

# POR TRÁS DAS FREQUÊNCIAS

VERA LÚCIA PAREDES P. SILVA

## 1 - INTRODUÇÃO

Uma discussão sempre atual é a que põe em pauta a relevância dos números na análise lingüística. Muitos contestam o tempo gasto na coleta e no exame de grandes massas de dados, necessários a uma análise quantitativa, e a importância atribuída aos resultados estatísticos.

Neste artigo, proponho-me a apresentar mais evidências de que a metodologia variacionista, utilizando o cálculo de probabilidades,<sup>1</sup> fornece um instrumental valioso à análise lingüística, conferindo-lhe a precisão desejável. Ao calcular o peso dos condicionamentos de um fenômeno variável segundo o postulado da independência dos fatores, ela permite que se confirmem hipóteses e se escolha com justeza, entre análises alternativas, aquela que melhor dê conta dos fatos. Como evidência, apresento os resultados obtidos em Paredes Silva (1988), onde foi estudada a expressão variável do sujeito da frase na língua escrita informal, representada por um *corpus* de cartas pessoais de cariocas jovens e adultos de ambos os sexos.

## 2 - A QUESTÃO

A possibilidade de explicitar-se ou não o sujeito da frase, característica de algumas línguas românicas como o português e o espanhol, é um caso típico de fenômeno variável e, como tal, já havia sido tratado por pesquisadores das duas línguas (Silva-Corvalán, 1980; Bentivoglio, 1980; Morales, 1980; Naro, 1981; Lira, 1982), sempre na modalidade oral.

Na quase totalidade desses trabalhos, há uma variável que sobressai como altamente condicionadora da presença do sujeito: a *mudança de referência* (MR) do sujeito da frase. Baseia-se na hipótese de que a manutenção do mesmo referente no sujeito da oração favorece a anáfora zero, enquanto a mudança de referência, ao contrário, promove sua explicitação, através de pronome (1ª e 2ª pessoas) ou ainda de nome (3ª pes-

Vera Lúcia Paredes P. Silva. Professora da Universidade Federal do Rio de Janeiro.

<sup>1</sup> Neste artigo é utilizada a expressão "probabilidade" seguindo uma tradição variacionista. A rigor, sabemos tratar-se de pesos relativos. De qualquer forma, a validade do instrumento de análise não é afetada.



soa). É o que se ilustra nos exemplos<sup>2</sup> de 1ª pessoa, abaixo:

- (1) Bom, hoje *acordei cedo e ME DISPUS* a estudar. (LE, p. 2)
- (2) Por favor, não me *interprete* mal quando *EU* te *DIGO* assim. (RC, p. 1)

Essa variável havia apresentado resultados significativos, bastante regulares, em todas as análises a que se aplicara, como se pode verificar na Tabela 1, que reúne, de um modo geral, os efeitos de MR nas análises multifatoriais de Silva-Corvalán, Bentivoglio e Lira, sendo que, nesta última, referentes a todas as pessoas gramaticais:

TABELA 1  
Probabilidade de ausência do sujeito de acordo com a *mudança de referência* no espanhol e no português.

	+MR	-MR
Los Angeles	.34	.66
Caracas	.34	.66
Rio de Janeiro	.33	.67

Diante de tantas evidências a favor da variável *mudança de referência*, era natural aplicar o critério ao *corpus* de cartas pessoais. No entanto, ao fazê-lo, deparei-me com alguns problemas.

De acordo com sua compreensão mais comum, esse critério exige o exame da oração anterior para verificar se seu sujeito tem ou não o mesmo referente da oração em questão. Não leva em conta, porém, a relação dessas duas orações no sentido mais amplo, da organização do discurso. Assim, a classificação torna-se um tanto mecânica. Por exemplo, a oração precedente pode ser a última de um parágrafo ou pode constituir um comentário parentético, o que decerto configura uma situação bastante distinta da identidade de sujeitos numa seqüência narrativa, por exemplo. Entretanto, numa classificação binária, não há como marcar a diferença.

