

A REPRESENTAÇÃO DO SUJEITO PRONOMINAL NO GRUPO ROMÂNICO: ESPANHOL E ITALIANO EM CONTRASTE COM O PORTUGUÊS

Juliana Espósito Marins*
Humberto Soares da Silva*

Resumo: Este trabalho visa a mostrar o comportamento do sujeito pronominal em variedades do espanhol e no italiano, línguas prototipicamente de sujeito nulo, no intuito de evidenciar o distanciamento do comportamento do português brasileiro no que se refere ao preenchimento do sujeito pronominal. Além disso, o trabalho mostrará que parece existir uma escala quanto ao modo como as línguas exibem as propriedades relacionadas ao Parâmetro do Sujeito Nulo, o que parece estar relacionado com o número de oposições que paradigma flexional do verbo exhibe em cada língua.

Palavras-chave: Parâmetro do Sujeito Nulo; sujeito pronominal; português brasileiro; espanhol; italiano.

Abstract: This work aims to show the behavior of the pronominal subject in varieties of Spanish and Italian, prototypically null subject languages, in order to evince the distance of the behavior of Brazilian Portuguese regarding the filling of the pronominal subject. Moreover, this work shows that there seems to be a scale of how languages exhibit properties related to Null Subject Parameter, which seems to be related to the number of oppositions that inflectional paradigm of the verb displays in each language.

Keywords: Null Subject Parameter; pronominal subject; Brazilian Portuguese; Spanish; Italian.

* Universidade Federal do Rio de Janeiro.

1. Introdução

Estudos sobre a representação do sujeito pronominal no português brasileiro (PB) vêm atestando alterações significativas quanto à remarcação do Parâmetro do Sujeito Nulo (PSN), ao contrário do que ocorre com o português europeu (PE) (DUARTE, 1995): no PB há uma preferência pela realização fonética do sujeito, em decorrência da redução do paradigma flexional do verbo, conforme hipótese de Duarte (1993, 1995), provocada por significativas mudanças no quadro pronominal (LOPES & DUARTE, 2003). Ou seja, dentro da perspectiva teórica gerativista (CHOMSKY, 1981), o PB estaria passando de língua positivamente marcada para língua negativamente marcada em relação ao PSN. Os exemplos em (1)¹ ilustram essa tendência:

- (1) a. *Mesmo que eu não fizesse o pré-vestibular, eu acho que eu passaria por causa da base que eu tinha.*
- b. *E quando vocês 'tão andando assim na vila vocês reconhecem (os astros do esporte)?*
- c. *Essa minha tia que mora aqui, ela é solteirona e eu acho que ela é super-feliz, sabe? Eu não acho que ela seria feliz assim... [...] Ela ficou solteira porque ela quis.*

Para embasar tais afirmações, no sentido de classificar o comportamento do PB como atípico no contexto das línguas românicas, era necessário que houvesse estudos contrastivos de outras línguas dentro de uma mesma perspectiva teórica. Assim, além do trabalho de Duarte (1995), que não apenas verificou uma expressiva preferência pelo preenchimento da posição de sujeito no PB, mas também atestou um comportamento no PE mais compatível com o de uma língua positivamente marcada para o PSN, utilizando-se de pressupostos da Teoria de Princípios e Parâmetros (CHOMSKY, 1981) e da Teoria da Variação e Mudança

¹ Exemplos retirados de Duarte (1995).

(LABOV, 1972), pode-se citar o trabalho de Soares da Silva (2006), que comparou os resultados de Duarte com o espanhol de Madri e de Buenos Aires. De modo geral, tais trabalhos apresentaram a seguinte situação: a variedade madrilena do espanhol foi a que apresentou o maior índice de sujeitos nulos, com 73%, seguida do espanhol de Buenos Aires, com 71%. Já o PE apresentou um percentual de 66% de sujeitos nulos, enquanto PB, apenas 29%. Esse quadro não só ilustra o afastamento de PB em relação às variedades do português e do espanhol, como também permite pensar que essas variedades não se comportam de modo idêntico quanto às propriedades relacionadas ao PSN.

Nesse sentido, o presente trabalho será desenvolvido em duas frentes: primeiramente, mostraremos o comportamento do italiano padrão (MARINS, 2009), em busca daquela que seria a língua mais próxima do protótipo de uma língua de sujeito nulo (LSN). Com isso, pretendemos mostrar quais são os fatores que podem levar uma LSN a preencher a posição estrutural de sujeito. Num segundo momento, trataremos da análise de outra variedade do espanhol (Porto Rico) para tentar compreender como podemos caracterizar as línguas em relação ao PSN. Para isso, proporemos um novo olhar sobre a questão da riqueza funcional do paradigma verbal (ROBERTS, 1993) e sua relação com a identificação e o licenciamento do sujeito nulo.

2. A análise do italiano padrão

Conforme Duarte (1995) e Soares da Silva (2006), para a análise do italiano padrão, foram associados pressupostos teóricos da Teoria da Variação e Mudança (WEINREICH, LABOV e HERZOG, 1968) e da Teoria de Princípios e Parâmetros (CHOMSKY, 1981). A amostra foi constituída através do *corpus* C-ORAL-ROM *Integrated corpora for spoken romance languages* (CRESTI & MONEGLIA, 2005). Foram investigadas 12 conversações, com falantes homens e mulheres distribuídos em três faixas etárias.

Foram excluídos da análise estatística os dados em que o sujeito é categoricamente pleno em LSN, a saber, os dados em que a expressão do sujeito está relacionada à ênfase, exemplificado em (2),² ou ao contraste, conforme se vê em (3):

- (2) *Ha detto lui: “sì, è vero. Però, forse, più che egoista sono individualista.” Ho detto: “no, IO sono individualista, TU invece sei egoista”.* [SND ifamcv21]
 (“Disse ele: ‘sim, é verdade. Mas, talvez, mais que egoísta sou individualista’. [Eu] Disse: ‘não, eu sou individualista, você ao contrário é egoísta.’”)
- (3) **LIA:** [...] *‘un [= non] mi ricordo indo’ ho messo le fotografie delle Dolomiti e della Germania. Oh, questa è bella. Eppure le presi...*
 (“não me lembro onde coloquei as fotografias dos Dolomíticos e da Alemanha. Oh, essa é bonita. Entretanto as tirei...”)
- ELA:** *Mi ricordo che TU me l’hai fatte vedere anche a me.* [ELA ifamcv01]
 (“Me lembro que você as mostrou também a mim.”)

