

## “NEGO” TAMBÉM INDETERMINA O SUJEITO EM SALVADOR

Valter de CARVALHO

Universidade do Estado da Bahia / Secretaria Estadual de Educação – Bahia  
valtinhodias@gmail.com

**Resumo:** A indeterminação do sujeito tem sido estudada no Brasil de maneira sistemática, até onde se sabe, desde 1982, em diversas regiões do país e considerando os mais variados aspectos linguísticos e extralinguísticos. Na Bahia, em particular, houve estudos realizados sobre o português falado na zona rural. Este trabalho se voltou para a fala urbana, mais especificamente a capital, Salvador. A investigação analisou não só as formas consideradas canônicas pelas gramáticas tradicionais, mas também outras estratégias, como as “formas nominais” (o cara, o sujeito, o pessoal, entre outras), objeto do presente estudo. O trabalho foi desenvolvido de acordo com o enquadramento teórico-metodológico da Sociolinguística Variacionista, buscando identificar os contextos extralinguísticos ou sociais e linguísticos que justificassem os usos encontrados. Foram analisados quarenta e quatro inquiridos do tipo DID (Diálogo entre Informante e Documentador) do Programa de Estudos sobre o Português Popular Falado de Salvador (PEPP) e do Projeto de Estudo da Norma Linguística Urbana Culta de Salvador (NURC-SSA). Os dados foram submetidos à quantificação através do pacote de programa Varbrul 2S e do GoldVarb X. Ao final, percebe-se que os falantes soteropolitanos fazem uso das mais diversas estratégias para indeterminar o sujeito, influenciados principalmente por fatores linguísticos.

**Palavras-chave:** indeterminação do sujeito; português; sociolinguística.

### 1 Introduz-se dizendo que...

A indeterminação do sujeito é o objeto de estudo deste trabalho e é entendido como a indeterminação do referente, ou seja, o agente da ação verbal que se encontra em um contexto externo ao linguístico, compartilhando do mesmo mundo concreto que os falantes participantes da interação verbal no ato de comunicação, conforme apontado por Menon (2006, p. 129):

Para mim, indeterminação do sujeito concerne os casos em que não se pode ou não se quer nomear o sujeito, na acepção de ‘referente extralinguístico’. No entanto, o referente é conhecido pelo locutor (e em certos casos, também do interlocutor, o que torna possível a compreensão mútua) e se ele quisesse ou se isso lhe fosse conveniente ou interessante, ele poderia nomeá-lo ou descrevê-lo. Nesse sentido, o referente pode ser recuperado pelo locutor a qualquer hora. Trata-se, antes de tudo, de uma maneira de escamotear o sujeito extralinguístico por meio de uma forma de expressão linguística, em função da situação de comunicação.

Analisando algumas gramáticas tradicionais, notou-se que não há uma explicação aceitável no contexto sintático para o uso do termo “indeterminado” atribuído ao sujeito. Ao observar as outras classificações do sujeito apresentadas por esses compêndios, tal como “oculto”, “simples”, “composto” entre outras, verificou-se que elas fazem referência à forma,

em oposição à classificação de sujeito “indeterminado” que é semântica (MILANEZ, 1982, p. 11).

Almeida (2005, p. 414), em sua “Gramática Metódica da Língua Portuguesa”, diz que “o sujeito é indeterminado quando de impossível identificação”, podendo ser empregado com verbos tidos como “ativos, acidentalmente impessoalizados na 3ª pessoa do plural”, conforme exemplo 01, ou com verbos “acidentalmente impessoalizados na passiva”, como o exemplo 02.

(01) “Dizem que ele vem.” (ALMEIDA, 2005, p. 414)

(02) “Assim se vai aos céus.” (ALMEIDA, 2005, p. 414)

Rocha Lima (2008, p. 235, grifos do autor), mantendo o mesmo postulado que o gramático anterior, diz que é sujeito indeterminado “se não pudermos ou não quisermos especificá-lo”.

O gramático completa sua definição dizendo que

Vale-se a língua de um dos dois expedientes:

- 1) Empregar o verbo na 3ª pessoa do plural, sem referência anterior ao pronome *eles* ou *elas*, e a substantivo no plural [cf. exemplos 03 e 04];
- 2) Usá-lo na 3ª pessoa do singular acompanhado da partícula *se*, desde que o verbo seja intransitivo, ou traga complemento preposicional [cf. exemplos 05 e 06]. (ROCHA LIMA, 2008, p. 235)

Há ainda exemplos citados por Rocha Lima (2008, p. 235) para elucidar melhor sua explicação:

(03) “Falam mal daquela moça.”

(04) “Mataram um guarda.”

(05) “Vive-se bem aqui.”

(06) “Precisa-se de professores.”

De forma bem sucinta, Cegalla (2008, p. 325, grifos do autor) afirma que o sujeito indeterminado é “quando não se indica o agente da ação verbal”. Para marcá-lo, o gramático apresenta três possibilidades, diferenciando-se dos gramáticos citados anteriormente:

Em português, assinala-se a indeterminação do sujeito de três modos:

- a) usando-se o verbo na 3ª pessoa do plural, sem referência a qualquer agente já expresso nas orações anteriores. Exemplos:  
Na rua *olhavam-no* com admiração.  
“*Bateram* palmas no portãozinho da frente.” (Josué Guimarães)  
“De qualquer modo, foi uma judiação *matarem* a moça.” (Rubem Braga)
- b) com um verbo ativo na 3ª pessoa do singular, acompanhado do pronome *se*. Exemplos:  
Aqui *se vive* bem.  
Devagar *se vai* ao longe.  
Quando *se é* jovem, a memória é mais vivaz.  
*Trata-se* de fenômenos que nem a ciência sabe explicar.  
“E *passou-se* a falar em internacionalização da Amazônia.” (Tiago de Melo)  
[...]
- c) deixando-se o verbo no infinitivo impessoal. Exemplos:  
Era penoso *carregar* aqueles fardos enormes.

É triste *assistir* a estas cenas repulsivas.

Em sua “Gramática de usos do português”, Neves (2000, p. 463-465, grifos da autora) apresenta outras formas tidas como mais inovadoras<sup>1</sup> para marcar o sujeito indeterminado, inclusive lexicalmente preenchido, conforme os exemplos 07, 08 e 09, sendo que esse último ela considera como indeterminação parcial, uma vez que ele “só abrange o universo das terceiras pessoas”.

- (07) “**VOCÊ** vai lá, fica dois dias fazendo curso, eles te catequizam, fazem **VOCÊ** comprar uma tonelada de sabão e abrir o seu negócio. (OMT)”
- (08) “**EU** vou lá, fico dois dias fazendo curso, eles **ME** catequizam, **ME** fazem comprar uma tonelada de sabão e abrir o meu negócio.”
- (09) “Sabe como é, quando a gente se acostuma com uma coisa, **ELES** inventam outra. (E)”

Além disso, Neves também apresenta outros exemplos nos quais o sujeito não está lexicalmente preenchido. Os exemplos 10 e 11 estão em consonância com os ditames tradicionais da gramática, enquanto os exemplos 12 e 13 caracterizam-se como mais inovadores, uma vez que não são contemplados nas gramáticas normativas. Inclusive a autora faz menção ao exemplo 16 como sendo “menos comum e de registro mais popular”. Já o exemplo 17 é classificado como marca de indeterminação parcial, pois “pelo menos uma referência é determinada, porque sempre está incluído o falante (o eu)”.

