

SOBRE O ALCANCE DA SOCIOLINGUÍSTICA NO ESTUDO DA MUDANÇA PARAMÉTRICA: UMA PERSPECTIVA INTERLINGUÍSTICA

ON THE RANGE OF SOCIOLINGUISTICS IN THE STUDY OF PARAMETRIC CHANGE: AN INTERLINGUISTIC PERSPECTIVE

Humberto Soares da Silva

Professor da Universidade Federal do Rio de Janeiro

humba6@gmail.com

RESUMO: Comparam-se resultados de análises de regra variável sobre a representação do sujeito pronominal de referências definida no italiano, no espanhol europeu, argentino e porto-riquenho e no português brasileiro. Embora conduzidos por diferentes pesquisadores, esses trabalhos podem ser comparados porque têm metodologias semelhantes e partem dos mesmos pressupostos teóricos: a Gramática Gerativa e a Sociolinguística, numa associação denominada Variação Paramétrica. Assume-se a visão de parâmetro como uma entidade contínua, havendo infinitas possibilidades de comportamento que pode ser apresentado por diferentes línguas em relação ao fenômeno. Com base na comparação das análises, as línguas são ordenadas segundo diferentes aspectos: (a) as taxas gerais de sujeitos nulos encontradas, (b) a intensidade da atuação do grupo de fatores ‘condição de referência’, relacionado à acessibilidade de um sujeito ao seu antecedente, (c) a estabilidade, “medida” pela quantidade de grupos de fatores apontados como condicionamentos para a seleção entre o apagamento e o preenchimento do pronome e (d) o grau de riqueza dos paradigmas flexionais verbais, considerando apenas a oposição entre as três pessoas gramaticais. A ordenação das línguas segundo cada um desses critérios coincide, mas isso não deve ser interpretado como casual ou aleatório – ao contrário, indica que todos esses aspectos estão relacionados à capacidade, na língua, do núcleo I de identificar sujeitos nulos, sendo propriedades de um mesmo parâmetro.

PALAVRAS-CHAVE: Espanhol; Italiano; Português brasileiro; Parâmetro; Sujeito nulo.

ABSTRACT: This work compares results of variable rule analyzes on the representation of the subject pronoun with definite reference in Italian, Spanish (European, Argentinian and Puerto Rican), and Brazilian Portuguese. Although conducted by different researchers, these works can be compared because they have similar methodologies and depart from the same theoretical assumptions: Generative Grammar and Sociolinguistics, an association called “Parametric Variation”. I assume a parameter to be a continuous entity, with endless possibilities of behavior that may be shown by different languages regarding the phenomenon. Based on the comparison of the analyzes, the languages are ordered according to different aspects: (a) the overall rates of null subjects found, (b) the intensity of the effects of the factor group ‘reference condition’, which concerns the accessibility of a subject to its antecedent, (c) the stability, “measured” by the quantity of factor groups identified as constraints to the selection between the deletion and the filling of the pronoun, and (d) the degree of richness of verbal inflectional paradigms, considering only the opposition between the

three grammatical persons. The ranking of languages according to each of these criteria coincide, but this should not be interpreted as casual or random – instead, it indicates that all these aspects are related to the ability, in a language, of the “head I” to identify null subjects, i.e., properties of the same parameter.

KEYWORDS: Spanish; Italian; Brazilian Portuguese; Parameter; Null subject.

INTRODUÇÃO

Este artigo apresenta uma comparação entre diferentes análises de regra variável relacionadas ao Parâmetro do Sujeito Nulo, mais especificamente no que se refere à identificação de sujeitos nulos pelo núcleo flexional I, no italiano, no espanhol europeu, argentino e porto-riquenho e no português brasileiro (estas duas últimas línguas¹, em processo de mudança paramétrica em direção ao preenchimento do sujeito). Embora conduzidos por diferentes pesquisadores, os trabalhos podem ser comparados porque seguem metodologias bastante semelhantes e se apoiam nos mesmos pressupostos teóricos, associando gerativismo e variacionismo, uma aproximação denominada Variação Paramétrica por Ramos (1999). O objetivo é demonstrar que a produtividade de sujeitos nulos nessas línguas é diretamente proporcional ao grau de riqueza de oposições entre as três pessoas gramaticais nos paradigmas verbais: quanto mais rico é o paradigma, maior é a probabilidade de o núcleo I identificar a referência dos sujeitos nulos.

Para isso, observa-se (a) a estabilidade das línguas em relação ao fenômeno, pela observação da quantidade de condicionamentos relacionados ao apagamento do pronome sujeito referencial e (b) a atuação da maior ou menor acessibilidade do sujeito em análise a um antecedente. Para a mensuração do grau de riqueza flexional, propõe-se uma maneira mais refinada de observar a existência de oposições entre as pessoas gramaticais nos paradigmas de conjugação verbal. Assumindo que parâmetros são escalas contínuas, e não dicotomias discretas, as cinco línguas comparadas são ranqueadas segundo a estabilidade, a atuação da acessibilidade e riqueza flexional – e conclui-se que a ordenação dessas línguas segundo esses aspectos é a mesma e, não coincidentemente, é idêntica à ordenação das mesmas pelas frequências gerais de sujeitos nulos.

O trabalho começa com considerações sobre os pressupostos teórico-metodológicos utilizados. Na Seção 2, apresenta-se a metodologia das análises de regra variável aplicadas às variedades do espanhol, que não se difere consideravelmente das metodologias das análises aplicadas às outras línguas por Duarte (1995) e Marins (2009), e em 3 apresento uma comparação de todas essas análises, com foco maior nos resultados referentes ao grupo de fatores ‘condição de referência’, que se refere à acessibilidade a um antecedente. Em 4, faço considerações sobre a estabilidade dessas línguas quanto à capacidade de I de identificar sujeitos nulos e, na seção 5, correlaciono

¹ Italiano, espanhol europeu, espanhol argentino, espanhol porto-riquenho e português brasileiro, neste estudo, são considerados diferentes línguas, porque o que está em foco são os diferentes comportamentos que exibem em relação ao Parâmetro do Sujeito Nulo. Apenas por uma questão de simplificação do texto, será usada a expressão “variedades do espanhol” quando for feita menção, em conjunto, ao espanhol europeu, ao argentino e ao porto-riquenho. Ainda assim, estas são consideradas três línguas distintas.

isso com a riqueza flexional que cada língua apresenta.

2 PRESSUPOSTOS

O trabalho foi realizado com base em pressupostos variacionistas e gerativistas. Essa associação foi apresentada como uma aproximação de duas correntes teóricas antes consideradas incompatíveis, com o nome de Sociolinguística Paramétrica, por Tarallo (1987), Tarallo e Kato (1989) – e produções posteriores dos mesmos autores. Podem-se citar, ainda, outros trabalhos que, assumindo a ideia de Tarallo e Kato, foram norteados pela Sociolinguística Paramétrica, ainda apresentada como uma união não tão natural, como os de Moino (1987), Nunes (1990), Pagotto (1992), Ramos (1992) e Cyrino (1994).

Críticas à Sociolinguística Paramétrica emergiram, apontando a utilização simultânea de duas correntes teóricas opostas – cf., por exemplo, os ensaios de Borges Neto (2004). Acontece que a Sociolinguística é um modelo de estudo da variação e da mudança, e não prescinde de uma teoria linguística, como a da Gramática Gerativa, o que permite a utilização da Sociolinguística Paramétrica, como nos trabalhos de Kato (2000), Duarte (2004), Barbosa, Duarte e Kato (2005), Kato e Duarte (2005) e Soares da Silva (2006, 2011) – que mostram que pressupostos gerativistas ajudam a estabelecer hipóteses e grupos de fatores para a análise variacionista, ao mesmo tempo em que os resultados dessas análises ajudam a atualizar assunções da teoria gerativa e a definir as propriedades dos parâmetros estudados. Neste trabalho, como o que está em foco é a observação da variação e da mudança e a conceituação de *parâmetro*, em vez de Sociolinguística Paramétrica, adoto o rótulo Variação Paramétrica, proposto por Ramos (1999).

