

REVISTA DA

# ABRALIN

---

ASSOCIAÇÃO BRASILEIRA DE LINGUÍSTICA



REVISTA DA  
**ABRALIN**  
ASSOCIAÇÃO BRASILEIRA DE LINGUÍSTICA

REVISTA DA ABRALIN	VOLUME 02	NÚMERO X	JUL/DEZ 2010.
--------------------	-----------	----------	---------------

**REVISTA DA  
ABRALIN**

---

ASSOCIAÇÃO BRASILEIRA DE LINGUÍSTICA

CONSELHO EDITORIAL

## **SUMÁRIO**

### **ARTIGOS**

O ARTIGO DEFINIDO FRENTE A PRONOMES POSSESSIVOS NA LITERATURA BRASILEIRA DO SÉCULO XIX.....	9
<i>Ane Schei - ....</i>	
A VARIAÇÃO NÓS / A GENTE NO DIALETO MINEIRO: INVESTIGANDO A TRANSIÇÃO .....	39
<i>Francisca Paula Soares Maia - Universidade Federal de Minas Gerais (UFMG)</i>	
CONTATO LINGUÍSTICO E MUDANÇA LINGUÍSTICA NO NOROESTE AMAZÔNICO: O CASO DO KOTIRIA (WANANO).....	65
<i>Kristine Stenzel- Universidade Federal do Rio de Janeiro (UFRJ)</i>	
RELACÕES PREDICATIVAS DAS PREPOSIÇÕES .....	97
<i>Márcia Barreto Berg - Universidade Federal de São João del Rei</i>	

PERCURSO DO ARTIGO DEFINIDO DIANTE DE N PRÓPRIO NA FALA DE CRIANÇAS EM FASE DE AQUISIÇÃO.....	113
<i>Patrícia Vargas Alencar - UNIRIO</i>	
 PADRÕES DE SELETIVIDADE NA PRODUÇÃO AGRAMÁTICA E DISTINÇÃO ENTRE MOVIMENTOS SINTÁTICOS NA COMPUTAÇÃO ON-LINE .....	135
<i>Ricardo Joseh Lima - Universidade Estadual do Rio de Janeiro</i>	
<i>Letícia M. Sicuro Corrêa - Pontifícia Universidade Católica do Rio de Janeiro</i>	
<i>Marina Rosa Ana Augusto - Universidade Estadual do Rio de Janeiro.</i>	
 VARIATION AND OPTIMALITY THEORY: REGRESSIVE ASSIMILATION IN VIMEU PICARD .....	165
<i>Walcir Cardoso</i>	

# **ARTIGOS**



# O ARTIGO DEFINIDO FRENTE A PRONOMES POSSESSIVOS NA LITERATURA BRASILEIRA DO SÉCULO XIX

ANE SCHEI

## RESUMO

*Neste artigo examina-se a variação entre o uso e o não uso do artigo definido antes de pronome possessivo seguido de substantivo em sete romances do século XIX. A análise mostra que há diferenças consideráveis entre os escritores, confirmando o resultado de outros estudos de que no português brasileiro o uso do artigo definido antes de possessivo varia bastante de um corpus para outro.*

## ABSTRACT

*This article examines the variation in the use of the definite article before possessive pronoun followed by a noun in seven 19<sup>th</sup> century novels. The analysis shows considerable differences between the authors, and thus confirms the results of other studies that in Brazilian Portuguese the use of the definite article before possessive pronoun varies from one corpus to another.*

## PALAVRAS-CHAVE

*artigo definido, pronome possessivo, língua literária, século XIX*

## KEY WORDS

*definite article, possessive pronoun, literary language, 19<sup>th</sup> century*

## Introdução

O uso do artigo definido antes de pronome possessivo seguido de substantivo é variável no português brasileiro (doravante PB): às vezes usa-se o artigo, outras vezes não:

- (1) *O seu trajo* era o comum em viagem: (Taunay, 61)
- (2) *Sua história* tem pouca coisa de notável. (Almeida, 67)

O presente trabalho tratará da variação entre o uso e o não uso do artigo antes de possessivo na literatura brasileira do século XIX ou, mais exatamente, em sete romances: faremos uma análise quantitativa para vermos com que frequência os sete escritores estudados usam o artigo. Outras pesquisas já apontaram alguns fatores que em maior ou menor grau afetam o uso do artigo, e aqui veremos se esses fatores atuam da mesma maneira no nosso *corpus*.

O *corpus* consiste nos seguintes romances (ano da primeira edição entre colchetes):

Joaquim Manuel de Macedo (1820-1882): <i>A Moreninha</i>	[1844]
Manuel Antônio de Almeida (1831-1861): <i>Memórias de um sargento de milícias</i>	[1854-1855]
José de Alencar (1829-1877): <i>Lucíola</i>	[1862]
Bernardo Guimarães (1825-1884): <i>O garimpeiro</i>	[1872]
Visconde de Taunay (1843-1899): <i>Inocência</i>	[1872]
Aluísio Azevedo (1857-1913): <i>O cortiço</i>	[1890]
Machado de Assis (1839-1908): <i>Dom Casmurro</i>	[1899]

O principal objeto de estudo será a narrativa dos livros; os diálogos só serão estudados resumidamente e à parte, na seção 5. Note-se que com ‘narrativa’ nos referimos não só à narrativa propriamente dita, mas também a trechos descriptivos e dissertativos, ou seja, ‘narrativa’ exprime o oposto de ‘diálogo’. Foram também deixados de lado diálogos indiretos, assim como prefácios, cartas, citações de livros e jornais, etc.,

ou seja, trechos em que o autor, por assim dizer, fala com voz diferente da usada na narração propriamente dita. As referências dos exemplos tirados dos livros examinados consistem no nome ou sobrenome do escritor e o número da página.

Na seção 2 faremos um breve resumo de algumas pesquisas anteriores, baseando-nos principalmente na tese de doutoramento de Giselle Machline de Oliveira e Silva: *Estudo da regularidade na variação dos possessivos no português do Rio de Janeiro* (1982). (Uma grande parte desta exposição já foi incluída num outro trabalho nosso (Schei 2007).) Aqui identificaremos os fatores que depois serão testados no nosso *corpus*. Na seção 3 definiremos alguns casos que não serão incluídos na análise por nunca ou quase nunca ocorrerem com o artigo; esses casos são excluídos para não enviesarem as análises quantitativas do material relevante, ou seja, os casos com variação livre. A análise da narrativa dos livros é apresentada na seção 4, e na seção 5 diremos algumas palavras sobre os diálogos. Na seção 6 faremos uns comentários finais.

Quanto às percentagens há algumas ligeiras diferenças entre as tabelas 2, 6 e 8 do presente trabalho e as tabelas correspondentes em Schei (2007); isso é devido a termos classificado alguns exemplos de maneira diferente nos dois trabalhos. Observe-se também que nas análises quantitativas arredondamos sempre as percentagens a números inteiros, não só nos cálculos nossos; também quando retomamos outros estudos arredondamos as percentagens daquelas pesquisas quando nelas se usam decimais.

Também temos que chamar a atenção para o fato de nem todas as tabelas deste trabalho, assim como outras referências a análises quantitativas, incluírem o mesmo tipo de dados. Nossa própria análise exclui os casos particulares apresentados na seção 3, mas as tabelas 4 e 5, baseadas na tese de Silva, incluem todas as ocorrências de possessivos (Silva 1982:356). Também os demais dados de Silva, mencionados no texto, diferem dos nossos da mesma maneira, mas julgamos que esse fato não impede que possamos comparar seus resultados com os nossos

sem grandes vieses. Acrescente-se que Silva baseia suas conclusões sobretudo em probabilidades, e só em menor grau em frequências relativas, mas para mais facilmente podermos comparar seus dados com os nossos, optamos por só apresentar suas frequências relativas.

Como acima foi dito, partes da seção 2 também se encontram no nosso artigo “Para o estudo do artigo definido antes de pronome possessivo no português brasileiro: algumas observações” no volume V de *Para a história do português brasileiro*. Infelizmente, o texto daquele artigo foi modificado pelo revisor de maneira que foram introduzidos vários erros. Entre outras coisas, o revisor alterou, inexplicavelmente, o texto em citações de Said Ali e Giselle Machline de Oliveira e Silva. Queremos aqui deixar claro que as citações transcritas no presente artigo são as corretas.

## 1 Panorama histórico e a variação no PB

Considerando primeiro a evolução diacrônica e as diferenças entre o português europeu (doravante PE) e o PB, verifica-se que nos primeiros documentos escritos em português o artigo era pouco usado, mas no PE passou a ser cada vez mais comum ao longo dos séculos, e hoje é, conforme Cunha & Cintra (1991:216), “praticamente obrigatório”, com exceção de alguns casos particulares. Um estudo de Said Ali nos dá a seguinte imagem da evolução<sup>1</sup>:

---

<sup>1</sup> Com o seguinte método e *corpus*: “Tal estatística sem pretensões a rigor absoluto foi por mim obtida, examinando, em páginas seguidas, todos os casos (em número de 100 a 150 para cada autor) não sujeitos a regras especiais e portanto parecendo permitir o emprego de possessivo com ou sem artigo. Ministraram exemplos: Fernão Lopes, *Crônica de D. João*, pág. 161 a 200; Camões, *Lusíadas*, cantos V a VIII; Vieira, *Sermões*, vol. 5, pág. 1 a 45; Herculano, *Eurico*, pág. 1 a 71.” (Said Ali 1964:96)

TABELA 1 - Baseada em Said Ali (1964:96-97): a evolução diacrônica do PE.

<b>corpus</b>	<b>frequência relativa de casos com artigo</b>
Fernão Lopes	5% aproximadamente
Camões	30%
Vieira	mais de 70%
Herculano	mais de 90%

A análise diacrônica do PE em Silva (1982:314) confirma a evolução descrita por Said Ali, chegando a uma percentagem de 90% em jornais do século XX, e uma análise nossa de alguns romances portugueses do século XIX<sup>2</sup> e do fim do século XX<sup>3</sup> mostra que o uso do artigo no PE de fato se generalizou (tabelas 2 e 3). Queremos lembrar que excluímos dos dados das nossas tabelas certos casos (apresentados na seção 3) nos quais a variante sem artigo é quase categórica, ou seja, o fato de Eça e os escritores da tabela 3 apresentarem uma percentagem de 100% não significa que não haja possessivos sem artigo nos livros em questão; existem ocorrências sem artigo, mas esses casos se encontram entre os excluídos por serem apostos, nomes de parentesco, vocativo, etc.

TABELA 2 - Romances portugueses do século XIX.

<b>autor</b>	<b>frequência relativa de casos com artigo</b>
Herculano	92%
Camilo	90%
Eça	100%

<sup>2</sup> Alexandre Herculano: *Eurico*; Camilo Castelo Branco: *Amor de perdição*; Eça de Queiroz: *O crime do Padre Amaro*.

<sup>3</sup> António Lobo Antunes: *Os cíus de Judas*; Américo Guerreiro de Sousa: *Os cornos de Cronos*.

TABELA 3 - Romances portugueses do século XX.

autor	frequência relativa de casos com artigo
Lobo Antunes	100%
Guerreiro de Sousa	100%

No PB, contudo, o uso do artigo não cresceu como no PE. Dado que a colonização do Brasil só começou no século XVI, o ponto de partida do PB já deve ter tido uma maior frequência do artigo do que a língua medieval, mas enquanto a frequência em Portugal aumenta cada vez mais até uma generalização, o PB muda muito pouco, conforme os dados de Silva, apresentados na tabela 4.<sup>4</sup> Como se vê, Vieira, que na análise do PE feita por Said Ali usa o artigo em mais de 70% dos casos, foi por Silva incluído entre os materiais brasileiros, apresentando uma percentagem de 41%. Silva (1982:356) comenta: “A discrepância [...] pode ser devida seja ao menor número de dados de Ali (no máximo 150, enquanto usamos 617) seja ao fato de termos usado todos os possessivos enquanto Ali eliminou os de ‘regras especiais’ (quais?).”

TABELA 4 - Baseada em Silva (1982:315): a evolução diacrônica do PB.

corpus	frequência relativa de casos com artigo
Vieira	41%
Documentos séc XVII	11%
Documentos séc XVIII	34%
Bernardo Guimarães	26%

<sup>4</sup> Padre Antônio Vieira: *Sermões*; Documentos da Câmara do Rio de Janeiro no século XVII; Documentos sobre a Inconfidência Mineira; Bernardo Guimarães: *A Escrava Isaura*.

A análise de Silva do PB de hoje,<sup>5</sup> resumida na tabela 5, confirma que no PB não houve uma generalização do artigo, e uma análise nossa de alguns romances do fim do século XX<sup>6</sup> (Tabela 6) mostra que a frequência das duas variantes pode variar bastante de um romance para outro.

TABELA 5 - Baseada em Silva (1982:311-312): o PB atual.

<b>corpus</b>	<b>frequência relativa de casos com artigo</b>
corpus oral	43%
jornais	29%
fotonovelas	35%
quadrinhos	59%

TABELA 6 - Romances brasileiros do século XX.

<b>autor</b>	<b>frequência relativa de casos com artigo</b>
Dourado	82%
Fonseca	37%
Luft	39%
Montello	48%
Queiroz	89%
Scliar	23%

Acrescente-se que no *corpus* do Projeto NURC, ou seja, na fala de informantes com formação universitária, a frequência do artigo é de 65% (Neves 1993:175).

<sup>5</sup> *Corpus* oral: dois grupos de jovens, universitários e alfabetizandos; jornais: *Jornal do Brasil* e *O Globo*; fotonovelas: *Amiga* e *Sétimo Céu*; quadrinhos: *Mônica*, *Cebolinha*, *Vaca Voadora*, *Pelezinho*, *Sítio do Pica-Pau Amarelo* e *Os Trapalhões*.

<sup>6</sup> Autran Dourado: *Confissões de Narciso*; Rubem Fonseca: *Vastas emoções e pensamentos imperfeitos*; Lya Luft: *Exílio*; Josué Montello: *Enquanto o tempo não passa*; Rachel de Queiroz: *Dôra*, *Doralina*; Moacyr Scliar: *Os voluntários*.

Diante dessa variação do PB no que diz respeito à frequência relativa do artigo pode-se perguntar se há diferenças regionais. Conforme Lucchesi (1993:91) há: “In Portugal and the south of Brazil, the definite article and the possessive normally co-occur, whereas in northern and northeast Brazil the article is normally absent.” Contudo, a análise do *corpus* do Projeto NURC, já mencionada, não mostra diferenças notáveis entre as cinco cidades estudadas: Recife 66%, Salvador 64%, Rio de Janeiro 68%, São Paulo 65% e Porto Alegre 64%. Se observarmos os romances do fim do século XX da tabela 6, vemos que, ao contrário das cinco cidades do Projeto NURC, há diferenças consideráveis entre os seis escritores, mas se essas diferenças são devidas a origem regional contradizem o que diz Lucchesi, já que Queiroz, com a frequência mais alta, é do Nordeste, enquanto Scliar e Luft, com frequências baixas, são do Rio Grande do Sul.

Existe também a possibilidade de a variação no uso do artigo ser devida a algum tipo de pressão normativa, mas conforme Silva (1982:274) muitas pessoas nem sequer estão conscientes dessa variação e, por conseguinte, não há uma consciência sobre como se ‘deve’ fazer. Apesar disso, Silva (1982:385) afirma que fatores sociais podem influir: “Quanto às classes notou-se que tanto os falantes provenientes de pais mais instruídos quanto os mais instruídos eles mesmos mostraram significativo desfavorecimento do artigo”. No entanto, como constatamos acima, os informantes com formação universitária do Projeto NURC usam o artigo em 65% dos casos, o que consideramos uma frequência relativamente alta.

Há, contudo, fatores linguísticos que afetam o uso do artigo. As gramáticas costumam apontar alguns casos em que normalmente não se usa o artigo, por exemplo aposto e nome de parentesco (para maiores detalhes, ver seção 3), mas no estudo detalhado de Silva são detectados mais alguns fatores linguísticos que podem influir em maior ou menor grau: número, o próprio pronome possessivo, se o possessivo vem logo depois de pausa ou não, se o possessivo é precedido de preposição, e sobretudo o fator ‘especificidade’.

Vejamos primeiro como Silva define o termo ‘especificidade’ e como este fator atua: com ‘especificidade’ Silva (1982:268) não quer “afirmar, em termos absolutos, que o artigo frente ao possessivo especifica o que equivaleria a dizer que, sem ele, o possessivo é inespecífico, indefinido, o que sabemos ser *irreal*”; o que aquele termo quer dizer é que “os [casos] sem artigo são aqueles em que o falante não necessita destacar elemento(s) do conjunto: a informação ou não é necessária ou, pelo contrário, já está bem clara” (Silva 1982:271). A função especificadora do artigo é assim descrita em outro estudo de Silva (1996:125):

Há leve diferença semântica entre os exemplos ‘seu livro’ e ‘o seu livro’. O primeiro pode ser parafraseado como ‘um dos seus livros’ enquanto o segundo tenderá a ser interpretado como ‘esse seu livro específico’. O uso do artigo está ligado à questão da especificidade. Quando o falante destaca apenas um dentro de um conjunto de elementos possuídos, há maior especificidade e o uso do artigo é maior.

Nos *corpora* sincrônicos e em quase todos os diacrônicos analisados por Silva (cujas frequências foram apresentadas nas tabelas 4 e 5) “os possuídos ‘específicos’ sempre favorecem a presença do artigo, enquanto os ‘não específicos’ a desfavorecem” (Silva 1982:358). A especificidade também explicaria o fato de nomes de parentesco (p.ex. *pai*, *irmão*) raramente terem artigo, já que são, por sua natureza, suficientemente específicos. Quanto ao fator ‘número’, o artigo é menos frequente no plural do que no singular, o que também seria um efeito do fator especificidade: “Havíamos postulado que o plural já trazia em si uma definição intrínseca, que fizesse com que o artigo especificado se tornasse inútil. O plural refere-se ao conjunto inteiro. Não necessitando de artigo para expressar veladamente a idéia de *todos*, o artigo torna-se redundante.” (Silva 1982:367). A especificidade também explica o fato de o aposto raramente ter artigo: “um aposto é uma definição, uma especificação e, mais do que o campo semântico de parentesco, deve ser

‘super determinado’, tornando o artigo redundante” (Silva 1982:272).

É importante notar que na tese de Silva a especificidade só favorece o emprego do artigo; não o torna obrigatório. Se, por exemplo, o *corpus* oral da tabela 5 for dividido em dois grupos – possuídos específicos e não específicos – o artigo ocorre em 49% dos casos específicos e em 30% dos não específicos, e se forem juntados os três *corpora* escritos da tabela 5, as percentagens correspondentes são 43% e 32%, respectivamente (Silva 1982:298-299). O mesmo vale para outros fatores examinados por Silva: quanto a número, como já foi dito, o artigo é menos frequente com plural do que com singular; quando o possessivo vem depois de uma pausa, o artigo é menos frequente do que quando o possessivo vem precedido de outra palavra; uma preposição antecedendo o possessivo favorece o artigo; quanto aos diferentes pronomes possessivos, o artigo é menos comum com *seu* do que com os demais. No entanto, em nenhum desses casos o artigo é categórico com uma das variáveis; veja-se, por exemplo, o fator ‘pausa’ no *corpus* escrito: com possessivo antecedido de pausa o artigo ocorre em 20% dos casos, e quando não há pausa antes do possessivo o artigo ocorre em 44% dos casos (Silva 1982:308).

A influência do próprio pronome possessivo não pôde ser muito bem analisada por Silva na obra de 1982 porque nem todos os possessivos ocorreram em número representativo no seu *corpus*, pelo que a questão foi retomada em um estudo posterior, de 1996, no qual também o fator ‘inerente’ foi estudado. O *corpus* da pesquisa de 1996 consiste em língua falada do *corpus* NURC/RGS e do *corpus* Leda Bisol, assim como o livro *O Retrato* de Érico Veríssimo.<sup>7</sup> Já que não havia ocorrências de *seu* da terceira pessoa no *corpus* oral, uma comparação entre este pronome e os demais só pôde ser feita no romance de Veríssimo, e a análise confirmou o resultado de Silva (1982) de que o possessivo de terceira pessoa – *seu* – inibe o artigo, ao contrário dos demais pronomes possessivos: a frequência do artigo era de 66% com *meu* e *nossa*, 55% com *teu*, *vostra* e

<sup>7</sup> Essa apresentação do *corpus* de Silva (1996) é dada em Silva (1986:240).

*seu* da segunda pessoa (*seu* = *de você(s)*), e apenas 16% com *seu* da terceira pessoa (*seu* = *dele(s)/ dela(s)*). Uma análise em separado de *seu* da segunda pessoa mostra que este pronome tinha artigo em 53% dos casos, o que confirma que a diferença entre *seu* da terceira pessoa e os demais pronomes é semântica e não fonológica (Silva 1996:136).

Depois de ter constatado no estudo de 1982 que nome de parentesco inibe o artigo, no trabalho de 1996 Silva faz uma divisão mais detalhada dos substantivos que não significam parentesco para estudar o fator ‘inerente’, dividindo o material nos seguintes quatro grupos: relações humanas outras que parentes; possuídos não-inerentes; possuídos inerentes; partes do corpo. A definição de possuído inerente e não-inerente é expressa assim (Silva 1996:132): “Foram considerados possuídos não-inerentes os objetos que só eventualmente são possuídos (revista, cadeira) bem como características mais abstratas que não são obrigatórias (nervoso, opinião). Já características e objetos inerentes são os obrigatoriamente possuídos na nossa cultura ou que têm pelo menos grande expectativa de sê-lo (vida, alma, pente, casa).”

Silva não encontrou nenhuma descrição sobre o que seria considerado inerente na cultura brasileira, pelo que teve que fazer sua divisão baseando-se na sua intuição, mas observa que conserva “à parte a categoria *partes do corpo* como controle, já que estas são inerentes por natureza” (Silva 1996:133). As relações humanas, por seu lado, são não-inerentes. O resultado dessa análise – que foi feita no *corpus* oral – é que com os substantivos não-inerentes (relações humanas outras que parentes; possuídos não-inerentes) o artigo é menos usado do que com os inerentes (possuídos inerentes, outros que partes do corpo; partes do corpo), isto é, inerente favoreceria o artigo.

Como se vê, outros estudos já observaram vários fatores que em maior ou menor grau favorecem ou inibem o artigo definido antes de pronome possessivo. Queremos lembrar que enquanto alguns desses fatores só afetam a escolha entre as duas variantes em escala relativamente modesta, outros determinam um uso categórico ou

quase categórico de uma das variantes. No PB não há nenhum caso que exige um uso categórico do artigo (exceto certas expressões fixas), mas existem certas construções com um emprego nulo ou muito reduzido do artigo. Por isso, antes de passarmos à nossa análise na seção 4, vamos excluir alguns tipos de construções do nosso *corpus*: aqueles que conforme as gramáticas e outras pesquisas apresentam um uso categórico ou grandemente majoritário da variante sem artigo; a inclusão desse tipo de exemplos enviesaria as análises quantitativas, visto querermos estudar o uso nos casos de variação livre. Na seção 3 faremos uma breve apresentação dos tipos que sempre ou em geral omitem o artigo, assim como de mais alguns casos que decidimos deixar de lado. Quando depois fizermos a análise do *corpus*, apresentaremos primeiro as frequências relativas das ocorrências com e sem artigo no total dos casos em cada um dos sete livros, e depois testaremos alguns dos fatores de Silva, para vermos se atuam da mesma maneira no nosso *corpus*: número, pausa antes do possessivo, preposição antes do possessivo, e o próprio pronome possessivo. A especificidade, no entanto, não será considerada, dado que nos parece difícil definir, em muitos casos, se o substantivo deve ser definido como específico ou não. Em outras palavras, especificidade não é um fator nítida e inequivocamente definível da mesma maneira que número, pessoa, etc., dado que depende da intuição do/a pesquisador/a, intuição essa que pode variar de uma pessoa para outra e portanto dificultar a comparação entre diferentes estudos. Aliás, a própria Silva (1982:289) diz que às vezes é difícil saber como classificar certas ocorrências, e que em certos casos se absteve de classificar os dados. Pelo mesmo motivo não faremos a divisão em quatro grupos de substantivos inerentes e não-inerentes de Silva (1996); só analisaremos os fatores ‘humano não-parente’ e ‘partes do corpo’, cuja definição não é problemática.

## 2 Casos deixados de lado

Para compormos a lista dos casos em que o artigo em geral não é usado, recorremos a algumas gramáticas. Não só elas mencionam um maior número de tipos do que os estudos empíricos já referidos mas, além disso, um levantamento do que afirmam as gramáticas também nos permite ver até que ponto a norma gramatical condiz com o uso no nosso *corpus*.

Tanto Cunha & Cintra (1991:217) como Neves (2000:425-427) dizem que o artigo é em geral evitado em fórmulas de tratamento (exemplo 3), quando faz parte de um vocativo (4), quando o possessivo vem precedido de um demonstrativo (5), e em certas expressões feitas (6). Além desses casos, Cuesta & Luz (1980:465-466) acrescentam nomes de parentesco (7) e a palavra *casa* (8). No artigo “Sobre o emprego do artigo com pronomes possessivos em português”, Meier (1973:7), referindo-se à gramática normativa, menciona, além de vocativo e expressões feitas, aposto (9), emprego indefinido do possessivo – ou seja, quando o possessivo tem o significado ‘algum’, ‘certo’ – (10) e predicativo (11).

(3) - Sr. tenente-coronal, disse-lhe ele, *V.S. [Vossa Senhoria]* já me livrou de uma que não era culpa minha; (Almeida,121)

(4) - Ah! *meu pai*, não me obrigue a semelhante sacrifício; por piedade! (Bernardo,108)

(5) Creio que José Dias achou desusado *este meu falar*. (Machado,61)

(6) Inocência, *por seu lado*, encostou a fronte ao ombro do amante, (Taunay,142)

(7) Só então *seu pai* reconheceu que o amor de *sua filha* não era uma simples veleidade de criança, (Bernardo,163)

(8) já os pés me pruriam para tomar o caminho de *sua casa*. (Alencar,94)

(9) Fora Leonardo algibebe em Lisboa, *sua pátria*; (Almeida,67)

(10) sobre o padrinho houve *suas dívidas* (Almeida,244)

(11) O cocheiro, que era *nossa escravo*, (Machado,217)

Antes de deixarmos esses tipos de construções de lado, veremos se no nosso *corpus* eles são, de fato, usados sem o artigo. Note-se que tratamento e vocativo quase só ocorrem nos diálogos, mas quanto aos demais tipos acima exemplificados, o texto a seguir só se refere às ocorrências encontradas na narrativa.

No *corpus*, tratamento, vocativo e possessivo precedido de demonstrativo se destacam dos demais tipos por uma ausência absoluta do artigo. Quanto às expressões, há cerca de 200 ocorrências, mas é um grupo heterogêneo, com muitas expressões diferentes, e apesar de essas expressões até certo ponto serem estereotipadas, Silva (1982:275) observa que no *corpus* dela há certa variação no uso do artigo, e também que algumas expressões “adquiriam sentidos algo diferentes, desde que tivessem ou não artigo”. Uma análise detalhada não nos parece viável dentro dos limites do presente trabalho, dado que grande parte dessas expressões ocorrem apenas uma ou duas vezes no *corpus*. No entanto, podemos notar que enquanto algumas sempre ocorrem sem artigo (p.ex. *por sua vez*, que com 23 casos ao todo é de longe a expressão mais frequente no *corpus*) e outras sempre são precedidas de artigo (p.ex. *ao seu alcance*, com quatro casos ao todo), há também algumas expressões que aparecem ora com artigo, ora sem, p.ex. *de minha/sua parte*: três

ocorrências sem e cinco com o artigo. De qualquer forma, a tendência geral no conjunto das expressões é para o não uso do artigo, que só ocorre em 29% dos casos.

TABELA 7 - Parentesco.

	<b>Total</b>	<b>com artigo</b>	<b>sem artigo</b>
Macedo	15	1 (7%)	14 (93%)
Almeida	52	3 (6%)	49 (94%)
Alencar	16	1 (6%)	15 (94%)
Bernardo	47	1 (2%)	46 (98%)
Taunay	5	0 (0%)	5 (100%)
Aluísio	37	27 (73%)	10 (27%)
Machado	167	6 (4%)	161 (96%)

Com nome de parentesco é evidente a preferência para o não uso do artigo nos livros do *corpus*: com exceção de Aluísio, em nenhum dos escritores o artigo é usado em mais de 7% dos casos. Como se vê, comparando tabela 7 com tabela 8 (seção 4), isso é muito abaixo da média geral. A diferença entre parentesco e os casos com variação livre é especialmente nítida em Machado, que nas ocorrências da tabela 8 usa o artigo quase sempre, enquanto só o emprega em 4% dos casos de nome de parentesco. Em Aluísio o artigo ocorre em 73% dos casos, e esta percentagem alta é devida ao fato de em *O cortiço* a palavra *homem* (no sentido de ‘marido’ e no sentido de ‘amante’) ser usada muito frequentemente, e quase sempre o artigo é usado com essa palavra. Contudo, mesmo se excluirmos *homem* dos cálculos, Aluísio usa o artigo em 27% dos casos de parentesco, ou seja, muito mais do que os demais escritores. Ao mesmo tempo, e independentemente de incluirmos *homem* ou não, também em Aluísio o uso do artigo é menos frequente com nome de parentesco do que com os substantivos ‘normais’ da tabela 8.

Nos demais tipos (exemplos 8-11) há poucas ocorrências, mas em todos os quatro tipos há variação, com tendência para o não uso do artigo.

Em suma, essas construções, que conforme as gramáticas inibem o uso do artigo, também no nosso *corpus* apresentam proporções bem reduzidas de artigo definido antes de possessivo. Em alguns casos o artigo é até categoricamente ausente.

Os tipos acima comentados não serão mais considerados; dado que mostram nítida tendência para uma das variantes, devem ser deixados de lado para não enviesar a análise quantitativa a seguir. Além disso, decidimos excluir também mais alguns contextos que talvez possam influir no uso do artigo: os casos com um adjetivo antes ou depois do substantivo, assim como aqueles com as palavras *todo/a(s)* e *próprio/a(s)* e os nomes próprios precedidos de possessivo:

(12) No seu *farto* cabelo (Aluísio,58)

(13) Vazei nelas *toda* a minha alma (Alencar,190)

(14) segundo suas *próprias* camaradas (Macedo,159)

(15) Agora que o nosso *Leonardo* está instalado em quartel seguro (Almeida,247)

Deixamos de lado também os casos de coordenação, quer dizer, as poucas ocorrências de um único possessivo seguido de dois substantivos coordenados, caso em que o gênero e o número do pronome possessivo concordam com o primeiro elemento coordenado, como em:

(16) Suas mãos e pés (Bernardo,15)

Também excluímos os SNs possessivos coordenados, sindética ou assindeticamente, como no exemplo (17), já que admitimos a possibilidade de o uso ou não uso do artigo com o primeiro possessivo poder influenciar o(s) seguinte(s). De fato, nosso material parece confirmar essa hipótese: de um total de 94 casos de coordenação deste

tipo, há apenas dois em que um dos elementos vem com artigo e ou outro sem artigo, como em (18).

(17) Assim pois, seu amor, suas esperanças, sua riqueza, sua felicidade, tudo isso fora uma ilusão (Bernardo,138)

(18) os seus olhos encandeados e sua garganta abrasada (Taunay,27)

Ficamos, assim, com um material bastante homogêneo para a análise a seguir, o que julgamos importante, já que parece haver muitos fatores capazes de influir no uso do artigo; excluindo os acima comentados será mais fácil detectar o uso dos escritores nos casos em que há uma variação pelo menos relativamente livre.

### 3 A análise

Os exemplos que nos restam, depois de termos deixado de lado os tipos comentados na seção 3, são quantificados na tabela 8:

TABELA 8 - Os pronomes possessivos.

	<b>Total</b>	<b>com artigo</b>	<b>sem artigo</b>
Macedo	194	66 (34%)	128 (66%)
Almeida	278	172 (62%)	106 (38%)
Alencar	409	280 (68%)	129 (32%)
Bernardo	250	53 (21%)	197 (79%)
Taunay	114	84 (74%)	30 (26%)
Aluísio	339	310 (91%)	29 (9%)
Machado	267	263 (99%)	4 (1%)

Há diferenças consideráveis entre os escritores: enquanto Macedo e Bernardo só empregam o artigo em 34% e 21% dos casos,

respectivamente, Aluísio e Machado o empregam em mais de 90% dos casos; em Machado o uso do artigo quase chega a ser categórico. Entre esses dois extremos encontram-se Almeida, Alencar e Taunay, com 62%, 68% e 74%, respectivamente.

Vejamos agora o efeito de alguns dos fatores tratados por Silva, começando por ‘número’ e ‘pausa’. Os casos classificados como ‘pausa’ são aqueles em que o possessivo se encontra diretamente depois de ponto, vírgula, ou outro tipo de pontuação (exemplo 19); ‘sem pausa’ são aqueles em que o pronome vem antecedido de outra palavra (exemplo 20).

(19) Mas o que então se passou em mim, lhe parecerá incrível; *a minha cólera* precisava desabafar-se contra alguém, (Alencar, 98)

(20) O médico deu por terminada *a sua visita*. (Macedo, 255)

Nos *corpora* de Silva o artigo é mais frequente no singular do que no plural, e no nosso *corpus* acontece o mesmo, com exceção de Machado. Convém notar, contudo, que tanto em Machado como em Taunay e Aluísio as diferenças são pequenas. Quanto a pausa antes do possessivo, conforme Silva este fator inibe o uso do artigo, mas uma diferença clara neste sentido só se observa em Bernardo, Taunay e Aluísio; em Almeida, Alencar e Machado as diferenças são pequenas, e em Macedo a frequência do artigo é mais alta depois de pausa.

TABELA 9 - Número: frequência relativa de casos com artigo.

	<b>singular</b>	<b>plural</b>
Macedo	48% (58/122)	11% (8/72)
Almeida	69% (138/199)	43% (34/79)
Alencar	71% (216/306)	62% (64/103)
Bernardo	25% (46/187)	11% (7/63)
Taunay	75% (58/77)	70% (26/37)
Aluísio	92% (242/263)	89% (68/76)
Machado	98% (177/181)	100% (86/86)

TABELA 10 - Pausa: frequência relativa de casos com artigo.

	pausa	sem pausa
Macedo	50% (6/12)	33% (60/182)
Almeida	60% (6/10)	62% (166/268)
Alencar	72% (26/36)	68% (254/373)
Bernardo	9% (2/23)	22% (51/227)
Taunay	38% (3/8)	76% (81/106)
Aluísio	75% (18/24)	93% (292/315)
Machado	94% (17/18)	99% (246/249)

Ainda em Silva (1982) o artigo é mais comum quando o pronome possessivo é da primeira pessoa do que da terceira. Dado que só os livros de Alencar e Machado são escritos na primeira pessoa é só nesses dois que podemos estudar a primeira pessoa do singular, mas em todos os sete livros há ocorrências da primeira pessoa do plural (quando os escritores falam sobre ‘o nosso herói’, ‘nossos agricultores’, etc.). Analisando este fator juntamos as ocorrências de masculino e feminino, singular e plural, ou seja, falando por exemplo da primeira pessoa do singular nos referimos a todas as ocorrências de *meu*, *minha*, *meus* e *minhas*, colocando só a forma do masculino singular na coluna correspondente da tabela 11. Na terceira pessoa nem sequer fazemos diferença entre a terceira pessoa do singular (*seu=dele/dela*) e a terceira pessoa do plural (*seu=deles/delas*), já que as formas *seu*, *sua*, *seus*, *suas* são as mesmas. Aliás, os casos de *seu(s)* e *sua(s)* que equivalem a *deles/delas* são poucos: perfazem apenas 6% do total (85 em 1353).

TABELA 11 - O pronome possessivo:  
frequência relativa de casos com artigo.

	<i>meu</i>	<i>nossa</i>	<i>seu</i>
Macedo	-	94% (15/16)	29% (51/178)
Almeida	-	96% (23/24)	59% (149/254)
Alencar	75% (151/201)	64% (14/22)	62% (115/186)
Bernardo	-	33% (1/3)	21% (52/246)
Taunay	-	100% (6/6)	72% (78/108)
Aluísio	-	0% (0/1)	92% (310/338)
Machado	99% (169/171)	98% (50/51)	98% (42/43)

Machado emprega o artigo quase categoricamente, tanto nas primeiras pessoas do singular e do plural como na terceira pessoa, mas em Alencar a frequência do artigo é de fato mais alta com a primeira pessoa do singular do que com a terceira pessoa: o artigo é usado em 75% e 62% dos casos, respectivamente. A primeira pessoa do plural, porém, tem praticamente a mesma frequência da terceira: 64%. Bernardo, Taunay e Aluísio têm tão poucas ocorrências da primeira pessoa do plural que não vale a pena tentar uma comparação com a terceira, mas em Macedo e Almeida, com um número maior de ocorrências, a primeira pessoa do plural tem mais frequentemente o artigo do que a terceira pessoa; em Macedo a diferença é até muito grande: 94% com *nossa/a(s)* e apenas 29% com *seu(s)/sua(s)*.

Os fatores até aqui examinados produzem quase os mesmos resultados da tese de Silva: o artigo é mais comum no singular do que no plural, e com os pronomes da primeira pessoa emprega-se o artigo com maior frequência do que com os pronomes da terceira pessoa. O fator pausa, contudo, só inibe o artigo em alguns dos escritores do nosso *corpus*. Convém notar que Machado usa o artigo quase categoricamente, e por isso esses fatores não são muito relevantes neste escritor.

Vejamos agora o efeito de preposição antes de possessivo. Em Silva (1982) a presença de uma preposição favorece o artigo nos *corpora* sincrônicos, mas não nos *corpora* diacrônicos do PB. O fato de a preposição favorecer o artigo sobretudo nos *corpora* orais levou Silva a crer que o que realmente afeta são as preposições que se contraem com o artigo. Por isso, dividimos nosso material nos seguintes três grupos: preposição que contrai (exemplo 21), preposição que não contrai (22), e preposição nenhuma antes do possessivo (23).

(21) Cirino fez repentina parada *nas suas explicações*. (Taunay,149)

(22) Saíram e encaminharam-se *para o seu destino*. (Almeida,175)

(23) E expôs *o seu projeto*: (Aluísio,169)

TABELA 12 - Preposição: frequência relativa de casos com artigo.

	<b>contrai</b>	<b>não contrai</b>	<b>sem prep.</b>
Macedo	33% (28/86)	6% (1/17)	41% (37/91)
Almeida	58% (89/153)	42% (11/26)	73% (72/99)
Alencar	56% (124/221)	84% (26/31)	83% (130/157)
Bernardo	10% (13/124)	31% (4/13)	32% (36/113)
Taunay	71% (36/51)	75% (3/4)	76% (45/59)
Aluísio	93% (155/166)	94% (31/33)	89% (124/140)
Machado	98% (128/130)	100% (11/11)	98% (124/126)

Comparando preposição que contrai com preposição que não contrai, observa-se que preposição que contrai favorece o artigo em Macedo e Almeida, mas inibe em Alencar e Bernardo. Em comparação com os casos sem preposição, preposição que contrai não favorece o artigo em nenhum desses quatro escritores. Em Taunay, Aluísio e Machado as diferenças são mínimas. Este fator, portanto, não tem o mesmo efeito no nosso *corpus* como nos *corpora* sincrônicos de Silva (1982).

Outros fatores que também não atuam da mesma maneira no *corpus* de Silva e no nosso são ‘humano não-parente’ e ‘partes do corpo’. Em Silva (1996:135) o artigo ocorre em 40% dos casos de relações humanas outras que parentes e em 58% dos casos de partes do corpo, ou seja, com palavras que designam partes do corpo o artigo é mais frequente do que com palavras que se referem a seres humanos. No entanto, não é isso que se verifica no nosso *corpus*.

TABELA 13 - Humano, corpo: frequência relativa de casos com artigo.

	<b>humano</b>	<b>corpo</b>
Macedo	67% (26/39)	9% (4/44)
Almeida	75% (38/51)	38% (6/16)
Alencar	60% (9/15)	63% (58/92)
Bernardo	26% (8/31)	0% (0/33)
Taunay	93% (14/15)	63% (17/27)
Aluísio	95% (18/19)	71% (22/31)
Machado	100% (39/39)	100% (21/21)

Deixando de lado Machado, que não tem variação, vemos que em Alencar quase não há diferença, mas nos demais cinco escritores o artigo é bem mais frequente com humano do que com corpo, ou seja, no estudo de Silva (1996), corpo favorece o artigo em relação a humano, mas no nosso *corpus* dá-se o contrário. Confirma-se, portanto, o que vimos acima quando da análise dos fatores pausa e preposição antes do possessivo: nem todos os fatores atuam da mesma maneira em todos os *corpora*.

Por fim, convém dizer algumas palavras em separado sobre Alencar. Como já foi observado por Silva (1982), Alencar comenta o uso do artigo definido antes de possessivo no pós-escrito à segunda edição de *Iracema*, defendendo “certa parcimônia no emprêgo do artigo definido”, explicando isso como “uma reação contra o abuso dos escritores

português” (Alencar 1953a:194). No entanto, depois de uma análise de um outro romance de Alencar, *Diva*, Silva (1982:356-358) afirma que este autor não parece ser coerente com suas teorias:

Em resumo, Alencar defende a ausência do artigo (portanto consciente), como libertação da língua de Portugal. Bem alto ( $325/687 = 47,3\%$ ), o *overall* de *Diva* não se coaduna com o sentimento de Alencar sobre o assunto já que ultrapassa os dados não só do século XIX mas até do século XX. Esse fato confirma quão inconsciente é o uso dessa regra e quão longínquo pode estar o uso da atitude.

Comparando o resultado da sua análise de *Diva* com os seus *corpora* dos séculos XIX e XX (cf. tabelas 4 e 5), Silva considera a frequência de Alencar alta, mas se compararmos Alencar com os demais escritores do nosso *corpus* (tabela 8), vemos que apesar de a percentagem de *Diva* (47%) ser mais alta do que as percentagens de Macedo e Bernardo, por outro lado, em relação a Almeida, Taunay, Aluísio e Machado não é particularmente alta. Nem sequer os 68% de *Lucíola* (o livro de Alencar na tabela 8) nos parece uma frequência muito alta em comparação com os demais escritores brasileiros. Na nossa opinião, a recomendação de Alencar de ‘certa parcimônia’ deve ser interpretada em relação ao uso do PE que, como acima foi dito, já tinha praticamente generalizado o artigo definido (cf. tabela 2), e tanto os 47% em *Diva* como os 68% em *Lucíola* podem ser, de fato, uma reação contra o ‘abuso’ do PE. Em Schei (2007) analisamos anúncios e cartas de leitores em jornais do século XIX (o *corpus* do Projeto para a História do Português Brasileiro), e neste material chegamos ao seguinte resultado:<sup>8</sup>

<sup>8</sup> O resultado de Schei (2007) ao qual aqui fazemos referência é aquele apresentado nas tabelas 8 e 9 daquele artigo, ou, mais exatamente, na tabela 9 e no que deveria ter sido a tabela 8. Como já foi dito na introdução, no nosso artigo de 2007 há vários erros que foram introduzidos pelo revisor. Um dos erros mais graves foi a substituição de duas tabelas nossas (tabelas 7 e 8) por tabelas de um outro artigo de outra autora no mesmo livro. Por conseguinte, dos dados que aqui apresentamos na tabela 14, tirados do nosso artigo de 2007, só metade de fato consta daquele artigo: as frequências relativas dos casos com artigo nos anúncios dos jornais do século XIX

TABELA 14 - Jornais do século XIX, 1<sup>a</sup> e 2<sup>a</sup> metade do século:  
frequência relativa de casos com artigo.

<b>corpus</b>	<b>1<sup>a</sup> metade séc. XIX</b>	<b>2<sup>a</sup> metade séc. XIX</b>
anúncios	37% (49/131)	47% (238/504)
cartas de leitores	54% (214/397)	61% (407/666)

Em comparação com este material, as frequências de Alencar também não são particularmente altas.

#### 4 Os diálogos

TABELA 15 - Os diálogos: os pronomes possessivos.

