

**VARIAÇÃO E MUDANÇA LINGÜÍSTICA:
FLUXOS E CONTRAFLUXOS NA COMUNIDADE DE FALA**

**ANTHONY J. NARO
MARIA MARTA PEREIRA SCHERRE
Universidade Federal do Rio de Janeiro**

É comum os lingüistas considerarem que a mudança diacrônica, e conseqüentemente a variação sincrônica, caminha em uma dada direção ao longo do eixo do tempo na comunidade de fala. Então, uma forma pode ser vista como "entrando na língua" ou "caminhando para a extinção". O quadro geral é o de pontos finais categóricos mediados por estágios intermediários sucessivos de variação que impulsionam os processos em uma direção única através da comunidade.

Vamos demonstrar a seguir que em alguns casos este quadro não ocorre porque a comunidade de fala pode estar caminhando em diversas direções, no sentido que alguns grupos de falantes podem estar num processo de aquisição da forma enquanto outros estão, ao mesmo tempo, perdendo a forma. Alguns grupos podem estar ainda estáveis, mostrando padrões típicos de variação estável, de tal forma que para eles não há, portanto, processo de mudança em curso. Todavia, a diversidade de direções não é refletida por qualquer tipo de separação social entre os grupos, que estão em interação face-a-face e continuam a participar da variação lingüística corrente. De fato, um dos problemas mais difíceis é precisamente identificar as dimensões sociais que definem os grupos independentemente dos fatos lingüísticos puros.

A primeira análise deste tipo para a comunidade de fala carioca foi a de Naro (1981) através da variável "orientação social". Ele sugeriu que no caso da regra de concordância sujeito/verbo há dois grupos distintos numa amostra de 20 falantes semi-analfabetos estudados. Todos estes falantes freqüentavam, na época, o MOBREAL - curso de alfabetização de adultos do governo brasileiro. Tendo em vista a variável em questão, foi possível verificar que um dos grupos evidenciava uma mudança lenta em direção a um sistema não marcado enquanto o outro resistia à mudança,

restaurando este mesmo sistema. A base desta análise foi a mera observação empírica de que alguns falantes apresentavam taxas de concordância acima da média da comunidade na qual eles viviam ou haviam crescido. Embora eles se assemelhassem aos outros de sua comunidade com relação a todos os aspectos materiais, eles evidenciavam uma visão cultural ampla e eram capazes de interagir bem com os lingüistas de classe média que os entrevistaram. Variáveis tradicionais tais como sexo, faixa etária e origem rural ou urbana se mostraram irrelevantes na delimitação dos dois grupos. Todavia, quando eles foram separados um apresentava uma taxa global de concordância de aproximadamente 65%, enquanto o outro se aproximava de 40%. O critério social usado para definir os grupos independentemente de suas taxas de concordância foi a sua reação às novelas de televisão. É de domínio público que estes programas são predominantemente direcionados à classe média e são exclusivamente baseados nos valores da classe média, sendo completamente irrelevantes para a cultura e o estilo de vida da classe baixa. Isto é tão claro que a maioria dos falantes com taxas mais altas de concordância eram os que assistiam às novelas e demonstravam entender realmente as cenas, enquanto os do outro grupo declaravam que não viam novelas e que não eram capazes de perceber qualquer coerência entre as cenas. É importante enfatizar que o hábito do falante de ver novelas foi usado apenas como um critério independente do comportamento lingüístico para a constituição dos dois grupos. Não foi estabelecida nenhuma relação causativa entre o tipo de linguagem usada nas novelas e a linguagem usada pelos falantes que assistem a este tipo de programa uma vez que ambos os grupos são igualmente expostos à linguagem padrão dos programas de rádio, dos filmes e de vários outros programas de televisão.