Assim, além de ser possível, para a análise pretendida aqui, a utilização da Variação Paramétrica é necessária. Estudos de base gerativista costumam usar dados provenientes da intuição de falantes², que não é totalmente confiável (cf. CHOMSKY, 1964). Em testes de percepção e de produção, falantes frequentemente afirmam usar o que efetivamente não usam; outras vezes, não percebem estruturas que conscientemente rejeitam.

Portanto, é desejável fazer uma análise com base em dados reais de fala. Por outro lado, uma análise apenas baseada em dados reais não elícita dados agramaticais, importantes para a descrição de um sistema – e, para obter dados agramaticais, precisa-se da intuição. Por isso, esta pesquisa, como a de Kato e Nascimento (2009), recorre à evidência positiva dos dados gramaticais coletados de fala real e à evidência negativa através de dados agramaticais produzidos pela intuição linguística.

A consequência de observar e comparar análises de regra variável é a dificuldade de classificar todas as línguas naturais por meio das discretas marcações positiva e negativa do Parâmetro do Sujeito Nulo. Na verdade, cada língua analisada apresentou um comportamento diferenciado para as propriedades relacionadas ao parâmetro. Por isso,

² Há, porém, diversos estudos realizados sob uma perspectiva gerativa com análise de dados reais, inclusive alguns que incluem observação quantitativa. Para um exemplo, consulte-se Novaes (1996).

adota-se a escala contínua de Soares da Silva (2011), na qual as línguas podem ser localizadas, segundo o comportamento exibido, entre os extremos que representam os comportamentos prototípicos [- sujeito nulo] e [+ sujeito nulo]³:

(1) [- sujeito nulo] |—————| [+ sujeito nulo]

A visão contínua do Parâmetro do Sujeito Nulo diverge do que sempre pregou a teoria gerativa. Mas a noção de parâmetro não foi totalmente abandonada. As propriedades associadas às línguas de sujeito nulo (e ausentes nas línguas de sujeito pleno), como, por exemplo, a possibilidade de inversão do sujeito e a subida de clítico, são mantidas, mas propõe-se observar as frequências e as tendências, que podem revelar que, em relação ao parâmetro analisado aqui, uma língua X pode ser “mais de sujeito nulo” do que uma língua Y (e que as línguas X e Y podem ser “menos de sujeito nulo” do que uma língua Z), quando duas ou mais línguas são comparadas.

Sobre isso, Tarallo e Kato (1989, p. 20) já afirmavam: “a linguística das propriedades paramétricas parece agir no sentido do TUDO ou NADA; a das probabilidades, na direção do MAIS ou MENOS” (grifos do original). Os autores dão a entender que a visão contínua pode ser uma forma de explicar e interpretar qualquer parâmetro ou fenômeno, não ficando restrita à realização do sujeito. Mais adiante, exemplificam (TARALLO; KATO, 1989, p. 24):

Três línguas podem ser agrupadas como pertencentes a um mesmo parâmetro por compartilharem uma mesma propriedade, mas a abordagem quantitativa poderá aproximar duas delas contra a outra em função do grau de incidência de um fenômeno.⁴

Ao descrever diferenças entre línguas ou a mudança de acordo com o modelo varacionista, é possível chegar a uma descrição completa de cada língua analisada, fazer uma comparação entre elas e até observar cada passo da mudança observada, porém o trabalho não será muito mais do que uma descrição – que não responde, totalmente, às questões propostas por Heinrich *et al.*, (1968). O gerativismo é adotado, pois, neste trabalho, como a teoria linguística necessária para tratar do fenômeno em pauta. A necessidade de uma teoria linguística já era apontada por Weinreich *et al.*, (1968), em sua Teoria da Variação e Mudança.

3 METODOLOGIA DE ANÁLISE DOS DADOS DO ESPANHOL

Os dados do espanhol europeu, do argentino e do porto-riquenho provêm das amostras de fala culta⁵ das cidades de Madri, Buenos Aires e San Juan, respectivamente, que

³ Ao analisar um parâmetro considerando a variação, é difícil tratá-lo como binário e discreto, já que a variação é contínua. Mesmo assim, há uma polaridade, representando binariamente os comportamentos mais prototípicos que as línguas podem apresentar, localizados nos limites do segmento de reta proposto.

⁴ O texto citado leva a crer que, em lugar de “como pertencentes a um mesmo parâmetro”, se deva ler: “como apresentando a mesma marcação (ou o mesmo comportamento) em relação a um parâmetro”.

⁵ A “fala culta” refere-se à fala dos indivíduos que cursavam ou já haviam cursado o nível superior.

fazem parte do *Macrocorpus de la norma lingüística culta de las principales ciudades de España y América* (SAMPER PADILLA *et al.*, 1995). O *Macrocorpus* foi gravado nos anos 1970 e contém amostras da fala culta do espanhol de 12 cidades (nove americanas e três espanholas), cada uma composta de 14 entrevistas com aproximadamente meia hora de duração cada uma. Os falantes de cada localidade estão estratificados em gêneros e três faixas etárias:

Quadro 1 Distribuição comum das entrevistas na amostra de cada cidade do *Macrocorpus*

FAIXA ETÁRIA	HOMENS	MULHERES	TOTAL (14)
DE 20 A 35 ANOS	2 entrevistas	2 entrevistas	4 entrevistas
DE 36 A 54 ANOS	3 entrevistas	3 entrevistas	6 entrevistas
MAIS DE 55 ANOS	2 entrevistas	2 entrevistas	4 entrevistas

Fonte: Autor

Vê-se, no Quadro 1, que há mais informantes na segunda faixa etária. Para equilibrar a quantidade de falantes de cada célula, foram desconsideradas duas entrevistas da faixa etária central (uma com homem e uma com mulher) de cada cidade. A distribuição dos informantes do espanhol europeu, do argentino e do porto-riquenho – identificados pela sigla da língua e numerados de 01 a 14 –, então, pode ser vista no Quadro 2.

Quadro 2 Inquéritos usados para as análises do espanhol europeu, do argentino e do porto-riquenho

CIDADE	FAIXA ETÁRIA	HOMENS	MULHERES
MADRI, ESPANHA – ESPAÑHOL EUROPEU (EE)	DE 20 A 35 ANOS	EE01 e EE02	EE03 e EE04
	DE 36 A 54 ANOS	EE06 e EE07	EE09 e EE10
	MAIS DE 55 ANOS	EE11 e EE12	EE13 e EE14
BUENOS AIRES, ARGENTINA – ESPAÑHOL ARGENTINO (EA)	DE 20 A 35 ANOS	EA01 e EA02	EA03 e EA04
	DE 36 A 54 ANOS	EA05 e EA07	EA09 e EA10
	MAIS DE 55 ANOS	EA11 e EA12	EA13 e EA14
SAN JUAN, PORTO RICO – ESPAÑHOL PORTO-RIQUENHO (EP)	DE 20 A 35 ANOS	EP01 e EP02	EP03 e EP04
	DE 36 A 54 ANOS	EP06 e EP07	EP08 e EP10
	MAIS DE 55 ANOS	EP05 e EP12	EP13 e EP14

Fonte: Autor

Foi definida uma quantidade máxima de 50 dados do informante de cada entrevista para cada uma das três pessoas gramaticais: a coleta de dados, para cada pessoa gramatical, foi feita considerando-se o trecho que vai do início da entrevista até o 50º dado do entrevistado (nos casos em que não houve 50 dados de alguma pessoa gramatical em determinada entrevista, foram computados todos os dados encontrados). Para possibilitar a análise das orações interrogativas e dos pronomes de segunda pessoa, foram considerados também dados da fala dos entrevistadores, coletando, de cada pessoa gramatical, os dados localizados no mesmo trecho da entrevista de onde já haviam sido coletados os dados do informante. A quantidade de dados levantados é a seguinte (em cada campo, o primeiro número refere-se à quantidade de dados retirados da fala dos entrevistados e o segundo inclui os dados dos entrevistadores)⁶:

⁶ Os números apresentados aqui se referem aos dados coletados no total, inclusive os que foram posteriormente excluídos para permitir a realização das análises de regra variável, conforme as justificativas que serão apresentadas. Confrontem-se esses números com os da tabela de resultados.