- (10) “**Jogaram** alguém na piscina; a velha cena da festinha em que todo mundo cai na piscina. (BL)”
- (11) “Falava-**SE** de Pedro. (A)”
- (12) “Lá **tira** título de eleitor, documento. (HO)”
- (13) “**NÓS**, todos **NÓS**, o ser humano não suporta o sucesso de outro ser humano, **NÓS** odiamos o Pelé. (OMT)”

Uma visão mais inovadora sobre a indeterminação do sujeito, que inclusive é apontada por alguns trabalhos sobre esse objeto, já é encontrada na gramática de Azeredo (2008, p. 225), elaborada em homenagem ao filólogo Antônio Houaiss:

Orações de sujeito indeterminado são empregadas por motivos cognitivos ou discursivos variados, e a língua oferece a seus usuários diferentes meios para indeterminar, dissimular ou mesmo ocultar a identidade do ser humano a quem o sujeito da oração se refere. A razão cognitiva óbvia é o desconhecimento da identidade do ser de que se fala. As razões discursivas, por sua vez, são variadas: a conveniência ou oportunidade da omissão da identidade do sujeito é uma delas, o registro de linguagem empregado ou o gênero de texto produzido é outra.

Esse autor cita os processos mais tradicionais apontados pelos autores anteriormente citados, mas não deixa de incluir também outras formas para marcar a indeterminação do sujeito, além das formas canônicas apresentadas. Contudo, ele faz uma ressalva que essas outras formas fazem parte de uma “interação mais espontânea”, como o verbo na terceira

<sup>1</sup> Aqui, inovador é todo recurso utilizado para marcar a indeterminação do sujeito que não sejam as formas canônicas apontadas pelas gramáticas tradicionais - GT.

pessoa do singular sem o “se” e sintagmas nominais com significação genérica ou indeterminadora, tal como “a gente, muita gente, todo mundo no papel de sujeito simples”.

Sobre essa fala mais “espontânea”, Said Ali (2006, p. 162, grifos do autor) já dizia que

o falar coloquial não se opõe obstinadamente ao que está em uso na linguagem literária e entre as pessoas cultas. Tanto se serve da expressão *a gente*, como do verbo acompanhado do pronome *se* para dar a entender que o sujeito psicológico do verbo é pessoa ou são pessoas indeterminadas.

Embora as inovações linguísticas em relação à indeterminação do sujeito tenham sido citadas apenas pelas gramáticas descritivas abordadas aqui, Haury (1986, p. 58) já mencionava outras obras que apresentavam algumas das inovações levantadas, além, é claro, dos padrões apontados. Há, por exemplo, registro do uso da forma “a gente” na “Gramática fundamental da língua portuguesa” de Gladstone Chaves de Melo, na qual o gramático cita um trecho do livro “Três romances” de Raquel de Queiroz como exemplo de marcação do sujeito indeterminado, tal como as gramáticas consideradas mais tradicionais fazem uso dos grandes escritores da língua portuguesa tidos como clássicos:

- (14) “Passando num meio-dia quente, ao trote penoso do cavalo, *a gente* pára ali...”  
(HAURY, 1986, p. 58 *apud* MELO, 1970, p. 194, grifo do autor)

O trabalho inicial que possibilitou fazer este recorte de estudo identificou nos *corpora* observados um total de 2595 dados, distribuídos entre as 11 variantes selecionadas. A tabela 1, a seguir, mostra o número de casos, como também sua porcentagem em relação ao total.

VARIANTES	Nº	%
Você	873	33,6
A gente	668	25,7
Formas Nominais	338	12,5
Ø+V3PS	185	7,1
Ø+VINFIN	171	6,6
Nós	115	4,4
Ø+V+SE	78	3,0
Ø+V3PP	79	3,0
Eles	68	2,6
Eu	18	0,7
PASSINT	2	0,1
<b>TOTAL</b>	<b>2595</b>	

**Tabela 1:** Frequência dos recursos de indeterminação levantados nos *corpora*.

A variante formas nominais, doravante FN, foi a terceira que mais ocorreu nos *corpora* analisados, figurando 12,5% do total de estratégias utilizadas pelos falantes de Salvador para marcar a indeterminação do referente do sujeito. Ela se destaca como objeto deste trabalho por não serem classificadas como pronominais, pois não exercem essa função, tão pouco serem comparadas com outras estratégias não-pronominais, formadas principalmente por um verbo sem sujeito lexical expresso, sendo que elas ocupam a função de sujeito preenchido.

O termo “formas nominais” foi empregado por Menon (1994; 2006) para designar as composições formadas por artigo definido mais substantivo, como “o camarada, o cara, o cidadão, o homem, o indivíduo, a pessoa, o público, o sujeito, a turma”.

Nesta pesquisa, além das estratégias citadas, foram encontradas as formas “nego” e “neguinho”, que também foram incluídas nesta variante, apesar de não serem constituídas de artigo definido, mas serem bastante representativas para a história e cultura baiana, muito presentes em Salvador, capital do Estado. Daí a importância de marcar essa estratégia no título deste trabalho.

A forma “a gente” não consta entre as FN por já constituir isoladamente uma variante, uma vez que deixou de ser sintagma nominal, formado por seu núcleo mais um artigo definido, gramaticalizando-se em um pronome de primeira pessoa do plural, bastante utilizado também como uma estratégia de indeterminação do sujeito.

Após a caracterização do que se considera sujeito indeterminado e apontado o recorte realizado na pesquisa inicial, parte-se agora para explicar como se deu o desenvolvimento do estudo.

## 2 A gente realizou da seguinte maneira:

A pesquisa que ora se apresenta foi realizada de acordo com o enquadramento teórico-metodológico da Sociolinguística Variacionista ou Laboviana, com base nas formulações de Labov (1994), buscando identificar os contextos linguísticos e extralinguísticos (ou sociais) que justificam os usos encontrados acerca do sujeito indeterminado, objeto do presente estudo.

Os estudos da Sociolinguística Variacionista têm como base o vernáculo, ou seja, a língua falada em contextos reais de uso, sinalizando as diferentes formas linguísticas que os falantes utilizam para expressar seus pensamentos, ideias, que se correspondem de alguma forma, caracterizando-as como variantes linguísticas; dito isso de outra maneira, as formas diferentes para tratar mais ou menos da mesma coisa.