Na pesquisa original, feita em 1981, os padrões lingüísticos dos dois grupos sociais descritos acima não foram analisados separadamente. Era então impossível determinar em que aspectos eles poderiam diferir entre si. Além disso, como somente dois grupos etários foram considerados, a gradação etária não pôde ser vista de forma mais detalhada. Todavia, quando todas as variáveis eram consideradas conjuntamente, através do uso de métodos computacionais tais como VARBRUL (SANKOFF, 1975; NARO & VOTRE, 1980) ou ANOVA (NIE et alii, 1975), foi possível perceber uma forte gradação etária, com falantes mais velhos usando formas com concordância mais freqüentemente do que falantes mais jovens. Agora, quando analisamos os grupos sociais separadamente através do programa de regra variável os resultados das restrições estruturais se mostram muito semelhantes para os dois grupos, mas o mesmo não ocorre com as restrições sociais convencionais. Com o objetivo de ver com mais clareza a

curva do tempo aparente, estabelecemos três grupos etários: (1) menos de 19 anos, (2) 19-39 anos e (3) acima de 40 anos. Este foi o único fator, ao lado da variável orientação cultural, que foi estatisticamente significativo para a amostra como um todo. O padrão etário geral é típico de uma forma que está desaparecendo da língua, com pessoas jovens usando menos concordância do que pessoas de meia idade, que, por sua vez, usam menos concordância do que pessoas mais velhas, como pode ser visto na Tabela 1 a seguir.

Variável social significativa: faixa etária

	Frequência	Peso relativo
< 19 anos	840/1685=50%	0,39
19/39 anos	1084/2403=45%	0,48
>39 anos	1078/2221=49%	0,63

**Tabela 1 - Amostra MOBRAL completa
Concordância verbo/sujeito**

As razões para a falta de concordância entre as frequências e os pesos relativos (probabilidades) são as mesmas que as explicadas em detalhes em Naro (1981).

Analisando separadamente os dados das pessoas que assistem a novelas, obtivemos os seguintes resultados para a faixa etária:

Variável social significativa: faixa etária

	Frequência	Peso relativo
< 19 anos	613/1131=56%	0,36
19/39 anos	421/523=81%	0,68
> 39 anos	96/141=68%	0,46

**Tabela 2:- Amostra MOBRAL: Telespectadores de novelas
Concordância verbo/sujeito**

É importante lembrar que era este o grupo que originalmente parecia estar restaurando a regra de concordância. Diferentemente, os novos resultados da Tabela 2 sugerem que este grupo apresenta um padrão típico de variação estável, com um aumento no uso da concordância após a adolescência e com um decréscimo na faixa etária mais velha.

O restante da amostra MOBRAL, mostra um padrão etário, pro-

vavelmente típico da maioria dos analfabetos, que pode ser visto através dos resultados da Tabela 3 a seguir.

Variável social significativa: faixa etária

	Frequência	Peso relativo
< 19 anos	209/5544=38%	0,49
19/39 anos	663/1880=35%	0,39
> 39 anos	982/2080=47%	0,62

Tabela 3 - Amostra MOBREAL: Não telespectadores de novelas
Concordância verbo/sujeito

Aqui encontramos novamente o padrão dos falantes mais velhos usando mais a regra do que os falantes mais jovens, embora o padrão não seja tão claro por causa de um leve abaixamento no peso relativo associado ao grupo do meio.

De forma geral, a análise da mostra MOBREAL sugere que um grupo de falantes mostra um padrão de variação estável enquanto o outro evidencia um padrão típico de eliminação da forma. Isto confirma a idéia que os grupos se movem em diferentes direções na amostra de analfabetos de falantes cariocas.

No caso da concordância entre os elementos do sintagma nominal, Scherre (1988) encontrou uma ampla gama de taxas de concordância, exatamente como Naro (1981) para a concordância sujeito/verbo. Mesmo entre os falantes com somente a educação primária, por exemplo, a taxa de marcas (excluindo os determinantes) varia de 7% a 90%. Todavia, a amostra com a qual Scherre estava trabalhando era mais diversa do que a de Naro, consistindo de 48 falantes estratificados em função do sexo, faixa etária e anos de escolarização (analfabetos e universitários não fazem parte da amostra). Os resultados para variáveis sociais do tipo faixa etária se mostraram instáveis no sentido que o agrupamento de falantes levemente diferentes ou a inclusão de variáveis diferentes na análise provocou resultados significativamente diferentes. Naturalmente, isto leva à conclusão de que deve haver grupos diferentes se movendo em direções opostas dentro da amostra. Todavia, o padrão dominante para a faixa etária, considerando toda a amostra, se mostrava curvilíneo, típico da variação estável, no qual os mais jovens e os mais velhos mostravam frequências de concordância mais baixas enquanto o grupo da faixa intermediária mostrava níveis de concordância mais altos.

