

O SUJEITO PRONOMINAL NO PORTUGUÊS FALADO E ESCRITO

Solange de Azambuja Lira
UFSC

Lira (1982) estudou um aspecto morfo-sintático do português — a variação entre o sujeito pronominal e o sujeito zero. Este é um tópico de interesse nas línguas românicas, pouco estudado até hoje, Podemos exemplificar esta variação com as seguintes sentenças tiradas de gravações com falantes nativos do português:

(1) Nós curtimo o namoro, mas legal mesmo! \emptyset Ia pro Parque Lage, \emptyset escrevia naquelas pedras nosso nome. Olha: \emptyset Ficou lá anos. Eu panhava lata de tinta do meu pai, \emptyset levava dentro da bolsa e pincel, aí \emptyset pintava nosso nome lá. (MC 12 A/200)

Cunha (1972:204) afirma que os pronomes sujeitos são normalmente omitidos no português pois as desinências verbais bastam para indicar a pessoa a que se refere o predicado como também o número gramatical. Ele argumenta que a referência do sujeito é clara na flexão sujeito-verbo e que o uso do pronome é devido às seguintes razões: ênfase; para opor duas pessoas diferentes; para diferenciar duas formas morfologicamente ambíguas. Quicoli (1976) em um enfoque gerativo-transformacional discute a regra que ele chama de apagamento do sujeito pronominal em português. Ele argumenta que esta regra apaga somente pronomes sujeitos que são redundantes, não-enfáticos, não-contrastivos. Entretanto a análise da fala de 30 falantes - 15 minutos de cada falante - mostrou a seguinte situação:

TABELA 1 - Frequência dos sujeitos pronominais e zero.

	N	%
Pronominal	5024	56
Zero	3900	44
Total	8924	100

A análise mostrou que a variação sujeito pronominal e zero é grande e a análise multivariada usando o programa de regra variável de David Sankoff (1975), demonstrou que os sujeitos pronominais têm maior probabi

lidade de serem empregados no português falado para:

1) Esclarecer o referente de uma situação potencialmente ambígua como por exemplo:

Quando o referente do sujeito for um de 2ª pessoa. Sem a presença do pronome a referência fica ambígua pois no português do Rio de Janeiro, o verbo é flexionado para a 3ª pessoa.

(2) Você é psicóloga, né? Você sabe que uma das coisas que gosto de falar é Ana. (C 41C?168)

Quando o referente do sujeito não é o mesmo do que o da oração anterior:

(3) Quando eu voltei da Europa, ela foi embora para a casa do pai. (Os 48 C/165)

Quando a oração for relativa e o pronome relativo não for o sujeito da oração:

(4) Um rapaz que ela estava namorando... (IL 21A/246)

A análise também mostrou que o sujeito pronominal é mais tópico do que o sujeito zero. Ele quase ocorre com sujeitos não-humanos. Um referente humano é mais tópico do que um não-humano. Um sujeito pronominal tende a ocorrer mais frequentemente quando o seu referente acabou de ser mencionado com em:

(5) O povo aqui me sustenta, e eu não fui... (ZM 28C/487)

O sujeito pronominal é usado para enfatizar o tópico da conversação (Nós estamos falando de X). O pronome sujeito ocorre mais frequentemente com referentes de 1ª e 2ª pessoas do que com os de 3ª pessoa. Os referentes de 1ª e 2ª pessoas são mais tópicos do que os de 3ª que favorece o sujeito zero.

A pesquisa em desenvolvimento a ser relatada aqui, parte dos resultados acima discutidos e tem a finalidade de comparar a língua falada com a língua escrita, inicialmente em relação aos sujeitos pronominais e zero e posteriormente a outros aspectos sintáticos que serão aventados na análise.

Para a modalidade escrita pretendo utilizar cartas familiares, composições de universitários, artigos de jornal e revista e romances modernos. Neste trabalho gostaria de apresentar os dados referente às car-

tas familiares, pois acho que esta modalidade é a que mais se aproxima da amostra da língua falada que utilizarei para a comparação. A amostra da língua falada são entrevistas gravadas por mim com cinco pessoas do sexo feminino, naturais do Rio de Janeiro, conhecidas da pesquisadora, pertencentes à classe média alta. A amostra de língua escrita são cartas familiares dirigidas à pesquisadora por quatro pessoas do sexo feminino, naturais do Rio de Janeiro, pertencentes à classe média alta. Em ambas as amostras há o conhecimento prévio entre os participantes, sendo que nas cartas, a relação entre os participantes é mais próxima. A hipótese inicial é que a frequência de sujeitos pronominais é mais alta na modalidade falada pelo fato de na modalidade escrita haver acesso ao referente do sujeito no texto. As hipóteses quanto aos fatores lingüísticos são as mesmas utilizadas em Lira (1982) e serão expostas no decorrer da análise. A partir dos resultados pretendo construir novas hipóteses para uma análise posterior da língua escrita e comparação com a falada.

