, Gesoel Mendes²

^{1,2} Departamento de Linguística – Universidade Federal do Paraná (UFPR)

maxguimaraes@ufpr.br, gesoelmendesjr@gmail.com

Abstract: Here we entertain the possibility of analyzing the diminutive morpheme ‘*inh*’ of Brazilian Portuguese (BP) as an infix, rather than as a suffix. Therefore, we argue against the tradition of (descriptive and analytical) morphological studies in BP. We show that, surprisingly, our unorthodox hypothesis not only accounts for the run-of-the-mill data more naturally than the competing analyses, but also accounts for some non trivial additional data that has been ignored so far.

Keywords: Morphology; Diminutive; Brazilian Portuguese; Suffixation; Infixation.

Resumo: Consideramos aqui a possibilidade de se analisar o morfema de diminutivo ‘*inh*’ do Português Brasileiro (PB) como um infixo, e não como um sufixo. Posicionamo-nos, assim, contra a tradição dos estudos morfológicos (descritivos e analíticos) em PB. Mostramos que essa hipótese heterodoxa surpreendentemente, além de acomodar os dados triviais de forma mais natural que as hipóteses concorrentes, também dá conta de dados adicionais não-triviais, até agora ignorados.

Palavras-chave: Morfologia; Diminutivo; Português Brasileiro; Sufixação; Infixação.

Introdução: dados iniciais e hipóteses concorrentes

Começamos pelos dados mais triviais em (1a) e (1b), formados respectivamente por raiz intrinsecamente masculina e feminina; e (2a) e (2b), formados por raízes subespecificadas para gênero às quais são atribuídos traços de gênero (masculino e feminino, respectivamente). Esses casos exibem claramente o morfema de diminutivo ‘*inh*’ posicionado entre a raiz e a vogal final (VF).²

(1)	a:	carr-o	a’:	carr-inh-o	<i>raiz intrinsecamente</i> [-FEM]
	b:	bol-a	b’:	bol-inh-a	<i>raiz intrinsecamente</i> [+FEM]

¹ * Agradecemos à CAPES pelo apoio financeiro, e a Gabriel Araújo, Marcelo Ferreira, Andrew Nevins e Maria Cristina Figueiredo Silva, por comentários feitos sobre o tema em diferentes estágios deste projeto.

² Para Mattoso Camara (1970, 1972), a VF /a/ em ‘gat-a’ é morfema de gênero (MG), sendo este o caso marcado, enquanto a VF /o/ em ‘gat-o’ é vogal temática (VT), sendo este o caso não-marcado, e não morfema de masculino. No par ‘professor/professora’, o masculino é ‘professor’, e não ‘professor-o’. O plural de ‘professor’ é ‘professor-e-s’, e o plural de ‘mulher’ é ‘mulher-e-s’. Assim, ou aceitamos que existe uma VT /e/ subjacente ao final da raiz (o chamado /e/-teórico), ou que ‘es’ é alomorfe de ‘s’ como morfema de plural dos nomes masculinos terminados em /r/. Consideramos, no entanto, que tal questão é de menor relevância para o nosso trabalho. Conforme mostraremos a seguir, as generalizações adequadas podem ser feitas a partir do conceito mais abrangente de VF, o que quer que constitua essa VF em última instância. Portanto, adotamos aqui VF como um “termo guarda-chuva” para VT ou MG, ou mesmo um *portmanteau* de VT+MG (ou VT+MG+GRAU, como na hipótese **HS-R** adiante).

(2)	a:	gat-o	a':	gat-inh-o	raiz [-FEM] por atribuição
	b:	gat-a	b':	gat-inh-a	raiz [+FEM] por atribuição

Se seguirmos a tradição e tomarmos *'inh'* como um sufixo, há pelo menos duas implementações técnicas possíveis, conforme ilustrado na **Figura 1**. Pela *Hipótese Sufixativa Conservadora (HS-C)*, a diminutivização se faz por uma regra não-transformacional (livre de contexto) de inserção de *'inh'* num “molde” de nódulos não-terminais sufixais ordenados entre si de acordo com as regras da língua, o que quer que elas sejam.³ Pela *Hipótese Sufixativa Radical (HS-R)*, a diminutivização se faz pelas seguintes regras transformacionais de reescritura de VF (sensível(is) ao contexto do gênero da raiz): (i) VF → *inha* / [raiz (+fem)] ___; e (ii) VF → *inho* / [raiz (-fem)] ___.

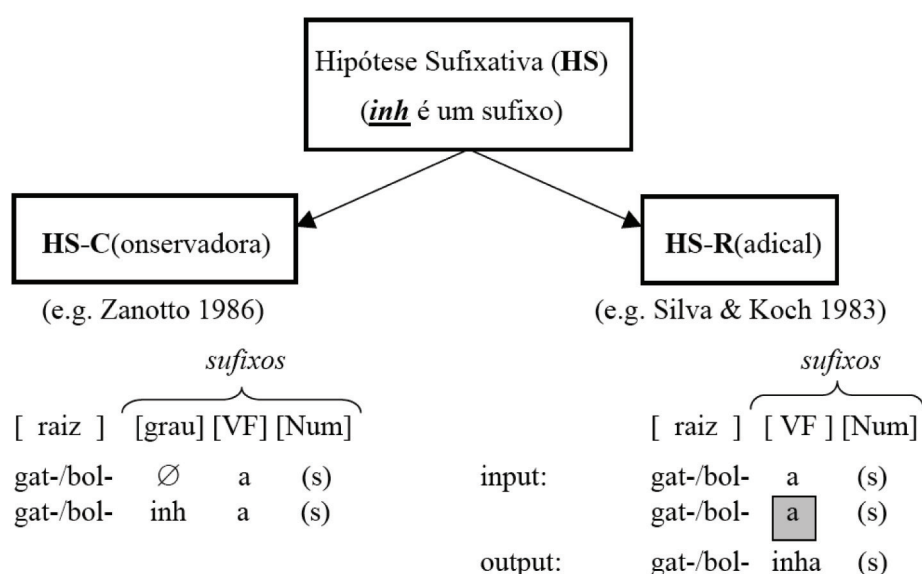


Figura 1. Duas implementações técnicas da Hipótese Sufixativa

Alternativamente, pode-se considerar *'-inh'* como um infixo. Pela *Hipótese Infixativa (HI)*, raiz e VF combinam-se formando uma base complexa cuja estrutura interna torna-se opaca no ciclo derivacional subsequente, havendo inserção de *'-inh'* no interior dessa base, de acordo com os requerimentos fonotáticos relevantes.⁴

³ Dizer que a inserção de *'inh'* para diminutivo (ou *'ão/on(a)'* para aumentativo, ou *∅* para grau normal) se dá por mecanismos não-transformacionais não significa que toda a morfologia da palavra seja construída de modo livre de contexto. Presumivelmente, consoante **HS-C**, a inserção do morfema específico que preenche a posição VF no molde é sensível a traços idiossincráticos de classe da raiz.