As formas nominais foram selecionadas, conforme mencionado anteriormente, dentre um grupo de doze, classificadas em dois grupos: com sujeito lexical (a gente, eles, eu, formas nominais, nós e você) e sem sujeito lexical ( $\emptyset$ +V+SE – verbo na terceira pessoa do singular mais a partícula “se”,  $\emptyset$ +V3PS – verbo na terceira pessoa do singular,  $\emptyset$ +V3PP – verbo na terceira pessoa do plural,  $\emptyset$ +VINFIN – verbo no infinitivo, VPSA – voz passiva sem agente e VPASSINT – passiva sintética<sup>2</sup>). Essas foram estabelecidas a partir de estudos prévios realizados (CARVALHO, 2009a; 2009b) e também por observação de outros trabalhos existentes sobre o sujeito indeterminado (MILANEZ, 1982; ROLLEMBERG, 1991; CUNHA, 1993; MENON, 1994; SETTI, 1997; CAVALCANTE, 1999; GODOY, 1999; SANTANA 2006; PONTE, 2008).

A variante FN foi analisada sob os pontos de vista sociais e linguísticos. Verificou-se se os fatores sociais gênero/sexo, escolaridade e faixa etária afetaram as escolhas de uso de uma ou de outra forma por parte dos falantes. A mesma observação foi tida com as variáveis linguísticas forma antecedente, mudança / manutenção do referente, tempo e modo verbais, tipo de oração, tipo de verbo, preenchimento do sujeito e grau de indeterminação, buscando perceber se haveria alguma influência desses fatores nos usos encontrados.

### 2.1 Os corpora

<sup>2</sup> Esta estratégia fazia parte dos objetivos da pesquisa, mas não foram encontradas ocorrências.

Os dados para o presente estudo foram coletados a partir de quarenta e quatro (44) inquéritos: trinta e dois (32) do PEPP — Programa de Estudos sobre o Português Popular Falado de Salvador — e doze (12) do NURC/SSA — Projeto Norma Urbana Culta de Salvador.

O PEPP surgiu diante da necessidade de investigação do português falado por falantes com níveis de escolaridade Fundamental (até a antiga 5ª série ou atual 6 ano) e Médio, e moradores, em sua maioria, de bairros mais distantes do centro da cidade, tidos como de maior concentração de pessoas de baixa renda, com raras exceções, caracterizando, portanto, o termo “popular” empregado em sua denominação.

Segundo Lopes (2009, p. 17), o PEPP adotou os mesmos critérios do NURC/SSA para seleção de informantes, uma vez que seria possível fazer um estudo comparativo entre os níveis de escolaridade de ambos os projetos, que se completam, tendo em vista a classificação de escolaridade estabelecida no Brasil. Assim, em relação aos informantes, exigiu-se que

eles fossem naturais de Salvador e, além disso, que tivessem permanecido nesta cidade a maior parte de suas vidas. Para excluir a interferência de outros dialetos, seus pais também deveriam ser de Salvador, ou que tivessem vindo para esta cidade ainda muito pequenos.

O Projeto NURC, por sua vez, tem por objetivo descrever os padrões reais de uso na comunicação oral adotados pelo estrato social composto por indivíduos de escolaridade superior. Para atingir tal objetivo, ele buscou, e vem buscando, dispor de material sistematicamente levantado, superar o empirismo na aprendizagem da língua-padrão, conhecer as normas tradicionais que estão vivas e quais as superadas, e oferecer dados para que se pudesse corrigir distorções do esquema tradicional da educação brasileira.

O *corpus* desse projeto é constituído de elocuições formais (EF), formadas por aulas e conferências, diálogos entre informante e documentador (DID) e diálogos entre dois informantes (D2). Deveriam fazer parte do *corpus* também gravações secretas de diálogos espontâneos (GS), os quais não chegaram a ser realizados. A duração das gravações varia entre 40 e 80 minutos no total.

Os inquéritos estão classificados segundo o gênero/sexo dos informantes (masculino e feminino) e em relação à faixa etária (de 25 a 35 anos, de 36 a 55 anos e acima de 56 anos).

O projeto é desenvolvido por equipes regionais, localizadas na Universidade Federal da Bahia (NURC/SSA), Universidade Federal de Pernambuco (NURC/PE), Universidade Federal do Rio de Janeiro (NURC/RJ), Universidade Federal do Rio Grande do Sul (NURC/POA) e pelas Universidades de São Paulo e Estadual de Campinas (NURC/SP).

Os *corpora* apresentados se organizam de acordo com cada variável social citada anteriormente. Para a variável escolaridade foram considerados o Ensino Fundamental e Médio (com os inquéritos do PEPP) e o Ensino Superior (com os inquéritos do NURC). Em relação ao gênero/sexo, foram considerados o masculino e o feminino<sup>3</sup>. E, por fim, a variável faixa etária que foram estabelecidas em número de quatro faixas etárias: faixa etária 1 ou F1 (dos 15 aos 24 anos), faixa etária 2 ou F2 (dos 25 aos 35 anos); faixa etária 3 ou F3 (dos 45 aos 55 anos) e faixa etária 4 ou F4 a partir dos 56 anos<sup>4</sup>.

<sup>3</sup> Embora os inquéritos originalmente tratem de “homem” e “mulher” para se referir à variável gênero/sexo, aqui foi adotada essa nomenclatura para não fazer menção ao caráter biológico dos seres, mas, sim, aos papéis sociais desempenhados por eles.

<sup>4</sup> A nomenclatura adotada para as faixas etárias (F1, F2, F3 e F4) foi adaptada, modificando a classificação original do NURC, uma vez que esse projeto não contempla a primeira faixa etária, devido aos poucos informantes encontrados, e englobando as classificações adotadas pelo PEPP. Além disso, o NURC preenche a F4 com informantes acima de 56 anos, enquanto o PEPP com informantes acima de 65 anos. Dessa forma, manteve-se a informação “mais de 56 anos” por contemplar os dois *corpora*.

Tendo em vista o exposto, os informantes foram equitativamente distribuídos em vinte e duas células, contendo dois informantes cada uma, conforme pode ser verificado no quadro logo a seguir:

	ENSINO FUNDAMENTAL (PEPP)		ENSINO MÉDIO (PEPP)		ENSINO SUPERIOR (NURC)	
	Masculino	Feminino	Masculino	Feminino	Masculino	Feminino
<b>F1</b> (15 – 24 anos)	2	2	2	2	-	-
<b>F2</b> (25- 35 anos)	2	2	2	2	2	2
<b>F3</b> (45 – 55 anos)	2	2	2	2	2	2
<b>F4</b> (mais de 56 anos)	2	2	2	2	2	2

**Quadro 1:** Distribuição dos informantes nas células.

## 2.2 Parâmetros de seleção das ocorrências

As entrevistas, objetos de análise deste estudo, foram realizadas na década de 90 do século passado e transcritas grafematicamente, o que possibilitou a leitura dos textos transcritos e digitalizados diretamente na tela de um computador, descartando a necessidade de audição do material gravado.

A seleção das ocorrências passou por alguns critérios para estabelecer se elas eram consideradas uma variante das estratégias de indeterminação do sujeito ou não. Assim, algumas restrições foram necessárias, tais como:

- i) Uso de dois recursos de indeterminação imediatamente um após o outro, caracterizando uma correção em relação ao item utilizado;
- ii) Recursos apresentados na fala de terceiros, algo como uma citação que ocorre na escrita;
- iii) Repetições da fala do documentador, o que marca o chamado “efeito gatilho”, pois o informante foi condicionado pela fala do outro;
- iv) Em interrupções, marcadas pelo uso das reticências na transcrição, pois não se sabe ao certo qual foi a intenção do informante, podendo revelar indecisão quanto ao uso pretendido;
- v) Estruturas refeitas, inclusive com mudança do verbo e seu tempo;
- vi) O primeiro uso de uma estratégia em estruturas idênticas e subsequentes.

