

## USO DAS FORMAS PRONOMINAIS TU, VOCÊ E A GENTE COMO ESTRATÉGIA DE INDETERMINAÇÃO DO SUJEITO: o que mostram os dados do ALiMA\*

USAGE OF THE PRONOUN FORMS TU, VOCÊ AND A GENTE AS A STRATEGY FOR SUBJECT INDETERMINATION: what is shown by ALiMA's data

EL USO DE LAS FORMAS PRONOMINALES "TU, VOCÊ Y A GENTE" COMO ESTRATEGIA DE INDETERMINACIÓN DEL SUJETO: lo que muestran los datos del ALiMA

Alana Brito Barbosa  
Conceição de Maria de Araujo Ramos

**Resumo:** Esta pesquisa trata do uso dos pronomes *tu*, *você* e *a gente* como estratégia de indeterminação do sujeito no português falado no Maranhão. Analisou-se, com base nos princípios teórico-metodológicos da Sociolinguística Variacionista o falar de pessoas de 11 municípios maranhenses que integram a rede de pontos do Atlas Linguístico do Maranhão – ALiMA –, a fim de verificar-se quais variáveis linguísticas e sociais estariam condicionando o uso dos pronomes *tu*, *você* e *a gente* indeterminados. A amostra analisada foi obtida com base na aplicação das questões 25 e 26 do questionário morfosintático, das perguntas metalinguísticas e dos discursos semidirigidos a 48 sujeitos distribuídos igualmente entre os dois sexos/gêneros, duas faixas etárias e dois níveis de escolaridade. Os resultados evidenciaram que o fenômeno analisado é influenciado por fatores linguísticos e sociais, sendo o pronome *a gente* o mais utilizado, seguido do pronome *você* e do pronome *tu*. Com esta pesquisa, portanto, pretende-se oferecer uma reflexão e possíveis encaminhamentos ao ensino/aprendizagem da língua portuguesa, de modo a colocar à disposição dos estudiosos da língua portuguesa um suporte científico para o embasamento dos seus trabalhos.

**Palavras-chave:** Pronomes. Indeterminação do sujeito. Português falado no Maranhão.

**Abstract:** This research deals with the use of pronouns *tu*, *você* and *a gente* as a strategy of indeterminacy of the subject in Maranhão's spoken Portuguese. It was analyzed, based on theoretical and methodological principles of Variationist Sociolinguistics, the speech of people from 11 cities that comprise the network points of Atlas Linguístico do Maranhão – ALiMA – in order to check up social and linguistic variables which would be conditioning the use of the pronouns *tu*, *você* and *a gente* in an undetermined manner. The sample was obtained based on the application of the questions 25 and 26 of the morphosyntactic questionnaire, metalinguistic questions and semi-directed speeches to 48 subjects equally distributed between the two sexes/genders, two age ranges and two levels of schooling. The results showed that the phenomenon analyzed is influenced by social and linguistic factors, being *a gente* the most used pronoun, followed by the pronoun *você* and *tu*. With this research, therefore, it is aimed to provide a discussion and possible referrals to the teaching/learning of the Portuguese language, in order to make available to students a scientific support to the foundation of their researches.

**Keywords:** Pronouns. Subject indetermination. The Portuguese language spoken in Maranhão State.

**Resumen:** Esta investigación trata del uso de los pronombres *tu*, *você* y *a gente* como estrategia de indeterminación del sujeto en el portugués hablado en Maranhão. Se analizó, con base en los principios teórico-metodológicos de la Sociolingüística Variacionista el habla de personas de 11 municipios marañenses que integran la red de puntos del Atlas Lingüístico del Maranhão – ALiMA – con la finalidad de verificarse cuales de las variables lingüísticas y sociales estarían condicionando el uso de los pronombres *tu*, *você* y *a gente*. La muestra fue obtenida con base en la aplicación de las preguntas 25 y 26 del cuestionario morfosintático, de las preguntas metalingüísticas y de los discursos semidirigidos a 48 sujetos distribuidos igualmente entre los dos géneros, dos grupos de edad y dos niveles de escolaridad. Los resultados evidenciaron que el fenómeno analizado es influenciado por factores lingüísticos y sociales, siendo el pronombre *a gente* el más utilizado, después el pronombre *você* y, por último, el pronombre *tu*. Con esta investigación, por lo tanto, se pretende ofrecer una reflexión y posibles encaminamientos a la enseñanza/aprendizaje de la lengua portuguesa, de modo de poner a disposición de los estudiosos de la lengua portuguesa, un soporte científico para el embasamiento de sus trabajos.

**Palabras clave:** Pronombres. Indeterminación del sujeto. Portugués hablado en Maranhão.

Trabalho premiado durante o XXIV Encontro do SEMIC, realizado na UFMA entre os dias 05 a 08 de novembro de 2012.

\*Artigo recebido em dezembro 2012

Aprovado em fevereiro 2013

## 1 INTRODUÇÃO

Com o objetivo de contribuir para o entendimento e a descrição do português brasileiro (PB) e, de maneira mais específica, para a descrição do português falado no Maranhão, neste trabalho, investigamos o uso dos pronomes *tu*, *você* e *a gente* como estratégias não-canônicas<sup>2</sup> de indeterminação do sujeito.

Para a análise, partimos dos pressupostos teórico-metodológicos da Sociolinguística Variacionista e da Geolinguística – que abordam a língua nos seus contextos reais de uso e de acordo com a sua distribuição espacial – de modo a observar em que medida os fatores linguísticos e sociais influenciam as escolhas dos falantes por uma forma em detrimento de outras nas situações comunicativas.

Segundo a Sociolinguística, “é comum que uma língua tenha diversas maneiras alternativas de se dizer a ‘mesma’ coisa” (LABOV, 2008, p. 221). Essas maneiras alternativas, denominadas de variantes linguísticas, são constitutivas da *variação linguística* e evidenciadoras da natureza social da língua.

