

PORTUGUÊS BRASILEIRO E PORTUGUÊS EUROPEU: UMA SÓ LÍNGUA?¹

Clara Villarinho (Bolsista de IC da FAPERJ / UERJ)
Renê Forster (Bolsista de Extensão / UERJ)

RESUMO: Este trabalho compara o sujeito nulo em Português Brasileiro (PB) e o Português Europeu (PE) em *corpus* constituído pelas versões brasileira e portuguesa do livro “Toda Mafalda”, uma coletânea de tiras em quadrinhos. O arcabouço teórico tem base na Teoria Gerativa, mais especificamente, no modelo de Princípios e Parâmetros, que diz que as línguas existentes podem ser classificadas de acordo com a possibilidade para a realização do sujeito: em algumas, o sujeito é preferencialmente nulo e, em outras, é obrigatoriamente pleno. Como apontado em estudos anteriores, o PB parece estar sofrendo uma mudança, passando de uma língua de sujeito nulo a uma de sujeito pleno, diferenciando-se assim da variedade europeia da língua. Este trabalho testou essa hipótese, analisando uma estrutura pouco investigada: as interrogativas. Os resultados confirmaram que o número de ocorrências de sujeitos plenos é muito maior do que o de nulos no PB. Ainda assim, foi observada no *corpus* uma ocorrência, apesar de pequena, de sujeitos nulos. Por isso, não foi possível classificar rigorosamente o PB como uma língua de sujeito pleno. Contudo, foi possível constatar que os mecanismos pelos quais o PB licencia os nulos são diferentes dos mecanismos pelos quais o PE o faz.

1) Introdução

O nível de distanciamento observado entre o português falado no Brasil e o Português Europeu, há muitas décadas, vem sendo alvo de inúmeros debates dentro e fora do meio acadêmico. Até hoje, muitas hipóteses foram levantadas no intuito de se responder à questão: “Portugueses e brasileiros ainda falam a

mesma língua?" Porém, os debates, em geral, não estabelecem um consenso.

A grande polêmica que há a respeito desse assunto se dá, em grande parte, pela escolha de diferentes critérios para se estudar um mesmo objeto. Na maioria dos casos, os critérios utilizados para se diferenciarem as línguas estão voltados, principalmente, para a compreensão que pode haver entre seus falantes. Dessa maneira, a partir desse critério pode-se acabar por inferir que há apenas uma língua, compartilhada pelos dois países com apenas algumas variações dialetais.

No entanto, a compreensão não parece ser um critério abalizado, visto que diz respeito, sobretudo, aos aspectos fonológicos e lexicais de cada língua, além de remeter a uma habilidade pessoal (pessoas com maior ou menor facilidade de compreender outras línguas). Há línguas, como o espanhol e o italiano, que possuem uma origem latina em comum com o português e, por esse motivo, apresentam uma base fonológica, além de uma grande quantidade de itens lexicais, extremamente semelhantes aos do português e, nem por isso, supõe-se que estas sejam apenas variações de uma mesma língua. É possível que haja uma aproximação nesse nível entre as diferentes línguas e, mesmo que seja quantitativamente relevante, nem sequer se cogita a possibilidade de se conjecturar que sejam a mesma língua. Além disso, uma hipótese baseada em critérios fonológicos e lexicais teria que dar conta das variações linguísticas (dialetos) sem que estas sejam consideradas línguas diferentes.

Por outro lado, outros estudiosos apóiam seus argumentos em uma hipótese facilmente verificável empiricamente: a de que cada língua possui características estruturais específicas que a distinguem e a individualizam. Assim, o critério utilizado para se caracterizar, identificar e diferenciar as línguas passa do nível léxico / fonológico ao nível mórfico / sintático (mais precisamente, os aspectos sintáticos e algumas influências dos morfológicos).

De tal maneira, toda a discussão a respeito do distanciamento entre o português falado no Brasil (doravante PB) e o português

européu (doravante PE) torna-se fruto e, ao mesmo tempo, motivação de pesquisas comparativas que procuram encontrar características estruturais distintas nas duas línguas, para que se possa concluir, fundamentadamente, se PB e PE são ou não duas línguas. O presente trabalho apresenta resultados de uma pesquisa que visa contribuir para esse debate, baseando-se na hipótese de que as línguas podem ser diferenciadas através de estruturas sintáticas comuns ou distintas.

Neste caso, a comparação se refere ao tipo de sujeito utilizado nas sentenças do PB e PE, por acreditar-se na hipótese de que as línguas possam ser classificadas de acordo com a realização fonética do sujeito: há línguas em que a expressão fonética é obrigatória e há línguas em que o sujeito é preferencialmente nulo, ou seja, não é realizado (Kato e Negrão, 2000). Estudos anteriores (Duarte 1993, 1995) demonstram como uma transformação ocorrida no PB, relativa à redução de seu paradigma flexional, pode ter contribuído para uma mudança estrutural que distinguiria esta língua do PE.

Para verificar tais hipóteses, enquadrámos nossa pesquisa no campo da sociolinguística paramétrica e analisamos quantitativa e qualitativamente dados reais retirados de um *corpus* formado pelas versões brasileira e portuguesa do livro de tiras em quadrinhos “*Toda Mafalda*”, de Quino.

2) O modelo de Princípios e Parâmetros, o processo de aquisição da linguagem e a questão do sujeito nulo

A Teoria Gerativa (Chomsky, 1957) veio, desde sua concepção inicial, se desenvolvendo em alguns modelos. Um desses modelos é o de Princípios e Parâmetros (Chomsky, 1981), segundo o qual a língua seria organizada em dois conjuntos distintos de características: aquelas universais, chamadas de *princípios*, e aquelas próprias de cada língua, chamadas *parâmetros*.

As características universais seriam propriedades comuns a todas as línguas e são, por isso, chamadas de princípios, já que são princípios aos quais todas as línguas devem *a priori* obedecer. Os

princípios formariam, no seu conjunto, a chamada Gramática Universal (GU), que é, segundo a Teoria Gerativa, o estágio inicial de aquisição da língua, ou seja, informações sobre língua que, antes mesmo da experiência lingüística, estão codificadas na mente do indivíduo. São, em uma palavra, propriedades *inatas*, informações que nascem com todos indivíduos da espécie humana. Essas informações servem de base para processo de aquisição da linguagem, durante o qual o indivíduo irá adquirir aquelas características próprias de sua língua, os parâmetros.

Os parâmetros são desdobramentos das características universais, dos princípios. De acordo com o modelo em questão, cada princípio possuiria duas possibilidades de variação e cada uma destas possibilidades é um parâmetro. Desta forma, se existe um princípio X, esse princípio poderá assumir duas variações, ou melhor, dois parâmetros: ou X+, ou X-. Então, se tivermos, por exemplo, os princípios X, Y e Z, teremos, para cada um deles, duas marcações de parâmetros distintas, resultando em várias possibilidades combinatórias, que formarão conjuntos diferentes: [X+, Y+, Z+], [X-, Y-, Z-], [X-, Y+, Z-] e assim por diante. Esses conjuntos diferentes de parâmetros constituirão a gramática própria de cada língua, ou, como se costuma dizer, a *gramática particular* de cada língua.