Com o objetivo de inferir fluxos e contrafluxos na concordância

entre os elementos do sintagma nominal, a amostra de 48 falantes foi subdividida conforme as mesmas observações empíricas que motivaram o reagrupamento de falantes no caso da concordância sujeito/verbo: alguns falantes tinham médias de concordância divergentes da média de seu grupo social ou do ambiente no qual eles cresceram. Considerando que esta amostra é socialmente variada, englobando profissionais de classe média e trabalhadores pobres, foi primeiro necessário dividir os falantes em função de sua origem social. Três categorias *ad hoc* foram usadas: níveis socioeconômicos baixo, alto e intermediário, dentro dos limites da amostra. Os critérios usados foram estritamente informais, embora baseados nas observações de uma das entrevistadoras e de uma socióloga que leu as entrevistas. A idéia central era agrupar aqueles falantes de origens "simples" (através de índices do tipo vizinhos de classe baixa, falta de utensílios modernos, moradia inadequada, alimentação fora dos padrões, dificuldade de acesso a transporte, etc.) em oposição aos falantes que viveram sob condições típicas de classe média.

Dentro de cada nível socioeconômico, as taxas de concordância foram ordenadas da mais alta para a mais baixa e um ponto divisório natural foi encontrado. Como resultado, obtivemos grupos de frequência de concordância alta e baixa em cada nível. O número de falantes do nível mais alto com taxas de concordância baixa foi muito pequeno e bastante enviezado com relação a outras variáveis sociais, não permitindo um estudo quantitativo adequado. Os grupos intermediários foram eliminados porque eles incluíam falantes que não foram classificados de forma clara por causa de falta de informação.

Denominamos "crioulizado" o grupo com a taxa de concordância mais baixa no nível socioeconômico mais baixo por causa da linha tradicional de raciocínio na filologia brasileira (SILVA NETO, 1976; GUY, 1981) que defende que muitas mudanças lingüísticas no português brasileiro são devidas a um estágio de crioulização da língua entre africanos e descendentes africanos durante o período da escravidão, estágio este não atestado até o presente momento. Seguramente, se uma real crioulização existiu no Brasil, este é o grupo que pode apresentar índices de sua sobrevivência. Por razões semelhantes, nós usamos o rótulo "descruiolização" para aqueles falantes de nível socioeconômico baixo que apresentavam uma taxa alta de concordância. Os falantes do grupo socioeconômico mais alto não apresentam status claro de crioulização na literatura pertinente. Embora certamente não descendentes de escravos africanos, alguns estudiosos afirmam que a língua dos grupos de nível socioeconômico baixo, ou que a de seus ancestrais, era significativamente influenciada pela crioulização. De qualquer forma, nós usamos estas três categorias com ró-

tulos tradicionais sem qualquer aceitação de seu conteúdo.

Os dados da concordância entre os elementos do sintagma nominal para os grupos descritos acima foram analisados, através do programa de regra variável, com todas as variáveis estruturais relevantes bem como com as variáveis sociais convencionais tais como faixa etária, anos de escolarização e sexo. O programa utilizado apresentou a seleção das variáveis que eram estatisticamente relevantes no nível de 0,05 ou menos, e rejeitou outras.

Para o nível socioeconômico mais alto, a única variável social selecionada pelo programa como significativa foi a faixa etária. Os seus resultados, apresentados na tabela 4, mostram um padrão típico bem definido de variação estável.

