

## ELEMENTOS PARA A INVESTIGAÇÃO DA SEMÂNTICA DO CLÍTICO SE NO PORTUGUÊS BRASILEIRO

ESMERALDA VAILATI NEGRÃO  
Universidade de São Paulo/CNPq  
EVANI VIOTTI  
Universidade de São Paulo

**RESUMO:** Este artigo tem como objetivo trazer uma contribuição para as discussões relativas à eventual tendência do português brasileiro à perda do clítico se em alguns contextos, e ao aumento de seu uso, em outros. De maneira geral, os estudos sobre o assunto concentram sua atenção nos aspectos morfossintáticos relacionados à presença ou à ausência do clítico. Nossa intenção é mostrar que existem também fatores semânticos envolvidos no uso versus não-uso do se que dizem respeito à conceitualização dos eventos que estão sendo codificados linguisticamente. A presença do clítico parece sempre envolver a conceitualização de uma causa indutora do evento, por mais indeterminada que seja. Sendo assim, o aumento da frequência de não-uso do clítico em contextos em que ele seria esperado implica uma preferência semântica do português brasileiro por conceitualizar certos eventos sem sua causa indutora. Por outro lado, o aumento do uso do clítico junto a verbos no infinitivo, contexto em que a presença de se não seria esperada, sugere que a codificação da força indutora do evento serve para acentuar as características propriamente verbais do infinitivo que, de outra maneira, estaria mais próximo de uma forma nominal.

**Palavras-chave:** clítico se; conceitualização de eventos; infinitivo.

**ABSTRACT:** This article aims at contributing to the discussions about the possible tendency of Brazilian Portuguese toward the loss of the clitic se in some contexts, and toward the increase of its use, in other contexts. Studies on this subject matter usually focus on the morphosyntactic aspects related to the presence or absence of the clitic. Our intention is to show that there are also semantic factors involved in the use versus non-use of se, which are related to the conceptualization of the events being linguistically codified. The presence of the clitic seems to involve the conceptualization of the event's cause, however indeterminate it may be. The increase in frequency of non-use of the clitic in contexts in which its presence would be expected implies a semantic preference for the non-conceptualization of the initial cause of certain events. On the other hand, the increase of the cliticization of se with infinitival verbs – a context in which the presence of se is not expected – suggests that the codification of the initial cause of the event denoted by the verb serves the purpose of stressing the verbal characteristics of the infinitival, which would otherwise be closer to a nominal form.

**Keywords:** clitic se; conceptualization of events; infinitivals.

## INTRODUÇÃO

Uma das questões fundamentais para uma teoria gramatical é explicitar os princípios que explicam a variação entre as línguas no modo como representam,

na sintaxe, a estrutura de argumentos relacionados aos verbos, ou seja, o modo como as línguas codificam os eventos descritos por suas sentenças. Mecanismos morfológicos, em geral, marcam essas diferenças de codificação. O fenômeno descrito na teoria gramatical como *voz verbal* congrega um conjunto de estratégias para, até mesmo dentro de uma mesma língua, marcar os diferentes modos de codificação dos participantes do evento na estrutura sentencial.

Estudos diacrônicos que analisam os processos históricos envolvidos nas mudanças morfológicas ocorridas no desenvolvimento das línguas românicas mostram a extensão do uso do pronome reflexivo *se* do Latim para construções com a semântica própria ao domínio médio (Kemmer, 1993). Este artigo tem como objetivo analisar o papel semântico desempenhado pelo pronome *se* no português brasileiro.

## BREVE REVISÃO DA VASTA LITERATURA

Tomando como base a hipótese inacusativa de Perlmutter (1978), segundo a qual a classe dos verbos monoargumentais não é homogênea, Burzio (1986) a sustenta trazendo para a literatura linguística evidências sólidas. Os verbos *arrivare* ‘chegar’ e *telefonare* ‘telefonar’ em italiano precisam apresentar diferenças estruturais importantes para explicar as assimetrias na distribuição do clítico *ne* e no uso de verbos auxiliares observadas nos pares (1) e (2) e (3) e (4) a seguir:

1. *Ne arrivano molti*
2. \**Ne telefonano molti*
  
3. *Giovanni è arrivato*
4. *Giovanni ha telefonato*

A assimetria de gramaticalidade na distribuição do clítico *ne* é elegantemente explicada pela generalização que estabelece que *ne-cliticization* (*ne-cl*) só é possível quando o argumento posposto ao verbo está relacionado a um objeto direto. E os dados em (5) e (6) exemplificam essa observação:

5. \**Ne esamineranno il caso molti*
6. *Giovanni ne inviterà molti*

A agramaticalidade de (5) se deve ao fato de que *ne* retoma parte de um argumento que é na verdade um sujeito, ao passo que, em (6), o clítico retoma parte de um argumento objeto. Essa elegante generalização só é possível se assumirmos que em (1) o único argumento de *arrivare* é um objeto. Portanto, apesar da aparente similaridade estrutural entre as sentenças monoargumentais, o comportamento com relação à cliticização por *ne* fala em favor da hipótese de que essa classe de verbos, na realidade, constitui-se em duas classes: a dos verbos intransitivos

e a dos verbos ergativos.<sup>1</sup> Ergativos são verbos monoargumentais cujo único argumento é um objeto direto. A divisão da classe dos verbos monoargumentais é algo observado em diferentes línguas, embora os argumentos empíricos que dão sustentação para essa divisão variem.<sup>2</sup>

Na sua análise da cliticização por *ne*, Burzio observa uma uniformidade de seu comportamento em certos domínios, além das construções transitivas: passivas (como em (7)), uma variante de construções impessoais – a construção impessoal com *se* (como em (8)) – e os verbos da alternância transitivo/ergativo (como em (9)).

7. Ne saranno invitati molti
8. Se ne leggerà alcuni
9. Ne affondarono due

Os três domínios têm, em comum, propriedades que contribuíram para a formulação do que ficou conhecido como a generalização de Burzio: nesses domínios não há atribuição de papel temático ao argumento externo e não há atribuição de Caso acusativo ao objeto direto.