Ao contrário da GU, a gramática particular é dada em função da experiência lingüística do indivíduo. Os parâmetros serão marcados de acordo com os dados lingüísticos aos quais as crianças serão expostas durante o seu período de aquisição da linguagem. A tarefa da mente da criança, neste processo de aquisição, será a de, a partir dos princípios, previamente fornecidos, identificar os parâmetros da língua a qual a criança está sendo exposta para formar, assim, a gramática de sua língua.

2.1) A questão do sujeito nulo

Como já foi dito, o modelo de Princípios e Parâmetros postula que existe um princípio segundo o qual todas as línguas,

em todas as suas orações, devem projetar uma posição para o sujeito. Esse é o chamado *Princípio da Projeção Estendida*. Esse princípio assume duas possibilidades de variação, dois parâmetros: algumas línguas expressam obrigatoriamente esse sujeito de forma explícita, atribuindo a ele uma cadeia fonética, já outras línguas, as quais a Teoria Gerativa chama de línguas *pro-drop* (*pronoun-drop*), determinam uma preferência por sujeitos não expressos foneticamente, configurando os sujeitos que a tradição gramatical chama de *sujeitos nulos*.

Um exemplo da primeira classe de línguas é o inglês. Observamos nessa língua uma obrigatoriedade em se expressar o sujeito, mesmo no caso de verbos que, semanticamente, não possuem um sujeito, como, por exemplo, os verbos que expressam fenômenos naturais. Na oração "*It rains.*" (=Chove.), por exemplo, pode-se perceber que antes do verbo *rains* (*chove*), existe um elemento, mais especificamente, um pronome, marcando a posição de sujeito desta oração.

Em línguas como o inglês, o número de ocorrências de sujeitos foneticamente expressos (sujeitos plenos) será sempre elevado e orações que não tragam o sujeito explícito poderão ser consideradas sintaticamente impossíveis. Certamente, não se pode converter a oração "*I am a student.*" (=Eu sou um estudante.), em uma oração correlata sem a presença explícita do sujeito, "**Ø Am a student*", pois esta última não seria entendida por um falante nativo, pelo menos não sem algum estranhamento, já esta sentença não respeita os parâmetros sintáticos do inglês.²

Já no segundo conjunto de línguas, nas línguas de sujeito nulo, a ocorrência de sujeitos não expressos foneticamente será sempre mais elevada. Por exemplo, no espanhol, que é classificada como uma língua *pro-drop*, a utilização do pronome em algumas orações, apesar de possível, é muito incomum. Certamente a seqüência "*Ø Soy un estudiante.*" (=Sou um estudante.) é mais comum e mais aceita entre os falantes do que a seqüência "*Yo soy un estudiante.*" (=Eu sou um estudante.). Em algumas outras

orações, a expressão do sujeito seria mesmo impossível, como na oração “Ø *Llueve.*” (=Chove.).

Durante o processo de aquisição, o parâmetro do sujeito nulo é um dos parâmetros que a criança terá que desvendar para poder definir os limites de sua língua. Ao adquirir uma língua, a mente de uma criança irá, a partir dos dados lingüísticos aos quais será exposta, definir uma, dentre as duas estratégias possíveis de apresentação do sujeito, como legítima para a língua que está sendo adquirida. Assim, o parâmetro do sujeito nulo é um dos parâmetros tomados como referência pela mente para diferenciar dada língua de uma outra.

Sendo assim, podemos considerar que este parâmetro possa ser também uma variável útil para verificar a unidade de um determinado sistema lingüístico. Se quisermos saber se determinadas variantes de um sistema lingüístico são ou não a mesma língua, podemos considerar que um dos procedimentos possíveis será realizar uma checagem comparativa das estratégias de apresentação do sujeito. Se as estratégias forem iguais, não se terá muito a dizer, pois ainda será necessário proceder a verificação das outras estratégias sintáticas da língua. Mas, se elas forem diferentes, poderemos dizer de pronto que se trata de línguas distintas, pois se, entre dois conjuntos de regras, tem-se um elemento que não é igual, pode-se dizer então que os conjuntos são diferentes, ou mais propriamente, neste caso, línguas diferentes.

3) O comportamento do PB e do PE em relação ao sujeito nulo

Vejamos agora, algumas hipóteses sobre o comportamento do Português Europeu e do Português Brasileiro em relação ao sujeito nulo e quais os resultados obtidos até agora com a utilização do parâmetro do sujeito nulo para caracterizar a relação entre os dois.

Duarte (1993) é uma das primeiras autoras a abordar a questão do sujeito nulo no Português Brasileiro, sob o modelo de

Princípios e Parâmetros. Neste trabalho, ela constata uma mudança no mecanismo de preenchimento do sujeito no PB. Baseando-se em peças populares de teatro, escritas a partir do final da primeira metade do século XIX até o final do século XX, a autora constata que o PB, que até o começo do século XX mostrava uma clara preferência pelos sujeitos nulos—o número de ocorrências de nulos nas três primeiras peças estudadas (1845, 1882, 1918) era de cerca de 75%—passou, a partir da terceira década do século XX, a mostrar uma redução crescente do número de nulos, chegando finalmente a uma preferência significativa pelos plenos na última peça estudada, na qual a ocorrência de sujeitos nulos foi de menos de 30%.

A autora faz ainda uma correlação desse progressivo aumento de sujeitos plenos com a redução do paradigma flexional registrada no PB. Em função de duas outras mudanças ocorridas na língua (primeiramente, a perda dos pronomes *vós* e *tu*, e, em segundo, a substituição da forma *nós* pela forma *a gente*), o PB, de um paradigma flexional composto por seis formas, passou a um paradigma composto por apenas três formas, como mostra a Tabela 1. Essa progressiva redução operada no paradigma se deu, segundo a autora, concomitantemente, com a preferência do PB pelos sujeitos plenos. Com base nisso, a autora sugere que o PB teria passado a preferir os sujeitos plenos em função dessa redução do paradigma.

	Paradigma 1		Paradigma 2	
1. ^a pessoa do sing.	eu	amo	eu	amo
2. ^a pessoa do sing.	tu	amas	você	ama
3. ^a pessoa do sing.	ele	ama	ele	ama
1. ^a pessoa do plur.	nós	amamos	a gente	ama
2. ^a pessoa do plur.	vós	amais	vocês	amam
3. ^a pessoa do plur.	eles	amam	eles	amam

Tabela 1: A evolução do paradigma flexional no Português Brasileiro

Para fazer essa correlação, a autora se baseia numa hipótese segundo a qual as línguas com um paradigma flexional mais diversificado tendem a preferir os sujeitos nulos, enquanto as línguas com paradigma flexional pouco diversificado tendem a preferir sujeitos plenos. Tal fato acontece porque na primeira classe de línguas, a flexão verbal seria por si só capaz de identificar o sujeito, enquanto, na segunda, a flexão seria incapaz de realizar tal identificação, deixando-a a cargo do pronome. Podemos citar, como exemplos, o espanhol e o inglês. Neste o paradigma flexional é pobre (ver Tabela 2) e há uma preferência pelos sujeitos plenos, naquele o paradigma flexional é rico (ver Tabela 2) e a preferência é pelos sujeitos nulos.