O Brasil, país que apresenta uma grande extensão territorial e um vasto número de centros urbanos e aglomerados populacionais, retrata bem esse panorama. Há, na realidade, uma grande variação no âmbito linguístico, uma vez que as línguas são “[...] também inevitavelmente voltadas para a mudança, pois os grupos humanos são dinâmicos, e as línguas que eles falam precisam adaptar-se às novas situações históricas” (CASTILHO, 2010, p. 197).

Levando em consideração o caráter heterogêneo da língua, atinamos para a questão de como a Gramática Tradicional (GT) tem tratado as variações e mudanças pelas quais o PB tem passado. Destacamos, por exemplo, a realidade pronominal que, apesar de ser um assunto muito presente na realidade linguística dos falantes, tem sido abordada de forma parcial por grande parte dos autores de gramáticas, já que não incluem, no subsistema dos pronomes pessoais, as formas mais inovadoras *você* e *a gente*, ainda que seja notório o perfeito enquadramento desses pronomes nesse subsistema.

Outro ponto que também é tratado de forma parcial pela GT é a indeterminação do sujeito. A GT postula que a língua portuguesa dispõe de apenas duas estratégias de indeterminação do sujeito: o uso da terceira pessoa do plural com o pronome oculto (V3PP) e o uso da terceira pessoa do singular com o pronome *se* (V3PS+SE), não englobando, desse modo, a totalidade das estratégias que se fazem presentes no uso real da língua.

Muitas pesquisas têm-se encarregado do estudo da indeterminação do sujeito em diferentes regiões brasileiras e os resultados têm mostrado a utilização de outras estratégias além daquelas prescritas pela GT. Godoy (1999), com base nos dados do Projeto Varsul, investigou esse fenômeno no falar de três municípios paranaenses – Irati, Londrina e Pato

Branco – e registrou o uso das seguintes formas de indeterminação: *a gente*, *ele(s)*, *eu*, *formas nominais*, *nós*, *se*, *tu*, *você(s)*, *voz passiva sem agente*, *voz passiva sintética*, *V3PS+SE* e *V3PP+SE*. Setti (1997), tendo também como base os dados do Projeto Varsul, observou o fenômeno da indeterminação do sujeito nas três capitais do sul do Brasil e apontou a utilização de 12 estratégias: *a gente*, *eles*, *eu*, *formas nominais*, *nós*, *se*, *tu*, *você(s)*, *voz passiva sem agente*, *voz passiva sintética*, *V3PS+SE* e *V3PP+SE*.

Ao tratarmos dos pronomes *tu*, *você* e *a gente* como recursos de indeterminação – não prescritos pela GT –, contribuimos também para a desmistificação do preconceito linguístico que envolve, sobretudo, a língua falada, e oferecemos subsídios ao ensino/aprendizagem da língua portuguesa e a uma melhor interpretação do português do Brasil, de modo a beneficiar os estudiosos da língua portuguesa, os autores de livros didáticos e os alunos.

Para alcançarmos nossos objetivos, estruturamos este estudo da seguinte forma: no tópico 2, enfocamos metodologia adotada; no tópico 3, apresentamos uma síntese das distintas visões acerca da indeterminação do sujeito, dando ênfase aos conceitos e estratégias abordados pelos estudiosos do assunto; no tópico 4, apresentamos os resultados e as análises; no tópico 5, as considerações finais.

## 2. METODOLOGIA

Para a realização desta pesquisa seguimos as seguintes etapas:

- a) pesquisa bibliográfica no âmbito da geo-sociolinguística, com enfoque nas questões morfosintáticas e semânticas;
- b) delimitação e seleção do *corpus*;
- c) transcrição e análise estatística dos dados por meio do uso do programa estatístico *Varbrul*<sup>3</sup>.

Como esta pesquisa se insere em um projeto mais amplo – o ALiMA – os procedimentos metodológicos seguiram a orientação do projeto do qual esta pesquisa é subsidiária. Os dados utilizados foram selecionados do banco de dados do ALiMA. Os pontos linguísticos foram escolhidos de modo a contemplar as cinco mesorregiões do Estado e a possibilitar-nos uma visão mais ampla acerca do fenômeno em análise. Selecionamos os seguintes municípios: Alto Parnaíba<sup>4</sup>, Araioses, Bacabal, Balsas, Carolina, Imperatriz, Pinheiro, São João dos Patos, São Luís, Tuntum e Turiangu.

Utilizamos, pois, para a constituição do *corpus*, os inquiridos de 48 informantes (sendo quatro informantes por localidade, exceto na capital do Estado, São Luís, em que foram entrevistados oito). Os informantes, de ambos os sexos/gêneros, foram distribuídos, igualmente, em duas faixas etárias: 18 a 30 anos (faixa

etária I) e 50 a 65 anos (faixa etária II); e em dois níveis de escolaridade: Ensino Fundamental (no máximo, até o 5º ano) e Ensino Superior completo (apenas em São Luís, onde esse nível de ensino já faz parte da realidade sociocultural do município há bastante tempo, o que permite considerá-lo como um dos traços característicos do perfil do informante).

Em nossa análise foram considerados tanto fatores sociais como fatores linguísticos, pois, como postula Labov (2008), as variáveis linguísticas e sociais exercem influência na escolha que os falantes fazem das variáveis dependentes (neste estudo, os pronomes *tu*, *você a gente*). Os fatores sociais selecionados foram: *sexo/gênero, faixa etária, escolaridade e localidade*; e os fatores linguísticos: *tempo e modo verbal, tipo de verbo, tipo de oração, paralelismo formal e grau de indeterminação do sujeito*.

### 3. O QUE SE FALA SOBRE A INDETERMINAÇÃO DO SUJEITO

#### 3.1 Visão tradicional

No âmbito das obras didáticas adotadas na educação básica, a *Gramática sem Segredos*, de Menón (2008, p. 379), defende que o sujeito é indeterminado quando “não identificável porque não se sabe qual é ou não se quer identificá-lo”, havendo, pois, dois recursos de que a língua portuguesa se vale para indeterminar o sujeito: V3PP, sem fazer qualquer referência a um elemento que possa ocupar tal posição, e V3PS+SE. A autora atenta, ainda, para o caráter generalizador, indefinido do referente, já que não se sabe quem é esse referente e ressalta que o conhecimento compartilhado pelos interlocutores possibilita, em determinados casos, a identificação do sujeito.

