

A perda do sujeito nulo no português do Brasil: uma abordagem discursivo-funcional

Hella Olbertz (Universiteit van Amsterdam / Universidade de Santiago de Compostela)

1. Introdução¹

Na maioria das línguas românicas, a especificação do sujeito é prescindível quando o referente é conhecido, uma propriedade que as línguas românicas herdaram do latim. Uma exceção notável é o francês, que perdeu essa propriedade entre os séculos XIII e XVI (Harris 1988: 231), exigindo, assim, a expressão lexical ou pronominal do sujeito em qualquer contexto. Ainda que no português brasileiro não seja obrigatório explicitar o sujeito, chama a atenção a frequência com a qual é usado o pronome pessoal de sujeito. O exemplo a seguir ilustra este fenômeno:

- (1) a. *Su hijo* (x_i) *vive en Suíza donde* (x_i) *trabaja en un banco.* (EP)
- b. *O filho deles* (x_i) *mora na Suíça onde* (ele) (x_i) *trabalha num banco* (PB)
- c. *Leur fils* (x_i) *habite en Suisse où* il (x_i) *travaille dans une banque* (F)

No espanhol peninsular (EP), seria estranho o uso de pronome *él* em (1a) para se referir à pessoa mencionada imediatamente antes. No português do Brasil (PB), o uso do pronome *ele* não causaria estranhamento, mas em (1b) seria igualmente natural não o usar. Entretanto, no francês (F), o pronome *il*, em (1c), é obrigatório, ficando agramatical sua omissão.

O português brasileiro tem ainda outra propriedade atípica para uma língua com sujeito nulo, que fica ilustrada em (2):

- (2) *cebola, alho... deixo dourar um pouquinho... e coloco carne moída e vou fritando ela ali até ela mudar de cor para ficar dourada... assim que ela começa a dourar...* (BP, Iboruna 80, 29f, M)

Além da alta frequência do pronome *ela* nesse exemplo, também é notável o fato de o pronome referir-se a uma entidade concreta inanimada, *carne moída*. No espanhol, isso não é possível, já que o pronome pessoal de sujeito quase sempre faz referência a entidades animadas (Real Academia 2009: 1180).² Já no francês, sendo a especificação do sujeito lexical ou pronominal obrigatória, não existe tal restrição.

Dada a alta frequência do pronome de sujeito no português, é lógico que se pergunte se o português brasileiro está se aproximando do francês. Essa pergunta é discutida em detalhe por Kaiser (2009), que conclui ser negativa a resposta. Esta conclusão pode ser devida parcialmente ao fato de ele tratar o problema de um ponto de vista gerativo, como também se vê nos numerosos estudos publicados no Brasil e no exterior sobre o tema do uso do pronome de sujeito (Duarte 1995; Barbosa, Duarte & Kato, 2005; Marins & Soares da Silva 2012; Camacho 2016; entre outros). A perspectiva gerativa tem a restrição importante de considerar só os fatos sintáticos, sem levar em conta os aspectos pragmáticos. O presente trabalho estuda o fenômeno

¹ Agradeço a Marize Mattos dall'Aglio Hattnher, Kees Hengeveld, Taísa Peres de Oliveira e Monielly Serafim sua leitura crítica de versões anteriores deste trabalho. Os erros que ficam são meus.

² Isso não é limitado ao espanhol peninsular. Mesmo no espanhol falado nos EUA o uso do pronome de sujeito com referência a entidades inanimadas é marcado. Otheguy (2014: 380) encontra que em mais de 90% dos casos os falantes usam o pronome pessoal para se referir a pessoas.

da perda do sujeito nulo do PB a partir de uma perspectiva funcional, no sentido de considerar os fatos linguísticos em primeiro lugar a partir das funções que eles cumprem no uso. Somente quando o uso não explica o fato linguístico é que soluções formais são levadas em consideração. No caso extremo em que todas as propriedades gramaticais de uma língua podem se explicar exclusivamente pelo uso, então essa língua é considerada completamente funcional, ou seja “transparente” (cf. seção 2.1 para mais detalhes). De fato, parece que tais línguas não existem (Leufkens 2015; Hengeveld & Leufkens 2018). No presente contexto, interessa-nos apenas a relação entre o uso do pronome pessoal de sujeito e a concordância verbal. Uma língua como o francês, que possui tanto pronomes pessoais obrigatórios como concordância verbal, é opaca a respeito desse traço específico, porque a combinação obrigatória dos dois não é funcional, uma vez que há duas marcas para expressar a referência a só uma entidade. A mudança em curso no português brasileiro pode significar que o PB esteja também se desenvolvendo na direção do francês, ou seja, numa direção mais opaca. Portanto, a pergunta central de este estudo é se o português brasileiro está se tornando mais opaco ou não.

Para responder a essa pergunta, o presente artigo estudará o uso dos pronomes que têm concordância verbal de terceira pessoa do singular. Essa seleção tem dois motivos: (i) a terceira pessoa pode ter referência não humana e (ii), no português do Brasil, pode se usar também na referência ao ouvinte (segunda pessoa do singular) e à primeira pessoa do plural. As línguas a serem estudadas são o português, o espanhol e o francês. No que diz respeito ao português e ao espanhol, são estudadas duas variedades, o português europeu (PE) e o português brasileiro (PB), e, para o espanhol, serão consideradas a variedade peninsular (EP) e a de Porto Rico como representante da variedade caribenha (EC). A ideia de olhar para uma variedade caribenha do espanhol foi inspirada por Marins e Soares da Silva (2012),³ que corretamente constatarem que essas variedades apresentam o uso mais frequente do pronome de sujeito no mundo hispânico.

Este trabalho está estruturado da seguinte maneira. Na seção 2, o conceito de transparência será explicado em mais detalhes, fazendo uso do arcabouço da Gramática Discursivo-Funcional, seguido da apresentação de uma breve visão tipológica do que significa transparência em relação ao uso do pronome de sujeito e da concordância verbal. A seção 3 apresenta a metodologia da pesquisa para o presente artigo. Na seção 4, apresentarei primeiro as motivações funcionais que pode ter o uso do pronome de sujeito de terceira pessoa do singular (3sg), com exemplos principalmente do espanhol, para depois considerar alguns usos sem motivação funcional, principalmente do PB, e investigar as explicações relacionadas ao paradigma verbal reduzido do PB. A seção 5 estará dedicada à comparação do PB com o francês para, na seção 6, serem discutidas algumas predições dos possíveis desenvolvimentos do PB. Na seção 7 apresentarei minhas conclusões.

2. Tipos de referência e concordância

Esta seção consiste em duas subseções, na primeira das quais (2.1), refinarei o conceito de transparência fazendo uso da Gramática Discursivo-Funcional. A segunda (2.2) está dedicada à aplicação desse conceito a três línguas que representam diferentes tipos.

³ Os autores incluem em seu estudo também o italiano, que segundo eles tem um uso ainda menos frequente do pronome de sujeito do que o espanhol peninsular. A comparação dos dados italianos com os espanhóis no C-Oral-Rom não confirma essa diferença: num estudo preliminar encontrei no italiano 4,62% (2.953 / 64.797) de sujeitos pronominais por verbo finito em orações declarativas contra só 4,34% (2.368 / 53.939) no espanhol peninsular. (Além da restrição à ilocução declarativa, foram contados só os pronomes de sujeito que precediam o verbo, separados dele no máximo por clíticos e/ou partículas). Sendo a diferença entre as duas línguas mínima, não considero o italiano no presente trabalho.

2.1 Referência, concordância e transparência

Para explicar o que é transparência, é necessário explicar alguns conceitos básicos da Gramática Discursivo-Funcional (GDF) (Hengeveld & Mackenzie 2008). Esse arcabouço consiste de uma Gramática como centro de uma teoria mais ampla de interação verbal, tal como fica apresentada de uma maneira simplificada na Figura 1.

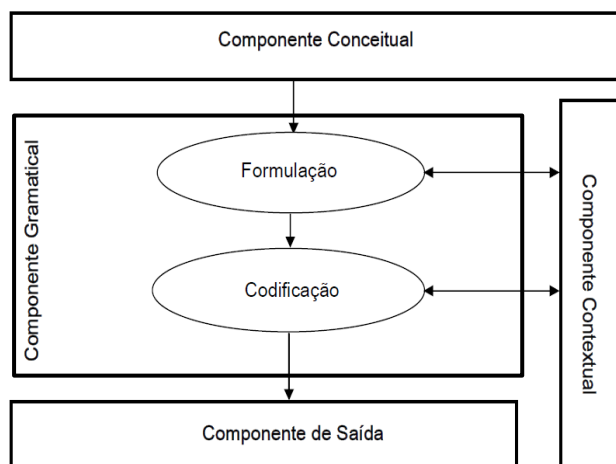


Figura 1: O componente gramatical como centro de uma teoria de interação verbal (Hengeveld & Mackenzie 2012: 44)

O Componente Conceitual é responsável pela conceptualização pré-linguística do que o falante pretende comunicar. O Componente Contextual contém a informação situacional e contextual relevante para uma interação eficaz (ou ‘feliz’ em termos de Searle 1969). Finalmente, o Componente de Saída transforma o *input* da gramática em realizações fonéticas, ortográficas ou de sinais. As flechas na figura indicam a direção da interação da gramática com os três componentes não gramaticais; só no caso do Componente Contextual essa interação é bidirecional, uma vez que as informações situacionais e contextuais servem como *input* da Gramática e que a produção da Gramática, por sua vez, entra no Componente Contextual. Finalmente, é importante notar que apenas o Componente Gramatical é considerado como linguístico no sentido estrito do termo.