	<b>Total</b>	<b>com artigo</b>	<b>sem artigo</b>
Macedo	196	115 (59%)	81 (41%)
Almeida	35	28 (80%)	7 (20%)
Alencar	188	132 (70%)	56 (30%)
Bernardo	157	86 (55%)	71 (45%)
Taunay	132	111 (84%)	21 (16%)
Aluísio	23	21 (91%)	2 (9%)
Machado	57	56 (98%)	1 (2%)

Se compararmos os diálogos com a narrativa, vemos, observando as tabelas 15 e 8, que em Alencar, Aluísio e Machado as percentagens são praticamente as mesmas, mas em Macedo, Almeida, Bernardo e Taunay o artigo ocorre mais frequentemente nos diálogos. Supondo que os diálogos de alguma maneira refletem a língua falada, pelo menos em maior grau do que a narrativa, e dado que o PE (que por muitas pessoas naquela época era visto como a norma da língua) usava o artigo quase sempre enquanto no PB o artigo era mais raro, o maior uso do artigo nos diálogos é um pouco surpreendente.

Nos diálogos há um número menor de ocorrências, pelo que é mais difícil examinar os diferentes fatores testados na análise da narrativa.

---

são tiradas da tabela 9 em Schei (2007:279). As frequências nas cartas de leitores, porém, não constam mais do artigo de 2007, por causa da substituição da tabela 8 daquele artigo.

Por isso, nos limitaremos a analisar o pronome possessivo, já que os diálogos nos oferecem vários tipos; ao contrário da narrativa, em que só se encontram exemplos de *meu*, *nosso* e *seu* da terceira pessoa, nos diálogos há também pronomes da segunda pessoa: *teu*, *vosso* e *seu* da segunda pessoa.

TABELA 16 - Os diálogos: o pronome possessivo:  
frequência relativa de casos com artigo.

	<b><i>meu</i></b>	<b><i>nosso</i></b>	<b><i>teu</i></b>	<b><i>vosso</i></b>	<b><i>seu 2<sup>a</sup> p.</i></b>	<b><i>seu 3<sup>a</sup> p.</i></b>
Macedo	71% (55/77)	58% (7/12)	84% (16/19)	46% (11/24)	62% (16/26)	26% (10/38)
Almeida	80% (16/20)	-	-	-	89% (8/9)	67% (4/6)
Alencar	79% (63/80)	57% (4/7)	63% (27/43)	-	67% (18/27)	65% (20/31)
Bernardo	51% (30/59)	67% (12/18)	38% (5/13)	100% (2/2)	57% (25/44)	57% (12/21)
Taunay	85% (53/62)	100% (11/11)	67% (2/3)	100% (1/1)	80% (33/41)	79% (11/14)
Aluísio	88% (7/8)	100% (2/2)	100% (2/2)	-	83% (5/6)	100% (5/5)
Machado	95% (20/21)	100% (16/16)	100% (3/3)	-	100% (10/10)	100% (7/7)

Como vimos na seção 2, no *corpus* de Silva (1996) a frequência relativa de casos com artigo era muito mais baixa com *seu* da terceira pessoa do que com os demais pronomes. O mesmo se dá em Macedo, mas não nos demais. Em Aluísio e Machado, *seu* da terceira pessoa tem artigo em 100% dos casos enquanto alguns pronomes da primeira e segunda pessoas têm percentagens mais baixas, e em Alencar e Taunay *seu* da terceira pessoa não difere muito dos demais pronomes. Em Almeida a percentagem de *seu* da terceira pessoa é de fato mais baixa da de *meu* e *seu* da segunda pessoa, mas as diferenças não são muito grandes, e em Bernardo é um pronome da segunda pessoa, *teu*, que tem a percentagem mais baixa, 38%, enquanto tanto *seu* da segunda como *seu* da terceira

pessoa têm 57%. Em suma, a pessoa gramatical do pronome possessivo não é um fator que atua de maneira uniforme nos sete escritores.

Resumindo esta breve análise dos diálogos, a frequência relativa dos casos com artigo em alguns dos sete escritores é mais alta do que na narrativa, e o estudo dos diferentes pronomes não deu o resultado esperado conforme o estudo de Silva (1996); *seu* da terceira pessoa nem sempre inibe o artigo se comparado com os demais pronomes.

## 5 Palavras finais

Neste trabalho apresentamos os resultados do exame de um *corpus* de romances brasileiros oitocentistas, assim como um resumo de alguns outros estudos. O que é, então, que todas essas análises nos dizem sobre o uso do artigo antes de possessivo seguido de substantivo? Na verdade, pouca coisa. A única coisa que podemos dizer com certeza é que as frequências relativas das duas variantes variam de um *corpus* para outro. Quanto aos fatores analisados, os efeitos de muitos deles são apenas tendências e alguns dos fatores produzem efeitos diferentes em diferentes *corpora*, o que levanta a questão de até que ponto eles realmente são relevantes. Além disso, com tantos fatores coexistindo ao mesmo tempo é muito difícil dizer como eles interagem, ou seja, qual/ is dos fatores realmente determina/m a escolha entre artigo ou não artigo num determinado caso. No entanto, há algumas observações que podem ser feitas no que diz respeito ao nosso *corpus*.

Nossa análise, como foi resumida na tabela 8, assim como outros estudos empíricos referidos na seção 2, mostraram que a frequência do uso do artigo antes de possessivo varia consideravelmente no PB. Apesar disso, tiramos a conclusão que o uso praticamente categórico do artigo em Machado não constitui uma variante natural brasileira, mas é devido ao fato de Machado em *Dom Casmurro* ter seguido a norma europeia,<sup>9</sup> que já no século XIX tinha generalizado o uso do artigo.

<sup>9</sup> Em outros livros anteriores, Machado usa menos o artigo: em *Memórias póstumas de Brás Cubas* o artigo ocorre em 84% dos casos, em *A mão e a luva* em 60%, e em *Helena* em apenas 34%.

É possível que também a alta percentagem em Aluísio seja devida à mesma causa. Quanto aos demais cinco escritores do *corpus*, Macedo e Bernardo se distinguem por um uso bastante reduzido do artigo, enquanto Almeida, Alencar e Taunay usam o artigo em 62-74% dos casos. Talvez essa diferença seja devida ao fato de Macedo e Bernardo usarem um modelo mais antigo; dado que o PB parece ter evoluído de um estágio em que o artigo era usado em escala bastante modesta para um uso mais frequente, não é impossível que Macedo e Bernardo ainda estejam usando um modelo mais antigo, com poucos casos de artigo definindo antes de possessivo, ao passo que Almeida, Alencar e Taunay já estejam usando um modelo mais moderno – mas sempre brasileiro e não lusitano – com maior uso do artigo. Seja como for, confirmou-se mais uma vez que o uso do artigo definido antes de possessivo no PB varia bastante de um livro para outro.

## Referências

### 1. Obras citadas

ALENCAR, José de. **Iracema e Ubirajara**. In: **Obra completa**, vol. **VIII**. 2. ed. Rio de Janeiro: José Olympio, 1953a.

CUESTA, Pilar Vázquez & LUZ, Maria Albertina Mendes da. **Gramática da língua portuguesa**. Lisboa: Edições 70, 1980

CUNHA, Celso Ferreira da & CINTRA, Luís F. Lindley. **Nova gramática do português contemporâneo**. 8. ed. Lisboa: Edições João Sá da Costa, 1991.

LUCCHESI, Dante “**The article systems of Cape Verde and São Tomé creole Portuguese: general principles and specific factors**”. *Journal of Pidgin and Creole Languages*, 1993. 8:1, p. 81-108.

MEIER, Harri. “**Sobre o emprego do artigo com pronomes possessivos em português**”. *Littera* 3, 1973. p. 5-14.

NEVES, Maria Helena de Moura. “**Possessivos**”. In: CASTILHO, Ataliba Teixeira de (org.), **Gramática do português falado, vol. III**. Campinas: Editora da UNICAMP, 1993. p. 149-211.

\_\_\_\_\_. **Gramática de usos do português**. São Paulo: Editora UNESP, 2000.

SAID ALI, Manuel. **Gramática histórica da língua portuguesa**. In: **Gramática secundária e Gramática histórica da língua portuguesa**, 3. ed. Brasília: Editora Universidade de Brasília, 1964.

SCHEI, Ane. “**Para o estudo do artigo definido antes de pronome possessivo no português brasileiro: algumas observações**”. In: RAMOS, Jânia M.; ALKMIM, Mônica A. (orgs.), **Para a história do português brasileiro, vol. V: Estudos sobre mudança linguística e história social**. Belo Horizonte: Editora FALE/UFMG, 2007. p. 265-281.

SILVA, Giselle Machline de Oliveira e. **Estudo da regularidade na variação dos possessivos no português do Rio de Janeiro**. Rio de Janeiro: UFRJ, 1982. Universidade Federal do Rio de Janeiro, 1982. Tese (Doutorado)

\_\_\_\_\_. “**Artigo frente a possessivos e nomes próprios**”. Projeto subsídios sociolinguísticos do projeto censo à educação, vol II. Universidade do Rio de Janeiro, Faculdade de Letras, 1986. p. 232-255.

\_\_\_\_\_. “**Realização facultativa do artigo definido diante de possessivo e de patronímico**” In: SILVA, Giselle Machline de Oliveira; SCHERRE, Maria Marta Pereira (orgs.) **Padrões sociolinguísticos: análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro**. Rio de Janeiro: Tempo Brasileiro: Departamento de Línguística e Filologia, UFRJ, 1996. p.119-145.

## 2 Obras examinadas

ALENCAR, José de. **Luciola**. reprodução da 3. edição, revista pelo autor. In: **Obra completa, vol IV**, 2. ed. Rio de Janeiro: José Olympio, 1953b. (1862)

ALMEIDA, Manuel Antônio de. **Memórias de um sargento de milícias**. fixação de texto de Mamede Mustafa Jarouche. Cotia: Ateliê Editorial, 2000. (1854-1855)

ANTUNES, António Lobo. **Os cus de Judas**, 9. ed. Lisboa: Publicações Dom Quixote, 1983. (1979)

ASSIS, Machado de. **Dom Casmurro**, edição do texto de Ivan Teixeira. São Paulo: Martins Fontes, 1988. (1899)

\_\_\_\_\_. **Helena**, 18. ed. São Paulo: Editora Ática, 1994. (1876)

\_\_\_\_\_. **A mão e a luva**. 12. ed. São Paulo: Editora Ática, 1995. (1874)

\_\_\_\_\_. **Memórias póstumas de Brás Cubas**. Cotia: Ateliê Editorial, 1998. (1881)

AZEVEDO, Aluísio. **O cortiço**. 36. ed. São Paulo: Editora Ática, 2000. (1890)

BRANCO, Camilo Castelo. **Amor de perdição**. de acordo com a 5<sup>a</sup> edição de 1879, revista pelo autor. Lisboa: Editorial Comunicação, 1983. (1862)

DOURADO, Autran. **Confissões de Narciso**. Rio de Janeiro: Francisco Alves, 1997. (1997)

FONSECA, Rubem. **Vastas emoções e pensamentos imperfeitos**. 11. ed. São Paulo: Companhia das Letras, 1992. (1988)

GUIMARÃES, Bernardo. **O garimpeiro.** Rio de Janeiro: B.L. Garnier, 1872. (1872)

HERCULANO, Alexandre. **Eurico, o presbítero.** Edição crítica. Lisboa: Livraria Bertrand, 1944. (1844)

LUFT, Lya. **Exílio.** (1988)3. ed. São Paulo: Siciliano, 1991.

MACEDO, Joaquim Manuel de. **A Moreninha.** Edição crítica de Tânia Serra. Rio de Janeiro: Lacerda Ed, 1997. (1844)

MONTELLO, Josué. **Enquanto o tempo não passa.** Rio de Janeiro: Nova Fronteira, 1996. (1996)

QUEIROZ, Eça de. **O crime do Padre Amaro**, de acordo com a edição de 1880, revista pelo autor. Lisboa: Livros do Brasil, s.d. (1875)

QUEIROZ, Rachel de. **Dôra, Doralina.** In: **As três Marias e Dôra, Doralina; Obra reunida, v. II.** Rio de Janeiro: José Olympio, 1989. (1975)

SCLIAR, Moacyr. **Os voluntários**, 4. ed. Porto Alegre: L&PM, 1991. (1979)

SOUSA, Américo Guerreiro de. **Os cornos de Cronos.** 3. ed. Venda Nova: Bertrand Editora, 1989. (1980)

TAUNAY, Alfredo d'Escragnolle. **Inocência.** 5. ed. São Paulo: FTD, 1999. (1872)

# A VARIAÇÃO NÓS / A GENTE NO DIALETO MINEIRO: INVESTIGANDO A TRANSIÇÃO

Francisca Paula Soares MAIA

Universidade Federal de Minas Gerais (UFMG)

## RESUMO

*Neste artigo apresento verificações feitas sobre a questão: “nós / a gente” no dialeto mineiro, mudança ou variação? Utilizando dados de uma localidade urbana (Belo Horizonte), e uma rural (Pombal), estudo fatores linguísticos e extralingüísticos, com base na Teoria da Variação de Weinreich, Labov e Herzog. A análise desses fatores mostra a força da morfologia na variedade analisada. Por sua vez, a análise em tempo aparente constatou uma ‘mudança em progresso’, com ritmo lento em Minas Gerais, quando comparada à situação do Rio de Janeiro e de outras localidades.*

## ABSTRACT

*This paper is a report on the research I conducted on the alternation nós / a gente in Minas Gerais (Brazil), in order to establish whether the increasing usage of a gente is a case of sociolinguistic variation or a case of diachronic change. The framework of this research was the Variation Theory by Weinreich, Labov and Herzog. By comparing data from an urban locality (Belo Horizonte), and a rural one (Pombal), I evaluated the importance of several linguistic and extra-linguistic factors. My analysis pointed to morphology as the most important factor affecting this choice. My apparent time analysis detected a ‘change in progress’, slower in Minas Gerais, than, e.g. in Rio de Janeiro and others localities.*

## PALAVRAS-CHAVE

*Formas pronominais. Mudança linguística. Transição. Variação linguística.*

## KEY WORDS

*Brazilian Portuguese pronouns. Linguistic change. Linguistic variation. Transition.*

## Introdução

Meu objeto de estudo é o uso das formas pronominais **nós** e **a gente** na língua portuguesa falada no Brasil, mais especificamente, no dialeto mineiro. Por “dialeto mineiro” entenda-se aqui uma forma de falar da região central de Minas Gerais, conforme aparece em Nascentes (1953:17) e em Zágari (1998: 1).

No Português Brasileiro Padrão, a presença do pronome **nós** exige a desinência número-pessoal *-mos*; e a presença do pronome **a gente** exige a terminação de *3<sup>a</sup>. pessoa do singular*, conforme aparece abaixo:

- (1) a. **Nós éramos** cinco e **brigávamos** muito, recordou Augusto, olhos perdidos num ponto X, quase sorrindo. (Carlos Drummond de Andrade, CA: 5; *apud* CUNHA e CINTRA, 1985: 131)  
b. Não culpes mais o Barbaças, compadre! **A gente** só queria gastar um bocadito do dinheiro. (Fernando Namora, TJ: 165; *apud* CUNHA e CINTRA, 1985: 288)

Observa-se em (1) que as formas pronominais **nós** e **a gente** alternam-se como formas de expressão da *pessoa do falante + alguém* no discurso, conforme reconhecem os gramáticos tradicionais: “No colóquio normal, emprega-se **a gente** por **nós** (...).” (CUNHA e CINTRA, 1985: 288). Além dessa acepção, podem também expressar a voz do próprio falante:

- (2) a. Hoje em dia **nós tamo** assim... (P.S.,27,f1, BH)<sup>1</sup>  
b. **A gente** perdeu um pôco o contato (P.S. 27, f1, BH)

ou ainda ter referência indefinida (3 a-b):

---

<sup>1</sup> Visando preservar a identidade dos informantes e o acesso imediato aos dados, usei siglas aleatórias no lugar dos nomes próprios.

- (3) a. Quando **a gente** é menino... (G.J., 48, f2, BH)  
 b. **Nóis** planta é mio, feijão, arroiz é horta mesmo. (C.,31,f1, Pb)

No português não-padrão, cada uma dessas formas realiza concordância ora com o verbo na 1<sup>a</sup>. pessoa do plural, ora com o verbo na 3<sup>a</sup>. pessoa do singular. Comparem-se (1a-b) e (4a-b):

- (4) a. Verdura **nóis** come sim senhora... (D.L., 64, f2, Pb)  
 b. **A gente** duramo quase dois meses. (I.A.S., 22, f1, BH)

Além disso, a desinência verbal de primeira pessoa do plural pode ainda apresentar-se foneticamente reduzida (5 a-b):

- (5) a. **Nós** começamo só com três. (P.S.,27,f1, BH)  
 b. **Nóis** vão lanchá num lugá lá. (R.A., 20, f1, Pb)

As próprias variantes podem, por sua vez, ter as seguintes realizações fonéticas: **nós** pode ser ditongado e **a gente** pode sofrer queda de segmento inicial, medial ou final; ou ainda ser pronunciado com consoante aspirada. Comparem-se (6a-b) e (7a-e):

- (6) a. **Nós** fomos, descemos e varremo as sala... (G.J., 48, f2, BH)  
 b. **Nóis** tinha marcado o casamento... (R.C.,35, f1,Pb)
- (7) a. **A gente** fica amolado. (d.E., 76,f3, Pb)  
 b. **A /ente** plantava lá a meia, plantava a terça. (J.S.,72,f3,Pb)  
 c. **A g/te** vai na casa das pessoas... (E.M., 21,f1, BH)  
 d. **A gen/** namorava era assim, cunversando, era a corte... (Cl, 74, f3, BH)  
 e. **A bente** só roçasse lá, quemasse, plantava a terça... (J.S., 72, f3, Pb)

## 1 A variação *nós / a gente*

O uso de ***nós*** e ***a gente*** como variantes sociolinguísticas já foi estudado em alguns dialetos do Português: Recife, Salvador, Rio de Janeiro, São Paulo e Porto Alegre (Albán et alii, 1991; Abraçad, 1991; Menón (1994), Lopes (1999), Zilles (1999), não existindo ainda um trabalho dedicado exclusivamente ao comportamento dessas variantes pronominais no dialeto mineiro. Este trabalho veio preencher essa lacuna, possibilitando o estabelecimento de comparações entre diferentes dialetos brasileiros.

Estudos realizados sobre os pronomes ***nós*** e ***a gente*** dentro da abordagem variacionista para outras regiões do país mostram não se tratar somente de uma variação, mas de uma tendência à mudança (Machado, 1995; Lopes, 1999).

Meu propósito foi investigar duas comunidades linguísticas mineiras, uma rural (Pombal/Mariana) e uma urbana (Belo Horizonte), contrapô-las, e verificar se um processo de mudança está presente e qual a sua força. Essa contraposição decorre, por sua vez, do fato de as comunidades rurais e urbanas apresentarem ritmos de mudança distintos, o que permitiu a ampliação do espectro do fenômeno analisado, contemplando um período de tempo mais amplo. Meu foco de atenção foi a correlação entre ***nós*** e ***a gente*** e a morfologia verbal.

A bibliografia estudada, sem exceção, traça o perfil da variante ***a gente*** sem chamar a atenção para os contextos de resistência. Os trabalhos podem ser agrupados de acordo com duas hipóteses referentes ao gatilho da mudança:

Hipótese 1: a alteração no paradigma pronominal leva ao empobrecimento da morfologia verbal, defendida nos trabalhos de Menón (1995, 1996), Lopes (1999), Zilles (2002), Omena (2003).

Hipótese 2: alterações na morfologia verbal levam à mudança no paradigma pronominal. Defendida por Abraçado (1991). Esta é a hipótese que defendo neste trabalho sobre a variação das formas ***nós / a gente*** no dialeto mineiro. Numa linha diferente da dos trabalhos

anteriores, e semelhante à de Abraçado, testei a seguinte hipótese: o enfraquecimento da morfologia verbal foi o gatilho da mudança.

A análise quantitativa foi realizada com o propósito de responder às seguintes questões:

- (A) Que fatores linguísticos e sociais estariam condicionando o uso das variantes **nós / a gente** no Português falado em Minas Gerais?
- (B) Os dados configuram *mudança* ou *variação estável*?
- (C) A implementação da forma **a gente** é anterior ao processo de “simplificação do paradigma verbal” ou é seu gatilho?
- (D) Qual foi a *transição* dessa mudança no sistema linguístico, se é que houve mudança?

## 2 O modelo variacionista

Minha análise das ocorrências de **nós** e **a gente** no dialeto mineiro desenvolveu-se dentro da Teoria da Variação, que concebe a língua como um sistema heterogêneo do qual a variação é parte inerente (Weinreich, Labov & Herzog, 1968; Labov, 1972a, 1972b).

Uma variação pode manter-se estável ou evoluir para uma *mudança em progresso*. No segundo caso, a variação tem maiores índices de realização nos grupos sociais centrais,<sup>2</sup> resultando em representações gráficas de padrão curvilíneo. Além disso, os falantes mais jovens tendem a realizar a variante inovadora mais frequentemente que os falantes mais velhos (evidência de *tempo aparente*). A essas duas evidências, Faixa Etária e Padrão Curvilíneo, associa-se a observação das transformações ocorridas ao longo do tempo, denominada evidência de *tempo real*. Se há *variação estável*, uma variante ocorre mais nas classes mais altas, e a outra variante ocorre mais nas classes mais baixas. A representação gráfica

---

<sup>2</sup> Oliveira (1982) observa que não se deve padronizar a atuação dos fatores não-estruturais sobre as mudanças linguísticas, assim, nem sempre o padrão não-curvilíneo caracteriza variação estável, este varia de acordo com a atuação dos fatores sociais dentro da sociedade em estudo.

desse processo resulta em uma linha com vários picos, não havendo hierarquia entre as faixas etárias.

Na presente pesquisa, foi efetivada uma análise quantitativa com base no *tempo aparente*, evidenciado pelo comportamento linguístico de grupos de indivíduos de diferentes faixas etárias. Para a quantificação dos dados foi usado o programa GOLDVARB 2001 (Lawrence & Tagliamonte 2001), tendo como auxílio teórico Scherre, 1992;Sankoff, 1988).

### 3 Objeto de análise

Conforme já foi dito na secção anterior, foram utilizados dois *corpora* de entrevistas sociolinguísticas individuais. Um, com falantes de uma comunidade rural mineira: a comunidade de Pombal,<sup>3</sup> é representativo do dialeto mineiro rural<sup>4</sup> contemporâneo. O outro *corpus* reúne entrevistas sociolinguísticas individuais com 12 informantes de Belo Horizonte. São entrevistas representativas da fala urbana mineira e pertencem ao banco de dados da “fala belo-horizontina”<sup>5</sup>; três das doze entrevistas foram realizadas por mim.

---

<sup>3</sup> Cedidas pela Profa.Dra. Mônica G.R.Alkmin. Aparecem temas como futebol, perigo de vida, religião, etc.

<sup>4</sup> Essa amostra constitui-se de 12 informantes, selecionados de acordo com a faixa etária. A comunidade de Pombal situa-se a 35 km da sede do distrito de Mariana, mais precisamente na Serra dos Pretos, com um total de 500 habitantes, nas imediações da Fazenda da Vargem e próxima ao Córrego Pombal. Seus moradores têm baixíssimo poder aquisitivo e poucos contatos com a área urbana. Quando da coleta dos dados, nem televisão esse povoado possuía. Apresenta um baixo índice de alfabetização (apud Alkmim, 2001).

<sup>5</sup> A capital tinha, conforme dados do IBGE de 1997, área total de 330,93 km<sup>2</sup>. Sua população era então de 2.238.526 habitantes. (Fonte: IBGE, Base de Informações Municipais – Malha Municipal Digital 1997). Segundo Corrêa (1998,p.43) “Belo Horizonte constitui núcleo metropolitano, formado por, aproximadamente 250 bairros, agrupados em 9 regiões, e mais 3 cidades satélites em seu contorno (...). Liga-se a todo o país por aeroportos, rodovias e ferrovias. As principais atividades econômicas são: indústrias de minerais não metálicos, metalurgia, material de transportes, químicos, perfumaria, têxtil, vestuário, bebidas, grande comércio varejista e atacadista.”

Ao utilizar dois *corpora* tive por objetivo formar células que permitissem a análise do fator localização geográfica.

A variável dependente contém duas variantes, a saber: <**nós**> [**nós**] / [**a gente**] <sup>6</sup>. Cada uma das variantes aparece exemplificada abaixo:

- (8) *Nós* vê pra cá... (J.S., 72, f3, Pb)  
 (9) *A gente* num *sabe* quem é o certo. (P.S., 27,f1, BH)

Cada uma das variantes acima possui realizações foneticamente distintas. Conforme referido no Capítulo I, a variante *a gente* pode apresentar queda de segmentos, resultando [ɐ'ʒẽ], [ɐ'ʒetʃ] e [ɐ'ʒtʃ] ou alteração fonética [ɐ'χetʃ]. Já a variante **nós** pode realizar-se como ['nois]. A cliticização e a ditongação das variantes, entretanto, não serão tratadas aqui.

Foram consideradas cinco variáveis linguísticas e duas variáveis extralingüísticas, que aparecem enumeradas abaixo.

As variáveis linguísticas testadas foram:

- I. Pessoa Verbal;
- II. Tempo Verbal;
- III. Referência [ $\pm$ Genérica];
- IV. Realização Fonológica da Desinência de Número e Pessoa;
- V. Saliência Fônica.

As variáveis extralingüísticas foram:

- VI. Faixa Etária: de 20 a 35 (F1); de 36 a 65 (F2); > 65 (F3);
- VII. Localização Geográfica: zona rural (Pombal) e zona urbana (Belo Horizonte).

---

<sup>6</sup> Os colchetes foram aqui utilizados para distinguir as variantes ‘nós’/‘a gente’ da variável ‘nós’, colocada entre parênteses angulares.

<sup>7</sup> Utilizo aqui o Alfabeto Fonético Internacional (revisado em 2005). In: <http://weston.ruter.net/projects/ipa-chart/view/keyboard/>

As seguintes hipóteses orientaram a presente pesquisa:

- (A) A forma **a gente** é uma inovação;
- (B) O percurso da mudança que resultou na inserção de *a gente* no paradigma pronominal teria sido:
  - (i) **nós** V + -mos > **nós** V + -moØ > **nós** V + -ão > **nós** V + Ø > **a gente** V + Ø
- (C) Em (i) os tempos verbais inicialmente afetados foram:
  - (ii) V [- passado] > V [+passado]
  - (iii) V [- saliente] > V [+saliente] fonicamente

A seleção dessas variáveis foi feita tendo em vista a hipótese geral dessa pesquisa e os resultados de trabalhos já realizados sobre a variação em análise.

## 4 A análise

Farei a seguir um relato dos fatores verificados. Paralelamente apresentarei os resultados obtidos.

O primeiro fator estrutural testado foi a realização morfológica da pessoa verbal. Esse fator busca distinguir a realização do verbo na 1<sup>a</sup>. pessoa do plural ou na 3<sup>a</sup>. pessoa do singular. A fim de separar presença e ausência de desinência número-pessoal e tempo verbal, apenas a desinência foi observada. O propósito foi verificar se a desinência de 1PP (-mos; -moØ; -ão) inibe o uso da variante **a gente**.

Os valores obtidos (em percentual e peso relativo, doravante PR) indicam que é preferencial o uso de **a gente** com verbo na 3PS (61%; PR.61), havendo apenas 3% (PR.00) de ocorrências com verbo na 1PP. Já o maior uso do pronome **nós** é com verbo na 1PP (96%; PR.99), sendo que, na 3PS é de (38%; PR.30). Os valores probabilísticos obtidos confirmam os resultados percentuais. Chama a atenção aqui o quanto a morfologia verbal de 1<sup>a</sup>. PP inibe a variação (PR.99).

Além das ocorrências de **a gente** e **nós** com formas finitas (de 3PS e de 1PP), foram obtidas ocorrências desses pronomes com formas não-flexionadas (gerúndio),<sup>8</sup> perfazendo um percentual de 57% (PR.55) dos casos com **a gente**. Veja-se que esta porcentagem assemelha-se à da 3<sup>a</sup>.PS, que é de 61% (PR.69). A diferença entre os respectivos valores percentuais é de apenas 4%, o que seria indicativo de que a ausência de marcas morfológicas de pessoa favorece o uso dessa variante.<sup>9</sup>

O segundo fator testado foi tempo verbal. O objetivo foi distinguir a ocorrência da forma verbal no presente ou em outros tempos verbais. A observação desse fator considerou a hipótese de que o presente, por seu aspecto indeterminado, favoreceria o uso de um termo mais genérico, no caso, a variante **a gente**. Essa hipótese resultou das observações feitas por Alves (1998: 68) que, ao estudar as novas formas de indeterminação do sujeito, mais especificamente ocorrências com *você*, *oê*, *cê* encontrou 56% dos casos com verbos no presente, forma verbal “em que não há delimitação temporal para a realização de ação”.

Sendo a variante **a gente** usada com traço [+genérico] e o tempo presente indicador de indeterminação, a hipótese levantada é de que essa forma pronominal seja favorecida por esse tempo verbal. Teria sido esse tempo verbal a porta de entrada de **a gente** no paradigma pronominal?

As formas de infinitivo e gerúndio foram classificadas como [-passado], por compartilharem com o presente o aspecto [-conclusivo].

---

<sup>8</sup> As formas de infinitivo foram consideradas como de 3PS, por tratar-se, segundo Cunha & Cintra (1985: 473) de “infinitivo pessoal” e “terem sujeito próprio”.

(c) Hoje eu vim cá pra **nós** *averta...* (D.o., 80, F3, BH)

(d)...tudo isso influí também **a gente** *rotá* (GR, 60, f2, BH)

<sup>9</sup> Segundo BENVENISTE (1988; apud LOPES, op.cit., p.16), a 3PS é a “não-pessoa”, pois opõe-se, no discurso, à pessoa que fala (eu) e à pessoa que ouve (tu). GALVES (2000), no quadro gerativista, analisa a 3PS como não-pessoa. As formas impessoais do verbo também apresentam realização padrão e não-padrão: [andando] / [andanU]; [comprar] / [compra]. Entretanto, a variação na realização dessas formas não foi considerada.

Os resultados confirmaram que a variante **a gente** é favorecida pelo tempo verbal [-passado], 70% (PR.66). Esse resultado confirmou a expectativa de que o traço [-passado], de aspecto indeterminado e duração ilimitada, favoreceria o traço genérico da variante **a gente**. Por sua vez, a variante pronominal **nós** mostra-se mais favorecida pelo tempo verbal [+passado], correspondendo a 60% (PR.63).

O terceiro fator verificado foi referência, tendo em vista a hipótese de que contextos de referência [+genérica] favoreceriam a variante **a gente**, pois esta seria resultante de um processo de *gramaticalização* do item lexical *gente* (lat. *gens, gentis*), conforme Menón (1996); Omena & Braga (1996); Lopes (1999), e possuiria uma significação mais ampla.

O traço genérico da variante **a gente** é definido por Lopes (1999: 35) nos seguintes termos: “O **a gente** pronominal designa, mais comumente, um todo abstrato, indeterminado e genérico, representando o conjunto base “SER PESSOA (...) massa indeterminada de pessoas disseminada na coletividade.”

Dessa forma, é possível substituir **a gente** por “se”, por “qualquer pessoa”, ou por ”a pessoa”, que são expressões de indeterminação, sem que o enunciado deixe de ser adequado ao contexto, portanto, nesse contexto, é forma Referencial [+genérica].

As variantes **nós** e **agente** foram classificadas quanto à referencialidade em Referência [ $\pm$ genérica].

Os valores obtidos confirmam a hipótese de que a variante **nós** não é favorecida pelo traço [+Ref +Gen]: 31% dos casos (PR.34). Entretanto, esses valores revelam que há um espaço para a referência genérica no uso da variante **nós**. (Cf. Menón, 1994; Coveney, 2000), traço coincidente com a história da variante **a gente**, (Cf. Abraçado, 1991 e Lopes, 1999).

O quarto fator investigado foi a realização fonológica da desinência de número e pessoa. Busquei verificar se a forma padrão no Português do Brasil (PB) era realizada ou não. Com a quantificação desse fator, tive o

propósito de confirmar a ocorrência do processo descrito na sequência (i) da hipótese (C), repetido abaixo:

(i) **nós** V + -mos > **nós** V + -moØ > **nós** V+ -ão > **nós** V+ Ø > **a gente** V + Ø

Nesse processo a desinência de 1PP sofre erosão lenta, ratificando o enfraquecimento morfológico, inicialmente com a perda do *-s* desinencial e, depois, com a perda do *-mo*.

Em ambas as ocorrências, existe a mesma pessoa verbal, mas, do ponto de vista fonológico, há diferença.

Também é considerada não-padrão a seguinte ocorrência encontrada no *corpus* de Pombal:

(10).    **Nóis** *vão* lanchá num lugá lá. (R. A., 20, F1,Pb)

Essa forma parece ser uma variação fonética da forma padrão. Provavelmente, teria havido supressão do *-s* do morfema *-mos* resultando '*vamo*'. Em seguida, teria havido a formação de uma só sílaba, em vez de duas: *vamo* > *vão*, resultando um ditongo nasal.

Dessa forma, em (15) a forma *vão* seria mais um dos estágios de implementação da mudança, e não a forma verbal de 3<sup>a</sup> pessoa do plural como poderia parecer.

Esse fator é uma evidência a favor da hipótese de que a mudança afeta mais imediatamente o paradigma morfológico, via fonologia e, posteriormente, afeta o paradigma pronominal, dando-se a mudança representada por **nós** > **a gente**.

Os resultados mostraram que a desinência número-pessoal do verbo pode se realizar foneticamente como padrão ou não-padrão. Em termos quantitativos, a forma pronominal **a gente** ocorre em 92% dos dados (PR.92) com a desinência número-pessoal padrão, ou seja, tem maior realização na 3PS. Já a forma pronominal **nós** apresenta um comportamento intrigante. Tem apenas 7% das ocorrências

*-mos* (PR.07). A maior parte de suas ocorrências, 93% (PR.95) é com [-Padrão], isto é, há erosão fonética na maioria das ocorrências.

Após refinamento dos dados, os valores percentuais obtidos confirmaram que a desinência de 1<sup>a</sup>. PP é um fator que favorece o **nós**. Com o morfema padrão *-mos*, há 100% das ocorrências com a variante **nós** e 0% com a variante **a gente**; indicando que essa terminação favorece o **nós**. Com o morfema *-moØ*, há 96% das ocorrências com a variante **nós** e apenas 3% com a variante **a gente**. A terminação *-ão* mostrou favorecimento categórico em relação à forma *nós* (100%). Mesmo quando a desinência é *zero*, a porcentagem de **nós** é de (38%)! Desse modo, os resultados evidenciam que a desinência se desfaz bem lentamente e que um vestígio mínimo de 1<sup>a</sup>. PP é o bastante para favorecer a ocorrência do pronome **nós**.

Parece haver aqui uma evidência de que não é a entrada da forma lexical **a gente** que desencadeou o uso da 3PS. Se fosse, não teríamos um índice tão alto de **nós** com verbo na 3PS, representada por zero na tabela. Estamos aqui argumentando a favor de que formas de 3PS são, de fato, ambíguas, pois um item como *era* é, ao mesmo tempo, resultado de erosão fonética de *-mos*, a partir de *éramos*, como também é resultado de verbo *ser* + 3PS.

O quinto fator selecionado foi a saliência fônica, a partir da proposta de análise feita em Naro & Lemle (1977). Ao estudarem a ausência de concordância no Português do Brasil, os autores levantaram a hipótese de que tal ocorreria mais frequentemente com as formas verbais nas quais o morfema de plural é aplicado por último, ou seja, aquelas em que a forma singular difere pouco da forma fonética originada do verbo, sendo, portanto, raramente percebida.

Interessou-me capturar os casos de ambiguidade entre 3<sup>a</sup>. pessoa do singular (3PS) e realização zero da desinência *-mos*. Por exemplo, o par *come/comemos* seria ambíguo, já o par *foi/fomos* não seria. *Come* pode resultar de apagamento de *-mos* em *come+ -mos*, já *foi* não pode

resultar do apagamento de *-mos* em fomos, porque, se tivesse havido apagamento, o resultado seria *\*fo* e não *foi*.

De acordo com os resultados, o nível de saliência fônica que mais favorece o uso da variante ***a gente*** (70%; PR.66) é o que contém palavras nas quais o acréscimo do morfema de plural *-mos* ao radical altera o vocábulo quanto à posição da sílaba tônica ao ser usado com o pronome ***nós***.

O segundo nível a favorecer o uso de *a gente* (62%; PR.59) é o nível em que não há nenhuma alteração fônica na mudança de 3PS para 1PP.

O terceiro nível em termos de favorecimento do uso da forma ***a gente*** é o que contém palavras que sofrem completa alteração ao passarem de ***a gente*** para ***nós*** (é/somos), com um percentual de 61% (PR.58).

Em quarto lugar está o nível das formas nominais do verbo (gerúndios), com um percentual de 57% (PR.53).

Os casos de paroxítonas que passam a proparoxítonas (51%), e palavras que apresentam algum tipo de alteração (alterações vocálicas: acréscimo; perda; omissão; hiatização; ditongação; alçamento; metátese; etc) com percentual de 42%, com pesos relativos.<sup>47</sup> e.39, respectivamente, já não favorecem o uso da variante ***a gente***.

Pelos resultados obtidos, a saliência fônica não se mostra um fator quantitativamente significativo, tendo sido excluído pelo programa estatístico. Esse fator, entretanto, mostra resultados curiosos. A variante ***a gente*** tem maior peso relativo nas formas verbais menos salientes, como *dar* e *ver*, e nas formas ambíguas: formas verbais como *muda*, que pode ser resultado de *mudamos* > *muda* + *-mos* ou de *muda* + *zero*.

O sexto e o sétimo fatores são extralingüísticos. São *faixa etária* e *região geográfica*. O primeiro fator extralingüístico teve por objetivo a realização da análise em *tempo aparente*. Os entrevistados foram agrupados em três faixas etárias. A faixa 1 (f1) é formada por indivíduos jovens, de 20 a 39 anos. A faixa 2 (f2) é formada pelos medianos, indivíduos de 40 a

59 anos; e a faixa 3 (f3), pelos indivíduos mais velhos, com idade acima de 60 anos.

Os resultados mostram que os jovens (que chamo de Faixa Etária 1) estão usando mais a forma **a gente** (63%, PR.60) do que as pessoas da Faixa Etária 2 (59%, PR.55) ou da Faixa Etária 3 (41%, PR.38). Segundo Labov (1972), quando é alto o índice na faixa etária dos jovens, é porque está havendo *mudança em progresso*. Em termos probabilísticos, é confirmada a tendência ao uso da variante inovadora **a gente** no lugar da forma pronominal conservadora **nós**.

O gráfico abaixo permite visualizar o uso das variantes nas faixas etárias verificadas:

GRÁFICO 1 - Faixa etária

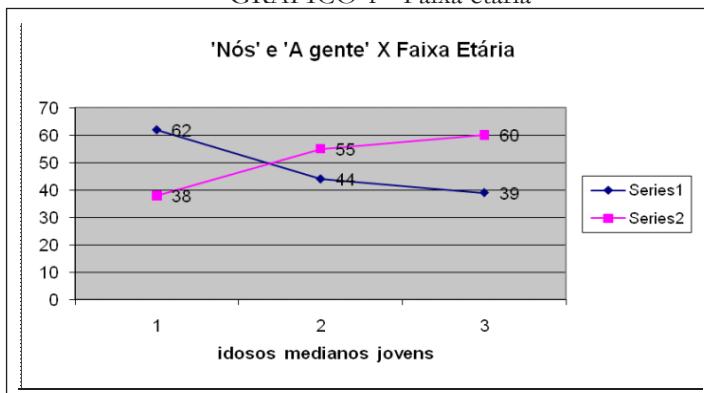
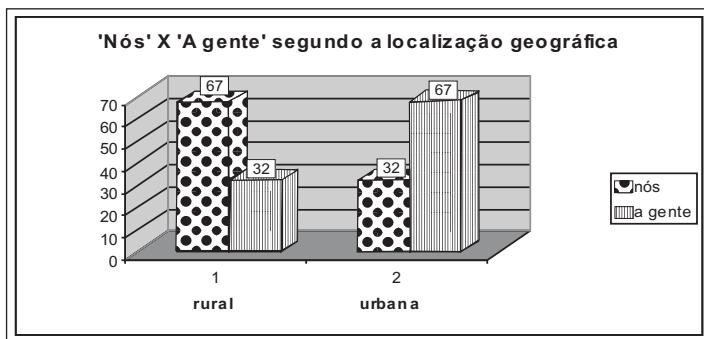


GRÁFICO 2 - Localização geográfica



## 5 A transição de *a gente*

Segundo Weinreich, Labov & Herzog (1968: 184) *transição* é “the intervening stage which defines the path by which Structure A evolved into Structure B”.

Os resultados acima apresentados nos permitem afirmar que a mudança **nós > a gente** teria tido o seguinte percurso:

(a) presença do pronome NÓS acompanhado de V (verbo) com a realização plena da desinência de número-pessoal da 1<sup>a</sup>. pessoa do plural → -MOS;

(11)...seja mais o nós, **nós** *podemos*. (A.H., 34, f1, BH)

Conforme vimos na tabela sobre a ocorrência da desinência de 1PP com as variantes **nós / a gente** é inexistente a ocorrência da terminação *-mos* com a variante **a gente** nos *corpora* analisados.

(b) presença do pronome NÓS / A GENTE acompanhado de V (verbo) com a realização parcial da desinência de número-pessoal da 1<sup>a</sup>. pessoa do plural → MOØ;

(12)      **Nós** *tamo* hoje no dia sete, né? (P.S., 27, f1, BH)

(13)      ... **a gente** *duramo* quase dois meses. (I.A., 22, f1, BH)

(c) presença do pronome NÓS / A GENTE acompanhado de V (verbo) com a realização parcial da desinência de número-pessoal da 1<sup>a</sup>. pessoa do plural → -ÃO;

(14)      **Nóis** *tão* tudo pricisano distrui dente aí, né... (C.,31,f1,Pb)

Não foi encontrado nenhum caso de **a gente** + *-ão* nos *corpora* analisados.

(d) presença do pronome NÓS acompanhado de V(verbo) com a ausência total da desinência de número pessoa. Neste momento, a forma pronominal **a gente** adquire força para ocupar o lugar da forma pronominal **nós**, pois a forma verbal é ambígua.

(15) No mais **nóis** *fica* é em casa mesmo... (C., 31, f1, Pb)

(16) **A gente** *ficava* sempre junto... (L.C., 69, f3, BH)

No exemplo (15) acima a forma verbal *fica* apresenta ambiguidade morfológica. Tanto pode ser forma de 3PS, portanto, apresenta morfema  $\emptyset$ , quanto pode ser forma de 1PP sem a realização fonológica da desinência *-mos*.

Observando o exemplo (16), vemos que o mesmo acontece. Entretanto, é com as formas de tempo [-passado] que a ambiguidade será maior. É constante na literatura linguística (Machado, 1995; Lopes, 1999) a hipótese de que formas verbais menos marcadas (presente do indicativo e infinitivo pessoal) condicionam o uso de **a gente**; e formas verbais morfológicamente mais marcadas favorecem o emprego de **nós** (pretéritos; futuro do indicativo e formas do subjuntivo). Entretanto, vimos que esse fator não foi considerado significativo pelo programa de análise multivariada utilizado. Apesar disso, é interessante observarmos a atuação desse fator na variação de **nós** e **a gente**.

Se observarmos o comportamento da forma verbal dos exemplos (15) e (16), nos tempos mencionados como menos marcados morfológicamente, temos que, tanto no presente do indicativo, quanto no infinitivo pessoal, essa forma, quando destituída da desinência de 1PP, apresenta ambiguidade: *fica* + *mos* > *fica* (3PS) / *fica* ( $\emptyset$ ).

E o que acontece nas formas morfológicamente mais marcadas?