	Inglês	Espanhol
1ª. pessoa do sing.	love	amo
2ª. pessoa do sing.	love	amas / ama
3ª. pessoa do sing.	loves	ama
1ª. pessoa do plur.	love	amamos
2ª. pessoa do plur.	love	amáis
3ª. pessoa do plur.	love	aman
TOTAL	2 formas	6 formas

Tabela 2: Paradigmas verbais do inglês e do espanhol: verbo amar no presente do indicativo

A partir daí, é possível concluir que o empobrecimento do paradigma flexional do português poderia ter feito eclodir no PB uma preferência pelos sujeitos plenos, já que o paradigma flexional não seria mais, por si, só capaz de identificar os sujeitos nulos.

O surgimento, no PB, dessa preferência pelos sujeitos plenos modifica radicalmente sua relação com o PE, já que temos neste uma clara preferência pelos sujeitos nulos, como poderá ser observado nos próprios resultados da presente pesquisa. A mudança do mecanismo de apresentação do sujeito ocorrida no PB parece não ter paralelo no PE, e por isso poderíamos apontar a

hipótese de que esses dois sistemas lingüísticos não possuem mais o mesmo conjunto de regras sintáticas, ou melhor, o mesmo conjunto de marcações de parâmetro, o que poderia levar à conclusão de que se trata de duas línguas diferentes. Ainda existem, no entanto, alguns detalhes que merecem atenção antes que se faça tal afirmação.

Apesar de haver uma grande preferência pelos plenos no PB, ainda existe um número razoável de ocorrências de sujeitos nulos. O número de nulos apontado, por exemplo, na referida pesquisa de Duarte, é um índice elevado, se comparado a outras línguas de sujeito pleno, como o Inglês ou o Francês. Para caracterizar, com segurança, o PB como uma língua de sujeito pleno, seria necessário que as pesquisas indicassem um número de nulos ainda menor, já que é teoricamente inviável postular que o parâmetro do sujeito nulo esteja marcado "pela metade" no PB.

Como já foi dito, só existem, para cada princípio, duas possibilidades de marcação paramétrica, sendo impossível que uma língua possua as duas marcações ao mesmo tempo. Assim, é de se estranhar que o PB, ao mesmo tempo em que apresenta uma preferência pelos plenos, apresente também uma disponibilidade para a ocorrência dos nulos, que é incondizente com o que se conhece sobre as línguas de sujeito pleno, nas quais as ocorrências de sujeitos nulos são restritas a contextos realmente muito específicos, sendo desta maneira incompatíveis com os quase 30% encontrados no PB por Duarte.

Existem, no entanto, algumas hipóteses que tentam explicar o fator que poderia estar licenciando essa ocorrência de sujeitos nulos no PB. Duarte (1995) oferece uma explicação para esta presença significativa de nulos no Português Brasileiro. Segundo ela, esses nulos que ainda se fazem presentes no PB não seriam frutos da atual gramática do PB, mas sim resíduos de uma gramática antiga, que é reforçada no processo de escolarização. A autora afirma isso, provavelmente, com base no fato de que a escola difunde alguns aspectos lingüísticos que não fazem mais parte da

gramática natural dos falantes, como, por exemplo, o relativo *cujo*. Então, a ocorrência dos nulos no PB poderia também ser um vestígio de uma gramática ultrapassada, que é mantida artificialmente, tendendo, desta forma, a desaparecer. Nestes termos, poderíamos concluir que o PB já é uma língua de sujeito pleno e, portanto, uma língua diferente do PE (Kato, 2005).

Em contraposição a Duarte (1995), Figueiredo Silva (1996) sinaliza outra hipótese que pode explicar a ocorrência dos nulos em PB. A autora aponta que tal ocorrência de nulos poderia ser explicada se fazendo recurso ao contexto discursivo. Em algumas línguas, como o chinês, o contexto discursivo desempenha um importante papel na identificação dos sujeitos. O Chinês, apesar de possuir um paradigma flexional pobre, é uma língua de sujeito nulo, na qual o contexto antecedente e/ou pragmático se encarrega de identificar os sujeitos.

Já que a flexão verbal já não é mais capaz de identificar o sujeito no PB, poderíamos supor que talvez o contexto pudesse desempenhar esse papel, assim como o faz no chinês. Neste caso, também poderíamos afirmar que houve uma mudança paramétrica, o que demonstra que o PB e o PE são sistemas lingüísticos distintos.

Nas seção 4, a seguir, tentamos mostrar como esta pesquisa pretende contribuir para que estas hipóteses sejam verificadas.

4) A pesquisa

Para a Teoria Gerativa, a língua é um conjunto de conhecimentos que cada (e todo) indivíduo tem de sua língua nativa (esse conhecimento é chamado Competência). Este é seu objeto de estudo: um sistema mental. Existem, porém, alguns estudos de base gerativista (principalmente aqueles que investigam características de uma determinada língua) que precisam ir além do nível de representações mentais e que, para isso, utilizam dados de fala reais em suas pesquisas. Isso é possível porque a língua é

um sistema abstrato que pode ser concretizado (porém com algumas alterações exteriores à faculdade da linguagem) através da fala (essa concretização é chamada Desempenho). A corrente da lingüística que une a Teoria Gerativa e os dados reais é chamada "Sociolingüística Paramétrica".

Desta maneira, seria incoerente para uma pesquisa que pretende investigar uma determinada língua baseando-se nesta Teoria utilizar um *corpus* retirado de fontes escritas formais, como textos literários, visto que estas estariam impregnadas pela norma culta, por isso, se afastariam grandemente dos dados reais da fala espontânea dos usuários de uma língua. Assim, o *corpus* utilizado nesta pesquisa foi retirado das versões brasileira e portuguesa de uma coletânea de tiras em quadrinhos (o livro "Toda Mafalda"), que mistura os registros formal e informal, aproximando-se mais da língua falada que da língua escrita.

Poderia-se argumentar que, para uma maior fidelidade aos estudos sobre PB e PE, os dados deveriam ser retirados de um *corpus* formado por atos reais de fala espontânea. No entanto, três importantes motivos justificam a escolha do *corpus* de "Toda Mafalda". O primeiro diz respeito às pesquisas anteriores sobre o preenchimento do sujeito no Português, que freqüentemente têm seus *corpora* retirados destas transcrições (ou seja, as tiras em quadrinhos são uma fonte ainda não explorada). O segundo argumento trata especificamente da comparação entre as línguas, já que os dados foram retirados das traduções brasileira e portuguesa deste livro, o que permite confrontar duas versões de um mesmo texto de partida. O terceiro e talvez mais importante motivo é o de algumas estruturas que dificilmente podem ser encontradas em outras manifestações de fala espontânea, como por exemplo entrevistas, poderem ser obtidas em número relevante nas tirinhas.