Tabela 1 Distribuição dos dados coletados da fala dos informantes e dos entrevistadores

PESSOA GRAMATICAL	PRIMEIRA	SEGUNDA	TERCEIRA	TOTAL
Espanhol Europeu	597; 646	68; 366	482; 582	1147; 1473
Espanhol Argentino	548; 600	42; 300	397; 520	987; 1393
Espanhol Porto-Riquenho	477; 485	40; 83	396; 411	894; 964
TOTAL	1622; 1731	150; 749	1275; 1513	3047; 3993

Fonte: Autor

A inclusão da fala dos entrevistadores, dos quais não estão disponíveis informações como idade ou gênero, poderia distorcer os resultados da pesquisa. Porém, como só é analisada a fala culta, e como os entrevistadores são falantes cultos e pertencem à mesma comunidade de fala de seus entrevistados⁷, imagina-se que a inclusão dos seus dados não prejudique muito significativamente os resultados e as generalizações. Os dados dos entrevistadores, porém, não se aplicam aos grupos de fatores sociais.

Foram coletados dados com sujeitos pronominais, nulos e plenos, com referência definida e em orações como verbo finito. Sujeitos que recebem Caso diferente de nominativo, como os de miniorações sem verbo, como exemplificado em (2), ou de orações infinitivas, como em (3), foram desconsiderados. Não foram coletados, também, os tipos de dados listados em (4), a seguir.

- (2) *Lo veo reflejado también en los salones de clase; el maestro...* (EP10)
Vejo-o refletido também nas salas de aula; o professor...
- (3) *De ahí decidí [volver ____ a Puerto Rico]* (EP02)
Daí decidi voltar para Porto Rico
- (4) a. Sujeito relativizado, representado por um relativo:
Hay algunos_i [que_i han t_j tenido contacto com el arte en otros...] (EP03)
Há alguns [estudantes] que tiveram contato com a arte em outros...

Em orações como a exemplificada acima, o sujeito, por ser relativizado, se move obrigatoriamente para specCP, deixando sempre um vestígio de movimento na posição de sujeito da oração relativa. Não há, portanto, variação quanto à representação do sujeito nesse contexto. O mesmo se pode dizer de dados como o exemplificado em (4b), abaixo, que também foram desconsiderados.

- b. Sujeito representado por um interrogativo:
no se sabe exactamente [qué_i ocurre t_i] (EP01)
não se sabe exatamente o que acontece
- c. Sujeito de uma oração coordenada que não seja a primeira da sequência:
Terminé mi cuarto año y empecé ____ a buscar trabajo. (EP02)
Terminei o quarto ano e comecei a procurar trabalho.
- d. Sujeito de primeira menção ao referente na entrevista

⁷ A exceção é a amostra de Porto Rico, que contém inquéritos conduzidos por espanhóis. Os dados desses entrevistadores não foram considerados. Assim, a quantidade total de ocorrências de segunda pessoa coletadas da fala porto-riquenha é bem menor do que a das outras duas variedades do espanhol.

Os sujeitos de primeira menção são majoritariamente expressos, já que há uma tendência ao preenchimento quando um novo referente é inserido no discurso (cf. DURANTI e OCHS, 1979). A inclusão dos sujeitos de primeira menção no envelope de variação reduziria as taxas de sujeitos nulos a ponto de ofuscar a atuação dos grupos de fatores controlados. Por esse motivo, só são considerados dados que contenham sujeito pronominal, nulo ou pleno, de retomada (que tenha um antecedente, mesmo que distante no discurso).

e. Sujeito em marcador discursivo cristalizado, como, por exemplo, em *tu sabes*

Expressões como a exemplificada acima (que pode ser traduzida para **você sabe**) são formas cristalizadas de um idioma. Não acompanham a mudança paramétrica quando a língua passa por uma. Tais dados são desconsiderados porque mascarariam os resultados, levando a falsas descrições e conclusões sobre a língua, como ocorre no trabalho de Hochberg (1986), conforme as críticas de Flores-Ferrán (2007).

As ocorrências coletadas foram codificadas segundo duas variáveis sociais (gênero e faixa etária) e 12 variáveis linguísticas. O Quadro 3 lista os grupos de fatores selecionados nas análises probabilísticas, na ordem de relevância apontada pelo VARBRUL. O grupo de fatores que apresentou maior atuação foi a condição de referência.

Quadro 3 Grupos de fatores selecionados nas análises das três variedades do espanhol (aplicação: sujeito nulo)

	Espanhol europeu	Espanhol argentino	Espanhol porto-riquenho
1	Pessoa gramatical	Condição de referência	Condição de referência
2	Condição de referência	Pessoa gramatical	Tipo de oração
3	Tipo de oração	Faixa etária do informante	Tempo e modo verbais
4	Faixa etária do informante	Elemento adjunto a IP	Pessoa gramatical
5	Elemento entre specIP e V	Elemento entre specIP e V	Gênero do informante
6	Forma verbal	Função sintática da oração	Elemento entre specIP e V
7		Estrutura de CP	Elemento adjunto a IP
8		Tipo de desinência	Faixa etária do informante
9		Gênero do informante	
	<i>input: 0,76</i>	<i>input: 0,69</i>	<i>input: 0,61</i>

Fonte: Autor

Passemos às análises de regra variável das variedades do espanhol. A essas análises, comparo resultados de estudos anteriores sobre o português e o italiano. A seção seguinte será mais dedicada aos resultados referentes à condição de referência⁸.

4 ANÁLISE COMPARATIVA

A porcentagem geral de sujeitos nulos no espanhol porto-riquenho é intermediária: 59%, que é uma taxa menor do que as encontradas para o espanhol europeu (76%) e o argentino (68%), mas ainda assim revela uma preferência geral pelo sujeito nulo sobre o pleno. Isso ocorre porque a língua não está deixando de ser de sujeito nulo (o sujeito

⁸ Resultados sobre os outros grupos de fatores controlados nas análises dessas línguas, inclusive os não selecionados, e a interpretação desses resultados podem ser consultados em Soares da Silva (2011).

nulo é licenciado pelo paradigma flexional), mas está perdendo uma das formas pelas quais a referência de um sujeito nulo é identificada. É como se uma língua de sujeito nulo cujo núcleo I não é capaz de identificar a referência dos sujeitos nulos sofresse uma dupla pressão: como língua de sujeito nulo, os sujeitos *têm que* ser nulos, mas, sem a referência identificada, os sujeitos *não podem* ser nulos, ou seja, o sujeito tende a ser nulo sempre que a sua referência puder ser facilmente identificada, e a ser pleno quando há maior dificuldade na identificação.