O Componente Gramatical integra de uma maneira dinâmica a realização do enunciado, que passa por duas etapas consecutivas, (i) a Formulação e (ii) a Codificação. Como se pode apreciar na Figura 1, essas duas etapas formam o centro do Componente Gramatical (que conta também com um Léxico e outros primitivos não indicados aqui). A Figura 2 mostra a estrutura interna da Formulação e da Codificação.

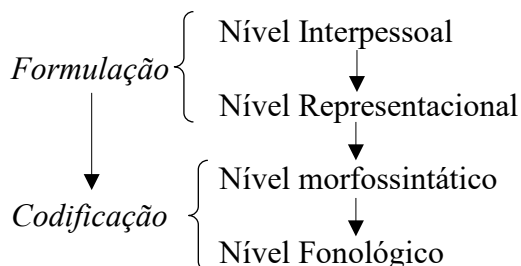


Figura 2: Formulação e codificação dos enunciados na GDF

A Formulação contém todos os elementos pragmáticos (Nível Interpessoal) e semânticos (Nível Representacional) que caracterizam um enunciado. Mais concretamente, o Nível Interpessoal é o nível da expressão das intenções comunicativas do falante, e o Nível Representacional é o nível do conteúdo descritivo do enunciado. Dito de outra maneira, o Nível Interpessoal é responsável por o que o falante *faz* no momento de falar, e o Nível Representacional é responsável por o que o falante *diz* no momento de falar. Como a intenção comunicativa do falante é considerada como primária em relação ao conteúdo que ele comunica, o Nível Interpessoal é hierarquicamente superior ao Nível Representacional. Depois da Formulação, o enunciado passa pela Codificação nos Níveis Morfossintático e Fonológico, nos quais já não é acrescentado nem subtraído conteúdo semântico e/ou pragmático.

A partir dessa conceptualização⁴ da formação dinâmica do enunciado, a transparência pode se definir como a correspondência biunívoca entre os quatro níveis.

Para concretizar essa ideia, consideremos um aspeto concreto, relevante no presente contexto: o Ato de Referência (R_1),⁵ que forma parte dos Atos Discursivos declarativos. O correlato do Ato de Referência no Nível Representacional é um “indivíduo”, ou seja, uma entidade com referentes concretos ou abstratos. No presente contexto, só nos interessam os indivíduos concretos, animados ou inanimados, que são representados na GDF como (x_1). Na Codificação, só nos interessa a realização linguística dos Atos de Referência dos indivíduos no Nível Morfossintático. Os exemplos inventados em (3) são ilustrativos das relações transparentes e opacas:

- (3) a. *Maria escreve bem.*
 b. *Escreve bem.*

Como em todas as línguas românicas de sujeito nulo, (3b) pode ocorrer no caso de o referente do sujeito ser conhecido por já ter sido mencionado anteriormente. Nas representações na GDF a seguir, a marca ‘[-S, -A]’ quer dizer que o referente não é nem o falante ($S = speaker$), nem o ouvinte ($A = addressee$), ou seja, é a terceira pessoa; as flechas horizontais indicam as relações hierárquicas entre os três níveis relevantes.

- (3’) a. (R_1 : Maria) \rightarrow ($^f x_i$) \rightarrow (Np) (V+3sg)
 (3’) b. ($^c R_1$: [-S, -A]) \rightarrow ($^f x_i$) \rightarrow (V+3sg)

A representação (3’a) do exemplo (3a) contém um Ato Referencial, a Maria,⁶ que, no Nível Representacional, corresponde a um indivíduo de sexo feminino, indicado pelo sobrescrito ‘f’. Na realização morfossintática, há duas marcas, o sintagma nominal (Np), que corresponde ao Ato Referencial (R_1), e a expressão da concordância verbal de terceira pessoa do singular (V+3sg). Ou seja, nesse caso, não existe uma relação biunívoca entre os níveis, porque há duas realizações morfossintáticas para um só item nos níveis da Formulação. Na representação (3’b), o sobrescrito ‘c’ no Ato Referencial (R_1) indica que o referente foi recuperado do contexto, ou mais precisamente, do Componente Contextual.⁷ Aqui sim parece haver uma relação transparente: há só um item em cada nível. Assim, no português parece haver expressões referenciais

⁴ Neste trabalho só são considerados, de uma maneira simplificada, os aspectos da teoria relevantes para o presente estudo. Uma apresentação mais detalhada da GDF escrita em português pode ser encontrada em Hengeveld & Mackenzie (2012). Já em inglês, Keizer (2015) representa uma introdução recente à teoria e uma introdução breve e parcial se encontra em Keizer & Olbertz (2018).

⁵ No final deste artigo há uma lista de abreviações usadas no texto e nas glosas.

⁶ O Nível Interpessoal contém só léxico sem significado descritivo. Por isso, os nomes próprios formam parte do léxico interpessoal.

⁷ O funcionamento do Componente Contextual e sua interação com a gramática são descritos em detalhe em Hengeveld & Mackenzie (2014).

tanto opacas como transparentes. Na seção a seguir, veremos que a transparência e a opacidade que dizem respeito à referência ao sujeito podem ser uma propriedade estrutural das línguas.

2.2 Tipologia

A presente seção começa com a discussão de dois tipos de línguas nas quais a especificação do referente do sujeito é obrigatória, o sueco e o francês, para depois considerar as línguas de sujeito nulo a partir de uma perspectiva tipológica.

As línguas escandinavas continentais (dinamarquês, norueguês e sueco) sempre têm um sintagma nominal para fazer referência ao sujeito, mas não existe a concordância verbal (Haugen 1990: 170).⁸ No sueco, a forma verbal do presente do indicativo é formada pelo acréscimo de um *-r* ao infinitivo que, exceto no caso de alguns poucos modais e auxiliares, termina em *-a*.

(4) Sueco (Haugen 1990: 171)

1sg	<i>jag</i>	<i>talar</i>	(eu)	falo
2sg	<i>du</i>	<i>talar</i>	(você)	fala
3sg	<i>han / hon</i>	<i>talar</i>	(ele / ela)	fala
1pl	<i>vi</i>	<i>talar</i>	(nós)	falamos
2pl	<i>ni</i>	<i>talar</i>	(vocês)	falam / (o senhor / a senhora) fala
3pl	<i>de</i>	<i>talar</i>	(eles / elas)	falam

O passado é um pouco mais complicado, porque existem duas classes de verbos, os assim chamados ‘fracos’ (regulares) e os ‘fortes’ (irregulares). Os fracos formam o passado pela adição da desinência *-de* ao infinitivo, e os fortes não têm desinência, mas mudam a vogal da raiz (Haugen 2000: 171), mas em nenhuma dessas formas há concordância. Em razão da inexistência da morfologia verbal de pessoa e número, a referência é expressa exclusivamente pelo sintagma nominal lexical ou pronominal no sueco moderno. Isso implica que, no sueco e nas outras línguas escandinavas continentais, a referência ao sujeito é estruturalmente transparente, pois, em princípio, há só uma realização morfossintática por Ato Referencial (R₁) e indivíduo (x₁):

$$(4') \quad (R_1) \rightarrow (x_1) \rightarrow (Np)$$

Denominarei o tipo de língua exemplificado aqui de ‘referência nominal única’.

O francês também requer um sintagma nominal lexical ou pronominal para a referência ao sujeito, mas, diferentemente das línguas escandinavas continentais, tem concordância verbal. Assim, o paradigma verbal do verbo francês *parler* ‘falar’,⁹ no presente do indicativo, é mais complicado do que o paradigma sueco de *talar* em (4).

⁸ Segundo Falk (1993: 156), a concordância verbal se perdeu na escrita no século XX, mas na fala deve ter deixado de existir bem antes.

⁹ O verbo *parler* representa a classe de conjugação mais numerosa e ainda produtiva, análoga à classe em *-ar* do português.

(5) Francês

1sg	<i>je</i>	<i>parle</i>	/parl/	(eu)	falo
2sg	<i>tu</i>	<i>parles</i>	/parl/	(você)	fala
3sg	<i>il / elle</i>	<i>parle</i>	/parl/	(ele / ela)	fala
1pl	<i>nous</i>	<i>parlons</i>	/par'lõ/	(nós)	falamos
2pl	<i>vous</i>	<i>parlez</i>	/par'le/	(vocês)	falam / (o senhor / a senhora) fala
3pl	<i>ils / elles</i>	<i>parlent</i>	/parl/	(eles / elas)	falam

No francês escrito, que apresenta o paradigma verbal do francês antigo, não adaptado às mudanças na pronúncia, há só um sincretismo entre o singular na primeira pessoa (1sg) e terceira pessoa (3sg). Na pronúncia, o paradigma é muito mais simples, consistindo de apenas três formas, que, contudo, não são referenciais, pois não podem ser usadas sem que esteja presente um sintagma nominal lexical ou pronominal. Portanto, a congruência verbal não é funcional: não serve a uma função comunicativa, mas é uma cópia da informação codificada no sintagma nominal que expressa a referência ao sujeito. Isso significa que a congruência tem uma função puramente morfossintática, e que, conseqüentemente, no francês, há uma relação opaca entre os níveis da formulação e da codificação:

(5') (R₁) → (x₁) → (Np) (V+desinência)

Seguindo Hengeveld (2012), vou me referir ao tipo de língua ilustrado aqui como 'congruência sintática'.