2.1. A análise dos dados: resultados gerais

Excluídos tais dados, chegamos a um total geral de 867 dados computados, dos quais 746 exibem sujeitos nulos (86%) e 121, sujeito preenchidos (14%). Desse modo, de acordo com as previsões feitas, na amostra analisada, o italiano apresentou-se predominantemente

² Os exemplos do italiano mencionados neste trabalho mantêm a grafia encontrada na transcrição da fala dos informantes, tal como aparecem em Cresti & Moneglia (2005). Como pudemos detectar, há elementos, sobretudo no âmbito lexical, oriundos do dialeto florentino, mas que não prejudicam a análise proposta (além disso, o significado desses elementos aparece esclarecido através das glosas em português que acompanham cada exemplo). É necessário esclarecer que o que acompanha cada exemplo em língua estrangeira não são propostas de tradução livre, mas glosas que presam, sempre que possível e tentando preservar ao máximo as propriedades da língua, por uma tradução “palavra por palavra”: com isso, temos o intuito de identificar os elementos gramaticais que compõem os enunciados em análise, e não propriamente captar o sentido essencial dos enunciados.

favorável ao apagamento do sujeito. Para estabelecer uma comparação com os resultados encontrados para as duas variedades do português, PE e PB (DUARTE, 1995), e para as do espanhol, Madri e Buenos Aires (SOARES DA SILVA, 2006), apresentamos abaixo a Tabela 1, com os resultados gerais para as cinco línguas românicas observadas:

TABELA 1
Distribuição de sujeitos nulos vs. plenos pelas línguas românicas

Língua românica	Nulos	Plenos	Total
Italiano sub- <i>standard</i> médio	746 (86%)	121 (14%)	867 (100%)
Espanhol europeu (Madri)	937 (76%)	422 (26%)	1238 (100%)
Espanhol argentino (Buenos Aires)	834 (68%)	414 (21%)	1221 (100%)
Português europeu	738 (66%)	378 (34%)	1116 (100%)
Português brasileiro	415 (29%)	1009 (71%)	1424 (100%)

Fonte: Marins & Soares da Silva (2006, p. 204)

Como se vê, apesar da diferença entre os números de dados analisados para cada variedade românica, é nítida a aproximação entre italiano, espanhol e PE, quanto à preferência pelo sujeito nulo, todos com índices superiores a 65%. Por outro lado, encontramos o PB, que revela um percentual absolutamente expressivo de sujeitos plenos (71%), comportamento diametralmente oposto ao das outras línguas românicas.

Esse resultado geral para o italiano era esperado, assim como ocorreu para o espanhol e para o PE, tendo em vista o Princípio do Subconjunto (RAPOSO, 1992), segundo o qual línguas positivamente marcadas em relação ao PSN, apesar de exibirem a forma nula como *default*, podem (ou precisam) apresentar, em determinadas circunstâncias, sujeitos foneticamente realizados. É o que afirma Rizzi (1982), sobre os casos de contraste e ênfase, além da questão da desambiguação, e até mesmo da introdução do falante ou da retomada de um sujeito distante na mudança de tópico, um fator discursivo que não pode ser ignorado. Entretanto, os 14% referentes aos sujeitos plenos da amostra analisada não contemplam os sujeitos enfáticos e contrastivos, excluídos dessa análise geral, o que nos fez proceder a um exame dessas estruturas

para tentar explicar os fatores que concorreram para que os sujeitos aparecessem foneticamente realizados nessa amostra.

2.2. Sujeitos plenos no italiano padrão: motivações discursivas

Apesar de os sujeitos plenos totais representarem uma parcela relativamente pequena (14%), é necessário fazer algumas considerações sobre os contextos discursivos em que se inserem os sujeitos foneticamente realizados em uma língua de sujeito nulo.

Primeiramente, o *corpus* utilizado para a constituição da amostra em análise é composto por conversações, em que há sempre mais de um participante. Nessas condições, constatamos situações particulares desse tipo de *corpus*, como (a) a introdução do participante na conversa, tanto no primeiro registro de sua presença, quanto em retomada de turno de fala, expressando opinião, ratificação ou reiteração de uma opinião/ideia anterior; (b) a referência direta a participantes, tanto em sua primeira menção, quanto nos casos de mudança de interlocutor; e (c) a introdução de um tópico discursivo e sua consequente progressão ou mudança do tópico discursivo. Desse modo, foi possível verificar que todas as ocorrências de sujeitos plenos poderiam ser enquadradas em algum dos itens de (a)-(c), o que nos permite dizer que, dentre os dados analisados, a presença de um pronome sujeito anteposto ao verbo exhibe outras funções discursivas, além das que já havíamos considerado. Observemos o Gráfico 1, dividido de acordo com os itens (a), (b) e (c):