Assim, a partir do *corpus* foram selecionadas sentenças interrogativas, para que se pudesse verificar nos dados empíricos as hipóteses iniciais. As interrogativas têm possível influência no

preenchimento do sujeito. Assim, a partir das interrogativas, foram separados sujeitos plenos e nulos e classificados de acordo com elementos mais gerais que notadamente influenciam o preenchimento, como a animacidade do sujeito (animado / inanimado). Posteriormente, as estruturas foram analisadas a partir desses critérios de preenchimento mais intrínsecos a cada uma.

No caso das interrogativas, o fator que influenciaria o preenchimento é a existência ou não de um elemento à esquerda do sujeito. Portanto, foram divididos e analisados separadamente os dados de interrogativas que possuíam (como em "Aonde você vai?") e os que não possuíam esses elementos (como em "Ø Você vai sair?").

Para que se apurasse, sem dúvidas, se haveria diferenças entre o mecanismo de reconhecimento que licencia o sujeito nulo no PB e o mecanismo de reconhecimento utilizado pelo PE, outros fatores mais gerais de influência também foram analisados em cada caso, de acordo com cada estrutura, como a menção (anterior ou posterior) ao sujeito (como em "Vai sair, Mafalda?"), a existência de um elemento entre o sujeito e o verbo (como em "Você já vai sair?"), a pessoa do discurso, etc.

Todos esses dados passaram por um processo de análise, separados pelos critérios mais gerais, utilizando o programa "Varbrul" para análise estatística, a fim de quantificá-los e, posteriormente, compará-los. Terminada a análise quantitativa, foram feitas análises qualitativas, com cada critério específico para cada tipo de estrutura.

5) Análise dos resultados

Nesta seção serão apresentados, através de gráficos, os resultados obtidos através da análise comparativa dos dados de PB e PE, tanto gerais quanto específicos, de acordo com cada tipo de estrutura e suas influências externas. Os resultados que se seguem são fruto de um estudo quantitativo, em primeiro plano, e qualitativo, em segundo, no qual procurou-se isolar quaisquer

tipos de fatores que pudessem causar uma má interpretação dos resultados, sendo estes fatores também estudados para uma análise mais profunda.

5.1) Interrogativas sem elementos à esquerda do sujeito

A divisão das sentenças interrogativas entre “com elemento à esquerda do sujeito” e “sem elemento à esquerda do sujeito” faz-se necessária, neste caso, por esta pesquisa ter como objetivo a verificação das hipóteses citadas. Esta divisão faz sentido ao se pensar no elemento à esquerda como uma barreira que impediria a ligação da sentença com o contexto, o que, segundo a hipótese de Figueiredo Silva (1996), favoreceria a utilização do sujeito pleno.

Aqui serão investigadas as orações sem elementos à esquerda do sujeito para que se possam obter resultados onde não houvesse esse tipo de influência e se pudesse verificar a hipótese de Duarte (1993, 1995), a respeito da impossibilidade de se identificar o sujeito nulo através das desinências verbais, que levaria a uma preferência pelo sujeito pleno no PB, encaminhando-o para uma mudança na marcação paramétrica.

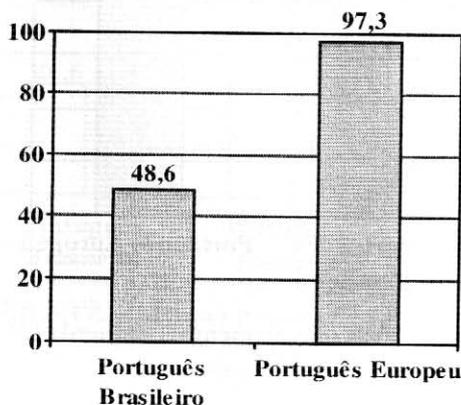


Figura 1: Interrogativas sem elemento à esquerda (dados gerais)
– % de nulos

A análise dos dados nos quais o sujeito começa a sentença, como “(PB) Eu te contei que meu irmãozinho já engatinha, Filipe? (t.1022)” X “(PE) Já te disse que o meu irmãozinho já gatinha, Filipe? (t.1017)”, mesmo em uma análise geral com todos os dados, já apresenta um resultado esperado, de acordo com a hipótese de Duarte. Enquanto no PE percebe-se uma preferência categórica pelo sujeito nulo (97,3%), o percentual de nulos encontrado no PB, embora ainda exista, é extremamente baixo se comparados a uma língua de sujeito nulo. Isto mostra que a flexão verbal, realmente, não é mais responsável pela identificação do nulo no PB, ao contrário do que acontece no PE.

Com a intenção de se aprofundar ainda mais a análise dos dados, foram focalizadas apenas as sentenças em que o sujeito fosse de segunda pessoa, por ser onde são encontradas as perguntas “clássicas”, ou seja, as perguntas direcionadas para uma pessoa, como em “(PB) Você está pensando só num pontinho raquítico? (t. 60, p. 20)” X “(PE) Ø Estás a pensar apenas num miserável pontinho? (t.60, p.72)”.

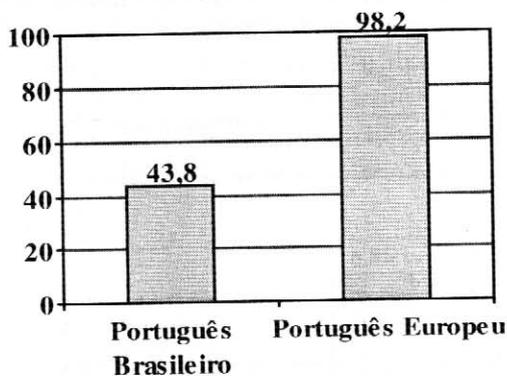


Figura 2: Interrogativas sem elemento à esquerda (2ª pessoa) – % de nulos

Novamente, o alto percentual de nulos do PE contrasta com uma grande tendência de se expressar o sujeito no PB. Em-

A análise dos dados nos quais o sujeito começa a sentença, como “(PB) Eu te contei que meu irmãozinho já engatinha, Filipe? (t.1022)” X “(PE) Já te disse que o meu irmãozinho já gatinha, Filipe? (t.1017)”, mesmo em uma análise geral com todos os dados, já apresenta um resultado esperado, de acordo com a hipótese de Duarte. Enquanto no PE percebe-se uma preferência categórica pelo sujeito nulo (97,3%), o percentual de nulos encontrado no PB, embora ainda exista, é extremamente baixo se comparados a uma língua de sujeito nulo. Isto mostra que a flexão verbal, realmente, não é mais responsável pela identificação do nulo no PB, ao contrário do que acontece no PE.