À primeira vista, as línguas de sujeito nulo parecem ser transparentes, como as línguas escandinavas, mas na direção oposta, no sentido de que é a desinência verbal a que é referencial, sem que seja necessário especificar um sintagma nominal. Se olhamos para o paradigma do espanhol peninsular, há de fato seis formas diferentes, uma para cada combinação de pessoa e número gramatical:

(6) Espanhol Peninsular

1sg	<i>hablo</i>	(eu)	falo
2sg	<i>hablas</i>	(você)	fala
3sg	<i>habla</i>	(ele / ela)	fala / (o senhor / a senhora) fala
1pl	<i>hablamos</i>	(nós)	falamos
2pl	<i>habláis</i>	(vocês)	falam
3pl	<i>hablan</i>	(eles / elas)	falam

Porém há um problema quando se trata da terceira pessoa, porque, nesse caso, para poder comunicar eficazmente, precisamos saber quem é o referente, e conseqüentemente é necessária sua especificação lexical em algum lugar. Dado esse conhecimento, é possível entender as referências que seguem pelo contexto anterior, ou seja, é suficiente a marca verbal. Geralmente isso se produz quando esse referente é topical e segue sendo o tópico. Considerem o exemplo construído a seguir:

- (7) a. *Pedro piensa en María.*
'Pedro pensa em Maria.'
b. *La echa de menos.*
'Está com saudade dela.'
c. *Está triste.*

Isso significa que, apesar do paradigma verbal diferenciado em (6), o espanhol não é uma língua estruturalmente transparente, mas há concordância na terceira pessoa. As representações das orações em (7) refletem isso:

(7') a. $(R_I: \text{Pedro})_{\text{Top}} \rightarrow ({}^m x_i) \rightarrow (\text{Np}) (\text{V}+3\text{sg})$

(7') b. $({}^c R_K: [-S, -A])_{\text{Top}} \rightarrow ({}^m x_i) \rightarrow (\text{V}+3\text{sg})$

(7') c. $({}^c R_L: [-S, -A])_{\text{Top}} \rightarrow ({}^m x_i) \rightarrow (\text{V}+3\text{sg})$

(7'a) claramente demonstra que temos aqui um caso de concordância: há duas realizações morfossintáticas para só um Ato de Referência (R_I) e um indivíduo (x_i). Portanto, (7'b) e (7'c) só parecem ser transparentes porque seguem a (7'a) e porque o referente pode ser resgatado do Componente Contextual. Contudo, como veremos na seção 3, pode haver várias situações que requeiram a especificação (pro)nominal do sujeito. Dado o papel do contexto com esse tipo de concordância, sigo Hengeveld (2012), denominando o tipo de sujeito nulo românico, exemplificado aqui pelo espanhol, de 'concordância contextual'.¹⁰

	marca verbal + SN opcional	SN obrigatório
Referência	–	<i>Referência nominal única</i> (Sueco)
Concordância	<i>Concordância contextual</i> (Espanhol)	<i>Concordância sintática</i> (Francês)

Tabela 1: Tipos de referência e concordância (na terceira pessoa do singular)

A Tabela 1 resume os resultados desta seção: a combinação do sintagma nominal obrigatório com a concordância dá um resultado completamente opaco (concordância sintática), ilustrado aqui pelo francês; quando o sintagma nominal é obrigatório mas não há concordância, o resultado é completamente transparente (referência nominal única), ilustrado aqui pelo sueco; quando, finalmente, há concordância em combinação com um sintagma nominal opcional, temos uma forma intermediária, parcialmente opaca (concordância contextual).

Obviamente, há mais línguas em cada um desses três tipos. A concordância sintática também se dá no alemão, mas é tipologicamente rara, sendo um dos poucos casos fora das línguas europeias a língua skou, falada na Nova Guinéia (Kees Hengeveld, comunicação pessoal). A concordância contextual é mais frequente, sendo um outro caso o turco (Hengeveld 2012: 475). A inexistência da concordância verbal (concordância nominal única) é um fenômeno tipologicamente muito frequente: Siewierska (2013) estudou mais de 300 línguas e achou 82 línguas do mundo carecendo de marca verbal de pessoa e número.

Esclarecidas estas preliminares teóricas, as três seções seguintes estarão dedicadas ao uso da concordância com a terceira pessoa do singular (3sg) no PB e outras línguas e variedades românicas. Mas antes de apresentar e analisar os resultados de meu estudo, é preciso explicar a maneira como obtive os dados, que apresentarei na seção a seguir.

3. Metodologia

Para o presente estudo, analisei um total de 828.145 palavras de (fragmentos de) entrevistas sociolinguísticas em cinco (variedades de) línguas. Os materiais são comparáveis no sentido de

¹⁰ No que diz respeito ao espanhol, é importante notar que a concordância contextual só é aplicável à terceira pessoa. Nas outras pessoas, a desinência verbal é plenamente referencial, ou seja, nesses casos não se trata de congruência mas de referência verbal.

abranger todas as faixas etárias a partir dos 26 anos, e de haver na faixa educacional relativamente poucos representantes de educação fundamental e predominantemente falantes com educação universitária.

Os fragmentos dos *corpora* usados para as duas variedades do espanhol formam parte do *Proyecto para el estudio sociolingüístico del español de España y de América* (PRESEEA). Os textos do espanhol peninsular (EP), de 167.366 palavras, foram gravados entre 1991 e 1998 em Alcalá de Henares (AdH). As entrevistas do espanhol caribenho (EC) de San Juan de Puerto Rico (SJPR) (fragmento de 166.747 palavras) foram gravadas entre 1990-1999. Os dados do português europeu (PE) (163.390 palavras) vêm de duas fontes: de entrevistas gravadas em 2001 com falantes com educação universitária provindo de C-Oral Rom, e de uma parte do *corpus Português falado* com falantes com faixas educacionais mais baixas, gravados entre 1990-1999. Os textos do Português brasileiro (PB) são transcrições simplificadas¹¹ do projeto Iboruna do norte do estado de São Paulo (fragmento de 171.948 palavras), gravados em 2006. Finalmente, as entrevistas em francês (F), (158.694 palavras), provêm de CES (*Un corpus d'entretiens spontanés*), gravado entre 1980 e 1989. As fontes dos *corpora* serão especificados no final deste trabalho.

Desses textos selecionei ao acaso aproximadamente 100 ocorrências em 3sg por língua e variedade, em presente e passado do indicativo, usando verbos frequentes: *dizer, falar, querer, ir* (só com significado de movimento), *poder, estar / ficar* (só designando lugar) e cognatos espanhóis e franceses. Selecionei exclusivamente orações declarativas nas quais o referente de sujeito é uma entidade de primeira ordem (ou seja, uma entidade concreta), e o sintagma nominal lexical ou gramatical precede o verbo. Essa última restrição é necessária para evitar (sobretudo no espanhol) que o sujeito tenha função focal. Além disso, exclui usos (parcialmente) cristalizados com esses verbos, tais como, por exemplo, *quer dizer*, ou (parcialmente) gramaticalizados, como, por exemplo, *diz que* (Casseb-Galvão 2010).

As duas seções seguintes estarão dedicadas aos resultados dessa pesquisa.

4. O uso do pronome pessoal de sujeito de terceira pessoa: espanhol e português

A presente seção começará com uma apresentação das motivações funcionais para o uso do pronome de sujeito nas línguas de concordância contextual, exemplificado com o espanhol (4.1), para logo considerar o uso funcionalmente não motivado, procurando a explicação no paradigma verbal (4.2).

4.1 Motivações para o uso do pronome de sujeito

O motivo mais óbvio para o uso do pronome é a mudança de referente de sujeito que, essencialmente, é uma motivação semântica, pois se trata aqui do conteúdo descritivo, não do Ato Referencial. Porém, a mudança de referência geralmente também implica uma mudança pragmática, na medida em que o referente de sujeito também tem função topical. Considerem o exemplo a seguir:

- (8) *y ocho días antes de la boda regañaron, y ella se ha casado con otro y él está en Canarias* (EP, AdH 2, 62f, F)¹²

¹¹ As transcrições originais refletem as propriedades fonológicas da fala; para facilitar a legibilidade dos exemplos, normalizei a ortografia.

¹² Nos exemplos há indicações específicas das fontes. À sigla da língua ou variedade e do *corpus* segue o número da entrevista, seguido da especificação da idade e do sexo do falante em minúscula. As letras em maiúscula indicam a faixa educacional: fundamental (F), média (M) e superior (S).