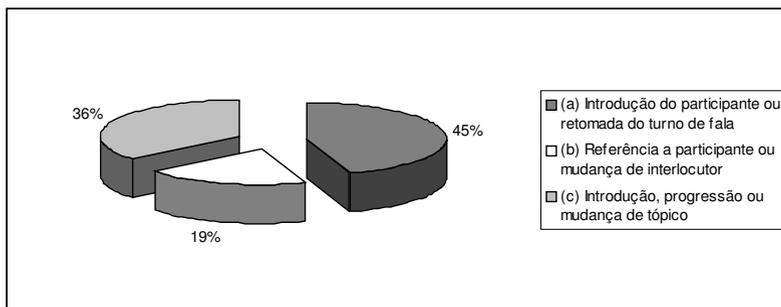


GRÁFICO 1: Distribuição de sujeitos plenos segundo a função discursiva

Como é possível observar, dos 121 dados de sujeitos plenos, 54 ocorrências (45%) representavam introdução do turno de fala do participante na conversação ou sua reintrodução, retomando o turno de fala, no sentido de expressar uma opinião, reiterar uma informação, confirmar um posicionamento ou ideia anteriormente expressos. No exemplo em (4), abaixo, representamos uma situação em que dois membros de uma banda de música discutiam a aquisição de novos equipamentos e uma das participantes, ausente da discussão momentaneamente, retoma o turno de fala para esclarecer o ponto de vista de um dos músicos:

(4) **SRE:** *No? Più o meno, tu lo sai come la pensano loro, più o meno. Loro non vogliono. [...]*

(“Não? Mais ou menos, você sabe como pensam eles, mais ou menos. Eles não querem.”)

GPA: *Io, magari, se si riusciva a dare un po' di tempo per fare... [...] Eh, io posso convincerli, [...] mi devo convincere anch'io [...] non posso nemmeno rompere i coglioni.*

(“Eu, talvez, se se conseguisse dar um pouco mais de tempo para fazer... [...] Eh, eu posso convencê-los, [...] devo me convencer também também eu [...] não posso nem sequer me aborrecer.”)

ANG: *No, io lo capisco. Cioè, la questione di principio, la capisco.*

(“Não, eu o entendo. Quer dizer, a questão de princípio, a entendo.”)

A referência direta a um participante da conversação ou a mudança de interlocutor, conforme o Gráfico 1, representa 19% (23 ocorrências) do total de sujeitos plenos. Abaixo, no exemplo em (5), ilustramos um caso de mudança de interlocutor.

(5) **GIA:** *Io non guardo Troisi, perché non lo capisco.*
(“Eu não assisto Troisi, porque não o entendo.”)

SIL: *Ma va, ma dai...*
(Como assim?)

GIA: *Quando parla, non capisco un cavolo!*
(“Quando fala, não entendo nadinha!”)

FAB: *Ma ridi lo stesso, non c'è problema se non lo capisci...*
 (“Mas [você – arb.]³ ri da mesma forma, não tem problema se [você – arb.] não o entende...”)

SIL: *Ma che non lo capisci... ma come fai a non capirlo!*
 (“Mas como [você] não o entende... mas como [você] faz para não entendê-lo!”)

GIA: *Non si capisce una parola di Troisi.* [Ele ri]
 (“Não se entende uma palavra de Troisi.”)

SIL: *Tu capisci?* [Falando com FRA]⁴ *Eh?*
 (“Você entende? Eh?”)

Por fim, a questão da introdução, progressão, retomada ou mudança de tópico discursivo, que correspondeu a 36% (45 ocorrências) do total, é ilustrada em (6), em que há a retomada do tópico introduzido por SND, que conversava com RSY, quando entra no recinto OLM, interrompendo a conversa momentaneamente.

(6) [Entra OLM. SND cumprimenta OLM e continua falando com RSY]

RSY: *Cioè, a ballare da sola, come vai?*
 (“Quer dizer, dançar sozinha, como [você] vai?”)

SND: *A ballare, no, in effetti.*
 (“Dançar, não, de fato.”)

OLM: *Cosa stavate bevendo? Il succo?*
 (“O que [vocês] estavam bebendo? Suco de caixinha?”)

RSY: *No, la spremuta.*
 (“Não, suco natural.”)

³ Uso da segunda pessoa do singular com referência arbitrária, o que é tradicionalmente chamado de *sujeito indeterminado*, como ocorre na seguinte ocorrência do PB (PEUL), em que o pronome *você* não se refere ao receptor da informação: “Você, quando você viaja, você passa a ser turista, aí você faz coisas que você nunca faria se você estivesse no Brasil”.

⁴ FRA é um informante que fazia parte da conversa num momento anterior, mas ainda não havia se manifestado quanto ao assunto ilustrado no exemplo. A referência a esse participante demonstra a reintrodução do interlocutor no discurso.

OLM: *Il succo nel senso della spremuta...*

("Suco no sentido de natural...")

SND: *Quindi, secondo me, lui è individualista, un po' egoista rispetto a me. Invece, secondo lui, dice che io non sono individualista. Sono più aperta, magari... lui è più chiuso.*

("Então, na minha opinião, ele é individualista, um pouco egoísta em relação a mim. Ao contrário, na opinião dele, diz que eu não sou individualista. Sou mais aberta, talvez... ele é mais fechado.")

2.3. Pessoa gramatical: a hipótese da Hierarquia Referencial

A análise da pessoa gramatical tem papel central na comparação entre o PB e o italiano, no sentido de demonstrar o afastamento entre as duas línguas. Em primeiro lugar, numa análise mais geral, notou-se, conforme a Tabela 2, que ilustra a distribuição de sujeitos nulos pelas pessoas gramaticais, que a amostra de italiano em análise apresentou um comportamento semelhante ao encontrado por Duarte (1995) para o PB – e também para o PE e nas duas variedades do espanhol (SOARES DA SILVA, 2006).

TABELA 2
Sujeitos nulos segundo a pessoa gramatical: italiano

Pessoa gramatical	Ocorrências	%	Pesos relativos
Primeira pessoa	213/280	76%	0.37
Segunda pessoa	129/159	81%	0.44
Terceira pessoa	288/312	92%	0.64
Total	630/751	84%	
Log likelihood: – 300,989	Significância:	0,012	Input: 0,87

A Tabela 2 mostra que, embora em todas as pessoas a preferência seja pela forma nula, a primeira e a segunda⁵

⁵ Os sujeitos plenos encontrados na amostra de italiano revelam algum tipo de ligação com aspectos pragmático-discursivos motivados pelo tipo de inquérito. Tais aspectos estão diretamente relacionados com as duas primeiras pessoas

apresentam índices de preenchimento ainda mais altos que os da terceira (que exhibe a maior porcentagem de sujeitos nulos), mesma hierarquia encontrada por Duarte (1995) para o PB, como se pode ver na Tabela 3.