Essas restrições se mantêm quando observamos os resultados numéricos obtidos na análise do sujeito de cartas pessoais. Se, em termos percentuais, o critério confirma nossas primeiras expectativas, com uma diferença de frequência sensível a favor de MR na omissão dos sujeitos, no entanto os números associados às probabilidades são desconcertantes, como se verifica nas Tabelas 2 e 3, que apresentam os resultados de 1ª e 3ª pessoa, respectivamente. Na primeira, números próximos a .50 indicam a indiferença do fator. Na terceira, a questão é mais séria: as expectativas chegam a inverter-se. Em ambos os casos, a variável não foi selecionada pelo programa computacional VARBRUL 2S.

<sup>2</sup> Os dados em questão nos exemplos estão grifados e em caixa alta. As ocorrências prévias consideradas são apenas grifadas.

TABELA 2  
Efeito da *mudança de referência* na ausência de sujeitos de 1ª pessoa (cartas pessoais)

	Apl/T	Freq.	Prob.
-MR	579/652	89%	.53
+MR	692.998	69%	.47

TABELA 3  
Efeito da *mudança de referência* na ausência de sujeitos de 3ª pessoa

	Apl/T	Freq.	Prob.
-MR	235/294	80%	.42
+MR	114/408	28%	.58

Para os variacionistas, é um truísmo afirmar que o cálculo de frequências pura e simplesmente, embora possa sugerir caminhos, pode também preparar-nos armadilhas, uma vez que não leva em conta inter-relações entre os fatores postulados. É meu objetivo apresentar novas evidências empíricas que comprovam a validade do princípio.

O fator específico em questão merece nossa insistência, já que, em análises anteriores do mesmo fenômeno, sua validade parecia irrefutável. Por outro lado, as questões acima levantadas, relativas à interferência da estrutura discursiva, desafiavam a uma abordagem nova. Passemos a ela.

### 3 - A NOVA PROPOSTA

Com o objetivo de absorver as distinções mencionadas, não captáveis numa classificação binária, foi elaborado um critério escalar, que chamei de *conexão do discurso*, compreendendo 6 graus. Ele atribui uma posição nessa escala ao sujeito de cada oração, levando em conta aspectos do contexto discursivo compreendido entre a ocorrência de um referente como sujeito e sua menção anterior - possíveis elementos interferentes, mudança de plano ou mesmo de tópico discursivo, enfim, aspectos capazes de afetar a conexão do discurso.

Meu ponto de partida foi a constatação de que, apesar de se poder manter o referente (particularmente o de 1ª pessoa, mas também o de 3ª) como sujeito de uma sucessão de orações numa carta,<sup>3</sup> havia casos a diferenciar: contextos em que havia maior expectativa de ocorrência de pro-

<sup>3</sup> Na tese a que este artigo se reporta, trabalho com as três pessoas do discurso, tendo a 2ª pessoa também se mostrado sensível ao fator. Porém, como suas ocorrências são pouco numerosas (exceto em cartas classificadas como de *conselho*) ela não será aqui considerada.



nomes do que outros. Minha hipótese era que tal variação devia estar relacionada a alguma propriedade do discurso, não apreensível apenas nas orações, quando consideradas por si. Portanto, era necessário estabelecer os níveis dessas diferenças, delimitando os graus da escala.

Entretanto, a tarefa se tornava mais complexa, dadas as peculiaridades da carta pessoal, que incorpora vários modos de condução discursiva – o narrativo, o descritivo, o persuasivo, o expressivo. Os trabalhos que, na literatura lingüística, já haviam tratado de questões semelhantes sempre foram centrados em narrativas, na 3ª pessoa.

Li & Thompson (1979), por exemplo, estudam a alternância entre pronomes de 3ª pessoa e anáfora zero em narrativas do chinês. Nessa língua, o mais freqüente é que o sujeito da oração não se realize (anáfora zero), sendo a sua ocorrência que deve ser explicada. Nesse sentido, há uma coincidência com a 1ª pessoa na variante do português que analiso: sendo a ausência desses sujeitos da ordem de 77%, são os 23% de sujeitos explícitos que necessitam de explicação.