‘e oito dias antes do casamento eles brigaram, e ela se casou com outro e ele está em Canárias’

A partir da não especificação do referente do sujeito na primeira cláusula, podemos concluir que se trata de um Ato de Referência topical. Consequentemente, nos outros Atos Referenciais também deve se tratar de Tópicos, desta vez, novos. Em termos discursivo-funcionais, os tópicos novos são descritos como Tópicos focais.¹³ Os três Atos de Referência, os Indivíduos e sua realização morfossintática poderiam ser representados da seguinte maneira:

- (8') a. *regañaron* (°R_I: [-S, -A])_{Top} → (pl^hX_i) → (3pl)
b. *ella se ha casado ...* (R_J: [-S, -A])_{TopFoc} → (sg^fX_j) → (ella) (3sg)
c. *él está en* (R_K: [-S, -A])_{TopFoc} → (sg^mX_k) → (él) (3sg)

Em (8'a) há um ato de referência contextual com função pragmática de tópico (indicado pelo subscrito Top). A descrição semântica indica que se trata de referente humano (indicado pelo 'h' sobrescrito) múltiplo. Em (8'b) e (8'c) os atos de referência têm função pragmática de tópico focal. Nas descrições representacionais, trata-se primeiro de um referente de sexo feminino (sobrescrito 'f') e, depois, de um referente masculino (sobrescrito 'm'). O efeito da função pragmática TopFoc e a mudança do referente em (8'b) e (8'c) é a dupla codificação dos referentes no Nível Morfossintático.

O *corpus* contém um exemplo parecido no PB, onde o uso do segundo pronome pessoal, *ela*, é motivado da mesma maneira:

- (9) *eles tiveram que deixar o país ... e pediram asilo ... à Suécia ... certo? então ... do Chile ... ela foi para Suécia* (PB, Iboruna 150, 57F, S)

Uma motivação exclusivamente interpessoal é o uso do pronome de sujeito para a ênfase:

- (10) *estudió derecho y medicina / al mismo tiempo para saber / qué le hacía [...] el alcohol / [...] él mismo quiso curarse* (EP, AdH 34, 65f, M)
‘estudou direito e medicina / ao mesmo tempo para saber / o que fazia o álcool nele / [...] ele mesmo quis se curar’

Nesse exemplo, a ênfase é realizada pela expressão do pronome pessoal em combinação com o modificador enfático *mismo*. Mas também pode ser expressa por meios fonológicos, como no exemplo a seguir:

- (11) *NUN: um comerciante?*
ANT: é // um velhote // vai fechar a loja no dia trinta e um de dezembro // especificamente //
NUN: por causa do euro?
ANT: ELE diz / que é por causa do euro / e por causa da velhice / e não sei qué
(PE, C-Oral Rom, conversações informais, adaptado)

Em termos discursivo-funcionais, a ênfase é uma propriedade do Nível Interpessoal. Quando expressa por meios gramaticais, é um operador no Ato Referencial (emph):

¹³ Seguindo as ideias da Perspectiva Funcional da Sentença da Escola de Praga (Firbas 1992), na GDF as dicotomias das funções pragmáticas mais importantes são Tópico – Comentário, Foco – Fundo (cf. Hengeveld & Mackenzie 2008: 88–102, 2012: 53), o qual permite a combinação das funções de Tópico e Foco, para representar o conceito de ‘tópico novo’.

(11') (emph R_i: [-S, -A]) → (sg ^mx_i) → (ele) (3sg)

Um caso diferente é a reativação do referente (topical), que fica ilustrado no seguinte exemplo:

(12) *tengo una hija que es maestra de inglés en la ... en una universidad en Barranquitas. No me acuerdo del nombre de la universidad. Yo sé que es una sucursal de una universidad de aquí en Puerto Rico, pero ella está en Barranquitas.*

(EC, SJPR 32, 50+m, M)

‘tenho uma filha que é professora de inglês na ... numa universidade em Barranquitas. Não me lembro do nome da universidade. Eu sei que é uma faculdade da universidade daqui de Porto Rico, mas ela está em Barranquitas.’

No exemplo (12), a falante interrompe seu enunciado sobre a filha para tentar dar uma explicação sobre o lugar de trabalho dela, o que acaba sendo complicado, com o efeito de ela ter que fazer referência a uma série de outras entidades. Assim, quando a falante volta ao enunciado inicial, resulta necessário se referir à filha explicitamente. Em termos discursivo-funcionais, a referência à filha já não pode ser resgatada do Componente Contextual (Hengeveld & Mackenzie 2014) e por isso é necessário renová-la.

Os três motivos mencionados aqui estão relacionados a vários aspectos funcionais, principalmente aos aspectos interpessoais da interação verbal, mas também à interação da gramática com o contexto e também com o aspecto semântico de esclarecer a identidade do referente ou, em outras palavras, indicar a que entidade extralinguística se está fazendo referência. Este último ponto é de crucial importância para a descrição do uso do pronome de sujeito no PB.

4.2 O uso do pronome sem motivação aparente

Como mencionei na introdução a este artigo, no PB a terceira pessoa do singular tem uma carga funcional considerável, o qual fica claro se comparamos o PB com o EP e as duas variedades do espanhol. Considere a Tabela 2.

	ESPAÑHOL			PORTUGUÊS		
		EP	EC		PE	PB
1sg	<i>yo</i>	<i>habla</i>	<i>hablo</i>	<i>eu</i>	<i>falo</i>	<i>falo</i>
2sg	<i>tú</i>	<i>hablas</i>	<i>hablas</i>	<i>tu</i>	<i>falas</i>	–
3sg	<i>él, ella, usted</i>	<i>habla</i>	<i>habla</i>	<i>ele, ela, você, a gente, o senhor / a senhora</i>	<i>fala</i>	<i>fala</i>
1pl	<i>nosotros, -as</i>	<i>hablamos</i>	<i>hablamos</i>	<i>nós</i>	<i>falamos</i>	<i>falamos</i>
2pl	<i>vosotros, -as</i>	<i>habláis</i>	–	–	–	–
3pl	<i>ellos, ellas, ustedes</i>	<i>hablan</i>	<i>hablan</i>	<i>eles/-as, vocês</i>	<i>falam</i>	<i>falam</i>

Tabela 2: Pronomes e paradigmas verbais no espanhol e português

Das quatro variedades, o EP tem o paradigma mais diferenciado. A concordância de 3sg é usada principalmente para a referência à terceira pessoa, e também para o tratamento formal, sempre menos comum no espanhol peninsular. O EC, como qualquer outra variedade do espanhol das

Américas, carece de uma forma específica para a segunda pessoa do plural, usando *ustedes* com a 3pl para o tratamento informal no plural. O paradigma do português europeu é análogo ao paradigma hispano-americano. O português do Brasil tem o paradigma verbal mais reduzido das quatro variedades ibero-românicas. No que diz respeito à concordância da 3sg, a maior diferença com relação ao português europeu é seu uso tanto para o tratamento informal (*você*) como para o tratamento formal (*o senhor / a senhora*). Assim, em princípio, não é de se estranhar que no PB o pronome pessoal se use com mais frequência do que nas outras três variedades das línguas ibero-românicas, porque existe uma necessidade maior de desambiguar os possíveis referentes. De fato, bastantes casos do uso do pronome de sujeito podem ser explicados dessa maneira:

- (13) *o gerente me chamou e falou – “ah... cê vai ter que pagar” – eu falei – “não... de forma alguma... não sou o caixa... pediram pra ir buscar... e pediram pra não contar... tá de prova que o o guarda foi junto... e acabamos não contando o dinheiro” – ... tá bom... aí ele falou – “bom... a gente vai ter que RESOLVER de uma outra forma... eu vou te mandar embora” – eu falei – “bom se cê for me mandar embora até por mim tudo bem” (PB, Iboruna 115, 50m, S)*

No exemplo (13), *ele* na quarta linha remete ao gerente mencionado na primeira linha, assim, o motivo para esse pronome é a reativação do referente, como já vimos no exemplo (12) do espanhol caribenho. O motivo da especificação do pronome *a gente* é a mudança de referente (mas ver seção 5.1). A especificação do pronome *cê* não é estritamente necessária para a desambiguação da referência, porque a identidade do referente fica clara pelo contexto, mas o fato de se tratar de um pronome de tratamento pode ser motivação suficiente para sua especificação, por motivos inerentes à pragmática do tratamento ao interlocutor.¹⁴

Por outro lado, existem muitos casos nos quais o uso do pronome de sujeito carece de funcionalidade, sobretudo quando se trata da referência à terceira pessoa.

- (14) *ela não quis... ela falou assim que ela achou que a criança que ela tava esperando... apesar que foi uma coisa violenta... ela achou que/ na cabeça dela...* (PB, Iboruna 72, 28f, F)

- (15) *assim e a porteira abriu sozinha... [...] e ele já ficou meio assim... aí ele passou... a hora que ele olhou pra trás... tinha uma mula-sem-cabeça... [...] como ele num tinha medo... ele falou assim que a hora que ele olhou... ele arrepiou... ele falou assim que o cabelo arrepiou assim ((mostrando))* (PB, Iboruna 114, 47f, S)

Tanto em (14) como em (15), o referente do pronome não muda. Nem existe alguma outra motivação funcional, tais como a ênfase, para a repetição frequentíssima do pronome de sujeito. Ainda que com menor frequência, o mesmo fenômeno acontece no espanhol caribenho:¹⁵

¹⁴ O uso do pronome *você* e sua variante *cê* é relativamente frequente no *corpus Iboruna* completo (407.269 palavras), formando 31% dos usos totais dos pronomes de sujeito de 3sg. A relevância da especificação do pronome de tratamento fica confirmada no espanhol de Medellín, que é comparável com o PB pelo fato de *usted* (com concordância de 3sg) ser o pronome de tratamento não marcado, pois é usado tanto para o tratamento formal como informal. No *corpus PRESEEA* de Medellín (578.467 palavras) há 1889 usos de *usted*, o que é equivalente a 56% dos usos de pronomes de sujeito de 3sg. Evidência da importância de *usted* no espanhol de Medellín é sua comparação com o uso dos pronomes de tratamento exclusivamente informais, dos quais há só 156 (33 de *tú* e 126 de *vos*) no *corpus*.