TABELA 3
Sujeitos nulos de acordo com a pessoa gramatical: PB

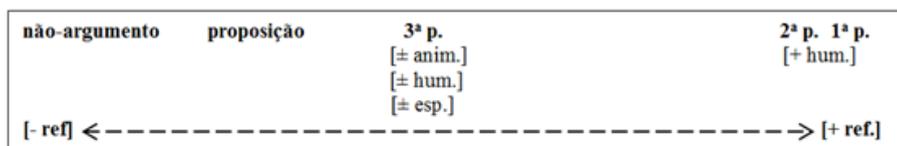
Pessoa gramatical	Ocorrências	%
Primeira pessoa	195/751	26%
Segunda pessoa	13/127	10%
Terceira pessoa	209/546	38%
Total	417/1424	29%

Fonte: Duarte (1995)

Por outro lado, um fator que marca nitidamente a diferença entre o italiano e o PB é o traço semântico do referente do pronome de terceira pessoa. Na amostra do italiano, os sujeitos de terceira pessoa que tinham um referente com o traço [- animado] foram categoricamente nulos, como nas amostras do espanhol, ao contrário do que ocorre no PB, em que esse traço admite a possibilidade de preenchimento – ainda que mais timidamente que o traço [+ animado]. Nesse sentido, o comportamento do italiano é mesmo o que se espera de uma LSN, o que nos permite afastar o PE do protótipo de LSN, já que este apresentou ainda algum preenchimento de sujeitos de terceira pessoa associados ao traço [- animado].⁶ Tais resultados estão de acordo com a hipótese da Hierarquia Referencial proposta por Cyrino, Duarte & Kato (2000), ilustrada na Figura 1 abaixo:

do discurso. Essa observação, no tocante à primeira pessoa, está de acordo com a hipótese de Duranti & Ochs (1979), que mostram que o falante tende a se introduzir no discurso através do pronome pleno. Parece que o mesmo ocorre com a segunda pessoa: o falante tende a introduzir o seu interlocutor no discurso através do uso de um pronome pleno.

⁶ Para esses percentuais, ver Marins (2009), Soares da Silva (2006) e Duarte (1995).



Fonte: Cyrino, Duarte & Kato (2000)

FIGURA 1: Escala de referencialidade

Segundo as autoras – e conforme se vê na Figura 1 – quanto mais referencial é o sujeito, maior é a probabilidade de que apareça preenchido, mesmo em LSN. Como a 1ª e a 2ª pessoas apresentam inerentemente o traço [+humano], elas tendem a ser as mais preenchidas, o que está de acordo com os resultados encontrados para todas as amostras analisadas nos trabalhos mencionados. Por outro lado, a 3ª pessoa seria a que apresentaria os menores índices de sujeitos plenos, comportando-se, desse modo, no caso do PB, como um contexto de resistência à mudança.

3. A representação do sujeito no espanhol: análise comparativa

As análises do espanhol europeu e do argentino foram feitas com o objetivo de descrever a representação do sujeito em línguas de sujeito nulo. Os resultados, porém, mostraram que essas línguas, embora apresentem preferência pelo sujeito nulo em todos os contextos controlados, se comportam de maneira diferente quanto ao PSN, além de não apresentar a prototipicidade observada no italiano e mostrada na seção anterior. Isso sugere que o comportamento das línguas em relação ao PSN extrapola as duas possibilidades previstas pelo referencial gerativista, as marcações positiva e negativa, motivo pelo qual classificamos as línguas, quanto ao parâmetro, em forma de escala, podendo, cada uma, estar mais próxima do protótipo de LSN ou do protótipo de língua de sujeito pleno.

Para a descrição de uma língua que, como o PB, também passa por mudança em direção ao preenchimento do sujeito, foi

feita a análise do espanhol porto-riquenho. Como será mostrado nessa seção, diferentemente do italiano, do espanhol europeu e argentino e do PE, o espanhol porto-riquenho apresenta alguns contextos de preferência pelo sujeito pleno, o que não é esperado para uma língua de sujeito nulo, embora a mudança esteja menos avançada que no PB, pois a frequência geral de sujeitos nulos é maior no espanhol porto-riquenho: 59% neste contra 29% no PB (DUARTE, 1995), em que o sujeito pleno é que se mostrou privilegiado em todos os contextos controlados.

Os dados do espanhol foram coletados das entrevistas que fazem parte do *Macrocorpus* de Samper Padilla, Hernández Cabrera & Troya Déniz (1995), todas realizadas no formato sociolinguístico “documentador-informante”. Foram selecionados 36 informantes, 12 de cada cidade: Madri (espanhol europeu), Buenos Aires (espanhol argentino) e San Juan (espanhol porto-riquenho). Nove diferentes grupos de fatores foram selecionados nessas três análises, e a interpretação dos resultados encontrados para todos eles é significativa para caracterizar e diferenciar essas três línguas: neste artigo, entretanto, enfatizaremos a condição de referência.⁷

O Gráfico 2 mostra os percentuais gerais de sujeitos nulos encontrados para o espanhol, junto com os índices do italiano (ver seção 2) e do PB (DUARTE, 1995). Essas taxas gerais indicam que a propriedade “sujeitos pronominais nulos” do PSN se manifesta com maior vigor no italiano e, bem mais timidamente, no PB, ficando as outras línguas comparadas aqui mais ou menos próximas de cada um desses extremos. Os resultados para a condição de referência, apresentados a seguir, corroborarão a postulação de diferentes comportamentos para essas línguas.

⁷ Consultem-se os resultados encontrados para todos os grupos de fatores selecionados nas análises do espanhol em Soares da Silva (2006, 2011).