Com a intenção de se aprofundar ainda mais a análise dos dados, foram focalizadas apenas as sentenças em que o sujeito fosse de segunda pessoa, por ser onde são encontradas as perguntas “clássicas”, ou seja, as perguntas direcionadas para uma pessoa, como em “(PB) Você está pensando só num pontinho raquítico? (t. 60, p. 20)” X “(PE) Ø Estás a pensar apenas num miserável pontinho? (t.60, p.72)”.

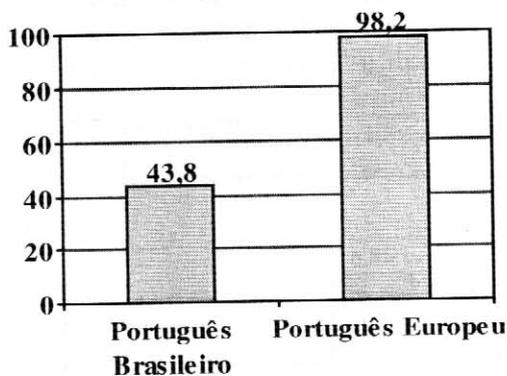


Figura 2: Interrogativas sem elemento à esquerda (2ª pessoa) – % de nulos

Novamente, o alto percentual de nulos do PE contrasta com uma grande tendência de se expressar o sujeito no PB. Em-

bora as diferenças sejam mínimas se as Figuras “geral” e “segunda pessoa” forem comparadas, nota-se uma redução ainda maior dos sujeitos nulos na segunda (quase 5%), talvez por ser nessa pessoa onde se encontra uma maior dificuldade de identificação do sujeito através da desinência verbal (os verbos na segunda pessoa, se tiverem seus sujeitos nulos, podem ser confundidos tanto com a terceira pessoa, quanto com a primeira do plural: “você/ ele/ a gente gosta”, “vocês / eles gostam”). Esse resultado, assim como o anterior, sugere que a flexão *não* é mais responsável pela identificação do nulo no PB.

Após essa seleção dos dados apenas de segunda pessoa, estes foram analisados de forma que se pudesse investigar se a opção pelo preenchimento ou não se devia a fatores não intrínsecos à estrutura da interrogativa, como a menção ao sujeito (como em “(PB) Você gosta da primavera, passarinho? (t.499)” X “(PE) Gostas da primavera, passarinho? (t.495)”) e a existência de um elemento entre o sujeito e o verbo (como em “(PB) Ø Já imaginou, Susanita? (t.310)” X “(PE) Já pensaste, Susanita? (t.308)”).

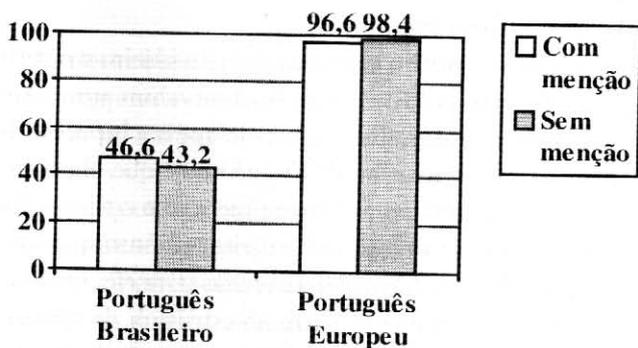


Figura 3: PB X PE – Menção posterior ao sujeito (2ª pessoa) – % de nulos

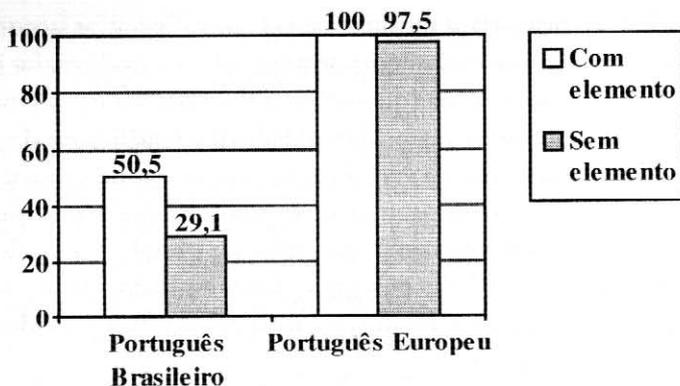


Figura 4: PB X PE – Elemento entre o sujeito e o verbo (2ª pessoa) – % de nulos

Os resultados da Figura 3 mostram uma diferença muito pequena no percentual de nulos com e sem menção, tanto no PB quanto no PE, revelando que, ao contrário do esperado, a menção ao sujeito não favoreceu o nulo (supunha-se que, tendo sido mencionado, a tendência era que se diminuísse a necessidade de se expressar o sujeito pronominal).

Já a existência de um elemento entre o sujeito e o verbo foi relevante apenas para o PB, em que se observa um aumento considerável na taxa de nulos. Esse resultado apóia a hipótese de que a exigência feita pela estrutura da língua não é que se preencha o sujeito pronominal em si, mas que se preencha o *espaço* destinado ao sujeito (como toda oração tem sujeito, mesmo que este não seja preenchido, há um lugar determinado para ele, uma categoria vazia), ou seja, o elemento entra na estrutura da sentença no lugar do sujeito (por exemplo, “Já foi ao mercado?”, “Não gosta de mim?”, etc.) (Duarte, 1995).

Analisando os resultados como um todo, foi possível observar que no PB, mesmo nas sentenças onde não havia outros fatores que favoreceriam o sujeito pleno e naquelas onde o nulo era esperado como opção mais provável, há uma forte tendência

de se expressar sempre o sujeito (mostrando que as desinências não podem mais identificá-lo). No PE nenhum fator alterou significativamente o alto percentual de nulos, resultado já esperado, por o PE se tratar de uma língua *pro-drop* que identifica os nulos através das desinências. Se o PB utiliza um mecanismo de identificação dos sujeitos nulos diferente do PE, já é suficiente para afirmar que estas teriam se tornado línguas distintas. Porém, ainda há a questão do parâmetro que não pode ser marcado “pela metade”. Se o PB deixou de ser uma língua *pro-drop*, é necessário que se verifiquem as hipóteses que podem explicar a existência ainda de sujeitos nulos.

5.2) Dados com elemento à esquerda do sujeito

Neste item da pesquisa, examinamos dados semelhantes aos Exemplos A e B:

Exemplo A – E o Paulo? Que menino ajudou na prova?

Exemplo B – E o Paulo? Ajudou que menino na prova?

O nosso objetivo era detectar a possível influência dos elementos à esquerda do sujeito, em orações interrogativas como a sublinhada no exemplo A (esses exemplos foram construídos e não retirados do *corpus*), onde se vê um pronome interrogativo funcionando como elemento à esquerda. No *corpus* foram encontrados não só dados com este tipo de elemento, mas também com outros tipos elementos à esquerda, como advérbios, objetos antepostos ao sujeito, adjetivos, interjeições, partículas do tipo “*será que*” e “*dá prá*”, vocativos e outros elementos que ocorreram em quantidades pouco significativas.