¹⁵ Embora seja difícil saber os motivos da alta frequência do uso do pronome de sujeito no espanhol caribenho, não pode se descartar a hipótese da influência do inglês no caso concreto do espanhol falado em Porto Rico.

- (16) *A: Tiene treinta años y ella, ella...*
B: ¡Es jovencita!
A: Porque ella cayó en una depresión mental, por dos ocasiones. Ella fue a, al hospital.
 (EC, SJPR 49, 20-34m S)
 ‘A: Tem trinta anos e ela, ela ...’
 B: É muito nova!
 A: Porque ela caiu numa depressão mental, duas vezes. Ela foi a, ao hospital.’

Porém, em relação com as outras quatro variedades ibero-românicas, é considerável a quantidade de usos do pronome pessoal de sujeito não motivados no português do Brasil, tal como está apresentado na Tabela 3.

LÍNGUA E VARIEDADE	PRONOME DE SUJEITO	NÃO MOTIVADO
Espanhol peninsular	9	0 0%
Português europeu	23	3 13%
Espanhol caribenho	26	6 23%
Português brasileiro	38	12 31%

Tabela 3: Uso do pronome de sujeito no espanhol e no português

Tanto o número absoluto do pronome de sujeito como o seu uso não motivado resultam ser relativamente altos no PB.

A Figura 3 mostra a relação entre as quatro variedades da perspectiva oposta, a da quantidade dos sujeitos nulos.¹⁶ Aqui, para demonstrar o contraste, foi introduzido o caso do francês, que já não possui sujeitos nulos.

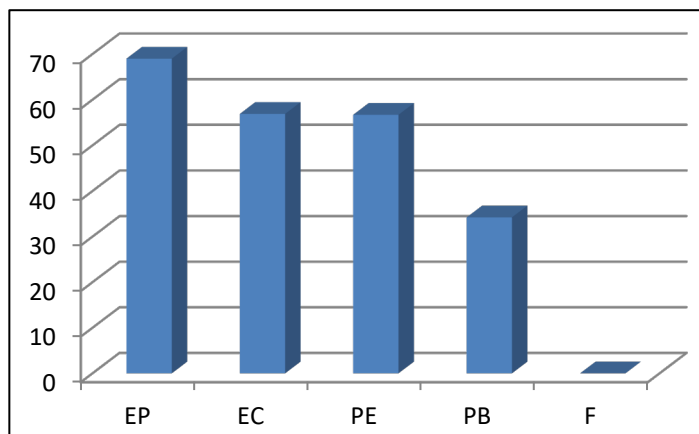


Figura 3: Sujeitos nulos com concordância de 3sg por Vfin

A Figura 3 indica que o espanhol peninsular (EP) tem quase 70% de sujeitos nulos, e que o espanhol caribenho (EC) e o português europeu (PE) têm aproximadamente 57%. A diferença dessas três línguas e variedades com o português brasileiro (PB) é considerável. De fato, com seus 34% de sujeitos nulos, o PB se encontra a meio caminho entre o extremo do EP por um

¹⁶ A Figura 3 é baseada nos dados a seguir: EP: 113 verbos finitos, 78 sujeitos nulos e 35 sujeitos explícitos, dos quais 26 são nominais e 9 pronominais; EC: 116 verbos finitos com 66 sujeitos nulos e 50 sujeitos explícitos, dos quais 24 são nominais e 26 pronominais; PE: 97 verbos finitos com 55 sujeitos nulos e 42 sujeitos explícitos, dos quais 19 são nominais e 23 pronominais; PB: 111 verbos finitos com 38 sujeitos nulos e 73 sujeitos explícitos, dos quais 21 são nominais e 52 pronominais; Francês (F): 124 verbos finitos com todos os sujeitos explícitos, dos quais 4 são nominais e 120 pronominais.

lado e o do francês (F) (0%), por outro. A seção a seguir se ocupará da relação que pode haver entre o PB e o francês.

5. O português do Brasil comparado com o francês

Na sua comparação do desenvolvimento do português brasileiro com o processo de perda do sujeito nulo no francês antigo, Kaiser (2009) menciona dois motivos para descartar a teoria da possível paralela entre as duas línguas românicas. O primeiro argumento guarda relação com o fato de o desenvolvimento de pronomes de sujeito clíticos pertencer à etapa da perda do sujeito nulo no francês (Kaiser 2009: 143–149). Segundo Kaiser (2009: 148), a variante *cê* do pronome *você* não é um clítico no sentido próprio, por admitir a inserção de alguns itens não-clíticos entre o pronome e o verbo, apesar de compartilhar importantes propriedades com os clíticos, como demonstram Martelotta & Cesario (2011: 731–732). Além de talvez ser motivado por argumentos de natureza interna à teoria gerativa, esse argumento só indica que o desenvolvimento do PB não terá um percurso absolutamente idêntico ao desenvolvimento do francês antigo. O segundo argumento de Kaiser é convincente: o português brasileiro não usa os pronomes pessoais em função expletiva, enquanto o uso expletivo no francês medieval já se desenvolveu antes de se perder o sujeito nulo (2009: 150–152). O que Kaiser não explicita é a função que tinha esse pronome expletivo: no francês moderno, o pronome *il* ‘ele’ é usado como expletivo tanto com verbos meteorológicos (17a) como na função existencial (17b):

- (17) a. *normalement il pleut très peu ici* (F, CES 30, 20f, M)
‘geralmente chove muito pouco aqui’
b. *Alors il y a quarante personnes avec lui a bord* (F, CES 23, 60+m, M)
‘Então tem quarenta pessoas com ele a bordo’

Como o uso de expletivos em expressões existenciais não está relacionado com o sujeito nulo,¹⁷ é provável que Kaiser esteja se referindo ao uso do pronome com verbos meteorológicos, pois o uso dos verbos meteorológicos sem sujeito é uma propriedade exclusiva das línguas de sujeito nulo, como demonstram Eriksen *et al.* (2010). Mas os autores também indicam que existem múltiplas maneiras de especificar um sujeito com os verbos meteorológicos. Uma maneira tipologicamente recorrente é o que Eriksen *et al.* (2010: 593) denominam *argument-predicate p[recipitation]-encoding*, correspondendo a ‘a chuva chove’. Essa forma de codificação meteorológica é usada pela língua skou (papua), uma das poucas línguas de concordância sintática fora das línguas europeias (Donohue 2004: 207).

Podemos concluir que, devido ao uso dos verbos meteorológicos sem sujeito expletivo, a perda do sujeito nulo é ainda incipiente no português do Brasil. Para conseguir alguma indicação das possíveis direções que pode tomar o português brasileiro, é interessante observar uma série de paralelos que existem entre o PB atual com o francês.

Na seção 5.1, ocupamo-nos de aspectos da redução (parcial) do paradigma nas duas línguas. A seção 5.2 é dedicada à referência inanimada e, relacionada a ela, a perda do impacto semântico do pronome e sua redução (parcial) a uma marca morfossintática. Finalmente, na seção 5.3, será discutida uma inovação funcional do pronome pessoal sujeito.

¹⁷ Mesmo o italiano, uma língua prototípica de sujeito nulo, faz uso da partícula expletiva *ci* em expressões existenciais paralelas ao exemplo francês em (17b).

5.1 A redução parcial do paradigma

No português brasileiro e no francês, é comum o uso de formas impessoais, ou seja, da 3sg para a referência à primeira pessoa do plural. No francês, é o pronome impessoal *on*, que já no final do século XIX, ainda estigmatizado (Hansen 2013: 128), começou a ser usado para substituir o pronome *nous*.

- (18) *A: Et cela vous plaît euh aller en boîte?*
C: Bon, nous, on ne peut pas encore y aller puisque...
B: On n'y va pas souvent
C: Quand on y va, c'est bien. (F, CES 1, 14m, F)
'A: E vocês gostam de ir na discoteca?
C: Bom, nós, a gente ainda não pode ir porque ...
B: A gente não vai com frequência.
C: Quando a gente vai, é gostoso.'

O pronome português *a gente* é bem menos frequente do que *on*,¹⁸ provavelmente pelo fato de ser mais novo: segundo Martelotta & Cezario (2011: 733), *a gente* se generalizou nos anos 70 do século XX no Brasil.