Essa generalização possibilitou que Burzio buscasse uma análise unificada para algumas construções com o pronome *se*, especialmente o que ele classificou como o *se* reflexivo (como em (10)),<sup>3</sup> o *se* ergativo (como em (11)) e o *se* inerente-reflexivo (como em (12)), uma vez que as três construções têm em comum a seleção do auxiliar *essere*:

10. Maria si guarda
11. Il vetro si rompe
12. Giovanni si sbaglia

Para ele, essas construções diferem radicalmente do que ele chamou de *se* impessoal, um clítico relacionado à posição de sujeito (exemplos (13) a (16) abaixo), e não à de objeto, como nos casos entre (10) e (12) acima:

13. Gli si telefona spesso
14. Si leggerà volentieri alcuni articoli
15. Alcuni articoli si leggeranno volentieri
16. Si leggeranno volentieri alcuni articoli

---

<sup>1</sup> A divisão entre verbos monoargumentais ergativos e intransitivos é mais conhecida na literatura gerativa como verbos inacusativos e verbos inergativos, respectivamente.

<sup>2</sup> Ver Perlmutter (1978), principalmente os dados de passivas impessoais.

<sup>3</sup> Essa proposta ficou conhecida como a análise inacusativa dos reflexivos, uma vez que segundo Burzio o reflexivo *se* é um afixo que absorve o papel temático do argumento externo. Essa análise é criticada por Reinhart e Siloni (2004), que argumentam que as construções com *se* reflexivo são na verdade inergativas.

Dobrovie-Sorin (2005) aponta que somente certas línguas românicas pro-drop, particularmente o italiano, o espanhol e o português apresentam o que ela classifica como o *se* nominativo/sujeito, o *se* impessoal de Burzio. A autora argumenta que apesar das tentativas feitas na literatura, o *se* nominativo e o *se* acusativo (inerente, incoativo, médio, passivo e reflexivo-recíproco) devem ser tratados como homônimos e não como realizações de um mesmo elemento. Apoiando-se em Naro (1976), que alega que construções do tipo *se* V NP, envolvendo o *se* médio passivo, foram reanalisadas como sequências SVO em que o *se* é um pronome sujeito, a autora propõe que a existência ou ausência do *se* nominativo seja um parâmetro regulando a variação entre as línguas românicas: italiano, espanhol e português têm ambos, o *se* acusativo e o *se* nominativo, enquanto romeno e francês têm só o *se* acusativo.

## O PRONOME *SE* ERGATIVO

O pronome *se* que Burzio chama *ergativo* é discutido com frequência na literatura tipológica que se interessa pelos fenômenos de natureza morfológica envolvidos na alternância causativa. O clítico *se* é entendido, de maneira geral, como uma marca de derivação morfológica associada à diminuição de diátese de um verbo causativo, a partir da qual ele se torna um verbo ergativo.<sup>4</sup> Esse fenômeno se verifica consistentemente em várias línguas românicas. Para além dos exemplos do italiano já apresentados (ver (11) acima), em (17) e (18) temos exemplos do espanhol e, em (19) e (20), do francês:

17. Juan rompió la copa.

18. La copa se rompió.

19. Jean a cassé le verre.

20. Le verre s'est cassé.

Marcas morfológicas que codificam a alternância causativa se encontram em diversas línguas, das mais variadas famílias. A questão que a literatura morfológica coloca diz respeito à direção da derivação da alternância: busca-se estabelecer qual é a forma básica e qual é a forma derivada. Ou seja, a questão é estabelecer se é a forma ergativa que deriva da forma transitiva, ou se é a forma transitiva que deriva da ergativa.

Segundo Haspelmath (1993), há uma grande variação translinguística relativamente à direcionalidade da alternância causativa. O primeiro passo para estabelecer essa direcionalidade é determinar qual é a forma do verbo que recebe

---

<sup>4</sup> Vamos manter aqui a terminologia de Burzio para caracterizar esse tipo de verbo. Outros autores preferem chamá-los incoativos. Para uma discussão sobre a terminologia relevante, ver Chagas de Souza (2000).

as marcas morfológicas. Essa será considerada a forma derivada. Russo, por exemplo, é uma língua em que a forma ergativa deriva da causativa, na medida em que é a forma ergativa que recebe o sufixo *-sja* para marcar a diminuição da diátese. Observem-se as formas em (21) e (22):

21. *rasplavit* ‘derreter’ (causativo)  
 22. *rasplavitsja* ‘derreter’ (ergativo)

Diferentemente, em mongol khalkha, a marca morfológica – o infixo *-uul-*, aparece na versão causativa; a ideia, então, é que a forma causativa deriva da ergativa:

23. *xajuulax* ‘derreter’ (causativo)  
 24. *xajlax* ‘derreter’ (ergativo)

Alternâncias como as do exemplo do russo em (21) e (22) são, então, chamadas *anticausativas*. A anticausatividade pode ser marcada quer por um afixo, como no caso do russo, quer por um auxiliar anticausativo, quer por uma modificação na raiz do verbo.

Alternâncias como as exemplificadas com o mongol khalkha em (23) e (24) acima são chamadas *causativas*. Verbos causativos são obtidos a partir de uma base intransitiva pela adição de um afixo transitivizador, de um auxiliar causativo, ou pela mudança da raiz.

Há, no entanto, línguas que apresentam alternâncias em que nem a forma não-causativa, nem a causativa é derivada da outra. Em japonês, por exemplo, há um subtipo de alternância não-direcional chamado *equipolente*: a derivação da forma causativa e da forma não-causativa ocorre pela adição de sufixos diferentes a uma mesma raiz verbal:

25. *atum-eru* ‘juntar’ (causativa)  
 26. *atum-aru* ‘juntar’ (não-causativa)

As línguas podem também apresentar alternâncias não-direcionais do tipo *supletivo*. Trata-se de alternâncias em que verbos com raízes diferentes são usados para expressar a transitividade ou intransitividade de uma eventualidade. O russo tem alternâncias desse tipo: *goret*’ denota o evento de ‘queimar’ sem que uma causa ou um agente seja especificado, e *zef*’ denota o mesmo evento, agora elaborando a causa que deu origem ao evento.