Paschoalin (2008, p. 241), em *Gramática: teoria e exercícios*, afirma que “o sujeito da oração é indeterminado quando sua existência é evidente, mas não há nenhum termo que o represente, nem mesmo em orações anteriores.” Seguindo a mesma linha de Menón (2008), a autora afirma que há dois motivos pelos quais o falante indetermina o sujeito: por não haver a intenção de determiná-lo ou por desconhece-lo, e aponta duas estratégias: V3PP, sem que seja feita referência a algum termo utilizado anteriormente, e V3PS+SE.

Abaurre, Abourre e Pontara (2008, p. 495) defendem que “somente os casos de indeterminação sintática (impossibilidade de identificar o termo a que se refere o verbo) é que definem **sujeitos indeterminados**”. Apontam, também, apenas dois recursos, delimitando, no entanto, a transitividade verbal inerente a cada um deles: verbo transitivo direto na 3ª pessoa do plural e verbo transitivo indireto, intransitivo ou de ligação na 3ª pessoa do singular acompanhado do pronome *se*. Vejamos os exemplos apresentados pelas autoras:

- (1) Incendiaram vários ônibus.
- (2) Precisa-se de vendedores.
- (3) Come-se bem na Itália.
- (4) Aqui se está feliz.

Conforme pudemos observar, há uma lacuna na abordagem feita pela GT quando o assunto é sujeito indeterminado, mesmo que esta noção seja, como salienta Castilho (2010, p. 297), “a propriedade semântica mais explorada na Gramática Tradicional”.

Vimos que, apesar de a GT tratar da indeterminação como uma questão semântica, prescreve apenas duas estratégias. Se há uma conceituação semântica, por que não incluir outras estratégias, já que estas podem representar, adequadamente, o sujeito indeterminado? Vejamos o exemplo a seguir:

- (5) “[...] porque *você* gostava muito de uma pessoa e a pessoa termina gostando de ôtro e isso se chama traição e *você* nunca esquece (...) *você* dorme, *você* acorda, toda vida *você* lembra, claro que antes *você* lembrava, doía mais porque *você* tava recente [...]” (MA14/01<sup>5</sup>)

A GT, supomos, diria que esse *você* é um sujeito simples e determinado. Contudo, como classificá-lo como tal se, apesar de ser um pronome de segunda pessoa, não está fazendo referência à segunda pessoa e sim a *qualquer pessoa*? O contexto não nos permite dizer ‘quem gostava muito de uma pessoa’, ‘quem nunca esquece’, ‘quem dorme e acorda sempre lembrando’. Há, portanto, a possibilidade de inclusão de *qualquer pessoa* nesse *você*. Trata-se, claramente, de um caso de sujeito indeterminado, visto que não há a possibilidade de resgatar a sua referência.

Menón (2008), ao tratar do sujeito indeterminado como sendo aquele não pode ser identificado por não se saber qual é ou por não se querer identificá-lo, trata-o sob um viés semântico, dando-nos a entender que a indeterminação do sujeito pode ser realizada por meio de todas as outras estratégias que o português brasileiro põe à disposição dos falantes – além das que a GT prescreve –, uma vez que as estratégias não-canônicas são igualmente legítimas para expressar um referente indeterminado, independentemente do motivo. Entretanto, a autora apresenta, em sua obra, apenas as duas estratégias canônicas.

Paschoalin (2008), embora também enfoque o sujeito indeterminado sob o viés semântico afirma, como vimos, que há evidências da existência do sujeito, mas que “não há termo que o represente”. Quando utiliza a palavra ‘termo’, desconsidera todas as estratégias lexicalmente preenchidas havendo, portanto, um conflito entre semântica e sintaxe. A respeito da posição de Paschoalin e Menón, fazemos a seguinte pergunta: por que não classificar como sujeito indeterminado os

casos que apresentam o sujeito lexicalmente preenchido se eles – assim como as estratégias canônicas – podem, em dados contextos, assumir um caráter genérico?

Abaurre, Abourre e Pontara (2008), por sua vez, explicitam que a questão da indeterminação dá-se no nível sintático, desconsiderando, assim, o nível em que se encontra a referência do sujeito – o semântico. De acordo com essa posição, as autoras classificariam, por exemplo, o sujeito do exemplo (05) como sendo simples e determinado.

Milanez (1982), ao examinar as abordagens do sujeito indeterminado propostas pelas autoras das GTs supracitadas, chama a atenção para o fato de que esses enfoques não seguem um critério de classificação, pois ora *situam* o sujeito no campo semântico, ora o abordam como uma questão pertinente ao campo morfossintático. Além disso,

[...] não se explicita o que seria, primeiramente, um sujeito 'determinado'. Ao contrário, o conceito de indeterminação já aparece diretamente, por ocasião da Classificação do Sujeito, misturado às classificações de *claro, oculto, simples, composto*, que são simplesmente formais, enquanto a de 'indeterminado' é semântica. (MILANEZ, 1982, p. 11).

A esse respeito, Menón (2008), apesar de tratar do sujeito indeterminado em oposição ao determinado, fá-lo de maneira superficial. *Postula*, sobre o sujeito determinado, que "o núcleo (ou núcleos) está expresso ou é identificável por meio da desinência verbal" (p. 379), não partindo, portanto, da ideia de determinação do *referente*, ou seja, do agente da ação verbal para, então, tratar da indeterminação do referente.

### 3.2 Visão não-tradicional

Para que a análise do fenômeno da indeterminação do sujeito seja feita de forma sistemática e aprofundada devemos esclarecer os conceitos de determinação, indefinição e impessoalização. Vejamos, primeiramente, as características que fazem um sujeito ser classificado como determinado ou indeterminado. Segundo Milanez (1982), quando analisamos a dicotomia determinação/indeterminação,

Estamos, portanto, diante de um processo da língua que permite ao falante passar do universo das três pessoas especificadas e identificáveis (que consideraremos com o nível da *determinação*) a um nível de generalização, que transcende o anterior por implicar numa referência de tal forma abrangente que pode envolver *qualquer* pessoa. É o fenômeno da *indeterminação*. (MILANEZ, 1982, p. 26).