## 2.3 O suporte quantitativo

Definidos os *corpora*, as variantes linguísticas, as variáveis extralinguísticas e linguísticas, torna-se necessário definir o suporte quantitativo que dará conta de fazer com que os dados linguísticos sejam processados e obtenham-se os dados numéricos estatístico-probabilísticos importantes para a análise qualitativa.

A quantificação dos dados e o cruzamento das variáveis são importantes para os estudos da variação linguística porque permitem ao pesquisador, segundo Guy e Zilles (2007,

p.73) “apreender sua sistematicidade, seu encaixamento linguístico e social e sua eventual relação com a mudança linguística.”

Para realizar tal procedimento, o pesquisador se utiliza do Varbrul (do inglês *Variable Rules*), que é um pacote ou conjunto de programas computacionais que realiza a análise multivariada, especialmente elaborado para tratar de dados relativos à variação linguística.

A análise multivariada, ainda segundo Guy e Zilles (2007, p.105), “permite investigar situações em que a variável linguística em estudo é influenciada por vários elementos do contexto, ou seja, múltiplas variáveis independentes.”

O pacote de programas Varbrul, em sua versão conhecida como Varbrul 2S, possui uma versão mais atualizada e inovadora que constitui em apenas um programa, chamado GoldVarb X, o qual desenvolve praticamente o mesmo trabalho de sua versão inicial, com exceção de análises ternárias e eneárias de até cinco variantes na variável dependente.

Obviamente, o programa (ou pacote de programas) precisa que o pesquisador em sociolinguística estabeleça códigos para cada fator das variáveis dependente e independentes. Assim, ele processará os códigos informados, o que permitirá cruzar todas as variáveis e estabelecerá a definição dos grupos de fatores que influenciam ou não influenciam nas seleções das variantes linguísticas.

Para este trabalho, foram atribuídos códigos distintos para cada fator de cada grupo, conforme lista abaixo:

- Variável dependente: A gente (g), Eles (l), Eu (e), Formas Nominais (@), Nós (n), Você (v), Ø+V+SE (\$), Ø+V3PP (#), Ø+V3PS (&), Ø+VINF (i), VPSA (s) e VPASSINT (a)<sup>5</sup>;
- Variáveis independentes:
  - Variável gênero/sexo: Masculino (M) e Feminino (F);
  - Variável faixa etária: Faixa etária 1 (1), Faixa etária 2 (2), Faixa etária 3 (3) e Faixa etária (4);
  - Variável escolaridade: Ensino Fundamental (f), Ensino Médio (m) e Ensino Superior (s);
  - Variável tempo e modo verbal: Presente do Indicativo (h), Pretérito Perfeito (d), Pretérito Imperfeito do Indicativo (i), Futuro do Presente (f), Futuro do Pretérito (c), Presente do Subjuntivo (j), Pretérito Imperfeito do Subjuntivo (e), Futuro do Subjuntivo (u), Infinitivo (n), Particípio (p), Gerúndio (g) e Imperativo (m);
  - Variável forma antecedente: Forma precedida de a gente explícito (g), Forma precedida de a gente implícito (G), Forma precedida de eles explícito (l), Forma precedida de eles implícito (L), Forma precedida de eu explícito (e), Forma precedida de eu implícito (E), Forma precedida de FN explícito (@), Forma precedida de FN implícito (A), Forma precedida de nós explícito (n), Forma precedida de nós implícito (N), Forma precedida de você explícito (v), Forma precedida de você implícito (V), Forma precedida de Ø+V+SE (\$), Forma precedida de Ø+V3PP (#), Forma precedida de Ø+V3PS (&), Forma precedida de Ø+VINF (i), Forma precedida de VPSA (s) e Forma precedida de VPASSINT (a);
  - Variável mudança/manutenção do referente: Referente igual à estrutura imediatamente anterior na série discursiva (I) e Referente diferente da estrutura imediatamente anterior na série discursiva (D);

<sup>5</sup> Cabe salientar que o trabalho foi desenvolvido com doze variantes, mas neste texto houve um recorte, tratando exclusivamente da variante formas nominais – FN.



- Variável preenchimento do sujeito: Sujeito pronominal explícito (e), Sujeito pronominal implícito (i) e as demais formas não pronominais receberam o código não se aplica (/);
- Variável tipo de verbo: Ligação (L), Transitivo (T) e Intransitivo (I);
- Variável grau de indeterminação: Parcial (p) e Completa (c); e
- Variável tipo de oração: absoluta (A), coordenada (C), subordinada (S) e principal (P).

Após o estabelecimento dos códigos, o pesquisador passa a analisar cada dado encontrado no *corpus* estudado. Os programas do pacote Varbrul ou o próprio GoldVarb X farão o processamento estatístico a partir da leitura dos códigos atribuídos.

No trecho a seguir, por exemplo, extraído do inquérito 18 do PEPP, será analisado o emprego da forma indeterminadora do sujeito “a gente”. Inicialmente, coloca-se a abertura do parêntese “(“ para informar ao programa que se trata de uma cadeia de códigos que ele deverá ler; é como se fosse um diálogo entre o pesquisador e o programa, a fim de garantir o processamento estatístico. Em seguida, colocam-se os códigos referentes aos grupos de fatores selecionados.

*Exemplo:*

**(gMlfhgIeTcC** Sempre tem, na Fonte Nova também sempre teve, e agora está tendo sobre o boxe, a gente troca a nota e **A GENTE** vai assistir.

A ordem de colocação dos códigos deve ser obedecida rigorosamente para cada dado que for analisado, para não misturar as informações e prejudicar toda a análise. No exemplo, logo após o “(“, encontra-se o “g” informando ao programa que a variante que está sendo observada é o “a gente” destacado no trecho. Os três códigos seguintes dizem respeito às variáveis sociais: M (masculino), 1 (faixa etária) e f (Ensino Fundamental). Depois, continua-se com os demais códigos relativos às outras variáveis estabelecidas neste trabalho, conforme mencionado anteriormente.

Depois de obtido toda codificação dos dados, processam-se as informações diretamente nos programas. O resultado será apresentado primeiramente através de número e porcentagens de ocorrências e, depois, nas chamadas rodadas binárias (ternárias ou eneárias, de acordo com a versão do programa utilizada e a quantidade de variantes linguísticas), serão fornecidos os pesos relativos, tomando como base o inter cruzamento de todas as variáveis informadas.

De posse de todos os resultados, a sociolinguística passará a interpretar os dados estatístico-probabilísticos, a fim de justificar os usos linguísticos que foram escolhidos como objeto de estudo.