Quanto à animacidade do sujeito, os resultados para as três variedades são categóricos: em todos os dados de sujeito com o traço [- animado] o sujeito é nulo, como exemplifica (5). Esse é o resultado esperado para uma língua de sujeito nulo, portanto é mais uma confirmação de que a mudança no espanhol porto-riquenho não é no sentido de deixar de ser uma língua de sujeito nulo: continua sendo uma, como o espanhol europeu e o argentino. Os sujeitos inanimados foram excluídos para que pudessem ser calculados os pesos relativos e feita a análise, e, por esse motivo, o grupo de fatores animacidade não aparece no Quadro 3.

- (5) a. *aunque originalmente [Federación Universitaria Pro Independencia] tenía _____ mayor independencia del movimiento pro independencia* (EP01)
embora originalmente tivesse maior independência do movimento pró-independência
b. *que [la ciencia] obtiene _____ por los nuevos medios de investigación* (EP06)
que obtém pelos novo meios de comunicação

Marins (2009, p. 91) observou o mesmo resultado no italiano, ao perceber que todos os 117 sujeitos com o traço [- animado] coletados são nulos; no português europeu, o resultado não é categórico, mas é quase categórico, já que 94% dos sujeitos inanimados são apagados (DUARTE, 1995). Isso aproxima o espanhol europeu, o argentino, o porto-riquenho e o português europeu, que manifestam comportamentos muito semelhantes como línguas [+ sujeito nulo], pelo menos no que se refere à representação dos sujeitos inanimados. No português brasileiro, por outro lado, a preferência pelo preenchimento já atinge os sujeitos com o traço [- animado], visto que 44% destes são nulos (DUARTE, 1995), o que é uma das evidências de que tal língua está deixando de ser [+ sujeito nulo]⁹.

Para a condição de referência, observa-se a função sintática do antecedente do pronome em análise (o último elemento, anterior, correferente a ele). Foram usados os quatro padrões sentenciais propostos por Barbosa, Duarte e Kato (2005), além de um quinto: o padrão E¹⁰. Segue a lista de definições de cada padrão, acompanhadas de exemplos do espanhol porto-riquenho:

- (6) a. **Padrão A:** o sujeito em análise está numa oração subordinada e o seu antecedente

⁹ Duarte (1993, 1995) defende a ocorrência de mudança paramétrica no português brasileiro, justificada, principalmente, pelas continuamente decrescentes taxas de apagamento do sujeito, em análise diacrônica. Consulte-se, também, Novaes (1996), que defende que a mesma língua se mantém [+ sujeito nulo] – provavelmente, baseado na Teoria do Subconjunto, apresentada por Raposo (1992).

¹⁰ Exemplos dos padrões D e E, coletados para esta análise do espanhol, seriam, se fosse seguida a metodologia de Barbosa, Duarte e Kato (2005), codificados como padrão IV, que reúne os dados de sujeitos com antecedente distante – na presença de oração(ões) interveniente(s). A diferença entre os padrões D e E reside no fato de o antecedente ter ou não, também, a função sintática de sujeito.

- é o sujeito da respectiva oração matriz (c-comandado pelo antecedente)
yo_i bromeo [porque digo _____i [que yo_i hice una concentración de [...]]] (EP02)
 eu brinco porque digo que eu fiz uma concentração de [...]
- b. **Padrão B:** o antecedente do pronome em análise tem a mesma função sintática, a de sujeito, na oração imediatamente anterior, sem haver relação de c-comando
yo bromeo porque digo que yo_i hice una concentración de idiomas extranjeros estando en el Colegio de Comercio. Estudié _____i dos años de francés (EP02)
 eu brinco porque digo que eu fiz uma concentração de idiomas estrangeiros estando no Colegio de Comercio. Estudei dois anos de francês
- c. **Padrão C:** o antecedente está na oração anterior, mas com outra função sintática
Esto me_i causó problemas, el primer año, en cuanto a los estudios; fracasé _____i en dos cursos; pero poco a poco me fui acostumbrando al uso del idioma (EP02)
 Isso me causou problemas, no primeiro ano, quanto aos estudos; fracassei em dois cursos; mas pouco a pouco fui me acostumando ao uso do idioma
- d. **Padrão D:** há pelo menos uma oração interveniente entre a oração em que se encontra o pronome em análise e a oração em que está o antecedente, que é sujeito
que yo_i era estudiante [que esa era la gran actividad de la universidad]. De ahí, en el sesenta y dos, comienza todo el movimiento de reforma. Yo_i llegué a la (EP01)
 que eu era estudante, quando essa [a FUPI] era a grande atividade da universidade. Daí, em [19]62, começa todo o movimento de reforma. Eu cheguei à
- e. **Padrão E:** há pelo menos uma oração interveniente entre a oração em que se encontra o pronome em análise e a do antecedente, que tem função diferente
pero lo que usted me_i estaba preguntando [de que aquí no hay... mucho material], verdaderamente aquí lo que hay es un buen museo para exhibir las cosas que hay. [...] Yo_i no sé si usted há estado en el Museo de Antropología de México. (EP04)
 mas o que você estava me perguntando de que aqui não tem muito material, verdadeiramente o que há aqui é um bom museu para exibir as coisas que há. Eu não sei se você já esteve no Museo de Antropología de México.

O resultado é gradual nas variedades do espanhol, com os primeiros padrões favorecendo o sujeito nulo e os últimos, desfavorecendo-o. Na tabela a seguir, apresento os resultados deste grupo de fatores para o espanhol e o italiano, facilitando a comparação. Os resultados da análise do italiano (MARINS, 2009, p. 95) mostram que a atuação da condição de referência na representação do sujeito referencial não é tão universal como se poderia supor comparando os resultados para as três variedades do espanhol presentes na tabela e os resultados para o português de Barbosa, Duarte e Kato (2005) e para o espanhol de Cameron (1993).

Tabela 2 O padrão sentencial no espanhol europeu, no argentino, no porto-riquenho e no italiano

	Espanhol europeu ¹¹			Espanhol argentino			Esp. p.-riquenho			Italiano ¹²		
	ocs.	%	PR	ocs.	%	PR	ocs.	%	PR	ocs.	%	PR
A	142/162	88	0,71	156/186	82	0,67	108/124	87	0,86	49/54	91	0,60
B	290/342	85	0,63	298/373	80	0,65	135/199	68	0,58	200/217	92	0,65
C	94/119	79	0,51	80/114	70	0,50	54/90	60	0,51	47/51	94	0,68
D	321/464	69	0,39	242/437	55	0,35	159/340	47	0,34	285/365	78	0,40
E	90/151	60	0,30	61/111	55	0,32	47/103	46	0,34	25/30	83	0,38
	937/1238	76		834/1221	68		503/856	59		630/751	84	

¹¹ Resultados para o espanhol europeu de Soares da Silva (2006).

¹² Marins (2009) acrescentou um padrão sentencial F: primeira menção. Omiti os resultados desse padrão porque não podem ser comparados com o espanhol, já que não foram coletados sujeitos de primeira menção. Todos os resultados sobre o italiano apresentados neste artigo foram retirados de Marins (2009).

Fonte: Autor

Começamos analisando o italiano, que apresentou o resultado mais discrepante dentre todas as línguas comparadas. Apesar de não seguir uma gradação uniforme ao longo dos cinco padrões, há uma diferenciação clara entre os padrões A, B e C, de um lado, e D e E, de outro. Os primeiros três padrões exibem frequências de sujeitos nulos entre 91% e 94%, uma diferença de apenas três pontos percentuais, que é estatisticamente irrelevante, além de mostrar favorecimento do sujeito nulo, com todos os pesos relativos na casa dos 0,60; os dois últimos padrões também exibem percentuais próximos (78% e 83%, respectivamente, com uma diferença desprezível de cinco pontos percentuais) e pesos relativos que indicam desfavorecimento do sujeito nulo (embora a preferência pelo sujeito nulo se manifeste em todos os contextos controlados), respectivamente, 0,40 e 0,38.