Um exemplo de alternância do tipo supletivo que se encontra em várias línguas é a dupla *matar* e *morrer* do português. O evento designado pelos dois verbos é substancialmente o mesmo, mas *matar* exige a codificação de uma causa, enquanto *morrer*, não.

O último subtipo de alternâncias cuja direcionalidade não pode ser estabelecida é chamada *lábil*. Trata-se de alternâncias em que a mesma forma

do verbo pode ser usada em uma sentença transitiva e em uma sentença ergativa. Inglês é uma língua cujas alternâncias são majoritariamente desse tipo:

27. John broke the glass.

28. The glass broke.

Haspelmath admite que a classificação dos verbos das línguas que ele estudou nessas cinco categorias – anticausativas, causativas, não-direcionais equipolentes, não-direcionais supletivas e não-direcionais lábeis – não é sem problemas. Para o autor, a direcionalidade é mais facilmente estabelecida quando um dos verbos do par contém algum segmento (um afixo, uma partícula, etc.), que não esteja presente no outro membro do par, como nos exemplos (22) e (23) acima. Entretanto, quando a relação de alternância é marcada por uma mudança na raiz, a situação se complica, como no seguinte par o hindi-urdu:

29. khul-naa ‘abrir’ (intransitivo)

30. khol-naa ‘abrir’ (transitivo)

Outra dificuldade está na classificação de pares de verbo que diferem entre a forma causativa e a não-causativa apenas pela escolha do auxiliar. Em alemão, por exemplo, o par em (31) e (32) se distingue apenas pela escolha do *sind* ‘ser’, no primeiro, e do auxiliar *haben* ‘ter’, no segundo:

31. Der Krug ist zerbrochen.

‘O jarro quebrou’.

32. Wer hat den Krug zerbrochen?

‘Quem quebrou o jarro?’

De um ponto de vista semântico, a caracterização da eventualidade a que verbos de alternância causativa podem se referir tem sido desafiadora. De maneira geral, os autores concordam que a situação básica denotada por esses verbos envolve uma mudança de estado do participante não-agentivo da eventualidade. Desse modo, poderiam participar de alternância causativa verbos como *quebrar*, *queimar*, *abrir*, entre outros, mas não verbos como *construir*, *cortar*, entre muitos outros.

Para além da mudança de estado de um dos participantes da eventualidade, Haspelmath propõe uma outra característica semântica dos verbos que aceitam a alternância causativa: a ausência de componentes de significação que sejam orientados para o agente (p. 93). O que Haspelmath quer dizer com isso é que, para que a alternância seja viável, a interpretação do membro ergativo do par de verbos alternantes terá que ser possível. Como esse membro ergativo necessariamente implica a ausência de um agente, elementos semânticos que sejam orientados pelo agente não poderão entrar no jogo. A ilustração dessa posição vem do contraste dos verbos *tear* (rasgar) e *cut* (cortar) do inglês. Enquanto o primeiro permite

a alternância (exemplos (33) e (34) abaixo), o segundo, que tem componentes semânticos orientados para o agente, não a permite, como mostram (35) e (36):

33. The girl tore her pants.

34. The pants tore.

35. The tailor cut the cloth.

36. \*The cloth cut.

A generalização semântica de Haspelmath, então, é colocada da seguinte maneira:

37. *A verb meaning that refers to a change of state or a going-on may appear in an inchoative/causative alternation unless the verb contains agent-oriented meaning components or other highly specific meaning components that make the spontaneous occurrence of the event extremely unlikely (p.94).*

Apesar de essa caracterização semântica poder sugerir que existam classes universais de verbos de alternância causativa, Haspelmath aponta que as línguas variam a esse respeito, podendo apresentar comportamentos diferentes do esperado. Ele sugere que duas interpretações são possíveis para esses casos: (i) a caracterização em (37) não seria universal, variando de língua para língua; ou (ii) os significados dos verbos variam de língua para língua. Como se sabe que, de fato, o significado de verbos correlatos de diversas línguas não é exatamente o mesmo, Haspelmath prefere atribuir a essa diferença a variação nas possibilidades de alternância causativa observadas translinguisticamente, preservando, assim, a generalização semântica em (37).

## **ALGUNS COMPORTAMENTOS DO CLÍTICO *SE* NO PORTUGUÊS BRASILEIRO**

Como é que o português brasileiro se encaixaria nesse panorama tipológico de base morfológica e semântica? Sabidamente, o português brasileiro tem um grupo de verbos que alternam entre uma leitura causativa e uma leitura ergativa (ou incoativa, como prefere Haspelmath), de maneira a fazer valer a generalização semântica em (37) acima. Aqui vão alguns exemplos já clássicos:

38. João quebrou o vidro.

39. O vidro (se) quebrou.

40. Ana abriu a porta.

41. A porta (se) abriu.

42. A esquadra inimiga afundou o navio.

43. O navio (se) afundou.

Esses verbos designam uma mudança de estado, e, em sua forma ergativa, designam eventos que ocorreram espontaneamente, sem que qualquer agente esteja envolvido na mudança de estado. A alternância entre a forma biargumental e a monoargumental está semanticamente associada à não conceitualização da causa que teria dado origem ao evento, fazendo com que ele seja interpretado como um evento espontâneo.