### **3 Nego, o cara, o sujeito, o pessoal entre outros**

A variante FN foi a terceira que mais ocorreu nos *corpora* analisados, conforme mencionado na introdução. Ela passou por um procedimento estatístico-probabilístico, através do qual foi possível obter sua frequência frente a algumas variáveis extralinguísticas e linguísticas, como também seu peso relativo referente a cada fator pertencente a essas variáveis.

Para que fosse realizada a rodada binária, algumas modificações no arquivo de condições foram necessárias para que os nocautes iniciais fossem eliminados. Primeiro, excluiu-se a forma “eles” explícito e implícito do grupo de fatores forma antecedente. Depois,

exclusão total da variável preenchimento do sujeito, uma vez que as FN só ocorreram preenchidas. Por fim, a amalgamação da oração principal composta por duas coordenadas com a oração coordenada da variável tipo de oração. Feito isso, o procedimento de obtenção da frequência e pesos relativos foi efetivado, cujos dados serão apresentados a seguir na mesma ordem de seleção do programa.

### 3.1 Variáveis selecionadas

**Forma antecedente** – este foi o primeiro grupo de fatores selecionado para mostrar o quanto é importante que uma forma nominal seja expressa para que outras continuem a ser usadas. Isso é evidenciado por ver que o peso relativo foi de 0.872 para as formas nominais antecedendo-as a si mesmas. Esse resultado se manteve ao verificar o percentual de frequência de 46,9% (conferir tabela 2).

<b>Forma antecedente</b> Significância = 0.001	<b>Apl./T</b>	<b>%</b>	<b>P.R.</b>
<b>Formas nominais</b>	91/194	46,9	0.872
<b>Formas sem sujeito lexical</b>	21/269	7,8	0.522
<b>A gente</b>	28/423	6,6	0.472
<b>Nós</b>	4/68	5,9	0.390
<b>Você</b>	36/605	6,0	0.375
<b>Eu</b>	1/13	7,7	0.302
<b>TOTAL 181/1572</b>			

**Tabela 2:** Aplicação da variante “Formas Nominais” em relação à variável forma antecedente.

Tentou-se realizar o cruzamento dessa variável com a mudança/manutenção do referente, contudo ocorreram vários nocautes, que até as formas nominais teriam que ser excluídas.

**Grau de indeterminação** - dentre os outros recursos utilizados pelos falantes de Salvador para marcar a indeterminação do sujeito, as formas nominais foram uma das poucas estratégias que apresentaram maior peso relativo para a indeterminação completa, com 0.661 (ver tabela 3), revelando que elas são mais empregadas quando de fato não há possibilidade, mesmo por meio de inferências, ter noção acerca do referente extralinguístico, ou seja, o sujeito referencial, isso por possibilitar que qualquer pessoa envolvida ou não no discurso possa ser concebida como sujeito.

<b>Grau de indeterminação</b> Significância = 0.001	<b>Apl./T</b>	<b>%</b>	<b>P.R.</b>
<b>Indeterminação Completa</b>	205/1011	20,3	0.661
<b>Indeterminação Parcial</b>	133/1584	8,4	0.395
<b>TOTAL 338/2595</b>			

**Tabela 3:** Aplicação da variante “Formas Nominais” em relação à variável grau de indeterminação.

**Tempo e modo verbal** - o gerúndio e o pretérito imperfeito do subjuntivo mostraram-se como os principais fatores favorecedores para o emprego das formas nominais, com pesos

relativos de 0.662 e 0.659 respectivamente. Em seguida, não tão distante deles, o pretérito perfeito do indicativo, com peso relativo de 0.649.

<b>Tempo e Modo Verbal</b> Significância = 0.001	<b>Apl./T</b>	<b>%</b>	<b>P.R.</b>
<b>Gerúndio</b>	23/97	23,7	0.662
<b>Pretérito Imperfeito do Subj.</b>	7/40	17,5	0.659
<b>Pretérito Perfeito do Ind.</b>	28/136	20,6	0.649
<b>Presente do Ind.</b>	172/1104	15,6	0.545
<b>Pretérito Imperfeito do Ind.</b>	67/653	10,3	0.514
<b>Futuro do Presente</b>	5/73	6,4	0.331
<b>Infinitivo</b>	20/405	4,9	0.293
<b>TOTAL 322/2513</b>			

**Tabela 4:** Aplicação da variante “Formas Nominais” em relação à variável tempo e modo verbal.

**Mudança/manutenção do referente** - o uso das formas nominais mostrou-se mais proeminente quando o referente é diferente, com peso relativo de 0.746. Sempre que se muda o referente, a tendência é mudar a estratégia para marcar essa mudança.

Infelizmente, não foi possível cruzar essa variável com a forma antecedente, na qual revelou que as formas nominais são preservadas nas cadeias oracionais, uma forma leva a outra forma.

<b>Mudança/Manutenção do Referente</b> Significância = 0.001	<b>Apl./T</b>	<b>%</b>	<b>P.R.</b>
<b>Diferente</b>	26/96	27,1	0.746
<b>Igual</b>	150/1512	9,9	0.483
<b>TOTAL 176/1608</b>			

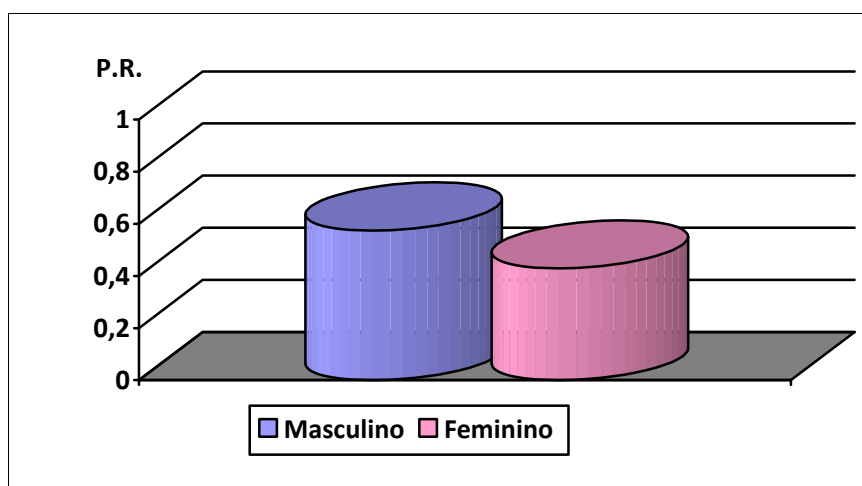
**Tabela 5:** Aplicação da variante “Formas Nominais” em relação à variável mudança/manutenção do referente.

**Gênero/sexo** - esta variável foi a única extralinguística selecionada para a aplicação das formas nominais. Nesse grupo de fatores, há uma leve tendência de favorecimento do gênero/sexo masculino, com peso relativo de 0.574, não tão distante do feminino que apresenta o peso de 0.429.

<b>Gênero/sexo</b> Significância = 0.001	<b>Apl./T</b>	<b>%</b>	<b>P.R.</b>
<b>Masculino</b>	207/1269	16,3	0.574
<b>Feminino</b>	131/1326	9,9	0.429
<b>TOTAL 338/2595</b>			

**Tabela 6:** Aplicação da variante “Formas Nominais” em relação à variável gênero/sexo.

Os resultados obtidos e apresentados na tabela anterior podem ser verificados de maneira ilustrativa no gráfico a seguir, quando se percebe claramente a distância existente entre os dois gêneros/sexos.