Comecemos pelo exemplo (16), que apresenta uma forma verbal de pretérito (Pretérito Imperfeito): *ficáva+ mos > ficava (3PS) / ficava (Ø)*. Observamos que, caso não haja a realização da desinência de 1PP, também teremos ambiguidade. Nos demais tempos verbais teremos:

- a) *Pretérito Mais- que-perfeito (simples): ficara+ mos > ficara (3PS) / ficara (Ø);*<sup>10</sup>
- b) *Pretérito Mais-que-perfeito (composto): tinha+mos ficado > tinha (3PS) ficado / tinha (Ø) ficado;*
- c) *Pretérito Perfeito: fica+ -mos ≠ ficou (3PS);* por sua vez, *fica (Ø)* é forma verbal do Presente.
- d) *Futuro do Indicativo (simples): ficar (e)+ -mos / ficar (á) ≠ ficar (Ø);* também neste tempo a ausência da desinência modo-temporal descaracteriza o tempo em análise;
- e) *Futuro do indicativo (composto): \*va + -mos ficar;* aqui a ausência da desinência modo-temporal não é aceitável;
- f) *Presente do Subjuntivo: fique+ -mos > fique (3PS) / fique (Ø);*
- g) *Pretérito Imperfeito do Subjuntivo: ficasse + -mos > ficasse (3PS) / ficasse (Ø);*
- h) *Futuro do Subjuntivo: ficar + -mos > ficar (3PS)/ ficar (Ø).*

Por esta visão rápida sobre o comportamento da forma verbal dos exemplos (30) e (31), observamos que o favorecimento da variante **a gente** ou da variante **nós** parece estar relacionada à possibilidade de ambiguidade morfológica verbal. Portanto, a hipótese de que formas dos Pretéritos, do Futuro do Indicativo e do Subjuntivo favoreceriam **nós**, não se sustenta. Conforme visto acima, as formas verbais dos Pretéritos não têm o mesmo comportamento em relação à ambiguidade morfológica, ou, à aceitabilidade da ausência da desinência de 1PP. O

<sup>10</sup> O Pretérito Mais-que-Perfeito simples foi aqui mencionado por estar no rol de todos os tempos verbais mencionados pelas gramáticas tradicionais. Ressalte-se o fato de que este tempo verbal não apresenta uso na língua portuguesa contemporânea falada no Brasil, há muito tendo sido substituído pelo Pretérito Mais-que-Perfeito composto.

gráfico a seguir permitirá visualizar o fenômeno das realizações da desinência número pessoal da 1<sup>a</sup> pessoa do plural em Pombal.

Observando-se a realização da desinência de 1PP em relação à forma pronominal **nós**, encontram-se as seguintes situações: a desinência *-mos* não ocorreu nenhuma vez nos dados de Pombal (0%). Em 12 ocorrências, a desinência *-mo* ocorreu 100%, o mesmo acontecendo com a desinência *-ão*, que teve 05 ocorrências; ou seja, em Pombal as terminações *-mo* e *-ão* têm ocorrência categórica com a variante **nós**. Por sua vez, a desinência *-Ø*, com percentual de 62% de ocorrências revela uma possível abertura para a substituição da variante **nós** pela variante **a gente**.

Todavia, se a inserção do pronome fosse o fator responsável pelo enfraquecimento da desinência verbal, não teríamos **nós** + *zero*. Em Pombal, a variante **nós**, embora não supere a variante **a gente** neste contexto, apresenta 37% de ocorrência, o que é um índice surpreendente, conforme evidenciado no gráfico abaixo:

Objetivando-se à reconstrução desse percurso, o fato de serem significativos os fatores: pessoa verbal e realização fonológica da desinência número-pessoal constitui evidência de que o gatilho da mudança **nós /a gente** foi a morfologia. Conforme visto na análise do fator pessoa verbal, mais bem detalhado pelo fator realização fonológica, a existência de um traço mínimo da desinência de 1PP é o suficiente para a não ocorrência da forma verbal com a forma pronominal **a gente**.

Os fatores referência e tempo verbal, por sua vez, forneceram evidências de que a aceitabilidade da forma **a gente** pela forma verbal já morfologicamente enfraquecida foi favorecida pelo traço [+genérico] carregado por essa forma desde sua origem, conforme bem visto em Lopes (1999). É por isso que os altos percentuais de ocorrência de **a gente** foram com o tempo verbal [-passado], 70% (PR.67; e com referência [+genérica], 68%, (PR.64).

De acordo com os resultados obtidos na investigação do Fator Faixa Etária, os jovens estão usando mais a forma **a gente** (PR.60) do que os medianos (PR.55) e do que os idosos (PR.38), o que significa *mudança em progresso*. Segundo Weinreich, Labov & Herzog (1968: 171) é preciso ‘distinguir-se *age grading* de *mudança em progresso*’. O primeiro fenômeno é uma variação linguística própria de uma faixa etária, com tendências a sumir; e não propriamente de uma variação linguística onde uma das formas que co-variaram acaba por impor-se após algumas gerações de falantes. O primeiro caso é descrito em Lopes (1999): apesar de a faixa etária dos jovens apresentar alto índice de uso da forma *a gente* (PR.77), os índices apresentados pelo grupo dos medianos e dos velhos são muito próximos (PR.39 e PR.40, respectivamente), caracterizando *age grading*.

Observando-se, neste trabalho, os pesos relativos atribuídos às três faixas etárias, é possível confirmar que se trata de *mudança em progresso*, uma vez que o peso relativo atribuído aos jovens apresenta um uso bastante significativo da variante **nós** (PR.60) e é bem alto em relação à geração dos medianos (PR.55) e dos idosos (PR.38); por sua vez, a geração dos medianos apresenta um uso bastante significativo da variante **nós** (PR.55), o que afasta a hipótese de *age grading*.

## 6 Contextualizando os fatos estudados na transição ‘nós’/‘a gente’

O cenário apresentado na seção anterior permite capturar um paralelo<sup>11</sup> entre a inserção de *você* e a inserção de **a gente** no sistema pronominal do PB: ambos decorrentes do enfraquecimento da morfologia verbal iniciada na fonologia. Em relação a *você* teria havido o

<sup>11</sup> Considero que a mudança ‘nós’/‘a gente’ envolve um conjunto de mudanças que se caracteriza, na sociolinguística laboviana como um caso de encaixamento (cf. Labov, 1994). Desse modo, os fatores determinantes do início e da continuidade das mudanças de *nós* / *a gente* não residem neles mesmos, mas estes itens mudam de modo simultâneo em consequência de outras mudanças que estão ocorrendo ou que já ocorreram em subsistemas linguísticos relacionados, conforme apontados nesta seção.

seguinte, segundo Oliveira & Ramos (2000): a perda do /d/ intervocálico que atingiu os morfemas número-pessoais da 2<sup>a</sup>. pessoa do plural no português medieval, favorecendo o aparecimento de duas gramáticas – uma que reconstitui o paradigma verbal com a inserção da semivogal e, outra que, resultado de uma crase, faz surgir um paradigma verbal com formas neutralizadas entre a 2<sup>a</sup>. pessoa do singular e a 2<sup>a</sup>. pessoa do plural. A primeira é a gramática do Português Europeu; a segunda, a do Português Brasileiro. É nessa última que a forma de tratamento *vocé* passa a fazer parte do paradigma pronominal. (Cf. Oliveira & Ramos, 2000).

Lembrando o que dizem Weinreich, Labov & Herzog (1968: 172) sobre o encaixamento:

“Linguists are naturally suspicious of any account of change which fails to show the influence of the structural environment upon the feature in question; it is reasonable to assume that the feature is embedded in a linguistic matrix which changes with it.”

Vejamos, pois, como a variante inovadora *a gente* ingressa no paradigma pronominal do dialeto mineiro.

Em relação a ***a gente*** temos na amostra analisada a documentação do enfraquecimento fonético da desinência *-mos*. Conforme visto na análise, a erosão da desinência da 1PP inicia-se no nível fonético e atinge o morfológico, processo evidenciado pelos fatores *Pessoa Verbal* e *Realização fonética da Desinência de Número e Pessoa*. O espaço do [+genérico] existente no uso da variante conservadora *nós*, demonstrado pelo fator *Referência*, também favoreceu a inserção da variante inovadora *a gente* no quadro pronominal.

## 7 A propósito do debate sobre paradigma verbal no Português Brasileiro

As descrições acima inserem-se num debate que atualmente ocupa os linguistas brasileiros, no qual se identificam duas posições aparentemente contrárias.

A primeira posição é a que é reforçada neste trabalho. O paradigma verbal sofre perdas, de início fonológicas, passando a morfológicas. Essas perdas, ou reduções nas formas verbais, ao se tornarem frequentes, abrem espaço para a inserção da nova forma pronominal, alterando assim o paradigma pronominal. Na mudança **nós/a gente** a realização zero da desinência número-pessoal de 1PP é, então, reanalisada pelo falante como desinência de 3<sup>a</sup> pessoa, fazendo com que a forma **a gente** comece a fazer parte do paradigma pronominal. (Cf. Abraçado, 1991).

A segunda posição teórica, largamente presente na literatura linguística, defende que o inverso ocorre. Surge uma forma pronominal. Esta entra em concorrência com outra já existente. Após certo tempo, a forma existente termina por perder seu lugar no sistema pronominal para a forma inovadora.

Portanto o presente trabalho mostra que, ao contrário do que argumentam vários linguistas, dentre eles Duarte (1993), Roberts (1993), Galves (1993), Faraco (1996), Menón (1994, 1995, 1996), Lopes (1999) e Zilles (2002), temos no estudo da mudança *nós / a gente* uma evidência de que alterações morfológicas na forma verbal favoreceram a alteração do paradigma pronominal.

## 8 Conclusão

A realização deste trabalho teve como objetivo principal verificar se no dialeto mineiro há variação entre **nós / a gente**, se essa variação configura *mudança em progresso*; em caso afirmativo, qual teria sido o percurso da mudança.

Diferentemente das abordagens anteriores, tivemos a preocupação de verificar fatores inibidores da variante **a gente**. Feita a análise quantitativa e qualitativa das ocorrências, podemos concluir que:

- 1 - Há variação entre as formas **nós / a gente**;
- 2 - A variação de **nós / a gente** configura *mudança em progresso*; portanto, a forma **a gente** é uma inovação;
- 3 - Enquanto há um traço mínimo da desinência de 1<sup>a</sup>. pessoa do plural no verbo o pronome **nós** resiste à ocorrência de **a gente**;
- 4 - Verbos utilizados no [-passado] favorecem a ocorrência da forma inovadora **a gente**;
- 5 - A forma **a gente** tem sua ocorrência favorecida pela referência [+genérica]; o que está diretamente relacionado à sua origem;
- 6 - Os fatores testados (Pessoa Verbal; Realização Fonológica da Desinência Número-pessoal; Número de Morfemas) comprovam a força da morfologia na variável analisada;
- 7 - A mudança apresenta ritmo mais lento em comunidades interioranas.

Espero que este trabalho tenha contribuído para explicitar parte de um fenômeno presente na gramática do dialeto mineiro, fornecer argumentos para o debate referente ao paradigma pronominal e, do ponto de vista metodológico, ter mostrado a relevância de tratar áreas geográficas distintas como amostras temporalmente diferentes, ao assumir que as mudanças não caminham num mesmo ritmo em comunidades dessemelhantes.

## Referências

- ABRAÇADO, A. M. J. **Mudanças no sistema pronominal do português brasileiro: causas e consequências**. Belo Horizonte: Faculdade de Letras da Universidade Federal de Minas Gerais, 1991. Dissertação (Mestrado).
- ALVES, Nilton. **Formas de indeterminação do sujeito**. Belo Horizonte: Faculdade de Letras UFMG, 1998. Dissertação (Mestrado).
- COVENEY, Aidan. **Vestiges of nous and the 1<sup>st</sup> person plural verb in informal spoken French**. Languages Science [Department of French, University of Exeter, Exeter], 2000. 22. p.447-481.
- CUNHA, C.; CINTRA, L. **Nova gramática do português contemporâneo**. Rio de Janeiro: Ed. Nova Fronteira, 1985.
- DUARTE, M. E. L. **Do pronome nulo ao pronome pleno**. A trajetória do sujeito no português de Brasil. In: ROBERTS, Ian & KATO, Mary A. (orgs). **Português Brasil: uma viagem diacrônica**. Campinas, Editora da Universidade Estadual de Campinas: 1993. p.107-128.
- FARACO, C. A. **O tratamento você em português**: uma abordagem histórica. In: **Fragments**. Curitiba: Editora da UFPR, 1999. 13: p. 51-82.
- GALVES, C. **O enfraquecimento da concordância no português Brasileiro**. In: ROBERTS, Ian & KATO, Mary A. (orgs). **Português Brasileiro: uma viagem diacrônica**. Campinas: Ed. Unicamp, 1996. p. 387-408.

- LABOV, Willian. **Sociolinguistic patterns**. [= Conduct and Communication, 4], Philadelphia: University of Pennsylvania Press, 1972.
- \_\_\_\_\_. **Principles of Linguistic Change: social factors**. Oxford, USA; Balckewell Publishers, 2001.
- \_\_\_\_\_. **Principles of Linguistic Change: internal factors**. Oxford: USA: Balckewell Publishers, 1994.

LOPES, C. R. dos S. **A inserção de “a gente” no quadro pronominal do português**: percurso histórico. Faculdade de Letras da Universidade Federal do Rio de Janeiro. 1999. Tese (Doutorado).

MACHADO, M. dos S. **Sujeitos pronominais “nós” e “a gente”**: variação em dialetos populares norte-fluminenses. Faculdade de Letras da Universidade Federal do Rio de Janeiro. 1995. Dissertação (Mestrado).

MENÓN, Odete da S. **Analyse sociolinguistique de l'indetermination du sujet dans le portugais parlé au Brésil, à partir des données du NURC**- São Paulo. Département de Recherches Linguistiques, Université de Paris VII. 1994. Tese (Doutorado).

\_\_\_\_\_. **O sistema pronominal do português do Brasil**. Curitiba: Ed. da UFPR, 1995. n 44, pp. 91-106.

\_\_\_\_\_. **A gente**: um processo de gramaticalização. Estudos linguísticos, [Anais de Seminários do GEL]. XXV Taubaté: UNITAU/ CNPq/ GEL, 1996. p. 622-628.

NARO, A. **Variação e Funcionalidade**. Revista de Estudos da Linguagem. Belo horizonte: Faculdade de Letras da UFMG, 1998. v. 7, n. 2, p. 109-120.

NARO, Anthony ; LEMLE, Miriam. **Syntactic diffusion**. Revista Ciência e Cultura. Rio de Janeiro: [Sociedade Brasileira para o Progresso da Ciência]. 1977. v. 29, n. 3, p. 259-268

NASCENTES, Antenor. **O linguajar carioca.** 2. ed. Rio de Janeiro: Organização Simões, 1953.

OLIVEIRA, M. A. **Sobre os reflexos sociais da mudança em progresso.** Ensaios de Linguística. [Belo Horizonte, Universidade Federal de Minas Gerais], 1982. n. 7, p. 71-89.

\_\_\_\_\_. RAMOS, J.M. **O estatuto de ‘você’ no preenchimento do sujeito.** No prelo. 2000.

OMENA, Nelize P.; BRAGA, M. L. **A gente está se gramaticalizando?** In: MACEDO, A. T.; RONCARATI, C.; MOLLICA, M.C. (orgs). Variação e Discurso. Rio de Janeiro: Tempo Brasileiro, 1996. p.75-83.

\_\_\_\_\_. **A referência à primeira pessoa do plural: Variação ou mudança?** In: PAIVA, M.C.; DUARTE, M.E.L. (orgs). **Mudança Linguística em Tempo Real.** Rio de Janeiro: Contra Capa / FAPERJ, 2003. p. 63-80.

\_\_\_\_\_. **A referência à primeira pessoa do discurso no plural.** In: SILVA, G. M.; SCHERRE, M. M.P. (org). **Padrões Sociolinguísticos: Análise de Fenômenos Variáveis do Português Falado na Cidade do Rio de Janeiro.** Rio de Janeiro: Ed. Tempo Brasileiro, 1996. p. 185-215.

RAMOS, Jânia M. **Avaliação de dialetos brasileiros:** o sotaque. Revista de Estudos Linguísticos, [Belo Horizonte, Faculdade de Letras da Universidade Federal de Minas Gerais], 1997. n. 5, p. 141-146.

\_\_\_\_\_. **A alternância entre “não” e “num” no Dialeto Mineiro-Um caso de mudança linguística.** In: COHEN, Maria Antonieta e Jânia M.Ramos (organizadoras). **Dialeto Mineiro e outras Falas:** estudos de variação e mudança linguística. Belo Horizonte: Editora da Universidade Federal de Minas Gerais, 2002. pp. 31-40.

\_\_\_\_\_. **O uso das formas você, ocê e ce no Dialeto Mineiro.** In: HORA (org). **Diversidade Linguística no Brasil.** João Pessoa: Idéia Editora, 1997. p. 43-60.

ROBERTS, I. **O Português no contexto das línguas românicas.** In: ROBERTS, I; KATO, M. A. (orgs) **Português Brasileiro. Uma viagem diacrônica.** Campinas: Editora UNICAMP, 1993. p. 409-421.

SCHERRE, M.M.P et alii. **Introdução ao pacote VARBRUL para microcomputadores.** UFRJ, Faculdade de Letras. Departamento de Linguística e Filologia. Projeto de Estudo sobre o uso da língua (PEUL) 1992. (mimeografado).

TAGLIAMONTE, S.; LAWRENCE, H. **GOLDVARB 2001:** A multivariate analysis Application for Windows. (impresso)

WEIREICH, U.; LABOV, W.; HERZOG, M. **Empirical foundations for a theory of language.** 1968. In: MALKIEL (eds). Perspective on historical linguistics. Amsterdam: Philadelphia, John Benjamins Publishing Company, 2001. p. 97-193.

ZÁGARI, Mario R. L. et alii. **Esboço de um Atlas linguístico de Minas Gerais.** Rio de Janeiro: Fundação Casa Rui Barbosa/Juiz de Fora, Universidade Federal de Juiz de Fora, 1977.

ZILLES, A. M. S. **Gramaticalization of ‘a gente’ in Brazilian Portuguese.** University of Pennsylvania Working Papers Linguistics, 2002. v. 8-3.

# CONTATO LINGUÍSTICO E MUDANÇA LINGUÍSTICA NO NOROESTE AMAZÔNICO: O CASO DO KOTIRIA (WANANO)<sup>1</sup>

Kristine STENZEL

Universidade Federal do Rio de Janeiro (UFRJ)

Elsa GOMEZ-IMBERT

CNRS/ERSS Université Toulouse-Le Mirail

## RESUMO

*Neste trabalho, examinamos a situação de contato linguístico Aruák/Tukano Oriental no noroeste amazônico e seus efeitos em línguas TO da sub-região do Uaupés. Consideramos também algumas hipóteses propostas em Aikhenvald 2002 quanto à natureza da difusão em situações de contato (múltilateral/unilateral, direta/indireta) e os resultados linguísticos (enriquecimento vs. empobrecimento / abandono de padrões).*

## ABSTRACT

*In this paper we examine the situation of Arawak/Eastern Tukanoan linguistic contact in northwest Amazonia and its effects in the ET languages of the Vaupés sub-region. We also consider certain hypotheses proposed in Aikhenvald 2002 regarding the nature of diffusion (multilateral/unilateral, direct/indirect) in contact situations and its linguistic consequences (enrichment vs. abandonment of patterns).*

---

<sup>1</sup> A pesquisa das línguas Kotiria (Wanano) e Wa'ikhana (Piratapuyo) recebeu apoio financeiro da Endangered Languages Fund, da Wenner-Gren Foundation for Anthropological Research, da National Science Foundation (0211206), da NSF/NEH Documenting Endangered Languages Program (FA-52150-05), do CNPq, e do Hans Rousing Endangered Languages Documentation Program - SOAS/University of London (MDP-0155), bem como apoio institucional e logístico no Brasil do Instituto Socioambiental, do Departamento de Linguística e do Programa de Pós-Graduação em Antropologia Social (PPGAS)/MN da Universidade Federal do Rio de Janeiro.

## PALAVRAS-CHAVE

*Contato lingüístico. Línguas Aruak. Línguas Tukano Oriental.*

## KEYWORDS

*Arawak languages. Eastern Tukanoan languages. Linguistic contact.*

## Introdução

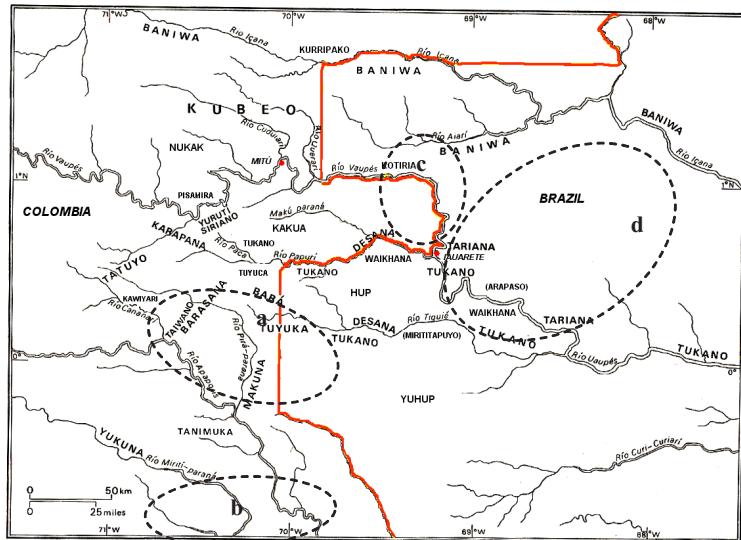
A região do Uaupés, no noroeste amazônico na fronteira entre Brasil e Colômbia, é conhecida por seu multilinguismo e seu sistema social baseado em exogamia linguística, em que os casamentos só são permitidos entre indivíduos que falam línguas diferentes (Sorensen 1967, 1985; Gomez-Imbert 1991, 1999; Jackson 1983; Aikhenvald 1999; Stenzel 2005a). Nesse sistema, cada indivíduo “fala” a língua do pai enquanto, por mais proficiente que seja, apenas “imita” outras, inclusive a língua da mãe. Essa distinção reflete o fato de que o uso ativo e exclusivo da língua do pai indica a identidade da pessoa como membro de um grupo social patrilinear e linguístico. Além da exogamia linguística, a norma da virilocalidade junta numa mesma maloca ou aldéia esposas de vários grupos exogâmicos. Assim, cada criança se forma num ambiente doméstico bilíngue e numa comunidade multilíngue, adquirindo primeiro a língua da mãe, mas depois adotando o uso exclusivo da língua do pai para afirmar a sua identidade social. Esses padrões de aquisição e uso linguístico tornam inevitáveis as influências mútuas entre as línguas em contato e criam um contexto rico para pesquisadores interessados em situações de mudança linguística.

No entanto, há poucos estudos que investigam as influências resultantes dessa situação de contato, que envolve quinze grupos exógamos da família Tukano Oriental (TO) e quatro grupos Aruák (AR)<sup>2</sup>.

---

<sup>2</sup> Os grupos TO são os Bará, Barasana, Desana, Eduria/Taiwano, Karapana, Kubeo, Makuna,

FIGURA 1. A região do Uaupés e os principais focos de contato TO/AR



Semelhanças resultantes do contato entre línguas de uma mesma família certamente existem, mas são dificilmente distinguidas dos traços compartilhados atribuíveis à herança genética. No entanto, há alguns estudos, por exemplo os de Gomez-Imbert (1991, 1993, 1999, no prelo-*b*; e Gomez-Imbert & Hugh-Jones 2000), que tratam de contato entre as línguas TO da região do Piraparaná. Alguns outros estudos focalizam influências observadas em algumas das principais situações de contato *entre* famílias. Gomez-Imbert (1996) analisa uma categoria grammatical do Kubeo (TO) que revela influência do Baniwa (AR, na

Siriano, Retuarã/Tanimuka, Tatuyo, Tukano, Tuyuka, Yurutí, Kotiria (Wanano), e Waikhana (Piratapuyo). Os grupos AR são os Baniwa/Kurripako, Kawiyari, Tariana, e Yukuna. Nas regiões interfluviais vivem quatro grupos da família Nadahup (Makú)—os Hup, Yuhup, Nukak e Kakua—que não participam do mesmo sistema de casamento, mas que mantêm outros tipos de relações sócio-culturais e econômicos com os grupos ribeirinhos (ver Jackson 1983: capítulo 8; Chernela 1993: capítulo 8; Ribeiro 1995; e Epps 2005a:12-27).

área *c* na figura 1) e atualmente esta mesma pesquisadora investiga a situação de contato entre os Tatuyo, Barasana/Edúuria (Taiwano) (TO), e Kawayari (Kabiyari) (AR) no alto Cananari e no alto Piraparaná (área *a*). Aikhenvald (1999, 2002) descreve o caso de influência massiva da língua Tukano (TO) sobre o Tariana (AR) (área *d*) e resume alguns dados referentes ao contato intenso entre o Tanimuka/Retuarã (TO) e o Yukuna (AR) no baixo Apaporis (área *b*)<sup>3</sup>.

Aikhenvald (2002:1-31) também propõe algumas hipóteses gerais relacionadas aos efeitos linguísticos do contato<sup>4</sup>. Ela afirma, em primeiro lugar, que situações de contato sem língua dominante se caracterizam pela difusão multilateral e gradual de traços, e resultam num enriquecimento mútuo de padrões. A “área linguística” do Uaupés, pelo menos até o século dezenove, seria um caso desse tipo, com convergência de características AR ⇔ ET e enriquecimento dos recursos linguísticos de línguas de ambas as famílias. Por outro lado, situações em que uma língua domina provocam mudanças descontínuas e rápidas, com difusão unilateral de padrões encontrados na língua dominante e o abandono de estruturas não encontradas nela. A língua Tariana (AR), principalmente no século vinte, teria passado por um processo de difusão unilateral da língua Tukano (TO), resultando numa série de isomorfismos estruturais e na criação de uma gramática “Tukanizada” (Aikhenvald 2002:277, tabela 12.1). No mais, Aikhenvald afirma que especificamente no contexto tradicional do Uaupés, a maioria das influências fonológicas, morfológicas e sintáticas atribuíveis ao contato são casos de difusão *indireta*, ou seja, difusão de categorias ou elementos de categorias. O uso de empréstimos lexicais—difusão *direta*—é um fenômeno menor, culturalmente reprimido e limitado ao uso mais freqüente de padrões raros.

<sup>3</sup> Além dos estudos de contato AR/TO, Epps (2005b, 2007) analisa uma série de mudanças estruturais diacrônicas em Hup sob influência de contato com Tukano.

<sup>4</sup> Ver resenhas dessa obra em (Meira & Gomez-Imbert 2005; Stenzel 2005b).

Neste trabalho, examinamos essas hipóteses à luz de dados provindos principalmente do Kotiria (Wanano),<sup>5</sup> uma língua TO falada na sub-região do alto Uaupés, na qual detectamos padrões e inovações atribuíveis à influência Aruák. No processo de análise dos dados, tivemos sempre em mente várias questões de teor metodológico (ver Heine & Kuteva 2005:21-34). Primeiro, consultamos análises de vários tipos—antropológicas, históricas e arqueológicas—para podermos entender o tempo de contato e avaliar a natureza das relações entre os grupos em questão. Segundo, para cada inovação que suspeitamos ser fruto de contato, pesquisamos a literatura linguística existente para verificar se o elemento se encontra em outras línguas TO sem o mesmo contato, e também avaliar se o elemento poderia ter se desenvolvido por outras vias que não a do contato. Finalmente, adotando uma perspectiva histórica em que a dinâmica das relações sócio-lingüísticas tem um papel fundamental, para cada inovação, consideramos a questão de direcionalidade AR ⇔ TO e a possibilidade de ter havido “correntes” de difusões ocorridas em períodos diferentes.

O trabalho tem a seguinte estrutura: a Seção 1 descreve a situação de contato entre os Kotiria e os grupos “cunhados” vizinhos: os Kubeo (TO), Baniwa e Tariana (ambos AR). A Seção 2 descreve as inovações fonológicas atribuídas ao contato, e a Seção 3 focaliza as inovações morfológicas.

---

<sup>5</sup> O grupo é conhecido por vários nomes na região e na literatura, entre outros, Wanano, Uanaño e Guanano. No entanto, em 2006, lideranças das comunidades, junto a diretores, professores e alunos da escola indígena, decidiram em assembléia pela adoção do uso exclusivo de seu nome tradicional, *Kotiria* “povo d’água”, para referência tanto ao povo quanto à língua, pedindo que pesquisadores e outros assessores de fora respeitassem essa importante decisão política e a apoiassem como expressão de autodeterminação e de valorização cultural. Outros grupos da região têm tomado decisões parecidas, entre outros, os Wa’ikhana (Piratapuyo). Acatamos esses pedidos neste artigo.

## 1 O contato entre os Kotiria e os grupos vizinhos

A população Kotiria é de aproximadamente 1.600 indivíduos, 68% dos quais vivem na Colômbia e 32% no Brasil (FOIRN & ISA 2006:43). A maior parte da população ainda vive no território tradicional dos Kotiria no alto do rio Uaupés, entre a cachoeira de Arara e a cidade colombiana de Mitú<sup>6</sup>. Documentos históricos analisados por Wright (2005:80-81) confirmam a ocupação desse território pelos Kotiria em 1740, e a história da região reconstruída através de narrativas orais e pesquisa arqueológica indica uma ocupação muito mais antiga, que antecede a migração dos Tariana ao Uaupés em aproximadamente 700 anos (Amorim 1987; Neves 1988:158, 206; Wright 2005:13).

A localização geográfica dos Kotiria fica na margem noroeste da região, longe do centro ocupado por outros grupos TO, e os mantêm em contato próximo com dois grupos AR com os quais formaram relações de aliança e casamento: os Tariana, cujo território no Uaupés começa pouco abaixo do território Kotiria, e os Baniwa, que vivem em comunidades no rio Aiarí ligadas a comunidades Kotiria por trilhas terrestres relativamente curtas (ver Koch-Grünberg 1995a:167-176, e Neves 1988:116-117). Subindo o Uaupés, entra-se no território dos Kubeo (TO), que historicamente também mantêm relações próximas com os Baniwa (ver Goldman 1963; Koch-Grünberg 1995b:68; Gomez-Imbert 1996:446; Wright 2005:11). É importante frisar que o *sib* Kubeo com o qual os Kotiria mantêm relações de casamento, os *Yurémara*, é também o *sib* que exibe a maior influência linguística AR; são descendentes de falantes da língua AR *Inkacha* que migraram para o território Kubeo no rio Querarí e lá assimilararam à população local<sup>7</sup>.

<sup>6</sup> Segundo estudos recentes do Instituto Socioambiental, 140 Kotiria (cerca de 9% do total) residem na comunidade de Iauaretê (Andrello, et. al. 2002) e 101 Kotiria (6% do total) vivem na cidade de São Gabriel da Cachoeira (FOIRN & ISA 2005:20). Não há estatísticas sobre o número de Kotiria que residem em Mitú.

<sup>7</sup> Informação fornecida por Simón Valencia López, linguista Kubeo e membro desse *sib*, em comunicação pessoal. Ver também (Valencia López 1994).

Vemos, então, que no contexto social Kotiria há uma notável influência Aruák. A história das relações entre os grupos TO Kotiria e Kubeo e os grupos AR Baniwa e Tariana se reflete tanto em evidência física quanto em tradições orais de vários tipos. Por exemplo, tanto Koch-Grünberg (1995b:63), no início do século vinte, como a expedição de Rondon de 1928<sup>8</sup> documentaram cerimônias Kotiria/Kubeo e Kotiria/Baniwa que indicam relações antigas entre os grupos e o compartilhamento de conhecimentos tradicionais, como a confecção de máscaras cerimoniais feitas de *tururi* tanto pelos Kubeo como pelos Kotiria<sup>9</sup>. Os Kubeo também figuram na literatura oral Kotiria, por exemplo, no mito que reconta que o nome *Kotiria*, “povo da água” foi dado pelos Kubeo (Stenzel 2004:22). Várias narrativas Baniwa identificam os Kotiria como aliados em guerras travadas contra grupos Kubeo e Sulia (um grupo do Papuri, agora extinto) no século dezenove (Wright 2005:89), e uma série de histórias Tariana fala de guerras, e mais tarde, alianças defensivas com os vizinhos Kotiria (Amorim 1987; Neves 2001).

De acordo com as normas do sistema Uaupés, até hoje, as alianças estabelecidas entre os Kotiria e os Kubeo, Tariana e Baniwa são refletidos na prática de casamento. Atualmente, cerca de 50% dos homens Kotiria são casados com mulheres desses três grupos<sup>10</sup>, e podemos supor que as três línguas figuram historicamente no “repertório” de línguas mais faladas pelas esposas nas comunidades Kotiria. De fato, um pequeno

---

<sup>8</sup> Num filme intitulado “No Rio Içana,” da coleção da Comissão Rondon no Museu do Índio, Rio de Janeiro.

<sup>9</sup> A produção especializada e troca de bens entre grupos linguísticos é um traço marcante da sociedade do Uaupés (Ribeiro 1995: 63-71). Por exemplo, os Baniwa tradicionalmente produzem raladores de mandioca, os Tukano fazem bancos de madeira, os Tuyuka fazem canoas, os Desana e Bará cultivam e preparam tinta corporal *caraiuru*, etc. O fato de tanto os Kubeo como os Kotiria produzirem máscaras cerimoniais feitas com a camada interna da casca da árvore *tururi* (*ficus brosimum*) indica relações extremamente próximas entre esse grupos.

<sup>10</sup> Dados sobre casamentos em 12 comunidades Kotiria foram coletados por Kristine Stenzel e Lucia Alberta Andrade de Oliveira, do Instituto Socioambiental, em setembro de 2004. Dados referentes aos homens Kotiria que residem em Iauaretê são de Andrello et al. (2002). Não há informações sobre as comunidades no Uaupés colombiano.

censo linguístico feito com 40 homens Kotiria mostra que o Baniwa e o Kubeo ainda figuram entre as línguas mais conhecidas, as outras são o Tukano, o Desana, e o Tuyuka (Stenzel 2005a:20-21).

O mesmo censo mostra que apesar do grande número de casamentos entre homens Kotiria e mulheres Tariana (mais de 25% do total), a língua Tariana não figura mais no “repertório” linguístico dos Kotiria, evidenciando o fato de que quase todos os Tariana falam Tukano atualmente. Não sabemos exatamente quando esse processo de mudança linguística começou, mas já no início do século vinte, Koch-Grünberg (1995b:23, 28), em viagem pela região, nota o declínio do uso de Tariana em Iauaretê e nas comunidades do médio Uaupés. Uma hipótese bastante plausível é que o processo teve seu começo após as invasões devastadoras das expedições escravagistas no meio do século dezoito, que resultaram numa redução drástica da população Tariana. Não sabemos com certeza, mas se as datas estimadas da migração Tariana ao Uaupés forem corretas, podemos presumir que houve contato linguístico Kotiria-Tariana durante 400-500 anos.

Constatamos assim que os Kotiria mantêm contato prolongado não só com dois grupos AR mas também com o sub-grupo dos Kubeo mais influenciado por contato com os Baniwa. Podemos então esperar encontrar evidências de influência AR maior na língua Kotiria do que em outras línguas TO da região, como Wa'ikhana e Tukano, cujo contato com línguas AR se restringe principalmente ao Tariana. Analisaremos agora dados comparativos para investigar quais inovações da língua Kotiria possam ser atribuídas a esse contato prolongado com falantes de línguas Aruák e considerar a hipótese de que a língua Kotiria tem sido mais influenciada do que outras línguas TO da região.

## 2 Inovações fonológicas

TABELA 1 - O inventário de consoantes em Kotiria

	LABIAL	CORONAL		VELAR	LARINGAL
OBSTRUÍNTES	b	+ant	-ant	g	
APPROX.	p	d		k	
	$p^h$	$t^h$		$k^h$	
		$s$	$tʃ$		
	w		$[r][l]j [dʒ]$		h

### 2.1 Oclusivas pós-aspiradas

O Kotiria tem o maior inventário fonêmico das línguas TO, com seis vogais e catorze consoantes, inclusive uma série de oclusivas surdas pós-aspiradas que ocorrem somente em posição inicial de palavra. Sem ser no Kotiria, oclusivas surdas pós-aspiradas ocorrem como fonemas apenas na língua Koreguaje, do ramo Tukano Ocidental (Wheeler 1992:40), enquanto no ramo Oriental, ocorrem foneticamente em Tukano e Wa'ikhana como resultado de processos de redução em fala rápida: raízes do tipo CVhV, como vemos em (1) são realizadas como [CHVV] (Ramirez 1997:39-40; Waltz 2002:161-4).

Algumas palavras em Kotiria são lexicalizações de reduções desse tipo, mas não há evidências que nos levem a supor que esse processo deu origem ao desenvolvimento completo de uma série contrastiva.

(1) <sup>1</sup>	TUKANO	WA'IKHANA	KOTIRIA
a.	<i>tobá</i>	<i>tobá</i>	<i>thóá</i> ‘voltar’
b.	<i>~páhí</i>	<i>pahí</i>	<i>phií</i> ‘ser grande’
c.	<i>di'píhi</i>	<i>~píhi</i>	<i>-~phi</i> classificador nominal de objetos pontiagudos

Oclusivas aspiradas ocorrem em 49% das raízes Kotiria e há dezenas de pares mínimos do tipo *ka* ‘macaco’ — *kha* ‘gavião’ que demonstram a natureza contrastiva das séries surdas aspiradas e não-aspiradas (ver também Waltz 2002:198-199). De fato, as oclusivas aspiradas constituem uma inovação notável, não encontrada nem mesmo na língua Wa'ikhana, a mais próxima do Kotiria dentro da família (ver anexo 1). Nos dados de Wa'ikhana, oclusivas aspiradas ocorrem em apenas 16% das raízes, e, como já foi mencionado, não têm estatus contrastivo (ver também Klump & Klump 1973:110; Waltz 2002:161). Se foi de fato o contato com línguas AR que estimulou o desenvolvimento de consoantes pós-aspiradas, como sugere Waltz (2002:162), não deveremos estranhar o fato de encontrá-las em maior escala em Kotiria do que em Wa'ikhana. Embora os Wa'ikhana também casem tradicionalmente com os Tariana, este seria o seu único grupo cunhado AR, os outros grupos preferidos sendo os TO Tukano e Desana.

Ao contrário das línguas TO, consoantes aspiradas são comuns nas línguas AR da região. Aikhenvald (2002:36, 297) inclui consoantes aspiradas surdas e sonoras (a coronal /dh/, três nasais /mh, mh, ñh/, e a semi-vogal /wh/) nos inventários fonológicos de ambas as línguas, Tariana e Baniwa (nesta última, inclui apenas as nasais e a semi-vogal), embora afirme na sua gramática de Tariana (Aikhenvald 2003:31, 46) que a maioria das ocorrências de consoantes sonoras aspiradas decorrem de processos de metátese. Ramirez, por seu lado, analisa todas as consoantes aspiradas do Baniwa — sonoras e surdas — como “clusters” que resultam de metátese com /h/ (2001:45). Sendo o estatus das consoantes aspiradas fonêmico ou fonético, os estudos

de Aikhenvald e Ramirez nos mostram que consoantes desse tipo são um traço saliente de ambas as línguas. Waltz (2002:162-166) sugere que o contato com essas línguas há muito tempo estimulou uma aspiração secundária em Kotiria, e que esse traço se disseminou por analogia a todas as oclusivas surdas em posição inicial. Mais tarde, um processo de perda de sílabas na margem esquerda da raiz (ver Seção 2.5) acabou criando os contrastes sincrônicos entre oclusivas surdas aspiradas e não-aspiradas em início de palavra.

## 2.2 A africada /tʃ/ e a variante alofônica [dʒ]

O Kotiria foi a única língua TO que desenvolveu um conjunto de quatro consoantes coronais [+anterior]—as oclusivas sonora, surda e surda pós-aspirada /d, t, tʰ/ e a fricativa /s/—além de um par contrastivo [-anterior]: a africada /tʃ/ e a aproximante /j/ (2), cujo alofone [dʒ] ocorre antes das vogais posteriores /a, o, ȫ, u/ (3).

- (2) [-anterior]
- |    |            |          |                |    |            |        |                       |
|----|------------|----------|----------------|----|------------|--------|-----------------------|
| a. | <i>da</i>  | [dáá]    | ‘ser pequeno’  | e. | <i>cha</i> | [tʃáá] | ‘comida<br>preparada’ |
| b. | <i>ta</i>  | [táa]    | ‘vir’          | f. | <i>ya</i>  | [jáa]  | ‘enterrar’            |
| c. | <i>tba</i> | [tʰáá]   | ‘grama’        |    |            |        |                       |
| d. | <i>~sa</i> | [sã□ ȫ□] | ‘estar dentro’ |    |            |        |                       |
- (3) a. *yoa* [j~dʒoá] ‘fazer’ d. *p̪hayt̪* [p̪ʰaj~dʒá] ‘muito(s)’  
 b. *yt̪k̪t̪* [j~dʒt̪ʰk̪t̪] ‘árvore’ e. *yai* [j~dʒáí] ‘onça’  
 c. *yuka* [j~dʒuʰká] ‘urubu’  
 cf. f. *yese* [jeʰsé] ‘porco’

Waltz (2002:168-169) analisa a africada /tʃ/ do Kotiria como tendo quatro fontes na proto-língua—\*s (em início de palavra), \*k, \*g e \*y (no interior da palavra)—e caracteriza o seu desenvolvimento como palatalização condicionada por uma vogal anterior /i, e/ adjacente. Os exemplos em (4) de fato sugerem uma correspondência de TO /k/ e Kotiria /tʃ/ no interior da palavra nesse ambiente condicionador.

(4)	BARASANA	TUKANO	WA'IKHANA	KOTIRIA
	<i>wekí</i>	<i>wekí</i>	<i>wekí</i>	<i>wach<sup>h</sup>tʃ</i> [wa <sup>h</sup> tʃú]
	<i>~yikí</i>	<i>~yekí</i>	<i>~yekí</i>	<i>~y<sup>h</sup>tʃ<sup>h</sup>ekí</i> [n <sup>h</sup> u <sup>h</sup> tʃ <sup>h</sup> ú]
	<i>riká</i>	<i>diká</i>	<i>d<sup>h</sup>uká</i>	<i>dicha</i> [di <sup>h</sup> tʃá]
	<i>~ribí</i>	<i>im<sup>h</sup>kóbó</i>	<i>dekó</i>	<i>dachó</i> [da <sup>h</sup> tʃó]

Embora processos de palatalização desse tipo, condicionados por qualidades de uma vogal contígua, sejam fonologicamente comuns, é curioso que somente duas línguas TO—Kotiria e Kubeo (Morse & Maxwell 1999:3)—pareçam ter sido sensíveis o suficiente para desenvolver um fonema /tʃ/. Waltz (2002:170) sugere, mas não aprofunda, a hipótese de que o contato com línguas AR tenha sido um dos fatores do desenvolvimento de /tʃ/ em Kotiria, um cenário que nos parece bastante provável tanto para essa língua quanto para o Kubeo, já que ambas as línguas AR da região têm o fonema /tʃ/ ou /ts/ (Ramirez 2001:43; Aikhenvald 2002:55), herdada genéticamente do proto-AR (Aikhenvald 2002:294). A africada sonora /dʒ/ também ocorre em Tariana e Baniwa; é analisada como fonema que se originou de \*/j/ em Baniwa (Ramirez 2001:43, 54) e como alofone de /j/ em Tariana (Aikhenvald 2002:30). O [dʒ] alofone de /j/ em Kotiria e Kubeo espelha essas ocorrências e reforça a hipótese de uma influência areal AR sobre essas duas línguas TO.

## 2.3 A variação [r] ~ [l]

Em Kotiria, a oclusiva /d/ e o seu alofone, o flap [ɾ] ocorrem em distribuição complementar, o primeiro em início de raízes, e o segundo somente no início de sufixos, e.g. *die-ro* ‘cachorro-sg’ ou no interior de raízes, e.g. o demonstrativo *a’ri*. Segundo Gomez-Imbert (no prelo), uma distribuição parecida tende a ocorrer em outras línguas da sub-região do Uaupés: em Desana, Tukano, Siriano e Wa’ikhana, /d/ ocorre em início de palavra e varia com [ɾ] internamente, e.g. Wa’ikhana [die-do]~[die-ro]. Sabendo que processos de lenição em posições internas de palavra, nesse caso “flapping”, são relativamente comuns (Kenstowicz 1994:35), não encontramos evidência de que a variação [d~r] nesse sub-grupo de línguas TO tenha sua origem em contato.