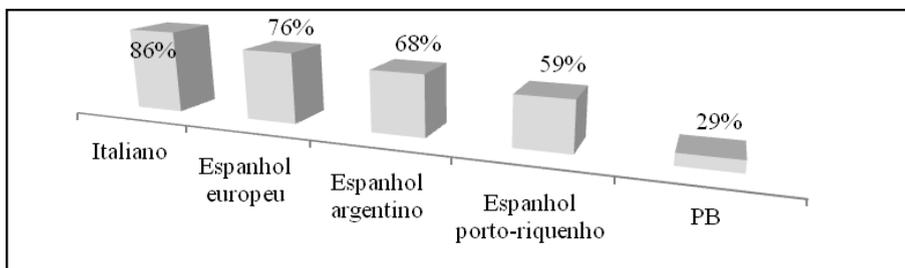


GRÁFICO 2: Percentuais gerais de sujeitos nulos em cinco línguas

Para investigar a influência da condição de referência na representação do sujeito, todos os dados coletados foram codificados segundo cinco padrões sentenciais, de acordo com a proximidade/distância do antecedente do pronome em análise e a identidade de função sintática. Esses padrões foram criados a partir dos padrões I, II, III e IV de Barbosa, Duarte e Kato (2001), através da cisão do padrão IV em dois, agora chamados de D e E, conforme a lista a seguir. Espera-se que o padrão A, que representa o antecedente mais acessível, favoreça o apagamento do sujeito, e que o resultado seja gradual, até o padrão E, que deve desfavorecer o apagamento – exceto se o paradigma de flexão verbal for suficientemente rico a ponto de ele mesmo ser capaz de identificar a referência dos sujeitos nulos, dispensando a coindexação com um antecedente.

a) Padrão A: O pronome em análise, nulo ou pleno, é classificado como padrão A quando está numa oração subordinada e a menção anterior ao mesmo referente, seu antecedente, é o sujeito da oração principal correspondente. No exemplo abaixo, o sujeito nulo de *digo* está em uma oração encaixada e seu antecedente, *yo*, sujeito de *bromeo*, está na respectiva oração matriz. Esse mesmo sujeito nulo é, ainda, antecedente para o sujeito *yo* da oração “que yo hice una...”, subordinada ao verbo *digo* – esse sujeito pleno *yo*, portanto, também representa o padrão A.

- (7) *yo_i bromeo [porque Ø_i digo [que yo_i hice una concentración en idiomas extranjeros estando en el Colegio de Comercio]]. Estudié dos años de francés (EP02)*
 (“eu brinco porque digo que eu fiz uma concentração em idiomas estrangeiros estando no Colegio de Comercio. Estudei dois anos de francês”)

b) Padrão B: O pronome é codificado no padrão B se o seu antecedente é o sujeito da oração imediatamente anterior, não havendo relação de subordinação entre a oração que tem o sujeito analisado e a que contém o antecedente. É o caso do sujeito nulo em (8a), abaixo, e do sujeito pleno em (8b).

- (8) a. *yo bromeo porque digo que yo_i hice una concentración en idiomas extranjeros estando en el Colegio de Comercio. Ø_i Estudié dos años de francés (EP02)*
 (“eu brinco porque digo que eu fiz uma concentração em idiomas estrangeiros estando no Colegio de Comercio. Estudei dois anos de francês”)
- b. *Yo_i no, no sé, no. Yo_i soy de Historia del Arte. (EP03)*
 (“Eu não, não sei, não. Eu sou de História da Arte.”)

c) Padrão C: O antecedente do sujeito em análise está na oração imediatamente anterior, mas tem outra função sintática.

- (9) a. *A mí me_i tiemblan las piernas cada vez [que Ø_i veo sangre]. (EE01)*
 (“Me tremem as pernas [= as minhas pernas tremem] cada vez que vejo sangue.”)
- b. *y me recordó a ella_i, Ella_i se murió cuando yo tenía once (EP04)*
 (“e [isso] me lembrou ela. Ela morreu quando eu tinha onze [anos]”)

d) Padrão D: O antecedente também é sujeito, mas há pelo menos uma oração interveniente entre a que contém o pronome em análise e a oração em que se localiza o seu antecedente. Nos exemplos abaixo, a(s) oração(ões) interveniente(s) aparece(m) entre chaves.

- (10) a. \emptyset_i *Considero también* { [*que la misma forma de... de rechazar el inglés psicológicamente del puertorriqueño lo vive el maestro*] . *En el aspecto del inglés, hay afortunadamente una serie de cosas que lo mantendrán relegado al plano que debe tener, que es de una segunda lengua...* } *Así que... a modo de un breve resumen de lo anterior, \emptyset_i considero que el inglés continuará siendo una segunda lengua* (EP10)

(“Considero também que o professor vive a mesma forma de rejeitar o inglês psicologicamente do porto-riquenho. No aspecto do inglês, há felizmente uma série de coisas que o manterão relegado ao plano que deve ter, que é de uma segunda língua... Assim, como um breve resumo do anterior, considero que o inglês continuará sendo uma segunda língua”)

- b. *con la época que yo_i era estudiante* { [*que esa era la gran actividad de la universidad*] } . *Yo_i llegué a la universidad en agosto*

(“com a época que eu era estudante, que essa era a grande atividade da universidade. Eu cheguei à universidade em agosto”)

e) Padrão E: Há oração(ões) intervenientes, como no padrão D, e o antecedente tem função diferente de sujeito.