**Gráfico 1:** Aplicação da variante “Formas Nominais” em relação à variável gênero/sexo.

A fim de conhecer melhor o gênero/sexo masculino que mais faz uso das formas nominais, realizou-se o cruzamento com as demais variáveis sociais, cujos resultados estão registrados na tabela 7.

Significância = 0.000		Ensino Fundamental			Ensino Médio			Ensino Superior		
Gênero/ Sexo	Faixa Etária	Apl./T	%	P.R.	Apl./T	%	P.R.	Apl./T	%	P.R.
<b>Masculino</b>	1	9/112	8,0	0.396	27/182	14,8	0.567	xxx	xxx	xxx
	2	14/84	16,7	0.600	36/164	22,0	0.679	13/104	12,5	0.518
	3	13/63	20,6	0.661	32/143	22,4	0.684	23/128	18,0	0.622
	4	12/77	15,6	0.581	11/122	9,0	0.427	17/90	18,9	0.636
<b>Feminino</b>	1	11/105	10,5	0.468	8/74	10,8	0.476	xxx	xxx	xxx
	2	9/65	13,8	0.547	17/88	19,3	0.643	28/227	12,3	0.514
	3	6/71	8,5	0.409	3/102	2,9	0.186	22/302	7,3	0.371
	4	15/99	15,2	0.573	2/82	2,4	0.158	10/111	9,0	0.426

**Tabela 7:** Cruzamento das variáveis sociais escolaridade, faixa etária e gênero/sexo em relação à aplicação da variante “Formas Nominais”.

Ao analisar os resultados apresentados nessa tabela, percebe-se que não se trata de variação influenciada pela escolaridade ou faixa etária, uma vez que os maiores pesos relativos se encontram no gênero/sexo masculino, quase que independente dos outros fatores.

**Tipo de oração** - a sexta variável selecionada, com maior peso relativo para a oração absoluta, seguida da subordinada, com pesos relativos de 0.639 e 0.633 nessa ordem.

<b>Tipo de oração</b> Significância = 0.001	<b>Apl./T</b>	<b>%</b>	<b>P.R.</b>
<b>Absoluta</b>	7/35	20,0	0.639
<b>Subordinada</b>	85/381	22,3	0.633
<b>Coordenada</b>	211/1835	11,5	0.489
<b>Principal</b>	35/344	10,2	0.395
<b>TOTAL</b>	<b>338/2595</b>		

**Tabela 8:** Aplicação da variante “Formas Nominais” em relação à variável tipo de oração.

**Tipo de verbo** - neste último grupo de fatores selecionado pela rodada binária com fator de aplicação “formas nominais”, o verbo de ligação mostrou-se como principal favorecedor no emprego dessa variante, cujo peso relativo é de 0.693, bem distante dos demais tipos de verbo, conforme pode ser verificado na tabela 9.

<b>Tipo de verbo</b> Significância = 0.001	<b>Apl./T</b>	<b>%</b>	<b>P.R.</b>
<b>Ligação</b>	40/129	16,9	0.693
<b>Intransitivo</b>	19/187	10,2	0.490
<b>Transitivo</b>	279/2239	12,5	0.485
<b>TOTAL</b>	<b>338/2595</b>		

**Tabela 9:** Aplicação da variante “Formas Nominais” em relação à variável tipo de verbo.

Ao se deter no estudo das formas nominais encontradas nos *corpora* observados, formados por falantes de níveis de escolaridade, idade e gênero/sexo distintos, houve necessidade de buscar entender como essas palavras passaram a fazer parte do universo da indeterminação do sujeito.

Assim, partindo da noção de gramaticalização e de que a “lexicologia é a ciência que estuda o léxico em todas as suas relações linguísticas, pragmáticas, discursivas, históricas e culturais” (ABBADÉ, 2006, p. 219), serão vislumbradas aqui algumas formas nominais que mais ocorreram, a fim de ilustrar como as palavras ganham novos conceitos e assumem novas funções gramaticais, cujas significações as gramáticas tradicionais não atualizam e não consideram as novas abordagens de usos da língua.

Para buscar atingir esse objetivo, foram utilizados quatro dicionários da língua portuguesa: Machado (1967), Cunha (2007), Houaiss (2001) e Ferreira (2004).

Cabe ainda ressaltar que as formas nominais estudadas foram encontradas com artigo, como foi o caso de “o pessoal”, ora sem artigo, como “nego”, “neguinho”, uma vez que a literatura sociolinguística já confirmou que é muito comum a não colocação do artigo diante de substantivos, de maneira geral, no Nordeste brasileiro.

### 3.2 As formas nominais

A seguir, são apresentadas algumas estratégias que constituem, de fato, o que se chama aqui de formas nominais:

**O cara** - a palavra de origem grega “kára” e latim tardio “cara” tem como significado comum o mesmo sentido atribuído a “rosto”, “face”, segundo Machado (1967, p. 537), Ferreira (2004, p. 398), Houaiss (2001) e Cunha (2007).

A ideia de seu emprego como sujeito indeterminado já é encontrado nos dicionários de Ferreira (2004, p. 398), especificando que se trata de uma “pessoa que não se conhece”, sinônimo de “indivíduo”, “sujeito”, também uma das formas analisadas, e do Houaiss (2001), atribuindo-lhe o sentido de “indivíduo qualquer; sujeito, pessoa”.

Os falantes de Salvador utilizam sempre a expressão “o cara”, com artigo, para marcar a indeterminação do referente do sujeito, sendo qualquer pessoa capaz de realizar ou sofrer qualquer ação.

- (14) Eu fiquei admirado outro dia perguntando aí na, na Independência, o que é Independência do Brasil, **o cara** não sabia o que era... [PEPP06-M4f]

**O indivíduo** - a palavra “indivíduo”, do latim “individuu” (FERREIRA, 2004, p. 1097) apresenta-se nos dicionários consultados, além de outras significações, como “exemplar de uma espécie qualquer”, “ser humano único considerado isoladamente na coletividade”. Comum também foi seu significado indeterminado (objeto do presente estudo), atribuindo-se este nome a “uma pessoa qualquer, cujo nome não se quer dizer” (FERREIRA, 2004, p. 1097) ou, como informou o Houaiss (2001), “homem anônimo, indeterminado”.

- (15) E **o indivíduo** tinha o direito mas aí pagava por fora o médico né, e nisso os médicos, muitos médicos ganharam muito dinheiro [...] [NURC013/R-F4s]

**O sujeito** - Ferreira (2004), no item sete, apresenta essa palavra como “indivíduo indeterminado, ou cujo nome se quer omitir”, justamente a mesma ideia apresentada pelas gramáticas tradicionais apresentadas anteriormente. Compartilha também dessa opinião o Houaiss (2001), dizendo que sujeito é a “pessoa indeterminada ou cujo nome não se enuncia”, mantendo, portanto, a ideia defendida aqui, como algo que está presente na vida social dos falantes, e não algo representativamente sintático.