Contudo, tanto em Kotiria como em Wa’ikhana, há um segundo processo de variação alofônica envolvendo o flap que provavelmente decorre de contato: nesse processo, morfemas em que [ɾ] precede uma vogal anterior /i, e/ são freqüentemente pronunciados com um alofone [l]. Assim, a pronúncia do demonstrativo *a’ri* varia entre [a’ri] e [a’li] e sufixos como o nominalizador *-ri* ou o marcador de objeto *-re*, podem ser realizados como [li] ou [le], i.e. *~dabo-ti-ri-ro* ‘casado’, literalmente ‘um que tem esposa’, realizado como [nãm̩̩tíiro] e *wese-p̩-re* ‘a/para a roça’ realizado como [we<sup>h</sup>sép̩le].

O contato com línguas AR é a fonte mais provável do desenvolvimento dessa lateral em Kotiria e Wa’ikhana (mas não em Tukano), já que as duas línguas, Tariana e Baniwa, têm flap e lateral cujas ocorrências são condicionadas pela qualidade da vogal adjacente. Em Tariana, o flap e a lateral ocorrem em variação livre no interior da palavra antes de vogais anteriores /i, e/ e.g. *-liphe*~*-riphe* ‘segurar com firmeza’ (Aikhenvald 2002:39-40). Em Baniwa, as líquidas também ocorrem mais freqüentemente em posição interna, com a lateral ocorrendo antes de /i/ e.g. *máali* ‘garça’ (Ramirez 2001:53).

A tendência do alofone [l] ocorrer antes de vogais anteriores em Kotiria e Wa’ikhana espelha em grande parte a sua ocorrência nessas

duas línguas AR; ou seja, é possível que a lateral nessas duas línguas TO tenha se desenvolvido de modo a refletir as condições da sua ocorrência em línguas AR. É interessante, porém, a afirmação de Aikhenvald de que a neutralização dos fonemas flap/lateral em posição interna de palavra em Tariana é “indicativa de uma tendência que leva à redução do sistema, de modo que fique mais parecido com os sistemas [fonológicos] TO” (2002:40, tradução nossa). Embora pareça que estamos diante de hipóteses conflitantes quanto à direcionalidade de mudanças decorrentes de contato AR-TO, de fato, podemos simplesmente estar lidando com dois processos distintos que ocorreram em momentos históricos diferentes. O desenvolvimento do alofone [l] nas línguas TO pode muito bem ter acontecido durante o período prolongado de contato com o Baniwa e (um ainda-robusto) Tariana. A redução do sistema em Tariana sob influência de Tukano, que nunca desenvolveu o alofone [l], é certamente um processo muito mais recente.

#### 2.4 A perda de V/σ em posição inicial de palavra

Uma característica notável das línguas TO da sub-região do Uaupés é a gradual elisão de material fonológico na margem esquerda das raízes, normalmente os *onsets* das sílabas iniciais com tom baixo. Esse processo é muito pronunciado em Kotiria, onde ocorreu a ampla elisão de sílabas inteiras, como vemos nos dados em (5) (ver Waltz 2002:176-177 para uma análise das restrições ativas no processo de elisão).

(5)	SIRIANO	DESANA	TATUYO	TUKANO	WA'IKHANA	KOTIRIA	
a.		<i>dekó</i>	<i>óko</i>	<i>akó</i>	<i>akó</i>	<i>kóo</i>	‘agua’
b.	<i>guikáru</i>	<i>g̊íkeri</i>	<i>opí</i>	<i>upí</i>	<i>upí</i>	<i>píi</i>	‘dente’
c.	<i>ítá(je)</i>	<i>ítá(je)</i>	<i>~ítá</i>	<i>~ítá</i>	<i>~ítá</i>	<i>~táá</i>	‘pedra’
d.	<i>aariri</i>	<i>ariri</i>	<i>atí</i>	<i>a'í</i>	<i>a'í</i>	<i>táa</i>	‘vir’
e.		<i>api</i>	<i>~ópe</i>	<i>~ó'pe</i>	<i>~o'pé</i>	<i>~péé</i>	‘seio’

Ainda não temos explicação sólida da origem dessa elisão em Kotiria (processo que também ocorre em palavras compostas por duas raízes em Tukano, segundo Ramirez 1997:48), mas encontramos um processo paralelo em Baniwa que sugere uma origem decorrente de contato. Em Baniwa, vogais simples e vogais longas são contrastivas. Ambas podem ocorrer em posição inicial de palavra, sendo que nessa posição as vogais simples são muito mais freqüentes. Ramirez (2001:85-86) afirma que vogais simples em posição inicial freqüentemente caem na fala rápida, e há muitos casos de raízes nas quais essa elisão tem sido completamente gramaticalizada. É possível que a elisão de sílabas V (muito mais comuns em línguas TO) em Kotiria tenha sido induzida por esse fenômeno do Baniwa, generalizando-se em seguida o processo na língua. É curioso que não encontramos evidências de uma ocorrência paralela em Kubeo, mas isso pode ser devido ao fato de que, de modo geral, houve menos elisão de consoantes iniciais em Kubeo do que em outras línguas TO. Como resultado, Kubeo ficou com um número menor de palavras iniciadas por V que poderiam ser alvos de elisão desse tipo. Precisamos de análises comparativas mais detalhadas para confirmar ou não essas hipóteses.

É mais uma vez interessante notar que, como no caso na neutralização r/l, Aikhenvald (2002:49) atribui a elisão de sílabas não-acentuadas no dialeto Tariana de Periquitos à influência de falantes de Kotiria (seus cunhados preferidos). Podemos de novo estar diante de uma corrente de mudanças induzidas por contato: Baniwa → Kotiria → Tariana.

### 3 Inovações morfológicas

Além das inovações fonológicas acima apresentadas, encontramos evidências que sugerem que há algumas inovações morfo-sintáticas em Kotiria que também poderiam ser analisadas como resultado de contato com línguas AR. Nesta seção, investigaremos três dessas inovações: o desenvolvimento de um paradigma de clíticos possessivos (derivados

de pronomes pessoais) em Kotiria e algumas outras línguas TO, o uso inusitado de classificadores de forma com nomes animados em Kubeo e Kotiria, e a ocorrência de um negador pré-verbal em Tukano e Kotiria.

### 3.1 O desenvolvimento de proclíticos possessivos

Nas línguas TO, as relações parte/todo são geralmente indicadas por simples justaposição de nomes, com o nome determinador precedendo o nome dependente.

(6)	KOTIRIA	TUKANO (Ramirez 1997:327)	TATUYO (Gomez-Imbert 1981:111)
	<i>y<sup>1</sup>k<sup>2</sup></i> ~ <i>phu-ri</i>	<i>wék<sup>1</sup></i> ~ <i>opekó</i>	<i>yai</i> <i>r<sup>1</sup>po-a</i>
	árvore folha-PL	vaca leite (lit:	onça cabeça-
	‘folhas de árvore’	líquido do seio)	CLS:redondo
	<i>wa'i-kiro</i> <i>di'i</i>	‘leite de vaca’	‘cabeça da onça’
	animal-SG carne		
	‘carne de animal’		

Em muitas, porém não todas, as línguas TO, encontramos determinantes pronominais na mesma construção: a forma completa do pronome pessoal antecede o nome para indicar uma relação de posse inalienável, como no caso de relações de parentesco.

(7)	TUKANO (Ramirez 1997:326)	WA'IKHANA	DESANA (Miller 1999:49)
	<i>y<sup>1</sup>í<sup>2</sup> pako</i>	~ <i>b<sup>1</sup>u<sup>2</sup></i> ~ <i>dabo-do</i>	<i>igo</i> <i>pago</i>
	1SG mãe	2SG esposa-SG	3SG mãe
	‘minha mãe’	‘tua esposa’	‘mãe dela’

No entanto, em algumas línguas TO, inclusive Kotiria e Kubeo, a posse pronominal é marcada por proclíticos, criados (fonologicamente e morfologicamente) por redução das formas completas independentes. A Tabela 2 mostra o paradigma em Kotiria.

TABELA 2 - Pronomes e proclíticos possessivos em Kotiria

		SINGULAR	PLURAL	
1a	proclítico	<i>y<sup>u</sup> [j<sup>u</sup>]</i>	<i>~bari</i> [mãri] (INCL)	<i>~sa</i> [sã] (EXCL)
	pronome completo	<i>y<sup>u</sup> u [y<sup>u</sup>?<sup>u</sup>]<sup>0</sup></i>	<i>~bari</i> [mãri <sup>00</sup> ]	<i>~sa</i> [sã]
2a	proclítico	<i>~b<sup>u</sup> [m<sup>u</sup>]</i>	<i>~b<sup>u</sup>sa</i> [m <sup>u</sup> <sup>h</sup> sã]	
	pronome completo	<i>~b<sup>u</sup> u<sup>0</sup> [m<sup>u</sup>u<sup>0</sup><sup>u</sup>]</i>	<i>~b<sup>u</sup>sa</i> [m <sup>u</sup> <sup>h</sup> sã]	
3a	proclítico	<i>to</i> [to]	<i>ti</i> [tí]	
	pronome completo	<i>tiro</i> [tíró] (MASC) <i>tikoro</i> [tíkóró] (FEM)	<i>ti-~da</i> [tí <sup>0</sup> nã]	

Notamos que para as categorias de primeira e segunda pessoa plural, as formas dos proclíticos não são reduzidas morfologicamente, como é o caso das formas singulares *y<sup>u</sup>* e *~b<sup>u</sup>*, bem como as formas da terceira pessoa *to* e *ti*. Há, no entanto, redução fonológica de todos os proclíticos: ocorrem com tom baixo e formam uma só palavra fonológica com o nome “possuído”, como vemos em (8a). Em comparação, pronomes independentes sempre têm pelo menos um tom alto. No mais, e em outros tipos de construções com nomes justapostos, como vemos em (8b), cada palavra fonológica contém um tom alto e há pausas entre palavras.

- (8) *to* *~pho'da* [top<sup>h</sup>õnã<sup>0</sup>]  
 a. 3SG.POSS filhos  
 ‘filhos dele/dela’  
*ti* *wa'i* *~waha-~ida* [tiwa<sup>0</sup>í wâhâ<sup>0</sup> ñã]  
 b. 3PL.POSS peixe matar-NOM:PL  
 ‘os peixes (que foram pescados/mortos) deles’

O desenvolvimento dessas formas prefixadas/proclítizadas é bastante surpreendente, já que tipologicamente, as línguas TO são quase que inteiramente “sufixais”. A hipótese de que a fonte dessa inovação tenha sido contato com línguas AR foi sugerido por Aikhenvald (2002:71) e há, de fato, paralelos óbvios entre essas formas e estruturas encontradas em todas as línguas AR da região, em que a posse inalienável é marcada por prefixos possessivos no nome possuído e os mesmos prefixos ocorrem como marcadores de pessoa em VPs (Aikhenvald 2002:62, 79; Ramirez 2001:105).

Em Kotiria, na verdade, observamos um pouco de flutuação quanto ao uso das formas completas e proclítizadas, principalmente em construções com as formas de primeira e segunda pessoa singular: às vezes, são usadas as formas completas — *y<sup>tt</sup>' tt* e *~b<sup>tt</sup>' tt* (ver exemplos em Stenzel 2004:194-195). No entanto, os falantes nativos são unâimes em afirmar que somente as formas reduzidas são “corretas” em construções possessivas. Ainda não sabemos como interpretar essa variação. Por um lado, poderia indicar que a mudança das formas completas para formas reduzidas ainda se encontra em vias de gramaticalização. Por outro lado, é possível que, como Aikhenvald afirma, o contato sem dominação resulte num enriquecimento de recursos linguísticos. No caso do Kotiria, houve gramaticalização das formas proclítizadas como o meio não-marcado de indicar posse, enquanto as formas completas foram mantidas e reinterpretadas como tendo uma função marcada—a de indicar foco no nome possuído em discurso. Aikhenvald afirma ainda que os pronomes podem aparecer em próclise ao verbo quando o sujeito não está em foco (2002:71), porém não encontramos exemplos desse fenômeno em nossos dados. No entanto, se a nossa análise dos dados do Kotiria de fato vier a revelar esse tipo de especialização funcional, isso seria uma evidência positiva de que a difusão areal das línguas AR pode ter acrescentado “dimensões adicionais à complexidade estrutural” das línguas TO em contato com línguas AR (Aikhenvald 2002:81, tradução nossa).

A hipótese de que o contato com línguas AR resultou no desenvolvimento desses proclíticos é reforçada por evidências encontradas em outras línguas TO. Em Kubeo, por exemplo, encontramos dois paradigmas interessantes: um conjunto de “pronomes possessivos” derivados das formas completas dos pronomes pessoais, e um conjunto de “adjetivos possessivos” proclitizados, usados na expressão de posse inalienável, como no caso das relações de parentesco (Morse & Maxwell 1999:81-82, 125-6). Formas proclitizadas de posse pronominal também se encontram em Tatuyo (Gomez-Imbert 1982:249), e em Desana, uma forma pronominal procliticizada pode ocorrer como argumento de um pós-posicional (Aikhenvald 2002:81).

(9)	KUBEÓ (Morse and Maxwell 1999:126)	TATUYO (Gomez-Imbert 2000:337)	DESANA
	<i>xi-pako</i> 1SG.POSS-mãe 'minha mãe'	<i>yɪ-pàkɪ</i> 1SG-pai 'meu pai'	<i>yɪ-~bera</i> 1SG-ACCOMP 'comigo'

As inovações paralelas mais marcantes são encontradas na língua TO Retuarã, que sofreu “influência substancial da língua Aruák vizinha [Yukuna]” (Strom 1992:72, tradução nossa). O Retuarã desenvolveu não só prefixos de “concordância” que indicam o possuidor em construções genitivas, mas também são usados como marcadores de sujeito nos VPs (em paralelo direto às estruturas AR). Como nas outras línguas TO em que encontramos marcadores parecidos, uma comparação dos paradigmas revela um processo de cliticização das formas pronominais completas (Strom 1992:35).

(10)	RETUARÃ	a. pronomé 3MS	b. possessivo (Strom 1992:126)	c. marcador de sujeito (Strom 1992:109)
		<i>i'ki</i>	<i>ki-wi'ia</i> 3MS-casa casa dele'	<i>i'sia-ha</i> aquilo- ADVZR Ele só come aquilo.'

### 3.2 Uso de classificadores de forma com animados

Tanto as línguas TO como as línguas AR têm sistemas complexos de classificação nominal. Nas línguas AR, a distinção maior indica uma dicotomia entre humanos e não humanos (Aikhenvald 2002:88-91) enquanto nas línguas TO a dicotomia maior cria as categorias básicas de animados e inanimados. Os nomes animados nas línguas TO são marcados quanto ao número por morfemas específicos, e, no caso de nomes animados referentes a humanos, há também especificação de gênero. Opcionalmente, animais como cachorros, antas ou onças, que são altamente salientes ou importantes para os humanos, também podem levar marcadores de gênero. Um sub-conjunto de nomes animados indicam “coletivos” de entes como abelhas, cupins e peixinhos; esses nomes são inherentemente plurais e é necessário acrescentar um morfema específico para indicar uma entidade individual.

A marcação morfológica de inanimados em línguas TO é bem mais complexa. O “plural” é geralmente marcado por um único morfema *-(V)ri* (Gomez-Imbert no prelo), enquanto o “singular” é marcado por morfemas classificadores, na sua maioria sufixos (mas também podem ser raízes nominais) indicando a forma geral (e.g. Kotoria *-ka* “redondo”, *-paro* “curvado” e *-da* “filiforme”, entre outros), a distribuição dos membros individuais de um conjunto (i.e. Kotoria *-thu* “empilhados”), ou tipo (i.e. Kotiria *~phu* “folha de X”, *~yo* “palmeira de X”, etc.). Esses mesmos classificadores ocorrem em modificadores como números, demonstrativos, e partículas anafóricas. Nomes genéricos e abstratos, bem como nomes que referem a objetos sem forma pré-definida levam um marcador  $\emptyset$ . Modificadores de nomes genéricos levam um classificador genérico *-ye/-e* na maioria das línguas TO, sendo que esse marcador não ocorre em Kotiria.

Em contraste com o sistema básico das línguas TO, um dos traços salientes da classificação nominal em Baniwa e outras línguas AR é o emprego de classificadores de forma tanto com nomes animados como inanimados, um traço que se disseminou a pelo menos duas línguas TO.

Gomez-Imbert analisa o desenvolvimento do uso de classificadores de forma e de morfemas de gênero com nomes animados em Kubeo, um “padrão classificatório importado para a língua por ancestrais falantes de [uma língua] Aruák” (1996:464).

Há evidências de difusão, porém num grau menos extenso, do uso de classificadores de forma com animados em Kotiria também, provocando um “ajuste” semântico das categorias existentes. Nos exemplos em (11), vemos que sincronicamente, os classificadores de forma são usados para se referir tanto a inanimados como a animados em Kotiria.

(11)

CLASSIFICADOR	OCORRÊNCIA	OCORRÊNCIA COM NOME ANIMADO
	COM NOME	
	INANIMADO	
a.		
-da	<i>yo'ga-ri-da</i> pescar-NOM- CLS:filiforme	<i>phi-ri-da</i> ser.grande-NOM-CLS:filiforme 'um ser comprido como corda' (na descrição de uma cobra)
	'filiforme'	'linha de pesca'
b.		
-ka	<i>~ta-ka</i> 'pedra-' CLS:redondo	<i>~su'i-ka</i> caramujo-CLS:redondo
	'redondo'	<i>~p<sup>u</sup>-ka</i>
		<i>siri</i> -CLS:redondo
		<i>khasipo-ka</i> barata da mata-CLS:redondo
		<i>wapa-ka</i> minhoca cascuda- CLS:redondo
c.		
-paro	<i>ho-paro</i>	<i>w<sup>u</sup>po-paro</i>
	'curvado'	lagarto-CLS:curvado
	banana- CLS:curvado	

Gomez-Imbert afirma que o uso de classificadores de forma com certos tipos de fauna (animados) decorre do fato de esses animados não serem “individualizados” por nenhum outro traço saliente. Lembramos que a hierarquização das categorias de animados em línguas TO se revela na marcação morfológica. Nomes das categorias mais altas levam mais marcação (i.e. distinções de gênero e número em nomes com referentes humanos), enquanto nomes que referem a entidades das categorias inferiores não levam especificação morfológica alguma; de fato, às vezes nem são classificados como entidades singulares (o caso dos “coletivos”). A adoção do uso “aruakiano” de classificadores de forma com essas categorias inferiores facilita a identificação dessas entidades no discurso, e enriquece os recursos linguísticos das línguas TO sem minar o sistema de classificação original. Nas palavras de Gomez-Imbert, havia nas línguas TO “uma categorização pré-existente das entidades inanimadas baseada na forma [que] simplesmente se estendeu a entidades animadas” (1996:464, tradução nossa).

### 3.3 O desenvolvimento do negador pré-verbal *~de(e)* ’

A negação nas línguas TO é canonicamente marcada no verbo por um morfema que ocorre depois da raiz e antes dos sufixos finais de modalidade oracional (i.e. evidenciais, interrogativos, imperativos, etc.). Em Kotiria, o morfema negativo é *–era*, como em *~basi-era-ha* ‘Eu não sei’. No entanto, há uma segunda construção negativa bastante frequente em Kotiria, na qual ocorre uma partícula negativa pré-verbal *~de* [ne□], como vemos nos exemplos em (12). A presença dessa partícula funciona para enfatizar a proposição negada, criando uma espécie de negação “absoluta”.

- (12) *~de to-p~~ti~~-re ti-p~~ti~~-re ~tidi-era-ha andar/*  
 NEG REM-LOC-OBJ ANPH-LOC-OBJ caçar-NEG-VIS. IMPERF. 1
- a. *Nunca vou caçar naquele lugar.*  
*~de ~bari pb~~ti~~k-~~ti~~ hi-era-ra*  
 NEG 1PL.INC(POSS) parente-MASC COP-NEG- VIS. IMPERF.  
 NON. 1
- b. *'Esse absolutamente não é nosso pai.'*  
*ti-re y~~ti~~'~~ti~~ ~bi-p~~ti~~-re ~de bo-era-~sidi-ka*  
 ANPH-OBJ 1SG agora-LOC- NEG esquecer-NEG-fazer.ainda-  
 TMP ASSERT:IMPERF
- c. *Jamais vou esquecê-las (as danças tradicionais dos Kotiria).'*

Além do Kotiria, na sub-região do Uaupés, construções paralelas ocorrem somente nas línguas Tukano e Tariana, mas, significativamente, não em Baniwa.

- (13) *TUKANO TARIANA*  
 (Ramirez 1997:154) (Aikhenvald 2002:135)
- néé ~~ta~~-tí-sa' Ne ma-na-kade-mha*  
 NEG querer-NEG-PRES.NON. NEG NEG-querer-NEG-PRES.NON.  
 VIS-NON.3 VIS  
 'Não quero nada.' 'Não quero nada.'

Aikhenvald (2002:136) afirma que o Tariana adquiriu a partícula negativa pré-verbal por difusão de Tukano, e que essa “dupla negação” foi reanalisada como uma construção enfática. Ela admite, no entanto, que as línguas AR geralmente têm marcadores negativos com segmento nasal e que esse pode ser um caso de “acomodação gramatical”. De fato, encontramos em Baniwa uma construção similar com a partícula negativa *ñame*, que Ramirez (2001:196-7) analisa como uma raiz não-flexionada ‘negar’ que pede como complemento uma oração sem outra marcação de negação.

(14)

**BANIWA****(Ramirez 2001:196)**

*Ñam Peduru i-iñaha-ka*  
 NEG Pedro CON-comer-SUB  
 'Pedro não comeu.'

Acreditamos que o desenvolvimento da construção com “dupla negação” em Kotiria e Tukano tem origem em estruturas AR desse tipo, que passaram por reanálise e ganharam um sentido requintado de negação enfática. Que tenha havido reanálise desse tipo não é de se estranhar, pois o padrão TO de negação morfológico no verbo ainda permanece. Ou seja, não houve equação com a estrutura AR (sem marcador negativo no verbo), e sim, o acréscimo de uma partícula pré-verbal, criando uma segunda opção de negação. É bem possível que esse seja mais um exemplo de uma “corrente de difusões”: uma construção AR encontrada em Baniwa (e provavelmente em Tariana antigo) se difundiu para duas línguas TO em contato. A construção adotada passou por reanálise nessas línguas e foi com forma e sentido reanalizados que a estrutura se difundiu mais recentemente da língua Tukano de volta para Tariana (junto com a morfologia de negação no verbo, conforme o padrão TO).

## Conclusões

Reconhecemos que a investigação dos efeitos linguísticos provenientes de contato entre línguas AR e línguas TO ainda está no início. Carecemos de pesquisa descritiva básica de várias línguas da região, o que torna qualquer trabalho comparativo muito mais difícil. No entanto, podemos arriscar algumas conclusões de teor provisório e ainda passíveis de reformulação e/ou maior aprofundamento.

Em primeiro lugar, podemos afirmar que há inovações em Kotiria e outras línguas TO para as quais uma hipótese de origem por contato com línguas AR parece ser bastante provável. Além disso, entre as línguas TO da sub-região do Uaupés, podemos arguir que a língua Kotiria exibe um número maior de inovações dessa natureza, frutos de sua situação de contato maior com dois grupos AR e uma outra língua TO bastante “aruakizada”. Em segundo lugar, podemos dizer que vários dos elementos analisados reforçam a noção proposta por Aikhenvald, de que o contato sem dominação tende a resultar em enriquecimento de padrões. No entanto, também reconhecemos quão difícil é analisar situações de contato como as encontradas na bacia do Uaupés—precisamos ter cuidado redobrado na hora de fazer afirmações gerais, e precisamos sempre complementar as avaliações de direcionalidade de influência com uma visão histórica, para podermos considerar possíveis correntes de difusão.

## Referências

- AIKHENVALD, Alexandra Y. **Areal diffusion and language contact in the Içana**-Vaupés basin, North West Amazonia. In: Robert, M. W.; Dixon & Alexandra Y. Aikhenvald (Eds.). *The Amazonian Languages*, Cambridge, CUP, p. 385-415, 1999.
- \_\_\_\_\_. **Language Contact in Amazonia**. New York: Oxford University Press, 2002.
- \_\_\_\_\_. **A Grammar of Tariana, from Northwest Amazonia**. Cambridge: CUP, 2003.
- ANDRELLO, Geraldo; BUCHILLET, Dominique; AZEVEDO, Marta. **Levantamento Sócio-Econômico, Demográfico e Sanitário de Iauaretê**. Iauaretê/Centro: Instituto Socioambiental, 2002.

BRANDÃO DE AMORIM, Antonio. **Lendas em nheengatu e em português.** Manaus: Fundo Editorial - ACA, 1987.

CHERNELA, Janet M. **The Wanano Indians of the Brazilian Amazon:** A Sense of Space. Austin: University of Texas Press, 1993.

EPPS, Patience. **A Grammar of Hup**, University of Virginia - PhD (dissertation), 2005a.

\_\_\_\_\_. **Areal Diffusion and the Development of Evidentiality:** Evidence from Hup. *Studies in Language*, 2005b. 29, p. 617-649,

\_\_\_\_\_. **The Vaupés melting pot:** Tukanoan influence on Hup. In: ROBERT M. W.; DIXON & ALEXANDRA Y. Aikhenvald (Eds.). **Grammars in Contact. Oxford:** Oxford University Press, 2007. p. 267-289

FOIRN & ISA. **Levantamento socioeconômico, demográfico e sanitário da cidade de São Gabriel da Cachoeira.** São Gabriel da Cachoeira: FOIRN/ISA, 2006.

\_\_\_\_\_. **Povos indígenas do alto e médio Rio Negro:** uma introdução à diversidade cultural e ambiental do noroeste da Amazônia brasileira. Brasília: MEC/SEF, 2006.

GOLDMAN, Irving. **The Cubeo: Indians of the Northwest Amazon:** Illinois Studies in Anthropology Urbana: University of Illinois Press, 1963. (n. 2). [S.l.]

GOMEZ-IMBERT, Elsa; HUGH-JONES, Stephen. **Introducción al estudio de las lenguas del Piraparaná (Vaupés).** In: GONZÁLEZ DE PÉREZ, María Stella; RODRÍGUEZ DE MONTES, María Luisa (Eds.). **Lenguas Indígenas de Colombia: una Visión Descriptiva.** Santafé de Bogotá: Instituto Caro y Cuervo, 2000. p. 321-356.

- GOMEZ-IMBERT, Elsa. **La Expresión de la Posesión en Tatuyo.** In: **Revista del Instituto Colombiano de Antropología**, XXIII, 1981. p. 111-124.
- \_\_\_\_\_. De la forme et du sens dans la classification nominale en tatuyo (langue Tukano orientale d'Amazonie colombienne), Ecole Pratique des Hautes Etudes - IVe Section, Université Paris-Sorbonne, Doctorat de Troisième Cycle, 1982.
- \_\_\_\_\_. Force des langues vernaculaires en situation d'exogamie linguistique: le cas du Vaupés colombien, Nord-ouest amazonien, *Cahiers des Sciences Humaines*, 1991. 27, p. 535-559.
- \_\_\_\_\_. Problemas en torno a la comparación de las lenguas Tukano orientales, Estado actual de la clasificación de las lenguas Indígenas de Colombia, Santafé de Bogotá, Instituto Caro y Cuervo, 1993. p. 235-267.
- \_\_\_\_\_. **When Animals Become “Rounded” and “Feminine”:** Conceptual Categories and Linguistic Classification in a Multilingual Setting. In: GUMPERZ, J. J.; LEVINSON S. C. (Eds.). **Rethinking Linguistic Relativity**, Cambridge, CUP, 1996. p. 438-469.
- \_\_\_\_\_. **Variations tonales sur fond d'exogamie linguistique.** In: **Cahiers de Grammaire**, 1999. 24, p. 67-94.
- \_\_\_\_\_. **Famille Tukano.** In: **Dictionnaire des langues du monde**. Paris: Presses Universitaires de France. (No prelo.)
- \_\_\_\_\_. *Fonología de dos idiomas tukano del Piraparáná: barasana y tatuyo*, Ameríndia, 2005. 29/30, p. 43-80.
- GONZÁLEZ DE PÉREZ, María Stella; RODRÍGUEZ DE MONTES, María Luisa (Eds.). **Lenguas Indígenas de Colombia:** una Visión Descriptiva. Santafé de Bogotá: Instituto Caro y Cuervo, 2000.
- HEINE, Bernd; KUTEVA, Tania. **Language Contact and Grammatical Change.** Cambridge: CUP, 2005.

JACKSON, Jean E. **The Fish People.** Linguistic Exogamy and Tukanoan Identity in Northwest Amazonia. Cambridge: Cambridge University Press, 1983.

KENSTOWICZ, Michael. **Phonology in Generative Grammar.** Cambridge/MA.: Blackwell, 1994.

KLUMPP, James; KLUMPP, Delores. **Sistema Fonológico del Piratapuyo.** Sistemas Fonológicos de Idiomas Colombianos. Lomalinda: Instituto Linguístico de Verano/Editorial Townsend, 1973. p. 107-120.

KOCH-GRÜNBERG, Theodor. **Dos años entre los indios; viajes por el noroeste brasileño 1903/1905.** Santafé de Bogotá: Editorial Universidad Nacional, 1995a. (Publicação original: Berlin, 1909). (v. 1).  
\_\_\_\_\_. **Dos años entre los indios; viajes por el noroeste brasileño 1903/1905.** Santafé de Bogotá: Editorial Universidad Nacional, 1995a. (Publicação original: Berlin, 1909). (v. 2).

MEIRA, Sergio; GOMEZ-IMBERT, Elsa. **Review of:** Aikhenvald, A. Language Contact in Amazonia, Studies in Language, 2005. 29, p. 168-78.

MILLER, Marion. **Desano Grammar:** Studies in the Languages of Colombia 6. Arlington: Summer Institute of Linguistics/University of Texas, 1999.

MORSE, Nancy L.; MAXWELL, Michael B. **Cubeo Grammar:** Studies in the Languages of Colombia 5. Arlington: Summer Institute of Linguistics/University of Texas, 1999.

- NEVES, Eduardo. **Paths in Dark Waters**: Archaeology as Indigenous History in the upper Rio Negro Basin, PhD (dissertation) - Northwest Amazon, Indiana University, 1988.
- \_\_\_\_\_. **Indigenous historical trajectories in the upper Rio Negro basin**. Unknown Amazon. In: MCEWAN, Colin; BARRETO, Cristiana; NEVES, Eduardo (Eds.). **London**: The British Museum Press, 2001. p. 266-286.
- RAMIREZ, Henri. **A fala Tukano dos Ye'pâ-Masa**. Manaus: CEDEM, 1997. (Tomo I, Gramática)
- \_\_\_\_\_. **Línguas Arawak**. Manaus: Universidade do Amazonas, 2001.
- RIBEIRO, Berta G. **Os índios das águas pretas**: modo de produção e equipamento produtivo. São Paulo: Companhia das Letras - EdUSP, 1995.
- SORENSEN, Arthur P. Jr. **Multilingualism in the Northwest Amazon**. American Anthropologist, 1967. 69, p. 670-684.
- \_\_\_\_\_. **An Emerging Tukanoan Linguistic Regionality**: Policy Pressures. South American Indian Languages: Retrospect and Prospect, Harriet E. Manelis Klein & Louisa R. Stark (eds), 140-56. Austin: University of Texas Press, 1985.
- STENZEL, Kristine. **A Reference Grammar of Wanano**, PhD (dissertation) - University of Colorado, 2004.
- \_\_\_\_\_. **Multilingualism**: Northwest Amazonia Revisited. In: II Congress on Indigenous Languages of Latin America CILLA, Austin, Texas, 2005a.
- \_\_\_\_\_. **Review**: Language Contact in Amazonia. Alexandra Y. Aikhenvald. In: International Journal of American Linguistics, 2005b. 71, p. 505-07.
- \_\_\_\_\_. **Glottalization and other suprasegmental features in Wanano**. In: International Journal of American Linguistics, 2007. 73, p. 331-366.

STROM, Clay. **Retuarã Syntax: Studies in the Language of Colombia** 3. Arlington: Summer Institute of Linguistics/University of Texas, 1992.

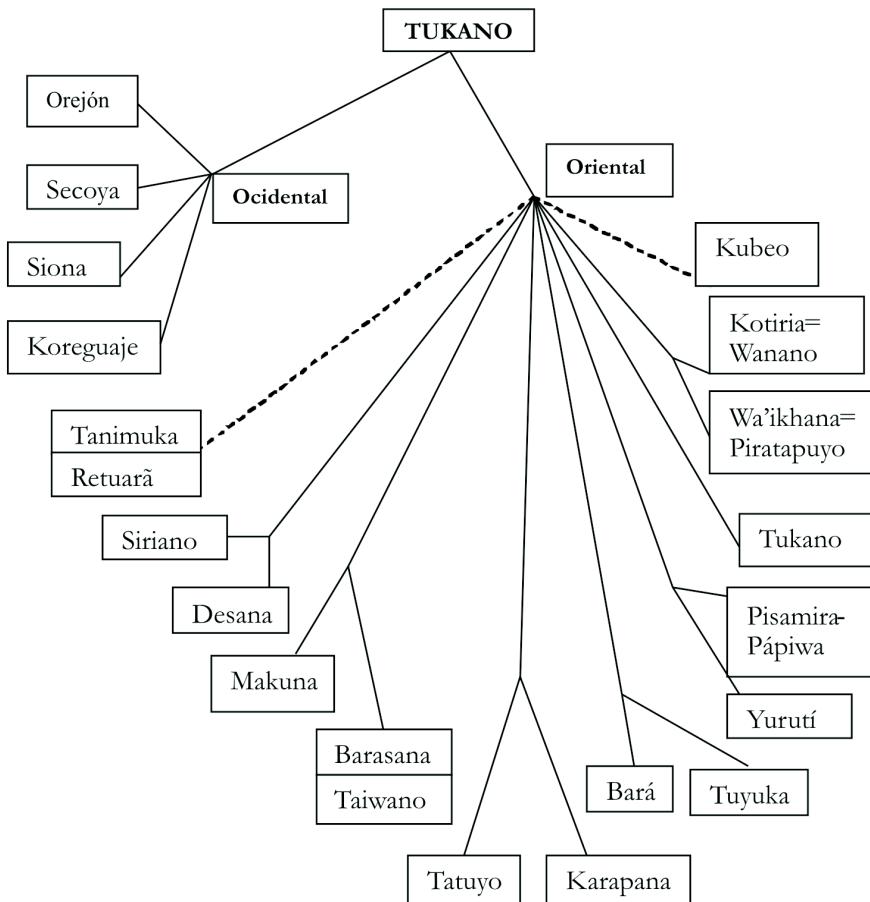
VALENCIA LÓPEZ, Simón. **Pamiene Toivaiyede Mahikaitukubo.** Cartilla para Aprender a Escribir. Pamie (Kubeo): Centro Experimental Piloto, 1994.

WALTZ, Nathan E. **Innovations in Wanano (Eastern Tucanoan) when compared to Piratapuyo.** In: International Journal of American Linguistics, 2002. 68, p. 157-215,

WHEELER, Alva. **Comparaciones Lingüísticas en el Grupo Tucano Occidental.** In: STEPHEN H., Levinsohn (Ed.). **Estudios Comparativos Proto Tucano.** Santafé de Bogotá: Editorial Alberto Lleras Camargo, [s.d.]. 1992. p. 17-53.

WRIGHT, Robin M. **História indígena e do indigenismo no Alto Rio Negro.** Campinas/São Paulo: Mercado de Letras/Instituto Socioambiental, 2005.

## ANEXO 1. A FAMÍLIA LINGUÍSTICA TUKANO



(Footnotes)

<sup>1</sup> Os dados do Tukano provêm de Ramirez (1997), os do Wa'ikhana são de Waltz (2002) ou, como todos os dados do Kotiria, da minha própria pesquisa de campo. Os dados de Barasana são de Gomez-Imbert, e os dados de outras línguas TO provêm das listas de palavras encontradas em (González de Pérez & Rodríguez de Montes 2000). Nos exemplos de Kotiria e Waikhana, a marca de nasalização precede o morfema, indicando nasalização morfêmica: a nasalização se manifesta em todos os segmentos sonoros do morfema (vogais, oclusivas sonoras e aproximantes). Em todos os exemplos, o tom alto é indicado com acento agudo (') e tom baixo não é marcado; a marca de apóstrofe (') representa uma oclusão glotal. Ver Stenzel (2007), para uma análise detalhada dos suprasegmentos em Kotiria: nasalização, tom e glotalização.



# RELAÇÕES PREDICATIVAS DAS PREPOSIÇÕES

Márcia Barreto BERG  
Universidade Federal de São João del Rei

## RESUMO

*Este trabalho trata da predicação das preposições, de acordo com Gramática Gerativa, propondo uma distinção entre preposições predicadoras e funcionais. As preposições predicadoras são aquelas que têm argumentos não acarretados pelos verbos e são elas próprias que atribuem papel temático ao seu argumento; já as preposições funcionais são aquelas que encabeçam os argumentos lógicos acarretados pelos verbos.*

## ABSTRACT

*This paper deals with the predication of prepositions according to the Generative Grammar. Following, it defines what we understand by predicating and it argues for a distinction between predicating and functional prepositions. Predicating prepositions have arguments which are not demanded by the verbs and they, themselves, attribute a thematic role to their arguments. Functional prepositions are those that introduce the logical arguments demanded by the verbs.*

## PALAVRAS-CHAVE

*Preposições funcionais. Preposições predicadoras*

## KEY WORDS

*Prepositions. Functional prepositions. Predicative prepositions,*

## Introdução

A pergunta inicial, neste trabalho, é se de fato, todas as preposições são predicadoras. Estamos assumindo, junto a FRANCHI e CANÇADO (1997, p. 95), que:

Entre quaisquer categorias lexicais (predicadoras em um sentido lógico) ou “predicadores”, para usar o termo distintivo sugerido por RAPOSO (1992) e seus argumentos se estabelecem relações predicativas a que chamamos “relações temáticas”. As funções determinadas por essas relações sobre os termos chamamos “diátese” e “papéis temáticos”. Os papéis temáticos assim caracterizados têm, portanto, um caráter derivado e não são meros termos primitivos da teoria.

Baseados nessa definição, vamos responder, pois, se a preposição tem sempre essa capacidade de estabelecer relações predicativas (relações temáticas) com seus argumentos. Quanto à diátese, de acordo com FRANCHI e CANÇADO (*ibid*: 95),

um item lexical-predicador (independente da categoria lexical a que pertence) contém em sua representação a diátese em que se fixou historicamente para seu uso atual, e que contribui para a estruturação da sentença e para sua interpretação, na medida em que determina um conjunto de argumentos explícitos ou implícitos que devem ser “instanciados” em determinadas posições sintáticas.

em (1), mostramos que algumas preposições determinam argumentos explícitos. Em geral, entende-se que uma categoria lexical que determina certos argumentos, exclui outros:

- (1) a. O réu falou perante *o juiz*.
- b. \*O réu falou perante o dia 15 de outubro.
- c. Roberto saiu com os amigos.
- d. \*Roberto saiu com seis horas da tarde.

É assumido na literatura que algumas preposições estabelecem relações predicativas (são elementos “predicadores”), por exemplo:

- (2) Ela fez tudo isso para seu namorado.

Acreditamos que há um consenso entre os autores em afirmar que a preposição para estabelece uma relação predicativa com seu argumento, seu namorado e atribui a ele o papel temático de beneficiário. Isso se dá a partir dos sentidos lexicais de para e seu namorado. Entretanto, existem outros contextos em que isso não ocorre. De acordo com CHOMSKY (1986) é esse o caso do conhecido exemplo:

- (3) the destruction [of the city].

CHOMSKY argumenta que nesse ambiente é o nome destruction que seleciona um complemento e, portanto, é o nome que atribui papel temático ao SN the city. Segundo a proposta de Chomsky, a preposição tem só a função sintática de realizar o caso na estrutura superficial, pois, para o papel temático ficar visível na sintaxe é necessário um caso para essa posição<sup>1</sup> LOBATO (1989), na mesma linha de CHOMSKY, argumenta que há dois tipos de preposições: verdadeiras (com conteúdo semântico: veio sem dinheiro, viajou com amigos, etc) e não verdadeiras (ou marcadoras de Caso), como por exemplo a preposição de em o seu desejo de sucesso. OUHALLA (1994), ao tratar dos papéis temáticos, admite que as preposições on e to, em Mary put the book on the shelf / Mary gave the book to John, não atribuem papel temático. Para o autor, nesses ambientes, o verbo marca tematicamente o SPrep e a preposição apenas transmite este papel ao SN. Na mesma linha de raciocínio de OUHALLA, outros autores analisam algumas preposições como invisíveis para a teoria temática, como MARANTZ (1984), WILLIAMS (1987), GRIMSHAW e WILLIAMS (1993) entre outros. O que há de comum, entre esses autores, é que eles apenas se preocupam com a atribuição de papel temático.

<sup>1</sup> Não é pertinente, para os propósitos deste texto, entrar em detalhes sobre o funcionamento da Teoria Gerativa. Por isso, remeto o leitor a textos que tratam do assunto, como CHOMSKY (1986, p. 192-193), RAPOSO (1992, p. 510-511), MIOTTO (1999), entre outros.

Também concordamos que em determinados ambientes, as preposições não atribuem papel temático, mas queremos também salientar que as preposições funcionais podem ter sentido. Por isso, estamos assumindo que, como comprova BERG (1996), independente do contexto sentencial, todas as preposições têm sentido. Vejamos quais são as nossas evidências a favor dessa afirmação.

Vamos dividir as preposições em dois tipos, predicatoras e funcionais, de acordo com as relações predicativas que elas estabelecem: as do primeiro tipo, predicatoras, são aquelas que atribuem papel temático a seu complemento. As do segundo tipo, funcionais, são aquelas que não atribuem papel temático ao seu complemento. Advertimos que predicatoras e funcionais não são classes distintas, mas funções. Distanciamo-nos, porém, da abordagem gerativa na explicitação de como os papéis temáticos são atribuídos.

Primeiramente, vejamos a noção de argumento aqui assumida. CANÇADO (2003) propõe que os papéis temáticos atribuídos por um verbo serão especificados a partir da estrutura conceitual semântica desse predicator, e não a partir da estrutura temática como assumem, geralmente, os gerativistas. Explicando melhor, os argumentos de um verbo, por exemplo, são todos os argumentos acarretados por esse verbo<sup>2</sup> no sentido lógico do termo argumento. Vamos chamar esses argumentos de argumentos lógicos, para não confundir com os argumentos atribuídos geralmente a um verbo (os argumentos que se projetam na posição de sujeito e complementos de um verbo). Por exemplo, estamos assumindo que um verbo como vender terá quatro argumentos lógicos, cujos papéis temáticos serão: desencadeador, objeto estativo, alvo e valor:

- (4) Ana vendeu seu apartamento para Pedro por dez reais.

---

<sup>2</sup> DOWTY(1989) define acarretamento lexical de um predicado como o grupo de todas as coisas que podemos concluir sobre *x* somente por saber que a sentença *x* *predicador* *y* é verdadeira.

Necessariamente, esses quatro argumentos (Ana, seu apartamento, Pedro, dez reais) são acarretados, no sentido de DOWTY (1989), pelo verbo vender, e é o verbo que lhes atribui papel temático, ou seja, quando pensamos quando pensamos no item lexical vender sempre está associado a ele o conhecimento de que existem um agente dessa venda, um objeto vendido, uma pessoa para quem vendemos e um valor. Veja que argumentos lógicos não equivalem às noções de complemento e adjunto e não há nenhuma explicitação do que deve ou não deve estar presente na sintaxe. Ainda tomando o verbo vender como exemplo, repetimos, em (5), o exemplo (4):

- (5) a. Ana vendeu seu apartamento para Pedro por dez reais.
- b. Ana vendeu seu apartamento para Pedro.
- c. Ana vendeu seu apartamento.
- d. E o apartamento? Ana vendeu.
- e. Apartamentos vendem bem.

