A gente adquire o vernáculo e *aprende-se o padrão*: sobre as construções de indeterminação no PB

Marco Antonio Martins (PGL-UFSC)

1. Introdução

O objetivo mais imediato deste artigo é, a partir de resultados estatísticos de um estudo sociolingüístico acerca das construções de indeterminação com *se* no Português do Brasil – PB – (cf. MARTINS 2005), evidenciar que a escola se depara (quase sempre sem estar consciente disso) com distintas gramáticas (o vernáculo e o padrão) no ensino de língua. Nossa proposta parte do pressuposto de que, nos reportando à comunidade de fala do PB sem distinções de gerações, há a possibilidade de alternância entre estruturas como (i) *vende(m)-se verdades*, (ii) *se vende(m) verdades* e (iii) *a gente/você/tu vende verdades*.

O artigo está assim dividido: Apresentaremos os resultados estatísticos de Martins (2005) em relação aos condicionadores estilístico-sociais das variantes acima descritas serão apresentados na terceira seção. Na segunda seção faremos uma inserção aos pressupostos do modelo teórico proposto por Anthony Kroch de gramáticas em competição (cf. KROCH 1989; 1994; 2001). Na quarta e última seção, enfim, correlacionaremos os resultados de Martins aos pressupostos de Kroch, tendo em vista evidenciar um fenômeno periférico do PB que reflita um caso de gramáticas em competição nessa língua. De modo que a competição está vinculada a fatores sócio-estilísticos e parece refletir diferentes gramáticas atreladas à variante vernacular (efetivamente adquirida) e à variante padrão (efetivamente – ou não tão bem assim – aprendida na escola).

2. Sobre Freqüências e Probabilidades de Uso: As Construções com *SE indeterminador*

Martins (2004, 2005) propõe uma análise derivacional, na esteira do Programa Minimalista de pesquisa (cf. CHOMSKY 1995), para as construções de indeterminação (com *se*) no PB. De acordo com a proposta de Martins, o *se indeterminador enclítico* possui propriedades morfológicas bastante específicas numa relação *composicional* com o verbo, de modo que o que

entra na derivação de objetos sintáticos no sistema computacional é um todo morfológico formado pelo verbo e o "morfema" *se indeterminador*. Observe-se o paradigma em (1) em que o *se indeterminador enclítico* se mantém na estrutura das construções com a possibilidade se ser elidido na estrutura linear (cf. 1b e 1c).

(1) Nesta terra <u>vendem-*se*</u> balaios.

(1a) Nesta terra vende-*se* balaios.

(1b) Nesta terra *pro expletivo* vende(-*se*) balaios.

(1b') Nesta terra *a gente* vende(-*se*) balaios.

(1b'') Nesta terra *você* vende(-*se*) balaios.

(1b''') Nesta terra *tu* vende(-*se*) balaios.

(1c) Nesta terra se vende(-*se*) balaios

Ancorado nesta proposta de análise formal para a estrutura em questão, Martins (2005) apresenta um estudo sociolingüístico sincrônico tomando por regra variável o processo que alterna as construções com *se indeterminador enclítico* realizado foneticamente (cf. (1a)), e as construções com *se indeterminador enclítico* elidido na estrutura linear, com *pro expletivo* (cf. (1b)) ou com a realização do *pronome a gente/você/tu* (cf. (1b', 1b'' e 1b''')) [1].

Apresentaremos, nesta seção, os resultados estatísticos do estudo de Martins acerca da realização/elisão de *se indeterminador enclítico* em estruturas de indeterminação no PB, conforme (1), (2) e (3) a seguir. Nosso foco de atenção se voltará às variáveis sócio-estilísticos em duas amostras analisadas pelo autor, a saber, uma amostra de língua falada com dados extraídos do banco de dados VARSUL[2] e outra de língua escrita com dados extraídos da revista Veja.