⁴ Salientamos que Kehdi (1999) se posiciona contra **HI** (questionando as conclusões de Back e Mattos (1972)), mas tal trabalho está mencionado aqui por se tratar de um dos poucos em que se considerou **HI** como uma possibilidade e em que se pesaram seus prós e contras. As análises de Back e Mattos (1972) e de Monteiro (1998) em favor de **HI**, embora inevitavelmente tenham algo em comum, são distintas tanto nas suas bases empíricas como na lógica de seus argumentos. Em ambos os casos, trata-se de descrições e reflexões bastante parciais do fenômeno, e que, portanto, não dão conta de uma série de casos de diminutivização da língua, que foram postos de lado. Nosso trabalho pretende ampliar significativamente (embora não exaustivamente) o escopo desses casos de diminutivização, para que cheguemos primeiramente a generalizações empíricas mais robustas, e, depois, a análises mais consistentes para elas.

Hipótese Infixativa (HI)

inh é um infixo

(e.g. Back & Mattos 1972, Monteiro 1998, Kehdi 1999)

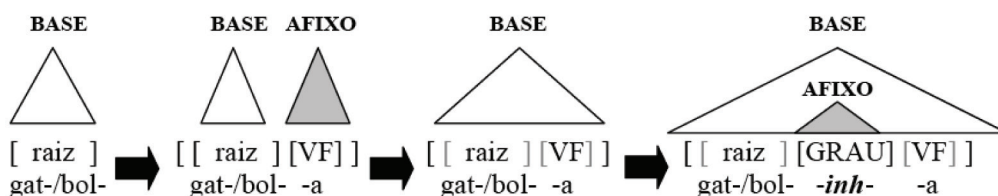


Figura 2. Hipótese Infixativa

Dados como (1) e (2) são facilmente acomodados por **HS-C**, **HS-R** e **HI** com igual sucesso, e com sistemas de regras de grau de complexidade equivalente. Tradicionalmente, assume-se alguma versão de **HS**. **HI**, no entanto, é uma possibilidade sequer mencionada pela maioria dos autores, discutida por muito poucos e efetivamente defendida (como modelo geral, até as últimas consequências) por (quase) ninguém.

Embora este trabalho discuta e proponha algumas ideias no plano teórico-analítico, ele se define como eminentemente descritivo.⁵ O objetivo aqui é, por um lado, contribuir com generalizações empíricas, chamando atenção para fatos que por muito tempo têm passado despercebidos, e, por outro lado, contribuir com *insights* analíticos acerca do valor e do poder heurístico de *famílias* de hipóteses/abordagens, que aqui correspondem a **HS** e **HI**, e que podem ser tecnicamente implementadas de várias maneiras, sem que isso altere significativamente as nossas principais conclusões. Obviamente, por mais que se queira abordar o objeto de estudo o mais ‘pré-teoricamente’ possível, é inevitável que haja um certo viés teórico de algum tipo, pois se faz necessária alguma metalinguagem abstrata que nos possibilite sistematizar as observações e formular as generalizações. Aqui, tal metalinguagem é pura e simplesmente o critério distribucional oriundo do Estruturalismo e presente, de uma forma ou de outra, em todas as teorias morfológicas contemporâneas, sejam elas lexicalistas, transformacionistas (pré- ou pós-sintáticas) ou mistas (o que, aliás, é uma distinção irrelevante nesse estágio inicial).

Desse modo, formulamos aqui algumas generalizações empíricas e algumas proto-hipóteses para a regra de afixação de ‘*inh*’, definidas numa metalinguagem semi-formalizada, o que tem a desvantagem de não ser algo matematicamente rigoroso o suficiente para que se atinja, de fato, adequação descritiva, mas com a vantagem de ser algo “aberto” e “intuitivo” o suficiente para ser passível de ser implementado em diversos quadros teóricos, e.g. *Morfologia Distribuída* (HALLE; MARANTZ, 1993), *Teoria de Otimalidade*

⁵ A partir da conceituação de Chomsky (1964, 1965), que classifica os níveis de adequação de uma teoria/análise linguística em (i) *observacional*, (ii) *descritivo* e (iii) *explicativo* (tal que (iii) pressupõe assimetricamente (ii), que por sua vez pressupõe assimetricamente (i)), nosso objetivo aqui é, numa terminologia rigorosa, atingir adequação *observacional* (i.e. identificar corretamente o que é ou não é uma expressão da língua), contribuindo com generalizações empíricas que possam lançar luzes sobre a adequação descritiva do fenômeno (i.e. uma descrição adequada da realidade psicológica da língua, modelando a representação simbólica da gramática na mente do(s) falante(s), de modo a prever a geração de todas e apenas as expressões da língua e dos significados a elas associados).

(MCCARTHY; PRINCE, 1986, 1993, 1999; PRINCE; SMOLENSKY, 1993), *Teoria de Multiprecedência & Linearização* (RAIMY, 2000). Aqui, não nos comprometemos com nenhum desses formalismos específicos, mas com proto-análises, a que chamamos de **HS-C**, **HS-R** e **HI**, e avaliamos os méritos e deméritos de cada uma delas.

Afixos ‘inh’ & ‘-zinh’

Esta seção tem por objetivo “separar o joio do trigo”, e isolar o que à primeira vista parece ser um subfenômeno do fenômeno que estamos investigando, mas que, após uma inspeção mais de perto, revela-se algo distinto, portanto fora do escopo da nossa pesquisa. Trata-se da existência de diminutivos em ‘**-zinh(o/a)**’, cuja distribuição é mais idiossincrática e mais complexa que a distribuição dos diminutivos em ‘**-inh(o/a)**’, tal como ilustrado em (3) e (4).

(3)	<i>grau normal</i>	<i>grau diminutivo</i>	(4)	<i>grau normal</i>	<i>grau diminutivo</i>
	a: bebê	b: bebê-zinho		a: pastéis	b: pastéis-zinhos
		c: *bebê-inho			c: ?paste[w]-zinhos
		d: *beb-inh-ê			d: *paste[w]-inhos
		e: *bebi-inho			e: *pastel-inhos ⁶

Em qualquer das três abordagens aqui consideradas, casos como (3) e (4) exigem mecanismos adicionais de composição e/ou epentetização. Argumentos contundentes em favor dessa conclusão encontram-se em Mattoso Camara (1970), Leite (1974), Moreno (1977), Brakel (1981), Lee (1992, 1995, 1999), Menuzzi (1993) e Bachrach e Wagner (2007), *inter alia*.