Com base nisso, veremos, nos exemplos elencados a seguir, que o sujeito se enquadra no campo da determinação, sendo claramente possível observar a união dos aspectos sintático, semântico e pragmático e a pontual identificação das pessoas do discurso:

- (6) *Tu* és uma pessoa muito educada.  
(2ª pessoa)

- (7) *Maria* sempre diz que a vida é curta.  
(3ª pessoa)  
(8) *Eu e mamãe* gostamos de viajar.  
(1ª pessoa e 3ª pessoa)

A determinação do sujeito não é, portanto, uma questão puramente sintática, como a tradição gramatical costuma postular. Há, na realidade, a identificação da *referência* do sujeito e, por sua vez, apenas uma possibilidade interpretativa. Por outro lado, há casos em que não há a possibilidade de recuperação pontual da referência do sujeito mesmo que este esteja preenchido lexicalmente – é o caso dos sujeitos indeterminados. Veremos a seguir três exemplos extraídos de nosso *corpus* que contemplam os pronomes *a gente*, *tu* e *você* com referência genérica e veremos que esses pronomes fazem referência a qualquer pessoa:

- (9) INF. - Basicamente isso, eu gosto de televisão, assistir e... de, de, de telejornal, sou uma pessoa que me considero atualizado. Hoje **pra gente** caminhá, *a gente* tem que corrê, né? Se *você* pará, fica pra trás. No mínimo, tem a informação é fundamental". (MA 26/07)

- (10) "[...] eu me formei, tô a esse um ano estudando, só que eu já estô... tô com minha carteira do conselho do contadô e, de vez em quando, eu faço uma perícia contábil, judicial ou extrajudicial. O juiz me nomeia perito... por exemplo, *tu* entras com uma ação contra ela e, aí, *tu* alegas um valô e ela... ela alega que vai recebê um valô e ela alega que... que ela tem que te pagá outro menô. [...];" (MA 26/05)

- (11) [...]. Hoje não, *você* nasce, *você* já tá dento d'ua cultura diferente, já tem alguém ensinado ãa cultura diferente pra *você*. Então *você* fala errado muitas vezes porque qué, mas que *você* sabe realmente o significado da palavra. (MA14/01)

Cunha (1993) evidencia que, no âmbito da indeterminação, há três níveis de indeterminação do referente: indeterminação parcial com referência explícita, indeterminação parcial com referência implícita e indeterminação completa.

A indeterminação parcial com referência explícita ocorre quando os pronomes mantêm uma relação com elementos lexicais *explícitos* no texto, não deixando de apresentar um caráter genérico, uma vez que não podemos, com exatidão, precisar a referência. A indeterminação parcial com referência implícita<sup>6</sup> ocorre quando os pronomes mantêm uma relação não com elementos lexicais, mas com o próprio contexto. Assim, mesmo que não haja palavras ou termos que resgatem a referência do pronome, podemos, por meio de inferência,



depreendê-la. Já na indeterminação completa, os pronomes “não se reportam diretamente aos participantes da situação comunicativa” (CUNHA, 1993, p.30), não se relacionam com elementos explícitos ou implícitos no contexto, de modo que não há a possibilidade de resgatar a referência do sujeito no contexto.

Assim, “quanto mais a interpretação do pronome depende de elementos contextuais (o texto, a situação comunicativa, o tipo de contrato estabelecido entre os participantes da conversa) mais próximo ele está de ser indeterminado” (CUNHA, 1993, p. 33).

Outro conceito que precisa ser esclarecido para que não haja imprecisão na classificação do sujeito indeterminado, é o de *indefinição*. É certo que há um caráter impreciso em ambos, no entanto, “a indefinição não está sujeita, como a indeterminação, à restrição de só abranger elementos humanos, podendo referir-se também aos não-humanos.” (MILANEZ, 1982, p. 38).

A indefinição abrange apenas itens lexicais de 3ª pessoa, como por exemplo, *tudo*, *nada*, *alguém*. Já a indeterminação do sujeito pode apresentar recursos que, ainda que sintaticamente, se remontem às três pessoas gramaticais, conservam um grau de generalização, de modo que sua referência não pode ser atribuída a uma pessoa específica.

No que diz respeito à *generalização*, na indeterminação ela é um fator *primordial*, enquanto na indefinição trata-se apenas de uma *possibilidade*. Como salienta Milanez (1982, p 38-39), “os indefinidos podem remeter tanto a uma referência *individualizada*, como em ‘Alguém me chama’, como a uma referência *generalizada* ‘Todos estiveram aqui’. Percebemos que há uma generalização, porém, limitada. Concluimos, então, que “os recursos de indefinição *sempre* pressupõem um conjunto ‘fechado’ de elementos que pode ser expresso ou na sua totalidade (através das ocorrências de *todos*, *tudo*), ou no seu esvaziamento (através dos indefinidos *alguns*, *uns*, etc.)” (MILANEZ, 1982, p.39). Vejamos os seguintes exemplos:

- (12) Na livraria Cultura lê-se livros interessantes.
- (13) Na livraria cultura *todos* leem livros interessantes.
- (14) Na livraria Cultura *alguém* lê livros interessantes.
- (15) Na livraria Cultura *ninguém* lê livros interessantes.

O exemplo (12) apresenta um caráter genérico e entendemos que *qualquer pessoa* pode fazer ‘leitura de livros interessantes’. Os outros três exemplos, apesar de também possuírem um caráter genérico, apresentam certo *traço quantificador* e apenas marcas de terceira pessoa – traços da indefinição.