Através de testes para preenchimento de lacunas, Li & Thompson verificaram a existência de considerável variação entre os informantes na escolha ou não de um pronome. Ainda assim, foi possível captar certas tendências que levaram à formulação de um princípio determinante da ocorrência de pronomes no chinês:

O grau de preferência pela ocorrência de um pronome numa oração corresponde inversamente ao seu grau de conexão com a oração precedente. (Li & Thompson, 1979, p. 330)

A conexão é algo que se define com base não só em propriedades sintáticas e semânticas das orações em causa, mas também leva em conta o conhecimento pragmático e o contexto discursivo em que se inserem as orações. Além disso, afirmam os autores que não é invariável ou absoluta, e, sim, uma questão de *grau*.

Ainda a partir dos resultados empíricos, Li & Thompson estabelecem três restrições à conexão (com conseqüente aumento na incidência de pronomes): a mudança de plano na narrativa (figura-fundo), a presença de expressões temporais ou de morfemas contrastivos e a troca de interlocutor no diálogo.

O ponto interessante dessa concepção de conexão está no fato de que ela permite captar tendências para o uso de pronomes, levando em conta algum tipo de informação que não é sintática no sentido estrito, mas decorre do próprio modo de organização do discurso.

Na tentativa de aplicar critério semelhante ao uso de pronomes em cartas, foram examinados outros trabalhos que lidam com mudança de plano (figura-fundo): o clássico de Labov (1975) sobre a constituição da narrativa, o de Hopper (1979) sobre aspecto e o de Hopper & Thompson (1980) sobre transitividade. Em todos, porém, a questão é mais uma vez

colocada em função de textos narrativos. E cartas, como já vimos, não são textos caracteristicamente narrativos.

A mesma dificuldade distancia-me dos trabalhos reunidos em Givón (1983). O autor ali postula que o domínio funcional da continuidade de tópico é coberto por uma série de mecanismos formais (os pontos de codificação da escala), que indicam a direção continuidade – descontinuidade, a partir da anáfora zero (continuidade máxima) até o SN indefinido (descontinuidade máxima), passando por concordância verbal, pronome pessoal, SN definido, entre outros. Essa hipótese é testada com sucesso na análise de textos narrativos de diversas línguas. O conjunto dos trabalhos acaba por demonstrar que, apesar de as línguas variarem na seleção dos mecanismos formais utilizados, a ordem hierárquica da escala de continuidade é mantida através de todas elas.

Para aferir a continuidade, Givón propõe três medidas discursivas, aplicadas igualmente a todos os textos que compõem esse estudo: distância referencial, interferência e persistência. Assinale-se, mais uma vez, que os textos submetidos a essas medidas são todos narrativos, em 3ª pessoa.

Por esses motivos, já seria impossível aplicar os critérios acima de um *corpus* de cartas pessoais – além de não serem apenas narrativas, são escritas em 1ª pessoa.

No entanto, a idéia, lançada por Li & Thompson (1979), de que há uma relação entre ocorrência de pronomes e grau de conexão pareceu perfeitamente compatível também com meus dados. Assim, foi tomada como ponto de partida para a escala aqui proposta, cujos graus passo a especificar, exemplificados na 1ª e na 3ª pessoa.

#### *Grau 1*

O grau mais alto na escala de conexão diz respeito a seqüências onde não só se mantém o mesmo referente como sujeito da oração, mas também o mesmo tempo e modo verbal, geralmente representando uma seqüência de eventos em torno do mesmo participante. Ou seja, não há mudança de tópico, no sentido discursivo e frasal. Note-se, ainda, que não importa se as orações estão no mesmo período ou não. Esses casos correspondem ao que Li & Thompson (1979) chamam de cadeia tópica. Observem-se os exemplos abaixo:

(03) Bom, (0) estudei alemão e (0) COMECEI a copiar o nosso trabalho (hâm, hâm). (0) TELEFONEI procê, (0) CONTINUEI a copiar o trabalho. (0) FIQUEI de saco cheio, (0) PAREI de copiar e (0) FUI JUNTAR-ME aos meus familiares na praia, onde (0) FIQUEI até as duas horas. (LE, p.2)



(04) *R. ficou*, ( $\emptyset$ ) *SAIU* pra levar A., ( $\emptyset$ ) *DEMOROU* um pouco e ( $\emptyset$ ) *VOLTOU* e ( $\emptyset$ ) *FICOU CONVERSANDO* com a gente. (TH, p.2)

Aí temos um caso de conexão ótima: há uma seqüência temporal de eventos, todos referentes ao mesmo personagem, no que se pode caracterizar como cadeia tópica. A taxa de ausência de sujeitos aqui é bastante alta, em todos os tipos de rodadas.