Entretanto, como vimos mostrando em uma série de artigos,<sup>5</sup> o português brasileiro vem apresentando um outro tipo de alternância que também se caracteriza pela não conceitualização da causa iniciadora do evento. A diferença é que, nesse tipo de alternância, estão envolvidos verbos que não designam mudança de estado, e que, em princípio, pareceriam ter componentes semânticos que teriam orientação para o agente. Alguns dados observados em situações de fala espontânea são os seguintes:

44. Nas minhas férias, você não acredita, meu carro roubou, mas ainda bem que achou no dia seguinte.
45. O carro vai carregar aqui ou na garagem?
46. O terreno já capinou e já preparou pra plantar.
47. Eu fiquei com tanto medo que o sanduíche nem engoliu.
48. Meu jardim destruiu todo com a reforma.
49. Este prédio tá construindo desde que a casa do vovô vendeu.
50. Eu só vou trocar o carpete depois que a casa acabar de pintar.
51. Não adianta correr que esse trem já perdeu.
52. Não tem nenhum concurso que anulou por causa de mérito.
53. O Luiz Omar nunca torceu tanto pra um saque errar como nesse ponto do jogo.

De um ponto de vista morfológico, observamos uma assimetria entre construções como as entre (38) e (43) e aquelas entre (44) e (53), e essa assimetria está relacionada com a possibilidade ou não de o clítico *se* aparecer na construção não-causativa: no caso das primeiras, a alternância entre a forma causativa e a forma ergativa ainda pode ser marcada morfológicamente pela presença do clítico, enquanto no caso das últimas, a presença do clítico é impossível (Negrão e Viotti, 2010).

De acordo com a tipologia de Haspelmath, o português brasileiro seria uma língua que ora seria vista como pertencente a uma categoria, ora a outra. As alternâncias que apresentam o clítico *se* seriam evidência de que o português brasileiro é uma língua do tipo anticausativo, ou seja, uma língua que tem a forma causativa como básica, e a ergativa como derivada. No caso em que as sentenças não apresentam o clítico *se*, a ideia, então, seria a de que o português brasileiro é uma língua do tipo lábil, em que a forma causativa e a ergativa do verbo são idênticas. De um jeito ou de outro, do ponto de vista semântico, Haspelmath consideraria os pares de verbos entre (38) e (43), com ou sem o clítico *se*, como

<sup>5</sup> Ver Negrão e Viotti (2008, 2010, 2014).



pares de alternância causativa, porque trata-se de verbos de mudança de estado, que não têm componentes especificamente orientados para o agente.<sup>6</sup>

Entretanto, a presença ou ausência do clítico *se* não deve ser considerada como um fenômeno de natureza exclusivamente morfossintática. Ainda que de maneira sutil, a presença ou ausência do clítico traz consequências semânticas para a interpretação da eventualidade sendo descrita pela sentença. Por exemplo, com o modificador *sozinho*, a sentença com o clítico parece menos aceitável que sem ele:

54. O navio afundou sozinho.

55. \*?O navio se afundou sozinho.

Ainda, a presença de um modificador que especifique qualquer faceta relacionada à causa do evento gera estranheza na sentença sem o clítico:

56. \*O navio afundou como uma estratégia do capitão para driblar o inimigo.

57. ?O navio se afundou como uma estratégia do capitão para driblar o inimigo.

Em sentido metafórico, parece que a ausência do clítico não é possível:

58. A crise levou muitas pessoas a se afundar em dívidas.

59. \*A crise levou muitas pessoas a afundar em dívidas.

Portanto, vemos que a presença ou ausência do clítico tem um impacto semântico – ainda que tênue – na interpretação da sentença. Isso significa que a classificação categórica de verbos e de línguas proposta por Haspelmath, separando anticausativas e lábeis apenas por critérios morfológicos, exige cautela. A propósito do uso de verbos lábeis do inglês, em sua versão intransitiva, como em (28), Langacker (1991: 335; 389) observa que há casos em que não se trata simplesmente da ausência de conceitualização da causa do evento, mas sim de uma maior ou menor saliência da causa.

60. The door opened very easily.

61. The door suddenly opened.

Ao contrastar as sentenças (60) e (61), Langacker observa que a presença de modificadores como *very easily*, em (60), e *suddenly*, em (61), atribuem a uma e a outra diferentes graus de saliência da causa do evento: em (60), há um agente implícito, para quem foi fácil abrir a porta; em (61) nada indica a presença de um agente.

Assumir uma única semântica para as alternâncias causativas e classificar línguas e verbos em categorias de natureza estritamente morfossintática, como

---

<sup>6</sup> Cf. a generalização em (37) acima.

quer Haspelmath, torna também difícil a caracterização do clítico *se* do espanhol. Segundo Maldonado (2006), esse clítico pode gerar duas semânticas diferentes. Quando ele se cliticiza a um verbo prototipicamente transitivo, o agente (ou causa) do evento denotado pelo verbo perde saliência. Observem-se os exemplos retomados a seguir:

62. Juan rompió la copa.

63. La copa se rompió.

Em (62), o verbo *romper* é um verbo transitivo que denota um evento em que um agente – *Juan* – causa a mudança de estado de um outro participante – *la copa*. Em (63), o participante afetado ganha saliência, e o participante agente perde saliência, embora continue a ser conceitualizado. Nos termos da Haspelmath, esse tipo de alternância sugeriria que o espanhol é uma língua do tipo anticausativa, na medida em que a versão ergativa do verbo parece derivar da versão causativa.<sup>7</sup>

Diferentemente, quando o *se* se cliticiza a um verbo prototipicamente intransitivo, como *caer*, ele faz com que a ação denotada pelo verbo seja conceitualizada como diferente da ação prototípica. Ao comparar as sentenças (64) e (65) abaixo, Maldonado observa que é natural que as folhas caiam no outono, mas não na primavera. Por isso, em (65), é necessário o uso do clítico *se*, assinalando um afastamento do padrão natural.

64. En el otoño, las hojas (\*se) caen de los árboles.

65. En la primavera, las hojas se (\*Ø) cayeron de los árboles.

Maldonado considera (65) uma *construção energética*, afirmando que alguma força não especificada, como uma tempestade ou uma praga, deve necessariamente fazer parte de sua conceitualização. Como no outono *se* espera que as folhas caiam, o uso do clítico *se*, em (64), não é possível. Dentro do quadro tipológico proposto por Haspelmath, esse tipo de construção do espanhol seria evidência de que essa língua seria do tipo causativo, ou seja, ela partiria de um verbo de semântica ergativa, para, a partir daí, derivar o verbo de semântica causativa.