- (19) *que eles não aprenderam em casa... a educação que a gente vai dar:... né?... que é a/ a pri/ o principal... eu... eu (pretendo assim) às vezes a gente... pensa assim nossa que a gente quer pôr os nossos filhos/ nossos filhos numa boa escola*
(PB, Iboruna 116, 36f, S)

Como se viu na Tabela 2, o uso de *a gente* para a referência à primeira pessoa do plural não é propriedade exclusiva do português do Brasil, mas é usado com maior frequência no PB.¹⁹

Uma das propriedades essenciais do pronome *a gente* é a obrigatoriedade de sua especificação. Sendo o referente menos frequente e diacronicamente mais recente e, por isso, o mais marcado dos três possíveis referentes com concordância de 3sg (tratamento formal e informal; terceira pessoa; primeira do plural), a referência à primeira pessoa do plural quase nunca acontece sem sujeito pronominal. A omissão desse pronome só é possível quando é mencionado no contexto anterior imediato, como nos dois exemplos a seguir:

- (20) *a gente vê e fala: "será que eles tão satisfeitos com isso?"* (PB, Iboruna 80, 28f, F)
(21) *e a gente não pode ser feliz... nem sabe se um dia eu vou ser feliz com ele...*
(PB, Iboruna 68, 27f, F)

No exemplo (20), há dois verbos coordenados imediatamente, construção que mesmo em línguas de concordância sintática não exige a repetição do pronome. No caso de (21), trata-se de duas orações sintaticamente independentes, mas o pronome fica bem próximo. Fora desse tipo de contextos, é improvável a omissão de *a gente*.²⁰

¹⁸ Nos meus dados franceses há 24 casos em 124 verbos com concordância de 3sg de *on* para referir à primeira pessoa do plural, ou seja aproximadamente 19,3 %. Nos dados brasileiros há 7 casos de *a gente* em 111 verbos de 3sg, o que corresponde a só 6,3 %.

¹⁹ No corpus do PB (171.948 palavras) há 662 casos de *a gente* e no corpus do EP (163.390 palavras) só 266.

²⁰ Zilles (2005: 40), que estuda o português falado em Porto Alegre, também indica que a frequência da omissão do pronome *a gente* é mais baixa do que a do pronome *nós*. Porém, ela apresenta um número ainda elevado de sujeitos nulos com *a gente*, mas infelizmente não dá nenhum exemplo.

Nos estudos das duas línguas, tem sido considerada a possibilidade teórica da 1pl desaparecer completamente a favor da alternativa de *on* e *a gente*. No caso do francês, isso significaria a perda da concordância, com a exceção da 2a pessoa de tratamento formal (Hansen 2003: 128–129; entre outros). Para o PB, significaria a perda da concordância da 1pl (Oliveira 2018; entre outros). Porém, meus dados indicam que ainda não é o caso de *a gente* estar substituindo a forma da 1pl completamente na fala informal no PB. Ao contrário, essa última é usada ainda mais do que a alternativa impessoal.²¹ É comum a mistura das duas formas, como no exemplo a seguir:

(22) *tivemos que ficar esperando esperan/... esperando um mo/ uma hora e meia mais ou menos... até que a enfermeira chamou a gente a gente entrôu... tivemos que esperar DE NOVO lá dentro* (PB, Iboruna 74, 26f, M)

Uma segunda propriedade relacionada com a redução parcial do paradigma verbal é a tendência à neutralização da diferença entre a 3sg e a 3pl. No caso do PB, trata-se do uso informal falado, não aceito de um ponto de vista normativo:

(23) *eles não pode invadir a área ... se eles invadir a área o juiz pode mandar voltar* (PB, Iboruna 119, 54m, S)

(24) *eles tava... ali perto da/ sabe? da... ali no plantão do América lá... porque tava sem documento tava sem nada... tava embriagado...* (PB, Iboruna 79, 27m, M)

Como os exemplos (23) e (24) demonstram, a falta de concordância de número com o sujeito no plural não é restrita aos falantes de um nível educacional baixo, embora possivelmente seja um fenômeno socialmente marcado.²² No francês, no entanto, não há escolha, pois nos paradigmas verbais regulares existe um sincretismo fonológico entre a 3sg e a 3pl, como já vimos em (5) na seção 2.2. Porém, esse sincretismo não acontece no perfeito composto, que funciona como tempo passado não marcado no francês, devido à complexidade morfológica do auxiliar:

(25) a. *Ils ont fait leur fortune à Paris* (F, CES 16, 55m, F)
 ‘Eles se fizeram ricos em Paris’²³
 b. *Il a fait sa fortune à Paris*
 ‘Ele se fez rico em Paris’

Contudo, podemos concluir que a neutralização entre o singular e o plural na terceira pessoa está bem mais avançada no francês, por ser uma propriedade inerente da língua, do que no português brasileiro, no qual essa neutralização constitui uma variação na fala.

²¹ No presente do indicativo dos verbos analisados (*dizer, estar, falar, ficar, ir, poder, querer*), há 48 formas com concordância de 1pl e 30 formas com *a gente* com concordância de 3sg.

²² No *corpus* Iboruna, a referência ao plural de terceira pessoa se expressa mais frequentemente pela 3sg na fala dos informantes com ensino fundamental. Esse tema é discutido exaustivamente tanto dentro da teoria “crioulista” da origem do PB, segundo a qual as simplificações na congruência verbal se devem ao contato com os escravos (Baxter 2009), entre outros, como do ponto de vista “anti-crioulista” variacionista, que defende a origem portuguesa das simplificações; ver exemplo Naro & Scherre (2007).

²³ Esse exemplo ilustra ainda outra propriedade do francês que pode servir para distinguir o plural do singular na terceira pessoa, que é de natureza fonológica, a *liaison* ‘ligamento’: *ils ont* é pronunciado /ilzɔ̃/, ou seja, a sibilante final dos pronomes de plural *ils* ‘eles’ e *elles* ‘elas’ não é muda mas é pronunciada (/z/) quando a palavra que segue começa com uma vogal. Obviamente, esse critério não tem relação com o perfeito composto: a pronúncia de *elles arrivent* ‘elas chegam’ é /ɛlzaˈvi:vɔ̃/ em contraste com *elle arrive* //ɛlaˈvi:vɔ̃/.

5.2 Dessemantização

Como já mencionamos no começo, o pronome de sujeito é usado no português do Brasil para a referência a entidades inanimadas, uma propriedade atípica para as línguas românicas de sujeito nulo:

(26) *Mirassol era [...] uma cidade BELAMENTE administrada... e daí pra cá ela deteriorou..*
(BP, Iboruna 151, 89m, S)

(27) *e aí a hora que ele tiver nessa consistência você desliga o fogo e mexe uns... dois minutinhos e cê continua mexendo pra ele num perder [...] essa consistência de creme...*
(PB, Iboruna 76, 28f, M)

O referente de *ela* em (26) é Mirassol, uma cidade no norte do Estado de São Paulo, e em (27) o referente de *ele* é um creme. Diferentemente dos casos descritos na seção 4, o pronome de sujeito aqui deixa de expressar o sexo do referente, pois as coisas não têm sexo. A única função que resta para o gênero gramatical do pronome é expressar a concordância com o gênero arbitrário do lexema que ocupa a função sintática do sujeito. Assim, a representação simples do primeiro *ele* em (27) teria a seguinte forma:

(27') (R_i: [-S, -A]) → (sg x_i) → (ele) (3sg)

O que distingue essa representação das anteriores é a falta de sobrescrito 'h' para referente humano, 'f' para referente feminino ou 'm' para referente masculino. O que motiva a forma *ele* é a concordância, uma propriedade inerentemente morfossintática (Hengeveld & Mackenzie 2008: 350), sem motivação funcional.

Enquanto no PB o pronome de sujeito de 3sg é usado incidentalmente com referentes não animados, a obrigatoriedade da marcação do sujeito no francês implica o uso sistemático do pronome de sujeito com referentes inanimados. Como, além disso, o pronome *il* é usado como expletivo (cf. os exemplos em (17) acima), os pronomes de sujeito do francês já não conservam semântica nenhuma, mas desempenham uma função exclusivamente morfossintática.

Assim, de novo, o francês resulta estar muito mais avançado no processo relacionado com a perda do sujeito nulo do que o português do Brasil.