### Grau 2

O grau seguinte na escala representa uma ligeira queda na conexão, por não se manterem todas as restrições do primeiro grau. daquelas, permanece a exigência de manutenção do tópico (discursivo e frasal), mas pode haver uma mudança no tempo, aspecto e/ou modo do verbo, reflexo da passagem de figura a fundo, do real para o irreal, ou de um fato para um comentário. Ou seja, estou ampliando o conceito usual de mudança de plano (aplicado à narrativa na distinção figura/fundo) para dar conta dos diferentes tipos de transposição que ocorrem nas cartas, sem haver, todavia, alteração do sujeito. Os exemplos (5) e (6) reúnem algumas dessas mudanças:

(5) Sabe, A., ultimamente *eu não tenho tido* muitas novidades pra te contar, porque *EU ESTAVA VIVENDO* em completa alienação, casa-faculdade, faculdade-censo. Como meu primeiro “trabalho” não ( $\emptyset$ ) *GOSTEI* não. Quase ( $\emptyset$ ) *PERCO* a matrícula na faculdade, quase perco o namorado, quase perco os amigos. (CL, p. 3)

(6) Acho que *ele também tinha* uma quedinha por mim, mas não ( $\emptyset$ ) *ousava* sequer sair do sério. *ELE É* um cara muito legal, não é? (MC, p. 2)

Em (5), embora o referente do sujeito se mantenha o mesmo ao longo de toda a seqüência, não temos o mesmo grau de conexão do item precedente. Não há uma narrativa, antes, trata-se de um pedido indireto de desculpas. Ao tempo da oração inicial se superpõe parcialmente, com mudança de aspecto, o da oração que expressa a justificativa, sendo essas propriedades (superposição temporal, descrição de estado necessária à compreensão de atitudes) algumas das apontadas por Hopper (1979, p. 216) para caracterizar o que está em segundo plano (fundo). Ainda centrado na 1ª pessoa, o discurso segue com um comentário avaliativo, representando outra mudança de plano – agora a emissora diz como se sentiu. Finalmente as orações que repetem o verbo *perder* (entre as quais a conexão é ótima) realizam também uma mudança de plano com relação à do verbo *gostar*; trata-se, mais uma vez, de apresentar motivos, justificativas.

Em (6), a seqüência escolhida exemplifica a mudança de plano – depois do relato de eventos passados, envolvendo o mesmo personagem como sujeito, é-lhe atribuída uma qualidade permanente, usando-se, por isso, o presente atemporal.

Neste grau da escala já se observa uma queda sensível da taxa de omissão do sujeito, e note-se que, numa medida de continuidade como a de Givón (1983), ou num tratamento através apenas de mudança de referência, a distinção entre os graus 1 e 2 não seria apreendida. Entretanto, confira-se nas Tabelas 4 e 5, que ela é bastante significativa.

### Grau 3

Se nos dois graus anteriores havia a exigência de mesmo referente no sujeito, aqui este elo se perde, enfraquecendo-se um pouco mais a conexão. No entanto, entre o sujeito em questão e sua menção prévia, ocorrem orações de curta extensão, geralmente impessoais, que não chegam a representar um corte, uma interrupção na seqüência do discurso, cujo tema permanece o mesmo. Isto é, o sujeito não é o mesmo da oração anterior. Conecta-se com uma menção mais distante, porém não é introduzido nesse intervalo nenhum participante distinto e pouca informação é acrescentada. Vejam-se os exemplos abaixo: no primeiro, a expressão interferente é uma oração impessoal, indicativa de tempo; no segundo, o tópico (discursivo e frasal) é estabelecido no início do parágrafo, através da expressão *quanto ao*, seguida de nome próprio. Não entra em cena nenhum concorrente à mesma função, embora haja, na seqüência, uma breve interferência de orações com outro sujeito.