Variável social significativa: faixa etária

	Frequência	Peso relativo
15/25 anos	681/804=85%	0,44
26/49 anos	934/1010=92%	0,64
50/71 anos	789/924=85%	0,41

Tabela 4 - Nível socioeconômico alto
Concordância no sintagma nominal

O grupo da faixa etária intermediária mostra um aumento significativo na frequência da concordância, provavelmente resultado de pressões sociais do mercado de trabalho sofridas por este grupo na sua vida profissional.

No nível socioeconômico mais baixo, a faixa etária foi também a única variável social escolhida como estatisticamente relevante para o grupo com frequência de concordância mais baixa. Relembramos que este grupo é o que está mais próximo do que temos hoje para apresentar evidências da língua supostamente crioula no passado.

Variável social significativa: faixa etária

	Frequência	Peso relativo
15/25 anos	276/581=47%	0,38
26/49 anos	295/546=54%	0,59
50/71 anos	427/812=53%	0,53

Tabela 5 - Nível socioeconômico baixo
Taxas mais baixas de marcas (crioulizado)
Concordância no sintagma nominal

Aqui podemos ver um decréscimo na freqüência para os falantes mais jovens, enquanto os outros dois grupos mostram uma taxa mais alta. Embora a gradação etária não seja forte, não mostrando diferenças significativas entre os dois grupos mais velhos, podemos considerar estes resultados como indicativos do padrão de perda da concordância ao longo do tempo.

Para os falantes do mesmo nível socioeconômico com as taxas de concordância mais altas, o grupo presumivelmente descrioulizado, somente o nível educacional foi escolhido como significante.

Variável social significativa: anos de escolarização

	Freqüência	Peso relativo
1/4 anos	398/529=75%	0,33
5/8 anos	460/569=81%	0,47
9/11 anos	351/381=92%	0,70

Tabela 6 - Nível socioeconômico baixo
Taxas mais baixas de marcas (descrioulizado)
Concordância no sintagma nominal

Estes resultados sugerem exatamente por onde a descrioulização pode vir. Surpreendentemente, o sistema educacional parece funcionar, mas é melhor ter em mente que nossos resultados mostram tão somente uma correlação e não necessariamente uma causa.

Acreditamos que os resultados apresentados sugerem fortemente que é fácil encontrar fluxos e contrafluxos envolvendo a variação e mudança na comunidade de fala carioca. O que está mudando para algumas pessoas pode estar estável para outras pessoas e o que está aumentando para alguns pode estar diminuindo para outros. Para algumas pessoas o mercado de trabalho pode ter efeito, enquanto o sistema escolar pode influenciar outras. Finalmente, podemos reafirmar que os fluxos e contrafluxos só parecem emergir com agrupamentos de indivíduos de forma não convencional. No futuro, esperamos ser capazes de avaliar este ponto em mais detalhes.

REFERÊNCIAS BIBLIOGRÁFICAS

- GUY, Gregory R. **Linguistic variation in Brazilian Portuguese: aspects of the phonology, syntax, and language history**. PH.D. Dissertation, University of Pennsylvania, 1981. 391p. mimeo.
- NARO, Antony J. The social and structural dimensions of a syntactic change. **Language**. LSA, 57(1):63-98, 1981.
- NARO, Antony J. & VOTRE; Sebastião J. **SWAVA: Sistema SWAMINC/VARBRUL (Manual do Usuário)**. Rio de Janeiro, UFRJ, 1980. mimeo.
- NIE, Norman H. et alii. **SPSS: Statistical package for the social sciences**. 2^a ed. New York, McGraw-Hill.
- SANKOFF. David. **VARBRUL Version 2**. Centre de Recherches Mathématiques, Université de Montréal, 1975. mimeo.
- SCHERRE, Maria Marta Pereira. **Reanálise da concordância nominal em português**. Tese de Doutorado. Rio de Janeiro, UFRJ, 1988. Em dois volumes, com 554p. mimeo.
- SILVA NETO, Serafim da. **Introdução ao estudo da língua portuguesa no Brasil**. 3.ed. Rio de Janeiro, Presença, 1976.