Isso mostra que o italiano é a língua que tem o comportamento mais prototípico de língua de sujeito nulo, em que o núcleo flexional I tende mais a ser capaz de identificar a referência dos sujeitos nulos¹³. Parece que este grupo de fatores foi selecionado apenas pela distância entre o pronome em análise e seu antecedente: no italiano, o pronome nulo é favorecido quando seu antecedente está mais distante no discurso – quando há oração(ões) interveniente(s) – o que é representado pelos padrões D e E, e favorecido se o antecedente estiver mais próximo (na mesma sentença ou na oração imediatamente anterior), ou seja, a existência ou a ausência de relação de c-comando para a identificação do sujeito nulo por um antecedente e o fato de o antecedente ter ou não a mesma função sintática não influem nos resultados. O comportamento prototípico do italiano, em comparação com as outras línguas, é reforçado pelo fato de a preferência pelo sujeito nulo ser muito mais intensa: a menor frequência de sujeitos nulos no italiano, na Tabela 2, é de 78%, enquanto as outras três línguas exibem, no padrão sentencial E, frequências de 60% ou até mais baixas.

Comparando os resultados do espanhol europeu e do argentino percebe-se que as duas gramáticas são iguais quanto à atuação das condições de referência, se for adotada a mesma interpretação que Cameron (1993) deu aos seus resultados. Os pesos relativos encontrados para cada padrão, observando a tabela horizontalmente, são muito semelhantes entre o espanhol europeu e o argentino: a maior diferença entre os pesos é a que se vê nos padrões A e D, que afastam as duas variedades por apenas 0,04. A gradação percentual e de pesos, ao longo dos cinco padrões, mostra que o comportamento do espanhol europeu e do argentino quanto à identificação de sujeitos nulos não é tão prototípico, já que todos os cinco fatores do grupo influem na representação do sujeito, o que indica que, além da distância do antecedente, também atuam a existência de uma relação de c-comando e a coincidência ou não coincidência de função.

No espanhol porto-riquenho, o único peso relativo que se diferencia consideravelmente do que foi encontrado para as outras variedades do espanhol é o do padrão sentencial A,

¹³ Consulte-se Toribio (1994), que mostra que esse núcleo funcional tem duas propriedades: I [+ lexical] licencia sujeitos nulos e I [- lexical] não licencia, ao passo que I [+ pronominal] é capaz de identificar a referência dos sujeitos nulos, enquanto um I [- pronominal] não tem essa capacidade.

que também se afasta bastante dos pesos relativos encontrados para os outros padrões. Esse é o padrão sentencial em que o antecedente é mais acessível para a identificação do sujeito nulo, pois tem a mesma função sintática e há uma relação de c-comando dentro da mesma sentença, sem a necessidade de um operador discursivo – numa língua em mudança em direção à perda dos sujeitos referenciais nulos, é esse o padrão sentencial que se espera que seja o mais resistente à mudança. Além disso, numa língua em mudança em direção aos sujeitos plenos, estes passam a ser preferidos em alguns contextos em dado momento do progresso da mudança: esse momento já aconteceu no espanhol porto-riquenho, visto que já há preferência pelos sujeitos plenos nos padrões D e E, que apresentam frequências de sujeitos nulos abaixo de 50%.

A mesma mudança também está em progresso no português brasileiro (PB), só que mais avançada: segundo Barbosa, Duarte e Kato (2005), as taxas de sujeitos nulos no padrão D, nessa língua, já chegam a 25%. O grupo de fatores ‘condição de referência’, então, identificou três tipos de língua. O italiano, o espanhol europeu e o argentino apresentam comportamentos que indicam que o núcleo I tende a identificar sempre a referência do sujeito nulo, sendo o italiano a língua mais prototípica. Espanhol porto-riquenho e português brasileiro passam por uma mudança no parâmetro, que está mais avançada no português brasileiro.

A tabela abaixo mostra os índices gerais de sujeitos nulos das análises de regra variável citadas neste trabalho, que só consideram sujeitos referenciais:

Tabela 3 – Índices gerais de sujeitos nulos em cinco línguas, em ordem crescente de percentuais

PB	Espanhol porto-riquenho	Espanhol argentino	Espanhol europeu	Italiano
29%	59%	68%	76%	84%

Fonte: Autor

Embora existam várias razões para preencher o sujeito referencial, como, por exemplo, marcar respeito, (re)inserir-se no discurso e desfazer ambiguidades, além de denotar contraste, ênfase e individualização do referente, o sujeito será obrigatoriamente pronunciado, em qualquer contexto, se a sua referência não puder ser identificada. E, embora o sujeito nulo possa ser identificado de outra maneira, na presença de um I capaz de identificá-lo ele sempre será identificado por esse núcleo. Dessa forma, se os comportamentos das línguas em relação à representação do sujeito se mostrarem coerentes com os percentuais de sujeitos de referência definida nulos, isso não deve ser interpretado como casual.

5 A ESTABILIDADE DAS LÍNGUAS

Uma língua com comportamento radical que só selecione I capaz de identificar a referência de um sujeito nulo terá os sujeitos nulos sempre identificados, em qualquer contexto. Quanto mais instável a língua for quanto à capacidade de I de identificar sujeitos nulos, mais contextos de (des)favorecimento do sujeito pleno aparecerão. Assim, interpreto uma maior quantidade de grupos de fatores selecionados nas análises de regra variável como indicadora de uma menor estabilidade da língua em relação ao parâmetro.

As análises de regra variável podem ser comparadas porque seguiram metodologias muito semelhantes, principalmente no que se refere ao estabelecimento dos grupos de fatores. No italiano, língua localizada mais à direita na Tabela 3, apenas quatro grupos de fatores foram selecionados como relevantes (MARINS, 2009), seguido pelo espanhol europeu, cuja análise só apontou seis grupos como significativos, conforme o Quadro 3. Já as línguas que exibem os menores percentuais de sujeitos nulos também são as línguas cujas análises selecionaram mais grupos de fatores: no português brasileiro, oito foram selecionados (DUARTE, 1995, p. 83); no espanhol porto-riquenho e no argentino, respectivamente, oito e nove grupos de fatores controlados são relevantes.

A seleção dos grupos de fatores sociais, que podem estar relacionados à avaliação das variantes do fenômeno, também é reveladora da estabilidade das línguas quanto à identificação do sujeito nulo. As análises das cinco línguas consideram duas variáveis sociais: faixa etária e gênero. Em todas as análises a variável faixa etária foi selecionada, mas o gênero, selecionado nas análises do português brasileiro, do espanhol porto-riquenho e do argentino, não é relevante no italiano e no espanhol europeu (EE), confirmando a maior estabilidade destas duas línguas frente àquelas.

As línguas podem, agora, ser localizadas no segmento de reta que representa o parâmetro, estando mais à direita as línguas com maior tendência a ter um I que pode identificar a referência de um sujeito nulo. Os pontos em que as línguas estão representadas no *continuum* do parâmetro, em (7), não são localizações exatas em escala. É possível dizer, por exemplo, que o espanhol porto-riquenho exibe um comportamento intermediário entre os observados para o português brasileiro (PB) e o espanhol argentino (EA) (cf. TARALLO; KATO, 1989, p. 24) –, mas não há meios de determinar a posição exata de cada língua em relação aos polos.