Como se pode ver pelas sentenças em (5) b-e, os argumentos lógicos que foram explicitados (5) podem não ocorrer, necessariamente, na estrutura superficial da sentença (em realidade, nem os argumentos, sujeito e complementos, normalmente atribuídos ao verbo vender, de acordo com os estudos gramaticais, precisam estar explicitados na sintaxe). Tendo em vista que os argumentos lógicos são acarretados pelo predicador a partir da semântica desse predicador e passando agora para as preposições, entendemos que elas são predicadores de natureza diferente dos verbos. Os verbos são predicadores por excelência e é possível especificar seus argumentos lógicos mesmo fora do contexto sentencial. Em geral, um mesmo verbo possui mais de um argumento lógico. Diferentemente dos verbos, com exceção da preposição entre, há apenas um argumento interno para cada preposição.

Assumindo, então, a idéia de argumento lógico, que daqui para frente chamaremos apenas de argumento, vejamos como as preposições serão classificadas de acordo com a sua função semântica.

## 1 Preposições predicatoras

De acordo com o proposto acima, as preposições que têm a função de predicatoras são aquelas que têm argumentos que não são acarretados pelos verbos e são elas próprias que atribuem papel temático ao seu argumento. Em

- (6) João viajou *entre* as bananas.

o SN *bananas* não é acarretado pelo verbo *viajar*, pois podemos afirmar que o verbo *viajar* não acarreta a noção de que o agente desta viagem tem que viajar necessariamente entre alguma coisa. Assim o SN *as bananas* não é um argumento do verbo *viajar*, mas é um argumento da preposição *entre*. Ressaltamos que não são somente as preposições fortes que têm a função de serem predicatoras. Também temos preposições fracas que têm essa função:

- (7) João viajou *com* sua namorada.

O argumento *sua namorada* não é acarretado pelo verbo *viajar*, pois podemos afirmar que o verbo *viajar* não acarreta a noção de que o agente desta viagem tem que viajar necessariamente com alguém. Assim o SN *sua namorada* não é argumento do verbo *viajar*, mas sim, da preposição *com*. Estamos assumindo que papel temático de um item lexical é a relação de sentido estabelecida entre esse item lexical e seu predicator. Para podermos estabelecer qual é o papel temático atribuído pela preposição *com* ao seu complemento, temos que nos valer da idéia de composição. Entretanto, a noção de composição em relação à preposição é um pouco mais complexa do que o que ocorre com os verbos. Para as preposições temos que o papel temático atribuído ao argumento dessa preposição só pode ser estabelecido a partir da composição do predicator com seu complemento, mas a composição desse predicator complexo com o verbo predicator da sentença; só

assim se pode estabelecer exatamente o conteúdo semântico do papel temático atribuído ao complemento da preposição *com*. Estamos assumindo que as preposições, nesses ambientes, são predicadores numa relação complexa de sentidos, isto é, o conteúdo semântico do papel temático vem da predicação componencial ou composicional<sup>3</sup>, segundo a proposta de FRANCHI (1997a)<sup>4</sup>. Explicando melhor, vejamos um exemplo do próprio FRANCHI (1997b, p.173):

- (8) João [[rasgou o bilhete] rapidamente]

Para FRANCHI,

- *rasgar* é o predicador de *o bilhete* (que, por sua vez, é argumento de *rasgar*)
- *rapidamente* é predicador da expressão complexa *rasgar o bilhete* (que, por sua vez, é argumento de *rapidamente*)
- A expressão complexa *rasgar o bilhete rapidamente* é predicador de *João* (seu argumento).

Vejamos em (9) um exemplo, com preposição, que ilustra a idéia de predicação complexa.

- (9) João viajou de chinelos.

---

<sup>3</sup> Sobre os termos “componencial” e “composicional” FRANCHI (1997a) assevera que “... com o primeiro refiro-me ao fato de que é o resultado da construção de uma expressão complexa  $X=[YZ]$  que se deve atribuir uma propriedade da relação de  $X$  a outra expressão  $W$ , e não exclusivamente a propriedades de  $Y$  e de  $Z$ . Com o segundo termo, refiro-me aos casos em que se dá um processo transitivo pelo qual, se expressões  $X$  e  $Y$  contratam uma relação  $R_iXY$ ,  $Y$  e  $Z$  contratam  $R_iYZ$ , então  $X$  e  $Z$  contratam uma relação  $R_iXZ$ .

<sup>4</sup> Para FRANCHI a Predicação não deve ser vista como uma relação local entre os predicados e seus argumentos, antes a Predicação é uma relação de sentido entre expressões singulares ou entre expressões complexas, correlata das operações construtivas que as combinam na derivação sintática. A relação de sentido entre as expressões lingüísticas é determinada exclusivamente por propriedades semânticas dos itens lexicais.

Seguindo as pegadas de FRANCHI, podemos analisar (9) como:

- *de* é predicador de *chinelos* (como seu argumento);
- A expressão complexa *de chinelos* é predicador de *viajou* (como seu argumento);
- A expressão complexa *viajou de chinelos* é predicador de *João* (como seu argumento).

Em (9), o sintagma preposicionado *de chinelos* vai receber as propriedades semânticas de estativo-modo. O conteúdo semântico do papel temático atribuído por uma preposição predicadora, principalmente se esta tiver um sentido fraco, depende da composição da preposição com o complemento, e, às vezes também, dessa expressão complexa com o verbo da sentença, isto é, de acordo com CANÇADO (2003), o papel temático de um argumento é derivado do grupo de propriedades atribuídas a esse argumento a partir das relações de acarretamento estabelecidos por toda a proposição em que esse argumento se encontra.

Passemos agora à atribuição de papel temático e preposições funcionais.

## 2 Preposições funcionais

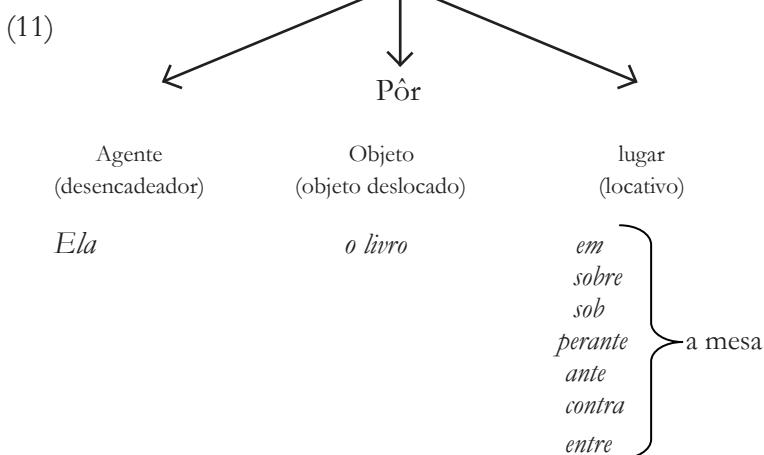
Seguindo algumas idéias já esboçadas em BERG (1996), vamos assumir que as preposições funcionais são aquelas que encabeçam os argumentos lógicos acarretados pelos verbos.

Essas preposições não atribuem papéis temáticos. Entretanto, assumimos que elas têm sentido e que seus sentidos têm que ser compatíveis semanticamente com o papel temático dos argumentos acarretados pelo verbo predicador:

(10) João jogou a bola para a cesta.

O papel temático do argumento *a cesta*, um alvo, é atribuído pelo verbo *jogar*, pois podemos afirmar que o verbo *jogar* acarreta as propriedades semânticas de que existe necessariamente um agente que joga, alguma coisa que é jogada e um alvo (locativo) para onde esse objeto é jogado. Nesses ambientes, as preposições funcionais não atribuem papel temático, sendo o verbo o predicador responsável pela atribuição de papel temático. De forma geral, exceto em certos ambientes específicos, na sintaxe do português, não é usual que mais de um argumento interno seja explicitado na sentença sem preposição. Talvez, por isso, é necessário que a preposição *para* seja introduzida na estrutura sintática de (10), no entanto, não é ela a responsável pela atribuição de papel temático.

Quanto ao sentido da preposição funcional, é necessário que ele seja compatível com o sentido do papel temático do argumento acarretado pelo verbo. Tomemos, por exemplo, o verbo *pôr*, na sentença, *ela pôs o livro na mesa*. Podemos afirmar que o verbo *pôr* acarreta as propriedades semânticas de que existe necessariamente um agente que pratica a ação de *pôr*, um objeto que é posto e um lugar onde o objeto é posto. Observem que as preposições que encabeçam o segundo argumento interno de *pôr* têm que ter o mesmo sentido de lugar, como mostra (11) abaixo:



Em (11) o sentido das preposições *em, sobre, sob, perante, ante, contra, entre* é compatível com o sentido do papel temático do argumento acarretado pelo verbo *pôr*<sup>5</sup> (=lugar). Nessa sentença, todas as preposições (*em, sobre, sob, perante, contra, entre*) que se enquadram na noção de lugar, isto é, que podem encabeçar o argumento que expressa o sentido de lugar, são funcionais, pois não são elas que estão atribuindo papéis temáticos, e sim o verbo *pôr*. O fato de as preposições, em (11), possuírem sentido, não desencadeia papéis temáticos duplos (não aceitos pelos gerativistas), pois, como discutimos mais detalhadamente em BERG (2005), a definição de papel temático que estamos assumindo aqui remete a “um grupo de propriedades semânticas”. Por exemplo, em *João jogou o livro sobre a mesa*, o argumento *a mesa*:

- não altera suas propriedades durante um intervalo de tempo;
- expressa o lugar onde o livro cai;
- expressa o alvo onde João joga o livro, etc

Esse “grupo” de propriedades é o papel temático. É o grupo de propriedades semânticas que tem estatuto teórico aqui, e não o papel temático por si, como é afirmado nas abordagens gerativistas. O sentido da preposição entra na relação de sentidos do argumento *a mesa*.

Por outro lado, entendemos que nem todas as palavras que têm sentido, necessariamente, têm que atribuir papel temático.

Já em (12), vemos que não é possível colocar outras preposições encabeçando o argumento *a mesa*, porque elas não são compatíveis com a noção de lugar acarretado pelo verbo *pôr*.

- (12) \*Ela pôs o livro *com/ de/ desde/ para/ sem* a mesa.

---

<sup>5</sup> Gostaríamos de ressaltar que as preposições não são acarretadas pelos verbos. Os verbos acarretam argumentos e as preposições podem encabeçá-los ou não, dependendo da compatibilidade dos seus sentidos com o sentido dos argumentos dos verbos.

As preposições *com*, *de*, *desde*, *para*, *sem*, nessa sentença, não expressam a noção de lugar. Essa é uma evidência de que também as preposições funcionais têm sentido, pois é necessário haver uma compatibilidade de sentidos com papel temático atribuído pelo verbo.

Existe ainda uma outra categoria de preposição funcional, conhecida na literatura como preposições inerentes (FILLMORE 1966, LAKOFF 1970, GRUBER 1970, NEELEMAN 1997)<sup>6</sup>. Essas preposições constituem um pequeno grupo e parecem ser incorporadas ao verbo.

- (13) a. Rosa *concorda com* a Maria.
- b. Bernardo *carece de* talento artístico.
- c. Rosa *confia em* seu namorado.

Uma característica das preposições funcionais-inerentes é que não podem ser trocadas por outra:

- (14) a. \*Rosa *concorda a/ ante/ em/ de* o casamento de seu filho.
- b. \*Bernardo *carece a/ em/ com/ para* talento artístico.
- c. \*Rosa *confia a/ de/ com/ para* seu namorado.

Já as preposições funcionais que não são inerentes podem ser trocadas por outra:

---

<sup>6</sup> FILLMORE (1966, p. 11-33) trabalhando com a noção de Caso (= actante), em seu texto, assume que algumas preposições são preenchidas desde o léxico, em geral as que possuem maior carga semântica (dentro, antes, depois). A preposição do Caso objetivo é geralmente vazia dependendo do verbo, pois, em sua opinião, algumas preposições são determinadas pelos traços sintáticos inerentes de verbos específicos. LAKOFF (1970, p. 116-117) considera que algumas preposições são introduzidas pelas transformações (na estrutura superficial). Para GRUBER (1970, 39-58) as preposições são “formativos” desde o nível pré-lexical, sendo substituídas depois pela forma fonológica do item lexical, na estrutura superficial. O autor considera que há verbos que já possuem uma preposição incorporada, como os verbos de movimento. NEELEMAN nomeia esses casos de “seleção de preposições idiomáticas”.

- (15) a. Ela pôs o livro na mesa/ sobre a mesa/ ante a mesa/ contra a mesa, etc..

Como já vimos anteriormente, em (15), as preposições *em*, *sobre*, *ante*, *contra*, atuam como funcionais, não necessariamente inerentes.

Estamos assumindo, nesta seção, que quem atribui papel temático aos argumentos encabeçados pelas preposições inerentes são também os verbos. Entretanto, diferentemente dos últimos autores citados, assumimos que as preposições inerentes não são destituídas de sentido. Assim como nas preposições funcionais vistas anteriormente, o sentido das preposições inerentes também tem que ser compatível com o papel temático atribuído pelos verbos aos seus complementos. Vejamos, em (16), os sentidos das preposições inerentes dos exemplos em (13) acima:

(16)	<b>Sentido da preposição</b>	<b>Papel temático do argumento</b>
<i>Rosa concorda com Maria.</i>	ser o objeto de referência	objeto estativo
<i>Bernardo carece de talento artístico.</i>	ser o objeto de referência	objeto estativo
<i>Rosa confia em seu namorado.</i>	ser o objeto de referência	objeto estativo

Os argumentos do verbo encabeçados pelas preposições *com*, *de*, *em*, nessas sentenças, têm o papel temático de objetos estativos<sup>7</sup>, as preposições *com*, *de*, *em*, nessas sentenças, têm o sentido de serem apenas um objeto de referência. Ressaltamos que as preposições que ocorrem em ambientes funcionais-inerentes são majoritariamente aquelas que se classificam como as de sentido fraco.

<sup>7</sup> Segundo CANÇADO (2003), objeto estativo será definido como algo ao qual se faz referência, sem que haja uma mudança de estado. Para mais detalhes, ver BERG (2005, cap. 3).

Concluímos este trabalho realçando que as preposições, tanto as predicatoras como funcionais podem ser fracas e fortes e que essa classificação, predicatoras e funcionais, não divide as preposições em dois grupos mutuamente exclusivos, pois, a preposição *com*, por exemplo, vai ser funcional em *Rosa concorda com Maria*, e predicatora em *João viajou com sua namorada*, ou seja, o rótulo, a função, predicatora ou funcional, não está na preposição enquanto item lexical, mas será dado, de acordo com o ambiente semântico em que ela ocorre: se acarretada pelo verbo será funcional, caso contrário, será predicatora.

Outro ponto a ser realçado é que todas as preposições têm sentido e que, apesar de terem sentido, elas nem sempre atribuem papel temático. Esta visão é diferente da tradição gerativista que atrela a noção de sentido à atribuição de papel temático, isto é, para alguns gerativistas, as preposições que não atribuem papel temático são aquelas consideradas de sentido esvaziado, ou mesmo sem sentido. A nossa proposta parece coerente com o comportamento de outras classes de palavras. Por exemplo, temos nomes que não selecionam argumentos e não atribuem papel temático (*cadeira, livro, etc*), e nomes que selecionam e atribuem papéis temáticos (*construção, elaboração, etc*); entretanto, todos os dois tipos de nome têm sentido.

### 3 Conclusão

As informações mais relevantes que podemos depreender respondem às perguntas que abrem esse trabalho: a) as preposições têm sentido? (b) quais são os sentidos das preposições? (c) as preposições são predicatoras? Vimos que:

- a) todas as preposições, em todos os ambientes, têm sentido, isto é, os falantes têm uma idéia ou um conceito sobre esses itens lexicais.

- b) Algumas preposições cobrem uma gama de sentido maior que outras. As que apresentam mais sentidos são denominadas de fracas, que são as seguintes: *a, com, de, em, para, por*. Já as que apresentam um sentido mais específico são denominadas de fortes: *ante, após, até, contra, desde, entre, perante, sem, sob, sobre*.
- c) As preposições que antecedem argumentos acarretados pelos verbos são denominadas de funcionais e não atribuem papel temático. As preposições que não são acarretadas pelos verbos são denominadas de predicatoras e atribuem papel temático de acordo com a predicação complexa de FRANCHI.
- d) As preposições fortes e fracas podem desempenhar a função predicatora e/ou funcional, dependendo do contexto semântico em que elas aparecem, portanto a classificação das preposições em funcionais e predicatoras não estabelece dois grupos mutuamente exclusivos.

## Referências

BERG, Márcia Barreto. **A Natureza Categorial da Preposição**. Belo Horizonte: FALE/UFMG, 1996. Dissertação (Mestrado em Letras)  
\_\_\_\_\_. **O Comportamento Semântico-Lexical das Preposições do Português do Brasil**. Belo Horizonte: FALE/UFMG, 2005. Tese (Doutorado em Letras)

CANÇADO, Márcia. **Um estatuto teórico para os papéis temáticos**. In: MÜLLER, Ana Lúcia; NEGRÃO, Esmeralda V.; FOLTRAN, Maria José (Org). **Semântica formal**. São Paulo: Ed. Contexto, 2003a.

CHOMSKY, Noam. **Knowledge of Language: Its Nature, Origin and Use**. New York: Praeger Publishers, 1986.

DOWTY, D. R. **On the content of the notion of thematic role.** In: CHIERQUIA; PARTEE (eds). **Properties, types and meaning. Studies in linguistic and philosophy 2.** Dordrecht: Kluver, 1989. p. 69-129. (Semantics Issues).

FILLMORE, Charles J. **A Proposal Concernning English Prepositions.** Report of The Seventeenth Annual Round Table Meeting on Linguistic and Language Studies, ed. Francis P. Dineen, s. J. Washington, 1966.

FRANCHI, Carlos. **Predicação.** Manuscrito publicado em CANÇADO, M. (org) **Predicação, Relações Semânticas e Papéis Temáticos:** anotações de Carlos Franchi. Revista de Estudos da Linguagem, Belo Horizonte: 1997a. v. 11, n. 2, p. 17-81, Jul/Dez. 2003.

\_\_\_\_\_. **Teoria da Adjunção:** predicação e relações temáticas. (1997b). Manuscrito publicado em CANÇADO, M. (org) **Predicação, Relações Semânticas e Papéis Temáticos:** anotações de Carlos Franchi. Revista de Estudos da Linguagem, Belo Horizonte: Jul/Dez. 2003. v. 11, n. 2, p. 155-176.

\_\_\_\_\_; CANÇADO, M. (1997). **Teoria generalizada dos papéis temáticos.** Manuscrito publicado em CANÇADO (Org). **Predicação, Relações Semânticas e Papéis Temáticos:** anotações de Carlos Franchi. Revista de Estudos da Linguagem, Belo Horizonte, v. 11, n. 2, p. 83-123, Jul/Dez. 2003.

GRIMSHAW, Jane e WILLIAMS, Edwin. **Nominalization and Predicative Prepositional Phrases.** In: PUSTJOVSKY, James (Ed.). **Semantic and The Lexicon.** Netherlands: Kluwer Academic Publishers, 1993. p. 97-105.

GRUBER, Jeffrey. **Studies in Lexical Relations.** Reproduced by the Indiana University Linguistic Club, 1970.

LAKOFF, George. **Irregularity in Syntax**. New York: Holt, Rinehart, e Winston, Inc., 1970.

LOBATO, Lúcia M. Pinheiro. **Advérbios e Preposições, Sintagmas Adverbiais e Sintagmas Preposicionais**. D.E.L.T.A, 1989. v. 5 (1): 101-120.

MARANTZ, A. **On the nature of grammatical relations**. Cambridge. Mass: MIT Press, 1984.

MIOTO, Carlos; SILVA, M. C. F.; LOPES, R. V. **Manual de Sintaxe**. Florianópolis: Insula, 1999.

NEELEMAN, A. D. PP-complements. **Natural Language and Linguistic Theory**. Netherlands: 1997. v. 15, p. 89-137.

OUHALLA, Jamal. **Introducing transformational grammar** - from rules to principles and parameters. Great Britain: Bristishlibrary, 1994.

RAPOSO, Eduardo Paiva. **Teoria da gramática**: a faculdade da linguagem. Lisboa: Caminho, 1992.

WILLIAMS, E. **Implicit Arguments, the Binding Theory, and Control**. Natural Language and Linguistic Theory, 1987. 5: 151-180.

# **PERCURSO DO ARTIGO DEFINIDO DIANTE DE N PRÓPRIO NA FALA DE CRIANÇAS EM FASE DE AQUISIÇÃO**

Patrícia Vargas ALENCAR  
UNIRIO

## **RESUMO**

*Este artigo focaliza a emergência do uso variável de artigo definido frente a N próprio por crianças que estão adquirindo o Português como L1. Procedemos a uma análise de situações reais de uso da língua, conjugando um estudo longitudinal e um estudo estratificado. Nossa análise evidenciou que a aquisição da variação do artigo definido se organiza em determinados padrões que são paralelos, em alguns aspectos, à fala do adulto.*

## **ABSTRACT**

*This article is focused on the upcoming variable use of the definite article before proper N by children who are acquiring Portuguese as L1. We proceed in the analysis of situations of real language use, putting together a longitudinal and stratified study as well. Our analysis put on evidence that the acquisition of the variation of the definite article is organized under determined patterns that parallel, in some aspects, the speech of the adult.*

## **PALAVRAS-CHAVE**

*Artigo definido. Aquisição. Primeira língua. Uso variável.*

## **KEY WORDS**

*Acquisition. Definite article. First language. Variable use.*

## Introdução

Neste trabalho, expomos parte da análise apresentada em nossa tese de Doutorado (ALENCAR, 2006) sobre o percurso aquisitivo do artigo definido no contexto variável “Art + N próprio”, conforme os exemplos “Cadê **a Quistina?**” (MA 2;02 p.2) / “Cadê **Ø** Quistina”. Trata-se de uma pesquisa calcada no arcabouço teórico-metodológico da Teoria da Variação de orientação Laboviana e que se orienta pelo estudo da aquisição baseada nas situações reais de uso da língua.

Nosso principal objetivo foi verificar a emergência e incorporação dos padrões que regulam a variação do artigo frente a N próprio no discurso da criança, de modo a confrontar a variação da fala infantil com os padrões de variação da fala dos adultos com os quais interagiram nas amostras analisadas para investigar até que ponto o discurso da criança reflete o *input* a que teve acesso.

Uma das hipóteses levantadas na pesquisa é que a variação no uso do artigo frente a N próprio segue uma trajetória que reflete, em grande parte, características do *input* variável a que a criança tem acesso durante o processo aquisitivo de L1. Assim, pressupomos que o artigo, no contexto variável focalizado, é incorporado primeiro nos contextos mais favoráveis à presença do determinante na fala do adulto, já que esses contextos estariam mais salientes no *input* a que a criança está exposta, e vai, gradativamente, se expandindo para outros contextos até que a fala da criança passe a espelhar a regularidade da variação observada na comunidade.

Na comparação entre a fala das crianças e a fala dos adultos, partimos da hipótese de que a variação no uso do artigo frente a N próprio no discurso infantil pode não refletir inteiramente o padrão observado na fala dos adultos, em função de fatores maturacionais, ou seja, de pressões associadas ao desenvolvimento linguístico e cognitivo mais geral da criança.

As principais questões que nortearam este estudo foram: (1) Estaria a fala da criança refletindo os padrões de variação da fala do adulto (*Child-*

*Directed Speech)?* Neste caso, como o *input* opera no processo aquisitivo? (2) Considerando que a trajetória aquisitiva do artigo definido frente a N próprio pode sofrer restrições em razão do próprio desenvolvimento cognitivo da criança, em que medida fatores de ordem maturacional podem estar motivando a variação?

## 1 Amostra e Metodologia:

Conjugamos dois tipos de estudo da fala de crianças em fase de aquisição da linguagem: um estudo longitudinal, através do qual acompanhamos o desenvolvimento linguístico de uma criança da amostra da Unicamp<sup>1</sup>, no período de 1;2 a 4;10 e um estudo estratificado baseado na fala de 10 crianças (7 da amostra da Puc-SP e 3 da amostra da Unicamp) subdividida em cinco pontos etários (1;6, 2;00, 2;6, 3;00 e 4;00).

Realizamos também a análise da variação na fala dos adultos presentes nas situações interacionais consideradas. Os procedimentos para a análise da fala da criança foram, então, aplicados na fala do adulto.

Comparamos ainda os resultados deste estudo aos de pesquisas que investigaram o uso do artigo definido frente a N próprio na comunidade, a fim de minimizar os possíveis efeitos da limitação imposta por um estudo que controla a fala dos adultos unicamente em situações de interação com as crianças analisadas.

---

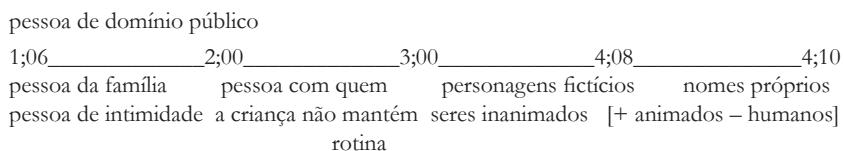
<sup>1</sup> Trata-se de uma amostra cujos dados constituem o corpus do “Projeto Aquisição da Linguagem Oral”, coordenado pela professora Cláudia Lemos, do Instituto de estudos da Linguagem (IEL)/UNICAMP. Com o objetivo de descrever e analisar o processo de aquisição do português como primeira língua em 12 crianças com idades que recobrem o período de 11 meses a 6 anos, a amostra foi composta por gravações de áudio e videotape realizadas com crianças paulistas da classe média, entre as décadas de 70 e 80, filhas de pais universitários. Grande parte dos entrevistadores é composta pelos professores do IEL que gravaram os seus próprios filhos. A amostra encontra-se organizada cronologicamente e apresenta informações sobre o contexto de gravação como, por exemplo, a data da realização, idade do falante, investigador, acompanhante, descrição física, entre outras.

## 2 Resultados mais significativos

### 2.1. Análise qualitativa

A trajetória da incorporação do artigo definido frente a nome próprio pode ser sintetizada e observada na figura 1 em que dispusemos, na parte superior, o tipo de nome próprio que ocorreu com ausência categórica de artigo a partir do ponto etário em que está localizado e, na parte inferior, os nomes que ocorreram com presença variável de artigo de acordo com os pontos etários a partir dos quais passaram a ocorrer. Importa ressaltar que a análise qualitativa esquematizada na figura 1 corresponde à fala de R, criança com cujo discurso fizemos o estudo longitudinal:

FIGURA 1: Expansão do uso do artigo conforme o tipo de N próprio



A figura 1 mostra evidências de que a idade de 3;00 é um marco para a trajetória percorrida pelo artigo definido diante de nome próprio. A fase anterior a 3;00 é marcada pela emergência do artigo nos contextos preferenciais, evidenciando certa restrição lexical vinculada a nomes mais familiares. A partir dos 3;00, observa-se uma expansão no uso do artigo com a sua extensão a um conjunto mais amplo de N próprios.

Ao longo dessa análise, puderam ser depreendidos alguns aspectos relevantes para a compreensão da trajetória aquisitiva do artigo definido. Ressaltou-se a prevalência da função dêitica do artigo, no início do processo aquisitivo, a estreita correlação entre presença da preposição e ocorrência do artigo, desde a sua emergência, e a importância do fator familiaridade. Através de uma análise multivariacional, procuramos

verificar a validade estatística dessas indicações, assim como o paralelismo entre os padrões variáveis depreendidos na fala da criança e os padrões verificados na fala do adulto, confrontando-os com os resultados obtidos para uso do artigo na comunidade. A seguir, apresentamos parte dos principais resultados de nossa pesquisa quantitativa.

## 2.2. Análise quantitativa

Nesta seção, focalizamos a incorporação dos padrões de variação no uso do artigo frente a N próprio, confrontando os padrões de variação identificados na fala infantil com os dados verificados na fala dos adultos. Através dessa comparação, buscamos verificar o paralelismo entre as amostras e a importância do *input* no processo aquisitivo, já que a fala a que a criança está exposta é considerada fator importante no desenvolvimento da linguagem e na incorporação de processos variáveis.

Discutimos o efeito de seis variáveis independentes que se mostraram relevantes para o uso do artigo antes de nome próprio. São elas: “Estrutura do SN”, “Função Sintática”, “Forma de recuperação do referente”, “Especificidade do referente”, “*Status Informacional*” e “Introdução do referente”. Analisamos a trajetória aquisitiva do artigo definido frente a nome próprio em cada grupo de fatores mencionado de acordo com a análise multivariacional possibilitada pelo pacote computacional VARBRUL (PINTZUK, 1988). Procedemos, ainda, a um cruzamento entre as variáveis “Faixa etária” e as demais mencionadas, a fim de identificar o momento em que a variação na fala da criança espelha a variação observada na fala do adulto. Essa análise permite traçar as regularidades que regem a emergência e ampliação no uso do artigo definido antes de N próprio na fala infantil.

Adotamos na discussão das variáveis relevantes o seguinte procedimento: partimos dos resultados do estudo longitudinal e comparamos tais resultados àqueles obtidos no estudo estratificado. Tal procedimento nos permite observar, num primeiro momento, os

contextos mais favoráveis para a emergência do artigo diante de N próprio, bem como o ponto mais específico do *continuum* aquisitivo em que esses contextos se sistematizam. A comparação com as tendências verificadas no estudo estratificado nos permite, assim, verificar a sistematicidade e a regularidade dos padrões constatados.

Como já dissemos, dada a impossibilidade de um estudo controlado do fenômeno na comunidade a que pertencem os informantes que compõem nossa amostra, tomamos os resultados apontados por outros trabalhos como um parâmetro indicativo do comportamento da variação na comunidade.

A comparação entre os padrões de variação na fala das crianças e na fala dos adultos está sujeita aos limites impostos pelos problemas inerentes à questão do *input* durante o desenvolvimento da linguagem. Estudos sobre as trocas comunicativas entre a criança em fase de aquisição e os adultos com quem ela interage evidenciam que a mediação entre a criança e a ação com o mundo, provocada pela maneira como o adulto se dirige à criança, é fator fundamental para o desenvolvimento linguístico (GUY e BOYD, 1990 e ROBERTS, 2002). É importante ressaltar, no entanto, que a comparação entre a fala infantil e a fala dos adultos apresenta limitações, uma vez que, como mostram diversos trabalhos (SNOW, 1994, PINE, 1994, LIEVEN, 1994 e RICHARDS, 1994), a fala que o adulto dirige à criança, principalmente em estágios iniciais de aquisição, em muitos aspectos, pode não corresponder ao seu uso linguístico nas interações com adultos ou mesmo com crianças mais velhas. Ao se dirigir a crianças em fase aquisitiva, o adulto usa uma espécie de fala infantilizada (“*Baby Talk*” ou “*Child-Directed Speech*”), expondo as crianças a amostras de fala que não correspondem inteiramente ao uso linguístico estabilizado.

Em função do espaço limitado comum a artigos, apresentamos aqui apenas os resultados mais significativos de dois grupos de fatores imprescindíveis para a análise da incorporação do artigo, a saber: “Estrutura do SN” e “Forma de Recuperação do Referente”. Indicamos

a leitura da tese (ALENCAR, 2006) para um acesso mais detalhado à análise.

### 2.2.1 Estrutura do SN

Um dos aspectos apontados como relevante pela análise multivariacional para o uso do artigo definido frente a nome próprio na fala infantil diz respeito à configuração estrutural do SN. Os resultados estatísticos referentes tanto à fala infantil quanto à fala adulta mostram que a presença da preposição é atuante na ocorrência do artigo, confirmando a tendência já apontada na análise qualitativa.

Controlamos, inicialmente, diferentes estruturas de SN, conforme os exemplos abaixo:

*Artigo + Nome Próprio*

(26) “**Cadê a Quistina?**” (MA 2;02 p.2)

*Artigo + Atributo<sup>2</sup> + Nome Próprio*

(27) “**Titiu Ricardo**” (BGL 2;00 p.1)

*Preposição + Artigo + Nome Próprio*

(28) “**Casa do Leandro**” (MA 2;02 p.1)

*O Preposição + Artigo + Nome Próprio*

(29) “**Eu di um besu O Tátia**” (MA 2;00 p.3)

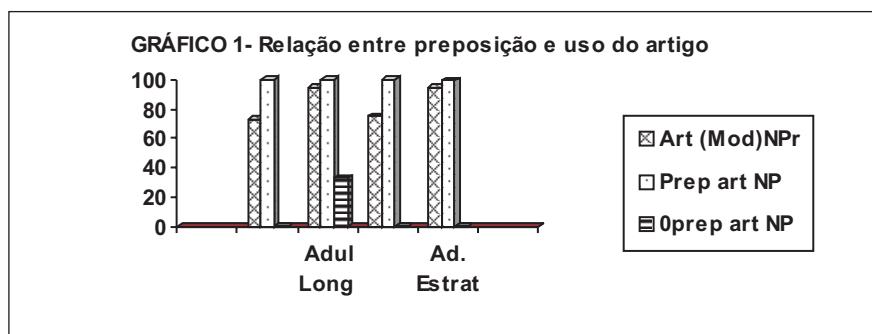
É necessário destacar, antes de tudo, que se fez necessária uma reorganização desse grupo de fatores, em função da própria natureza dos dados. Uma análise estatística prévia indicou maciça concentração dos dados das crianças na configuração estrutural mais simples “Art

---

<sup>2</sup> A palavra “atributo” diz respeito a qualquer elemento situado entre o artigo e o nome próprio que qualifica este último, como “titio”, “vó”, “Dona” etc.

“+ Nprop”<sup>3</sup>. Em função do número pouco expressivo de dados na estrutura “Art + Atributo + Nprop”<sup>4</sup> e das semelhanças na configuração dos SNs, reagrupamos os dados das duas estruturas em questão. Do amálgama de tais estruturas resultou a configuração “Art + (Atributo) + Nprop”, em que a ocorrência de “atributo” é facultativa.

Os resultados de nossa pesquisa permitem constatar a estreita correlação entre a presença da preposição e a presença do artigo. O gráfico 1 mostra que a distribuição de taxas de artigo na fala infantil é perfeitamente paralela à que se verifica na fala dos adultos.



O paralelismo entre a fala infantil e a fala dos adultos no que se refere à presença categórica de artigos em SN`s encaixados em sintagmas preposicionais leva a crer que o *input* atua no processo aquisitivo.

Outros trabalhos também já haviam apontado tal tendência no uso variável do artigo definido, principalmente os que se voltaram para o estudo da comunidade. Oliveira e Silva (1996), a partir da análise de amostras de fala do Rio de Janeiro, mostra que a preposição leva ao uso do artigo em grande parte dos dados, conduzindo a autora a excluí-los (Op. Cit: 128):

<sup>3</sup> Estudo Longitudinal da fala infantil: “Art + Nprop”: 141/187 = 75%  
Estudo Estratificado da fala infantil: “Art + Nprop”: 45/60 = 75%

<sup>4</sup> Estudo Longitudinal da fala infantil: “Art + Atributo + Nprop”: 1/7 = 14%  
Estudo Estratificado da fala infantil: “Art + Atributo + Nome”: 1/1 = 100%

Foram também eliminados da computação eletrônica todos os dados do tipo “O livro do Pedro está na minha mesa” em que os artigos são precedidos por preposição que possa contrair-se com os mesmos, já que estes casos, como vimos, mostram-se praticamente categóricos, no sentido de sempre favorecerem a presença do artigo.

Tendência semelhante foi encontrada em Callou e Oliveira e Silva (1997) e Callou (2000), sobre amostras de outras comunidades, o que sugere que essa correlação entre preposição e artigo definido é uma tendência mais geral e sistemática.

Uma hipótese possível é que a influência do input se faça mais presente em estágios incipientes de aquisição e, gradativamente, deixe de ser relevante, o que pode ser examinado através de uma análise que procede ao cruzamento da faixa etária<sup>5</sup> e da configuração estrutural do SN que pode ser vista na tabela 1.

TABELA 1: Cruzamento entre Idade e Estrutura do SN – Estudo Estratificado

<i>Idade</i>	Art + (Mod) + Nprop	Prep + Art + Nprop	Øprep + Art + Nprop	<i>Total</i>
1 ano e 6 meses	-	-	-	-
2 anos	$16/23 = 70\%$	$1/1 = 100\%$	$0/8 = 0\%$	32
2 anos e 6 meses	$6/9 = 67\%$	$1/1 = 100\%$	-	10
3 anos	$8/11 = 73\%$	$4/4 = 100\%$	-	15
4 anos	$16/18 = 88\%$	$3/3 = 100\%$	-	21

<sup>5</sup> Como os resultados provenientes do cruzamento entre a idade da criança e as demais variáveis do estudo longitudinal se revelaram pouco esclarecedores em função, principalmente, de haver células com poucos dados e até mesmo sem dado algum, consideramos, em grande parte da pesquisa, a correlação da idade e demais grupos de fatores do estudo estratificado cujos resultados revelaram com maior clareza a direcionalidade da aquisição.

Como já era esperado, os resultados da tabela 1, mostram que, assim que é incorporada na fala infantil, a preposição leva ao uso categórico do artigo definido diante de N próprio. De forma complementar, a ausência de preposição leva ao cancelamento categórico da preposição. É interessante salientar, no entanto, que a estrutura “ØPrep + Art + Nprop” está circunscrita à faixa correspondente a 2 anos, sinalizando que este tipo de estrutura pode ser decorrente de etapas do desenvolvimento da linguagem.

Na análise qualitativa, já destacamos, considerando o desenvolvimento linguístico de R, que essa estrutura específica fica mais concentrada no período anterior aos 2;00 (um dado em cada idade a saber: 1;06, 1;10 e 2;00), ocorrendo apenas um único caso de estrutura com omissão da preposição aos 4;08. As generalizações apontadas pelo estudo estratificado somadas ao maior detalhamento proporcionado pela análise qualitativa nos possibilitaram depreender, portanto, que a estrutura “ØPrep + Art + Nprop” ocorreu de modo mais expressivo por volta dos 2 anos de idade, podendo ser esta uma fase fundamental do processo aquisitivo em que se observa maior diversificação de estruturas.

Ainda de acordo com a tabela 1, podemos notar que a estrutura “Art + (Atributo) + Nprop” se caracteriza como contexto de maior variabilidade desde a sua emergência no processo aquisitivo. As frequências das faixas etárias correspondentes a 2;00, 2;06 e 3;00 praticamente se mantêm ao longo do desenvolvimento da linguagem: 70%, 67% e 73%, respectivamente. Por outro lado, aos 4;00, a fala infantil demonstra taxa (88%) próxima à da variação encontrada na fala dos adultos (94% nos dois estudos). Tal resultado pode ser uma evidência de que é por volta dos 4;00 que a fala infantil reflete de forma mais próxima os padrões de variação da fala do adulto.

Resultados de outros trabalhos apresentam evidências para essa conclusão. Assim, por exemplo, Roberts (1994), no seu estudo sobre o apagamento variável de /t/ e /d/ finais no inglês e da variação na

produção do -ing já havia verificado que os padrões de variação da fala do adulto já se encontram incorporados na fala de crianças em fase pré-escolar (entre 3 e 4 anos de idade). Ao analisar como as crianças adquirem os padrões de variação estabilizados na fala dos adultos, a autora conclui: (Op. Cit: 176):

Assim como para as regras categóricas, o período pré-escolar parece ser crítico para a aprendizagem das bases para as regras variáveis de apagamento de (t/d) e produção de (ing). Algumas das restrições sobre as regras são refinadas em anos posteriores, mesmo na fase adulta, como foi apontado por Guy e Boyd (1990). Na idade de três ou quatro, entretanto, muitas dessas restrições internas já foram adquiridas, incluindo as dialetamente específicas como a restrição colocada por uma pausa seguinte sobre o apagamento de (t/d), o que insere essas crianças como membros da comunidade de fala da Filadélfia. A aquisição de restrições sociais sobre a variação se inicia na infância remota, mas o volume de acesso a este conhecimento parece acontecer na idade de 4 anos.<sup>6</sup>

## 2.2.2 Forma de recuperação do referente

Nesta seção, investigamos a atuação do processo de recuperação do N próprio antecedido por artigo definido, embora tal grupo não tenha sido selecionado pelo programa estatístico, para checar até que ponto a referência dêitica, apontada como um dos fatores mais favoráveis para a incorporação do artigo definido nos períodos iniciais de aquisição, conforme já mencionamos, mostra-se estatisticamente relevante

---

<sup>6</sup> “As is the case for categorical rules, the preschool period appears to be the critical one for learning the foundations of the variable rules (-t,d) deletion and (ing) production. Some of the constraints on rules are refined in later years, even up through adulthood, as pointed out by Guy and Boyd (1990). By the age of three and four, however, the many of these internal constraints have already been acquired, including the dialect specific following pause constraint on (-t,d) deletion, firmly establishing these children as members of the Philadelphia speech community. The acquisition of social constraints on variation has its beginning in early childhood, but the bulk of this learning appears to take place the age of four.” (Roberts, 1994: 176)

durante o processo aquisitivo de L1. Torna-se relevante ressaltar que os estudos sobre a comunidade não levam em conta a questão da forma de recuperação do referente na análise do uso do artigo definido. Sob este aspecto, nosso trabalho se particulariza, já que testa qual das funções do artigo é mais atuante no contexto variável investigado tanto na fala infantil como na fala do adulto, pelo menos, no que se refere à maneira como se dirige linguisticamente à criança.

Conforme já assinalado na literatura, o artigo definido pode ocorrer diante de um elemento cuja referência é recuperada na situação de fala (extralingüística) ou no plano discursivo, através das relações anafóricas do texto (MOURA NEVES, 2000: 391-392 e LYONS, 1999). No primeiro caso, trata-se de uma referência exofórica e no segundo caso de uma referência endofórica. O artigo pode ser considerado, então, um item fórico já que instrui a busca de recuperação semântica no texto ou na situação.

Ressaltamos que, nesta oportunidade, não vamos nos aprofundar e nem desenvolver um estudo específico sobre *déixis* já que uma discussão dessa natureza fugiria aos propósitos deste trabalho. Vamos discutir apenas os aspectos pertinentes ao estudo da maneira como a forma de recuperação do referente atua no padrão de variação do artigo na fala infantil.

Há um consenso quanto ao fato de que a *déixis* ocupa um espaço central não só no desenvolvimento da linguagem pela criança, como no seu desenvolvimento cognitivo em geral (LEVINSON, 1992 e TOMASELLO, 2003). Um dos primeiros recursos de expressão utilizados pela criança é o de mostrar ou apontar referentes disponíveis na situação de fala seja para nomeá-los seja para chamar a atenção para determinados referentes ou buscar a satisfação dos próprios desejos. Dessa forma, a criança aponta/orienta, por meio linguístico, o elemento da situação imediata ao qual se refere fazendo intervir os objetos perceptíveis e os interlocutores. Segundo Tomasello (Op. Cit: 200), “mesmo antes de aprenderem qualquer língua, as crianças pequenas

são capazes de dirigir a atenção dos outros para objetos exteriores. Elas o fazem principalmente através de gestos, mais frequentemente apontando.”<sup>7</sup>

Como já vimos, o artigo desempenha igualmente importante papel na construção das relações anafóricas do discurso. Assim, neste estudo, controlamos duas situações possíveis no que se refere à forma de ancoragem da referência do N próprio:

Referência *exofórica* - Nome próprio cujo referente, introduzido no discurso pela primeira vez, pode ser identificado na situação de fala.

- (34) (A criança deita no chão ao lado da boneca Barbie)  
Criança: “**Babi** joga a bola”  
Adulto: “Joga!” (MGL 2; 06 p.14)

O nome “Barbie”, citado pela criança, está disponível na situação de fala, mostrando que o referente descrito pode ser identificado diretamente no contexto em que a interação ocorre. Salientamos que o artigo, neste caso, parece exercer a mesma função ostensiva que o demonstrativo “esta”, por exemplo, desempenharia: “**A/Esta Babi** joga a bola”.

Referência *anafórica* – Nome próprio que já foi citado no discurso. Trata-se da referência recuperada no plano linguístico.<sup>8</sup>

- (35) Adulto: “Como que chama o seu titio que operou seu pintinho, A?