À primeira vista, parece ser possível tratar ‘**-zinh(o/a)**’ como um alomorfe de ‘**-inh(o/a)**’, sendo ‘**-zinh(o/a)**’ o caso marcado, condicionado pela base terminar em sílaba tônica (ou pesada). Note-se, por exemplo, que ‘cadáver’, embora não-oxítono, tem como correspondente diminutivo ‘cadáver-zinho’, e não *‘cadaver-inh(o)’. Essa observação acerca do padrão acentual parece ser, de fato, uma generalização válida, e, de algum modo, as regras do sistema precisam codificá-la. Entretanto, tal generalização ainda precisa ser refinada. Quando a base é oxítona (como em (3a)), o diminutivo só pode ser feito por meio de ‘**-zinh(o/a)**’. As bases não-oxítonas (cuja última sílaba não é pesada), como ‘livr-o’ e ‘músic-a’, tanto podem ser diminutivadas como em ‘livr-inh-o’ e ‘music-inh-a’, como podem, também, ser diminutivadas como em ‘livr-o-zinh(o)’ e ‘music-a-zinh(a)’. Ao analisarmos os dados (3) e (4) com atenção, vemos que, ao que tudo indica, ‘**-inh(o/a)**’ e ‘**-zinh(o/a)**’ são itens lexicais distintos em vários aspectos. Por um lado, ‘**-inh(o/a)**’ parece ser um morfema canônico, ocupando uma posição dentro de uma base. Por outro lado, ‘**-zinh(o/a)**’ seria, em primeira instância, classificável como aquilo que Mattoso Camara (1970) chamava de *forma dependente*. Mais especificamente, ‘**-zinh(o/a)**’ parece ser uma espécie de *adjetivo enclítico*, que forma compostos com bases nominais oxítonas (ou terminadas em sílabas pesadas). Note-se, por exemplo, que, em (4), há morfologia flexional de número antes e depois de ‘**-zinh(o/a)**’. Além disso, é possível que ‘**-inh(o/a)**’ e ‘**-zinh(o)**’ co-ocorram na mesma palavra, mas sempre com ‘**-inh(o)**’ sendo mais interno

⁶ Segundo Maria Cristina Figueiredo Silva (comunicação pessoal), em Português Europeu, a forma diminutiva de (4a) é (4e), em vez de (4b), em contraste com o PB.

(mais próximo da raiz) que *'zinh(o/a)'*. Por exemplo, pode-se dizer 'livrinhozinho', mas não *'livrozinho'.

Nomes terminados em /e/ átono (em forma subjacente)

Observem-se os dados (5), (6) e (7) abaixo, sendo aqueles em (5) formados por raiz inerentemente [-FEM] e aqueles em (6) e (7) por raízes inerentemente [+FEM].

(5)	[-FEM]		(6)	[+FEM]		(7)	[+FEM]
<i>normal</i>		<i>diminutivo</i>	<i>normal</i>		<i>diminutivo</i>	<i>normal</i>	
a: tigrE		b: tigrinho	a: MatildE		b: Matildinha	a: colher(E)	b: colherinha
c: *tigrinhe			c: *Matildinhe			c: *colherinhe	
d: tigrezinho			d: Matildezinha			d: colherzinha ⁷	

Passemos agora às previsões feitas por cada uma das três hipóteses concorrentes em relação a esses dados. **HS-R** prevê corretamente a geração de (5b) e (6b) e o bloqueio de (5c), (6c) e (7c); não fazendo nenhuma previsão errada quanto a esses dados. **HS-C** prevê erroneamente o bloqueio de (5b) e (6b) e a geração de (5c), (6c) e (7c); não fazendo nenhuma previsão correta quanto a esses dados, exceto pela sua compatibilidade com a geração de (5d) e (6d), que se daria por mecanismos independentes (cf. seção anterior). **HS-C** também prevê erroneamente o bloqueio de (5b) e (6b) e a geração de (5c), (6c) e (7c); não fazendo nenhuma previsão correta quanto a esses dados, exceto pela sua compatibilidade com a geração de (5d) e (6d), que se daria por mecanismos independentes (cf. seção anterior).

Dados desse tipo são melhor acomodados por **HS-R**. Aplicando-se a regra geral de **HS-R** descrita acima a (5), teríamos o apagamento da VF /e/ de 'tigr-e' e a sufixação de *'inho'*, por se tratar de uma raiz masculina. Dados como o nome feminino 'Matilde/Matildinha' em (6) são igualmente triviais de acordo com **HS-R**, dispensando maiores explicações. Casos como (7) revelam-se um pouco mais complexos e demandam um pouco mais de abstração. Basicamente, parece inescapável postular que 'colher' é, em forma subjacente, bimorfêmica, incluindo uma VF correspondente ao /e/-teórico mattosiano (cf. nota 1 acima). Aplicando-se a regra geral de **HS-R** descrita acima a (7), teríamos o apagamento da VF /e/ de 'colher-e' e a sufixação de *'inha'*, por se tratar de uma raiz feminina.

Em contrapartida, HS-C e HI preveem, erroneamente, *'tigr-inh-e' e *'colher-inh-e'. No caso de HS-C, a regra geral (cf. seção introdutória acima) determina que o sufixo *'-inh'* ocupe, na sequência de posições vazias que se seguem à raiz, o lugar que lhe caberia segundo as regras da língua, e que seria depois da raiz ('tigr-' ou 'colher') e antes da VF /e/ presente em forma subjacente. Isso sobregeraria dados negativos como *'tigr-inh-e' e *'colher-inh-e', e subgeraria dados positivos como *'tigr-inh(o)' e *'colher-inh(a)'. O problema enfrentado por HI é tão semelhante que parece ser uma variante notacional. A regra geral (exposta na seção 1 acima) tomaria como *input* as bases 'tigr-e', 'Matild-e' e 'colher-e', e geraria como *output* os dados negativos *'tigr-inh-e', 'Matild-inh-e' e *'colher-inh-e', após ter infixado *'inh'* entre a raiz ('tigr-' ou 'Matild-' ou 'colher') e a VF /e/.

⁷ Por alguma razão, há uma (quase) sinonímia entre os membros do par 'colherinha/colherzinha' com diferenças muito sutis, se houver. Já em pares como 'mulherinha/mulherzinha', há contrastes de significado bem mais salientes.