(7)	[– sujeito nulo]	– PB	– esp. porto-riquenho	– EA	– EE	– italiano	–	[+ sujeito nulo]
	sujeitos nulos:	29%	59%	68%	76%	84%		
	grupos de fatores:	8	8	9	6	4	(significantes)	

Pela comparação das análises realizadas para o italiano, o português brasileiro e as variedades do espanhol consideradas, é possível ordenar essas línguas segundo o comportamento exibido em relação à identificação da referência do sujeito nulo, conforme (7). O português brasileiro é a língua que tem o comportamento mais próximo ao protótipo de uma língua que tem um I que nunca pode identificar a referência dos sujeitos nulos, seguido do espanhol porto-riquenho, que tende a ter mais I capaz de identificar a referência de um sujeito nulo do que o português brasileiro. Mais próximos do polo [+ sujeito nulo] estão o português europeu, o espanhol argentino, o europeu e o italiano, que se aproxima mais do comportamento de uma língua que tem I capaz de identificar a referência dos sujeitos nulos sempre.

Tal ordenação das línguas é sustentada porque elas são ranqueadas na mesma ordem quando observados diversos aspectos relacionados à identificação da referência de sujeitos nulos. Assim, percorrendo a escala em (7) de [– sujeito nulo] a [+ sujeito nulo], as línguas aparecerão em ordem (a) crescente de frequências gerais de sujeitos referenciais nulos, (b) decrescente de quantidade de grupos de fatores que influenciam na representação do sujeito referencial e (c) decrescente de grau de atuação dos grupos de fatores sociais na seleção entre sujeito nulo e pleno. Além disso, as línguas também

estão ordenadas de acordo com a influência da condição de referência na identificação do sujeito, começando pela influência mais intensa.

Dentre as cinco línguas ordenadas em (7), a ocorrência de sujeitos inanimados plenos só é observada no português brasileiro, a língua localizada mais à esquerda da escala – consultem-se, também, considerações sobre outros aspectos relacionados ao Parâmetro do Sujeito Nulo, nas análises apresentadas em Soares da Silva (2011), como, por exemplo, as taxas de sujeitos pós-verbais nessas línguas, que as põem na mesma ordem apresentada em (7). O ranqueamento de todas essas línguas de acordo com a manifestação de diversos fenômenos relacionadas à representação do sujeito é o mesmo, o que ajuda a sustentar que realmente existe uma escala de prototipicidade de língua de sujeito nulo, em que as línguas – ou pelo menos essas cinco línguas, aqui comparadas – podem ser ordenadas. A relação entre a riqueza flexional do verbo e a produtividade de sujeitos nulos, levantada desde que o Parâmetro do Sujeito Nulo foi postulado (CHOMSKY, 1981), leva à mesma ordenação: quanto mais à direita na representação em (7), mais rico é o paradigma, como se verá na seção seguinte – o que não deve ser interpretado como uma coincidência.

6 A RIQUEZA FLEXIONAL DE PESSOA

A interpretação dos resultados das análises de regra variável mostra que, em relação a identificação dos sujeitos nulos por I, os comportamentos das línguas consideradas seguem a mesma ordem das frequências gerais de sujeitos referenciais nulos. A observação do grau de riqueza dos paradigmas flexionais de cada uma dessas línguas confirma a ordenação delas, como mostrarei nesta seção.

Começamos observando o paradigma completo de conjugação verbal do italiano, que é o mais rico¹⁴. Os quadros abaixo apresentam os paradigmas flexionais de sete tempos verbais que são representados por formas simples nas cinco línguas comparadas, cada um separado em dois miniparadigmas, um para o singular e outro para o plural. Não é necessário incluir os tempos verbais compostos porque, nestes, o verbo auxiliar se apresenta sempre sob a forma morfológica correspondente a um dos tempos verbais simples.

Quadro 4 Paradigma flexional (indicativo) do italiano

Presente		Pret. imperfeito		Pret. perfeito		Futuro presente		Futuro do pretérito	
S.	Plural	S.	Plural	S.	Plural	S.	Plural	S.	Plural
parl	parliam	parlav	parlavam	parlai	parlamm	parler	parlerem	parlerei	parleremm
o	o	o	o	o	o	ò	o	o	o
parli	parlate	parlav	parlavate	parlas	parlaste	parler	parlerete	parleresti	parlereste
		i		ti		ai			
parl	parlano	parlav	parlavan	parlò	parlaron	parler	parlerann	parlerebb	parlerebbe
a		a	o	o	o	à	o	e	ro

Fonte: Autor

¹⁴ Os paradigmas apresentados nesta seção foram construídos a partir da lista de todas as formas verbais encontradas nos dados analisados, que geraram os resultados apresentados neste artigo, para as diferentes línguas. Os paradigmas não foram retirados de gramáticas descritivas ou normativas. Assim, formas verbais que não foram efetivamente usadas nas amostras analisadas não aparecem nos quadros.

Quadro 5 Paradigma flexional (subjuntivo) do italiano

Presente		Pretérito imperfeito	
Singular	Plural	Singular	Plural
parli	parliamo	parlassi	parlassimo
parli	parliate	parlassi	parlaste
parli	parlino	parlasse	parlassero

Fonte: Autor

Dividindo o paradigma completo em sete tempos verbais e redividindo-o em singular e plural, obtêm-se 14 miniparadigmas de pessoa. No italiano, os 10 miniparadigmas do indicativo apresentam uma riqueza plena para pessoa gramatical, já que, em cada um deles, cada uma das três pessoas é representada por uma forma verbal diferente, como se vê no Quadro 4. No subjuntivo, porém, são identificados dois miniparadigmas sem essa riqueza: os de singular dos dois tempos verbais, conforme o Quadro 5.

No espanhol europeu, o morfema zero neutraliza a primeira e a terceira pessoas do singular em dois tempos verbais do indicativo e dois do subjuntivo, conforme os dois quadros seguintes. Esses paradigmas de singular não têm riqueza plena para pessoa. O espanhol europeu, então, tem um paradigma que, como um todo, é menos rico do que o do italiano, já que a riqueza plena é observada em 10 dos 14 miniparadigmas, enquanto no italiano ela existe em 12 deles.

Quadro 6 Paradigma flexional (indicativo) do espanhol europeu

Presente		Pret. imperfeito		Pret. perfeito		Futuro presente		Futuro pretérito	
S.	Plural	S.	Plural	S.	Plural	S.	Plural	S.	Plural
habl	hablam	hablab	hablábam	hablé	hablam	hablar	hablarem	hablarí	hablaríam
o	os	a	os	os	os	é	os	a	os
habla	habláis	hablab	hablabais	hablas	hablaste	hablar	hablaréis	hablarí	hablaríais
s		as		te	is	ás		as	
habla	hablan	hablab	hablaban	habló	hablaro	hablar	hablarán	hablarí	hablarían
		a		n		á		a	

Fonte: Autor

Quadro 7 Paradigma flexional (subjuntivo) do espanhol europeu

Presente		Pretérito imperfeito	
Singular	Plural	Singular	Plural
hable	hablemos	hablara	habláramos
hables	habléis	hablaras	hablarais
hable	hablen	hablara	hablaran

Fonte: Autor

No espanhol argentino, além das neutralizações que também são observadas no europeu, existe, ainda, o sincretismo entre a segunda e a terceira pessoas do plural, causado pela perda do pronome *vosotros* (vós), que afeta todos os tempos verbais. Assim, 11 miniparadigmas no espanhol argentino são pobres, enquanto apenas três apresentam riqueza plena. Os quadros a seguir mostram os sincretismos.