A intenção ao analisar esses dados era a de verificar uma das hipóteses que tentava dar conta dos nulos ainda presentes no PB. Mais especificamente, o objetivo de testar essa variável foi o de verificar a hipótese apontada por Figueiredo Silva (1996) so-

bre a influência do contexto no preenchimento do sujeito. Se o contexto realmente estivesse sendo o licenciador dos sujeitos nulos no PB, então, os elementos à esquerda poderiam ser considerados como um fator de influência nesse preenchimento, já que eles poderiam interromper a ligação entre o contexto e o sujeito da oração, diminuindo, assim, o poder identificador do contexto, ou seja, a hipótese é a de que se houver um elemento à esquerda se torna mais difícil para o contexto licenciar o nulo na oração investigada.

Para proceder essa checagem, verificamos os elementos à esquerda, tanto nos dados da versão portuguesa, quanto nos dados da versão brasileira do livro "Toda Mafalda". Em seguida, cruzamos essas informações com o número de ocorrências de sujeitos nulos. Além das orações com elementos à esquerda, como a do exemplo A, fizemos também a verificação em um grupo de controle, constituído por orações como a sublinhada no exemplo B, na qual não existe um elemento à esquerda do verbo.

Note-se que, entre os exemplos A e B, em conformidade com a hipótese abordada aqui, parece ser mais difícil determinar o sujeito na oração A, já que nela existe um elemento à esquerda do sujeito. No exemplo B a identificação do sujeito do verbo *ajudar* parece ser mais fácil, já que não existe um elemento à esquerda impedindo a ligação entre o sujeito e o contexto.

Antes de prosseguirmos a análise dos dados, vendo um dos exemplos retirados do *corpus*, cabe ainda uma observação sobre os exemplos já citados. É nítido que, no exemplo A, não é apenas o elemento à esquerda que dificulta a identificação do verbo, mas também uma ambigüidade gerada por dois fatores. Primeiro, o de haver dois candidatos à classe de sujeito: "Paulo" e "menino". Em segundo lugar, o de o objeto "que menino" estar ocupando a posição canônica do sujeito. Mas, ainda sim, podemos ver o pronome interrogativo como um elemento que dificulta a identificação do sujeito, se tivermos em mente que o movimento feito pelo objeto, para se colocar na posição de sujeito, foi de caráter obriga-

tório e o será sempre, em situações semelhantes. Isso nos leva à conclusão de que tal ambigüidade se deve às modificações desencadeadas pela presença do pronome. Além disso, podemos considerar também que a ambigüidade não é gerada somente pelo fato do objeto estar ocupando a posição do sujeito, mas também pelo fato de esta oração não seguir a orientação geral do PB, que é a de preencher os sujeitos. Caso essa tendência fosse seguida nesta oração, não se observaria esta ambigüidade (“E o Paulo,? Que menino ele, ajudou na prova?”).

Vejamos agora um exemplo extraído do *corpus*, onde o elemento à esquerda, nesse caso um pronome interrogativo, pode ter funcionado como barreira para o contexto: “O que você gostaria de ser?”. Neste exemplo existe um elemento (“o *que*”) à esquerda da posição de sujeito, que se encontra preenchido pelo pronome *você*. Conforme a hipótese com a qual estamos trabalhando neste momento, podemos concluir que o sujeito pleno apresentado aí pode ter sido causado pela presença do elemento à esquerda. A frase antecedente “O que você quer ser quando crescer, Mafalda?”, possui duas informações, um pronome e um nome próprio. Estes indicam que o sujeito da oração “O que você gostaria de ser?” funcione como um mecanismo licenciador do nulo.

Passemos agora, na Figura 5, aos primeiros resultados obtidos ao cruzarmos a ausência ou a presença de um elemento à esquerda com a possibilidade do sujeito ser nulo ou pleno. Como a maioria das Figuras apresentadas neste trabalho, consideramos, para efeito de análise, somente os dados nos quais o sujeito era de segunda pessoa do singular. Isso porque estamos estudando as interrogativas e a segunda pessoa é a pessoa de ocorrência mais provável em uma interrogativa produzida num diálogo entre duas personagens.

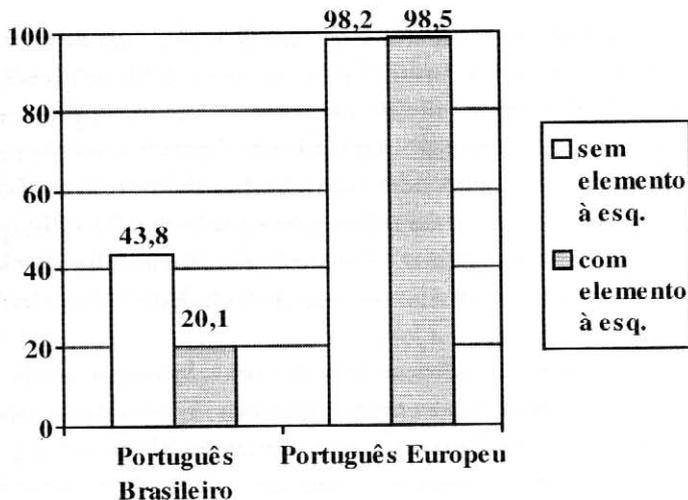


Figura 5: Percentual de nulos em função da presença de elementos à esquerda no PB e no PE

Observamos na Figura 5 que, no PB, o elemento à esquerda influenciou o preenchimento do sujeito, pois a sua presença diminuiu o número de nulos (20,1% de nulos; num total de 199 dados, com elemento à esquerda, 40 foram nulos), ao passo que a sua ausência permitiu que os nulos fossem maiores (43,8% de nulos; num total de 171 dados, sem elemento à esquerda, 75 foram nulos). Este resultado parece corroborar a hipótese de que o contexto pode ter assumido o papel de licenciar os nulos no PB.

No mesmo Gráfico 5, nas duas colunas à direita, que representam o PE, percebemos que os elementos à esquerda tiveram pouca influência sobre o elevado número de nulos, que se manifestou com ou sem os elementos à esquerda. A explicação para esse fenômeno é que, no PE, quem licencia o nulo é a flexão verbal e não o contexto, por isso os elementos à esquerda não influenciaram a ocorrência de nulos.

Na Figura 6, apresentamos os dados com elemento à esquerda em uma abordagem mais restrita. Para esta Figura consi-

deramos somente os elementos à esquerda representados por pronomes ou advérbios interrogativos.