- (16) Amigo a Prova de Bala, porque ele, com aquele, aquele fogo cruzado todo ele ainda tentou eh, me tirar do, mas não podia, mas aí ele ainda continu, conseguiu, eh, entrar no bonde já parado, quando **o sujeito** estava fugindo e o motorneiro alucinado aí ele me tirou botou num táxi e me levou em casa, esse foi um, foi um acidente interessante no bonde, bom, e aí era muito tranquilo porque o Farol da Barra era um “footing” contínuo NE [...] [NURC013/R-F4s]

**Nego** - a palavra “nego” é originária de “negro”, como redução fonética negro > nego. Sendo negro originário do latim *niger, nira, nigrum*; “diz-se de ou indivíduo de etnia negra” (CUNHA, 2007). Embora faça referência à etnia negra, seu conteúdo semântico se expandiu em caráter emocional na Bahia, mais especificamente em Salvador, sendo empregado também para qualquer pessoa a que se tenha estima ou então para falar de qualquer pessoa, não determinando quem seja.

- (17) Onde ela mora hoje e ficou um tempão lá de ... **nego** às vezes fazia de lixo aí , mas chegou um tempo que ela tomou a tenência de fazer quarto sala e cozinha, hoje em dia ela fez. [PEPP19-F2f]

- (18) Porque não tá como, o dia de hoje, hoje não, tá tudo uma violência danada aí, um de, um descontrole danado, um desrespeito, antigamente não era tanto assim, era, mas não era tanto, hoje você não pode pisar num pé, você dentro de um ônibus pisar num pé aí, **nego** já está querendo puxar a arma pra, pra querer,

querer, tá querendo chamar pra mão né, já está querendo tirar a vida da pessoa. [PEPP30-M3f]

**O pessoal, a(s) pessoa(s)** - a palavra “pessoa” vem do latim, “persona”, significando o papel desempenhado pelo ator, personagem (MACHADO, 1967, p. 1808; HOUAISS, 2001; CUNHA, 2007). As formas como aparecem nos dados coletados não correspondem exatamente ao uso “pessoa”, mas sempre com emprego do artigo, “a pessoa”, “as pessoas” ou “o pessoal”. No singular, o usuário demonstra ter em mente quem é o referente, mas ao utilizar “as pessoas” ou “o pessoal”, expressão que está no singular, mas que denota ideia plural, ele generaliza, atribuindo a ação verbal a qualquer pessoa, fazendo do referente um ser completamente genérico.

- (19) As vezes as, **o pessoal** briga porque gosta e as vezes é porque acontece, tem uns mesmo que vão pro pagode ficam bêbado ou drogado e aí já vai pra, procurar briga com os outros, tem gente que tem um temperamento muito alto não aguenta aí começa, aí começando uma briga, aí vem várias. [PEPP18-M1f]
- (20) Bom, eu acho que eu aprendi ler, assim, suficiente, escrever o, o necessário, e minhas professoras graças a Deus eu me dei bem com todas elas, com a diretora e tudo, eu me dava muito bem com as pessoas, me dava não, me dou, é uma coisa que, quando **a pessoa** toma antipatia por mim também não tem pra onde correr, mas quando simpatiza, geralmente eu consigo a simpatia das pessoas. [PEPP09-M2f]
- (21) **As pessoas** ainda sentam na porta para conversar, eu acho isso fantástico, é uma das coisas assim, porque eu não quero sair da Cidade Baixa, eu acho assim um clima que é completamente diferente do resto da cidade. [NURC010/N-M2s]

**O povo** - Houaiss (2001) no item 14 apresenta a palavra “povo” como sinônimo de “turma” e “gente. Já no item 16, ele diz que se trata também de “os seres humanos”, “a humanidade”. Percebe-se nessas acepções que a expressão “o povo” também é considerada uma forma genérica, que pode ser atribuída aos falantes de uma forma geral, sem precisar especificá-los, portanto, uma forma indeterminadora.

- (22) Problema todo ...Inint... botei em cima governo, porque o governo não, está todo mundo revoltado né, todo mundo revoltado, a maior parte de gente, do povo tá revoltado, a maior parte do povo né, tá revoltado, não tem emprego, não tem nada, não tem, não tem salário digno, não tem educação, não tem a saúde, não tem nada, **o povo** só tá vivendo assim mesmo, revoltado, qualquer coisa se revolta. [PEPP30-M3f]

**O público** - inicialmente, Houaiss (2001), trata a palavra “público” como um adjetivo referente a um povo, a coletividade, passando a informar depois outras concepções, como algo que é destinado às pessoas de maneira geral, algo que pode ser publicado, sinônimo também de platéia entre outras. O lexicógrafo também coloca a concepção que se tinha de “público” em 1720, como sendo “o homem comum, do povo”. Talvez seja essa a ideia que caracteriza essa forma como um item indeterminador do sujeito.

- (23) Hoje já não é mais realidade, minha mulher, por exemplo, sai às 8 da manhã, 9, às vezes um pouco mais tarde, porque ela, como é arquiteta, eh... é trabalha com **o público** também, que acorda mais tarde, não é? [NURC010/N-M2s]

**O homem** – segundo o conceito para “homem”, de acordo com Houaiss (2001), é de “espécie humana, humanidade”, ou seja, um caráter bastante generalizador, como se confirma no item 15, ao dizer que se trata de um “diacronismo” para designar “qualquer pessoa”, portanto, o sujeito indeterminado, essa qualquer pessoa que inventa algumas coisas, conforme o exemplo 91.

- (24) De... de... uma série de questão. Foi... tudo isso está aí, também. **O homem** inventa algumas coisas que ele não sabe até onde vai dar, quais são os reflexos. [NURC010/N-M2s]

**Gente** – a palavra “gente” foi usada isoladamente sem o artigo nos *corpora* estudados, não se confundindo com a forma pronominal “a gente”. Primeiramente, Houaiss (2001) informa que “gente” tem a mesma ideia de “multidão de pessoas, povo”. No item 4, ele especifica mais ainda, dizendo que se trata de um “número indeterminado de pessoas”.

- (25) Camiseta, é, malha de algodão, mas geralmente o que vai é isso, é a calça e a blusa solta por cima. Tem **gente** que vai já de short, direto, porque, hoje em dia, faz, e como é ginástica, não é aula de dança, o pessoal faz muito aula de short, barriguinha de fora, então você já vai com aquela roupa de lycra, bota uma blusa em cima, e já vai pronta assim. E anda na rua, assim mesmo. [NURC0011/R-F3s]

#### 4 Consideramos finalmente

Conforme pôde ser verificado ao longo deste trabalho, as diferentes estratégias utilizadas pelos falantes de Salvador para marcar a indeterminação do sujeito em seus momentos de fala extrapolam os ditames das gramáticas tradicionais, revelando que elas se encontram em constante desatualizam, uma vez que não conseguem acompanhar a evolução linguística imposta pelos usuários da língua portuguesa no Brasil, especialmente na Bahia.