Nomes masculinos terminados em /a/ átono

Vejam os dados (8) e (9) abaixo, sendo os primeiros (8) formados por raízes inerentemente [-FEM] e os últimos (9) por raízes [+FEM] especificadas como [-FEM].

(8)	a:	samb-a	a':	samb-inh-a	a'':	*samb-inh-o
	b:	problem-a	b':	problem-inh-a	b'':	*problem-inh-o
	c:	sistem-a	c':	sistem-inh-a	c'':	*sistem-inh-o
	d:	fantasm-a	d':	fantasm-inh-a	d'':	*fantasm-inh-o
	e:	poet-a	e':	poet-inh-a	e'':	*poet-inh-o
(9)	a:	o atlet-a	a':	atlet-inh-a	a'':	*atlet-inh-o
	b:	o camarad-a	b':	camarad-inh-a	b'':	*camarad-inh-o
	c:	o judoc-a	c':	judoqu-inh-a	c'':	*judoqu-inh-o
	d:	o obstetr-a	d':	obstetr-inh-a	d'':	*obstetr-inh-o
	e:	o pianist-a	e':	pianist-inh-a	e'':	*pianist-inh-o
	f:	o Mirand-a	f':	Mirand-inh-a	f'':	*Mirand-inh-o
	g:	o Lim-a	g':	Lim-inh-a	g'':	*Lim-inh-o

Dados desse tipo apresentam um paradigma extremamente coeso e são tão produtivos na língua que nos permitem fazer a seguinte generalização empírica: *Diminutivos de nomes masculinos terminados em '-a' átono terminam em '-inh-a(-s)'*.

HS-R prevê erroneamente o bloqueio de (8a'-8e') e (9a'-g') e a geração de (8a''-8e'') e (9a''-g'') não fazendo nenhuma previsão correta. **HS-C** e **HI** fazem ambas as mesmas previsões corretas quanto à geração de (8a'-8e') e (9a'-g') e o bloqueio de (8a''-8e'') e (9a''-g''), não fazendo nenhuma previsão errada.

Tomemos 'samb-a' como exemplo prototípico. Pela regra geral de **HS-C**, o sufixo '**-inh**' ocuparia, na sequência de posições vazias após a raiz, o lugar que lhe caberia segundo as regras da língua, e que seria depois de 'samb-' e antes de '-a', gerando 'samb-inh-a', como desejado. Como já comentamos, **HI** teria um mecanismo tão semelhante que parece ser uma variante notacional à primeira vista (veremos adiante que não é). A regra geral tomaria como *input* a palavra 'samb-a', e geraria como *output* 'samb-inh-a', após ter infixado '**-inh-**' entre a raiz ('samb-') e a VF /a/.

Por outro lado, **HS-R** prevê, erroneamente, a existência de *'sambinho', *'probleminho', etc. De acordo com a regra geral de **HS-R**, é irrelevante qual seja a VF. O que importa é somente o gênero da raiz ser masculino ou feminino, determinando se a VF será substituída por 'inho' ou por 'inha', respectivamente.

Nomes femininos terminados em /o/ átono

Atentemos agora para os dados em (10), que exibem variação dialetal quanto às possibilidades de diminutivização.

(10)	a:	fot-o	a':	fot-inh-o	a'':	fot-inh-a
	b:	mot-o	b':	mot-inh-o	b'':	mot-inha

Apesar da variação dialetal (cf. a’-a” & b’-b”), tratam-se de dados para os quais os informantes têm juízos de aceitabilidade bastante robustos. Em nossos estudos preliminares, identificamos a existência de ao menos três dialetos, descritos abaixo:

- (11) *Dialetos/Idioletos do tipo A:* ✓foto/✓fotinhA/*fotinhO
 ✓moto/✓motinhA/*motinhO
- (12) *Dialetos/Idioletos do tipo O:* ✓foto/✓fotinhO/*fotinhA
 ✓moto/✓motinhO/*motinhA
- (13) *Dialetos/Idioletos do tipo A/O:* ✓foto/✓fotinhA (uma foto qualquer, sem importância)
 ✓foto ✓fotinhO (uma foto pequena)
 ✓moto/✓motinhA (uma moto qualquer, sem importância)
 ✓motinhO (uma moto pequena)

Além desses pouquíssimos casos mais robustos, há muitos outros em que, na sua grande maioria, os falantes de todos os dialetos vacilam ao serem solicitados a emitir julgamentos de aceitabilidade acerca de qual seria a forma diminutiva da palavra.

- (14) a: libido a’: ?libid-inh-a a’’: ?libid-inh-o (‘zinh’?)
 b: tribo b’: ?trib-inh-a b’’: ?trib-inh-o (‘zinh’?)
 c: Yoko c’: ?Yok-inh-a c’’: ?Yok-inh-o (‘zinh’?)
 d: Keiko d’: ?Keik-inh-a d’’: ?Keik-inh-o (‘zinh’?)
 e: lot-o e’: ?lot-inh-o e’’: ?lot-inha (‘zinh’?)

Por que o subgrupo minúsculo (10) é tão regular, enquanto o subgrupo mais numeroso, representado por (14), apresenta-se de modo caótico? O que os distinguiria? Frequência estatística de uso das formas diminutivas? Seriam as formas diminutivas em (10) meramente lexicalizadas e não o produto de uma regra? Seriam as formas (14a-e) desprovidas de correspondentes em grau diminutivo expresso morfologicamente (sendo necessárias formas perifrásticas sintáticas, e.g. ‘tribo pequena’)? Por quê? Para cada uma dessas questões, cabe ainda perguntarmos o porquê de os fatos serem como são, tanto no nível da adequação descritiva como no nível da adequação explicativa.⁸ Reconhecemos, entretanto, que ainda precisamos refinar muito a nossa metodologia de coleta de julgamentos finos de aceitabilidade para podermos afirmar algo mais sólido a esse respeito.

⁸ Pode-se suspeitar, inicialmente, que isso tenha a ver com o fato de que ‘foto’ e ‘moto’ seriam formas truncadas de ‘fotografia’ e ‘motocicleta’, respectivamente, enquanto nada parecido possa ser dito sobre ‘libido’ ou ‘tribo’. Assim, talvez possamos falar em duas “classes naturais”. Entretanto, não é óbvio que todos os falantes consultados tenham ‘foto’ e ‘moto’ como reais formas truncadas de ‘fotografia’ e ‘motocicleta’, respectivamente, portanto membros de um mesmo paradigma (fora isso, fatos da diacronia são tipicamente irrelevantes na representação da gramática do falante em termos sincrônicos). Ademais, no exemplo (14e), temos ‘loto’, que se originou do truncamento de ‘loteria’, mas, por alguma razão, não se comporta como ‘moto’ e ‘foto’ quanto à diminutivização.