No que diz respeito à *impessoalização*, Milanez (1982, p. 47-48) afirma que se trata de “uma propriedade de certos verbos que,

por sua própria natureza, não atribuem a seus SN-sujeitos *nenhuma função referencial*. [...] Essa ausência de referência ao SN-sujeito diferencia as construções impessoais de todas as outras formas verbais da língua” (grifos originais). No exemplo “Enquanto *você* tá vendo televisão, *você*, tá distraído”, verificamos que nesse *você*, pode incluir-se *qualquer pessoa*. Já em “**Neva** muito em Bariloche”, não há referência a algo ou a alguma pessoa. Assim, o fato de não haver um sujeito implica que não há, também, a possibilidade de fazermos referência a ele.

Conforme vimos neste tópico, há a possibilidade de um pronome pessoal representar um sujeito indeterminado. Assim, o *você* e o *tu* podem ter a referência atribuída não somente à 2ª pessoa, mas também a qualquer pessoa. Desse modo, quando tratamos da indeterminação do referente, referimo-nos ao agente da ação verbal que está *situado fora* do contexto linguístico, de modo que esses e os outros pronomes podem assumir, em dados contextos, um valor genérico.

A respeito do caráter indeterminador dos pronomes *você* (*tu*) e *a gente*, Lopes e Rumeu (2007) afirmam que:

Além da referência definida, o uso de *você* (e também de *tu*) se expandiu para os contextos de referência indeterminada (também em concorrência com o *tu* em várias regiões do país) [...] O pronome *a gente* apresenta também um caráter indeterminador em oposição a uma nuance mais específica de *nós*. O falante se descompromete com o discurso, tornando-o mais vago e genérico, pois tal forma pode englobar as demais pessoas (“*eu + você(s) + ele(s) + todo mundo ou qualquer um*”) (LOPES; RUMEU, 2007, p. 420).

Milanez (1982, p. 81) destaca que o uso do *a gente* é diferente, por exemplo, do usos da forma *você*, pois ele inclui a primeira pessoa de modo real, enquanto o *você* envolve o ouvinte apenas hipoteticamente, “uma vez que a circunstância relatada nada tem a ver, necessariamente, com a situação pessoal do falante ou do ouvinte”.

No que tange à alternância dos pronomes de segunda pessoa do singular, os estudos têm apontado a predominância do uso do *você* como pronome de 2ª pessoa em grande parte do território brasileiro. Além disso, o pronome *você* é uma das estratégias de indeterminação do sujeito mais utilizadas pelos falantes do PB e talvez essa preferência se deva justamente ao que Setti (1997, p. 59) evidenciou com relação à utilização desse pronome. Segundo a autora, o uso de *você* “faz com que se estabeleça uma relação de maior cumplicidade entre o ouvinte e a informação transmitida”. Se tomarmos como base a bibliografia selecionada para a composição deste estudo, veremos que o pronome *tu* com referência genérica foi registrado somente nas pesquisas de Setti (1997) e de Godoy (1999).

Godoy (1999) verificou, a respeito ao uso de *tu*, que a cidade de Pato Branco, por ter sido colonizada por gaúchos e catarinenses –

povos que utilizam esse pronome em seu falar – foi a que registrou grande parte das ocorrências. Nas outras duas localidades pesquisadas – Londrina e Irati – esse pronome foi utilizado, porém em número pouco significativo. Setti (1997), por sua vez, verificou que, das três capitais do sul do Brasil, o *tu* foi encontrado apenas em Florianópolis e Porto Alegre, não havendo registro dessa forma em Curitiba – tanto com referência determinada, quanto indeterminada. Partindo dessa ideia, podemos supor que se o falante não utiliza o pronome *tu* com referência determinada, tampouco irá utilizá-lo como recurso de indeterminação do sujeito.

No que diz respeito ao português falado no Maranhão, vale destacar que Alves (2010) apontou um percentual de 61,6% de ocorrências de *você* enquanto o *tu* registrou apenas 38,4% das ocorrências, sendo a forma *você* mais recorrente inclusive com a referência genérica. No que tange à alternância nós/a gente, Ferreira (2012) apontou que o *a gente* é o pronome de 1ª pessoa do plural mais utilizado pelos falantes maranhenses.

#### 4 ANÁLISE DOS DADOS

Vimos que os pronomes *tu*, *você* e *a gente* passaram por um processo de ampliação semântica, no discurso, tendo desse modo o seu uso voltado para fazer referência a “qualquer pessoa”. Com base na discussão tecida a esse respeito, veremos, neste tópico, os resultados numéricos obtidos por meio do tratamento estatístico dos dados, o que nos possibilitará verificar em que medida os fatores sociais e linguísticos exercem influência na seleção que fazem os falantes maranhenses quando do uso dos pronomes enfocados neste estudo.

Para tanto, fizemos uma rodada geral considerando os seguintes grupos de fatores: sexo/gênero, faixa etária, localidade, grau de indeterminação do sujeito, tempo e modo verbal, tipo de verbo, tipo de oração e paralelismo formal. Dessas, foram selecionados pelo programa como relevantes o grau de indeterminação do sujeito, a localidade, a faixa etária e o sexo/gênero, sendo, portanto, um fator linguístico e três fatores sociais.

Fizemos uma rodada em separado com os dados de São Luís para verificar se a variável *escolaridade* (ensino fundamental até o 5º ano e superior completo) estaria condicionando a escolha dos falantes por um dos pronomes aqui analisados. O resultado, no entanto, apresentou *nocaute*, fato esse nos fez deixar essa análise para outro momento.

Os resultados da rodada geral foram obtidos por meio da análise de um corpus com 277 ocorrências, assim distribuídas: 199 ocorrências do pronome *a gente*, correspondendo a um percentual de 71,8%, 69 ocorrências da forma *você*, totalizando 24,9%, e 9 ocorrências da forma *tu*, correspondendo a 3,3% das realizações. Esses resultados apontam que os fa-

lantes maranhenses, dentre as três estratégias investigadas, têm utilizado com uma frequência bem maior a forma *a gente* para assumir a referência indeterminada, sendo seguida do *você* e, da forma *tu* (Tabela 1).

Nossos resultados corroboram os dados apresentados por Milanez (1982), Santana (2006) e Carvalho (2010), no que tange ao uso do *a gente*, que apontaram essa forma como a mais utilizada em São Paulo, no semi-árido baiano e em Salvador, respectivamente. Carvalho (2010) mostrou, ainda, ser a forma *você* a segunda estratégia mais utilizada em Salvador.