(7) Hoje ( $\emptyset$ ) *vou dormir* cedo pois ( $\emptyset$ ) *fiquei estudando* ontem até as 3:30h. AM. Ainda são 11hs. ( $\emptyset$ ) *VOU FICANDO* por aqui com meu coração cheio de “amor platônico” e “carinho armandônico”. (AR, p. 4)

(8) Quanto ao D., *ele apareceu* aqui no início de dezembro e estava na base da amizade, tudo indo bem, quando *ELE ENCRENCOU* e ( $\emptyset$ ) *embraveceu* novamente. (ZI, p. 2)

### Grau 4

Nos graus precedentes, a centralidade do referente como tópico/sujeito não foi questionada: mesmo no grau 3, onde havia elementos interferentes, eles não chegavam a criar uma interrupção. No grau 4, entretanto, o referente teve sua última menção em outra função sintática. Para a 1ª pessoa, isso significa ocorrência prévia como pronome oblíquo ou possessivo, do ponto de vista morfológico, e passagem de um papel secundá-



rio para o central, em termos discursivos. Como a 1ª pessoa predomina no papel central, nela as ocorrências deste grau são menos frequentes. O exemplo abaixo ilustra este grau de conexão.

(9) Desculpe-me pela *minha* reclamação infundada. *EU ESTAVA BRINCANDO.*(MA, p. 1)

Trata-se do período com que se inicia uma carta. Embora a 1ª pessoa já estivesse materialmente introduzida no texto, ainda não havia ocupado a posição de sujeito, ocorrendo, então, em sua forma explícita.

Por outro lado, é o grau que concentra o maior número de ocorrências de 3ª pessoa. Isso se explica porque entidades novas ou disponíveis no discurso são geralmente introduzidas em outra função que não a de sujeito. Quando elas permanecem e se tornam tópico da frase, podem ser substituídas lexicalmente, pronominalmente ou pela anáfora zero, como se vê a seguir:

(10) E a sua vinda para o B., ou melhor *para Washington?* Olha, vale a pena escavar. *A CIDADE É* ótima, gostosíssima de se viver, cheia de áreas verdes, e muita coisa para ver. (AR, p.4)

(11) Segundo, estou dando aula duas vezes por semana *a uma senhora.* *ELA VEM* aqui em casa. (AL, p. 2)

(12) Achei *um cara incrível*, mas (0)É noivo, daí tô partindo para um que é mais ou menos legal. (HE, p. 1)

#### Grau 5

Considero que, neste grau, a conexão é mais afetada por ter entrado em cena outro participante na função de sujeito, entre o sujeito em questão e sua última menção no texto, representando um interferente em potencial. A volta ao sujeito (de 1ª ou 3ª pessoa) é tida como uma retomada. O elemento interferente pode ser qualquer outro ser animado que apareça como concorrente ao papel de sujeito. De início, a interferência de cada pessoa foi computada em separado, mas os resultados foram tais que permitiram reuni-los, atribuindo-lhes a mesma pontuação na escala. No exemplo de 1ª pessoa que se segue, o emissor está comentando uma conversa telefônica que ele e a destinatária tiveram. Na parte interferente, além da referência à destinatária, através do SN *uma amiga tão especial*, surge um sujeito de terceira pessoa, expresso por outro SN pleno, *a menina da Embratel*. Assim, a volta do discurso ao âmbito da 1ª pessoa funciona como uma retomada. Observe-se, ainda, que não há mudança no tópico discursivo.

(13) Embora nem sempre (0) *consiga*, sempre (0) *tento alegrar* pessoas amigas minhas e principalmente *uma amiga tão especial*. Em dezembro *a menina da Embratel* (que sai de férias agora) voltará a seu cargo e então (0) *TENTAREI FAZER* novas chamadas para você. (AR, p. 3)

Já no exemplo (14), extraído de uma carta de conselho, a 3ª pessoa em questão deixa de ser tópico da frase para que se introduza um comentário generalizante da emissora, e volta depois de um imperativo (portanto, da interferência de um sujeito de 2ª pessoa), numa seqüência em que o tópico discursivo se manteve inalterado.