5.3 Pragmaticalização

Na medida em que o pronome de sujeito 3sg perde seu impacto semântico, ganha uma função pragmática. Considere o seguinte exemplo:

(28) *esse bairro ele num fica na cidade de Campinas ele fica na::... num fica na cidade fica... próximo à rodovia...* (PB, Iboruna 115, 50m, S)

À primeira vista parece se tratar de um argumento extraposto à esquerda com o pronome pessoal em aposição. Mas é preciso considerar o fato de não haver uma pausa entre o sujeito lexical e o pronome. Tal pausa ocorre, sim, na variante (28a):

(28) a. *esse bairro ele num fica na cidade de Campinas ele fica na... num fica na cidade fica... próximo à rodovia...*

Em (28a), *esse bairro* funciona como um constituinte extra-oracional que, de fato, é um Ato Discursivo parcial, introduzindo o tema sobre o qual o falante vai falar. Dik (1997: 388), denomina esse tipo de constituintes extra-oracionais ‘Tema’ com função de ‘Orientação’. Isso quer dizer que sua função é orientar o ouvinte sobre o que o falante deseja discorrer.²⁴ Agora, Hengeveld & Mackenzie (2008: 94–95) apontam que os Atos Discursivos de Orientação têm uma tendência de se integrar ao Ato Discursivo principal, sendo reinterpretado como Tópicos. É isso o que acontece no exemplo (28). Nesse tipo de caso, o pronome aparentemente redundante, que originalmente funcionava como expressão de referência, chega a ser interpretado como marcador de Tópico. Além de dar um exemplo de uma língua papua,²⁵ eles mencionam o seguinte exemplo do francês, enfatizando as diferenças fonológicas entre a variante com Ato Discursivo de Orientação (29a) e a variante com o pronome em função de marcador de Tópico (29b):

- (29) a. *Mon voisin, il est toujours malade.*
 /mõvwa'zẽ/ /iletuzurmalad/
 ‘Meu vizinho, ele está sempre doente.’
 b. *Mon voisin il est toujours malade.*
 /mõvwazẽiletuzurmalad/
 ‘Meu vizinho ele está sempre doente.’
 (Hengeveld & Mackenzie 2008: 433)

Como os autores demonstram, (29a) contém duas frases entoacionais, indicadas pelas barras oblíquas, enquanto (29b) forma uma frase entoacional única. Assim, podemos concluir que o pronome de sujeito que segue o Tópico sem pausa entre eles é usado como marcador de Tópico.

Um exemplo autêntico da reinterpretação do pronome como marcador de Tópico do francês é (30):

- (30) *donc le patron il peut dire a l'occasion: “Euh bon j'ai plus besoin de vous, vous pouvez partir. [...]”* (F, CES 16, 55m, P)
 ‘assim o patrão ele pode dizer nesse momento: “Bom, eu já não preciso de vocês, podem ir embora.”’

Nesse exemplo, o referente é humano, mas, como já vimos no exemplo (28), esse uso não depende da animação do referente. Nos exemplos do PB a seguir, o pronome funciona como marcador de Tópico para um referente inanimado (31) e para um referente humano (32).

- (31) *... a costela ela solta gordura então cê num pode pôr muito óleo né?*
 (PB, Iboruna 70, 27f, F)

- (32) *o goleiro ele pode pedir o tanto que quiser* (PB, Iboruna 119, 54m, S)²⁶

²⁴ Dik (1997: 389–392) demonstra, por meio de várias evidências, que o Tema, na verdade, é extra-oracional, entre outros porque (i) pode ter uma ilocução distinta da ilocução da oração e (ii) é em caso não marcado, independente do caso que o item correferente tem na oração que segue.

²⁵ “As Reesink has shown for Usan (Reesink 1987) with respect to units marked by the particle *eng*, the sequence of an Orientation Discourse Act and a Nuclear Discourse Act within a Move can come in time to be reinterpreted as a single Discourse Act, with the Orientation Discourse Act becoming the Topic of the Communicated Content of that Discourse Act.” (Hengeveld & Mackenzie 2008: 57)

²⁶ Pode ser mais frequente o uso dessa construção com referentes humanos, já que, geralmente, os referentes humanos são mais prominentes no discurso do que os outros referentes animados e inanimados (Comrie 1989: 185–200).

É importante observar que a condição para o pronome de sujeito funcionar como marcador de Tópico é a redundância semântica e sintática desse mesmo pronome. Assim, só quando tiver deixado de desempenhar as funções próprias dos pronomes pessoais de sujeito, descritas na seção 4, o pronome de sujeito pode chegar a funcionar como marcador de Tópico.

Considerando o desenvolvimento do pronome de sujeito de um ponto de vista da gramaticalização, podemos concluir que, no português do Brasil, o pronome de sujeito está se gramaticalizando de um pronome propriamente funcional, dentro do tipo de concordância contextual, a um tipo com uso de pronome de sujeito obrigatório que, perdendo sua função semântica, ganha uma função pragmática. Esse desenvolvimento do campo semântico ao campo pragmático também se encontra nos processos de gramaticalização nos domínios de aspecto, tempo, modalidade e evidencialidade descritos por Hengeveld (2017) dentro da teoria de Gramática Discursivo-Funcional.

6. Para onde está indo o português brasileiro?

Na seção anterior, demonstramos que o português brasileiro compartilha algumas propriedades com o francês, língua de concordância sintática: a redução parcial do paradigma, a redução da semântica do pronome de sujeito, e a pragmaticalização do pronome de sujeito. Por outro lado, ficaram claras as diferenças entre o francês e o PB, que reside essencialmente no fato de o PB não fazer uso do pronome de sujeito na expressão meteorológica. É possível que o PB, em lugar de se desenvolver na direção de um sistema opaco, venha a se desenvolver na direção do tipo transparente de referência nominal única como o sueco? Para poder discutir essa possibilidade, olhemos primeiro um paradigma hipotético do PB:

Referência	Pronome	Concordância	Forma verbal
primeira pessoa do sg	<i>eu</i>	1sg	<i>falo</i>
–	–	2sg	–
segunda pessoa do sg	<i>você</i> <i>o senhor / a senhora</i>	3sg	<i>fala</i>
terceira pessoa do sg	<i>ele / ela</i>	3sg	<i>fala</i>
primeira pessoa do pl	<i>a gente</i>	3sg	<i>fala</i>
–	–	1pl	–
segunda pessoa do pl	<i>vocês</i>	3pl	<i>fala</i>
terceira pessoa do pl	<i>eles / elas</i>	3pl	<i>fala</i>

Tabela 4: Paradigma hipotético do português brasileiro

Nesse paradigma hipotético, as formas sublinhadas apresentam a generalização das formas que agora são usadas incidentalmente: (i) a alternativa impessoal *a gente* com concordância de 3sg substituiu completamente a forma de 1pl; (ii) a distinção entre o singular e plural na 3a pessoa foi completamente neutralizada. Nesse caso, que demonstrei não ser provável a curto prazo, ficaríamos com só uma forma marcada por pessoa e número, a primeira do singular.

Voltamos agora à pergunta se o PB poderia perder a concordância completamente, para se tornar numa língua transparente do tipo de referência nominal única. Há um argumento essencial contra essa hipótese que, sem dúvida, é muito atraente: é muito pouco provável que a marcação da 1sg se perca. Na interação verbal, a primeira pessoa tem uma função muito importante. Independentemente da língua específica, a primeira pessoa do singular é a mais provável de receber a função pragmática de Tópico (Comrie 1989: 198–199). De fato, no meu *corpus* do PB, a primeira pessoa é a que mais frequentemente é especificada. Dos verbos analisados (*estar, ficar, ir* [só funções lexicais], *falar, poder, querer*) há 144 formas de

concordância de 1sg, versus só 88 formas da 3sg no presente do indicativo. A importância relativa da primeira pessoa do singular não é um caso específico do português brasileiro, como demonstra Posio (2013) em seu estudo do uso dos pronomes de sujeito de 1sg no espanhol e português europeus. Nossos dados confirmam a importância do pronome de 1sg.

	EP	EC	PE	PB	F
<i>Palavras</i>	<i>(167.366)</i>	<i>(166.747)</i>	<i>(163.403)</i>	<i>(171.948)</i>	<i>(158.694)</i>
pronome 1sg	1.114	2.011	1.773	3.381	2.119
outros pronomes	367	2.131	1.623	3.975	7.382
total de pronomes	1.481	4.142	3.396	7.356	9.501

Tabela 5: Pronomes pessoais de sujeito por língua e variedade

A Tabela 5 apresenta o total de pronomes de sujeito nos corpora. No espanhol peninsular e no português europeu, o pronome pessoal de sujeito é de modo absoluto o mais frequentemente especificado. Nas outras duas variedades, o pronome de 1sg é o relativamente mais frequente. Só no francês o pronome *il*, que tem várias funções expletivas, é um pouco mais frequente (2.155 casos) do que o pronome da 1sg. Além disso, não encontrei nenhum caso de falta de concordância na 1sg no PB. Tudo isso é uma forte indicação de que (i) a primeira pessoa é altamente relevante na interação, tanto do ponto de vista específico do PB como do ponto de vista tipológico; e que (ii) a concordância na 1sg é extremamente resistente. Assim, se o PB realmente mudar de tipo de concordância / referência, o mais provável é que se torne uma língua de concordância sintática, ou seja, que conserve a concordância pelo menos na primeira pessoa do singular (1sg), tornando-se obrigatória a especificação do pronome de sujeito.

7. Conclusão

No presente trabalho, mostrei que o português do Brasil é uma língua do tipo de concordância contextual como a maioria das línguas românicas. O PB possui uma série de propriedades atípicas das línguas de concordância contextual, que são parecidas com as línguas de concordância sintática, especificamente ao francês: (i) o pronome de sujeito é usado frequentemente, (ii) o paradigma verbal, já reduzido, está se reduzindo ainda mais, (iii) o pronome de sujeito da terceira pessoa do singular (3sg) está perdendo seu impacto semântico por ser usado com referentes inanimados, (iv) o pronome 3sg está se pragmaticalizando no seu uso em aposição imediata como marca de Tópico.