Tabela 1: Recursos de indeterminação levantados no corpus

| A gente | Você  | Tu   | TOTAL |
|---------|-------|------|-------|
| 199     | 69    | 9    | 277   |
| 71,8%   | 24,9% | 3,3% | 100%  |

Fonte: Elaborada pelos autores

Como podemos verificar, a tabela 1 evidencia a baixa ocorrência do *tu*, com apenas 3,3% do total das ocorrências. Considerando a baixa ocorrência e a consequente incidência de *nocautes* do pronome *tu*, não foi possível realizar uma análise estatística com essa variante e, por isso, decidimos excluí-la das rodadas seguintes por questões metodológicas, uma vez que não poderíamos obter os pesos relativos e, portanto, verificar sua relevância no que tange ao fenômeno em análise. Conforme vimos no tópico anterior, Alves (2010) apontou uma menor frequência na ocorrência do pronome *tu* em relação ao pronome *você* no falar maranhense, com 61,6% de ocorrências de *você* e 38,4% das ocorrências de *tu*.

Desse modo, o quadro geral das ocorrências passou a se configurar da seguinte forma:

Tabela 2: Frequência dos recursos de indeterminação *a gente* e *Você* levantados no corpus

| A gente | Você  | TOTAL |
|---------|-------|-------|
| 199     | 69    | 268   |
| 74,3%   | 25,7% | 100%  |

Fonte: Elaborada pelos autores

Tendo em vista a nova configuração do corpus, os grupos de fatores selecionados pelo programa, nessa rodada, como sendo relevantes foram, nesta ordem: grau de indeterminação do sujeito, localidade, faixa etária e sexo/gênero. Os outros fatores – *escolaridade*, *paralelismo formal*, *tempo* e *modo verbal*, *tipo de oração* e *tipo de verbo* – não foram considerados relevantes.

A tabela 3 apresenta a distribuição das ocorrências segundo o grau de indeterminação do sujeito – primeiro grupo de fatores selecionado. É, portanto, o que mais favorece a

variação *a gente/você* com referência genérica no PB falado no Maranhão.

Tabela 3: Distribuição dos dados segundo o grau de indeterminação

|                         |      | A gente | Você | TOTAL |
|-------------------------|------|---------|------|-------|
| Indeterminação Completa | Nº   | 62      | 63   | 125   |
|                         | %    | 49,6    | 50,4 |       |
|                         | P.R. | 0,15    | 0,85 |       |
| Indeterminação Parcial  | Nº   | 137     | 6    | 143   |
|                         | %    | 95,8    | 4,2  |       |
|                         | P.R. | 0,81    | 0,19 |       |
| Total                   | Nº   | 199     | 69   | 268   |
|                         | %    | 74,3    | 25,7 |       |

Fonte: Elaborada pelos autores

Quando os falantes fazem o uso do pronome *a gente*, utilizam-no de modo que podemos, por inferência, resgatar sua referência no contexto. Assim, seu uso é favorecido nos casos de indeterminação parcial, apresentando um peso relativo de 0,81. Já nos contextos em que a referência não pode ser depreendida – indeterminação completa – há um favorecimento da utilização da forma *você*, com o peso relativo de 0,85.

Resultados semelhantes foram encontrados por Cunha (1993) e Carvalho (2010) que verificaram que a forma *a gente* foi a mais recorrente nos contextos em que a indeterminação é parcial, com peso relativo de 0,59 e 0,58, respectivamente, enquanto a indeterminação completa mostrou-se como mais favorecedora do uso da forma *você*, com peso relativo de 0,77 e 0,64, respectivamente.

O segundo fator selecionado pelo programa foi a variável localidade:

Tabela 4: Distribuição das ocorrências de *a gente* por localidade

| LOCALIDADES        | Apl. /total | %    | PR   |
|--------------------|-------------|------|------|
| São João dos Patos | 18/19       | 94,7 | 0,92 |
| Balsas             | 9/10        | 90,0 | 0,74 |
| Pinheiro           | 39/47       | 83,0 | 0,56 |
| Imperatriz         | 18/26       | 69,2 | 0,55 |
| São Luís           | 36/53       | 67,9 | 0,47 |
| Araioses           | 6/8         | 75,0 | 0,47 |
| Bacabal            | 7/8         | 87,5 | 0,43 |
| Carolina           | 18/44       | 40,9 | 0,25 |
| Alto Parnaíba      | 6/11        | 54,5 | 0,14 |

|         |         |      |   |
|---------|---------|------|---|
| Turiagu | 31/31   | 100  | - |
| Tuntum  | 10/10   | 100  | - |
| TOTAL   | 157/226 | 69,5 | - |

Fonte: Elaborada pelos autores

Conforme aponta a tabela 4, São João dos Patos foi a localidade que mais favoreceu o uso da forma *a gente* (0,92), enquanto Alto Parnaíba foi a que mais favoreceu o uso da forma *você* (0,86). O resultado obtido em Alto Parnaíba corrobora o resultado de Alves (2010) no que concerne à preferência pela forma *você*, de modo que entendemos que, neste município, os falantes dão preferência a essa forma. Em São Luís, capital do Estado, verificamos que o peso relativo foi de 0,53 para o uso do *você*. Alves (2010) apontou que os falantes ludovicenses dão preferência ao uso da forma *você* em detrimento da forma *tu*.

Os municípios de Balsas, Pinheiro e Imperatriz indeterminam o sujeito muito mais com a forma *a gente*, enquanto os falantes de Araioses, Bacabal e Carolina favorecem o uso da segunda pessoa do singular – *você*, para a mesma situação.

Nessa rodada, obtivemos *nocautes* em duas localidades: Tuntum (10 ocorrências) e Turiagu (31 ocorrências), de modo que não houve possibilidade de verificarmos os pesos relativos correspondentes. Em ambas as cidades houve apenas a ocorrência do pronome *a gente*.