(14) Não entendi porque *ele ficou* tão zangado com sua resposta. Afinal quem ama realmente sempre tem uma esperança por mais remota que esta seja. Procure saber se *ELE ESTÁ* com algum problema e quer, sei lá, fugir dele procurando uma coisa nova, diferente na vida dele, ou seja, o casamento. (MC, p. 3)

As ocorrências incluídas no grau 5 da *escala de conexão do discurso* representam uma queda sensível com relação ao nível anterior (cf. Tabelas 4 e 5).

#### Grau 6

O último grau da escala traz os casos de conexão mais fraca: quando há mudança de tópico discursivo, do *assunto* de que se trata. Aqui é necessário fazer uma distinção entre as alternativas de 1ª e 3ª pessoa. Na 1ª pessoa, ainda que se mantenha o mesmo referente no papel de sujeito, a probabilidade de ocorrência de pronome aumenta em função da quebra na seqüência discursiva, que, neste caso, vai além de uma simples mudança de plano, ou da interferência de outros personagens: é o próprio assunto que muda, no todo ou em parte, como se verifica no exemplo abaixo:

(15) Preciso, além de ler, lavar, passar, cozinhar... eta vidinha de dona de casa. Campinas me ensinou a cozinhar, coisas que nunca *pensei* que *iria* fazer antes, às vezes *preciso* fazer...

Talvez *EU MONTE* uma república no próximo semestre, o aluguel aqui está caríssimo, o condomínio nem se fala. (IU, p. 1)

Observe-se que o referente do sujeito é mantido, o discurso continua centrado na 1ª pessoa, mas a mudança de tópico discursivo – dos afazeres domésticos para problemas de ordem econômica – favorece a presença do pronome sujeito.



Examinemos agora o exemplo da 3ª pessoa:

(16) Ah! O café que eu trouxe já acabou mas o *B. gosta* muito, se vocês puderem, mandem mais 1/2 quilo.

Sabem, ontem o *B. me DEU* um *poster* que ele fez no computador do hospital onde ele trabalha. (AD, p.3)

Na 3ª pessoa, o grau 6 da escala proposta caracterizou-se, em princípio, por dar preferência à explicitação do sujeito através do SN pleno. Por esse motivo, só foram identificadas 26 ocorrências consideradas passíveis de expressão variável. O natural, quando há quebras da conexão por mudança de assunto, é que surja outro referente de 3ª pessoa na posição de sujeito, ainda que não seja totalmente novo no discurso. Nessas circunstâncias não há, evidentemente, variação.

Sob esse aspecto, o exemplo acima é uma ocorrência rara: muda o tópico discursivo (do *café* para *o presente de B.*) mas o terceiro envolvido (a 3ª pessoa) permanece o mesmo. Apesar disso, o sujeito do novo parágrafo aparece na forma de um SN pleno. Parece, entretanto, que pelo menos a ocorrência do pronome *ele* seria admissível nesse contexto, embora essa seja uma avaliação puramente intuitiva. De qualquer forma, a exclusão dos 26 casos acima mencionados não afeta a argumentação, nem interfere na hierarquia da escala na 3ª pessoa. Simplesmente comprova que a conexão fraca no último grau pode obrigar ao preenchimento categórico do sujeito.

As Tabelas 4 e 5, abaixo, reúnem os resultados da escala proposta, na 1ª e na 3ª pessoa:

TABELA 4

Efeito da *conexão do discurso* na ausência de sujeitos de 1ª pessoa

	Apl/T	Freq.	Prob.
Grau 1	209/212	99%	.94
Grau 2	336/395	85%	.59
Grau 3	143/178	80%	.47
Grau 4	055/078	70%	.34
Grau 5	262/410	64%	.25
Grau 6	266/377	70%	.23

TABELA 5

Efeito da *conexão do discurso* na ausência de sujeitos de 3ª pessoa.

	Apl/T	Freq.	Prob.
Grau 1	165/176	94%	.93
Grau 2	079/120	66%	.70
Grau 3	021/055	38%	.53
Grau 4	077/243	32%	.37
Grau 5	006/082	7%	.21
Grau 6	001/026	4%	.15

#### 4 - INTERPRETAÇÃO DOS RESULTADOS

Comparemos agora as duas tabelas com as de *mudança de referência* (2 e 3) para melhor entender por que esta última não é selecionada pelo programa computacional.