A metodologia utilizada na pesquisa segue os pressupostos teóricos e metodológicos da teoria da variação lingüística desenvolvida por W. Labov (1966, 1972, 1980), D. Sankoff (1981) entre outros.

II) Frequência e Análise dos Grupos de Fatores

A tabela abaixo mostra a frequência dos dois tipos de sujeito nas modalidades falada e escrita.

TABELA 2 - Frequência dos sujeitos pronominais e zero na língua falada e escrita.

Sujeito	FALADA		ESCRITA	
	N	%	N	%
Pronominal	884	58	86	22
Zero	631	42	314	78
Total	1515		400	

A tabela acima mostra que a nossa hipótese inicial - menor frequência de sujeitos pronominais na língua escrita - foi confirmada. Evidentemente que temos que analisar outros estilos na língua escrita para ver

se esta hipótese se confirma. Note-se como a frequência dos sujeitos da Tabela 1 é praticamente a mesma da Tabela 2 na língua falada mostrando que é uma variação regular que não parece ser influenciada pelo contexto social distinto das duas amostras - uma, sendo a pesquisadora uma estranha e em outra sendo conhecida.

Como esta pesquisa ainda está na sua fase preliminar, para a análise dos dados iniciais da língua escrita vou medir somente a frequência dos fatores analisados para mais tarde ao chegar a um grupo de fatores mais significativos, usar o modelo matemático probabilístico (Cedergren e Sankoff 1974 e Sankoff 1975). A significação dos grupos de fatores será medida com o teste qui-quadrado sempre que necessário.

Os fatores escolhidos foram aqueles que maior influência tiveram na frequência maior ou menos dos sujeitos pronominais em Lita (1982): referência específica e generalizada, tipo de oração, informação nova e não-nova, referente da oração examinada mesmo ou diferente do da oração anterior.

1. Referência específica e generalizada - A categoria de pessoa é claramente definível com referência à noção dos papéis dos participantes no processo de comunicação. Benveniste (1976:286) diz que é na linguagem e pela linguagem que o homem se constitui como sujeito. Lyons (1977) acrescenta que a situação canônica do enunciado é egocêntrica. Tudo é relacionado através do ponto de vista do falante ou escritor. Assim, a terceira pessoa é negativamente definida com respeito à primeira e segunda pessoas; ela não se relaciona com nenhum papel participante no discurso. As primeiras e segundas pessoas têm basicamente uma referência exofórica — remetem à situação. Elas podem ter uma função anafórica no discurso direto dentro da narrativa (Lira 1982). A terceira pessoa tem principalmente uma função anafórica. Pode ter uma função dêitica na fala que é geralmente reconhecida pela entoação e gestos paralingüísticos. Um outro tipo de referência examinada foi a referência generalizada (impessoal). Os pronomes que a tem significam 'qualquer indivíduo humano', 'uma pessoa não-específica', 'as autoridades' ou 'um grupo particular de indivíduos que o falante deseja identificar-se' (cf. Halliday e Hasan 1976). Pelas características das referências pessoais a nossa hipótese inicial era de que como as 1ª e 2ª pessoas eram

deduzíveis do contexto situacional apareceriam mais como sujeito zero do que sujeito pronominal. A tabela abaixo mostra a frequência deste grupo de fator:

TABELA 3 - Frequência dos sujeitos pronominais e referências específicas e gerais¹.