Tentando Colocar um Pouco de Ordem no Caos: acomodando os dados às hipóteses através de mecanismos adicionais

Ajustando HI & HS-C

Vimos acima que **HI** e **HS-C** não dão conta de casos trivialmente explicáveis por **HS-R**, como ‘tigr-e/tigri-nh-o’ e ‘colher/colher-inh-a’. Uma maneira de sustentar **HS-C** ou **HI** e dar conta desses dados é emendar o modelo com uma regra morfológica adicional como em (15), que tem como consequência a realização do /e/ final (incluindo o /e/-teórico mattosiano) como [o] ou [a], conforme o gênero da raiz. Essa regra se aplicaria a um **caso marcado** do *output* da regra de diminutivização (por sufixação ou infixação). Tal ordenação se segue do fato de que a presença do morfema de diminutivo (DIM) é parte definitiva do contexto de aplicação da regra de ajuste em (15).⁹

- (15) $\{^{VF} /e/\} \rightarrow X / \{raiz\ nominal\ do\ gênero\ \alpha\} \{DIM\} ___$
onde $X = \{^{VF} /o/\} elsewhere$
 $X = \{^{VF} /a/\}$ se $\alpha = [+FEM]$

Vejamos alguns exemplos:

- (16) a: bol-a → bol-inha (caso não-marcado)
b: carr-o → carr-inho (caso não-marcado)
c: tigr-e → tigr-inh-e → tigr-inh-o (caso marcado)
d: colher-e → colher-inh-e → colher-inh-a (caso marcado)
e: Matild-e → Matild-inh-e → Matild-inha (caso marcado)

Adicionando-se (15) a um sistema do tipo **HI** ou **HS-C**, não há mais o problema de se sobregerar *‘tigr-inh-e’ e *‘colher-inh-e’ em lugar de *‘tigr-inh-o’ e *‘colher-inh-a’. O preço que se paga por isso é a postulação de uma regra adicional específica para esses casos.

No entanto, o ajuste promovido pela regra adicional em (15) não é suficiente para que as proto-hipóteses **HS-C** e **HI** possam dar conta da distinção entre os dialetos do tipo A (foto/fotinha) e do tipo O (foto/fotinho). Obviamente, é preciso reconhecer aqui que ainda é prematuro afirmar que tal variação dialetal não seria fruto de (mera) lexicalização. Somente estudos futuros (feitos com uma metodologia apropriada) poderão nos levar a uma conclusão nesse sentido. Além disso, mesmo que o contraste seja fruto de mecanismos gramaticais ligeiramente distintos em cada dialeto, tal ajuste adicional será necessário, sob alguma forma, em qualquer das três proto-hipóteses concorrentes. No que concerne a **HS-C** e **HI**, a ausência de tal ajuste adicional implica a previsão de que somente o padrão dos dialetos do tipo O (foto/fotinho) existiria. Se assumirmos **HS-C** e **HI**, e quisermos codificar a distinção entre os dois tipos de dialeto na gramática, e não no léxico, temos que postular que, nos dialetos do tipo A, tem-se a regra em (15), que afeta apenas a VF se ela for /e/, enquanto, nos dialetos do tipo O, tem-se a regra em (15’), que é análoga a (15), mas que define a VF a ser afetada em termos mais subespecificados (i.e. [-low]), afetando tanto /e/ como /o/.

⁹ Se tirássemos DIM da definição do contexto de aplicação da regra em (15), teríamos como consequência a sobregeração de *‘tigr-o’ e *‘colher-a’ em lugar de ‘tigr-e’ e ‘colher’.

- (15') $\{^{VF} / \text{[_-low]} / \} \rightarrow X / \{ \text{raiz nominal do gênero } \alpha \} \{ \text{DIM} \} \text{ ___}$
 onde $X = \{^{VF} / o / \}$ *elsewhere*
 $X = \{^{VF} / a / \}$ se $\alpha = [+FEM]$

O quadro a seguir apresenta uma série de exemplos da aplicação dessas duas regras aos mesmos dados. Cada uma prevê um dialeto.

Quadro 1. Aplicação da regra geral e de (15)/(15') aos dialetos do tipo O e dialetos do tipo A

Dialectos do tipo O			Dialectos do tipo A		
[HS-C ou HI] + regra (15)			[HS-C ou HI] + regra (15')		
carro	→ carrinho	(pela regra geral)	carro	→ carrinho	(pela regra geral)
bola	→ bolinha	(pela regra geral)	bola	→ bolinha	(pela regra geral)
samba	→ sambinha	(pela regra geral)	samba	→ sambinha	(pela regra geral)
tigre	→ tigrinho	(por (15))	tigre	→ tigrinho	(por (15'))
Matilde	→ Matildinha	(por (15))	Matilde	→ Matildinha	(por (15'))
foto	→ fotinho	(pela regra geral)	foto	→ fotinha	(por (15'))

Ajustando HS-R

Vimos que nomes masculinos terminados em /a/ invariavelmente formam diminutivos terminados em /a/ (e.g. ‘samb-a/samb-inh-a’). Tais dados são trivialmente explicáveis por **HS-C** e **HI**. Ao contrário, **HS-R** massivamente sobregera dados negativos como *‘sambinho’ e subgera dados positivos como ‘sambinha’. Uma maneira de sustentar **HS-R** e dar conta desses dados é emendar o modelo com uma regra adicional, segundo a qual **‘inha’** e **‘inho’** são excepcionalmente sufixados a nomes masculinos terminados em /a/ e a nomes femininos terminados em /o/, respectivamente. Esse mecanismo adicional de **HS-R**, sensível à identidade fônica da VF, geraria os dados desejados fazendo o sufixo de diminutivo “rimar”, por assim dizer, com a VF da forma que serve de *input* para a regra.¹⁰ Assim, nos casos não-marcados, a regra **HS-R** (definida como na **Figura 1**, do lado direito) seria aplicada. Nos casos marcados, aplica-se a regra em (17) abaixo, tal como exemplificado em (18). Isso então dá conta dos nomes masculinos terminados em /a/.