O VARBRUL 2S, operando com base no princípio da independência dos fatores, nos dá o peso combinado desses fatores, que, em última instância, definem a tendência ao uso de uma ou outra variante. Com esse objetivo, ele nos apresenta uma seleção dos fatores estatisticamente significativos para a aplicação da regra, na ordem de importância. Essa seleção é feita em cada nível da rodada, com base no teste de significância - um cálculo estatístico que mede a adequação dos resultados projetados para os dados observados. Quanto mais próximo de zero esse resultado, tanto melhor.

O programa atua iterativamente, comparando os diversos grupos de fatores entre si. Cada rodada compreende tantos níveis quantos forem necessários para realizar-se a seleção. No nível 0, calcula-se a probabilidade de *input*, que corresponde a uma probabilidade média de aplicação da regra, na ausência de restrições. No nível 1, ele toma cada grupo de fatores isoladamente, comparando-o com o *input* e efetua o cálculo. O grupo que melhor se sair no teste de significância será selecionado, e com ele serão comparados, no nível 2, todos os demais grupos, um por um, para que seja selecionado o segundo grupo, e assim sucessivamente até completar-se a seleção.

Ora, no caso em pauta, o critério de *conexão do discurso*, que compete com MR, é mais adequado aos fatos, na medida em que permite apreender distinções que de outro modo não aparecem (mudança de plano, papel dos elementos interferentes, sejam impessoais ou de outras pessoas gramaticais, importância da mudança de tópico discursivo, etc). Isso se reflete no processo de seleção, quando é o primeiro escolhido. A nova classificação cria, assim, uma escala cujos extremos são bastante po-



larizados (de .94 a .23 na 1ª pessoa e de .93 a .15 na 3ª.) e que estabelece diferenças mais finas, não captáveis num critério de *mudança de referência*.

Do ponto de vista estatístico, pode-se verificar como a variável *mudança de referência* é suplantada pela *conexão do discurso*, acompanhando-se seu comportamento através dos níveis de uma rodada do VARBRUL 2S. No nível 1, as duas variáveis aparecem com resultados polarizados e significância igual a 0.0, tanto na 1ª como na 3ª pessoa, mas é *conexão do discurso* a imediatamente selecionada. Assim, no nível 2, quando o efeito de *mudança de referência* é testado em combinação com *conexão do discurso*, as probabilidades associadas àquele fator na 1ª pessoa já se alteram sensivelmente (cf. Tabela 6). Nos níveis seguintes elas vão sofrendo pequenas quedas, até que, no último nível da seleção, aparecem os níveis já vistos na Tabela 2. Seguem-se as tabelas que mostram esse declínio:

TABELA 6  
Queda da variável *mudança de referência* na 1ª pessoa

	Nível 1	Nível 2	Nível 8
-MR	.65	.55	.53
+MR	.35	.45	.47
Sig	0.0	0.381	0.898

TABELA 7  
Queda da variável *mudança de referência* na 3ª pessoa

	Nível 1	Nível 2	Nível 5
-MR	.76	.47	.42
+MR	.24	.53	.58
Sig	0.0	0.553	0.260

Na 3ª pessoa, o fenômeno é ainda mais interessante. Uma vez selecionada a *conexão do discurso* no nível 1 da rodada, os resultados do nível 2 já mostram não só a proximidade de .50 (neutralidade do fator) mas uma inversão de expectativas: a probabilidade mais alta se associa à percentagem mais baixa. Ao final do processo de seleção, no nível 5, chegamos a encontrar resultados razoavelmente polarizados (.42 a .58), mas uma significância de 0.260, o que impede sua seleção.

Ainda no intuito de confirmar a precedência da variável *conexão do discurso* sobre a variável *mudança de referência*, foi feita uma rodada do VARBRUL 2S com os dados de 1ª pessoa, excluindo a *conexão*. Nesse caso, o fator *mudança de referência* é selecionado, com probabilidade .67 para a ausência de sujeito, quando o referente é mantido, contra .33, para quando é mudado, resultados muito próximos aos obtidos por outros investigadores, como se confere na Tabela 1. Ainda assim, não ocupa na seleção do programa o primeiro lugar (como sempre ocorre com *conexão do discurso*, se presente).