Fatores	Formas	Língua Falada			Língua Escrita		
		N	Total	%	N	Total	%
1ª pessoa	eu, você, a gente nós, a gente	502	764	65	32	239	15
2ª pessoa	você, vocês	47	62	76	22	37	59
3ª pessoa	ele, ela, eles, elas	304	640	47	25	101	24
generalizada	vocês, a gente	30	43	69	7	20	35
	eles	1	6	16	-	3	-
TOTAL		884	1515	58	86	400	21

A tabela acima mostra que o sujeito pronominal de maior frequência é o de 2ª pessoa nas duas modalidades. Apesar de ter referências distintas nas duas modalidades - na língua falada ele é essencialmente anafórico pois aparece em discurso direto e na modalidade escrita ele é exofórico - a frequência é a mais alta nas duas modalidades. Isto nos leva a concluir que o que parece influenciar mais a sua presença é a sua ambigüidade potencial pelo fato do verbo ser flexionado para terceira pessoa, tanto que em sua forma generalizada também tem uma incidência grande apesar da referência ser completamente distinta. A referência de 1ª pessoa apresentou um comportamento inverso nas duas modalidades. O uso total da referência de 1ª pessoa nas duas modalidades é bem próximo - no discurso falado 764/1515 (50%) e no discurso escrito (59%). Mas na língua escrita ela aparece mais com sujeito zero e na língua falada com o sujeito pronominal. Na língua falada

o caráter egocêntrico do discurso predomina e faz com que a referência de 1ª pessoa seja mais freqüente do que necessária para a compreensão do discurso. A referência de 3ª pessoa tem pouca influência nesta variação pronominal.

2. Tipos de oração - Uma oração é definida como um grupo de palavras contendo um e somente um predicado. Os tipos de orações que foram selecionadas como possíveis influências na presença ou não de pronome su jeito foram:

- a) Independentes - orações cujas estruturas podem aparecer sozinhas sem ter alguma relação com as orações vizinhas.
- (6) P: A tua mulher tem quantos filhos?
N: **Ela tem três.** (N 54A/26)
- b) Principais -
- (7) **Aí** ~~o~~ me perguntou se aquele lugar estava vazio. (Gui 27C/642)
- c) Subordinadas -
1. Adverbiais -
- (8) **Quando eu estou trabalhando**, eu não bebo. (AC 28A/439)
2. Substantivas -
- (9) Mario perguntou **se eu queria tomar um uísque.** (Os. 48C/525)
3. Relativas - A relativo em que o pronome relativo é o sujeito da oração não foi considerada pois no meu corpus o seu comportamento é quase categorico, no uso do sujeito zero, como no exemplo abaixo:
- (10) O marido que é da marinha morava na Bahia. (V 31A/705) As orações relativas consideradas são:
- (11) No dia **que eu tirei o gesso**, ele botou. (R. 64C/75)
- (12) Eu dei um ferro elétrico que eu tinha. (P 73B/279)
- d) Orações Coordenadas - A primeira oração coordenada foi separada das subsequentes.

- (13) Eu panhava plantas no parque Lage
 (14) e guardava,
 (15) botava data e pedrinhas. (MC 21A/220)

A oração (13) foi classificada como coordenada 1 e as orações 2 (14) e (15) como coordenadas 2.

TABELA 4 - Frequência de tipos de oração e sujeitos pro-
 nominais.

Fatores	Língua Falada			Língua Escrita		
	N	Total	%	N	Total	%
Independentes	64	119	53	9	42	21
Principais	176	248	70	17	64	26
Adverbiais	128	193	66	13	62	20
Substantivas	90	137	65	9	28	32
Relativas	88	96	91	7	32	22
Coordenadas 1	129	187	68	14	52	26
Coordenadas 2	209	535	39	17	120	14
Total	884	1515		86	400	

Na língua falada a oração relativa é um fator bem significativo para a presença do sujeito pronominal. Aqui novamente a razão principal é a significação ambígua do tipo de oração analisado. Na língua escrita ela já não apresenta este efeito. O teste χ^2 mostra que ela é não-significante ($P > 80$) o número das relativas na língua escrita é talvez insuficiente. A coordenada 2 é o único tipo de oração que parece ter um efeito semelhante nas duas modalidades - ela inibe a presença do sujeito pronominal. Este resultado era esperado pois o apagamento em estruturas coordenadas é universal (Hankamer 1979). Os outros tipos de orações não têm um efeito muito significativo.

3. Informação Nova e Não-Nova - Este fator refere-se à estrutura informacional do discurso. É o referente do sujeito novo ou não? - Um referente novo é um referente que não foi mencionado no discurso anterior. Concordância verbo-sujeito não foi considerada como tendo sido mencionada no discurso. Se o sujeito fosse classificado como não-novo investiguei aonde o

seu referente apareceu em uma, duas, três ou mais orações anteriores. A minha hipótese é de que os referentes novos e os referentes mencionados há mais de 3 orações fossem os que favorecessem os sujeitos pronominais. A tabela abaixo mostra os resultados deste grupo:

TABELA 5 - Referentes novos e não-novos e sujeitos pronominais.