- (17) $\{^{VF} / a / \} \rightarrow \{^{DIM} / iNa / \} / \{ \text{raiz nominal } [-FEM] \} \text{ ___}$ (N = nasal palatal)

- (18)
- | | | | |
|----|---------------|--------------------|---|
| a: | samb-a | → samb-inha | <i>raiz inerentemente</i> [-FEM] |
| b: | problem-a | → problem-inha | <i>raiz inerentemente</i> [-FEM] |
| c: | (o) judoc-a | → (o) judoqu-inha | <i>raiz inerentemente</i> [αFEM] <i>usada como</i> [-FEM] |
| d: | (o) pianist-a | → (o) pianist-inha | <i>raiz inerentemente</i> [αFEM] <i>usada como</i> [-FEM] |

¹⁰ Usamos a expressão ‘por assim dizer’, e colocamos “rimar” entre aspas duplas, porque, a rigor, isso não tem as características daquilo que conhecemos por rima, e que é parte do fenômeno da linguagem como um todo. A rima, tal como a concebemos, e, tal como ela é usada como recurso estilístico, ocorre sempre entre duas (ou mais) sílabas tônicas, tal que ambas estão presentes no enunciado (ou seja, no eixo sintagmático, não no paradigmático). O que supostamente estaria acontecendo em (17) é que a VF átônica de DIM “assemelha-se fonicamente” à VF átônica de uma forma abstrata subjacente, que pode ter estado (e provavelmente esteve) presente na derivação da forma de superfície, mas que não está presente no enunciado efetivamente falado. Enfim, o que temos aqui é um “rimar com o silêncio”, que, até onde se sabe, não encontra correlato em nenhum outro fenômeno das línguas naturais.

Esta proto-análise, baseada na proto-hipótese **HS-R**, no entanto, precisa de outros ajustes para dar conta da distinção entre os dialetos do tipo A (foto/fotinha) e do tipo O (foto/fotinho). Se assumirmos **HS-R**, e quisermos codificar a distinção entre os dois tipos de dialeto na gramática, e não no léxico, temos que postular a regra em (19), cuja ‘anatomia’ é análoga à de (17), e, num certo sentido, exhibe o padrão inverso (concordância fônica com /o/ se a raiz for feminina).

(19) {^{VF} /o/} → {^{DIM} / iNo / } / { raiz nominal [+FEM] } ____ (N = nasal palatal)

Desse modo, poderíamos dizer que a diferença entre os dois tipos de dialeto residiria na presença ou ausência da regra em (19) – ou algo equivalente – no sistema, em adição à regra geral e à regra (17), conforme o quadro abaixo:

Quadro 2. Aplicação da regra geral e de (17)/(19) aos dialetos do tipo O e dialetos do tipo A

Dialetos do tipo A	Dialetos do tipo O
<i>HS-R + regra (17)</i>	<i>HS-R + regra (17) + regra (19)</i>
carro → carrinho (pela regra geral)	carro → carrinho (pela regra geral)
bola → bolinha (pela regra geral)	bola → bolinha (pela regra geral)
foto → fotinha (pela regra geral)	foto → fotinho (por (19))
tigre → tigrinho (pela regra geral)	tigre → tigrinho (pela regra geral)
Matilde → Matildinha (pela regra geral)	Matilde → Matildinha (pela regra geral)
samba → sambinha (por (17))	samba → sambinha (por (17))

Intermezzo

HS-R versus [HS-C ou HI]

Embora a primeira divisão natural entre as três proto-hipóteses seja separar **HS-C** e **HS-R** de um lado, contra **HI** do outro, por se tratar de sufixação *versus* infixação, vimos que, até agora, em suas ‘formas puras’, **HS-C** e **HI** têm coberturas empíricas que contrastam com a cobertura empírica de **HS-R**. Descobrimos que, empiricamente, a comparação que se coloca, num primeiro momento, é entre **HS-R** de um lado e **HS-C** e **HI** do outro. Basicamente, **HS-R** e o par (**HS-C**, **HI**) estão em distribuição complementar. Uma facilmente acomoda padrões que a outra não acomoda, e vice-versa. Assim, ambas precisam de mecanismos adicionais para atingir a cobertura empírica (quase) total.

Dos dois mecanismos auxiliares estipulados, aquele utilizado em **HS-C** e **HI** (i.e. (15, 15’)) é mais natural, por ser, essencialmente, análogo à concordância de gênero, algo próprio da(s) língua(s). O mecanismo adicional estipulado para **HS-R** (i.e. (17, 19)) é anti-natural por ser, essencialmente, concordância fônica (com /a/ ou com /o/), algo sem correlato no resto da língua, e, principalmente, algo que contradiz a própria regra geral de **HS-R**, baseada em concordância de gênero. A partir de agora, descartamos, portanto, **HS-R**, e seguiremos considerando **HS-C** e **HI** como alternativas.¹¹

¹¹ Há uma implementação técnica alternativa para **HS-R** (que nos foi sugerida por Marcelo Ferreira, em comunicação pessoal) que nos parece bastante promissora, tendo uma cobertura empírica relativamente ampla, não necessitando de regras adicionais *ad hoc* de reajuste fonológico. Por ora, não discutiremos essa possibilidade, pois ainda não a testamos o suficiente.

HS-C versus HI

Prima facie, **HS-C** parece melhor que **HI** se pensarmos que algo só é um infixos se puder penetrar numa base monomorfêmica, o que não vemos nos dados típicos. Chegamos a esse ponto, portanto, com a impressão de que **HS-C** é superior a **HI**, levando vantagem devido a um detalhe técnico concernente à definição de infixos. Mostraremos na seção seguinte que, ao contrário do que possa parecer à primeira vista, **HI** supera empiricamente **HS-C**. Tecnicamente, entretanto, dizer que algo só é um infixos se puder penetrar numa base monomorfêmica é uma simplificação prematura. Tomar isso como premissa é questionável, pois isso depende de outras premissas (questionabilíssimas) sobre derivações e representações, e sobre a natureza das regras morfológicas envolvidas nas afixações (KIPARSKY, 1982; MOHANAN, 1986).¹² Ademais, poder penetrar em bases monomorfêmicas não implica logicamente não poder penetrar em bases plurimorfêmicas. Ampliemos, agora, nossas considerações empíricas.

Evidências e Argumentos Sugestivos em Favor de HI

Os dados abaixo mostram a inserção do morfema de diminutivo em dados canonicamente analisados como monomorfêmicos. Não há, portanto, como acomodar tais dados em nenhuma das versões de **HS**.

(20)	a:	Lucas	a':	Luqu-inh-a(s)
	b:	Carlos	b':	Carl-inh-o(s)
	c:	Marcos	c':	Marqu-inh-o(s)
	d:	(o) Santos	d':	Sant-inh-os
	e:	Matos	e':	Mat-inh-o(s)
	f:	Guarulhos	f':	Guarul-inh-o(s) ?
	g:	Barretos	g':	Barret-inh-o(s) ?