Quadro 8 Paradigma flexional (indicativo) do espanhol argentino

Presente		Pret. imperfeito		Pret. perfeito		Futuro presente		Futuro pretérito	
S.	Plural	S.	Plural	S.	Plural	S.	Plural	S.	Plural
habl	hablam	hablab	hablábam	hablé	hablam	hablar	hablarem	hablarí	hablaríam
o	os	a	os		os	é	os	a	os
habla	hablan	hablab	hablaban	hablas	hablaro	hablar	hablarán	hablarí	hablarían
s		as		te	n	ás		as	
habla	hablan	hablab	hablaban	habló	hablaro	hablar	hablarán	hablarí	hablarían
		a			n	á		a	

Fonte: Autor

Quadro 9 Paradigma flexional (subjuntivo) do espanhol argentino

Presente		Pretérito imperfeito	
Singular	Plural	Singular	Plural
hable	hablemos	hablara	habláramos
hables	hablen	hablaras	hablaran
hable	hablen	hablara	hablaran

Fonte: Autor

No espanhol porto-riquenho, além de todos os sincretismos existentes no argentino, observa-se ainda o apagamento do /s/ final, que causa neutralização entre a segunda e a terceira pessoas do singular (essa neutralização, em alguns tempos verbais, inclui ainda a primeira pessoa do singular). Dessa forma, o único miniparadigma com riqueza de pessoa gramatical é o do singular no pretérito perfeito (que permanece idêntico ao do espanhol europeu), em que a forma de segunda pessoa do singular não é afetada pelo apagamento do /s/ porque não termina em /s/. Os quadros a seguir mostram os sincretismos observados no espanhol porto-riquenho, assinalados em cor cinza:

Quadro 10 Paradigma flexional (indicativo) do espanhol porto-riquenho

Presente		Pret. imperfeito		Pret. perfeito		Futuro presente		Futuro pretérito	
S.	Plural	S.	Plural	S.	Plural	S.	Plural	S.	Plural
habl	hablam	hablab	hablábam	hablé	hablam	hablar	hablarem	hablarí	hablaríam
o	os	a	os		os	é	os	a	os
habl	hablan	hablab	hablaban	hablast	hablaro	hablar	hablarán	hablarí	hablarían
a		a		e	n	á		a	
habl	hablan	hablab	hablaban	habló	hablaro	hablar	hablarán	hablarí	hablarían
a		a			n	á		a	

Fonte: Autor

Quadro 11 Paradigma flexional (subjuntivo) do espanhol porto-riquenho

Presente		Pretérito imperfeito	
Singular	Plural	Singular	Plural
hable	hablemos	hablara	habláramos
hable	hablen	hablara	hablaran
hable	hablen	hablara	hablaran

Fonte: Autor

Por fim, o paradigma do português brasileiro é o mais pobre, pois a riqueza de oposições de pessoa não é observada em nenhum dos miniparadigmas. Por causa da

substituição dos pronomes **tu** e **vós** por **você e vocês**¹⁵, em todos os tempos verbais, no singular e no plural, a segunda e a terceira pessoas são representadas por formas verbais idênticas, como se vê nos dois quadros a seguir. E, nos tempos verbais em que a desinência zero também se associa à primeira pessoa do singular, as três formas dos miniparadigmas do singular são iguais.

Quadro 12 Paradigma flexional (indicativo) do português brasileiro (cf. DUARTE, 1995, p. 40)

Presente		Pret. imperfeito		Pret. perfeito		Futuro do pres.		Futuro do pretérito	
S.	Plural	S.	Plural	S.	Plural	S.	Plural	S.	Plural
falo	fala	falava	falava	falei	falou	falarei	falará	falaria	falaria
fala	falam	falava	falavam	falou	falaram	falará	falarão	falaria	falariam
fala	falam	falava	falavam	falou	falaram	falará	falarão	falaria	falariam

Fonte: Autor

Quadro 13 Paradigma flexional (subjuntivo) do português brasileiro (cf. DUARTE, 1995, p. 40)

Presente		Pretérito imperfeito	
Singular	Plural	Singular	Plural
fale	fale	falasse	falasse
fale	falem	falasse	falassessem
fale	falem	falasse	falassessem

Fonte: Autor

Nos quadros 12 e 13 observa-se, ainda, que a primeira pessoa do plural também é representada pelo morfema zero, consequência da introdução de **a gente** ao paradigma pronominal. Isso, porém, não influi no grau de riqueza de oposições de pessoa, já que, em cada miniparadigma de plural do português brasileiro, a forma de primeira pessoa permanece exclusiva, opondo-se às outras duas. O que causou a perda da riqueza de pessoa foi mesmo a generalização de **você(s)**, que afetou todos os tempos verbais de uma vez, desencadeada na década de 1950: assim, iniciou a mudança em direção ao I incapaz de identificar a referência dos sujeitos nulos, antes mesmo da gramaticalização de **a gente**, que é posterior e não tem efeito sobre a identificação¹⁶.

CONCLUSÃO

Fica demonstrado, portanto, que o grau de riqueza de oposições de pessoa se relaciona diretamente à produtividade de sujeitos nulos referenciais, ou seja, quanto mais rico para pessoa é o paradigma verbal, maior é a probabilidade de o núcleo I conseguir

¹⁵ O estudo diacrônico de Duarte (1993) revela o desaparecimento do pronome **tu** na metade do século XIX. No final do século, esse pronome volta a aparecer, porém não mais associado à desinência -s.

¹⁶ Mas tem efeito sobre o licenciamento. Em Soares da Silva (2011), defende-se, com base em Figueiredo Silva (1996), que a identificação do sujeito nulo está relacionada à riqueza de oposições entre as pessoas gramaticais e o licenciamento, à riqueza de oposições de número. O miniparadigma da primeira pessoa do pretérito imperfeito, no português brasileiro, por exemplo, não opõe os números: tanto singular quanto plural apresentam a forma **falava**; já o se segunda pessoa, no mesmo tempo verbal, apresenta oposição: **falava** vs. **falavam**. Em Soares da Silva (2011), é apresentada uma análise semelhante à apresentada aqui, para o licenciamento do sujeito nulo, com os mesmos paradigmas divididos em 21 miniparadigmas em que se observa a quantidade de oposições entre os dois números em cada uma das três pessoas, nos sete tempos verbais – e mostra-se que, após a generalização do uso de **a gente**, com consequente neutralização entre os números singular e plural na primeira pessoa, emergiram estratégias para o preenchimento da posição do sujeito expletivo, evidenciando que os sujeitos nulos começaram a deixar de ser licenciados.

identificar a referência do sujeito nulo. O italiano, que apresenta o comportamento mais prototípico de língua de sujeito nulo, tem, em seu paradigma verbal, riqueza plena de oposição entre as pessoas gramaticais em 12 dos 14 miniparadigmas, seguido de perto pelo espanhol europeu, com essa oposição em 10 miniparadigmas, e o espanhol argentino, com oposição plena em três miniparadigmas, ainda mantém um comportamento estável quanto à identificação do sujeito nulo por I. No espanhol porto-riquenho, em que as neutralizações restringiram a riqueza de pessoa apenas ao pretérito perfeito, e no português brasileiro, em que tal riqueza não existe mais, foi desencadeada uma mudança em direção ao I incapaz de identificar o sujeito nulo, reduzindo, conseqüentemente, as taxas de sujeitos pronominais nulos.

(8) repete o esquema apresentado em (7). Agora, foram acrescentadas as quantidades de miniparadigmas com riqueza plena de oposição entre as três pessoas gramaticais. Observe-se que elas servem para confirmar o ranqueamento das línguas já mostrado em (7).