Fatores	Língua falada			língua escrita		
	N	Total	%	N	total	%
Novo	10	23	43	6	42	14
1 oração	372	661	56	40	162	25
2 orações	162	293	55	18	69	26
3 ou mais orações	340	538	63	22	127	17
Total	884	1515		86	400	

A tabela acima mostra, que dentro das duas modalidades quando o referente do sujeito é novo os sujeitos pronominais não são favorecidos. Uma tabulação feita com os sujeitos novos do corpus de 8924 dados que apresentou os mesmos efeitos, mostrou que dos 138 sujeitos novos (pronominal e zero) 117 são generalizados. Este fato explica os resultados da língua falada. O sujeito generalizador não é importante para a inteligibilidade da mensagem. Se um falante diz:

(16) \emptyset Asfaltaram a avenida. (AC 28A/211)

Eu não vou precisar perguntar quem. A mensagem está completa.

Na língua escrita examinei todos os 42 sujeitos novos e constatei que com exceção de 4 sujeitos com referentes generalizadores, todos sujeitos zero, os outros todos poderiam ser classificados como o que Prince (1981) chama de 'inferíveis'. Inferíveis são entidades que o falante (ou escritor) acha serem deduzíveis pelo ouvinte (ou leitor). Assim todas as referências que eu classifiquei como novas são do tipo:

(17) Querida S.,

Depois de uma boa estadia que se estendeu a-

té Domingo, \emptyset chegamos finalmente. \emptyset Passamos o carnaval em MP com a casa cheia: B, \bar{O} . e filhos, G. com L., M. e nós. (N. 55C-1)

Os dois sujeitos zero foram classificados como novo. Kroch (1982:34) diz que na língua escrita os inferíveis comportam-se mais como entidades novas pois apesar deles fazerem parte do modelo do discurso, eles ainda não foram mencionados no texto. Há necessidade de uma análise mais qualitativa deste grupo de fatores para podermos ver se as diferenças básicas destas duas modalidades é a diferença como elas introduzem os referentes. Os dados referentes aos sujeitos novos na língua escrita não são suficientes ainda para se chegar a uma hipótese.

Um outro fato interessante é que o referente que acabou de ser mencionado tem maior probabilidade de aparecer com sujeito pronominal do que sujeito zero, contrariando a nossa hipótese inicial. Este fato talvez possa ser explicado por um paralelismo ou concordância na superfície. Poplack (1979) encontrou o mesmo fenômeno no estudo da concordância nominal no espanhol portoriquenho. Labov and Wiener (1977) no estudo da passiva e Schiffrin (1981) num estudo dos tempos verbais na narrativa, sugere que talvez haja uma tendência para certas formas gramaticais aparecerem juntas.

4. Referente - Mesmo ou diferente?

Outro aspecto que foi colocado como um fator importante para a presença do sujeito pronominal foi se o referente do sujeito da oração examinada é o mesmo ou diferente do referente do sujeito da oração anterior. A presença do sujeito pronominal será mais frequente se a referência for diferente como o exemplo abaixo:

(18) Eu botei \emptyset para fora no braço, aí **ele** não voltou não. (R. 63C/101)

Os resultados abaixo confirmam as hipóteses:

TABELA 6 - Referentes iguais ou diferentes e sujeitos pronominais.

Fator	Língua falada			Língua escrita		
	N	Total	%	N	Total	%
Igual	428	879	48	36	216	17
Diferente	456	636	71	50	184	27
Total	884	1515		86	400	

Em ambas as modalidades este grupo de fatores tem a mesma influência: se o referente for o mesmo inibirá a presença do sujeito pronominal, se for diferente cau será o aparecimento dele mais freqüentemente.

5. Discussão

Os fatores em ambas as modalidades que estimulam significativamente a presença do sujeito pronominal são a referência da 2ª pessoa, a referência generalizado-ra você e a diferença entre o referente do sujeito da oração analisada do da oração anterior. Os fatores que inibem a presença do sujeito pronominal são a informação nova do referente e a oração coordenada dois.

Os fatores que apresentam um efeito oposto nas duas modalidades, que em parte são responsáveis pela diferença de freqüência do sujeito pronominal, são os referentes de 1ª pessoa e os referentes não-novos que apareceram na fala ou no texto há mais de três orações.

As orações relativas causam quase que categoricamente o aparecimento do sujeito pronominal, mas nos dados de língua escrita analisados não mostrou efeito significativo.