- (11) a. *ese caso hizo mucha impresión en mí_i* { *No era un caso* } [*donde \emptyset_i estaba ganando dinero*] (EP12)

(“esse caso causou muita impressão em mim. Não era um caso onde eu estava ganhando dinheiro”)

- b. *pero a nosotros nos_i sirve mucho*, { [*porque es la práctica real en un centro de cómputos*] } . *Nosotros_i estamos como un empleado cualquiera* (EA04)

(“mas serve muito para nós, porque é a prática real em um centro de computação. Nós estamos [lá] como um empregado qualquer”)

A tabela a seguir mostra os percentuais de sujeitos nulos para cada um dos cinco padrões no espanhol europeu, no argentino e no porto-riquenho. Nas três línguas, o padrão A, que representa o antecedente mais acessível, privilegia o sujeito nulo, enquanto o padrão E, rótulo do sujeito que tem a identificação por um antecedente mais dificultada, por conta da maior distância

entre o pronome em análise e o antecedente e da diferente função sintática do antecedente, privilegia os sujeitos preenchidos. Observando cada coluna da tabela verticalmente, do padrão A ao E, percebe-se que as frequências de sujeitos nulos vão diminuindo de maneira gradual, o que indica que, nessas três línguas, os paradigmas verbais não têm riqueza plena, sendo necessária, para a identificação de um sujeito nulo, a coindexação com um antecedente.

TABELA 4
Sujeitos nulos no espanhol segundo o padrão sentencial

Padrão	Espanhol europeu	Espanhol argentino	Espanhol porto-riquenho
A	142/162 (88%)	156/186 (82%)	108/124 (87%)
B	290/342 (85%)	298/373 (80%)	135/199 (68%)
C	94/119 (79%)	80/114 (70%)	54/90 (60%)
D	321/464 (69%)	242/437 (55%)	159/340 (47%)
E	90/151 (60%)	61/111 (55%)	47/103 (46%)

Os percentuais exibidos na Tabela 4 são reproduzidos no Gráfico 3, que inclui, ainda, os resultados para os padrões sentenciais no italiano (MARINS, 2009) e no PB (BARBOSA, DUARTE e KATO, 2005). As barras mais escuras, que representam a língua italiana, podem ser separadas em dois grupos de percentuais semelhantes: os percentuais para os padrões A, B e C, que oscilam em torno dos 92% (MARINS, 2009), e as frequências de sujeitos nulos nos padrões D e E – 78% e 83%, respectivamente, segundo Marins (2009). O que diferencia os padrões A, B e C, de um lado, dos padrões D e E, de outro, é apenas a distância do antecedente.

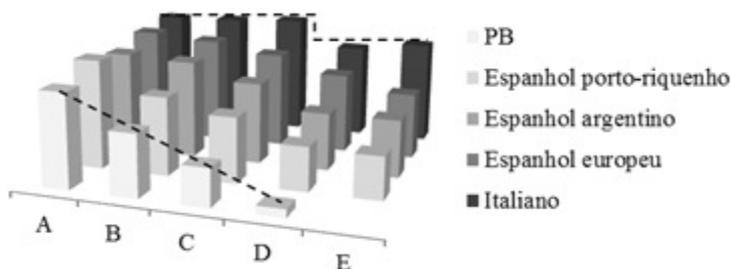


GRÁFICO 3: Percentuais gerais de sujeitos nulos em cinco línguas

Parece, portanto, que, no italiano, apenas a distância do antecedente influi na seleção entre sujeito nulo e pleno. Não importa a identidade de função sintática entre o pronome em análise e o antecedente nem a existência ou inexistência de relação de subordinação entre a oração que tem o pronome codificado e a oração que contém seu antecedente. Isso é compatível com uma língua que tem um paradigma rico, capaz de identificar a referência dos sujeitos nulos sem a necessidade de coindexação com um antecedente acessível.

No outro extremo, o PB apresenta um resultado mais gradual ao longo dos quatro padrões – o padrão IV do estudo de Barbosa, Duarte & Kato (2005), cujo resultado aparece na quarta barra do PB no gráfico, equivale a um universo que une os dados que classificamos, agora, como padrão D e padrão E. O PB é a língua, dentre as que aparecem no gráfico, que tem o paradigma mais pobre em oposições entre as pessoas gramaticais, como será demonstrado na seção seguinte. É de se esperar, portanto, que a atuação da condição de referência seja mais forte nessa língua, e os resultados mostrados no gráfico confirmam isso.

4. A riqueza flexional do verbo: uma proposta mais abrangente

Como demonstrado, ao investigar com dados reais de fala, fica muito difícil classificar todas as línguas quanto ao PSN apenas com as marcações + e -. Foi proposta, então, uma escala, em que as línguas podem ser ordenadas, localizadas mais próximas ou mais afastadas dos protótipos de LSN ou de língua de sujeito pleno, esquematizada na figura abaixo. E, se estamos assumindo que o parâmetro tem mais de duas possibilidades de realização, o mesmo faremos para a riqueza flexional: em vez de estabelecer um limite que separe o que é um paradigma rico do que é um paradigma pobre, como fizeram Chomsky (1981) e Roberts (1993), procuraremos ordenar as línguas comparadas aqui quanto ao grau de riqueza de seus paradigmas verbais.

[-] PB — espanhol porto-riquenho — espanhol argentino
— espanhol europeu — italiano [+]

FIGURA 2: Escala para o PSN

A ordenação das línguas na Figura 2, com o italiano mais próximo ao comportamento prototípico de LSN e o PB mais afastado, segue as frequências gerais de sujeitos nulos encontradas para essas línguas nas diversas análises e apresentadas no Gráfico 2. Está de acordo, também, com a maior/menor atuação dos padrões sentenciais na representação do sujeito referencial: esse grupo de fatores é mais atuante no italiano e menos atuante no PB, segundo a análise proposta para o Gráfico 3. Nesta seção, mostraremos que o grau de riqueza flexional também segue a mesma ordenação: o italiano é a língua, entre as cinco, que tem o paradigma verbal mais rico, o espanhol europeu o tem um pouco empobrecido, o argentino um pouco mais, o porto-riquenho mais ainda e, finalmente, o PB tem o paradigma mais pobre em oposições entre as pessoas gramaticais.