Além disso, note-se que *'inh'* também tem a capacidade de se encaixar em unidades monomorfêmicas não-nominais, como advérbios, e, em alguns dialetos/idioletos, formas verbais gerundivas, tal como em (21).

(21)	a:	O Gesoel e a Ana estão [[namor][a][nd[inh]o]] no sofá da sala.
	b:	Eu vou tomar banho [agor[inh]a] mesmo.
	c:	Eu vou ali na farmácia [[corr][e][nd[inh]o]], e já te trago o remédio.
	d:	Ele vem bajula[nd[inh]o] a Ana todos os dias.

O tratamento semântico adequado para esses casos ainda é um mistério. Mas há muitos paralelos extremamente sugestivos que podem ser traçados entre as ocorrências de *'inh'* em (21) e todas as demais. Além de diminutivos de tamanhos propriamente ditos, há diminutivos afetivos ((21a) e (22a)), encurtadores de intervalo temporal ((21b-c) e (22c)), depreciativos ((21d?) e (22b)), etc., e estamos longe de ter um tratamento unificado para tudo isso.

¹² cf. a postulação do *Bracket Erasure Principle* em Kiparsky (1982), e toda a discussão que isso gerou subsequentemente, e continua gerando até hoje. Mesmo em modelos representacionais como a *Teoria de Otimalidade*, que lidam com o problema de forma diferente, a questão continua existindo (o que motiva os praticantes dessa teoria a tentar derivar “efeitos de opacidade” (cf. MOHANAN, 1986) a partir de mecanismos puramente representacionais (cf. HERMANS; VAN OOSTENDORP, 1999).

- (22) a: A Flávia disse: “meu namoradinho é muito gentil”.
b: O ex-namorado da Flávia disse: “o namoradinho da Flávia é um babaca.”
c.: Aguarde um momentinho.

A respeito de (21b), podemos especular que o enunciador quer passar a ideia de que ele fará o possível para encurtar o intervalo de tempo entre o ato de enunciação da sentença e início do evento de tomar banho em si. Em (21c), o enunciador usa o infixos *‘inh’* encaixado na unidade monomorfêmica do gerúndio *‘-ndo’* (grosso modo, denotadora de aspecto progressivo, que “abre” um intervalo de tempo dentro do qual a ação vai acontecer) para expressar que ele fará o possível para encurtar o intervalo de tempo dentro do qual a corrida até a farmácia ocorrerá. Em suma, ele vai correr mais rapidamente que o normal, para que o doente sofra por menos tempo à espera do remédio. É interessante notar que não há contraste evidente entre (21c) e “Eu vou ali na farmácia correndão, e já te trago o remédio” quanto ao conteúdo proposicional (embora os conteúdos informacionais possam diferir ainda que sutilmente). Basicamente, no segundo caso, o enunciador também passa a mensagem de que vai tornar a corrida até a farmácia o mais breve possível. O aumentativo parece denotar o aumento da brevidade da corrida, o aumento da velocidade do modo como se corre, o que equivale, pragmaticamente, à diminuição do intervalo temporal que acabamos de mencionar. Intuitivamente, isso parece fazer sentido. O grande problema é saber como modelar isso formalmente, calculando, em cada caso, o significado do todo a partir do significado das partes. Especificamente, qual é o significado básico de *‘inh’*? Será que estamos diante do mesmo *‘inh’* em todos os casos?

Ainda explorando o (escorregadio/perigoso, porém fértil) território das especulações, podemos estabelecer um paralelo entre (22b) e os diminutivos de tamanho canônicos. Grosso modo, é possível parafrasear o enunciado do ex-namorado da Flávia como “a virilidade e o caráter desse homem são de magnitude muito pequena, insuficiente para ele merecer namorar uma mulher como a Flávia”.

Coda

Haveria vários *‘inh’* no léxico, cada um tendo sua própria denotação, sua própria categoria, sendo usado nos contextos que lhe são apropriados? Quer estejamos diante de um único *‘inh’* ou de mais de um, é prematuro dizer se todas essas nuances de significado são obtidas na semântica composicional ou no campo da pragmática; ou mesmo parte aqui, outra parte lá. Caso se conclua que há mais de um *‘inh’*, não é logicamente necessário que eles sejam todos sufixos ou todos infixos. Em última instância, isso é uma questão empírica, e que talvez possa variar de um dialeto/ídioteleto para outro. Além disso, a missão científica de dar uma explicação unificada para os sub-fenômenos do fenômeno não implica assumir sem questionamento que todos os dialetos/ídioteletos do PB são explicáveis através da mesma hipótese dentre as três discutidas aqui.¹³

Nesse estágio inicial da pesquisa, apesar da incerteza de estarmos diante de um mesmo fenômeno ou de vários, as regularidades dos paradigmas que detectamos – tanto em dados já bastante conhecidos como em dados que nós próprios “escavamos” – apontam

¹³ O fato de que, por exemplo, os “gerúndios diminutivizados” existem em alguns dialetos e não em outros pode sugerir uma solução lexicalista, mas isso não é uma necessidade lógica.

muito sugestivamente na direção de que todos esses casos podem efetivamente ser instanciações particulares de um mesmo mecanismo gramatical. Perseguindo essa ideia, comparamos três hipóteses concorrentes, constatamos que, em “estado bruto”, nenhuma delas é capaz de acomodar todos os casos mais prototípicos de diminutivização, sempre demandando postulações adicionais. Entretanto, a proto-análise infixativa, **HI**, revelou-se aquela cujos ajustes necessários são mais naturais, com menos custo teórico. Quando dados menos triviais são acrescentados à base empírica a ser descrita e analisada, **HI** mostra-se uma candidata ainda mais plausível para analisar unificadamente todos os fatos, posto que as duas proto-análises sufixativas alternativas parecem enfrentar problemas bem maiores diante dos dados menos prototípicos, mas que são igualmente reais, produtivos e relevantes (cf. (20) e (21)). Não obstante, ainda restam algumas lacunas à espera de investigação mais apurada (cf. (10) e (14)). Saliente-se, no entanto, que tais lacunas ainda permanecem inexplicadas, qualquer que seja a proto-análise que adotemos; além de outras lacunas não tratadas aqui e deixadas para serem estudadas em pesquisas futuras (cf. seção apêndice a seguir). Ademais, mesmo ainda não tendo encontrado explicações para tais questões, é importante destacar que muitos desses problemas só vieram a ser efetivamente detectados e explicitamente colocados a partir desses dados novos que foram “escavados” exatamente usando-se como “pá” a hipótese infixativa.