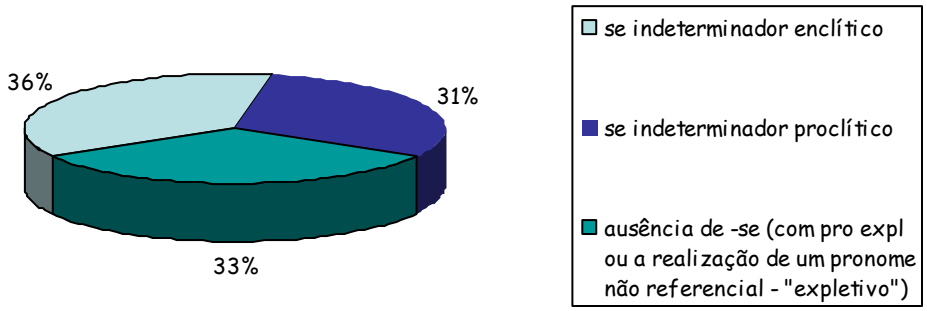
Em relação à amostra da revista Veja, houve um certo "equilíbrio" no uso das variantes em questão. Dos 160 dados da amostra, 57 (i. e. 36%) foram de construções com *se indeterminador enclítico* realizado foneticamente (cf. (1)); 50 (i. e. 31%) de construções com *se*

[1] Os resultados estatísticos apresentados por Martins apontam para o fato de que o estatuto destas construções de indeterminação com pronomes, como exemplifica (1b', 1b'' e 1b'''), parece ser o mesmo daquelas com o pronome *se proclítico*, conforme (1c). Deste modo, estatisticamente, o sistema estrutural das construções com -*se indeterminador enclítico* parece ser distinto daquele das construções com o pronome *se* proclítico. Fato este que corrobora a análise formal proposta por Martins (2004; 2005) para as construções em questão.

[2] Projeto de Variação e Mudança Lingüística no Sul do Brasil.

indeterminador proclítico (cf. (2)), também realizado foneticamente; e 53 (i. e. 33%) foram de construções *sem a realização do se indeterminador* (com ***pro*** *expl* ou a realização de um *pronome a gente/você/tu*) (cf. (3)), conforme ilustra o *gráfico 2*, a seguir[3].

Gráfico 2: Distribuição geral de aplicação da regra variável nos dados da revista Veja



(2) Antigamente dizia-**se** que o conhecimento humano dobrava a cada dezoito meses. *Veja 6 de ago. de 2003 – PV*

(3) Embora numa primeira observação **se** possa colocar tais decisões na conta do pragmatismo, ou do calor de um momento de crise, há algo mais nessas escolhas. *Veja 6 de ago de 2003 EF*

(4) Antigamente **você** precisava entender de mecânica para dirigir um carro. *Veja 6 de ago. de 2003 – PV*

Na amostra de língua escrita da revista Veja, ainda, a variável *Tipo de texto*[4] mostra-se um fator favorável ao uso da variante sem *se*, conforme observamos na *tabela Veja 1*, a seguir. Enquanto os editoriais apresentam 19% de aplicação das construções sem *se* e um peso de 0,32, as entrevistas apresentam um percentual de 38% de aplicação e peso de 0,57. Entendemos aqui que estes resultados refletem o fato de as entrevistas, apesar de passarem por edição, e, conseqüentemente, por revisões e todas os pressupostos que tal processo acarreta, apresentam

[3] Das 53 construções sem *se*, 47 são com a realização de um pronome não referencial ("expletivo"); destas, 40 ocorrências são com a forma *você* e 7 com o pronome *a gente*.

[4] Esta variável foi a segunda selecionada pelo programa estatístico como significativa no uso da construção sem *se* indeterminador enclítico.

uma maior correspondência com a língua falada o que caracterizaria uma maior recorrência de possibilidade de uso da variante inovadora sem *se*.

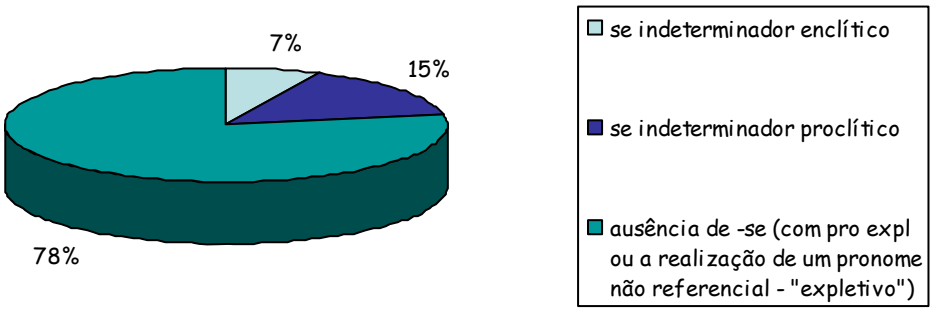
	Apl/total	%	PR.
Editoriais	8/42	19	0,32
Entrevistas	45/118	38	0,57
Total	53/160	33	