A *faixa etária* foi o terceiro fator com maior relevância. Isso evidencia que o aspecto diafórico também explica a escolha de tais variantes pelos falantes maranhenses.

Tabela 5. Distribuição das ocorrências segundo a faixa etária

|                 |      | A gente | Você | TOTAL |
|-----------------|------|---------|------|-------|
| I faixa etária  | Nº   | 101     | 26   | 127   |
|                 | %    | 79,5    | 20,5 |       |
|                 | P.R. | 0,63    | 0,37 |       |
| II faixa etária | Nº   | 98      | 43   | 141   |
|                 | %    | 69,5    | 30,5 |       |
|                 | P.R. | 0,38    | 0,62 |       |
| Total           | Nº   | 199     | 69   | 268   |
|                 | %    | 74,3    | 25,7 |       |

Fonte: Elaborada pelos autores

Ao observarmos a tabela 5, verificamos que os mais jovens favorecem o uso do *a gente* com peso relativo de 0,63, enquanto os mais velhos favorecem o uso do *você* (0,62). Nossos resultados se contrapõem aos encon-

trados por Godoy (1999), que observou que os falantes mais velhos (acima de 50 anos) utilizaram mais a forma *a gente* (0.54).

Alves (2010), ao analisar a alternância do *tu* e *você* no falar maranhense, verificou que o *você*, além de ter sido mais utilizado com referência genérica, foi também mais recorrente na fala dos mais velhos, corroborando, desse modo, nossos resultados.

O quarto e último fator relevante selecionado foi o fator sexo/gênero.

Tabela 6 - Distribuição das ocorrências segundo o sexo/gênero

|           |      | A gente | Você | TOTAL |
|-----------|------|---------|------|-------|
| Masculino | Nº   | 117     | 46   | 163   |
|           | %    | 71,8    | 28,2 |       |
|           | P.R. | 0,41    | 0,59 |       |
| Feminino  | Nº   | 82      | 23   | 105   |
|           | %    | 78,1    | 21,9 |       |
|           | P.R. | 0,63    | 0,37 |       |
| Total     | Nº   | 199     | 69   | 268   |
|           | %    | 74,3    | 25,7 |       |

Fonte: Elaborada pelos autores

Com relação ao sexo/gênero, observamos que as mulheres apresentaram um peso relativo de 0,63 para o uso do *a gente*, enquanto os homens apresentaram um peso relativo de 0,59 para o uso do *você*. São, portanto, as mulheres que favorecem a ocorrência de *a gente*, enquanto os homens favorecem o uso de *você*, corroborando os resultados encontrados por Setti (1997), Godoy (1999) e Santana (2006).

No que diz respeito às variáveis *tempo* e *modo verbal*, *tipo de oração*, *tipo de verbo*, *paralelismo formal* e *escolaridade*, vale destacar que elas não foram selecionadas como relevantes para a seleção dos pronomes *você* e *a gente* com referência genérica, no português falado no Maranhão.

## 5. CONSIDERAÇÕES FINAIS

A partir dos objetivos traçados para esta pesquisa, buscamos verificar quais os ambientes linguísticos e sociais que favorecem o uso das três estratégias não-canônicas de indeterminação do sujeito: *tu*, *você* e *a gente* no português falado no Maranhão.

Constatamos que a questão da indeterminação do sujeito não pode ser vista unicamente no nível sintático, uma vez que se trata de um "fenômeno linguístico explicável essencialmente nos níveis semântico e discursivo, visto que o que se indetermina é a referência do

sujeito, estando esta essencialmente relacionada ao contexto" (SANTANA, 2006, p. 140).

Com base nos resultados obtidos, vimos que a indeterminação pronominal do sujeito é um fenômeno condicionado por fatores sociais e linguísticos, confirmando que a escolha dos falantes pelo uso de uma dada forma não ocorre de maneira aleatória. Os fatores sociais (externos) localidade, faixa etária e sexo/gênero foram selecionados como sendo relevantes, e o grau de indeterminação foi o fator linguístico (interno) selecionado. A seleção deste fator interno mostra que a própria língua pode apresentar ambientes favoráveis a uma indeterminação completa ou a uma indeterminação parcial e, portanto, condicionar a escolha por determinada forma.

Desse modo, como postula a Sociolinguística, não há a possibilidade de dissociar o estudo da língua dos contextos e ambientes sociais nos quais ela está inserida. Sintetizando, concluímos que:

- o pronome *a gente* foi a estratégia mais utilizada pelos falantes maranhenses. Foi mais recorrente em contextos em que a indeterminação é parcial, na fala de informantes da faixa etária I, do sexo/gênero feminino, do município de São João dos Patos.
- o pronome *você* foi a segunda estratégia mais utilizada, tendo seu uso favorecido em contextos em que a indeterminação é completa; sendo mais frequente entre os sujeitos da faixa etária II, do sexo/gênero masculino, do município de Alto Parnaíba.

Vale destacar, em relação ao pronome *tu*, que, em virtude de sua baixa ocorrência (3,3%), não pudemos incluí-lo nas rodadas do Varbrul e obter os devidos pesos relativos. Essa baixa ocorrência pode dever-se ao fato de esse pronome já ter perdido certo espaço para o pronome mais inovador *você*, inclusive no que diz respeito ao seu uso mais genérico. Alves (2010), ao investigar a alternância do *tu* e *você*, observou o avanço do *você* em direção ao *tu* no português falado no Maranhão. Não descartamos, contudo, no que tange ao uso do *tu*, a necessidade de ser feito um estudo com base em um *corpus* que contemple outros contextos de interação verbal, de modo que possamos observar este fenômeno com uma maior amplitude.

Reiteramos, ainda, a importância deste estudo e de outros que partem do uso real da língua, para que seja observada a realidade linguística atual e possam ser elaborados materiais didáticos mais condizentes com a realidade linguística dos falantes. Ao afirmarmos isso, não desconsideramos a importância de a GT trabalhar com as formas canônicas; ao contrário, acreditamos que quanto mais o falante conhecer e souber, portanto, explorar as pos-



sibilidades que a língua lhe oferece, mais preparado estará para apropriar-se dos recursos comunicativos necessários para o desempenho das tarefas linguísticas com as quais se defronta em seu cotidiano.