Também testei o cruzamento dos dois fatores, *conexão do discurso* e *mudança de referência* na 1ª pessoa, tendo sido obtidos os resultados que estão na tabela abaixo:

TABELA 8  
Efeito conjugado de *mudança de referência* e *conexão do discurso* na 1ª pessoa.

	Grau 1	Grau 2	Grau 3	Grau 4	Grau 5	Grau 6
-MR	.95	.66			.30	.32
+MR			.55	.40	.31	.28

Por esses resultados, vê-se que os pontos mais baixos estão, de fato, nos graus 5 e 6 da *escala de conexão do discurso*, independente de o referente ser o mesmo ou ter mudado, o que mais uma vez confirma a melhor posição desse fator, ao mesmo tempo demonstrando que ele não representa uma simples subdivisão dos casos de *mudança de referência*. Esclareça-se, ainda assim, que, nessa rodada conjunta, o fator combinado teve também o primeiro lugar na seleção do programa.

## 5 - CONCLUSÃO

Em síntese, vimos que, embora a variável *mudança de referência* se mostre significativa nas análises em que é aplicada, uma variável que analise mais detalhadamente o comportamento da seqüência de sujeitos, levando em conta a estruturação do discurso, pode apreender diferenças mais sutis, revelando melhor os aspectos do uso de pronomes. Além disso, vimos que os números resultantes de um simples cálculo de frequências não conseguem indicar qual a melhor alternativa. Somente a avaliação do peso combinado dos fatores e sua seleção pelo programa computacional podem dar respaldo a nossa escolha da variável *conexão do discurso* como um condicionamento mais adequado ao fenômeno estudado.



## REFERÊNCIAS BIBLIOGRÁFICAS

- BENTIVOGLIO, Paola A. *Why 'canto' and not 'yo canto'?* The problem of first person subject pronoun in spoken Venezuelan Spanish. M. of Arts thesis – University of California, 1980. 72 p. mimeo.
- GIVÓN, Talmy (Ed). *Topic continuity in discourse*; quantitative cross-language studies. Amsterdam: John Benjamins, 1983.
- HOPPER, Paul J. Aspect and foregrounding in discourse. In: GIVÓN, T. (Ed.) *Syntax and semantics*. New York: Academic Press, 1979. v. 12: Discourse and syntax. p. 251-99.
- & THOMPSON, Sandra A. Transitivity in grammar and discourse. *Language LSA*, v. 56, n. 2, p.251-299, 1980.
- LABOV, William. *Sociolinguistic patterns*. 3. ed. Philadelphia: University of Pennsylvania Press, 1975. 343 p.
- LI, Charles N. & THOMPSON, Sandra A. Third-person pronouns and zero-anaphora in Chinese discourse. In: GIVÓN, T. (Ed.). *Syntax and semantics*. New York: Academic Press, 1979. v. 12: Discourse and syntax. p.311-35.
- LIRA, Solange de Azambuja. *Nominal, pronominal and zero subject in Brazilian Portuguese*. University of Pennsylvania Ph.D. Dissertation. University Microfilms International, 1982. 227 p.
- MORALES, Amparo. La expresión de sujeto pronominal, primera persona, en el español de Puerto Rico. *Boletín de la Academia Puertorriqueña de la Lengua Española*, San Juan: Academia Puertorriqueña de la Lengua Española, v. 8, n. 2, p. 91-102, 1980.
- NARO, Anthony J. Morphological constraints on subject deletion. In: SANKOFF, David & CEDERGREEN, Henrietta (Eds.). *Variation Omnibus*. Canada: Linguistic Research, 1981. p. 351-8.
- PAREDES SILVA, Vera Lúcia. *Cartas cariocas: a variação do sujeito na escrita informal*. Tese (Doutorado em Letras) – UFRJ, 1988. 330 p.
- SILVA-CORVALÁN, O. Subject expression and placement in Mexican-American Spanish. In AMASTAE, Y. & ELIAS OLIVARES, L. (Eds.). *Spanish in the United States: Sociolinguistic aspects*. Cambridge: Cambridge University Press, 1982. p. 93-120.