Como Haspelmath admite que as línguas nem sempre são bem-comportadas, e que muitas vezes verbos diferentes da mesma língua apresentam direcionalidades de alternância diferentes, essa particularidade do espanhol não seria problemática. Entretanto, o modelo de Haspelmath não daria margem para uma discussão da diferença semântica causada pela presença da marca morfológica – o clítico *se* – junto a um verbo prototipicamente transitivo ou a um verbo prototipicamente intransitivo.

---

<sup>7</sup> De qualquer maneira, uma diferença semântica entre a visão de Maldonado e a de Haspelmath permaneceria: para Maldonado, a sentença (63) continua a ter um agente conceitualizado, embora ele não seja saliente; para Haspelmath, o verbo ergativo em (63) não implica a conceitualização de um agente.

É justamente essa a diferença que estamos sugerindo que existe nas alternâncias do português brasileiro. A presença do clítico *se* indica que o evento ao qual o verbo se refere está sendo conceitualizado como um evento energético, ou seja, como um evento causado por alguma fonte de energia, ainda que essa fonte seja indeterminada, arbitrária, ou vaga. A ausência do clítico, por outro lado, significa que o que está sendo conceitualizado é apenas o resultado de um evento, não sua origem ou causa.

A observação dessa diferença semântica entre a presença e a ausência do clítico *se* em pares de sentenças como aqueles entre (38) e (43), e a investigação sobre sentenças como as que se encontram entre (44) e (53) acima nos levou à conclusão de que o que tem sido chamado construção ergativa (que é a versão intransitiva de verbos como *abrir, fechar, afundar, quebrar*, entre outros), e que pode variar entre a versão com ou sem o clítico *se*, são, na verdade, duas construções semanticamente distintas: a construção com *se* é uma construção energética (para usar o termo de Maldonado), que codifica a conceitualização de uma força indutora da ação, ainda que essa força não seja especificada; e a construção sem o clítico é a codificação de uma conceitualização que não envolve a presença de uma força indutora. Seguindo Langacker (1991), estamos chamando esse segundo tipo de sentença de *construção absoluta*.

Uma consequência dessa visão a respeito desse fenômeno do português brasileiro é a necessidade de ampliar as explicações sobre a preferência pelo não uso do clítico *se* com verbos como *abrir, fechar, afundar, quebrar*, entre outros. De modo geral, essas explicações se prendem a fatores de natureza morfossintática. O que estamos dizendo é que, de um ponto de vista semântico, o não uso do *se* revela uma preferência por uma conceitualização dos eventos denotados por esses verbos como eventos espontâneos. A mudança morfossintática relacionada à diminuição do uso do clítico está atrelada a essa mudança semântica.

Paralelamente à busca por uma maior abrangência para as razões que têm levado falantes do português brasileiro a preferir não usar o clítico *se* em determinados contextos, deve-se também alargar o escopo de busca dos motivos pelos quais o uso do clítico *se* é maior do que o esperado em outros contextos, nomeadamente junto a verbos na forma infinitiva. Cavalcante (2006; 2010) nos mostra que há uma grande diferença entre a porcentagem de uso de *se* com infinitivo em português europeu e português brasileiro. Seus achados são os seguintes:

- Para o português europeu:
  - Tanto para as amostras de fala quanto para as de escrita: 8% de *se* junto a infinitivos
- Para o português brasileiro:
  - Para as amostras de fala: 20% de presença de *se* junto a infinitivos
  - Para as amostras de escrita: 50% de uso do clítico junto a infinitivos

A explicação dada para essa assimetria está na proposta de que há uma diferença entre o núcleo Agr de uma língua e de outra. O português europeu teria um Agr forte (caracterizado por um conjunto completo de traços  $\phi$ ), capaz de

licenciar a posição de sujeito e de identificar categorias vazias que a ocupem. Essa seria, então, uma típica língua de sujeito nulo, que tenderia a preencher a posição de sujeito com DPs foneticamente realizados apenas quando pragmaticamente necessário. Para a baixa porcentagem de *se* junto a infinitivos, essa explicação seria válida, na medida em que, segundo Cavalcante, o clítico aparece apenas em contextos em que é necessário evitar que PRO, sujeito do infinitivo, seja controlado por algum outro constituinte, de modo a poder receber a interpretação arbitrária desejada. A presença do clítico *se* tem, portanto, uma função desambiguadora.

Diferentemente, o português brasileiro teria um Agr fraco (caracterizado por apresentar traço de [pessoa] defectivo), capaz de licenciar a posição de sujeito, mas não de identificar categorias vazias que lá se encontrem. Essa seria a explicação para a tendência geral do português brasileiro de preencher a posição de sujeito. No caso de sentenças com verbos no infinitivo, a explicação é que, como o sujeito do infinitivo não pode ter sua referência identificada por Agr fraco, ele tende a ir buscar sua referência em algum antecedente. Para evitar isso e conseguir a interpretação arbitrária, o clítico *se* é usado, associando-se a uma categoria vazia na posição de sujeito.<sup>8</sup>

A proposta de Cavalcante para a cliticização do *se* a verbos no infinitivo no português brasileiro está em consonância com uma série de trabalhos que investigam a posição de sujeito e suas relações com a concordância verbal nessa língua.<sup>9</sup> Entretanto, a ampliação do escopo do estudo do *se* aqui proposta, que pretende incluir a semântica que está envolvida em sua ausência ou presença em construções do português brasileiro, nos leva a considerar algumas questões que, até onde tenhamos percebido, não foram ainda exploradas.