As formas nominais formam um grupo de variantes bastante diversificado, de inesgotável valores culturais e sociais, que marcam a história de um pouco e seus costumes, como é o caso de “nego” e “neguinho” para os falantes soteropolitanos, que além de fazerem uso para tratar os seus interlocutores, recorrem também para as marcas de indeterminação, comparando-se com as estratégias canônicas das GT, talvez até em maior evidência, tendo em vista o seu uso abundante.

Essa estratégia mostrou-se favorável quando se trata da variável forma antecedente, ou seja, uma vez utilizada, o falante busca mantê-la no discurso, a fim de dar continuidade à indeterminação do sujeito.

Em relação às variáveis extralinguísticas, não houve evidências de que influenciariam muito nas escolhas pelos usuários da língua. A diferença apresentada pela única variável selecionada não foi grande, pois os pesos relativos para os gêneros masculino e feminino estão próximos.

Por fim, observou-se que os diversos usos que os falantes fazem das palavras, modificando os conceitos originais e atribuindo novos significados, fizeram com que as estratégias estudadas se ampliassem, não se limitando à apenas as formas de indeterminação registradas nas GT.



## Referências

- ABBADE, Celina Márcia de Souza. O estudo do léxico. In: TEIXEIRA, Maria da Conceição R.; QUEIROZ, Rita de Cássia R. de; SANTOS, Rosa Borges dos. **Diferentes perspectivas dos estudos filológicos**. Salvador: Quarteto, 2006. p.213-225.
- ALMEIDA, Napoleão Mendes de. **Gramática metódica da língua portuguesa**. 45.ed. 2.tir. São Paulo: Saraiva, 2005.
- AZEREDO, José Carlos de. **Gramática Houaiss da língua portuguesa**. São Paulo: Publifolha, 2008.
- CARVALHO, Valter de. Neguinho dizia assim que é um estudo sociolinguístico sobre a indeterminação do sujeito em Salvador. In: CONGRESSO INTERNACIONAL DA ABRALIN, 6., 2009, João Pessoa. **Anais...** João Pessoa: Ideia, 2009a. v. 2, p. 4440 - 4445. Disponível em: <<http://www.abralin.org/publicacao/ABRALIN2009VOL2.zip>>. Acesso em: 21 fev. 2010.
- \_\_\_\_\_. O sujeito indeterminado rodando por Salvador. In: IV SEMINÁRIO DE ESTUDOS FILOLÓGICOS, 4., 2009, Salvador. **Anais...** Salvador: UCSAL, 2009b. p. 908 - 914. CD-ROM.
- CAVALCANTE, Silvia Regina de Oliveira. **A indeterminação do sujeito na escrita padrão: a imprensa carioca dos séculos XIX e XX**. 1999. 177 f. Dissertação (Mestrado). Departamento de Letras Vernáculas da Faculdade de Letras, Universidade Federal do Rio de Janeiro, 1999.
- CEGALLA, Domingos Paschoal. **Novíssima gramática da língua portuguesa**. 48.ed. São Paulo: Companhia Editora Nacional, 2008.
- CUNHA, Antônio Geraldo da. **Dicionário etimológico da língua portuguesa**. 3.ed. 2.impr. Rio de Janeiro: Lexikon, 2007.
- CUNHA, Cláudia de Souza. **Indeterminação pronominal do sujeito**. 1993. 79 f. Dissertação (Mestrado) - Pós-Graduação da Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1993.
- FERREIRA, Aurélio Buarque de. **Novo dicionário Aurélio da língua portuguesa**. 3.ed. rev. e atual. Curitiba: Positivo, 2004.
- GODOY, Maria Alice Maschio. **A indeterminação do sujeito no interior paranaense: uma abordagem sociolinguística**. 1999. 128 f. Dissertação (Mestrado em Letras e Linguística), Universidade Federal do Paraná, Paraná, 1999.
- GUY, Gregory R.; ZILLES, Ana. **Sociolinguística quantitativa: instrumental de análise**. São Paulo: Parábola Editorial, 2007.
- HAUY, Amini Boainain. **Da necessidade de uma gramática-padrão da língua portuguesa**. 2.ed. São Paulo: Ática, 1986.
- HOUAISS, Antônio et al. **Dicionário eletrônico Houaiss de língua portuguesa**. Versão 1.0. [S.I.]: Objetiva, 2001. CD ROM.
- LABOV, William. **Principles of linguistic change**. Oxford/Cambridge: Blackwell, 1994.
- LOPES, Norma da S. O PEPP: histórico e caracterização. In: LOPES, Norma da S.; SOUZA, Constância Maria B. de; SOUZA, Emília Helena P. M. de (Org.). **Um estudo da fala popular de Salvador – PEPP**. Salvador: Quarteto, 2009.
- MACHADO, José Pedro. **Dicionário etimológico da língua portuguesa**. 2.ed. Lisboa: Confluência, 1967.
- MENON, Odete Pereira da Silva. **Analyse sociolinguistique de l'indetermination du sujet dans le portugais parle au Bresil, a partir des donnees du NURC/SP**. 1994. 397 f. Tese (Doutorado) - Département de Recherches Linguistiques, Université de Paris VII, Paris, 1994.

- \_\_\_\_\_. A indeterminação do sujeito no português do Brasil: NURC-SP e VARSUL. In: VANDRESEN, Paulino. **Variação, mudança e contato linguístico no Português da Região Sul**. Pelotas: Educat, 2006. Cap. 7, p. 125-167.
- MILANEZ, Wânia. **Recursos de indeterminação do sujeito**. 1982. 149 f. Dissertação (Mestrado) - Departamento de Linguística do Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 1982.
- NEVES, Maria Helena de Moura. **Gramática de usos do português**. São Paulo: Editora UNESP, 2000.
- PONTE, Vanessa. **A indeterminação do sujeito no português rural do interior do estado da Bahia**. 2008. 125 f. Dissertação (Mestrado) - Programa de Pós-Graduação em Letras e Linguística do Instituto de Letras, Universidade Federal da Bahia, Salvador, 2008.
- ROCHA LIMA, Carlos Henrique da. **Gramática normativa da língua portuguesa**. 47.ed. Rio de Janeiro: José Olympio, 2008.
- ROLLEMBERG, Vera et al. Os pronomes pessoais sujeito e a indeterminação do sujeito na norma culta de Salvador. **Estudos Linguísticos e Literários**. Salvador, n. 11, p. 53-74, 1991.
- SAID ALI, Manoel. **Investigações filológicas**. 3.ed. rev. e ampl. Estudo e organização de Evanildo Bechara. Rio de Janeiro: Lucerna, 2006[1976].
- SANTANA, Neila Maria Oliveira. **A indeterminação do sujeito no português rural do semi-árido baiano**. 2006. 153 f. Dissertação (Mestrado) - Programa de Pós-Graduação em Letras e Linguística do Instituto de Letras, Universidade Federal da Bahia, Salvador, 2006.
- SETTI, Adriane Cristina Ribas. **A indeterminação do sujeito nas três capitais do Sul do Brasil**. 1997. 116 f. Dissertação (Mestrado) - Curso de Pós-graduação em Letras, Setor de Ciências Humanas, Letras e Artes, Universidade Federal do Paraná, Curitiba, 1997.