Começamos pelo paradigma flexional do italiano, que apresenta, no modo subjuntivo, neutralização entre a primeira e a segunda pessoa do singular no pretérito (*io/tu parlassi*) e entre as três pessoas do singular no presente (*io/tu/lui parli*). Porém, Marins (2009) identificou, em seu estudo, que os falantes usam algumas estratégias para contornar essa falta de oposição entre pessoas gramaticais: a principal delas é evitar o uso do subjuntivo, substituindo-o pelo modo indicativo. Portanto, em todos os tempos verbais efetivamente usados na fala, os do indicativo, as formas verbais correspondentes a cada pessoa gramatical são diferentes (como, por exemplo, no pretérito imperfeito: *parlavo, parlavi, parlava, parlavamo, parlavate, parlavano*), o que nos permite dizer que o paradigma flexional do italiano é plenamente rico.

O paradigma do espanhol europeu é um pouco mais pobre porque, diferentemente do do italiano, tem formas idênticas para a primeira e a terceira pessoas do singular em dois tempos verbais do indicativo – pretérito imperfeito (*yo/él hablaba*) e futuro do pretérito – e dois do subjuntivo (presente e pretérito). Seguindo para a esquerda na escala da Figura 2, chegamos ao espanhol argentino, que, além de apresentar o mesmo sincretismo observado no europeu, tem um paradigma um pouco mais empobrecido, por conta da substituição do pronome de segunda pessoa do plural *vosotros* por *ustedes*. Essa mudança gerou o sincretismo que neutraliza a segunda e a terceira pessoas do plural em todos os tempos verbais do indicativo e do subjuntivo (como, por exemplo, no pretérito imperfeito: *ustedes/ellos hablaban*).

No espanhol porto-riquenho também se observam os sincretismos apontados para o espanhol argentino, mas, além disso, apaga-se o /s/ final não só de verbos, mas também de substantivos. Nos verbos, esse apagamento não afeta a riqueza flexional quando atinge uma forma de primeira pessoa do plural, como *hablábamos* (pretérito imperfeito do indicativo), pronunciada *hablábamo* (forma verbal que ainda se distingue das demais), mas cria um sincretismo entre a segunda e a terceira pessoas do singular em todos os tempos verbais, exceto o pretérito perfeito do indicativo, cuja forma de

segunda pessoa do singular não termina em /s/ (*bablaste*). E, nos quatro tempos verbais em que primeira e terceira pessoas do singular são idênticas no espanhol, o apagamento do /s/ da segunda faz com que as formas verbais que representam as três pessoas do singular não se diferenciem na pronúncia.

O PB, por fim, localizado no extremo esquerdo da escala proposta por conta do comportamento apresentado para o PSN, tem o paradigma mais pobre das cinco línguas descritas neste trabalho. Além da substituição de *vós* por *vocês* (análoga à substituição de *vosotros* por *ustedes*, no espanhol das Américas, que neutraliza a segunda e a terceira pessoas do plural: *vocês/eles falavam*), também é usado o pronome *você* no lugar de *tu*, o que cria um sincretismo entre a segunda e a terceira pessoas do singular em todos os tempos verbais (*você/ele falava*), inclusive no pretérito perfeito, diferentemente do que foi observado no espanhol porto-riquenho, porque no PB a mudança é no quadro pronominal, e não fonológica. A generalização do uso do pronome *a gente* (associado a formas verbais idênticas às de terceira pessoa do singular), em substituição a *nós*, acaba por neutralizar, no pretérito imperfeito e no futuro do pretérito do indicativo, e em todo o subjuntivo, as três pessoas do singular mais a primeira do plural (*eu/você/ele/a gente falava*).

Dessa forma, o ranqueamento das línguas proposto na Figura 2, com base nas frequências gerais de sujeitos nulos, traduz as diferentes “nuances comportamentais” das cinco línguas analisadas em relação à representação do sujeito, como, por exemplo, a maior ou menor atuação do grupo de fatores *condição de referência*, explorado neste artigo. A mesma ordenação das línguas na escala é obtida da análise dos graus de riqueza dos paradigmas verbais. A relação entre a riqueza flexional e o PSN, da maneira como apresentamos aqui, parece explicar unificadamente os diferentes comportamentos dessas línguas em relação ao parâmetro: a ampliação da comparação, incluindo a análise de mais línguas, poderia sustentar a validade dessa explicação.

5. Considerações finais

Partindo de estudos anteriores sobre a mudança paramétrica em direção aos sujeitos plenos no PB (DUARTE, 1993, 1995), buscamos, através de análises do italiano, do espanhol europeu e do argentino, descrições empíricas do comportamento das línguas positivamente marcadas em relação ao PSN, e, no espanhol porto-riquenho, a observação de uma língua que também passa por mudança em direção ao preenchimento do sujeito, como o PB. Para permitir uma melhor caracterização da mudança no PB, as análises dessas línguas, todas feitas com base em dados de fala real e em metodologia semelhante à aplicada aos estudos sobre o PB, e associando pressupostos gerativistas e variacionistas, foram comparadas, o que acabou revelando diferenças entre todas essas línguas quanto ao comportamento delas em relação ao PSN. Começamos apresentando a análise de alguns aspectos que apontam o italiano como uma LSN prototípica e, em seguida, passamos à análise comparativa do italiano e do espanhol, com foco na condição de referência, grupo de fatores codificado com base na maior ou menor acessibilidade do antecedente do sujeito em análise.

A comparação das análises levou à proposta de uma escala para o PSN, em que as línguas podem ser localizadas e ordenadas em relação aos polos que representam os comportamentos prototípicos de língua [+ sujeito nulo] e língua [- sujeito nulo], já que, trabalhando com coleta de dados de fala real, resultou difícil agrupar todas as línguas investigadas apenas com as marcações positiva e negativa do PSN, como prega a literatura gerativista. As línguas inicialmente foram ordenadas, no *continuum*, segundo as porcentagens gerais de apagamento do sujeito referencial, ficando o italiano mais próximo do protótipo de LSN e o PB, mais afastado. A análise da condição de referência, por meio de um grupo de fatores que codifica cada dado se sujeito nulo ou pleno em um dos cinco padrões sentenciais, mostrou baixíssima influência dos fatores na representação do sujeito no italiano e alta atuação no PB, confirmando a ordem proposta na escala.