Uma delas diz respeito ao infinitivo flexionado do português. Para o caso do português europeu, Cavalcante assume que há dois tipos de infinitivo: o flexionado, associado a Agr, e o não flexionado, que se comporta como o infinitivo de outras línguas românicas.<sup>10</sup> No caso do português brasileiro, a autora propõe que Agr não finito compartilha, com o Agr finito, o traço [pessoa] defectivo. A partir dessa proposta, ela atrela o preenchimento da posição de sujeito do infinito ao elevado índice de preenchimento da posição sujeito observado nas sentenças finitas (Duarte 1995).

Essa posição dá margem ao entendimento de que o português brasileiro, diferentemente do português europeu, não teria uma flexão de infinitivo robusta,

<sup>8</sup> Na verdade, Cavalcante afirma que o clítico *se*, ele mesmo, preenche a posição de sujeito do infinitivo, podendo variar com *pro*, *a gente* ou *você* (Cavalcante 2010: 65; 70-71). Como, no entanto, essa proposta se baseia no trabalho de Cinque (1988), preferimos manter a ideia de que *se* se cliticiza ao verbo e forma uma cadeia com Agr e um *pro* na posição de sujeito. No que diz respeito ao português europeu, Cavalcante assume que, nessa língua, a posição de sujeito do infinitivo pode ter ou PRO ou o clítico *se*, em casos especiais (p. 65).

<sup>9</sup> Ver, em especial, Galves (2001).

<sup>10</sup> Cavalcante (2006, 2010) propõe que, no português europeu, o clítico *se* está associado ao infinitivo flexionado, e PRO está associado ao infinitivo não flexionado. Para um tratamento do infinitivo em outras línguas românicas e das diferenças que o separam do infinitivo do português, ver Cinque (1988).

na medida em que seu Agr é defectivo para o traço [pessoa]. Entretanto, esse não parece ser o caso, pelo menos para o infinitivo. Em um estudo de língua em uso realizado a partir de um corpus de português brasileiro escrito de mais de 11 milhões de palavras, Canever (2012) observa que há uma alta frequência de infinitivos flexionados em contextos em que a flexão é considerada opcional: 75% para orações-adjunto; 94,5% para orações-complemento de nome; e 89% para orações-complemento de adjetivo. Em contextos em que a flexão é considerada agramatical, como em complementos de verbos aspectuais ou modais, a flexão do infinitivo aparece em uma frequência de 5% e 1%, respectivamente.<sup>11</sup> Alguns exemplos encontrados por Canever são:

• Orações-adjunto:

66. ...embora alguém pretenda que, ao pensarmos as palavras, falamos interiormente...
67. ...os sujeitos, ao interagirem por meio da linguagem, vão...
68. Citamos esses dois livros didáticos por pensarmos...
69. Tarefa que não podemos recusar, especialmente para entendermos...

• Orações completivas de nome e adjetivo:

70. ...categorias referidas eram muito vagas, difíceis de serem controladas...
71. ...sua fonte abriga passagens com temas diversos capazes de participarem de projetos...
72. ...na tentativa de fazerem prevalecer seus pontos de vista.
73. ...contam, sob o pretexto de trazerem um presente, com o qual fingia que o rei pudesse ser...

• Perífrases aspectuais e modais:

74. As virtudes começam a serem tratadas no capítulo 8.
75. ...em antigos manuscritos hebraicos e que eram tão importantes que não poderiam ser esquecidas.

Outra questão que deve ser colocada é a seguinte: se o português brasileiro vem mostrando uma tendência à simplificação de seu paradigma flexional, por que, no caso do infinitivo, a tendência parece justamente ir no sentido oposto? Talvez porque, apesar de o paradigma flexional do português brasileiro ser inquestionavelmente mais simples do que o do português europeu, seu sistema

---

<sup>11</sup> O corpus a partir do qual o estudo foi feito é formado por 180 teses e dissertações de vários programas de pós-graduação em Letras da Universidade de São Paulo. Trata-se, então, de um corpo de dados da norma urbana culta, o que, por um lado, pode explicar o alto índice de flexão. Por outro lado, o uso excessivo da flexão do infinitivo em contextos opcionais está em desacordo com o que prescreve a gramática normativa. Portanto, a expectativa seria, justamente, a de que em textos acadêmicos, escritos e supostamente revisados por profissionais de Letras, a flexão do infinitivo não fosse tão alta como é.

de concordância possa ser ativado, justamente para identificar clara e diretamente a fonte iniciadora do evento denotado pelo verbo, mesmo que seja uma fonte arbitrária.

Essa sugestão de explicação para a alta frequência da flexão do infinitivo em português brasileiro parece se assemelhar à que Cavalcante elabora para justificar a frequência de uso do clítico *se* com infinitivos. Como visto, a autora propõe que o uso do clítico é uma estratégia para evitar que o sujeito do infinitivo seja controlado por algum antecedente e possa, então, receber uma interpretação arbitrária. Em ambos os casos, as marcas morfossintáticas – flexão de número e pessoa, em um caso; o clítico *se*, no outro – parece que surgem justamente para prover um conteúdo semântico mais claro e menos ambíguo relativamente ao referente que corresponde ao iniciador do evento denotado pelo verbo.

Se essa sugestão estiver no caminho certo, ela corrobora a explicação que elaboramos em Negrão & Viotti (2010), e que retomamos acima para discutir o trabalho de Haspelmath: a presença ou ausência do clítico *se* nas sentenças do português brasileiro tem um impacto semântico. O uso do clítico é uma tática do português brasileiro relacionada à caracterização semântica dos argumentos do verbo. Quando ele diz respeito ao participante do evento denotado pelo verbo que corresponde a seu iniciador, ele dá a esse participante uma interpretação pouco específica, que pode ser indefinida (como nos casos em (39), (41) e (43) acima), ou arbitrária (como nos casos em que ele aparece junto a infinitivos).

Voltando aos comentários que fizemos sobre Haspelmath (1993), a proposta semântica que estamos esboçando vai contra a ideia de que a direcionalidade que relaciona alternâncias, como a ergativa, seja um fator definidor de uma tipologia linguística, ou de uma tipologia de verbos dentro de uma mesma língua. Como visto, em construções ergativas do português brasileiro, a presença do *se* não indica a supressão do participante relativo à causa do evento. Ela apenas torna essa causa inespecífica. A causa só é de fato suprimida em construções absolutas, em que o clítico não aparece.