LOG LIKELIHOOD: -91.174 SIGNIFICANCE: .032

Tabela Veja 1: Freqüência e probabilidade de *pro expl (pronome a gente/você/tu) V*, segundo a variável *Tipo de texto*

Do total de 366 dados extraídos de dezesseis entrevistas do Banco de dados do VARSUL, podemos observar que a realização do *se indeterminador enclítico* ou *proclítico* é bastante inferior à realização das construções sem *se*. Enquanto o uso desta soma 285 dos 366 dados da amostra (78%), o que inclui aquelas construções com o *pronome a gente/você/tu* [5], a ocorrência daquelas soma, apenas, 81 sentenças (22%) do total conforme ilustra o *gráfico 3*, a seguir.

Gráfico 3: Distribuição geral de aplicação da regra variável nos dados do VARSUL



[5] Destas 285 construções sem *se*, 206, o que equivale 72% do total, são com a realização de um pronome não referencial ("expletivo"); ou seja, com o preenchimento lexical por pronome, tal como *a gente*, *você* e *tu*. Das 206 construções com pronomes, 110 são ocorrências com a forma *a gente*; 49 com a forma *você;* e 47 com a forma *tu*.

Tal distribuição, quando confrontamos os resultados obtidos nas duas amostras (da revista Veja e do VARSUL), parece ser um indício de que o processo de mudança no que concerne às construções de indeterminação com *se* está bastante adiantado na fala[6]. Em outras palavras, a freqüência de uso da variante inovadora *Ausência do clítico se indeterminador*, sendo que em 72% do total das construções com ausência do *se* são de estruturas com a realização de um *pronome a gente/você/tu* (cf. nota 7), é significativamente superior ao uso das demais variantes (i. e. das construções com *se enclítico* e/ou *proclítico*).

Na alternância entre uma gramática e outra, ou entre o uso do *se* e o uso *pro expl* (ou a realização de um *pronome a gente/você/tu*)[7], a variável selecionada nas três rodadas foi a *Escolaridade*, ou seja, uma variável social. Nos valendo da proposta de Kroch (1989, 2001), parece que as construções com e sem *se* indeterminador apresentam-se como opções gramaticalmente incompatíveis, alternando-se, apenas, em diferentes registros sociais marcados pela escolaridade.

	Apl/total	%	PR.
Primário	183/210	87	0,63
Colegial	102/156	65	0,33
Total	285/366	78	

LOG LIKELIHOOD: -166.802 SIGNIFICANCE: .050

Tabela VARSUL 1: Freqüência e probabilidade de *pro expl (ou um pronome a gente/você/tu) V*, segundo a variável *Escolaridade* [8]

Como observamos nos resultados da tabela VARSUL 1 acima, a tendência ao uso da variante inovadora, ou das construções sem *se* (com uma freqüência maior de realização do *pronome a gente/você/tu*), segundo a variável social *Escolaridade* é significativa. Ou seja, no uso

[6] E nos referimos aqui a um processo de mudança de caráter estrutural envolvendo construções, uma vez que, apesar de superficialmente tais construções parecerem pertencer a uma mesma gramática (i.e. uma mesma estrutura), são bastante distintas.

[7] O pressuposto aqui é o de que tais variantes (com e sem *se* indeterminador) pertencem a distintas gramáticas, muito embora superficialmente pareçam a mesma, como veremos com a discussão dos resultados das rodadas estatísticas mais adiante. Na análise dos dados, mesmo sem termos controlado cada informante separadamente, observamos a alternância de uso na fala de um mesmo indivíduo entre as variantes aqui controladas.