(8) [- sujeito nulo] - PB - esp. porto-riquenho - EA - EE - italiano - [+ sujeito nulo]						
sujeitos nulos:	29%	59%	68%	76%	84%	
grupos de fatores:	8	8	9	6	4	(significantes)
riqueza de pessoa:	0	1	3	10	12	miniparadigmas

As cinco línguas comparadas neste trabalho foram ordenadas segundo (a) a frequência geral de sujeitos pronominais de referência definida nulos, (b) a estabilidade relacionada à capacidade do núcleo I de identificar sujeitos nulos sem a necessidade de um antecedente e (c) a riqueza flexional. O fato de esses ranqueamentos serem semelhantes, e compatíveis com a intensidade da atuação da condição de referência, não deve ser interpretado como uma coincidência casual. Ao contrário, é um forte indício de que todos esses aspectos estão relacionados à identificação dos sujeitos nulos e, conseqüentemente, à sua produtividade – motivo pelo qual se defende que um parâmetro é uma escala contínua, em que línguas podem ser localizadas e ordenadas.

REFERÊNCIAS

- BARBOSA, P.; DUARTE, M.; KATO, M. Null subjects in European and Brazilian Portuguese. *Journal of Portuguese Linguistics*, v. 4, n. 2, p. 11-52, 2005.
- BORGES NETO, J. *Ensaio de Filosofia da Lingüística*. São Paulo: Parábola, 2004.
- CAMERON, R. Ambiguous agreement, functional compensation, and nonspecific tú in the Spanish of San Juan, Puerto Rico, and Madrid, Spain. *Language Variation and Change*, Nova Iorque, n. 5 (1994), p. 305-334, 1993.
- CHOMSKY, N. Current issues in linguistic theory. In: FODOR, J.; KATZ, J. *The structure of Language: readings in the Philosophy of Language*. Englewood Cliffs: Prentice Hall, 1964, p. 50-118.
- _____, N. *Lectures on Government and Binding*. 2. ed. (1982). Dordrecht: Foris, 1981.
- CYRINO, S. *O objeto nulo no português do Brasil: um estudo sintático-diacrônico*. 1994. Tese (Doutorado) – Instituto de Estudos da Linguagem, UNICAMP, Campinas, 1994.

DUARTE, M. Do pronome nulo ao pronome pleno: a trajetória do sujeito nulo no português do Brasil. In: ROBERTS, I.; KATO, M. (org.). *Português brasileiro: uma viagem diacrônica*. Campinas: UNICAMP, 1993, p. 107-28.

_____, M. *A perda do princípio “Evite Pronome” no português brasileiro*. 1995. Tese (Doutorado) – Instituto de Estudos da Linguagem, UNICAMP, Campinas, 1995.

_____, M. On the embedding of a syntactic change. *Language variation in Europe: papers from ICLaVE2*, Uppsala, p. 144-155, 2004.

DURANTI, A.; OCHS, E. Left-dislocation in Italian conversation. In: GIVÓN, T. *Syntax and semantics: discourse and syntax*, v. 12. Nova Iorque: Academic Press, 1979, p. 377-415.

FIGUEIREDO SILVA, M. *A posição do sujeito no português brasileiro: frases finitas e infinitivas*. Campinas: UNICAMP, 1996.

FLORES-FERRÁN, N. A bend in the road: subject personal pronoun expression in Spanish after 30 years of sociolinguistic research. *Language and Linguistic Compass*, n. 1, 2007.

HOCHBERG, J. Functional compensation for /s/ deletion in Puerto Rican Spanish. *Language: Journal of the Linguistic Society of America*, v. 62, n. 3, p. 609-621, 1986.

KATO, M. The partial pro-drop nature and the restricted VS order in Brazilian Portuguese. In: KATO, M.; NEGRÃO, E. *Brazilian Portuguese and the Null Subject Parameter*. Madri: Iberoamericana/Frankfurt: Vervuert, 2000, p. 223-258.

KATO, M.; DUARTE, M. *(Micro) parametric variation between European (EP) and Brazilian Portuguese (BP): similarities and differences related to ongoing changes in Latin American Spanish*. Monterrey: ALFAL, 2005.

KATO, M.; NASCIMENTO, M. (org.). *Gramática do português culto falado no Brasil*. Campinas: UNICAMP, 2009.

MARINS, J. *O Parâmetro do Sujeito Nulo: uma análise contrastiva entre o português e o italiano*. 2009. Dissertação (Mestrado em Letras Vernáculas) – Programa de Pós-Graduação em Letras Vernáculas, UFRJ, Rio de Janeiro, 2009.

MOINO, R. *Passivas no discurso oral e escrito*. 1987. Dissertação (Mestrado) – PUC-SP, São Paulo, 1987.

NOVAES, C. *Representação mental das categorias vazias: o caso do sujeito nulo no português do Brasil*. 1996. Tese (Doutorado em Linguística) – Faculdade de Letras, UFRJ, Rio de Janeiro, 1996.

NUNES, J. *O famigerado se: uma análise sincrônica e diacrônica das construções com se apassivador e indeterminador*. 1990. Dissertação (Mestrado) – UNICAMP, Campinas, 1990.

PAGOTTO, E. *A posição dos clíticos em português: um estudo diacrônico*. 1992. Dissertação (Mestrado) – UNICAMP, Campinas, 1992.

RAMOS, J. *Marcação de Caso e mudança sintática no português do Brasil*. 1992. Tese (Doutorado) – Instituto de Estudos da Linguagem, UNICAMP, Campinas, 1992.

_____, J. “Sociolinguística Paramétrica” ou “Variação Paramétrica”? In: HORA, D.;

CHRISTIANO, E. *estudos linguísticos: realidade brasileira*. João Pessoa: Ideia, 1999, p. 83-94.

RAPOSO, E. *Teoria da gramática: a faculdade da linguagem*. Lisboa: Caminho, 1992.

SAMPER PADILLA, J.; HERNÁNDEZ CABRERA, C.; TROYA DÉNIZ, M. (org.). *Macrocorpus de la norma lingüística culta de las principales ciudades del mundo hispánico*. CD-ROM. Las Palmas de Gran Canaria: Universidad de Las Palmas de Gran Canaria, 1995.

SOARES DA SILVA, H. *O Parâmetro do Sujeito Nulo: confronto entre o português e o espanhol*. 2006. Dissertação (Mestrado em Letras Vernáculas) – Programa de Pós-Graduação em Letras Vernáculas, UFRJ, Rio de Janeiro, 2006.

_____, H. *Evidências da mudança paramétrica em dados da língua-E: o sujeito pronominal no português e no espanhol*. 2011. 126f. Tese (Doutorado em Língua Portuguesa) – Programa de Pós-Graduação em Letras Vernáculas, UFRJ, Rio de Janeiro, 2011.

TARALLO, F. Por uma Sociolinguística Românica “Paramétrica”: fonologia e sintaxe. *Ensaios de Linguística*, n. 13, p. 51-84, 1987.

TARALLO, F.; KATO, M. Harmonia trans-sistêmica: variação intra- e inter-lingüística. *Diadorim: Revista de Estudos Lingüísticos e Literários*, n. 2 (2006), p. 13-42, 1989.

TORIBIO, A. Dialectal variation in the licensing of null referential and expletive subjects. In: PARODI, C. *et al* (orgs.). *Aspects of Romance linguistics*. Washington: Georgetown University, 1994, p. 409-32.

WEINREICH, U.; LABOV, W.; HERZOG, M. *Fundamentos empíricos para uma teoria da mudança lingüística*. Tradução de Marcos Bagno (2006). São Paulo: Parábola, 1968.

Recebido: 16/06/2013

Aceito: 31/08/2013