Ampliando a perspectiva de modo a poder incluir a presença do clítico *se* junto a infinitivos, frequente em português brasileiro, escassa em português europeu, e impossível em vários contextos de outras línguas românicas, reforçamos a ideia de que inserir o *se* traz uma contribuição semântica que não pode ser ignorada. Uma evidência clara dessa contribuição vem de sentenças do tipo *tough-movement*, como (76) abaixo, retirada do Corpus NILC-São Carlos:

76. *par=Ilustrada--94a-2*: Para ela, todo relacionamento é **difícil de se terminar**.

Sem o clítico, essa sentença seria ambígua: em uma leitura, *todo relacionamento* seria o sujeito de *terminar*, como em uma construção absoluta (*Todo relacionamento termina com alguma dificuldade*); em outra, *todo relacionamento* seria complemento de *terminar*, como em uma sentença transitiva (*Ela termina todo relacionamento com alguma dificuldade*). Com o clítico, como acima, a sentença só pode ter uma leitura: o *se* está associado ao sujeito

de *terminar*, dando a ele uma interpretação arbitrária, e *todo relacionamento* é o complemento do verbo.

Entretanto, há contextos em que, com ou sem o *se*, a interpretação é arbitrária, mas, mesmo assim, o clítico é usado. Observem os seguintes dados do Corpus NILC-São Carlos:

77. *par=14590*: Nos poucos momentos em que as câmeras de TV chegaram mais perto de Jackson, **deu para se notar** uma espécie de buraco no lado esquerdo do seu nariz.
78. *par=Cotidiano--94b-2*: O diretor do Hemoacre disse que a situação da coleta de sangue no Estado ‘tá tão caótica que **«é necessário se repensar** os conceitos de transfusão»
79. *par=Esporte--94b-2*: «Por isso que no treino tático não **é necessário se colocar** o mesmo time que vai jogar a partida. “
80. *par=22351*: Lugares onde **é possível se passar** uma noite inteira sem ouvir sequer um acorde de música caipira e o clima está...
81. *par=44520*: No exemplo dado acima, **é possível se chegar** a uma conclusão válida simplesmente construindo um modelo mental.

Esses dados exemplificam contextos em que a única interpretação possível para o participante iniciador do evento denotado pelo verbo, mesmo que o se estivesse ausente, seria a interpretação arbitrária. A opção pelo uso do *se*, portanto, não pode ficar atrelada apenas à necessidade de evitar que ele seja interpretado como ligado por algum antecedente e garantir a interpretação arbitrária, como sugere Cavalcante (2006; 2010).

Os dados entre (77) e (81) reforçam nossa ideia de que o uso do clítico *se* traz informação semântica para o conteúdo da sentença, na medida em que traz para a superfície o participante iniciador do evento denotado pelo verbo.

É interessante notar que o estudo de direcionalidade de derivações morfossintáticas de Haspelmath (1993) é fundado na correlação entre marcas formais e relações semânticas, conhecida como *iconicidade diagramática*.<sup>12</sup> Se essa correlação for, de fato, pertinente, ela deve ser universal. O problema que Haspelmath vai tentar contornar é justamente o de que as marcas formais associadas à direcionalidade das relações de derivação das línguas variam de língua para língua, e mesmo dentro de uma mesma língua. A maneira encontrada pelo autor para solucionar essa questão foi a de recorrer à ideia de Givón (1991) de que a iconicidade na língua deve ser entendida em termos cognitivos. Para Givón, categorias que são cognitivamente marcadas tendem a ser também estruturalmente marcadas (p. 106). A proposta que estamos delineando aqui para o clítico *se* pode ser considerada tributária da iconicidade diagramática, mas no sentido inverso do da proposta de Givón: categorias que são estruturalmente marcadas tendem a ser cognitivamente marcadas. Ou seja, marcas formais existem para codificar algum

<sup>12</sup> Ver Haiman (1980), entre outros.

tipo particular de conceitualização. Se o português brasileiro, nas construções com verbo no infinitivo, tem tendido a apresentar marcas estruturais associadas ao participante que dá origem ao evento denotado pelo verbo, uma necessidade semântico-cognitiva deve estar em jogo. Nossa sugestão é a de que essa necessidade é a de codificar o iniciador do evento denotado pelo verbo, tornando a forma infinitiva menos nominal e mais verbal.

## À GUIA DE CONCLUSÃO

O clítico *se* das línguas românicas, apesar de já ter sido objeto de um número muito grande de análises em diferentes vertentes teóricas, exibe propriedades que ainda desafiam os pesquisadores envolvidos no estudo da gramática dessas línguas. Por trás de aparentes similaridades na distribuição do clítico *se* e em suas características sintáticas e semânticas, escondem-se particularidades que distinguem um uso de outro, uma vez que essa partícula está intrinsecamente ligada a propriedades gerais da gramática de cada uma dessas línguas. Daí sua resistência a análises unificadoras, tanto intra quanto interlinguísticas.

No caso particular da gramática do português brasileiro, as propriedades reguladoras de sua ocorrência e os contextos desencadeadores de sua realização ainda merecem novas investigações. Neste artigo analisamos a contribuição semântica desse clítico, argumentando que a sua ocorrência codifica a conceitualização da força indutora do evento denotado pelo verbo. Essa propriedade semântica está intimamente relacionada à expansão, na gramática do português brasileiro, das construções em que eventos são conceitualizados como ocorrendo espontaneamente, mesmo com verbos que trazem consigo uma semântica orientada para o agente. Essa proposta tem consequências diretas para o tratamento da alternância causativa. Ao clítico *se* do português brasileiro não pode ser atribuída a função de marcador morfológico da forma anticausativa da alternância, como se propõe para outras línguas românicas.