## NOTAS

1. Trabalho premiado no XXIV Seminário de Iniciação Científica da Universidade Federal do Maranhão, na categoria Painel – Ciências Humanas. A pesquisa foi financiada por meio do Programa Institucional de Bolsa de Iniciação Científica – PIBIC/CNPq.
2. Tratamos como não-canônicas as estratégias de indeterminação que não são prescritas pela Gramática Tradicional.
3. O *Varbrul*, primeiramente chamado de *Variável Rules*, é um programa computacional criado em 1998 por David Sankoff e Pascale Rousseau, com a finalidade de fazer o cruzamento matemático de variáveis (previamente analisadas e codificadas), que possibilite observar quais fatores linguísticos e sociais são relevantes para a escolha de cada uma das variáveis dependentes enfocadas em um determinado estudo, apresentando, como produto final, os pesos relativos de cada variável. A respeito dos pesos relativos, Guy e Zilles (2007, p. 211) afirmam que estes "(...) calculam os efeitos dos fatores de cada grupo em relação ao nível geral de ocorrência das variantes e resultam de uma análise multivariada. O efeito, assim calculado, pode ser neutro (0,50), favorecedor (acima de 0,50) ou desfavorecedor (abaixo de 0,50) em relação à aplicação da regra em estudo.". É importante, portanto, fazer uma análise sociolinguística utilizando o *Varbrul*, pois, por meio dele, obtemos os devidos pesos relativos de cada variável e damos ao estudo um caráter mais preciso e científico. Nesta pesquisa, utilizamos a versão *GoldVarbX*, que foi desenvolvida por Sankoff, Tagliamonte e Smith (2005).
4. A realização da pesquisa em Alto Parnaíba, Balsas e Imperatriz foi financiada pelo CNPq, processo no 402408/2006-3, bem como a de São João dos Patos, Tuntum e Turiaçu, processo no 401119/2009-2.
5. Por questões éticas e para viabilizar o trabalho com os dados, o Projeto ALiMA adotou a seguinte codificação: a sigla do Estado do Maranhão – MA – seguida do número correspondente ao município pesquisado, o qual vem separado, por uma barra inclinada, do número que indica o perfil dos informantes. Com relação ao número que codifica os informantes, temos o que segue: os números pares foram atribuídos às mulheres e os ímpares, aos homens; os números de 1 a 4 foram utilizados para os sujeitos com escolaridade de nível fundamental e os números de 5 a 8 para os que têm o ensino superior completo; os números 1, 2, 5 e 6 correspondem aos informantes da faixa etária I (18 a 30 anos) e os números 3, 4, 7 e 8 indicam os informantes da faixa etária II (50 a 65 anos).
6. Vale observar que, para este trabalho, incluímos na indeterminação parcial tanto os casos de *indeterminação parcial* com o referente implícito no contexto, quanto aqueles com o referente explícito, visto que o grau de indeterminação não se constitui como o ponto central deste trabalho.

## REFERÊNCIAS

- ABAURRE, Maria Luiza M.; ABAURRE, Maria Bernadete M.; PONTARA, Marcela. *Português: contexto, interlocução e sentido*. São Paulo: Moderna, 2008.
- ALVES, Cibelle Corrêa Béliche. *O uso do tu e do você no português falado no Maranhão*. 2010. 143f. Dissertação (Mestrado em Linguística) – Programa de Pós-Graduação em Linguística, Universidade Federal do Ceará, Fortaleza, 2010.
- BENVENISTE, Émile. *Problemas de linguística geral*. 4. ed. São Paulo: Pontes; Campinas: Ed. da Universidade Estadual de Campinas, 1995. 199p. v.1.
- CARVALHO, Valter de. *Você, a gente et alia indeterminam o sujeito em Salvador*. 2010. 198f. Dissertação (Mestrado em Linguística) – Programa de Pós-Graduação em Estudo de Linguagens, Universidade Estadual da Bahia, Salvador, 2010.
- CASTILHO, Ataliba T. de. *Nova gramática do português brasileiro*. São Paulo: Contexto, 2010. 768p.
- CUNHA, Cláudia de Souza. *Indeterminação pronominal do sujeito*. 1993. 111 p. Dissertação (Mestrado em Linguística) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1993.
- FERREIRA, Camilla Maramaldo. A *representação da primeira pessoa do plural: o que mostram os dados do ALiMA*. São Luís: Universidade Federal do Maranhão, 2012. 54p. Relatório do Programa Institucional de Bolsa de Iniciação Científica.
- GODOY, Maria Alice Maschio. A *indeterminação do sujeito no interior paranaense: uma abordagem sociolinguística*. 1999. 128f. Dissertação (Mestrado em Letras e Linguística), Universidade Federal do Paraná, Curitiba, 1999.
- GUY, Gregory R.; ZILLES, Ana. *Sociolinguística quantitativa: instrumental de análises*. São Paulo: Parábola Editorial, 2007.
- LABOV, William. *Padrões sociolinguísticos*. Tradução de Marcos Bagno, Maria Marta Pereira Scherre, Caroline Rodrigues Cardoso. São Paulo: Parábola Editorial, 2008.
- LOPES, Célia Regina dos Santos Lopes; RUMEU, Márcia Cristina de Brito. O quadro de pronomes pessoais do português: as mudanças na especificação dos traços intrínsecos. In: CASTILHO, Ataliba T. de;

et al. (Org.) *Descrição, história e aquisição do português brasileiro*. São Paulo: Pontes Editores; Campinas: Fafesp, 2007.

MENÓN, Lorena. *Gramática sem segredos*. São Paulo: Escala Educacional, 2008.

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.

PASCHOALIN, Maria Aparecida. *Gramática: teoria e exercícios*. São Paulo: FTD, 2008.

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. 125 f. Dissertação (Mestrado em Linguística) - Curso de Pós-Graduação em Letras, Universidade Federal do Paraná, Curitiba, 1997.