[8] Nas duas rodadas subseqüentes, Rodadas 2 e 3, em que a aplicação da regra é a construção com *se* enclítico "confrontada" com o *se* proclítico e com o *se* proclítico somado às construções com *pro expl* sem *se*, os percentuais e as probabilidades desta variável, respectivamente são: na Rodada 1, 15% e 0,25 para a variante *Primário* e 39% e 0,64 para a variante *Colegial*; na Rodada 2, 2% e 0,29 para a variante *Primário* e 13% e 0,77 para a Variante *Colegial*. Ou seja, enquanto a variante *Primário* tende a ser um ambiente favorável a aplicação da variante inovadora sem se, mostra-se inibidora das construções com *se* enclítico.

destas construções a variante *Primário*, com um peso de 0,64, parece ser uma variável favorecedora da aplicação à regra da variante inovadora, enquanto que no uso das demais variantes com *se*, com pesos de 0,25 e 0,29 conforme nota 10 (i. e. um menor grau de escolaridade – *Primário*), não.[9]

Ainda em relação à freqüência de uso e probabilidade da variante inovadora e aos fatores sociais, na Rodada 1 cuja aplicação da regra foi a construção sem *se* (com *pro expl* ou *pronome a gente/você/tu*) a variável *Idade* também se mostrou significativa. Tal variável foi a quarta selecionada pelo VARBRUL na amostra do VARSUL. Como esperado[10], a probabilidade de uso da variante inovadora está atrelada à faixa etária mais jovem, com 87% de aplicação dos dados e um peso de 0,62 (cf. *tabela VARSUL 2*, a seguir) [11].

	Apl/total	%	PR.
De 25 a 49 anos	88/101	87	0,62
Mais de 50 anos	197/265	74	0,45
Total			

LOG LIKELIHOOD: -166.802 SIGNIFICANCE: .050

Tabela VARSUL 2: Freqüência e probabilidade de *pro expl (ou o pronome a gente/você/tu) V*, segundo a variável *Idade*[12]

O que mostram tais resultados é que na língua falada o uso das variantes com *-se indeterminador enclítico* realizado foneticamente e com a elisão de *-se indeterminador enclítico* na estrutura linear é, acima de tudo[13], estilisticamente marcado, no sentido de que enquanto o uso da variante inovadora sem *se* está atrelado a uma fala vernacular, sem o monitoramento da escola (pensada aqui como apregoadora de uma gramática normativa), o uso da variante conservadora, com o clítico *se* (*enclítico*, essencialmente, porque, a realidade de *se proclítico* assemelha-se àquelas construções com a realização do *pronome a gente/você/tu*), é favorecido pela influência da escola, ou por um maior nível de escolaridade – *Colegial* (cf. os dados na nota 11).

[9] Nunes (1990) em seu estudo sobre a reanálise de *se* apassivador para *se* indeterminador nas passivas pronominais destaca que a queda da concordância naquelas estruturas é sensível à escolaridade.
[10] Os estudo sociolingüísticos vêm apontando que a variante inovadora está, quase sempre, atrelada às faixas etárias mais jovens (cf., dentre muitos outros, os artigos em MOLLICA & BRAGA 2004).
[11] Estudos de uma amostra de faixas etárias inferiores àquelas utilizadas na análise, a fim de se verificar se a curva da mudança se manteria, ou seja, que os mais jovens utilizam mais a variante inovadora, seriam bastante interessantes.
[12] Cf. também os resultados de Duarte (1995), nos quais o uso do *se* indeterminador se limita apenas à fala de informantes de uma faixa etária acima de 50 anos.
[13] Digo "acima de tudo" porque variáveis lingüísticas também foram selecionadas pelo programa em duas das rodadas (cf. MARTINS 2005).

Segundo Duarte (1999, p.110), se referindo ao *se indeterminador*,

> enquanto decresce o uso do pronome *se*, que alcança, na fala de informantes mais velhos, percentuais próximos aos da construção sem o pronome, na fala de informantes mais jovens, aumenta o uso de formas pronominais plenas, entre as quais destaca-se o uso de *você*.

De acordo com os resultados apresentados nesta seção, nota-se que enquanto na língua falada, conforme os dados da amostra do VARSUL, o uso do *se* indeterminador já é bastante pequeno (e quando este é enclítico, o seu uso é ainda mais restrito), na língua escrita, representada pela amostra da revista Veja, há uma distribuição um tanto que "equiparada" das variantes em questão. Tais resultados nos apontam para uma realidade bastante significativa em relação aos clíticos como um todo, e em especial ao *se* indeterminador, no PB. Em outras palavras, há alguns resquícios de ênclise em construções com *se* indeterminador na língua falada, aparentemente, condicionados por determinados fatores sociais, atrelados a uma gramática específica, muito provavelmente pelo monitoramento de uma gramática normativa, como L2 [14], enquanto o seu uso na língua escrita se faz mais presente.