Encore

Conforme admitimos ao final da seção de conclusão acima, a escavação está apenas começando. Abaixo, apresentamos uma pequena (mas significativa) amostra de fatos que ainda jazem no subterrâneo, e que pedem por serem trazidos à tona em investigações futuras.

(23)	a:	Lucas	a':	Luqu-inh-as	a'':	Luqu-inh-a
	b:	Carlos	b':	Carl-inh-os	b'':	Carl-inh-o
	c:	César	c':	*Ces-inh-ar	c'':	Ces-inh-a
	d:	Rubens	d':	Ruben-zinho	d'':	Rub-inh-o
	e:	Mercedes	e':	Mercede-zinha	e'':	Merced-inh-a
	f:	Gonçalves	f':	?	f'':	?
	g:	Torres	g':	?	g'':	?
	h:	Lopes	h':	?	h'':	?
	i:	Gomes	i':	?	i'':	?
	j:	Tales	j':	?	j'':	?
	k:	Mercedes	k':	?	k'':	?

(24)	a:	colher	→	colherinha	(colherzinha)
	b:	mulher	→	mulherinha	(mulherzinha)
	c:	flor	→	*florinha	(florzinha)
	d:	cor	→	*corinha	(corzinha)
	e:	mar	→	*marinho	(marzinho)
	f:	par	→	*parinho(parzinho)	

(25)	a:	chanceler	→	*chancelerinho/*chancelerinha/(✓chancelerzinh(o/a))
	b:	militante	→	*militantinho/*militantinha/*militantinho/(✓militantezinh(o/a))

Já numa primeira inspeção rápida, o leitor pode facilmente apreciar a magnitude e a complexidade dos problemas que esses dados representam para todas as hipóteses aqui consideradas, inclusive para a própria **HI**. Eles demandam uma extensa discussão que extrapolaria os limites deste artigo, forçando-nos a reformular/refinar cada uma das (proto-) hipóteses aqui consideradas, bem como a formular e testar novas hipóteses. Deixamos tal tarefa para trabalhos futuros.

REFERÊNCIAS BIBLIOGRÁFICAS

- BACK, E.; MATTOS, G. *Gramática construtural da língua portuguesa*. São Paulo: FTD, 1972.
- BACHRACH, A.; WAGNER, M. Syntactically Driven Cyclicity vs. Output-Output Correspondence: the case of adjunction in diminutive morphology. *UPenn Working Papers in Linguistics*, Philadelphia, v. 10, n. 1, p. 157-173, 2007.
- BRAKEL, A. Boundaries in a Morphological Grammar of Portuguese. *Word*, New York, v. 32, p. 193-212, 1981.
- CHOMSKY, N. *Current issues in linguistic theory*. Mouton: The Hague, 1964.
- _____. *Aspects of the theory of syntax*. Cambridge: The MIT Press, 1965.
- HALLE, M.; MARANTZ, A. Distributed Morphology and the pieces of inflection. In: HALE, K.; KEYSER, S. J. (Orgs.). *The View from Building 20*. Cambridge: MIT Press, 1993. p. 111-176.
- HERMANS, B.; VAN OOSTENDORP, M. *The Derivational Residue in Phonological Optimality Theory*. Amsterdam: John Benjamins, 1999.
- KEHDI, V. O problema do infixos em português. *Filologia e Linguística Portuguesa*, São Paulo, n. 3, p. 191-196, 1999.
- KIPARSKY, P. 1982. Lexical morphology and phonology. In: YANG, I.-S. (Org.). *Linguistics in the Morning Calm*. Seoul: Hanshin, 1982. p. 3-91.
- LEE, S.-H. Fonologia Lexical do Português. *Cadernos de Estudos Linguísticos*, Campinas, v. 23, p. 3-91, 1992.
- _____. *Morfologia e Fonologia Lexical do Português do Brasil*. 1995. Tese (Doutorado em Linguística, Área de concentração: Morfofonologia) Universidade Estadual de Campinas, Campinas.
- _____. Sobre a formação do diminutivo do português brasileiro. *Revista de estudos da linguagem*, Belo Horizonte, v. 8, p. 113-123, 1999.
- LEITE, Y. *Portuguese Stress and Related Rules*, 1974. Tese (Doutorado em Linguística, Área de concentração: fonologia). University of Texas, Austin.
- MATTOSO CAMARA JR., J. M. *Estrutura da Língua Portuguesa*. Rio de Janeiro: Vozes, 1970.
- _____. Considerações sobre o Gênero em Português. In: _____. *Dispersos*. Rio de Janeiro: Fundação Getúlio Vargas, 1972. p. 115-129.

- MCCARTHY, J.; PRINCE, A. *Prosodic Morphology*. Amherst, New Brunswick: University of Massachusetts; Rutgers University, 1986. (ms.)
- _____. *Prosodic Morphology I: Constraint interaction and satisfaction*. Amherst, New Brunswick: University of Massachusetts; Rutgers University, 1993. (ms.)
- _____. Faithfulness and identity in Prosodic Morphology. In: KAGER, R.; VAN DER HULST, H.; ZONNEVELD, W. (Orgs.). *The prosody-morphology interface*. Cambridge: Cambridge University Press, 1999. p. 218–309.
- MENUZZI, S. *On The Prosody of the Diminutive Alternation -inho/-zinho in Brazilian Portuguese*. Leiden: HIL/University of Leiden, 1993.
- MOHANAN, K. P. *The theory of lexical phonology*. Dordrecht: Reidel, 1986.
- MONTEIRO, J. L. Quem disse que não há infixos em português? In: CONGRESSO NACIONAL DE LINGUÍSTICA E FILOGIA, II, 1998, Rio de Janeiro. *Anais...*, Rio de Janeiro, Círculo Fluminense de Estudos Linguísticos e Filológicos, 1999, <http://www.filologia.org.br/anais/anais_iicnlf50.html> Acesso em: 28 jun. 2011.
- MORENO, C. *Os Diminutivos em -inho e -zinho, e a Delimitação do Vocábulo Nominal em Português*. 1977. Dissertação (Mestrado em Linguística, Área de concentração: morfologia). IL/UFRGS, Porto Alegre.
- PRINCE, A.; SMOLENSKY, P. *Optimality Theory: Constraint Interaction in Generative Grammar*. New Brunswick: Rutgers University Center for Cognitive Science, 1993.
- RAIMY, E. *The phonology and morphology of reduplication*. Berlin: Mouton, 2000.
- SILVA, M. C. P. S.; KOCH, I. V. *Linguística Aplicada ao Português: Morfologia*. São Paulo: Cortez, 1983.
- ZANOTTO, N. *Estrutura Mórfica da Língua Portuguesa*. Caxias do Sul: EDUCS, 1986.