A visão contínua do parâmetro nos inspirou a rever a relação entre a ocorrência de sujeitos nulos e a riqueza de oposições do paradigma verbal, estabelecida por Chomsky (1981) e refinada por Roberts (1993). Esta última proposta estabelece um limite de empobrecimento do paradigma (a saber, a presença de um sincretismo entre as formas verbais de duas pessoas gramaticais) que ainda permite a identificação dos sujeitos nulos, que, se ultrapassado, faz com que a flexão não mais identifique a referência dos sujeitos nulos. A partir do elenco de todas as formas verbais utilizadas pelos informantes das amostras das cinco línguas analisadas, verificamos que as línguas que têm o paradigma verbal mais empobrecido (por conta de alterações no quadro pronominal e/ou queda fonológica de segmentos finais) são as que têm as menores taxas de sujeitos nulos, portanto a ordenação das línguas na escala proposta para o parâmetro também reflete o grau de riqueza dos paradigmas verbais dessas línguas.

Assim, é possível estabelecer, como causa para a mudança paramétrica em direção ao preenchimento do sujeito no PB, o empobrecimento das oposições entre as pessoas gramaticais no paradigma verbal, iniciado com a substituição dos pronomes *tu* e *vós* por *você(s)* e acentuado, posteriormente, pela generalização do uso de *a gente*, em substituição a nós. Essas mudanças têm, como consequência, neutralizações nos paradigmas de todos os tempos verbais. Esse empobrecimento faz com que a flexão verbal perca a capacidade de identificar a referência dos sujeitos nulos, o que impulsiona a tendência ao preenchimento da posição.

Referências

- BARBOSA, P.; DUARTE, M. E. L.; KATO, M. A. A distribuição do sujeito nulo no português europeu e no português brasileiro. *In: ENCONTRO NACIONAL DA ASSOCIAÇÃO PORTUGUESA DE LINGÜÍSTICA*, 16, 2001, Lisboa. *Atas...* Lisboa, 2001, p. 539-550.
- BARBOSA, P.; DUARTE, M. E. L.; KATO, M. A. Null subjects in European and Brazilian Portuguese. *Journal of Portuguese Linguistics*, v. 4, n. 2, p. 11-52, 2005.
- CHOMSKY, N. *Lectures on government and binding*. 2. ed. Dordrecht: Foris, 1982. [1. ed. ing. 1981]
- CRESTI, E.; MONEGLIA, M. (Orgs.) *C-ORAL-ROM: integrated reference corpora for spoken Romance languages*. Amsterdam/Philadelphia: John Benjamins, 2005. (Studies in Corpus Linguistics, 15)
- CYRINO, S. M. L.; DUARTE, M. E. L.; KATO, M. A. Visible subjects and invisible clitics in Brazilian Portuguese. *In: KATO, M. A.; NEGRÃO, E. V. (Orgs.) Brazilian Portuguese and the Null Subject Parameter*. Madrid: Iberoamericana/Frankfurt: Vervuert, 2000. p. 55-73.
- DUARTE, M. E. L. Do pronome nulo ao pronome pleno: a trajetória do sujeito nulo no português do Brasil. *In: ROBERTS, I.; KATO, M. A. (Orgs.) Português brasileiro: uma viagem diacrônica*. Campinas: UNICAMP, 1993. p. 107-28.
- DUARTE, M. E. L. *A perda do princípio “evite pronome” no português brasileiro*. Tese (Doutorado) – Universidade Estadual de Campinas, Campinas, 1995.
- DURANTI, A.; OCHS, E. Left-dislocation in Italian conversation. *In: GIVÓN, T. (Org.) Syntax and semantics: discourse and syntax*. New York: Academic, 1979. p. 377-415.
- LABOV, W. *Sociolinguistic patterns*. 11. ed. Philadelphia: University of Pennsylvania, 1991. [1. ed., 1972]
- LOPES, C. R. dos S.; DUARTE, M. E. L. De *Vossa Mercê* a *você*: análise da pronominalização de nominais em peças brasileiras e portuguesas setecentistas e oitocentistas. *In: BRANDÃO, S.; MOTA, M. A. (Orgs.) Análise contrastiva de variedades do português: primeiros estudos*. Rio de Janeiro: In-Fólio, 2003. p. 61-76.

MARINS, J. E. *O parâmetro do sujeito nulo: uma análise contrastiva entre o português e o italiano*. Dissertação (Mestrado em Letras Vernáculas) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2009.

MARINS, J. E. ; SOARES DA SILVA, H. O comportamento das línguas românicas em relação ao parâmetro do sujeito nulo. *Signum: Estudos de Linguagem*, Curitiba, v. 12, p. 191-216, 2009.

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

RIZZI, L. *Issues in Italian syntax*. Dordrecht: Foris, 1982.

ROBERTS, I. *Verbs and diachronic syntax*. Dordrecht: Kluwer, 1993.

SAMPER PADILLA, J. A.; HERNÁNDEZ CABRERA, C. E.; TROYA DÉNIZ, M. (Orgs.). *Macrocorpus de la norma lingüística culta de las principales ciudades del mundo hispánico*. Las Palmas de Gran Canaria (Espanha): Universidad de Las Palmas de Gran Canaria/ Associação de Linguística e Filologia da América Latina, 1995. CD-ROM.

SOARES DA SILVA, H. *O parâmetro do sujeito nulo: confronto entre o português e o espanhol*. Dissertação (Mestrado em Letras Vernáculas) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2006.

SOARES DA SILVA, H. *Evidências da mudança paramétrica em dados da língua-E: o sujeito pronominal no português e no espanhol*. Tese (Doutorado em Língua Portuguesa) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2011.

WEINREICH, U.; LABOV, W.; HERZOG, M. I. *Fundamentos empíricos para uma teoria da mudança lingüística*. Tradução de Marcos Bagno. São Paulo: Parábola, 2006. [1. ed. ing., 1968]

Recebido para publicação em 31 de agosto de 2012

Aprovado em 13 de novembro de 2012