---

<sup>7</sup> “Even before they learn any language, young children are able to direct the attention of others to outside objects. They do this mainly gesturally, most often by pointing.” (Tomasello, 203: 200).

<sup>8</sup> É importante ressaltar que o nome próprio que foi mencionado pela segunda vez na cadeia linguística foi classificado como referente anafórico mesmo que estivesse fazendo parte do contexto situacional.

Criança: “Aimô”

Adulto: “Como?”

Criança: “Nocho”

Adulto: “Tio o quê?”

Criança: “Tio”

Adulto: “Gustavo”

Criança: “Foi o **Gustão**”

Adulto: “Foi o Gustavo! O Dr. Gustavo, né?! (GEM 3;02 p.12)

No exemplo acima, o nome “Gustavo” já havia ocorrido na cadeia discursiva. Neste caso, o SN destacado na fala da criança retoma anaforicamente o referente que já havia sido mencionado pelo adulto.

Esperamos encontrar, como já assinalamos, um número maior de ocorrência de artigo definido diante de N próprios com referência exofórica, cujo referente pode ser recuperado na situação comunicativa imediata. Nossa hipótese é que, no início do processo aquisitivo, a função dêitica se torne mais saliente já que a referência está relacionada diretamente com o que está em volta da criança. Trata-se de seres nomeados que estão disponíveis fisicamente. Neste caso, pode-se esperar que a dêixis seja determinante para a aquisição do artigo definido frente a N próprio, dada a sua importância na organização dos sistemas linguísticos e no desenvolvimento da linguagem, como mostra Crystal (2000: 75), para quem:

A noção de dêixis mostrou-se produtiva em diversas áreas da Linguística, principalmente na Pragmática, e nos estudos de aquisição da linguagem, onde os pesquisadores consideram o aprendizado destes itens pelas crianças como um traço significativo do desenvolvimento precoce.

Nossa hipótese encontra evidências também no estudo da aquisição de outros fenômenos variáveis. Assim, Roncarati (2000: 173) mostra,

no seu estudo sobre a gênese e a variação da negação com base em corpora longitudinais, que, na fase pré-linguística, há estreita correlação entre a negação e a dêixis. A autora ressaltou que “uma das estratégias adaptativas mais empregadas pelas crianças é simultaneamente chamar a atenção do interlocutor e apontar para um referente intencionado presente no ambiente imediato” (Op. Cit: 176). Trata-se, segundo a autora, de um recurso comum no estágio icônico no qual a negação é codificada, entre outras coisas, por gesto caracterizante. Para a autora, “a NEG é iconicamente codificada: exibe estreita correlação com a referenciação dêitica do contexto imediato.” (Op. Cit: 173)

Também na aquisição de aspectos não variáveis da gramática, como o uso de preposições, a dêixis desempenha um papel saliente. Ramos (2005;73)<sup>9</sup>, no seu estudo sobre a emergência e incorporação da preposição ‘de’ na fala infantil destaca que:

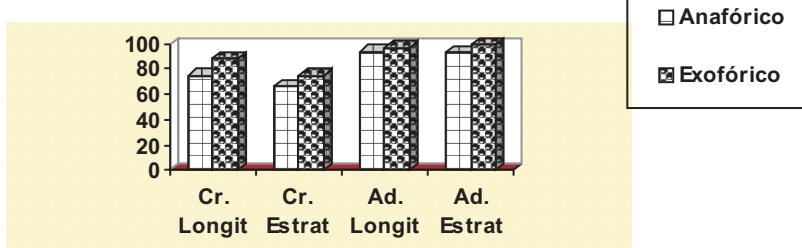
No momento em que emergem as primeiras ocorrências da preposição DE na fala de R (estudo longitudinal) verifica-se um conjunto importante de “progressos” linguísticos. Com 1;07, a criança observada longitudinalmente passa a fazer maior remissão aos elementos do contexto extralinguístico e o gesto de apontar objetos se faz acompanhar mais frequentemente do uso do demonstrativo ‘esse’. Verifica-se também nessa fase a intensificação no uso da negação.

As direções encontradas na escala aquisitiva refletem o padrão encontrado na fala do adulto no tocante à influência da dêixis na ocorrência do artigo, como se pode constatar no gráfico 2.

---

<sup>9</sup> A análise de Ramos (2005) é baseada nos mesmos *corpora* usados em nossa pesquisa (Cf. seção destinada à amostra).

**GRÁFICO 2- Efeito da forma de recuperação do referente**



O paralelismo observável no gráfico 2 pode conduzir à seguinte questão: a atuação da dêixis é decorrente de mera reprodução do input ou trata-se de restrições comuns a etapas do desenvolvimento linguístico da criança?

É interessante ressaltar que o fato de o artigo ocorrer em maior escala na fala dos adultos diante de N próprio cujos referentes são ancorados deiticamente pode ser uma característica da forma como os adultos se dirigem à criança. No entanto, a falta de resultados relativos ao efeito desse fator na comunidade dificulta verificar essa afirmação. Considerando, no entanto, as evidências encontradas em nosso trabalho e nas pesquisas retomadas, há indícios de que, pelo menos em parte, a aquisição no uso do artigo antes de N próprio é influenciada por etapas aquisitivas, o que explica por que um fator que possui efeito menos nítido na fala do adulto intervém de forma mais relevante no padrão de variação da linguagem infantil.

Uma análise do cruzamento entre as variáveis idade e forma de identificação do referente, mostrada na tabela 2, permite ressaltar de forma mais nítida a importância da função dêitica do artigo nos estágios mais incipientes de aquisição. De acordo com as considerações feitas acima e as indicações da análise qualitativa, podemos esperar que, nas faixas etárias iniciais, o artigo ocorra predominantemente antes de nomes com referência exofórica para depois se expandir a núcleos nominais com referência anafórica, de modo a apresentar uma

distribuição de uso ao longo do continuum etário. Essa expectativa encontra evidências favoráveis nos resultados dispostos na tabela 2.

TABELA 2: Cruzamento entre Idade, forma de recuperação do referente e realização do artigo antes de N próprio.

Idade	Referência anafórica	Referência exofórica	Total
1 ano e 6 meses	-	-	-
2 anos	$13/24 = 54\%$	$8/13 = 62\%$	37
2 anos e 6 meses	$3/5 = 60\%$	$4/5 = 80\%$	10
3 anos	$8/12 = 67\%$	$7/8 = 88\%$	20
4 anos	$15/17 = 88\%$	$3/3 = 100\%$	20

A distribuição das frequências da tabela 2 permite depreender a trajetória de uso do artigo em termos da oposição exofórico/anafórico, mantida a cautela necessária em função do desequilíbrio no número de dados por célula. Pode-se observar que o artigo definido emerge por volta de 2;00 com percentual considerável quando se trata de referência exofórica (62%). Na fase correspondente a 2;06, o fator “exofórico” favorece significativamente o uso do artigo (80%) e continua atuando de forma significativa aos 3;00 (88%) para, culminar, aos 4;00, na realização categórica do artigo definido frente a N próprio.

Para o artigo que antecede N próprio que já foi mencionado, por outro lado, o percentual mostra ainda grande variação no início do processo aquisitivo (54%, aos 2;00). Com 2;06, o fator “anafórico” já se mostra um pouco mais relevante para a presença do artigo (60%) e aos 3;00, mantém-se taxa equivalente (67%). Somente no final da escala aquisitiva, aos 4;00, notamos percentual de ocorrência do artigo próximo ao dos adultos (88%).

As tendências verificadas acima evidenciam uma graduação na aquisição dos padrões de variação do artigo de tal modo que ele começa a ser incorporado diante de N próprio com referência exofórica,

funcionado como marcador de referente que faz parte da situação imediata, para, ao longo do *continuum* aquisitivo, expandir seu uso para outros contextos, assumindo, portanto, outras funções.

A análise qualitativa desenvolvida permite acrescentar, no entanto, um aspecto importante: ao que tudo indica, no início do processo aquisitivo, destaca-se a função ostensiva do artigo diante de N próprio. Com o progresso da maturação na língua materna, a criança passa a dominar outras funções realizadas pelo artigo e o papel de “mostração” perde em relevância para dar lugar à ampliação contextual da variação.

Nesta seção, mostramos a forte correlação entre a presença do artigo e a dêixis durante o processo aquisitivo. As evidências encontradas em nossa análise quantitativa confirmam os resultados apontados na análise qualitativa, mostrando haver similaridade entre o comportamento linguístico das crianças do estudo longitudinal e da criança estudada longitudinalmente. Vimos que há indicações de que o artigo é incorporado incialmente como um recurso de ancoragem dêitica, principalmente de ostensão. A dêixis atua, portanto, tanto qualitativa quanto quantitativamente durante a emergência e a variação do artigo definido na fala de crianças em processo de aquisição de L1.

Um certo paralelismo entre os padrões identificados na fala das crianças e na fala dos adultos, embora a forma de recuperação do referente pareça ser menos relevante entre esses, leva a crer que, em situações interativas entre adultos e crianças, que estão adquirindo os padrões da língua materna, a dêixis sobressai como um dos recursos que caracterizam o CDS – Chid-Direct Speech (PINE, 1994). Mas, se, por um lado, o *input* responde por alguns aspectos refletidos na fala da criança, por outro, o efeito significativo da dêixis não pode ser atribuído somente à influência exercida pelo *input*. Questionamos até que ponto o CDS é simplesmente reproduzido pela criança ou se determinados padrões de variação encontrados na fala infantil refletem etapas do processo aquisitivo mais geral, (PINE, Op. Cit), como a ancoragem dêitica, que coloca em relevo o contexto de fala.

### 3 Considerações finais

Nossa pesquisa pôde evidenciar que a direcionalidade apontada na análise qualitativa foi, em grande parte, confirmada pela análise multivariacional. Houve sistematicidade acentuada entre as tendências dos estudos longitudinal e estratificado, bem como houve correspondências entre os padrões de variação da fala do adulto e os padrões de variação da fala infantil em diferentes fases aquisitivas.

Vimos que a presença da “preposição” foi determinante para o uso do artigo já nos estágios iniciais de aquisição, seguindo a tendência dos adultos, de tal modo que torna-se difícil precisar até que ponto o fato de a criança usar “preposição + artigo” pode ser atribuído à falta de segmentação comum ao período inicial de aquisição linguística, ou se a criança está reproduzindo uma tendência disponível no *input*. Verificamos que houve algumas divergências entre os padrões da fala infantil e da fala dos adultos que, ao que tudo indica, evidenciam que outros fatores ligados à aquisição intervêm durante o percurso aquisitivo de processos variáveis. Assim, a função dêitica do artigo (“Forma de recuperação do referente”) assume maior importância no início da escala aquisitiva.

Os resultados de nossa análise puderam evidenciar que há uma expansão gradativa na fala infantil, conforme o avanço da faixa etária, das funções do artigo encontradas no discurso do adulto. Tal reflexo parcial da fala adulta no percurso aquisitivo mostra que a interferência do *input* no processo de aquisição não se limita a uma representação exata da fala do adulto, confirmando conclusões de outros trabalhos que investigaram a aquisição de fenômenos variáveis.

O paralelismo entre os contextos de uso do artigo na fala da criança e os contextos de uso da fala do adulto confirma apenas parcialmente a hipótese da atuação do *input* já que o *continuum* de expansão do artigo parece ser determinado também por fatores maturacionais, como já foi assinalado em pesquisas que se inserem no referencial teórico adotado neste trabalho.

## Referências

ALENCAR, Patrícia Vargas. **Direcionalidade da aquisição do artigo definido frente a N próprio em contexto de input variável**. Rio de Janeiro: UFRJ, 2006. 166 fls., Tese Doutorado em Linguística, Faculdade de Letras, Universidade Federal do Rio de Janeiro.

CALLOU, Dinah. **A variação no Português do Brasil: O uso do artigo definido diante de antropônimo**. Faculdade de Letras da UFRJ, Rio de Janeiro, 2000. Série Conferências, v. 9.

\_\_\_\_\_ & OLIVEIRA e SILVA, Giselle Machline de. **O uso do artigo definido em contextos específicos**. In: HORA, Demeval da (org.). **Diversidade Linguística no Brasil**. João Pessoa: Idéia, 1997. p.11–27.

CRYSTAL, David. **Dicionário de Linguística e Fonética**. Rio de Janeiro: Jorge Zahar, 2000.

GUY, R. Gregory & BOYD, Sally. **The development of a morphological class**. In: KROCH, A. et all (ed). **Language Variation and Change**. Cambridge: Cambridge University Press, 1990. v. 2, n. 1. p. 1 – 18.

LEVINSON, S. C. Deixis. **Pragmatics**. Cambridge: Cambridge University Press, 1992. p. 54 – 96.

LIEVEN, Elena V. M. **Crosslinguistic and crosscultural aspects of language addressed to children**. In: GALLAWAY, Clare & RICHARDS, Brian. **Input and interaction in language acquisition**. Cambridge: Cambridge University Press, 1994. p. 56–73.

LYONS, Christopher. **Definiteness**. Cambridge: University Press, 1999.

MOURA NEVES, Maria Helena de. **Gramática de Usos do Português**. São Paulo: UNESP, 2000.

OLIVEIRA e SILVA, Giselle Machline de. **Realização facultativa do artigo definido diante de possessivo e de patronímico**. In: OLIVEIRA e SILVA, Giselle Machline de & SCHERRE, Maria Marta P. (orgs.) **Padrões Sociolinguísticos** – Análise de fenômenos variáveis do Português falado na cidade do Rio de Janeiro. Rio de Janeiro: Tempo Brasileiro, 1996. p. 119–145.

PINE. Julian M. **The language of primary caregivers**. In: GALLAWAY, Clare & RICHARDS, Brian. **Input and interaction in language acquisition**. Cambridge: Cambridge University Press, 1994. p. 15–37.

PINTZUK, Susan. **Varbrul programs, 1988**. Mimeo

RICHARDS, Brian J. “**Child-direct speech and influences on language acquisition: methodology and interpretation**”. In: GALLAWAY, Clare & RICHARDS, Brian. **Input and interaction in language acquisition**. Cambridge: Cambridge University Press, 1994. p. 74–106.

ROBERTS, Julia Lee. **Acquisition of variable rules: (-t,d) deletion and (ing) production in preschool children**. Tese de Douotrado. Faculties of the University of Pennsylvania, 1994.

RAMOS, Jacqueline Varela Brasil. **Aquisição da preposição “DE” em L1**. Rio de Janeiro: UFRJ, 2005. 198 fls. Tese de Doutorado em Linguística, Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.

RONCARATI, Cláudia. **A gênese das variantes da negação.** **Gragoatá:** revista do Programa de Pós-graduação em Letras, n. 9, Rio de Janeiro: Universidade Federal Fluminense, 2000. p. 171-192.

SNOW. Catherine E. **Beginning from Baby Talk:** twenty years of research on input interaction. In: GALLAWAY, Clare & RICHARDS, Brian. **Input and interaction in language acquisition.** Cambridge: Cambridge University Press, 1994. p. 3-12.

TOMASELLO, Michael. **Constructing a language:** a usage-based theory of language acquisition. Harvard University Press, 2003.

# **PADRÕES DE SELETIVIDADE NA PRODUÇÃO AGRAMÁTICA E DISTINÇÃO ENTRE MOVIMENTOS SINTÁTICOS NA COMPUTAÇÃO *ON-LINE*<sup>1\*</sup>**

Ricardo Joseh LIMA

Universidade Estadual do Rio de Janeiro

Letícia M. SICURO CORRÊA

Pontifícia Universidade Católica do Rio de Janeiro /

LAPAL - Laboratório de Psicolinguística e Aquisição da Linguagem

Marina Rosa Ana AUGUSTO

Pontifícia Universidade Católica do Rio de Janeiro /

LAPAL - Laboratório de Psicolinguística e Aquisição da Linguagem /

Universidade Estadual do Rio de Janeiro.

## **RESUMO**

*Dados do agramatismo são revisados, no que concerne a seletividade do déficit em operações de movimento. Discutem-se duas hipóteses atuais sobre o desempenho agramático. Argumenta-se que as previsões decorrentes de um Modelo Integrado da Computação on-line (Corrêa & Augusto 2006; 2007) são as que mais se ajustam aos perfis agramáticos.*

## **ABSTRACT**

*Current data on agrammatic production are reviewed, concerning the selectivity of the deficit regarding movement operations. Two current hypotheses on agrammatic performance are discussed. It is argued that the predictions derived from an Integrated Model of On-line Computation (Corrêa & Augusto 2006; 2007) can better fit the agrammatic profiles.*

## **PALAVRAS-CHAVE**

*Agramatismo. Movimento. Psicolinguística. Sintaxe.*

---

<sup>1</sup> \*Este artigo foi submetido para publicação em 2007, vinculado ao Grant CNPq (308.713/2005-2), concedido ao terceiro autor.

## KEY WORDS

*Agrammatism. Movement. Psycholinguistics. Syntax.*

## Introdução

A ausência de fluência e o caráter telegráfico e pausado da fala, privilegiando palavras de conteúdo como substantivos e verbos, levaram os primeiros estudiosos dos afásicos de Broca (Pick 1913 *apud* Caplan 1987) a classificá-los como *agramáticos*, baseando-se na idéia de que esses afásicos não possuíam uma gramática que lhes permitisse a fluência e a utilização plena de itens gramaticais, como artigos e conjunções. Essas descrições da fala agramática, entretanto, eram muito genéricas e não apontavam para uma possível seletividade na perda da gramática por esses afásicos. Apenas com a base em uma Teoria Linguística puderam os pesquisadores apresentar uma caracterização mais informativa da natureza das síndromes afásicas. O trabalho pioneiro de Jakobson (1941/1971) foi seguido por outros que, apoiados em modelos linguísticos, propuseram análises do desempenho linguístico de pacientes diagnosticados como agramáticos (Kean 1977, Lapointe 1983, Grodzinsky 1984).

Apesar de todos esses estudos apontarem para a seletividade do déficit na fala agramática, ainda havia a possibilidade de refinar mais essa noção. Tanto Kean (1977) quanto Lapointe (1983) e Grodzinsky (1984), por exemplo, propunham que todo tipo de morfema flexional estaria igualmente afetado na fala agramática. Apenas posteriormente, um estudo conduzido em Hebraico com uma paciente agramática revelou haver distinções entre o grau de comprometimento de morfemas de tempo e morfemas de concordância nos verbos, estando os primeiros caracteristicamente comprometidos (Friedmann e Grodzinsky 1997). A descoberta de que a paciente estudada também possuía problemas com interrogrativas e encaixadas levou os autores a proporem que esse déficit seria mais bem caracterizado fazendo-se menção à estrutura sintática da sentença e não apenas às representações sintáticas dos elementos

morfológicos (como em Grodzinsky 1984).

O trabalho de Friedmann e Grodzinsky (1997) deu início, na literatura sobre produção agramática, a um debate sobre qual seria a melhor caracterização desse déficit. Esses autores basearam sua explicação na estrutura da sentença, propondo que parte dela estaria indisponível na produção agramática, dando origem à *Hipótese da Poda da Árvore (HPA)*. Segundo essa hipótese, o comprometimento afetaria os nós de uma árvore sintática de acordo com sua posição hierárquica, prevendo que os nós mais altos estariam afetados nessa síndrome afásica. Um outro caminho foi focalizar a *operação de movimento sintático*, como na *Hipótese do Problema da Ordem Derivada (HPOD)* de Bastiaanse e van Zonneveld (2005), segundo a qual apenas os movimentos que alteram a ordem básica dos elementos constituem o problema na produção agramática.<sup>2</sup>

Este artigo tem como objetivo avaliar em que medida as manifestações da síndrome do agramatismo disponíveis na literatura encontram-se em conformidade com essas hipóteses, focalizando particularmente, as estruturas que envolvem movimento sintático para projeções funcionais.<sup>3</sup>

---

<sup>2</sup> Há outras hipóteses disponíveis na literatura sobre produção agramática. Essas hipóteses, entretanto, não serão aqui discutidas, por serem específicas para determinados dados (Lee, 2003) ou por não focalizarem construções com movimento, mesmo que indiretamente. Um exemplo desse tipo de hipótese é a Hipótese da Sub-Especificação de Traços, de Wenzlaff e Clahsen (2004). Embora essa hipótese aponte para um mecanismo distinto – a sub-especificação de traços – que possibilitaria uma caracterização dos comprometimentos na produção agramática em diferentes domínios, os trabalhos até então desenvolvidos se restringiram à observação do traço de tempo, o que inviabiliza considerações mais amplas pertinentes a movimento. Vale salientar, ainda, que desde a submissão deste artigo, a hipótese de Bastiaanse e van Zonneveld (2005) foi revista por Burchert et al. (2008). Esta reformulação foi levada em conta em artigo recentemente submetido para publicação (Lima, Corrêa e Augusto, mimeo).

<sup>3</sup> É necessário ressaltar, contudo, que não há um referencial teórico totalmente compartilhado entre os estudos que deram origem aos dados em discussão, ainda que todos remetam a um arcabouço gerativista. Assim sendo, noções pertinentes à estruturação sintática, como a natureza dos núcleos, ordem canônica e movimento, incorporadas nas hipóteses e análises a serem discutidas, podem ser distintas entre estas.

Os dados aqui discutidos são também considerados à luz de um Modelo Integrado da Computação on-line (MIC) (Corrêa & Augusto 2006; 2007<sup>4</sup>), originalmente aplicado a manifestações do Déficit Especificamente Linguístico (DEL) (Corrêa & Augusto, no prelo), tendo em vista que este permite prever dificuldades de acesso a informação gramatical representada no léxico assim como dificuldades na condução de operações com custo computacional alto no processamento de sentenças em tempo real, podendo, portanto, ser aplicado aos padrões de perda e preservação na fala agramática no que concerne aos movimentos sintáticos.

Na seção 2, as hipóteses selecionadas – HPA e HPOD, serão apresentadas. A revisão da base de dados será feita na seção 3. A seção 4 verifica a compatibilidade dos padrões de desempenho agramático identificados com as previsões decorrentes das hipóteses acima. Na seção 5, apresentamos o MIC e demonstramos que este permite dar conta dos padrões de agramatismo pertinentes a movimento. A seção 6 apresenta as considerações finais.

## 1 Produção agramática: contrastando hipóteses

Pesquisas recentes vêm delineando os padrões de produção agramática e algumas hipóteses têm sido elaboradas para dar conta dos fatos observados. Nosso objetivo principal nesta seção é identificar o que essas hipóteses fazem prever com relação ao desempenho de pacientes agramáticos em estruturas que envolvem movimento sintático.

### 1.1 A Hipótese da Poda da Árvore

Como mencionado anteriormente, Friedmann e Grodzinsky (1997) basearam sua hipótese em dados de uma paciente agramática falante do

---

<sup>4</sup> Este modelo foi originalmente denominado Modelo Integrado da Competência Linguística, mas, por conta de mal-entendidos, as autoras optaram por alterar o nome em publicações subsequentes (Corrêa e Augusto no prelo; Rodrigues, Corrêa e Augusto, 2008).

Hebraico que, em situações experimentais, teve bom desempenho na produção de concordância verbal, mas não na produção de morfemas de tempo. Uma explicação morfológica foi descartada pelos autores, que identificaram problemas sintáticos na fala espontânea da paciente, tendo em vista a ausência de complementizadores e a dificuldade na produção de interrogativas e de encaixadas. Dados adicionais de produção eliciada revelaram dificuldades também com sujeitos plenos (em uma tarefa de repetição, apenas 36% dos sujeitos foram pronunciados, um desempenho incompatível com a gramática do Hebraico, que exige a realização fonética do sujeito nas frases testadas) e com a ordenação de cópula e de negação em uma tarefa de anagrama. Desse modo, Friedmann e Grodzinsky propuseram que os nós sintáticos relativos a essas construções estariam afetados na produção agramática. Utilizando a estrutura sintática defendida por Pollock (1989), com um nó para concordância (Agr) situado abaixo do nó de tempo (T), foi proposto que a perda dos nós sintáticos se daria em função de uma métrica de severidade e de hierarquia. Nós mais baixos da árvore estariam preservados e, se um determinado nó estivesse afetado, todos os nós acima deste também estariam. Os autores aproveitaram o termo árvore sintática para sugerir que parte desta estaria ‘podada’ na produção agramática. Tal proposta dava conta do padrão observado na paciente estudada: a concordância verbal estava intacta (o nó Agr está em uma posição baixa na árvore), a produção de morfema de tempo, a produção de sujeito e a ordenação de cópula e negação estavam afetadas (todas envolvendo o nó T, responsável pelo morfema de tempo, por um local para o sujeito, em seu Spec, e por ser o local da cópula). Além disso, a produção de sentenças encaixadas e de interrogativas também estava comprometida (ambas envolvendo o nó C). A Hipótese da Poda da Árvore (HPA) é, portanto, uma proposta relativa à disponibilidade de locais para relações sintáticas e movimentos – a operação movimento não estaria afetada na produção e somente apresentaria problemas na fala se envolvesse algum nó afetado pela poda.

Trabalhos posteriores (Friedmann 2001, 2002; Ruigendijk et al. 2004) foram realizados a fim de verificar se a HPA estava no caminho certo. Friedmann (2001) observou que agramáticos do Hebraico e do Árabe tinham problemas sistemáticos em repetir e construir sentenças encaixadas que envolvem C. Friedmann (2002) investigou a produção de interrogativas por agramáticos falantes de Inglês e de Hebraico. O paciente agramático cuja língua era o Inglês tinha problemas na produção de interrogativas QU-- e de interrogativas sim/não; os afásicos cuja língua era o Hebraico demonstraram problemas apenas na produção de interrogativas QU-. Tal distinção de padrões pode ser explicada pela análise das construções interrogativas nessas línguas: em ambas, a interrogativa QU- envolve movimento do elemento QU- para Spec C, mas apenas no Inglês a interrogativa sim/não envolve o nó C, com movimento do auxiliar (Do you like pizza?). No Hebraico, esse tipo de interrogativa é produzido sem o envolvimento do nó C (Você gosta de pizza?). Essa diferença faz com que os agramáticos falantes de Hebraico não tenham problema com esse tipo de interrogativa, o que se verifica com o falante de Inglês<sup>5</sup>. Ruigendijk et al. (2004) investigaram a produção de interrogativas por agramáticos falantes de Holandês. Nessa língua, tanto a interrogativa QU- quanto à interrogativa sim/não envolvem o nó C (diferentemente do Inglês em que o auxiliar está em C, no Holandês é o verbo principal que se encontra nessa posição nas interrogativas – Gosta você de pizza?). Os resultados, conforme esperado pela HPA, mostraram que os agramáticos tiveram problemas com os dois tipos de interrogativa.

Assim, em relação a movimentos sintáticos, a HPA prevê que somente os que envolvem os nós afetados pela poda estejam comprometidos. A

<sup>5</sup> É importante observar-se que o nó C incorpora traços semânticos pertinentes à força ilocucionária da sentença, o que implica que mesmo que a interrogativa não envolva movimento, a expressão da força ilocucionária na entonação pressupõe um nó C. A autora reconhece esse ponto e propõe algumas alternativas para dar conta do bom desempenho dos agramáticos do Hebraico nessa construção: a presença do operador de interrogativa em um nó mais baixo nessa língua, ou a possibilidade de os agramáticos alocarem esse operador em um nó mais baixo ou mesmo de não o representar.

hipótese possui um espaço para variação na medida em que prevê uma escala de severidade: poda em C, agramático menos severo, poda em T, o perfil mais comum ou poda em Agr, o caso mais severo.

## 1.2 A Hipótese do Problema da Ordem Derivada

Bastiaanse e colaboradores (Bastiaanse e van Zonneveld 1998, Zuckerman, Bastiaanse e van Zonneveld 2001) inicialmente conduziram uma série de experimentos com agramáticos falantes do Holandês para verificar se o problema que estes agramáticos possuíam seria com a produção da flexão, com o movimento de verbo ou com o local do verbo na árvore sintática. O Holandês é uma língua SOV que possui movimento de verbo nas sentenças matrizas declarativas, resultando em uma ordem SVO. Nas orações encaixadas, o verbo permanece *in situ* e a ordem é SOV. Os agramáticos não demonstraram qualquer dificuldade na produção da flexão de verbos nas encaixadas, revelando que o problema não está no nível morfológico. As dificuldades se restringiram à produção da flexão em situações em que o verbo está movido. Em tarefa de preenchimento de lacuna, em que o alvo era a produção de verbos flexionados na posição medial (SVO), ou seja, em sentenças matrizas, os verbos foram produzidos na forma não finita. Em tarefa de finalização de enunciado a partir de preâmbulo, os pacientes optaram por produzir verbos flexionados em posição final (SOV), quando a posição medial seria a esperada. Esses resultados são interpretados como decorrentes de dificuldade com movimento e não com a flexão de tempo, o que levou à conclusão de que o movimento de verbo é a operação afetada na produção agramática.

Um estudo posterior (Bastiaanse e Thompson 2003) revelou que agramáticos falantes de Inglês não tinham problemas com a produção de auxiliares em T, o local onde são gerados, mas apenas com auxiliares movidos para C. A disponibilidade de nós que a HPA prevê que estejam afetados não foi o único problema que esses autores encontraram para essa hipótese. Bastiaanse, Koekkoek e van Zonneveld (2003)

observaram que agramáticos não conseguiam realizar com sucesso uma tarefa que exigia a produção de sentenças com o objeto movido para Spec  $\text{Agr}_O$  (*scrambling*) em Holandês. Duman, Aygen e Bastiaanse (2005) encontraram um perfil semelhante para agramáticos falantes de Turco: as sentenças relativas que envolviam movimento do sujeito para uma posição baixa da árvore<sup>6</sup> apresentaram dificuldade para os agramáticos. Crucialmente, as sentenças matrizes, cuja ordem é SOV sem movimento de verbo, foram produzidas sem dificuldades, revelando que o nó T está intacto, já que o sujeito é pronunciado, tem Caso estrutural e o verbo está flexionado. Assim, tanto a previsão da HPA de que os nós mais baixos da árvore devem estar intactos como a de que, uma vez um nó estando afetado, todos os acima também estão não foram confirmadas: os movimentos afetados no Turco e no Holandês envolviam nós mais baixos ( $\text{Agr}_O$  e Asp). Além disso, constatou-se que, no Turco, o problema com o nó mais baixo (Asp) não se refletiu em um nó mais alto (T).

Bastiaanse e van Zonneveld (2005) reformularam a hipótese que identificava o movimento de verbo como central para o déficit na produção agramática, estendendo-a para qualquer movimento que afetasse a ordem básica da língua. Essa reformulação, chamada de *Hipótese do Problema da Ordem Derivada* (HPOD), conseguiria dar conta: dos problemas dos agramáticos falantes de Holandês relativos à ordem SVO da matriz e nos movimentos *scrambling*; das dificuldades dos agramáticos do Inglês com o auxiliar movido nas interrogativas, resultando em AuxSVO; e dos problemas observados nas relativas do Turco. Assim, em relação a movimentos sintáticos, a HPOD prevê que apenas os movimentos que afetem a ordem básica da língua sejam problemáticos para os agramáticos.<sup>7</sup>

<sup>6</sup> Esse nó provavelmente é Asp, mas crucialmente não T ou C, já que nas relativas do Turco assume-se que não há T nem C, uma vez que os verbos estão na forma não-finita e não há complementizador (Aygen, 2004).

<sup>7</sup> Deve-se, no entanto, notar que o que os autores consideram como ordem básica não necessariamente corresponde ao conceito usual de *ordem canônica*. Este conceito, mais utilizado em

## 2 Tipos de movimentos sintáticos e seus padrões de seletividade na produção agramática

Nesta seção, vamos apresentar os padrões de seletividade na produção agramática referentes a movimentos sintáticos. Primeiramente, abordaremos os movimentos de núcleo e em seguida os movimentos de sintagma.

### 2.1 Movimento de Núcleo

#### 2.1.1 Movimento de V para T

Os dados mais claros sobre esse tipo de movimento vêm do estudo de Lonzi e Luzzatti (1993) com agramáticos do Italiano e do Francês. Nessas línguas, o movimento de V para T é aberto e resulta na ordenação do verbo antes da negação assim como antes do advérbio (“*Jean embrasse souvent Marie*”, “*Jean n'aime pas Marie*”). Por meio de um teste de anagrama feito com cartões, os agramáticos ordenaram corretamente verbo, negação e advérbio em 41 de 42 situações. Para corroborar o padrão obtido, Lonzi e Luzzatti investigaram a fala espontânea de outros agramáticos do Francês e do Inglês e problemas com a ordenação do verbo em relação aos elementos citados não foram encontrados.

O Grego também é uma língua que exibe movimento aberto de V para T e a análise da fala espontânea de dois agramáticos falantes do Grego revelou que eles não tiveram qualquer problema com esse tipo de movimento (Stavrakaki e Kouvava 2003). Dados de fala espontânea do Grego também confirmam que movimento de V para T está intacto.

---

estudos psicolinguísticos, diz respeito à ordenação fixada como semanticamente neutra em sentenças assertivas e pode levar em conta fatores como frequência de uso na língua. *Ordem básica*, tal como considerada pelos autores, é a ordem que não adviria de movimento sintático. No caso do Holandês, a ordem básica seria a das sentenças **encaixadas**, e a das matrizes seria derivada daquela. É importante observar, contudo, que essa assunção remonta a um estado da teoria linguística anterior a Kayne (1994), quando uma ordem SVO básica e universal passa a ser assumida, implicando que qualquer que seja a ordem da língua esta é derivada de movimento.

Apenas Friedmann e Grodzinsky (1997) relatam um teste em que uma paciente agramática do Hebraico teve problemas com a ordenação de elementos envolvendo T. No entanto, o teste envolveu cópulas, que não são verbos movidos para T, mas sim inseridos em T. Além disso, os resultados obtidos não foram claros: algumas cópulas no Hebraico são precedidas pela negação, assim como os verbos. O teste aplicado com verbos não apresentou problemas (4% de erros) e não deveria apresentar também nos casos de cópulas que vêm depois da negação, mas esse detalhamento não foi mostrado (apenas a taxa de erros na ordenação de cópulas, 76%). Considera-se, portanto, que *o movimento de V para T está intacto na produção agramática*.

### 2.1.2 Movimento de V para C em declarativas

O movimento de V para C em sentenças declarativas tem sido um dos tópicos mais discutidos na literatura sobre produção agramática. Estudos divergem quanto à preservação desse movimento em agramáticos do Holandês e do Alemão. Nessas línguas, considera-se que a ordem básica é SOV, encontrada em sentenças declarativas encaixadas (“*Ik denk dat de boer de koe melkt*”, ‘Eu acho que o camponês ordenha a vaca’), e que a ordem nas matrizes declarativas, SVO, é resultado do movimento do verbo para C (“*De boer melkt de koe*”, ‘O camponês ordenha a vaca’).

Os dados de fala espontânea do Holandês e do Alemão (Kolk e Heeschen 1992, de Roo 2001, Bastiaanse e van Zonneveld 1998, Burchert *et al.* 2005) indicam que quando um verbo flexionado é produzido, ele aparece na segunda posição da sentença, ou seja, movido de sua posição de origem, que é a final. Dados de teste com falantes do Alemão (Wenzlaff e Clahsen 2005) e do Holandês (Kok *et al.* 2006) apontam para o mesmo padrão. Ou seja, esses resultados sugerem que o movimento de V para C se encontra preservado no quadro do agramatismo.

São os estudos de Bastiaanse e van Zonneveld (1998), Zuckerman, Bastiaanse e van Zonneveld (2001) e Bastiaanse e Thompson (2003) que apontam para um padrão inverso. Os resultados obtidos com agramáticos falantes do Holandês foram interpretados como indicativos de problemas associados a esse tipo de movimento sintático, uma vez que não houve evidência de verbos flexionados em posição medial. Diferenças na metodologia de avaliação podem, no entanto, explicar os resultados conflitantes. Wenzlaff e Clahsen (2005) e Bastiaanse e van Zonneveld (1998) usaram o paradigma do preenchimento de lacunas. Enquanto, contudo, no estudo de Wenzlaff e Clahsen (2005), os agramáticos eram solicitados a completar sentenças em que duas lacunas possibilitavam uma escolha com relação ao posicionamento do verbo, no estudo de Bastiaanse e van Zonneveld (1998), apenas uma lacuna era apresentada em posição medial. Assim sendo, a posição estrutural do verbo produzido no estudo de Bastiaanse e van Zonneveld (1998) não refletia a efetiva formulação do enunciado pelo paciente, o que impede que se tomem esses resultados no contraste de hipóteses sobre o comprometimento de movimento de verbo para posição medial em pacientes agramáticos falantes de Holandês.

Diferenças metodológicas em tarefas de finalização de sentenças também dificultam uma comparação entre os resultados obtidos por Kok et al. (2006) e por Zuckerman, Bastiaanse e van Zonneveld (2001). Enquanto na tarefa utilizada por Kok et al. (2006), o paciente deveria ordenar cartões com palavras/sintagmas correspondentes a elementos de uma oração coordenada (SVO) ou de uma oração encaixada à direita (SOV), na tarefa de finalização de sentenças de Zuckerman, Bastiaanse e van Zonneveld (2001), o paciente deveria produzir oralmente uma oração coordenada ou encaixada correspondente ao evento descrito numa figura, a partir de um formato de enunciado fornecido pelo experimentador. Assim sendo, a produção ficou menos restritiva, no estudo de Zuckerman, Bastiaanse e van Zonneveld (2001), dando margem à elipse do verbo (em posição medial) nas orações coordenadas.

Elipses do verbo corresponderam, de fato, à maior parte das respostas para este tipo de sentença e foram consideradas erros, ainda que correspondessem a possibilidades gramatical e pragmaticamente válidas no contexto do experimento. Não é possível, pois, concluir que esses pacientes tenham comprometimentos vinculados a movimento de verbo para posição medial.

Em suma, consideramos aqui, como o fazem Kok *et al.* (2006), que os resultados encontrados até então não apresentam evidência sugestiva de que o movimento de V para C em declarativas esteja afetado na produção agramática. É possível assumir, portanto, que *o movimento de V para C em declarativas está intacto na produção agramática*.

### 2.1.3 Movimento de V-Aux para C em interrogativas

Tanto o Inglês quanto o Holandês são línguas que apresentam movimento aberto para C em sentenças interrogativas do tipo sim-não (“*Do you like pizza?*”, “*Melkt de boer de koe?*”). No Inglês, o movimento é feito pelo auxiliar e no Holandês pelo verbo principal. Três estudos (Friedmann 2002, Bastiaanse e Thompson 2003, Ruigendijk *et al.* 2004) investigaram esse tipo de construção nessas línguas e encontraram o mesmo padrão: dificuldade na produção do elemento movido. Respostas frequentes dos agramáticos foram perguntas na ordem SVO, sem movimento, e fragmentos de perguntas, ora sem o verbo, ora sem sujeito ou objeto. A ausência de dificuldades por parte de agramáticos falantes de Hebraico com sentenças interrogativas sim-não (Friedmann 2002) pode ser explicada pelo fato de esse tipo de interrogativa ser construído sem movimento, apenas com entonação ascendente, tal como em Português.

Burchert *et al.* (2005) investigaram a produção de interrogativas sim-não do Alemão, uma língua que se comporta como o Holandês. Dos oito agramáticos, apenas três não tiveram problemas com essa construção. Apesar de não ser um resultado tão consistente quanto os anteriores, este também indica que o movimento para C em interrogativas é problemático

para a produção agramática. Assim sendo, considera-se que *o movimento de V para C em interrogativas está afetado na produção agramática*.

## 2.2 Movimento de Sintagma

### 2.2.1 Movimento de DP para Spec

Não existem estudos na literatura sobre produção agramática especificamente voltados para o movimento de DP para Spec T, o movimento realizado pelo DP de sua posição originária em Spec *v* e que é feito para apagar o traço-EPP de T (Chomsky 1999). Friedmann e Grodzinsky (1997) observam que o fracasso da paciente agramática falante do Hebraico em repetir sentenças SVO, omitindo, de modo ilegal na gramática do Hebraico, o sujeito em 36% dos casos, seria um indicador de que a posição de especificador de T não estaria disponível na produção agramática. Entretanto, se tal afirmação fosse aplicada de modo geral a todos os agramáticos, poderíamos esperar pelo menos três situações: (i) uma alta taxa de sujeitos nulos, mesmo em sentenças finitas, inclusive em línguas que não permitem sujeitos nulos; (ii) a designação de Caso *default* ao DP sujeito, visto que o Caso estrutural não estaria disponível (ver Wexler 1998 para análise de aquisição) e (iii) problemas na ordenação do DP sujeito com verbo, advérbio e negação em línguas em que V se move para T.

A previsão (i) é apenas verificada em sentenças não-finitas. de Roo (2001) observou uma alta taxa de sujeitos nulos no Holandês, mas apenas quando o verbo não está flexionado; nas sentenças finitas, a taxa de produção do sujeito é normal. Agramáticos de todas as línguas investigadas até o momento sem dúvida produzem menos sujeitos explícitos do que indivíduos neurologicamente saudáveis, mas nenhum estudo estabelece como característica central do agramatismo a omissão de sujeitos. A previsão (ii) não é verificada nos agramáticos falantes de Inglês e de Francês, por exemplo. Nessas línguas, o Caso *default* do sujeito difere do Caso estrutural, sendo este Acusativo no Inglês e

Dativo no Francês. Contrariamente ao que se encontra em estudos de aquisição, não se verificam formas como “*Him go*” nem “*Moi vais*” na produção de agramáticos falantes de Inglês e Francês, respectivamente. De alguma forma, portanto, o Caso estrutural está disponível para o DP sujeito, e provavelmente a atribuição de Caso deve ser feita via T, com o sujeito em Spec T. Por fim, em 3.1.1 já observamos que agramáticos não possuem problemas com ordenação de elementos envolvendo T, como previsto em (iii), sendo o caso do Francês significativo: a construção “*je m'en rappelle plus*” dificilmente pode ser analisada sem que se assuma que o verbo está em T e o sujeito em Spec T. Assim sendo, considera-se que o movimento de DP para Spec T está intacto na produção agramática.

## 2.2.2 Movimento QU- em sentenças interrogativas

A dificuldade de afásicos agramáticos em produzir, em fala espontânea ou em testes, sentenças interrogativas com elemento QU-movido é atestada em vários estudos e em várias línguas. Friedmann e Grodzinsky (1997) e Friedmann (2002) investigaram esse fenômeno no Hebraico; Milman, Dickey e Thompson (2004), em Inglês; Ruigendijk et al. (2004) no Holandês; van der Meulen, Bastiaanse e Rooryck (2002), no Francês e Burchert, Swoboda-Moll e de Bleser (2005) no Alemão. Todos esses estudos investigaram sentenças com QU- objeto se movendo para a primeira posição da sentença e também adjuntos como elemento interrogativo. Não houve diferenças significativas em relação à produção de QU- objeto e de adjuntos.

A situação é menos clara em relação às interrogativas com QU- sujeito. Apenas Friedmann (2002) detalhou esse tipo de construção, indicando que também está afetada. Burchert et al. (2005) trabalharam com essa construção mas não remetem a ela nos resultados, deixando implícito que o padrão para esse tipo de construção se assemelha ao padrão das demais interrogativas com elemento QU-. Os demais estudos citados não explicitam se trabalharam ou não com interrogativa QU- sujeito.

Os trabalhos de Penke (Penke 2001, Neuhaus e Penke 2003) apresentam um panorama diferente para interrogativas com elemento QU- em agramáticos falantes de Alemão. Contrariamente aos estudos acima citados, o desempenho dos pacientes nas tarefas de pergunta-resposta utilizadas foi normal. Como a metodologia empregada foi semelhante à dos demais estudos, uma possível explicação para esse perfil destoante dos demais pode estar nos critérios utilizados para seleção e classificação de pacientes. Os pacientes investigados por Penke produzem não só interrogativas QU- como também sentenças encaixadas, com movimentos de sintagmas, e outras estruturas complexas. Esse é um perfil de desempenho que não é encontrado nos agramáticos dos demais estudos, o que levou ao questionamento dos critérios para o diagnóstico de agramatismo para esses pacientes (cf. Wenzlaff e Clahsen, 2005).<sup>8</sup>

Uma vez que a maior parte dos casos sugere comprometimento no movimento QU-, assumiremos aqui que o movimento QU- está afetado na produção agramática.