3. A Arquitetura Teórica: *Gramáticas em Competição*

Um dos pressupostos principais dos estudos em gramática gerativa é o de que os indivíduos são geneticamente dotados de uma Gramática Universal (GU) (numa visão biológica), o que implica um modelo formal de competência lingüística pautado em regras categóricas. Anthony Kroch (1989; 1994; 2001) oferece um modelo de variação e mudança lingüística ancorado no pressuposto de que falantes podem operar com mais de uma gramática, num tipo de "diglossia internalizada". De modo que um período de mudança pode ser pensado como uma competição entre gramáticas, ou entre opções gramaticalmente incompatíveis: uma conservadora e uma inovadora.

De acordo com Kroch, a diglossia e o bilingüismo são evidências de que um falante possui uma competência que abarca o domínio de diferentes sistemas lingüísticos; isto é, o

[14] Seguindo a proposta de Kato (a sair).

domínio de mais de um sistema, ou de mais de uma gramática. A variação lingüística pode ser entendida, muitas vezes, como o reflexo de *"escolhas"* que são somente probabilisticamente influenciadas por fatores tais como o contexto e a situação.

Um dos argumentos do autor a favor de sua proposta é que toda mudança segue a forma da curva em "S"; ou seja, a mudança se processa via competição entre opções gramaticalmente incompatíveis com a substituição de uma forma pela outra no uso, efetivamente. Um outro argumento apresentado por Kroch é o fato de que o ser humano é sensível às freqüências em suas mais variadas atividades, o que inclui a atividade lingüística. De modo que há no processo de aquisição, pelo qual a criança adquire a língua vernacular e as outras opções lingüísticas que ela adquire/aprende em outras instâncias (sociais, regionais etc), uma assimetria (pensada aqui como uma "falha" de transmissão. Em outras palavras, por algum motivo cuja origem é sócio e psicolingüística a freqüência de uso de uma dada forma (atrelada às feições de uma determinada gramática) é interpretada erroneamente, o que motiva uma pequena e constante assimetria entre um sistema (o vernacular) [15] e outro (adquirido tardiamente).

Dentro desta proposta, sempre que a freqüência superficial do uso de uma dada forma/estrutura estiver mudando, os falantes são hábeis (i. e. competentes) para associar corretamente tais mudanças com suas origens (causas) gramaticais[16]. A proporção da mudança, deste modo, segue uma escala logística (freqüência de uso) que motiva um efeito de cadeia – pensado aqui como o problema do *encaixamento* lingüístico (cf. proposta de WEINREICH, LABOV & HERZOG 1968). Adiante, uma das formas pode se tornar obsoleta por preferências estilísticas ou flutuações estatísticas ou as duas formas podem se tornar estáveis no sistema devido a diferenças de sentido e/ou de propriedades gramaticais.

Conforme as asserções de Kroch, o conjunto de contextos que mudam ao mesmo tempo na estrutura de um sistema lingüístico não é definido pelo agrupamento de uma propriedade superficial, como o aparecimento de uma palavra ou de um morfema particular, mas pela estrutura sintática, cuja existência pode somente ser o produto de uma análise gramatical abstrata independente dos falantes. Nessa perspectiva, a reflexão sobre o processo de mudança lingüística

[15] O vernáculo é a língua sempre mais automatizada, mais psicologicamente rápida e mais acessível nos processos de produção e processamento da fala. Para Labov (2003 [1969], p. 241), um estilo mais formal está, necessariamente, associado a um processo mental atrelado a uma maior atenção à fala que aquele de um estilo mais casual e espontâneo do uso da linguagem, para o qual, segundo o autor, o mínimo de atenção é dado.