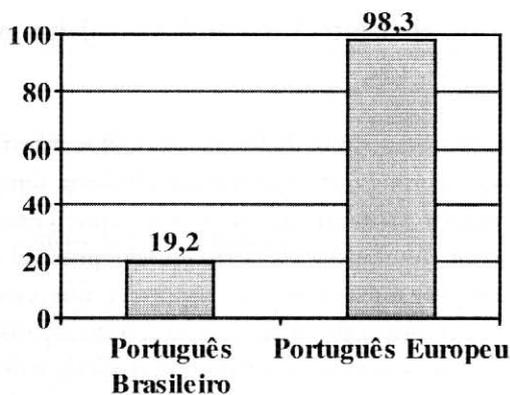


Figura 6: Porcentagem de nulos em função da presença de um elemento interrogativo à esquerda do sujeito

Na Figura 6, vemos repetida a tendência do PB em registrar uma influência dos elementos à esquerda no preenchimento do sujeito. Num total de 166 dados, 32 foram nulos, ou seja, 19,2%, como representado na Figura 6. Em comparação com a taxa de nulos dos dados sem elementos à esquerda, apresentada na Figura 6, a taxa de nulos dos dados com elemento interrogativo foi cerca de 100% menor, no PB, comprovando a influência do contexto. Novamente, no PE, a presença dos elementos à esquerda pareceu insignificante. Os exemplos C e D foram retirados do *corpus* de pesquisa e representam orações deste último tipo.

Exemplo C – *Versão em PB*: “O que você gostaria de ser?” (tira 68)

Exemplo D – *Versão em PE*: “Que Ø gostavas de ser?” (tira 68)

Uma última observação que pode ser feita quanto a estes dados é em relação à robustez do *corpus*. Obtivemos, em relação aos elementos à esquerda, uma quantidade considerável de

interrogativas que, dificilmente, poderia ser obtida em outros *corpora*, como em entrevistas.

5.2.1) Dados com elemento à esquerda do sujeito: menção ao sujeito

Por motivos práticos, nos dados apresentados anteriormente, partimos do pressuposto de que nas tiras haveria sempre um elemento que pudesse fornecer uma pista sobre quem era o sujeito das orações com elemento à esquerda, pois é possível afirmar que o contexto pragmático irá, na maioria dos contextos discursivos, oferecer pistas de quem seja o sujeito de uma dada oração. Como já foi esclarecido, a idéia, então, seria a de que os elementos à esquerda iriam dificultar o aproveitamento desta pista, pois, na presença deles, a mente do falante seria obrigada a processar mais um dado, antes de processar a posição do sujeito, sobrecarregando, assim, a memória do falante.

A partir de um dado momento, decidimos nos apoiar não somente na pressuposição de que o contexto estava presente, fornecendo pistas, mas decidimos checar também se este contexto era enunciado de forma explícita, através de uma menção clara, em alguma parte da tira, ao sujeito da oração investigada. Para isso, procedemos, no PB, a verificação de uma amostra, dentro dos dados que continham elementos à esquerda. Consideramos como menção ao sujeito elementos que fizessem referência a ele de forma explícita, como nomes próprios, e alguns substantivos comuns (como mamãe, papai e etc). Além de verificarmos se havia uma menção ou não, verificamos também onde essa menção era feita, se antes do sujeito, numa outra oração ou frase, se era feita antes, na mesma oração, ou ainda se depois do sujeito. Finalmente, cruzamos essas variáveis com o número de ocorrências de nulos.

O esperado era que uma menção explícita ao sujeito aumentasse o número de nulos. E, como pode ser observado na

Figura 6, esse aumento realmente aconteceu, mas só quando a menção era feita depois do sujeito.

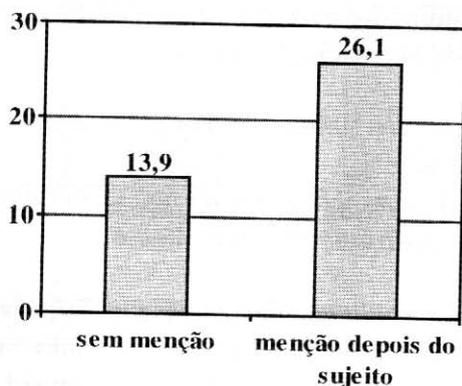


Figura 7: Dados com elemento à esquerda no PB: percentual de nulos em função da menção ao sujeito feita depois do sujeito

Como é possível observar, a taxa de nulos quase dobrou, no PB, com a menção após o sujeito. A taxa de nulos de 13,9% (6 nulos num total de 43 dados), obtida quando não havia menção, passou para 26,1% (12 nulos em um total de 46 dados), quando a menção era feita depois do sujeito. Isso provavelmente se deve ao fato de que, nestes dados com elemento à esquerda, a ligação entre a posição do sujeito e o contexto ficou livre à direita, ou seja, depois do sujeito, possibilitando que este tipo de menção contribuísse para o aumento do número de sujeitos nulos, ao passo que, nas menções feitas antes do sujeito, não houve aumento, já que os elementos à esquerda impediriam o processamento da pista fornecida pelo contexto. Esse resultado favoreceu, assim, a hipótese de que o PB pode estar usando o contexto para licenciar os seus nulos, já que a taxa de nulos aumenta quando a influência do contexto é, de alguma forma, favorecida na oração.

Os exemplos E e F ilustram a tendência geral encontrada nos dados. O exemplo E apresenta um sujeito preenchido, num

dado no qual não havia nenhuma menção ao sujeito. O exemplo F apresenta um sujeito nulo em um dado no qual uma menção ao sujeito era feita à direita. Note-se que ambos os exemplos, como todos os dados utilizados para essa verificação, possuem um elemento à esquerda da posição do sujeito.

Exemplo E – *sem menção*: “... Porque você está perguntando” (tira 456)

Exemplo F – *menção depois do suj.*: “Com quem \emptyset está falando, mamãe?...” (tira 783)

É interessante notar também, na Figura 7, que a influência da menção feita depois do sujeito se mostrou muito mais significativa se comparada aos dados sem elemento à esquerda, mas com o mesmo tipo de menção, obtidos anteriormente na Figura 3.

Nesta etapa da pesquisa não foi feita nenhuma verificação em relação ao PE, visto que este já tinha se mostrado muito avesso à influência do contexto.

5.2.2) Dados com elemento à esquerda do sujeito: vocativos

Ainda abordando os elementos à esquerda, analisamos também um tipo especial de elemento à esquerda, os vocativos. Diferentemente dos outros elementos à esquerda, como os pronomes interrogativos, que se constituíram como barreiras para o contexto, os vocativos deveriam favorecer o contexto. Isso acontece porque, sendo uma pista explícita de quem é o sujeito e estando imediatamente à esquerda do sujeito, o vocativo se torna um elemento que, ao invés de atrapalhar a identificação do sujeito, pode ajudar, aumentando, assim, a ocorrência de nulos. As porcentagens obtidas com este tipo de estrutura estão representadas na Figura 8.