### **2.2.3 Sentenças relativas**

Os dados disponíveis sobre produção de sentenças relativas por agramáticos são bem mais escassos, se comparados com os dados sobre produção de interrogativas com elemento QU-. Os estudos de fala espontânea observam a ausência desse tipo de construção na fala agramática e apenas Friedmann (2002) e Duman, Aygen e Bastiaanse (2005) investigaram-na em situações experimentais. As tarefas de Friedmann (2002) envolviam repetição de sentenças e finalização de enunciados. Nos dois casos, os agramáticos, falantes do Hebraico, tiveram problemas em construir sentenças relativas, optando às vezes por não completar a sentença, por omitir o complementizador ou

---

<sup>8</sup> É possível, no entanto, conceber-se que as manifestações comportamentais do agramatismo não sejam completamente homogêneas entre pacientes, considerando-se que a área da lesão ou o grau de comprometimento do cérebro dificilmente serão idênticos entre estes.

por produzir apenas fragmentos das sentenças. Duman *et al.* (2005) utilizaram a mesma metodologia de Zuckermann *et al.* (2001) com agramáticos falantes do Turco, que obtiveram um fraco desempenho na tarefa com as relativas. Deve-se ressaltar, no entanto, que essa metodologia apresenta problemas, os quais foram abordados em 3.1.2.

Assim, considera-se que *sentenças relativas estão afetadas na produção agramática*.

Em suma, os padrões observados no comportamento de pacientes agramáticos, com relação a tipos de movimento sintático estão esquematizados no Quadro 1 abaixo.

QUADRO 1: Padrões da produção agramática em estruturas que envolvem movimento sintático

Tipo de movimento	Tipo de desempenho	
	Intacto	Afetado
(1) movimento V-T	✓	
(2) movimento V-C (declarativas)	✓	
(3) movimento V-C (interrogativas)		✓
(4) movimento de DP para Spec T	✓	
(5) movimento QU-		✓
(6) movimento em relativas		✓

### 3 Verificando as previsões das hipóteses e os padrões de seletividade obtidos

A partir dos padrões de produção agramática discutidos na seção anterior, é possível proceder a uma análise dos mesmos em função das previsões que podem ser formuladas a partir das hipóteses aqui consideradas.

### **3.1 Hipótese da Poda da Árvore**

A Hipótese da Poda da Árvore prevê graus de severidade distintos, podendo dar conta de padrões que aparentemente seriam contrários à sua proposta. O caso do movimento de V para T é um desses padrões. Observando os dados que levaram à formulação original da hipótese de Friedmann e Grodzinsky (1997), notamos que a manifestação mais comum de agramatismo teria a árvore podada em T. Desse modo, espera-se que o movimento de V para T e construções envolvendo T estejam afetados, o que contraria o padrão discutido na seção 3 e observado no Quadro 1. A HPA, ao lidar com graus de severidade, permite, não obstante, supor a existência de agramáticos que têm a árvore podada apenas em C. Nesse caso, o movimento de V para T e as construções envolvendo T estariam intactos, em conformidade com o padrão encontrado. O padrão (2), movimento intacto de V para C em declarativas, é, no entanto, problemático para a HPA, que prevê que pelo menos o nó C esteja afetado na produção agramática. Essa hipótese, por outro lado, dá conta do padrão (3), movimento afetado de V/Aux para C em interrogativas. O raciocínio empregado para analisar o padrão (1) pode ser aplicado na análise do padrão (4). Os padrões (5) e (6), por envolverem o nó C, podem ser previstos pela HPA, para pacientes agramáticos com qualquer grau de severidade.

### **3.2 A Hipótese do Problema da Ordem Derivada**

A HPOD tem dificuldade de dar conta dos padrões (1) e (2). São casos de movimento aberto, cujo resultado é uma ordem distinta da ordem inicial da derivação da sentença. A HPOD preveria que esse tipo de movimento estaria afetado na produção agramática, o que não está de acordo com aqueles padrões. Já os padrões (3), (5) e (6), que também apresentam movimento aberto, são explicados pela HPOD. Essa hipótese não permite, contudo, que se façam previsões relativas ao padrão (4).

O Quadro 2 sintetiza as relações entre padrões de agramatismo, no que concerne à operação de movimento, e hipóteses sobre agramatismo estabelecidas nesta seção.

QUADRO 2: Compatibilidade entre os padrões de desempenho agramático relativo a movimento sintático e hipóteses sobre agramatismo

Padrões de desempenho agramático	Compatibilidade	
	HPA	HPOD
(1) De V para T intacto	*	Não
(2) De V para C em declarativas intacto	Não	Não
(3) De V para C em interrogativas afetado	Sim	Sim
(4) De DP para Spec T intacto	*	--
(5) QU em interrogativas afetado	Sim	Sim
(6) QU em sentenças relativas afetado	Sim	Sim

\* Dependente do grau de severidade

#### 4 Dois tipos de movimento sintático

Corrêa & Augusto (2006, 2007) apresentam um modelo da computação sintática conduzida em tempo real, no qual uma derivação linguística concebida na ótica do Programa Minimalista (PM) é caracterizada de modo a ser incorporada em modelos de produção e de compreensão de sentenças, no que concerne à formulação e ao parsing de enunciados linguísticos. Essa proposta retoma a discussão acerca da possibilidade de unificação entre o processador linguístico e um modelo de língua, re-introduzida com base na arquitetura do sistema cognitivo da língua proposta pelo PM (cf. Harkema, 2001; Phillips, 2003; Fong, 2005). A partir dos argumentos de Corrêa (2005a, b, 2008), o modelo proposto busca solucionar dois tipos de dificuldade que impediriam uma completa identificação entre derivação linguística, parser, e formulador sintático: (i) a direcionalidade da derivação, uma

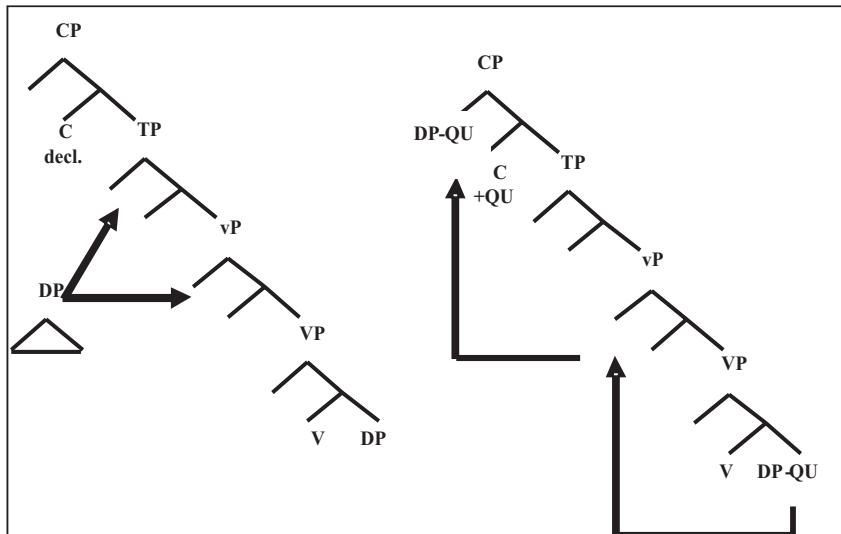
vez que na derivação minimalista adota-se uma perspectiva bottom-up, deflagrada pelas exigências dos núcleos sintáticos, o que se mostra pouco compatível com a idéia de processamento incremental, da esquerda para a direita, ao longo do tempo e (ii) a necessidade de se distinguir alguns tipos de movimento sintático, uma vez que parece não haver custo computacional mensurável associado a alguns destes.

Em relação à questão da direcionalidade, o MIC se caracteriza como um modelo de natureza mista. Na formulação de enunciados, assume-se que o acesso a traços de elementos de categorias funcionais, em função de um estado mental responsável por uma intenção de fala, dá origem a uma derivação top-down de estruturas com núcleos funcionais que definem o esqueleto sintático dos domínios sentencial (CP), verbal (TP) e nominal (DP) ao mesmo tempo em que codificam linguisticamente força ilocucionária, referência no tempo e referência a entidades. Nestes se acoplam estruturas constituídas de forma bottom-up a partir de informação relativa à estrutura argumental de elementos pertencentes a categorias lexicais, mantendo-se o pressuposto de que projetar a sintaxe do léxico garante a satisfação de restrições de seleção na derivação linguística (Baker, 1988; Chomsky, 1970; 1981). Assim sendo, NPs construídos bottom-up se acoplam a estruturas correspondentes a projeções máximas de D, construídas em espaços derivacionais paralelos. Os DPs resultantes se acoplam às posições de complemento ou de sujeito na árvore derivada top-down a partir de C, contendo T como núcleo mais baixo, o qual toma vP (derivado a partir da divisão dos traços de V). O modelo propõe que esses DPs sejam posicionados de acordo com a ordem canônica da língua, uma vez que conhecimento relativo a esta (representado pelo traço EPP) se encontra imediatamente disponível para o falante da língua.

Esse tipo de derivação mista implica distinção entre movimentos sintáticos, que é feita em função de sua natureza e do custo computacional mensurável associado a estes. Os movimentos são, assim, subdivididos em dois grupos: (i) aqueles que dão origem à ordenação característica

de uma dada língua, tomados como expressão formal das operações correspondentes à fixação de parâmetros relativos a ordem na aquisição da linguagem, cuja fixação precoce (Christophe et al., 2003) implicaria o acionamento automático de informação relativa à posição canônica dos constituintes na árvore sintática. Esse tipo de movimento não precisaria, portanto, ser computado a cada emissão linguística; (ii) aqueles que alteram a ordem canônica da língua e que são motivados por demandas específicas de uma dada situação de discurso. Em função dessas demandas específicas, sua computação deve se realizar na formulação/ parsing do enunciado em questão, isto é, on-line. Assim sendo, enquanto os primeiros (movimento de núcleo, de sujeito para SpecT, de objeto para Specv nas línguas OV) descrevem operações que transcorrem no curso inicial da aquisição da língua e não parecem acarretar custo computacional mensurável no processamento linguístico por parte do adulto, os movimentos do segundo tipo (que envolvem formação de interrogrativas e relativas, construções de foco e de topicalização, assim como passivas) descrevem operações que transcorrem em tempo real e que têm custo de processamento mensurável (Wanner & Maratsos, 1978; Zurif et al., 1993; Fiebach, Schlesewsky & Friederici, 2002; Felser, Clahsen & Munte, 2003). De modo a formalizar essa distinção, o MIC apresenta os primeiros por meio de cópias simultâneas do constituinte movido para sua posição canônica. No caso do DP sujeito, essas cópias simultâneas garantem a atribuição de papéis temáticos pelo verbo ao mesmo tempo em que possibilitam o posicionamento imediato do sujeito na posição em que este se apresenta em sentenças assertivas neutras da língua. Os movimentos do segundo tipo, por outro lado, são caracterizados no modelo por meio de cópias sequenciadas do elemento movido, de modo a explicitar o custo computacional da operação.

FIGURA 1: Cópias simultâneas no posicionamento do sujeito e cópias sequenciadas no movimento QU-



Para a discussão aqui empreendida, é relevante explorarmos a extensão que a proposta de distinção entre tipos de movimento alcançaria para a caracterização de comprometimentos no agramatismo. Na perspectiva do PM, segundo a qual todo o conhecimento adquirido encontra-se representado no léxico, as perdas agramáticas podem ser, em princípio, atribuídas a uma dificuldade de acesso a informação pertinente aos traços formais da língua, ou à própria implementação da computação sintática, considerando-se o custo operacional da mesma. Diante da distinção entre tipos de movimento sintático, é possível formular a hipótese de que as operações de movimento conduzidas on-line, as quais implicam custo computacional, estão afetadas nessa síndrome afásica. Nesse caso, podem-se encontrar perfis de agramatismo distintos em diferentes línguas, visto que um dado constituinte pode ser movido para uma determinada posição em função da fixação de parâmetros de ordem numa língua e em função de demandas discursivas

em outra. Por exemplo, o movimento de verbo para C no Alemão (ou Holandês) nas sentenças matrizes estaria preservado, segundo essa hipótese, tendo em vista que se trata de movimento pertinente à fixação de parâmetros de ordem – a ordem canônica das sentenças matrizes. Por outro lado, é possível que movimento de verbo para C esteja afetado nas interrogativas do Português Europeu, tendo em vista que este seria resultado de uma exigência das propriedades do traço de QU nessa língua, recuperado em função de demandas discursivas específicas. Observe-se que as previsões do MIC distinguem-se das oriundas da HPOD em dois sentidos: em primeiro lugar, considera-se que toda a ordem canônica advém da fixação de parâmetros, sendo, portanto, derivada ou de uma ordem universal, como proposto por Kayne (1994), ou de uma ordem inicial subspecificada, tendo em vista que ordem apresenta-se como uma exigência da interface fônica. Em segundo lugar, considera-se que a língua pode definir ordens canônicas distintas para sentenças encaixadas e matrizes, as quais são identificadas no curso da aquisição da linguagem.

Consideramos que um entendimento das manifestações da produção de agramáticos requer que se conceba um modelo integrado da competência linguística, no qual se distinga aquilo que pode ser tomado como potencial gerativo e conhecimento adquirido (mantido de forma estável na memória) e a computação efetivamente conduzida em função de uma intenção de fala.

Os padrões apresentados no Quadro 1 parecem estar em conformidade com essa previsão ao mesmo tempo em que nenhuma das hipóteses aqui discutidas consegue dar conta de todo o tipo de evidência já obtida (Ver quadro 3) .

QUADRO 3: Compatibilidade entre os padrões de desempenho agramático relativo a movimento sintático e hipóteses sobre agramatismo e o MICL

Padrões de desempenho agramático	Compatibilidade		
	HPA	HPOD	MIC
(1) De V para T intacto	Dependente do grau de severidade	Não	Sim
(2) De V para C em declarativas intacto	Não	Não	Sim
(3) De V para C em interrogativas afetado	Sim	Sim	Sim
(4) De DP para Spec T intacto	Dependente do grau de severidade	—	Sim
(5) QU em interrogativas afetado	Sim	Sim	Sim
(6) QU em sentenças relativas afetado	Sim	Sim	Sim

## 5 Conclusão

Este artigo teve como objetivo avaliar em que medida a seletividade nas manifestações da síndrome do agramatismo relativas a movimento sintático pode ser prevista por hipóteses que buscam explicar a produção agramática. Demonstrou-se que a Hipótese da Poda da Árvore (Friedmann e Grodzinsky 1997) assim como a Hipótese do Problema da Ordem Derivada (Bastiaanse e van Zonneveld 2005) não dão conta dos padrões de comprometimento de movimentos sintáticos, quando se contrastam dados de diferentes línguas. Essas hipóteses partem do pressuposto de que as operações sintáticas caracterizadas em uma gramática gerativa são efetivamente computadas na produção e na compreensão de enunciados. Consideramos, entretanto, que para

explicar o desempenho linguístico à luz de um modelo de língua é necessário que este seja integrado em modelos de processamento linguístico. Nesse sentido, o Modelo Integrado da Computação on-line (Corrêa e Augusto 2006; 2007), ao incorporar uma derivação minimalista na formulação sintática e no *parsing* de enunciados na produção e na compreensão de sentenças, respectivamente, permite que se considere em que medida operações descritas em uma derivação linguística precisam ser efetivamente conduzidas em tempo real. Um modelo de língua explicita numa derivação linguística operações de movimento que visam a dar conta da variabilidade das línguas com relação à ordenação de constituintes sintáticos, preservando pressupostos relativos à universalidade das propriedades de suas gramáticas. A informação relativa à ordem canônica da língua é incorporada pela criança em estágio bem inicial da aquisição da mesma, tornando-se o acesso a esta automático e sem custo computacional mensurável. Um modelo que integre uma derivação gramatical no processo de formulação de sentenças ou de parsing de enunciados tem de dar conta deste fato. Desse modo, torna-se necessário distinguir nesse modelo integrado operações que, no modelo de língua, são tratadas da mesma forma.

A distinção entre movimento motivado discursivamente (com custo computacional mensurável) e movimento necessário para fixação de parâmetro de ordem, ou seja, que não precisa ser conduzido on-line, permitiu explicar a seletividade na produção agramática no que concerne a movimento. Esse modelo, inicialmente explorado em casos de DEL (Corrêa e Augusto, no prelo), deve ser estendido a outros fenômenos relacionados ao agramatismo a fim de seu potencial explanatório ser mais amplamente avaliado.

## Referências

- AYGEN, G. **Finiteness, Case and Clausal Architecture**. Massachusetts: Harvard University, 2004. Doctoral Dissertation, Harvard University, Massachusetts, USA, 2004.
- BAKER, M. C. **Incorporation: A theory of grammatical function changing**. Chicago: University of Chicago Press, 1988.
- BASTIAANSE, R.; THOMPSON, C. **Verb and auxiliary movement in Dutch and English agrammatic speakers**. *Brain and Language* 84, 2003. p. 286–305.
- BASTIAANSE, R., VAN ZONNEVELD, R. **On the relation between verb inflection and verb position in Dutch agrammatic aphasia**. *Brain and Language* 64, 1998. 165–181.
- \_\_\_\_\_. **Sentence production with verbs of alternating transitivity in agrammatic Broca's aphasia**. *Journal of Neurolinguistics* 18, 2005. p. 57-66.
- \_\_\_\_\_. KOEKHOEK, J., VAN ZONNEVELD, R. **Object scrambling in Dutch Broca's aphasia**. *Brain and Language* 86, 2003. p. 287-299.
- BURCHERT, F.; SWOBODA-MOLL, M.; DE BLESER, R. **The left periphery in agrammatic clausal representations: evidence from German**. *Journal of Neurolinguistics* 18, 2005. p. 67-88.
- CAPLAN, D. **Neurolinguistics and linguistic aphasiology: an introduction**. Cambridge, Mass. : Cambridge University Press, 1987.
- CHOMSKY, N. (1970). **Remarks on nominalization**. In: Jacobs, R. & Rosenbaum, P.(Eds.). **Readings in English transformational grammar**. Waltham, Mass: Ginn.

- \_\_\_\_\_. (1981). *Lectures on Government and Binding*. Foris, Dordrecht.  
\_\_\_\_\_. (1999). Derivation by Phase. *MIT Occasional Papers in Linguistics*, 18, Cambridge, Mass: MIT Working Papers in Linguistics.

CHRISTOPHE, A. et al. **Prosodic structure and syntactic acquisition:** the case of the head-complement parameter. *Developmental Science*, 6, 2003. p. 213-222.

CORRÊA, L.M.S. **Possíveis diálogos entre Teoria Linguística e Psicolinguística:** questões processamento, aquisição e do Déficit Específico da linguagem. In: N. Miranda & M. C. L. Name (Eds.). *Linguística e Cognição*. Juiz de Fora: Editora da UFJF, 2005a

\_\_\_\_\_. **Uma hipótese para a relação entre processador linguístico e gramática numa perspectiva minimalista.** In: Anais do IV Congresso Internacional da ABRALIN, 2005b. p. 353-364, disponível em: <<http://www.abralin.org/publicacao/abralin2005.pdf>>

\_\_\_\_\_. **Relação processador linguístico – gramática em perspectiva:** problemas de unificação em contexto minimalista. *D.E.L.T.A.* 24, 2008. p.231-282.

\_\_\_\_\_; AUGUSTO, M.R.A. **Computação linguística no processamento on-line:** em que medida uma derivação minimalista pode ser incorporada em modelos de processamento? Trabalho apresentado na Mesa Inter-GTs Teoria da Gramática e Psicolinguística no XXI Encontro Nacional da ANPOLL, PUC-SP. São Paulo: PUC, 2006.

\_\_\_\_\_; \_\_\_\_\_. **Computação linguística no processamento on-line: soluções formais para a incorporação de uma derivação minimalista em modelos de processamento.** *Cadernos de Estudos Linguísticos* 49, 2007. p. 167-183.

\_\_\_\_\_; \_\_\_\_\_. (no prelo). **Possible loci of SLI from a both linguistic and psycholinguistic perspective.** *Lingua*.

DE ROO, E. **Root nonfinite and finite utterances in child language and agrammatic speech.** *Brain and Language* 77, 2001. p. 398–406.

- DUMAN, T.; AYGEN, G.; BASTIAANSE, R. **Syntactic movement in Turkish agrammatic production.** Brain and Language 95, 2005. p. 161-162.
- FELSER, C.; CLAHSSEN, H. & MUNTE, T.F. **Storage and Integration in the Processing of Filler Gap Dependencies: An ERP Study of Topicalization and *WhI*- Movement in German.** Brain and Language 87, 2003. p. 345-354.
- FIEBACH, C.; SCHLESEWSKY, M.; & FRIEDERICI, A. **Separating syntactic memory costs and syntactic integration costs during parsing: The processing of German WH-questions.** Journal of Memory and Language 47, 2002. p. 250-272.
- FONG, S. **Computation with Probes and Goals: A Parsing Perspective.** In: A. M. Di Sciullo & R. Delmonte (Eds.) **UG and External Systems.** Amsterdam: John Benjamins, 2005.
- FRIEDMANN, N. **Agrammatism and the psychological reality of the syntactic tree.** Journal of Psycholinguistic Research 30, 2001. p. 71-90.
- \_\_\_\_\_. **Question production in agrammatism: The tree pruning hypothesis.** Brain and Language 80, 2002. p. 160-187.
- FRIEDMANN, N., e GRODZINSKY, Y. **Tense and agreement in agrammatic production: pruning the syntactic tree.** Brain and Language 56, 1997. p. 397-425.
- GRODZINSKY, Y. **The syntactic characterization of agrammatism.** Cognition 16, 1984. p. 99-120.
- HARKEMA, H. **Parsing Minimalist Languages.** University of California, 2001. Unpublished PhD. Dissertation.

JAKOBSON, R. **Studies on Child Language and Aphasia.** Paris: Mouton, 1941/71.

KAYNE, R. **The Antisymmetry of Syntax.** Cambridge, Mass: MIT Press, 1994.

KEAN, M. **The linguistic interpretation of aphasic syndromes.** Cognition 5, 1977. p. 9-46.

KOLK, H., HEESCHEN, C. **Agrammatism, paragrammatism and the management of language.** Language and Cognitive Processes 7, 1992. p. 89-129.

KOK, P.; KOLK, H.; HAVERKORT, M. **Agrammatic sentence production: Is verb second impaired in Dutch?** Brain and Language 96, 2006. p. 243-254.

LAPOINTE, S. **Some issues in the linguistic description of agrammatism.** Cognition 14, 1983. p. 1-41.

LEE, M. **Dissociations among functional categories in Korean agrammatism.** Brain and Language 84, 2003. p. 170-188.

LONZI, L.; LUZZATTI, C. **Relevance of adverb distribution for the analysis of sentence representation in agrammatic patients.** Brain and Language 45, 1993. p. 306-317.

MILMAN, L.; DICKEY, M.; THOMPSON, C. **Production of functional categories in agrammatic narratives: An Item Response Theory Analysis.** Brain and Language 91, 2004. p. 126-127.

NEUHAUS, E.M.; PENKE, M. **Wh-question production in German Broca's aphasia.** Brain and Language 87, 2003. p. 59-60.

- PENKE, M. **Controversies about CP:** A comparison of language acquisition and language impairments in Broca's aphasia. *Brain and Language* 77, 2001. p. 351–363.
- PHILLIPS, C. **Linguistics and linking problems.** In: M. Rice; W. Warren (Eds.), **Developmental Language Disorders:** From Phenotypes to Etiologies. Mahwah, N.J.: Lawrence Erlbaum, 2003.
- POLLOCK, J-Y. **Verb movement, universal grammar, and the structure of IP.** *Linguistic Inquiry* 20, 1989. p. 365-424.
- RODRIGUES, E. dos S.; CORRÊA, L.M.S.; AUGUSTO, M.R.A. **Concordância sujeito-verbo em um modelo integrado misto (top-down/bottom-up) da computação on-line.** *Veredas on-line* 2, 2008. p. 76-91.
- RUIGENDIJK, E.; KOUWENBERG, M.; FRIEDMANN, N. **Question production in Dutch agrammatism.** *Brain and Language* 91, 2004. p. 116-117.
- STAVRAKAKIA, S.; KOUVAVA, S. **Functional categories in agrammatism:** Evidence from Greek. *Brain and Language* 86, 2003. p. 129-141.
- VAN DER MEULEN, I.; BASTIAANSE, R.; ROORYCK, J. **Wh-movement in French agrammatism.** *Brain and Language* 83, 2002. p. 184-187.
- WANNER, E. & MARATSOS, M. **An ATN approach to comprehension.** In: M. Halle; J. Bresnan; G. A. Miller. (Orgs.). **Linguistic Theory and Psychological Reality.** Cambridge: MIT Press, 1978.

WEXLER, K. **Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage.** Lingua 106, 1998. p. 23-79.

WENZLAFF, M., CLAHSEN, H. **Tense and agreement in German agrammatism.** Brain and Language 89, 2004. p. 57-68.

\_\_\_\_\_. **Finiteness and verb-second in German agrammatism.** Brain and Language 92, 2005. p. 33-44.

ZUCKERMAN, S.; BASTIAANSE, R.; VAN ZONNEVELD, R. **Verb Movement in Acquisition and Aphasia:** Same Problem, Different Solutions – Evidence from Dutch. Brain and Language 77, 2001. p. 449-458.

ZURIF, E. et al. **An on-line analysis of syntactic processing in Broca's and Wernicke's aphasia.** Brain and Language 45, 1993. p. 448- 464.

# **VARIATION AND OPTIMALITY THEORY: REGRESSIVE ASSIMILATION IN VIMEU PICARD**

Walcir CARDOSO

Concordia University (Montreal, Canada), Centre for the Study of Learning and Performance

## **RESUMO**

*Este artigo apresenta uma avaliação das propostas pelas quais a variação tem sido analisada dentro da Teoria da Optimidade (TO). Através da análise do processo fonológico de assimilação regressiva entre palavras em picardo conclui-se que uma versão estocástica da TO é mais adequada para a investigação de fenômenos variáveis.*

## **ABSTRACT**

*This article presents an assessment of how variation has been analyzed within the framework of Optimality Theory (OT). Analyzing the phonological process of acrossword regressive assimilation in Picard, we conclude that a stochastic version of OT is better suited for the investigation of variable phenomena.*

## **PALAVRAS-CHAVE**

*Assimilação Regressiva. Picardo de Vimeu. Variação. Teoria da Optimidade.*

## **KEY WORD**

*Optimality Theory. Regressive Assimilation Variation Vimeu Picard*

## **Introduction**

With the emergence of Optimality Theory (OT – Prince and Smolensky 1993) and the consequent demise of variable rules (Labov

1972, Cedergren 1973, Cedergren and Sankoff 1974, Guy 1975) in favor of constraint interaction for the analysis of variability (e.g. Reynolds 1994, Anttila 1997; see also Fasold 1996 and Bergen 2000 for a critique of variable rules), it has been argued that intra-language variation can be satisfactorily accounted for via constraint interaction (e.g. Reynolds 1994, Anttila 1997, Taler 1997, Cardoso 2001). In a constraint-based approach like OT, variability can be expressed without resorting to a separate grammar for each variant or, in the case of a process with more than two variants, without the postulation of more than one rule for a single phenomenon: the framework allows for variation to be encoded in (and therefore predicted by) a single constraint hierarchy. On the other hand, the theory also allows for the assignment of separate grammars for cases in which variation truly involves different grammars (e.g. different dialects, different registers).

Several approaches have been proposed for the analysis of variation in OT. In this paper, we will discuss and evaluate four of these proposals: (1) Kiparsky's (1993) grammars in competition, (2) Reynolds' (1994) floating constraints, (3) Anttila's (1997) partial grammars, and (4) Boersma's (1998) and Boersma and Hayes' (2001) stochastic OT. To level the playing field and to provide empirical evidence for my claims and arguments, the present investigation focuses on one single phenomenon: Across-Word Regressive Assimilation (AWRA) in Vimeu Picard (spoken in northern France), a phonological process that operates variably depending, among other factors, on the geographical distribution of its speakers. The data come from a database consisting of an oral fieldwork corpus of tape-recorded interviews with Picard speakers, and written documents (e.g. private letters, unpublished stories, etc.).

This study offers variationist OT accounts for AWRA within these four approaches. The general goals are: (1) to provide an overview of how variation has been analyzed within the framework of Optimality Theory, and (2) to assess four different approaches that

have been proposed for the analysis of variable phenomena in OT. More specifically, the study addresses the following research question: Which of these approaches is better capable of accounting for sociolinguistically-grounded variation, under the traditional variationist assumption that the grammar must include quantitative information, and that the manipulation of frequency is part of a speaker's linguistic competence (e.g. Guy 1975, 1997, Labov 1969, Cedergren and Sankoff 1974)?

This article is composed of four main sections. In section 1, I provide the data that illustrate AWRA, its domain of application, and the variation patterns that characterize the phenomenon. Section 2 presents the data collection procedures and a discussion of the relevant quantitative VARBRUL results for one of the variables considered in the study: The geographic distribution of the participants (originally reported in Cardoso 2001, 2003). In section 3, four different approaches to variationist OT are presented and evaluated vis-à-vis their abilities to satisfactorily account for the variable patterns that characterize AWRA in Picard. It is also shown how a stochastic, GLA-based approach to the analysis of variation is preferable to standard, ordinal OT to analyze the type of variation described in this study: It presupposes a simpler (and more precise) grammar, governed by the same constraints and principles that govern categorical phenomena. Finally, section 4 concludes the study.

## 1 The data: Across-Word Regressive Assimilation

Across-Word Regressive Assimilation (AWRA henceforth) is a phonological process of Vimeu Picard that operates exclusively at the domain juncture of a (CV)l shape clitic (*fnc* in (1)), followed by a consonant-initial lexical word (*lex*) (e.g. (dol)fnc bibin)lex → [dob bibin] 'some brandy'; (al)fnc (pɛk)lex → [ap pɛk] 'at the fishing'). When both phonological and morphosyntactic contexts are met, the root node of

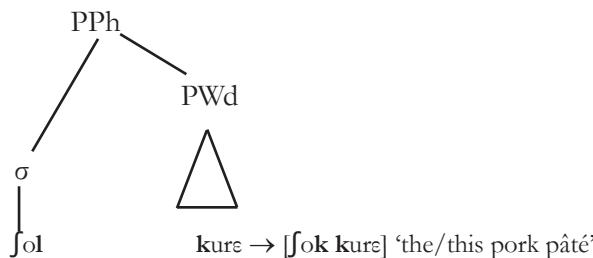
the lexical word's initial consonant associates to the timing slot of the preceding clitic-final /l/, resulting in a geminate across the two words (CV tier is used for illustrative purposes only):

(1) The AWRA process

C	V	C	C	V	C	→	[ʃɔf fɛt]	'the / this party'
		ʃ					(ʃ o l)fnç (f ε   t)lex	

As implied in the discussion and representation above, AWRA is a domain-sensitive phenomenon that applies exclusively at the domain juncture of an /l/-final syllable and the following consonant-initial Prosodic Word, within the Phonological Phrase domain (it does not operate within words – e.g. /kalfa/ → \*[kaf.fa], z[kal.fa] 'cauker' – and across words in higher prosodic domains – e.g. /bel de bel/ → \*[bed de bel], z[bel de bel] 'the very last match'). Following Nespor and Vogel's (1986) Prosodic Phonology approach to domain-sensitive phenomena, I will refer to this domain juncture as  $\Phi$  (see Cardoso 1999, 2003 for a comprehensive analysis of the prosodic domain of AWRA).

(2) The Prosodic Domain of AWRA in Picard:  $\Phi$



Contrary to what is illustrated in (1) and implied in the representation in (2), AWRA does not apply categorically. In this prosodic context, three distinct patterns can be observed: (a) faithfulness of input /l/ (/l/-preservation); (b) Across-Word Regressive Assimilation (AWRA); and (c) /l/-deletion. For convenience, I will use the acronym AWRA as both a general term for the phenomenon (to which I will occasionally refer as the “AWRA phenomenon”), and as a term for one of its variants, the one illustrated in (3b) below.

(3) Variants of AWRA

a. /l/-preservation

/ʃol kurə/	→	[ʃol kurə]	‘the/this pork pâté’
/dol tart/	→	[dol tart]	‘some pie’
/al kāt/	→	[al kāt]	‘she sings’

b. AWRA

/ʃol kurə/	→	[ʃok kurə]	
/dol tart/	→	[dot tart]	
/al kāt/	→	[ak kāt]	

c. /l/-deletion

/ʃol kurə/	→	[ʃo kurə]	
/dol tart/	→	[do tart]	
/al kāt/	→	[a kāt]	

In the following section, I will show that one of the factors that determine the patterns observed above is the geographic distribution of the speakers. These results will then serve to demonstrate how the four approaches proposed in OT can be used for the analysis of sociolinguistically-grounded variation.

## 2 AWRA and geographical distribution: a variationist investigation

In this section, I present a discussion of the data collection procedures that were adopted in order to obtain samples of non-categorical data such as those illustrated in (3) above. To illustrate the variable phenomenon in (3) across four different approaches to variation that have been proposed in OT, the discussion will focus exclusively on the geographic distribution of the speakers (geographic location factor). For a comprehensive discussion of the variationist study, see Cardoso (2001, 2003).

In brief, the study consisted of 2,783 tokens of variants of AWRA collected in the field by Julie Auger for the Picard project during the summers of 1996 and 1997, which were further transcribed by four research assistants. The data collected were stratified among six independent variables and later analyzed by the VARBRUL 2 program (Pintzuk 1988): Three extralinguistic factor groups: (1) level of formality, (2) speaker, and (3) geographic location; and three linguistic factor groups: (1) grammatical status of the l-clitic, (2) place of articulation of the following consonant, and (3) manner of articulation of the following consonant. As indicated above, I will only report and discuss the results for the geographic location factor in this paper.

The participants (Speakers 1-8) were eight male adult native speakers of Picard, with an average age of more than 70 years old; they inhabited five villages in the Picardie region in northern France: Feuquières, Fressenneville, Bienfay, Bouillancourt and Nibas. Women and younger speakers were not included in the investigation because the vast majority of native speakers of Picard who still use the language routinely are older men.

In order to collect tokens from a wide range of stylistic levels, the data collection methodology used in this study provides a three-level distinction in a formality hierarchy: (1) informal interview, (2) formal

interview, and (3) collection of written documents. (1) The informal interview consisted of tape-recorded conversations between the field worker and the interviewee or between the interviewee and other native speakers of Picard. (2) The formal interview consisted of an audio-recorded translation task (designed for the purpose of this study) in which the participants were asked to orally translate French sentences into Picard.<sup>1</sup> (3) The collection of written documents consisted of the selection of such documents from at least one speaker from each region investigated. These documents were extracted from articles from the Picard magazine *Ch'Lanchron*, and unpublished material (including short stories, articles and a few private letters).

From all the linguistic and extralinguistic factors that were initially included in the investigation, VARBRUL's probabilistic results indicate that the external variables level of formality and geographic location and the internal variable status of the *l*-clitic have significant conditioning effects on determining the output of the AWRA phenomenon. Focusing solely on the results for geographic location, two general patterns can be observed with regards to the AWRA phenomenon: One in which AWRA is favored (.48) as opposed to /*l*/-preservation (.28) and /*l*/-deletion (.24) (2 participants), and another pattern in which all three variants are relatively equally distributed (average around .33) (6 participants). The VARBRUL results (in probabilities) for this factor group are illustrated below (from Cardoso 2001, 2003).

(4)

---

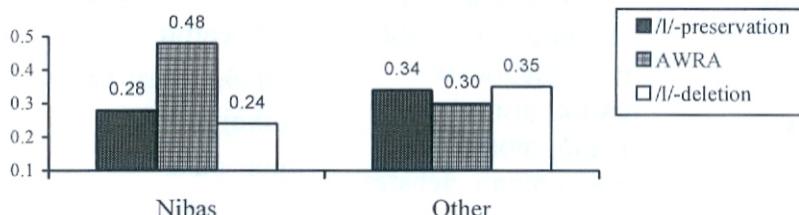
<sup>1</sup> The only way to elicit more formal oral data was through the translation task because Picard, as a dying language, is characterized by monostylistism (e.g. Dressler 1972, Dorian 1977, Dressler 1988).

TABLE 1: Probabilities in five geographic locations

	/l/-preservation	AWRA	/l/-deletion
Nibas	.28	.48	.24
Feuquières	.31	.30	.39
Fressenneville	.38	.29	.32
Bienfay	.30	.32	.38
Bouillancourt	.38	.30	.32

For ease of exposition (and because exactly two patterns can be observed), this factor group is regrouped into two major categories: Nibas and Other (which includes all the remaining villages). The results (in probability) are illustrated in Figure 1.

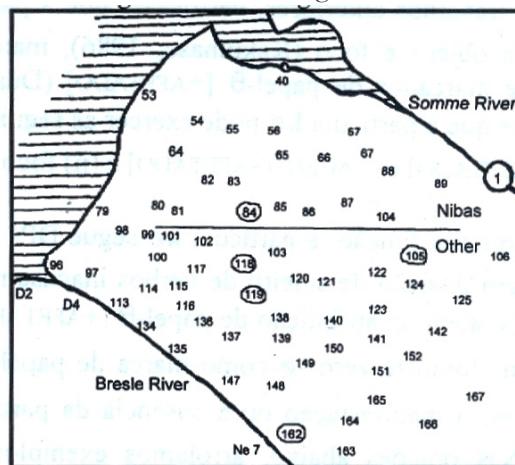
FIGURE 1: AWRA and Geographic Location



Some readers could argue that the variation pattern illustrated in Figure 1 merely demonstrates intraspeaker variation, especially in the context of a limited number of tokens and participants. A brief look at previous studies on regressive assimilation in Picard leads me to conclude that, even though intraspeaker variation is a logical alternative for describing the AWRA phenomenon in the language, more needs to be said about the effect of geographic variation as a determining factor in the results observed above. In the introduction to his *Dictionnaire des parlers picards du Vimeu* and in his grammar *Grammaire des parlers picards du Vimeu (Somme)*, Vasseur (1963, 8 and 1996, 7-8 respectively) refers to the region in which regressive assimilation applies “always and

without exception" as the "region of Nibas". He also acknowledges that assimilation occurs in other regions, but to a lesser extent and sometimes only involving the determiner /Sol/. Likewise, Debrrie (1981, 455) observes that regressive assimilation "is concentrated to its maximum intensity in Nibas and in other neighboring villages: 65 [Arrest], 66 [Mons], 86 [Franleu], 87 [Quesnoy] and 122 [Toeufles] (see Figure 2). It is no coincidence that our results display a relatively similar pattern for the AWRA phenomenon: Of the five villages included in this investigation, it is in Nibas that the AWRA variant is more likely to appear. In Figure 2 below (adapted from Dubois 1957), I show the geographic location of the five villages in the region of Vimeu. The numbers that relate to the villages investigated are circled: 84 = Nibas; 105 = Bienfay (which belongs to the commune of Moyenneville); 118 = Fressenneville; 119 = Feuquieres; and 162 = Bouillancourt. The straight line on the map indicates an isogloss-like geographical boundary between two probable dialects: One in which AWRA is highly favored (represented by the region of Nibas – 84), and one in which the three variants are equally likely to appear (represented by the other villages).

FIGURE 2: Five villages in Vime



As has been proposed for the analysis of distinct dialectal varieties (e.g. Selkirk 1997, Alber 2001, Boersma 2001), I argue that these two sets of villages define separate dialects, which are formally represented by two grammars: One for the village of Nibas, in which the AWRA variant is favored as opposed to the other two variants, and one for the other villages (Other) in which the three variants are equally predicted. The establishment of these two distinct variable grammars will allow us to analyze and evaluate the variable AWRA phenomenon within four OT approaches for the analysis of variation. This will be the topic of the following section.

### 3 Optimality Theory and variation: AWRA

In this section, I demonstrate how the variable AWRA phenomenon can be analyzed within four approaches that have been proposed for the analysis of variation in OT: (1) Grammars in competition (section 3.1), (2) Crucial nonranking: Floating constraints (section 3.2), (3) Crucial nonranking: Partial grammars (section 3.3), and (4) Stochastic OT (section 3.4).<sup>2</sup> After an evaluation of these four approaches in the context of the AWRA quantitative results, I will argue that a stochastic version of OT, the one proposed by Boersma (1998) and Boersma and Hayes (2001), is better suited for the analysis of variable phenomena.

To account for the variable aspects of AWRA, I adopt the following well-established OT constraints in (5). I assume that the reader is well versed in OT and knows the basic mechanisms involved in constraint-based analyses.

---

<sup>2</sup> A fifth approach for the analysis of variation in OT was proposed by Coetze (2006): A rank-ordering model of EVAL. The model, however, was left out from the discussion and analyses because it is unable to make precise quantitative predictions: the approach is restricted to predicting the *relative* frequencies of variants. Due to space limitations, I will not provide details of Coetze's rank-ordering model; suffice it to say that in the context of the AWRA phenomenon presented here, this model does not differ considerably from Kiparsky's (1993) grammars in competition approach, which will be discussed in forthcoming section 3.1.

## (5) Constraint definitions

<b>FAITH-Lex</b>	The outputs of lexical words are faithful to their inputs.
<b>MAX-IO</b>	Every segment of the input has a correspondent in the output.
<b>Linearity</b>	The input reflects the precedence structure of the output, and vice versa.
<b>NoCoda-Rt</b>	A Coda cannot license a Root node.

The constraint FAITH-Lex, which should be interpreted as a cover term for a set of constraints on correspondent elements (e.g. MAX-Lex, DEP-Lex, Linearity-Lex), expresses the cross-linguistic tendency for preservation of information contained in lexical words other than in function words. It was proposed by Casali (1997) and Pulleyblank (1997) (under ‘Faith-Stem’ – see also Trubetzkoy 1939, Steriade 1995, Casali 1996 and Beckman 1997), although implicit in McCarthy and Prince (1995) under the Root-Affix Faithfulness Metaconstraint: Root-Faith  $\gg$  Affix-Faith. In the context of AWRA, FAITH-Lex predicts the directionality of AWRA, and thus prevents cases of *progressive* assimilation (e.g. /ʃɔl kure/  $\rightarrow$  /ʃɔl lure/ ‘the pork pâté’). Due to space limitations, I will not illustrate unattested cases of progressive assimilation in the forthcoming rankings and discussions. It must be assumed, however, that FAITH-Lex is ranked at the higher, undominated end in the constraint hierarchy that characterizes the grammar of Picard.

MAX-IO is a constraint that militates against deletion and is violated in cases in which the clitic-final /l/ is deleted from the output; e.g. /ʃɔ kure/ violates MAX-IO because the input /l/ does not have a correspondent in the output.

The Linearity constraint rules out candidates in which the sequence of input segments is reversed or otherwise not obeyed in the surface

representation. In cases of regressive assimilation, the precedence relation of  $S_1 /l-k/$  is not reflected in  $S_2 [k-k]$ :  $/l/$  precedes  $/k/$  in  $S_1$  but the correspondent of  $/l/$  does not precede the correspondent of  $/k/$  in the output.

Finally, NoCoda-Rt is member of a family of constraints that captures the crosslinguistic observation on syllabic well-formedness that coda segments are marked. As originally proposed by Prince and Smolensky (1993) (i.e. Syllables do not have Codas), the general version of the constraint is inadequate to account for the range of behavior that coda consonants display cross-linguistically, since languages impose different types of restrictions on codas (in OT, see McCarthy and Prince 1993, Benua 1995, Lombardi 1995, Kawasaki 1998, among others). Observe that NoCoda-Rt is formulated in terms of licensing; consequently, a syllable final consonant can only surface without incurring a violation of this constraint if all of its features are linked to and therefore licensed by a following onset (cf. Piggott 2003). In languages in which NoCoda-Rt is highly ranked, the only codas permitted will be geminates. This is exactly the behavior observed in some dialects of Inuit (e.g. Kalaallisut and Labrador – Bobaljik 1996). In Picard, NoCoda-Rt rules out forms in which the clitic-final coda  $/l/$  bears and therefore licenses its own Root node, e.g. [ʃɔl vak] – see (6a). In cases of assimilation, however, NoCoda-Rt is not violated because the assimilated coda's segmental content (i.e. Root node) is licensed by the onset of the following word, e.g. [ʃov vak] – see (6b). This is shown in the representations below, using standard Onset-Rhyme theory (segments stand for Root nodes).