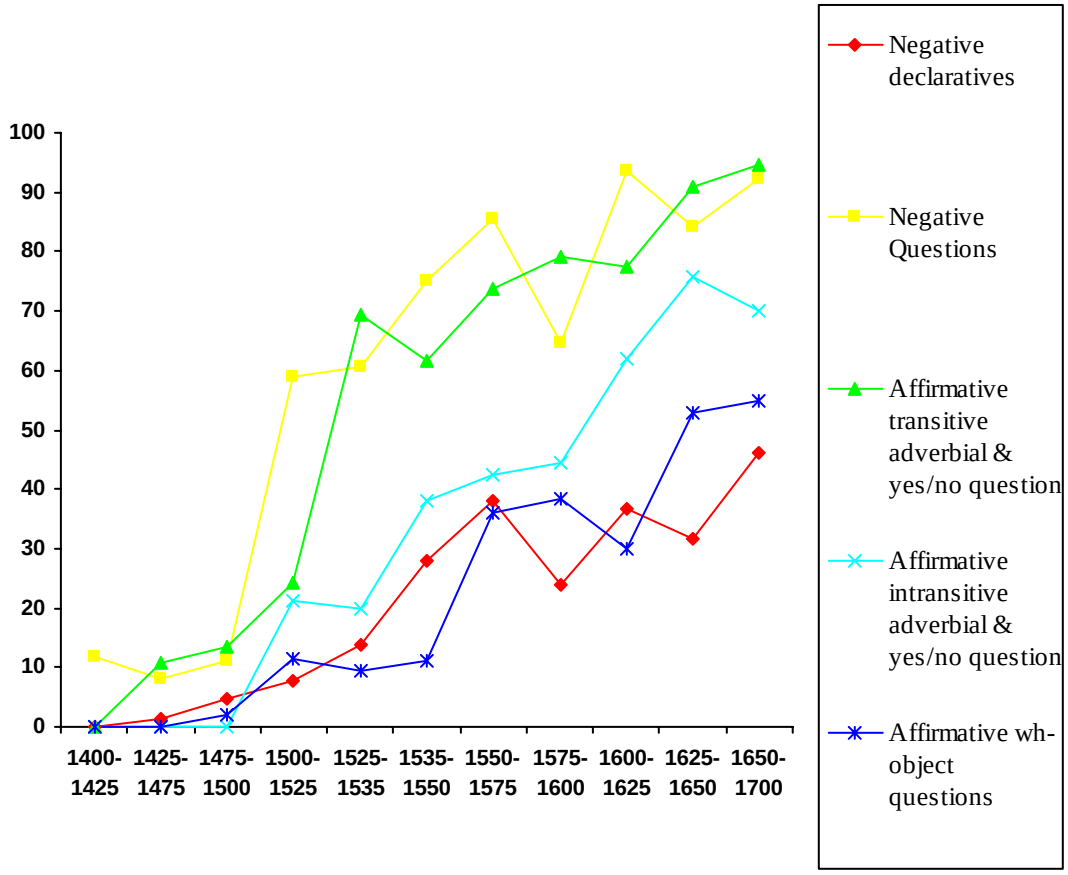
[16] Segundo Kroch, a freqüência de uso a que uma criança está exposta quando em processo de aquisição de língua é fundamental para que ela marque os traços paramétricos dessa língua.

é um convite à “garimpagem” de tais fenômenos “superficiais” em busca da origem, ou da fonte, de tal “desvio” no curso do sistema. Deste modo, a análise gramatical que define o contexto de uma mudança lingüística deve ser bastante abstrata.

De acordo com a proposta de Kroch, uma criança tem acesso direto somente aos dados da língua em uso, não à gramática que o falante usa para produção dos dados de fala; e o processo de inferência através do qual a criança tira conclusões frente aos dados está sujeito a erros (falhas, possíveis de serem formadas pelo sistema – processos formativos sintáticos e morfológicos, conforme Kroch (1994)).

Para o autor, em todo contexto lingüístico cujo sistema superficial reflete uma dada mudança sintática, a alteração na freqüência de uso se dá numa mesma proporção e de maneira ordenada, tendo em vista que todo processo de mudança lingüística pressupõe variação entre formas lingüísticas velhas e inovadoras. Em outras palavras, a proporção da mudança lingüística em diferentes contextos superficiais refletindo uma simples mudança paramétrica é, predominantemente, a mesma. Esse fenômeno é denominado por Kroch de *Efeito da Proporção Constante*. Kroch traz evidências empíricas com dados acerca da entrada do *Do perifrástico* no sistema do Inglês Médio em vários contextos (cf. gráfico 1, a seguir).

Gráfico 1: The rise of periphrastic do (adaptado de KROCH 1989, p. 22)



Segundo Kroch, baseado em resultados de Ellegard (1953), a curva logística da freqüência de uso dos vários contextos reflete uma mudança gramatical (ou paramétrica) na estrutura desta língua (de modo que a proporção da entrada do *Do* é a mesma em várias construções na gramática superficial).