Este estudo preliminar das modalidades falada e escrita além de nos apontar alguns fatores importantes para o condicionamento do sujeito pronominal, nos demonstrou a necessidade de se continuar a pesquisa afim de se encontrar parâmetros mais apropriados para a análise e comparação destas duas modalidades. Os fatores que devem ser re-examinados seriam:

1. Tipos de oração - As orações relativas devem ser estudadas com detalhe afim de podermos caracteri-

zã-las em relação às duas modalidades. Existem tipos diferentes nas duas modalidades? Quais são e até que ponto são importantes para a variação pronominal?

Seria interessante examinar as seqüências de orações coordenadas em termos de tamanho nas duas modalidades. Os outros tipos de oração devem ser repensados e talvez eliminados.

2. Informação nova e não-nova - Pelos resultados da pesquisa preliminar, conclui que é necessário se fazer uma análise mais exaustiva deste aspecto pois este me parece uma das variáveis mais importantes para caracterizar a sintaxe da fala e da escrita. As estratégias que os falantes usam para apresentar informação nova e não-nova parecem ser diferentes e cruciais para definirmos as duas modalidades. Há necessidade de se adotar uma classificação mais detalhada de novo e não-novo. Prince (1981) oferece uma boa alternativa quando propõe que entidades do discurso sejam classificadas como 'new', 'evoked' e 'inferrable'. O termo 'new' refere-se à entidades que estão sendo apresentadas pela primeira vez no discurso, as entidades 'evoked' referem-se àquelas entidades que já foram mencionadas no texto ou aquelas entidades que foram invocadas situacionalmente. As entidades 'inferrable' são entidades que o falante/escritor supõe que possam ser deduzíveis pelo receptor. Prince (1981) apresenta subdivisões deste modelo básico que também devem ser examinadas. Para Prince (1981) 'a text is a set of instructions from a speaker to a hearer on how to construct a particular discourse model'. Assim deveríamos tentar ver qual é este relacionamento nas duas modalidades e até que ponto o relacionamento entre os participantes - íntimo, 1º contato, impessoal - influenciam a coordenação destes modelos de discurso.

Finalizando acho que é necessário estabelecer parâmetros distintos para os vários estilos escritos mencionados na introdução.

NOTAS

¹Os números nas tabelas serão apresentados da seguinte forma: N = número dos sujeitos pronominais; T = sujeitos pronominais e zero; % = percentagem dos sujeitos pronominais.

²Existem poucos exemplos do pronome cópia nos meus dados. *Molli-*

ca (1977) e Tarallo (1983) estudaram este tipo de pronome nas relativas.

BIBLIOGRAFIA

- CEDERGREN, H. and D. Sankoff. Variable rules: performance as a statistical reflection of competence. *Language*, 50:333-55, 1974.
- CUNHA, C. *Gramática do português contemporâneo*. Belo Horizonte, Editora Bernardo Alvares S.A., 1972.
- HAWKINS, J.A. *Definiteness and Indefiniteness*. NJ: Humanities Press, 1978.
- KROCH, A. *A Quantitative Study of the Syntax of Speech and Writing*. Final Report to the NIE on grant n° G78-0169. Philadelphia: University of Pennsylvania, 1982.
- LABOV, W. *The Social Stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics, 1966.
- LABOV, W. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press, 1972.
- LIRA, S. de A. *Nominal, Pronominal and Zero Subject in Brazilian Portuguese*. Unpublished Ph.D. Dissertation, University of Pennsylvania, 1982.
- MOLLICA, M.C. *Estudo da cópia nas construções relativas em português*. M.A. Thesis PUC/RJ, Rio, 1977.
- POPLACK, S. *Function and process in a variable phonology*. Unpublished University of Pennsylvania Ph.D. dissertation, 1979.
- PRINCE, E.F. *Toward a Taxonomy of given/new information*. In: COLE, P. (ed.) *Radical Pragmatics*. New York, 1981.
- SANKOFF, D. *Varbrul version 2*. Unpublished program and documentation, 1975.
- SANKOFF, G. *The Social Life of Languages*. Philadelphia: University of Pennsylvania Press, 1980.
- SHIFFRIN, D. *Tense variation in Narrative*. *Language*, 57: 45-62, 1981.

TARALLO, F. **Relativization strategies in Brazilian Portuguese.** Ph.D. Dissertation. Univ. of Penn., 1983.

WEINER, E.J. and W. Labov. **Constraints on the agentless passive.** 39th LSA Summer Meeting, Honolulu, 1977.