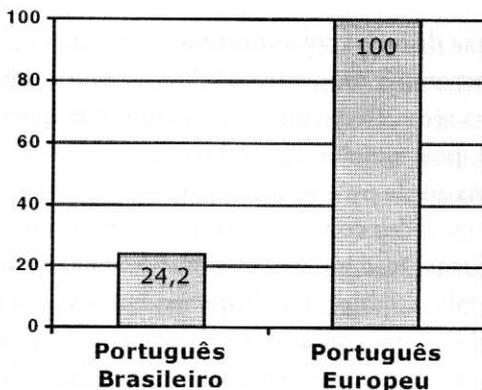


Figura 8: Porcentagem de nulos em função da presença de vocativos à esquerda do sujeito no PB e no PE

Exemplo G – *Versão em PB*: “Pai, você já ganhou medalha por alguma coisa?” (tira 1816)

Exemplo H – *Versão em PE*: “Papá, Ø nunca ganhaste uma medalha?” (tira 1811)

Como é possível observar, a Figura 8 mostra que a taxa de nulos no PB é aproximadamente quatro vezes menor do que a taxa de nulos no PE, que atingiu 100% de nulos (22 nulos em 22 dados) com a presença dos vocativos. Mas o PB, apesar de apresentar a mesma tendência de preenchimento do sujeito, apresentou uma tendência inversa àquela apresentada nos dados anteriores. A expectativa era que, nestes dados, a taxa de nulos fosse maior do que nos dados apresentados anteriormente, pois, se o contexto realmente favorecesse os nulos, seria esperado que um “indicador de contexto” tão próximo à posição de sujeito fizesse a taxa de nulos crescer significativamente, no entanto, não foi isso o que aconteceu. Ao contrário, no PB, a taxa de nulos, que foi de 24,2% (8 nulos num total de 33 dados), sofreu apenas um leve aumento se comparada com a taxa de nulos obtida nos dados com pronome interrogativo, que foi de 19,2%. Se compararmos

essa mesma taxa de nulos com vocativos, com a taxa de nulos dos dados sem elemento à esquerda (43,8%) veremos que a taxa sofreu ainda uma redução, quando o esperado seria que ela fosse no mínimo igual, pois nos dois casos o contexto é favorecido, sendo que talvez mais ainda no caso do vocativo.

Assim, os dados com vocativo mostraram um quadro desfavorável à afirmação de que o contexto licencia o nulo no PB, favorecendo, talvez, a hipótese de que os nulos sejam apenas resquícios de uma gramática antiga e não possuam mais nenhum mecanismo que os mantenha vivos na língua. Essa afirmação não pode, no entanto, ser feita sem receio, pois, devido à especificidade desta última estrutura, obtivemos, aqui, um número de dados muito menor do que o número obtido com as outras estruturas. Para que se observe melhor a comportamento deste tipo de estrutura é necessário que pesquisas com *corpora* maiores sejam feitas. O que se pode reafirmar com estes dados é a tendência do PB em preencher sempre os seus sujeitos, mesmo em ocasiões favoráveis ao nulo, como esta.

6) Considerações finais

Não é possível afirmar categoricamente, a partir da presente pesquisa, que o PB tenha-se modificado estruturalmente a tal ponto que pudesse ser considerado uma nova língua, sem que debates teóricos sejam levantados. Esta pesquisa foi realizada no intuito de contribuir com algumas novas informações que ajudassem a dissolver esse impasse, sem, contudo, propor-se a resolvê-lo definitivamente.

É inegável que a redução do paradigma flexional do PB, no último século, causou uma série de mudanças morfológicas e estruturais, como a preferência por se preencher o sujeito pronominal, entre outras, e isso o teria afastado de sua língua de origem, o PE. Sem que houvesse pesquisas e teóricos interessados na investigação desse fato, dificilmente seria possível chegar

a conclusões como as alcançadas aqui, firmemente amparadas por um arcabouço teórico e pela comprovação empírica, de que entre ser uma língua igual ao PE, na qual as diferenças seriam meramente dialetais, e ter se tornado uma nova língua totalmente independente, o PB se encontra muito mais próxima do último que do primeiro nível.

Os resultados obtidos nesta pesquisa indicam uma mudança paramétrica que traria uma resposta a respeito do PB como uma língua diferente do PE, já que esses resultados mostram que o PB, diferentemente do PE, não utiliza as desinências para identificar e, assim, licenciar os sujeitos pronominais nulos. Portanto, os resultados, mostrando mecanismos diferentes de identificação do nulo, parecem evidenciar que estas línguas não são a mesma.

É importante notar que o problema causado por resultados obtidos ao se estudar o PB, que podem não fornecer indícios suficientes a respeito de uma mudança total do Parâmetro do Sujeito (já que ainda existem sujeitos nulos), além de dificultar a delimitação que separa o português do “brasileiro”, também geram um desafio para a teoria, segundo a qual um parâmetro só pode ser marcado ou positiva ou negativamente, sem estágios intermediários. O PB é, nesse sentido, uma língua com particularidades que sugerem uma necessidade de uma possível reformulação deste parâmetro, na qual a marcação binária não estivesse ligada à questão do sujeito ser preferencialmente nulo ou pleno, mas à outra característica.

Independente disso, mesmo tendo cada estrutura estudada indicado mecanismos diferentes de licenciamento dos sujeitos nulos ainda existentes em quantidade relativamente elevada no PB, se comparado a uma língua não *pro-drop*, e tendo esse fato nos impedido de chegar a uma conclusão definitiva a respeito de qual das hipóteses testadas seria a mais condizente com a realidade, não se pode negar que o mecanismo é diferente do utilizado pelo PE e que, assim sendo, o Português brasileiro é muito mais que uma mera variante do Português europeu.

7) Referências

CHOMSKY, N. *Syntactic structures*. The Hague: Mouton, 1957.

_____. *Lectures on government and binding*. Dordrecht: Foris, 1981.

DUARTE, M. E. L. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In: ROBERTS, I.; KATO, M. A. (Org.). *Português Brasileiro: uma viagem diacrônica*. Campinas: EdUNICAMP, 1993, p. 107-128.

_____. *A perda do princípio "Evite Pronome" no português brasileiro*. 1995. Tese (Doutorado em Linguística) – Instituto de Linguística, Universidade Estadual de Campinas, Campinas.

FIGUEIREDO SILVA, M. *A posição sujeito no Português Brasileiro – frases finitas e infinitivas*. Campinas: EdUNICAMP, 1996.

KATO, M.; NEGÃO, E. (Orgs.). *Brazilian Portuguese and the null subject parameter*. Madrid: Iberoamericana, 2000.

KATO, M. A. A gramática do letrado: questões para a teoria gramatical. In: MARQUES, M. A.; KOLLER, E.; TEIXEIRA, J.; LEMOS, A. S. (Org.). *Ciências da Linguagem: 30 anos de investigação e ensino*. Curitiba: Braga, 2005, p. 131-145.

¹ Esta pesquisa faz parte do projeto de estágio voluntário "Sujeito nulo no Português Brasileiro: interrogativas e encaixadas no *corpus* Mafalda", realizado nos anos de 2004-2005.

² Apesar de, no inglês, existirem alguns contextos nos quais os nulos são possíveis, esses contextos são bastante restritos. "Woke up late", "Drank a cup of milk", "Left in a hurry", orações de sujeito nulo, são exemplos do que se chama *diary-drop*, uma espécie de registro usado em diários. Este é um contexto no qual os nulos são frequentes, mas trata-se de um contexto muito restrito, particular.