4. Enfim, *a gente adquire o vernáculo* e *aprende-se o padrão*

De acordo com os resultados aqui delineados, parece que o PB está operando com dois sistemas (ou duas gramáticas, na proposta de Antony Kroch) aparentemente bastante distintos: de um lado, as construções com *-se indeterminador enclítico* elidido na estrutura linear e com a realização do *pronome a gente/você/tu*, em estruturas tais como *(i) a gente vende casas nesta cidade*; e, de outro lado, as construções com a realização de *-se indeterminador enclítico* em

estruturas como *(ii) vende-se casas nesta cidade*, atreladas a uma língua conservadora como confirma a variável *Escolaridade*.

Retomando alguns resultados, observamos que (i) a construção sem *-se indeterminador enclítico (*realizado foneticamente) com os pronomes *a gente/você/tu* possui alta freqüência de uso (78% na amostra do VARSUL e 33% na amostra de LÍNGUA ESCRITA da revista Veja) – o que caracteriza um ambiente propício como *input* para aquisição; (ii) o uso de *-se indeterminador enclítico* é fortemente condicionado pelo fator *Escolaridade* (0,64 – tabela VARSUL 2; 0,77 – tabela VARSUL 3).

Esta situação parece ganhar cores mais estáveis na amostra de língua falada, conforme podemos observar nos dados do VARSUL, de modo que o uso da variante conservadora com *-se indeterminador enclítico* está sempre associado a uma L2, influenciada por um nível de escolaridade sempre alto, assim como a falantes de mais idade.

Referências bibliográficas

CHOMSKY, Noam.. (1995). *The minimalist program*. Cambridge. Mass: MIT Press.

DUARTE, Maria Eugênia Lamoglia. *A perda do princípio "evite pronome" no Português Brasileiro*. Tese de Doutorado. Campinas: Unicamp. 1995.

_____. A sociolingüística paramétrica: perspectivas. In: In: HORA, Dermeval da; CHRISTIANO, Elizabeth. (orgs.) *Estudos Lingüísticos:* realidade brasileira. Idéia: João Pessoa, 1999, p. 107-114.

KATO, Mary. A gramática do letrado: questões para a teoria gramatical. In: M. A. Marques, E. Koller, J. Teixeira & A. S. Lemos (orgs). *ciências da linguagem: trinta anos de investigação e ensino.* braga, cehum (u. do minho), a sair.

KROCH, Anthony. Reflexes of Grammar in Patterns of Language Change. *Language Variations and Change*, 1: 199-244, 1989.

_____. Morphosyntactic Variation. In: K. Beals et al. (eds.), *Papers from the 30th Regional Meeting of the Chicago Linguistic Society* Vol 2: The Parasession on Variation in Linguistic theory. 1994.

_____. Syntactic Change In: Baltin & Collins (eds.). *The handbook of contemporary syntactic theory*. Massachusetts. USA: BlackWell. 2001.

LABOV, William. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press. 1972.

_____. Building on empirical foundations. In: Winfred P. Lehmann & Yakov Malkiel (eds.). *Perspectives on Historical Linguistics*. Amsterdam/Philadelphia: John Benjamins. P. 17-92, 1982.

LABOV, William. Some Sociolinguistic Principles. In: Paulston, C. B.; Tucker, G. R. (org.) *Sociolinguistics: The essential Readings*. Blackwell Publishing, 2003 [1969].

MARTINS, Marco Antonio. Sobre o *se* indeterminador no Português do Brasil: uma proposta morfológica. *Working Papers em Lingüística*. Número 7, pp. 41-58. 2004.

_____. *Entre estrutura, variação e mudança: uma análise sincrônica das construções com -se indeterminador no Português do Brasil.* Florianópolis. UFSC: 2005. Dissertação de mestrado.

MOLLICA, Maria Cecília; BRAGA, Maria Luiza. (orgs.) *Introdução à Sociolingüística:* o tratamento da variação. 2. ed. São Paulo: Contexto, 2004.

NUNES, Jairo M. *O Famigerado se: uma análise sincrônica e diacrônica das construções com se apassivador e indeterminador*. Campinas. UNICAMP: 1990. Dissertação de Mestrado.

WEINREICH, Uriel; LABOV, William; HERZOG, Marvin. Empirical foundations for a theory of language change. In: Winfred P. Lehmann & Yakov Malkiel (eds.). *Directions for Historical Linguistics*. Austin: University of Texas Press. pp. 97 -195. 1968.