Por possuir concordância da primeira pessoa do singular, que é a pessoa pragmaticamente mais importante na interação verbal, é improvável que o paradigma verbal do português do Brasil se perca completamente e se desenvolva numa língua de referência nominal única. É possível, assim, que português do Brasil se converta no futuro numa língua de concordância sintática, tornando-se mais opaco.

Abreviações

1	primeira pessoa
2	segunda pessoa
3	terceira pessoa
A	ouvinte (<i>Addressee</i>)
ac	acusativo
Aff	afixo

PB	português brasileiro
c	presente no Componente Contextual
EC	espanhol caribenho
emph	ênfase
EP	espanhol peninsular
F	francês
f	feminino
h	humano (= feminino ou masculino)
inf	infinitivo
m	masculino
Np	sintagma nominal
obj	objeto
PB	português brasileiro
PE	português europeu
pl	plural
ps	passado
R	Referente
S	falante (<i>Speaker</i>)
sbj	sujeito
sg	singular
Top	Tópico
x	indivíduo

Corpora

- [AdH] Moreno, Francisco, Ana Cestero, Isabel Molina & Florentino Paredes. 2002–2007. *La lengua hablada en Alcalá de Henares. Corpus PRESEEA - Alcalá*. 3 vols. Alcalá de Henares: Universidad de Alcalá, Disponível <http://preseea.linguas.net/>; acesso em 15/09/2017.
- [C-Oral-Rom] Cresti, Emanuela & Massimo Moneglia (eds.) 2005. *C-ORAL-ROM: Integrated reference corpora for spoken Romance languages*. Amsterdam: Benjamins.
- [CES] Beeching, Kate. [s.d.] *Un corpus d'entretiens spontanés*. Bristol: UWE. Disponível em <https://www2.uwe.ac.uk/faculties/CAHE/ELC/Documents/iclru/corpus.pdf>. acesso em 15/09/2017.
- [Iboruna] Gonçalves, Sebastião Carlos Leite (org.) [s.d.] *Banco de dados Iboruna: amostras eletrônicas do português falado no interior paulista*, Disponível em <http://www.alip.ibilce.unesp.br/iboruna>; acesso em 15/09/2017.
- [SJPR] Ortiz, Luis & Amparo Morales, Amparo (orgs.) [s.d.] San Juan de Puerto Rico, Corpus PRESEEA. Disponível em <http://preseea.linguas.net/>; acesso em 15/09/2017.

Referências bibliográficas

- Barbosa, Pilar, Maria Eugênia L. Duarte & Mary Kato. 2005. Null subjects in European and Brazilian Portuguese. *Journal of Portuguese Linguistics* 4: 11–52.
- Baxter, Alan. 2009. Concordância verbal. In: Tânia Lobo & Klebson Oliveira (eds.), *África à vista*. 317–337. Salvador: EDUFBA.
- Camacho, José. 2016. The null subject parameter revisited: the evolution from null subject Spanish and Portuguese to Dominican Spanish and Brazilian Portuguese. In: Mary Kato &

- Francisco Ordoñez (eds.), *The morphosyntax of Portuguese and Spanish in Latin America*. 27–48. Oxford: Oxford University Press.
- Casseb-Galvão, Vânia Cristina. 2010. A reanálise fonético-fonológica na gramaticalização de [diskⁱ]. *Signótica* 22/2: 479–494.
- Comrie, Bernard. 1989. *Language universals and linguistic typology*. Second edition. Oxford: Blackwell.
- Dik, Simon C. 1997 *The theory of Functional Grammar*, edited by Kees Hengeveld. Part 2: *Complex and derived constructions*, Berlin: Mouton de Gruyter.
- Donohue, Mark. 2004. *A grammar of the Skou language of New Guinea*. Manuscrito, disponível em www.academia.edu, acesso em 8/11/2017.
- Duarte, Maria Eugênia L. 1995. *A perda do princípio “evite pronome” no português brasileiro*. Tese de doutorado. Campinas: IEL-Unicamp.
- Eriksen, Pål, Seppo Kittilä & Leena Kolehmainen. 2010. Linguistics of weather: cross-linguistic patterns of meteorological expressions. *Studies in Language* 34/3: 565–601.
- Falk, Cecilia. 1993. Non-referential subjects and agreement in the history of Swedish. *Lingua* 89: 143–180.
- Firbas, Jan. 1992. *Functional Sentence Perspective in written and spoken communication*. Cambridge: Cambridge University Press.
- Hansen, Anita Berit. 2013. Diaphasic variation and change in French pronouns: different uses of *on* with special reference to the *on / nous* alternation. In: Kirsten Jeppesen Kragh & Jan Lindschouw (eds.), *Deixis and pronouns in Romance languages*. 125–144. Amsterdam: Benjamins.
- Harris, Martin. 1988. French. In: Martin Harris & Nigel Vincent (eds.), *The Romance languages*. 209–245. New York: Oxford University Press.
- Haugen, Einar. 1990. Danish, Norwegian, Swedish. In: Bernard Comrie (ed.), *The world's major languages*. 157–179. Oxford: Oxford University Press.
- Hengeveld, Kees. 2012. Referential markers and agreement markers in Functional Discourse Grammar. *Language Sciences* 34: 468–479.
- Hengeveld, Kees. 2017. A hierarchical approach to grammaticalization. In: Kees Hengeveld, Heiko Narrog & Hella Olbertz (eds.), *The grammaticalization of tense, aspect, modality and evidentiality: a functional perspective*. 13–38. Berlin: De Gruyter Mouton.
- Hengeveld, Kees & Sterre Leufkens. 2018. Transparent and non-transparent languages. *Folia Linguistica* 52/1: 139–175.
- Hengeveld, Kees & J. Lachlan Mackenzie. 2008. *Functional Discourse Grammar: a typologically-oriented theory of language structure*. Oxford: Oxford University Press.
- Hengeveld, Kees & J. Lachlan Mackenzie. 2012. Gramática discursivo-funcional (trad. do inglês por Marize Mattos dall’Aglío Hattner). In: Edson Rosa Francisco de Souza (ed.), *Funcionalismo linguístico: novas tendências teóricas*. 43–85. São Paulo: Contexto.
- Hengeveld, Kees & J. Lachlan Mackenzie. 2014. Grammar and context in Functional Discourse Grammar. *Pragmatics* 24/2: 203–227.
- Kaiser, Georg A. 2009. Losing the null subject: a contrastive study of (Brazilian) Portuguese and (Medieval) French. In: Georg A. Kaiser & Eva Maria Remberger (eds.), *Proceedings of the Workshop “Null-subjects, expletives, and locatives in Romance”*. (Arbeitspapier 123). 131–156. FB Sprachwissenschaft, Universität Konstanz.
- Keizer, Evelien. 2005. *A Functional Discourse Grammar for English*. Oxford: Oxford University Press.
- Keizer, Evelien; Olbertz, Hella. 2018. Functional Discourse Grammar: a brief outline. In: Evelien Keizer & Hella Olbertz (eds.), *Recent development in Functional Discourse Grammar*. 1–15. Amsterdam: Benjamins

- Leufkens, Sterre C. 2015. *Transparency in language: a typological approach*. Utrecht: LOT Publications.
- Marins, Juliana E. & Humberto Soares da Silva. 2012. A representação do sujeito pronominal no grupo românico: espanhol e italiano em contraste com o português. *Caligrama* 17/2: 91–114.
- Martelotta, Mário Eduardo T. & Maria Maura Cezario. 2011. Grammaticalization in Brazilian Portuguese. In: Bernd Heine & Heiko Narrog (eds.), *The Oxford handbook of grammaticalization*. 729–739. Oxford: Oxford University Press.
- Naro, Anthony Julius & Maria Marta Pereira Scherre. 2007. *Origens do português brasileiro*. São Paulo: Parábola.
- Oliveira, Taísa Peres de. 2018. Referential subject expression in Brazilian Portuguese. In: Evelien Keizer & Hella Olbertz (eds.), *Recent developments in Functional Discourse Grammar*. 207–232. Amsterdam: Benjamins.
- Otheguy, Ricardo. 2014. Remarks on pronominal perseveration and functional explanation. In: Andrés Enrique Arias, Manuel J. Gutiérrez, Alazne Landa & Francisco Ocampo (eds.), *Perspectives in the study of Spanish language variation: papers in honor of Carmen Silva-Corvalán*. 373–396. Santiago de Compostela: Universidade de Santiago de Compostela (*Verba*, Anexo 72).
- Posio, Pekka. 2013. The expression of first-person-singular subjects in spoken Peninsular Spanish and European Portuguese: semantic roles and formulaic sequences. *Folia Linguistica* 47/1: 253–291.
- Real Academia = Real Academia Española y Asociación de Academias de la Lengua Española. 2009. *Nueva gramática de la lengua española*, 3 vols. Madrid: Espasa.
- Reesink, Ger. 1987. *Structures and their functions in Usan, a Papuan language of Papua New Guinea*. Amsterdam: Benjamins.
- Searle, John R. 1969. *Speech acts*. Cambridge: Cambridge University Press.
- Siewierska, Anna. 2013. Verbal person marking. In: Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Disponível em wals.info/chapter/102; acesso em 15/09/2017.
- Zilles, Ana Maria S. 2005. The development of a new pronoun: the linguistic and social embedding of *a gente* in Brazilian Portuguese. *Language variation and change* 17: 19–53.