(6)

a. Violation of NoCoda-Rt	b. Satisfaction of NoCoda-Rt
$\sigma$	$\sigma$
... Coda	Onset
l	v

Recall from section 1 that AWRA is a domain-sensitive phenomenon that operates exclusively at the domain juncture  $\Phi$  (see (3) above). As such, we cannot assume that the constraints discussed above have the same weight in the grammar of Picard. For instance, while the constraints MAX-IO, Linearity and NoCoda-Rt can be assumed to be operative at the domain juncture  $\Phi$ , it is clear that this assumption does not hold at other domains, since variable AWRA is unattested in domains lower or higher than  $\Phi$  (see discussion under (1) above). For the analysis of domain-driven phenomena such as AWRA, I adopt Cardoso's (2003, 2005) domain-specific constraint approach, in which each constraint is decomposable into their domain-specific (e.g. Phonological Phrase – PPh, Prosodic Word – PWd) and locusspecific (i.e. juncture, limit, and span) counterparts, each of which may be ranked independently within a single grammar to yield the alternations observed across domains. Influenced by insights from Prosodic Phonology (Selkirk 1972, 1997, and Nespor and Vogel 1986), the decomposition of constraints within this approach is restricted to those established by this theory (e.g. there are no direct references to morphosyntactic domains or words). In the forthcoming discussions, the constraints in (5) will appear specified for  $\Phi$  whenever relevant (e.g. MAX-IO <sub>$\Phi$</sub> , Linearity <sub>$\Phi$</sub> ). Constraints that have an Utterance span effect (i.e. those that operate within the entire span of the Utterance domain), on the other hand, will not be specified for a domain (e.g. MAX-IO, Linearity), implying that they operate across all domains of the grammar.

Let us now address how the different approaches proposed for the analysis of variation in OT are able to account for AWRA and its variation patterns.

### 3.1 Grammars in competition

The first account of variation within the OT framework was that of Kiparsky (1993). Within Kiparsky's approach, which follows a stricter view of constraint domination (i.e. a view in which constraints are

crucially ranked), variation is seen as a result of competing grammars (or distinct constraint rankings). For instance, in order to account for t/d deletion in English, he assigns a separate constraint ranking for each set of environments favoring the application of the phenomenon. Adapting Kiparsky's approach to the AWRA context, his view would require the assignment of three separate grammars to account for the variation patterns encountered in Nibas and in the other villages (Other):

- (7) Kiparsky's approach to variation and AWRA in Other and Nibas

Grammars (constraint rankings)	Output
(a) MAX-IO <sub>Φ</sub> , Linearity <sub>II</sub> >> NoCoda-Rt <sub>Φ</sub>	/l/-preservation
(b) MAX-IO <sub>Φ</sub> , NoCoda-Rt <sub>Φ</sub> >> Linearity <sub>Φ</sub>	AWRA
(c) Linearity <sub>Φ</sub> , NoCoda-Rt <sub>Φ</sub> >> MAX-IO <sub>Φ</sub>	/l/-deletion

Kiparsky's approach to variation resembles the cophonology approach (e.g. Itô and Mester 1995ab, Orgun 1996, Inkelas et al. 1997, Inkelas and Zoll 2000), which also appeals to separate grammars to account for different types of variation (e.g. those triggered by different morphological or prosodic constituents, by a class of specific morphemes). Consequently, his approach inherits one of the shortcomings of the use of cophonologies, namely, the proliferation of grammars. Furthermore, Kiparsky's approach to variation is unable to predict the likelihood of occurrence of each variant involved in AWRA. Based on the rankings shown in (7), each variant of AWRA in Nibas is equally likely to appear, which is inconsistent with the results illustrated in Figure 1 in section 2: In this village, AWRA is more likely to apply (.48) than the other two variants (.28 and .24 for /l/-preservation and /l/-deletion respectively).

Assuming the traditional variationist view that the grammar must include quantitative information and that the manipulation of frequency is part of a speaker's linguistic competence (e.g. Guy 1975, 1997, Labov 1969, Cedergren and Sankoff 1974), if the quantitative aspect of a variable grammar is ignored, variation is reduced to random selection, similar to the notion of "free variation". In standard, non-variationist OT, this is commonly claimed to result from the interaction of "freely-ranked" constraints (e.g. Clements 1997, 315), the antithesis of variationist linguistics. This is exactly what Kiparsky's approach implies: That variation such as that found in Nibas is merely the result of the random selection of grammars.

### 3.2 Crucial non-ranking: floating constraints

In an effort to account for variation by assuming the existence of a single grammar that allows the inclusion of both variable outputs and quantitative information, Reynolds 9

(1994) and Anttila (1997, previously published as a manuscript in 1995) pursued an idea hinted at by Prince and Smolensky (1993) in a footnote, about the possibility of crucial nonranking of constraints. In the early stages of OT, it was not evident why the crucial nonranking of constraints, an essential assumption for the concept that variation can be encoded within a single grammar, should be tolerated in a framework that advocates a *strict dominance hierarchy* (i.e. that each constraint must have absolute priority over all the constraints lower in the hierarchy). In the context of constraint ranking in OT, there could exist a situation in which a constraint set imposes crucial non-dominance (i.e. nonranking) of its components. When a given grammar is unable to categorically yield one of two or more rankings allowed by a set of constraints, the result is the possibility of two or more acceptable forms or outputs in that grammar, i.e. variation *per se*.

Based on the notion of crucial nonranking, two different proposals have been made in the OT literature: (1) Reynolds' (1994) floating

constraint approach (discussed in this section); and (2) Anttila's (1997) partial ranking of constraints approach (discussed in section 3.3). In this section, the focus is on Reynolds' "floating constraints" approach. In Reynolds' view, a variation grammar consists of variably ranked constraints (or *floating* constraints, using the author's terminology). In this approach, the grammar is defined by a single constraint hierarchy, in which one or more constraints may float with respect to another constraint or set of constraints. For example, in a constraint set (call it  $S$ ), some subset  $S'$  may float with respect to some other subset  $S''$ . Within each subset, constraints may float with respect to each other, as is the case in subset  $S''$  below.

(8) Reynolds' floating constraints

$$\{A \gg \boxed{\{B\}S' \{C D\}S''} \gg E\}S$$

From the number of rankings allowed by a set of variably ranked constraints, distinct outputs can be predicted. For instance, from the variable ranking of  $S'$  and  $S''$  above, four different rankings and therefore potentially different outputs are expected:

(9) Rankings allowed by a set of floating constraints

A >>	<b>B</b> >> C >> D	>> E
A >>	<b>B</b> >> D >> C	>> E
A >>	C >> D >> <b>B</b>	>> E
A >>	D >> C >> <b>B</b>	>> E

Anttila (1997) demonstrates that the probability of each variant's occurrence is the result of the number of rankings for which each variant wins, divided by the total number of rankings (or tableaux) generated by the variably ranked constraints. This is formalized in (10):

## (10) Variant probabilistic prediction (Anttila 1997)

- (a) A candidate is predicted by the grammar *if and only if* it wins in some tableaux.
- (b) If a candidate wins in  $n$  tableaux and  $t$  is the total number of tableaux, then the candidate's probability of occurrence is  $n/t$ .

To illustrate, suppose that in a given grammar, GRAM, two constraints B and C float with respect to each other (11a). This is indicated by the semi-colon (to distinguish crucial nonranking from cases of indeterminate ranking, indicated by a comma) between the two constraints involved, with the curly brackets delimiting the set of floating constraints. As a result, two different rankings are possible as illustrated in (11b):

## (11) A variably ranked grammar

- (a) Constraint ranking:  $A >> \{B; C\} >> D$
- (b) Ranking possibilities:  $A >> \mathbf{B} >> \mathbf{C} >> D$   
 $A >> \mathbf{C} >> \mathbf{B} >> D$

Imagine that two optimal forms are possible in GRAM, i.e. Cand1 and Cand2. Cand1 is selected when B is ranked higher than C, while Cand2 is selected in the reverse situation. This is illustrated in the two tableaux in (12).

## (12) Tableau 1: Variation in GRAM

Tableau (a) =  $A >> B >> C >> D$ 

	A	B	C	D
► Cand1			*	
Cand2		*!		

Tableau (b) =  $A >> C >> B >> D$ 

	A	C	B	D
Cand1			*!	
► Cand2			*	

Following Anttila's (1997) variant probabilistic prediction, the variable ranking of constraints B and C results in a pattern in which two outputs are possible, and the probability of each output occurrence can be predicted by (10). For example, candidates 1 and 2 in (12) win in exactly one tableau each ( $n=1$ ), and two is the total number of tableaux ( $t=2$ ).  $n/t = 1/2 = 0.5$  or 50%. Each candidate's probability of occurrence is thus 0.5 and each variant is likely to occur 50% of the time in the same grammar.

The constraint ranking in (11a) and the tableaux in (12) emphasize a crucial distinction in the context of variation in OT, i.e. the distinction between grammars and tableaux. While in (11a) one ranking or grammar yields two tableaux and consequently two outputs (see Tableau 1 in (12)), a categorical grammar yields only one tableau and consequently only one output (i.e. no variation).

Reynolds' approach to variation can be straightforwardly applied to the investigation of AWRA.<sup>3</sup> For instance, to account for the variation patterns observed in the village of Nibas (where AWRA is more likely to apply (.48) than /l/-preservation or /l/-deletion – .28 and .24 respectively), two subsets of domain-specific floating constraints would be required:

### (13) Floating constraints and AWRA

Nibas: { { MAX-IO<sub>Φ</sub>; NoCoda-Rt<sub>Φ</sub> }; Linearity<sub>Φ</sub> }  
 Other: { MAX-IO<sub>Φ</sub>; NoCoda-Rt<sub>Φ</sub>; Linearity<sub>Φ</sub> }

Consider the village of Nibas, where the hierarchy in (13) yields four rankings, two of which select AWRA as the optimal candidate, while the other two rankings select either /l/-preservation or /l/-deletion as the output. The application of Anttila's (1997) variant probability prediction ( $n/t$ ) in (10) yields probabilistic results that tightly match the

<sup>3</sup> In fact, Reynolds' floating constraints approach was adopted in the analysis of variation in AWRA in an earlier version of this investigation, as reported in Cardoso (2001).

ones observed in the village of Nibas. The same applies to the other villages (Other) in (15).

(14)

TABLE 2: Floating constraints – output selection for Nibas

Possible Rankings	Output	Observed	Predicted
$\text{MAX-IO}_\phi >> \text{NoCoda-Rt}_\phi$ $>> \text{Linearity}_\phi$	AWRA		
$\text{NoCoda-Rt}_\phi >> \text{MAX-IO}_\phi$ $>> \text{Linearity}_\phi$	AWRA	.48	$2/4 = .50$
$\text{Linearity}_\phi >> \text{MAX-IO}_\phi >>$ $\text{NoCoda-Rt}_\phi$	/l/- preservation	.28	$1/4 = .25$
$\text{Linearity}_\phi >> \text{NoCoda-Rt}_\phi >>$ $\text{MAX-IO}_\phi$	/l/-deletion	.24	$1/4 = .25$

(15)

TABLE 3: Floating constraints – output selection for Other

Possible Rankings	Output	Observed	Predicted
$\text{MAX-IO}_\phi >> \text{NoCoda-Rt}_\phi$ $>> \text{Linearity}_\phi$	AWRA	.30	$1/3 = .33$
$\text{Linearity}_\phi >> \text{MAX-IO}_\phi >>$ $\text{NoCoda-Rt}_\phi$	/l/- preservation	.34	$1/3 = .33$
$\text{Linearity}_\phi >> \text{NoCoda-}$ $\text{Rt}_\phi >> \text{MAX-IO}_\phi$	/l/-deletion	.35	$1/3 = .33$

In sum, from an empirical perspective, Reynolds' approach can satisfactorily account for the AWRA phenomenon in Picard, as indicated above. The approach, however, is flawed from a conceptual perspective. Most importantly, the model is too permissive in the possibilities of rankings allowed within the grammar. For instance, to account for a variation pattern in which four constraints (A, B, C, D) interact to yield two distinct variants (X, Y), several possibilities of

rankings (from which I include only five) are possible within Reynolds' approach (assume that X violates A and B, while Y violates C and D) (adapted from Taler 1997). In addition, the frequencies predicted by the variably-ranked hierarchy are always in small integer fractions (e.g. 1/2, 1/3, 2/3). While I have shown that this is exactly what is found in variable AWRA (e.g. in Other, the three variants are predicted 1/3 of the time), other studies have shown that this is not always the case in variationist linguistics (e.g. Cardoso 2007; see also the critique of the /kelyia/ results in Nagy and Reynolds' 1994 analysis of word-final deletion in Faetar, discussed in Guy 1997, 136-8).

(16)

TABLE 4: Reynolds' floating constraint approach: a permissive model

Possible Constraint Rankings: <i>several</i>	Predictability ( <i>n/t</i> )
a. { A; B; C; D }	X = .5; Y = .5
b. { { A; B }; { C; D } }	X = .5; Y = .5
c. { { A; C }; { B; D } }	X = .5; Y = .5
d. { { A; B } >> C; D }	X = .5; Y = .5
e. { { B; A } >> C; D }	X = .5; Y = .5

### 3.3 Crucial nonranking: partial grammars

The third approach to variation in OT was proposed by Anttila (1997). Instead of the use of sets of floating constraints, each of which may contain one or more constraints (see (13) and (16b-e)), Anttila's model accounts for variation by means of a more restricted version of crucial nonranking. In his approach, the only partial rankings allowed are those composed of single constraints. For instance, to account for the variation pattern illustrated in (16), only the crucial nonranking of all of the constraints A, B, C and D (i.e. (16a)) is permitted in an Anttila-like approach: {A; B; C; D}.

To account for the disparity of results observed involving the factor *geographic location* within Anttila's approach, consider the two

distinct variable grammars below (from Cardoso 2003), composed of domain-specific constraints (those not specified for a domain should be assumed to be operative over the span of the Utterance domain, as indicated earlier): (1) one grammar for the village of Nibas, in which the crucial nonranking of five constraints yields 120 tableaux, and (2) one grammar for Other, in which the nonranking of three constraints yields 6 tableaux.

(17) Geographic location and AWRA

<b>a. Nibas Grammar:</b>
{MAX-IO <sub>φ</sub> ; MAX-IO; NoCoda-Rt <sub>φ</sub> ; NoCoda-Rt; Linearity <sub>φ</sub> } >> Linearity
<b>b. Other Grammar:</b>
{MAX-IO <sub>φ</sub> ; NoCoda-Rt <sub>φ</sub> ; Linearity <sub>φ</sub> } >> MAX-IO, NoCoda-Rt, Linearity

The application of Anttila's variant probability prediction in (10) yields the results illustrated in Table 5. Observe that under each variant, the left column (under Pred) indicates the predicted probability of each variant's occurrence, calculated by  $n/t$ , and the parenthesized numbers illustrate the number of rankings (or tableaux) in which that candidate is the winner for each subset of villages (i.e. Nibas or Other). The values in the right column (Obs), on the other hand, indicate the actual VARBRUL weight established for each variant (values from Figure 1).

(18)

TABLE 5: Predicted & observed probability of variant occurrence by geographic location

Geographic location	Total # of tableaux	/1/-preservation		AWRA		/1/-deletion	
		Pred	Obs	Pred	Obs	Pred	Obs
Nibas	120	.23 (28)	.28	.53 (64)	.48	.23 (28)	.24
Other	6	.33 (2)	.34	.33 (2)	.30	.33 (2)	.35

In sum, the crucial nonranking of the constraints  $\text{MAX-IO}_\phi$ ,  $\text{MAX-IO}$ ,  $\text{NoCoda-Rt}_\phi$ ,  $\text{NoCoda-Rt}$ , and  $\text{Linearity}_\phi$  in Nibas yields a pattern in which the AWRA variant is more often favored (.53) in relation to the other variants (.23 for both /1/-preservation and /1/-deletion). In the grammar assigned for Other, on the other hand, the crucial nonranking of  $\text{MAX-IO}_\phi$ ,  $\text{NoCoda-Rt}_\phi$  and  $\text{Linearity}_\phi$  results in a pattern in which each of the three variants of the AWRA phenomenon is equally expected to surface (probability .33). As shown in (18) under “Obs”, the predictions made here closely correspond to the VARBRUL results obtained.

Comparing the options that are possible in the two approaches that appeal to the crucial nonranking of constraints, Anttila’s is more advantageous for the analysis of variation. Firstly, Anttila’s model is more constrained because it is less permissive on the possibilities of rankings allowed by the grammar. In fact, Anttila’s approach constitutes a subset of Reynolds’, as implied at the outset of this section. Secondly, Reynolds’ model presents problems from a learnability perspective because the range of options that the language learner will entertain when confronted with the numerous ranking possibilities that his model predicts is too vast. In other words, the hypothesis space in Reynolds’ approach is too large in comparison to Anttila’s. A desirable effect of Anttila’s approach in comparison to Reynolds’ is that it reinforces the notion that different rankings produce different results. More importantly, his model determines the shape of a variable grammar – a partial order composed exclusively of unranked constraints.

Even though Anttila’s approach constitutes an improvement over Kiparsky’s and Reynolds’ approaches to variation in OT, his proposal is exceedingly restrictive in the ranking possibilities allowed in the grammar: Only partial grammars are allowed, similar to what was illustrated in (17). It has been shown that this is not always possible without resorting to (sometimes) random constraints for the mere sake of matching probabilities (see Cardoso 2007). In other words, the

partial grammar approach encourages the proliferation of constraints (see Guy 1997, 139 for a critical assessment of the consequences if “OT allows unlimited decomposition of its putative universal constraints”). As was the case with Reynolds’ floating constraints, another serious shortcoming of this approach to variation is that it predicts that frequencies should be in small integer fractions (e.g. 2/3, 1/2, 1/3). While this is certainly the case in the present study (e.g. each variant is likely to occur 1/3 of the time in Other), other studies have shown that this is not always the case (Cardoso 2007; see also Pater and Werle 2001 for a similar view).

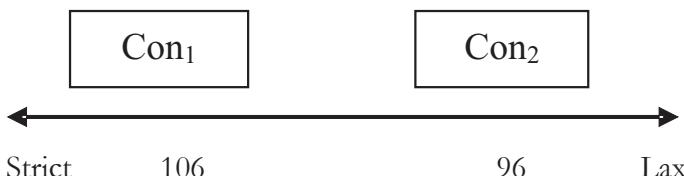
In the following section, I present another approach to variation in OT that addresses the limitations of the three proposals discussed thus far.

### 3.4 Crucial ranking: Stochastic Optimality Theory

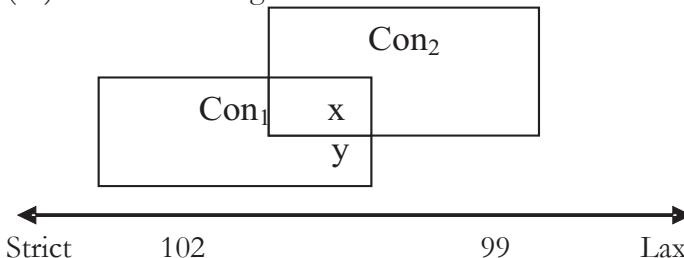
The fourth approach proposed for investigating and representing variability in the framework of Optimality Theory is that of Boersma’s (1998) and Boersma and Hayes’ (2001): Stochastic OT (StOT). StOT employs an associated learning algorithm: the Gradual Learning Algorithm (GLA). Within the StOT approach, variation and gradient wellformedness are accounted for by a probabilistically determined reranking of constraints at certain intervals during evaluation time (i.e. during the process of speaking). Briefly, StOT postulates a continuous scale of constraint strictness in which constraints (e.g.  $Con_1$  and  $Con_2$  in (19)) are annotated with arbitrary numerical strictness values established by a GLA (e.g. Boersma and Weenink’s (2000) Praat program, Hayes et al.’s (2003) OTSoft software). The probability of reranking (i.e. variation) is determined by the distance between  $Con_1$  and  $Con_2$  on the strictness scale and by the amount of evaluation noise (i.e. standard deviation, typically 2.0) added to the strictness values. A standard deviation of 2.0, the amount of noise used in this study, ensures that the distance of 10 or more values between two constraints will yield categorical rankings, as the hypothetical hierarchy in (19) illustrates. Under StOT, constraints

not only dominate other constraints (as is the case in standard OT), but they are also specific distances apart. The two figures in (19) and (20) below illustrate the distinction between a categorical grammar (in which crucially ranked constraints are distant on the strictness scale) and a variable one (in which crucially ranked constraints overlap in their distribution). Note that the superposition of  $Con_2$  above  $Con_1$  in (20) is merely to illustrate the boundaries that separate the overlapping constraints, whose distances are computed in the horizontal dimension only.

(19) Categorical ranking



(20) Variable ranking



In the context of a variable ranking, as shown in (20), the grammar might select for evaluation any point within the overlap of  $Con_1$  and  $Con_2$ . Most likely, the grammar will select the ranking  $Con_1 >> Con_2$  because of the higher ranking of  $Con_1$  over  $Con_2$ . However, it is also possible for the grammar to select a point within the leftmost (higher ranked) area of  $Con_2$  (i.e.  $x$ ) and the rightmost (lower ranked) area of  $Con_1$  (i.e.  $y$ ). In this case,  $Con_2$  is ranked higher than  $Con_1$  (i.e.  $Con_2 >> Con_1$ ) and a different candidate is selected. In sum, the distance between two or more constraints determines variability (e.g. outputs

*a* and *b*) and, more importantly, it encodes predictability of variant occurrence into the grammar (e.g. output *a*'s likelihood of occurrence is 12% while output *b*'s is 88%).

I will now illustrate how StOT is able to account for the geographical distribution of AWRA across the two regions in Vimeu. With the same set of constraints used in the preceding analyses at hand (along with a set of inputs, surface forms and their respective quantitative values established by VARBRUL, incorrect rival candidates, constraints, and constraint violations – just like in a standard OT analysis), a series of computer simulations was performed using the OTSoft 2.1 software package (Hayes et al 2003). In brief, the simulations proceeded as follows: As indicated in section 2, the data set investigated represents two distinct grammars: Nibas and Other. Each of these grammars was individually “learned” by OTSoft, which was supplied with the following information (all numbers are arbitrary) for the learning simulation to take place:

(1) Number of times to go through forms (or total of learning trials): 1,000,000. This number indicates how many learning trials the GLA will perform. The higher the number, the more likely the observed and predicted probabilities will match.

(2) The initial state: 100 for both markedness and faithfulness (the actual value is not essential). By default, the initial state is set with the arbitrary number of 100 for markedness and faithfulness constraints, a value that ensures that the ranking values will be always positive. This value can be manipulated by the researcher depending on his/her views regarding the initial state in language learning. For instance, s/he might decide that the learning process starts with a grammar in which markedness (e.g. 100) is ranked above faithfulness (e.g. 50) – a standard hypothesis for first language acquisition (e.g. Smolensky 1996, Davidson, Jusczyk, and Smolensky 2004, Hayes 2004).

(3) Initial/final plasticity: 2/.002 respectively, which are the default values in OTSoft 2.1. They serve to adjust the GLA results by comparing the outcome of the learning algorithm with the results entered for each pair of input-output. In Hayes' (2004, 21) own words, “[p]lasticity is the size of the change in the grammar that the GLA makes every time its own guess [does not] match the learning datum it encounters.” Note that the algorithm will only make adjustments to the simulated grammar if it detects discrepancies between what is observed and what it predicts – it is *error-driven*.

(4) Number of times to test grammar: 2,000 cycles (default). This number indicates the number of times the GLA will repeat the process of stochastic evaluation and compare the results to the relative frequencies that were observed in the data (the VARBRUL results in our particular case). As will be shown later, the predictions established by the algorithm closely match the frequencies observed in the corpus analyzed.

At the end of the simulations, the algorithm arrived at a final grammar that attempted to mimic the relative frequency of variants in the data, by assigning a ranking value for each of the constraints included in the analysis. Note that there is a degree of randomness in the learning, suggesting that separate GLA simulations will never generate the exact same ranking values and match-up to input frequencies. The procedure described above was repeated for each one of the two grammars established in the investigation.

For the AWRA analysis within StOT, consider the results obtained in Nibas, in which the AWRA variant is more likely to occur (.48) than l-preservation (.27) and l-deletion (.24). These frequencies in the data were learned by the GLA (OTSoft), which generated the following ranking values for Nibas (note that because the constraints overlap in their distribution – similar to what was illustrated in (20) above – the outcome is variation). The ranking values for these constraints are shown in (21). Note that FAITH-Lex (not illustrated in (21) due to

space limitations and its irrelevance to the variable results) is assumed to be ranked at the higher end (approximately 10 values higher than MAX-IO) in the hierarchy of the language. Using standard OT, the hierarchy can be represented as: FAITH-Lex<sup>110</sup> >> MAX-IO<sub>Φ</sub><sup>100.42</sup> >> NoCodaRt<sub>Φ</sub><sup>100.24</sup> Linearity<sub>Φ</sub><sup>99.34</sup> (where the superscripted numbers indicate the ranking value assigned by the GLA – OTSoft).

(21) Nibas: Ranking values

Constraint	Ranking value
MAX-IO <sub>Φ</sub>	100.42
NoCodaRt <sub>Φ</sub>	100.24
Linearity <sub>Φ</sub>	99.34

Because the three equally ranked constraints overlap in their distribution, it is predicted by the values established by the GLA that the set comprised of MAX-IO<sub>Φ</sub> and NoCoda-Rt<sub>Φ</sub> will outrank Linearity<sub>Φ</sub> most of the time (i.e. 48% of the time) since the latter has a lower ranking value; consequently, the outcome will be AWRA (see (22a) for the rankings that select AWRA as the optimal candidate). The GLA also predicts that NoCoda<sub>Φ</sub> will be over-ranked by the two other constraints 28% of the time and l-preservation will result (22b), while MAX-IO<sub>Φ</sub> will occasionally rank at the lower end of the hierarchy (24% of the time) whose consequence is l-deletion (22c). In (22), observe that the values established by the GLA simulations for each AWRA variant (under GLA) generate predictions that closely match those observed in the corpus analyzed (under Obs).

(22)

TABLE 6: Predicted &amp; observed probability of variant occurrence in Nibas

Rankings	Output Selection			Results	
	[I]	AWRA	$\emptyset$	GLA	Obs
a. $\text{MAX-IO}_\phi >> \text{NoCoda-Rt}_\phi$ $>> \text{Linearity}_\phi$		✓		.48	.48
$\text{NoCoda-Rt}_\phi >> \text{MAX-IO}_\phi$ $>> \text{Linearity}_\phi$		✓			
b. $\text{Linearity}_\phi >> \text{MAX-IO}_\phi$ $>> \text{NoCoda-Rt}_\phi$	✓			.28	.27
$\text{MAX-IO}_\phi >> \text{Linearity}_\phi$ $>> \text{NoCoda-Rt}_\phi$	✓				
c. $\text{NoCoda-Rt}_\phi >> \text{Linearity}_\phi$ $>> \text{MAX-IO}_\phi$			✓	.24	.24
$\text{Linearity}_\phi >> \text{NoCoda-Rt}_\phi$ $>> \text{MAX-IO}_\phi$			✓		

The same procedures described above were applied to the results obtained for the other villages. After learning the frequencies entered into the simulation, the algorithm determined the following ranking values for the constraints responsible for the AWRA phenomenon in Other:

(23) Other: Ranking values

Constraint	Ranking value
$\text{Linearity}_\phi$	100.19
$\text{NoCodaRt}_\phi$	99.99
$\text{MAX-IO}_\phi$	99.82

For convenience sake, I will not illustrate or discuss the rankings that select the three AWRA variants in Other. Instead, I summarize in (24) the stochastic analyses for the two grammars that characterize the phenomenon of Across-Word Regressive Assimilation in Picard.

Note that what distinguishes these two grammars from each other is the distance between the same three constraints. For instance, while in Nibas Linearity<sub>Φ</sub> is ranked relatively lower in the hierarchy, in Other, the three constraints are somewhat equidistant (with Linearity<sub>Φ</sub> slightly over-ranking all other constraints). Furthermore, observe under each variant in (24) that each grammar learned by the GLA generates output frequencies that are strikingly close or identical to those observed in the data, as was illustrated in Figure 1.

(24)

TABLE 7: Predicted & observed probability of variant occurrence by geographic location

Grammars of Vimeu Picard	/1/		AWRA		∅	
	Pred	Obs	Pred	Obs	Pred	Obs
<b>Nibas:</b> MAX-IO <sub>Φ</sub> <sup>100.42</sup> >> NoCodaRt <sub>Φ</sub> <sup>100.24</sup> >> Linearity <sub>Φ</sub> <sup>99.34</sup>	.28	.27	.48	.48	.24	.24
<b>Other:</b> Linearity <sub>Φ</sub> <sup>100.19</sup> >> NoCodaRt <sub>Φ</sub> <sup>99.99</sup> >> MAX-IO <sub>Φ</sub> <sup>99.82</sup>	.35	.34	.30	.30	.36	.35

In comparison with the three approaches that have been proposed for the analysis of variable phenomena in OT and, more specifically, in the context of variable AWRA, StOT has proved to be superior: the approach is able to account for the same variable phenomenon (AWRA) via a simpler grammar, with fewer and well-motivated constraints (cf. Reynolds 1994, Anttila 1997, Cardoso 2001, 2003, but see Cardoso 2007 for a similar approach), and predict even more accurate frequency data.<sup>4</sup>

<sup>4</sup> I am aware that StOT has its disadvantages. For instance, the approach is mathematically so powerful that it is possible to match any frequencies with a very small number of reasonable constraints. Whether this is a positive or negative aspect of the approach requires further studies based on both empirical and theoretical grounds.

## 4 Concluding remarks

The study's main goals were: (1) to provide an overview of how variation has been analyzed within the framework of Optimality Theory, and (2) to assess four different approaches that have been proposed for the analysis of variable phenomena in OT. To level the playing field and to provide empirical evidence for my claims and arguments, the overview and the assessment involved the investigation of a single phenomenon: Across-Word Regressive Assimilation in Vimeu Picard, a phonological process that operates variably depending, among other factors, on the geographical distribution of its speakers – while in the village of Nibas the variant AWRA is more likely to occur than l-deletion and l-reservation, in the other villages the three variants are relatively equally likely to occur. While three of the approaches (namely Kiparsky's grammars in competition, Reynolds' floating constraints, and Anttila's partial grammars) were able to adequately capture certain aspects of the phenomenon, they also presented serious limitations that have been minimized or eradicated with the advent of a new way of doing variation in OT: Via stochastic evaluation, in which constraints are not only crucially ranked (as is the case in standard OT), but they occupy specific ranking positions in the hierarchy. Within this approach, variation is the mere result of when two or more constraints overlap in their distribution within a given strictness scale.

For the analysis of variation, this paper argued in favor of a stochastic version of the framework of Optimality Theory: The Gradual Learning Algorithm proposed by Boersma (1998) et seq. and Boersma and Hayes (2001). I have argued in the context of variable data from AWRA that this approach is superior in comparison with its predecessors because it accounts for variation and its frequency effects via the same linguistic constraints and principles that govern categorical phenomena (e.g. the crucial ranking of constraints). The result is the stipulation of simpler and more accurate grammars, with fewer constraints (cf. Reynolds 1994, Anttila 1997, Cardoso 2001, 2003): “[a] grammar with fewer constraints

should in principle be preferred to a grammar with more constraints, providing they make identical predictions" (Asudeh 2001, 9).

## References

- ALBER, Birgit. **Regional variation at edges**: Glottal stop epenthesis and dissimilation in Standard and Southern varieties of German. Ms., University of Padova. To appear in *Zeitschrift für Sprachwissenschaft*. 2001.
- ANTTILA, Arto. **Deriving variation from grammar**: A study of Finnish genitives. In: HINSKENS, F.; HOUT R. van; WETZELS L. (eds.), **Variation, Change and Phonological Theory**. Amsterdam: John Benjamins, 1997. p. 35-68. (First appeared as a manuscript in 1995, Stanford University).
- ASUDEH, Ash. **Linking, optionality, and ambiguity in Marathi**. In: SELLS, Peter (ed.), **Formal and empirical issues in optimality-theoretic syntax**. Stanford, CA: CSLI Publications. 2001. p. 257-312.
- BERGEN, Benjamin K. **Probability in phonological generalizations**: Modeling French optional final consonants. Ms., University of California, Berkeley and International Computer Science Institute. 2000.
- BENUA, Laura. **Identity effects in morphological truncation**. In: BECKMAN, J.; DICKEY, L. Walsh-;URBANCZYK, S. (eds.), **University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory**. Amherst, MA: GLSA, 1995. p. 77-136.
- BECKMAN, Jill. **Positional faithfulness, positional neutralization and Shona vowel harmony**. *Phonology* 14, 1997. p. 1-46.

BOBALJIK, Jonathan David. **Assimilation in the Inuit languages and the place of the uvular nasal.** International Journal of American Linguistics 62, 1996. p. 323-350.

BOERSMA, Paul. **Functional phonology.** Formalizing the interactions between articulatory and perceptual drives. LOT 11. The Hague, Netherlands: Holland Academic Graphics, 1998.

BOERSMA, Paul. **Review of Arto Anttila:** Variation in Finnish phonology and morphology. GLOT International 5, 2001. p. 31-40.

BOERSMA, Paul; HAYES, Bruce. **Empirical tests of the gradual learning algorithm.** Linguistic Inquiry 32, 2001. p. 45-86.

BOERSMA, Paul; WEENINK, David. **Praat, a system for doing phonetics by computer.** 2000. Disponível em: <<http://www.praat.org>>

CARDOSO, Walcir. **The domain of across-word regressive assimilation in Picard** – an optimality theoretic account. Southwest Journal of Linguistics 17(2), 1999. p. 1-22.

\_\_\_\_\_. **Variation patterns in regressive assimilation in Picard.** Language Variation and Change 13, 2001. p. 305-341.

\_\_\_\_\_. **Topics in the phonology of Picard.** PhD thesis, McGill University. McGill Working Papers in Linguistics, 2003.

\_\_\_\_\_. **An integrated approach to variation in Optimality Theory: Evidence from Brazilian Portuguese and Picard.** In: GEERTS; GINNEKEN van; JACOBS (Eds.), **Romance Languages and Linguistic Theory.** Amsterdam and Philadelphia: John Benjamins, 2005. p. 1-15.

\_\_\_\_\_. **The variable development of English word-final stops by Brazilian Portuguese speakers:** A stochastic optimality theoretic account. Language Variation and Change 19 (3). 2007.

- CASALI, Roderic. **Resolving hiatus**. PhD dissertation, UCLA. 1996.
- \_\_\_\_\_. **Vowel elision in hiatus contexts**: Which vowel goes? *Language* 73, 1997. p. 493-533.
- CLEMENTS, George N. **Berber syllabification**: Derivations or constraints? In: ROCA, I. (ed.), **Derivations and Constraints in Phonology**. Oxford: Oxford University Press, 1997. p. 289-330.
- CEDERGREN, Henrietta. **The interplay of social and linguistic factors in Panama**. PhD dissertation, Cornell University, 1973.
- CEDERGREN, Henrietta; SANKOFF, David. **Variable rules**: Performance as a statistical reflection of competence. *Language* 50, 1974. p. 333-355.
- COETZEE, Andries W. **Variation as accessing ‘non-optimal’ candidates**. *Phonology* 23, 2006. p. 337-385.
- DAVIDSON, Lisa; JUSCZYK, Peter; SMOLENSKY, Paul. **The Initial and Final States**: Theoretical Implications and Experimental Explorations of Richness of the Base. In: KAGER, R.; PATER, J.; ZONNEVELD, W. (Eds.), **Constraints in phonological acquisition**. Cambridge: Cambridge University Press, 2004. p. 321-368.
- DEBRIE, René. **Problèmes posés par la présence de l'assimilation régressive dans le sud-ouest du domaine picard**. *Revue de Linguistic Romane* 45, 1981. p. 422-464.
- DORIAN, Nancy C. **The problem of the semi-speaker in language death**. *International Journal of the Sociology of Language* 12, 1977. p. 23-32.

DRESSLER, Wolfgang. **On the phonology of language death.** In: PERANTEAU, P. (ed.), **Papers from the Eighth Regional Meeting of the Chicago Linguistic Society.** Chicago: Chicago Linguistic Society, 1972. p. 448-457.

\_\_\_\_\_. **Language death.** In: NEWMEYER, F. (ed.), **Linguistics: The Cambridge Survey**, v. IV. Cambridge: Cambridge University Press, 1988. p. 184-92.

FASOLD, Ralph. **The quiet demise of variable rules.** In: SINGH, R. (ed.), **Towards a critical sociolinguistics.** Amsterdam: Benjamins, 1996. p. 79-98.

GUY, Gregory. **Use and application of the Cedergren-Sankoff variable rule program.** In: FASOLD, R.; SHUY, R. (eds.), **Analyzing Variation in Language.** Papers from the Second Colloquium on News Ways of Analyzing Variation. Washington, D.C.: Georgetown University Press, 1975. p. 59-69.

\_\_\_\_\_. **Competence, performance and the generative grammar of variation.** In: HINSKENS, F.; HOUT, R. van; WETZELS, W. (eds.), **Variation, Change and Phonological Theory.** Amsterdam: John Benjamins, 1997. p. 125-143.

HAYES, Bruce. **OTSoft: Constraint Ranking Software – Manual.** University of California, Los Angeles (UCLA), 2004/2007.

\_\_\_\_\_; TESAR, Bruce; ZURAW, Kie. **OTSoft 2.1** software package. 2003. Disponível em: <<http://www.linguistics.ucla.edu/people/hayes/otsoft>>

- INKELAS, Sharon. **Phonotactic blocking through structural immunity.** Ms., University of California, Berkeley, 1997.
- \_\_\_\_\_; ORGUN, Orhan; ZOLL, Cheryl. **The implications of lexical exceptions for the nature of grammar.** In: ROCA, I. (ed.), **Derivations and Constraints in Phonology.** Oxford: Oxford University Press, 1997. p. 393-418.
- \_\_\_\_\_; ZOLL, Cheryl. **Reduplication as morphological doubling.** Ms., University of California, Berkeley and Massachusetts Institute of Technology, 2000.
- ITÔ, Junko; MESTER, Ralf-Armin. **The core-periphery structure of the lexicon and constraints on reranking.** In: BECKMAN, J.; DICKEY, L. Walsh-; URBANCZYK, S. (eds.), **University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory.** Amherst, MA: GLSA. 1995a. p. 181-209.
- \_\_\_\_\_; \_\_\_\_\_. **Japanese phonology.** In: GOLDSMITH, J. (ed.), **The Handbook of Phonological Theory.** Oxford: Blackwell, 1995b. p. 817-838.
- KAWASAKI, Takako. **Codaconstraints:Optimizing representations.** PhD dissertation, McGill University, 1998.
- KIPARSKY, Paul. **Variable rules.** Paper presented at the Rutgers Optimality Workshop, New Brunswick, N.J. 1993.
- KIRCHNER, Robert. **Contrastiveness is an epiphenomenon of constraint ranking.** Proceedings of the Berkeley Linguistics Society 21, 1995. 198-208. [ROA-51].
- KROCH, Anthony. **Morphosyntactic variation.** In: BEALS, K. et al. (eds.), **Papers from the 30th Regional Meeting of the Chicago Linguistic Society v. 2: The Parasession on Variation in Linguistic Theory.** 1994. p. 180-201.

LABOV, William. **Contraction, deletion, and inherent variability of the English copula.** *Language* 45, 1969. p. 715-762.

\_\_\_\_\_. **Sociolinguistic patterns.** Philadelphia: University of Pennsylvania Press, 1972.

LOMBARDI, Linda. **Laryngeal neutralization, alignment, and markedness.** In: BECKMAN, J.; DICKEY, L. Walsh-; URBANCZYK, S. (eds.), **University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory.** Amherst, MA: GLSA, 1995. p. 225-248.

MCCARTHY, John; PRINCE, Alan. **Faithfulness and reduplicative identity.** In: BECKMAN, J.; DICKEY, L. Walsh-; URBANCZYK, S. (eds.), **University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory.** Amherst, MA: GLSA, 1995. p. 49-384.

\_\_\_\_\_; \_\_\_\_\_. **Generalized alignment.** In: BOOIJ, G.; MARLE, J. van (eds.), **Yearbook of Morphology.** Dordrecht: Kluwer Academic, 1993. p. 79-153.

NESPOR, Marina; VOGEL, Irene. **Prosodic phonology.** Dordrecht: Foris.20, 1986.

ORGUN, Cemil Orhan. **Sign-based morphology and phonology with special attention to Optimality Theory.** PhD dissertation, University of California, Berkeley, 1996.

PATER, Joseph; WERLE, Adam. **Typology and variation in child consonant armony.** In: FÉRY, C.; GREEN, A. Dubach; VIJVER, R. van de (eds.). **Proceedings of HILP5.** University of Potsdam, 2001.

- PIGGOTT, Glyne. The phonotactics of a 'Prince' language: a case study. In: **Living on the Edge**: 28 Papers in honour of Jonathan Kaye. Mouton de Gruyter: Berlin, 2003. p. 401-425.
- PINTZUK, Susan. **VARBRUL programs [computer program]**. Philadelphia: University of Pennsylvania Department of Linguistics, 1988.
- PRINCE, Alan; SMOLENSKY, Paul. **Optimality Theory**: Constraint Interaction in Generative Grammar. Oxford/Cambridge: Blackwell Publishers, 1993/2004.
- PULLEYBLANK, Douglas. **Optimality Theory and features**. In: ARCHANGELI, D.; LANGDOEN, D. T. (eds.), **Optimality Theory – An Overview**. Oxford: Blackwell, 1997. p. 59-101.
- REYNOLDS, William. **Variation and phonological theory**. PhD thesis, University of Pennsylvania. Selkirk, Elisabeth 1972. The phrase phonology of English and French. PhD thesis, MIT. Distributed in 1981 by the Indiana University Linguistics Club, Bloomington, Indiana, 1994.
- SELKIRK, Elisabeth. **Prosodic domains in phonology: Sanskrit revisited**. In: ARONOFF, M.; KEAN, M.-L. (eds.), **Juncture**. Saratoga, California: Anma Libri, 1980. p. 107-129.
- \_\_\_\_\_. **The prosodic structure of function words**. In: MORGAN J.; DEMUTH, K. (eds.), **Signal to Syntax**: Bootstrapping from speech to grammar in early acquisition. Mahwah, NJ: Lawrence Erlbaum, 1997. p. 187-213.

SMOLENSKY, Paul. **The initial state and ‘richness of the base’ in Optimality Theory**. Technical Report JHU-CogSci-96-4, Cognitive Science Department, Johns Hopkins University. Rutgers Optimality Archive: #154, 1996.

STERIADE, Donca. **Underspecification and markedness**. In: GOLDSMITH, J. (ed.), **The Handbook of Phonological Theory**. Cambridge, Massachusetts: Blackwell, 1995. p. 114-174.

TALER, Vanessa. **S-Weakening in the Spanish of San Miguel, El Salvador**. MA thesis, McGill University. 1997.

TRUBETZKOY, Nikolai. **Grundzuge der phonologie**. Göttingen: Vandenhoeck und Ruprecht. 1939. Appears in English translation as *Principles of Phonology*, translated by C. Baltaxe. Berkeley and Los Angeles: University of California Press, 1969.

VASSEUR, Gaston. **Dictionnaire des parlers picards du Vimeu (Somme)**. Collection de la Société de Linguistique Picarde. Amiens, France: Musée de Picardie, 1963.

\_\_\_\_\_. **Grammaire des parlers picards du Vimeu – Avec considération spéciale du dialecte de Nibas**. Abbeville, France: Le Conseil Régional de Picardie, F. Paillart, 1996.

Revista  
ABRALIN

Normas para apresentação dos trabalhos



ASSINATURAS / SUBSCRIPTION

Revista  
ABRALIN

Fonte AGaramond, corpo 12/15, impressa pela Ideia Editora Ltda.  
(ideia.editora@uol.com.br), em papel Chamoix 80gr. (miolo),  
capa em Supremo 250 gr.