Com o intuito de dar um outro passo na investigação semântica do *se*, voltamos para a evidência de que o português brasileiro tem uma maior frequência de uso do clítico com o infinitivo do que o português europeu. Apesar de estarmos de acordo com Cavalcante quanto ao fato de que, em muitos exemplos, o clítico aparece para garantir uma interpretação arbitrária para o participante iniciador do evento denotado pelo verbo, em muitos outros casos, a interpretação seria arbitrária com ou sem o *se*. Portanto, a pergunta sobre as razões semânticas que levam à realização desse *se* permanece. Para respondê-la, nossa sugestão é a de que o *se* vem, como no caso das construções ergativas, codificar a conceitualização de uma fonte indutora do evento, dando ao infinitivo um caráter mais propriamente verbal do que nominal. Sugerimos também que essa razão está igualmente por trás da tendência do português brasileiro à flexão do infinitivo em contextos em que ela seria opcional.

Entretanto, é preciso observar que as observações feitas aqui são apenas a ponta de um enorme iceberg. A semântica do clítico *se* deve envolver aspectos



que vão muito além da codificação da fonte iniciadora do evento. Um exemplo de questão que nos parece pendente é a seguinte. Cinque (1988) aponta que, no italiano, em orações infinitas complemento de *sarebbe meglio*, independentemente de qual seja o tipo de verbo, a presença do *si* causa agramaticalidade:

82. \*Sarebbe meglio scoprirsi il colpevole.  
 Seria melhor descobrir o culpado  
 83. \*Sarebbe meglio lavorarsi un po' di più.  
 Seria melhor trabalhar um pouco mais.

Curiosamente, o português brasileiro parece se comportar exatamente como o italiano: não foi encontrado, em uma busca rápida no Corpus NILC-São Carlos, um único caso de presença do clítico *se* junto a infinitivos no mesmo contexto indicado por Cinque. Em casos como os de (84) a (86), uma explicação semântica que poderia ser aventada é a de que a interpretação do sujeito do infinitivo não corre o risco de ser dada por alguma relação de controle. Portanto, não haveria a necessidade de uso do clítico.

84. *par=48064*: Mas aqui se trata menos de avaliar estratégias semióticas do que detectar o nascimento (o renascimento, **seria melhor dizer**) de uma arte, ou pelo menos de um sistema expressivo, em que a dimensão heurística passa a ser o dado fundamental.  
 85. *par=Brasil--94b-2*: Consideram que, para enfrentar as mesmas dificuldades, **seria melhor passar** diretamente do cruzeiro real para o real sem a URV .  
 86. *par=Brasil--94a-2*: Na primeira barreira, um grupo de olheiros avisou que **seria melhor não ir** ao alto do morro .

Entretanto, há outros casos, como aqueles entre (87) e (90), em que o sujeito do infinitivo poderia ser interpretado quer como controlado, quer como arbitrário, mas que, segundo nossa intuição, não aceitariam a presença do *se* nem mesmo para evitar a ambiguidade.

87. *par=23804*: Champion se recusou a dar outras informações sobre o caso, porque o comitê executivo da KPMG se reuniu na semana passada e chegou à conclusão de que **seria melhor** não falar nada, enquanto a questão não for esclarecida.  
 88. *par=115963*: Quando ia a beber, cogitei se não **seria melhor** esperar que Capitu e o filho saíssem para a missa;  
 89. *par=130694*: Olhe lembra-me até que **seria melhor** desistir de tal compra da fazenda...  
 90. *par=134776*: Campos respondeu: «Não **seria melhor** mandar o trem pagador?»

Cinque oferece a explicação sintática padrão para explicar a impossibilidade da presença do clítico nas orações não finitas do italiano, complementos de *sarebbe meglio*: sendo um elemento lexical de natureza nominal, o clítico precisa ser

parte de uma CADEIA à qual Caso tenha sido atribuído; como essas construções são estruturas de controle, a posição de sujeito do infinitivo, que seria parte da CADEIA da qual participa o clítico, não recebe Caso. Portanto, haveria uma violação do Filtro de Caso.

Mas, nossa questão é: qual seria a razão semântica para a impossibilidade da presença do *se* em orações como essas? Se, como sugerimos para outros casos, o *se* aparece para codificar o participante iniciador de um evento, por que é que nos casos entre (84) e (90), ele parece não poder aparecer?

Essas e outras questões indicam que o mapeamento do comportamento semântico do clítico *se* no português brasileiro está por ser feito. Este artigo teve o objetivo de apontar a necessidade de inclusão desse item na agenda da linguística brasileira.

---

## REFERÊNCIAS BIBLIOGRÁFICAS

- BURZIO, L. (1986). *Italian Syntax: a Government-Binding approach*. Dordrecht: D. Reidel Publishing Company, 468p.
- CANAVER, F. (2012). *Evidências para um modelo baseado no uso: o infinitivo flexionado no português brasileiro*. Dissertação (Mestrado em Linguística) – Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, São Paulo, 148p.
- CAVALCANTE, S. R. O. (2006). *O uso de se com infinitivo na história do português: do português clássico ao português europeu e brasileiro modernos*. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem da Universidade Estadual de Campinas, Campinas, 206p.
- CAVALCANTE, S. R. O. (2010). O uso de *se* com infinitivo do português clássico ao português europeu e brasileiro modernos. *Estudos da Língua(gem)*, Vitória da Conquista, v. 8, n.1, p.55-82.
- CHAGAS DE SOUZA, P. (2000). *A alternância causativa no português do Brasil: defaults num léxico gerativo*. Tese (Doutorado em Linguística) – Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, São Paulo.
- CINQUE, G. (1988). On *Si* constructions and the Theory of *Arb*. *Linguistic Inquiry*, Cambridge, v. 19, n.4, p.521-581.
- DOBROVIE-SORIN, C. (2005). The SE-Anaphor and its Role in Argument Realization. In: EVERAERT, M.; VAN RIEMSDIJK, H. (Eds.). *The Blackwell Companion to Syntax*. Oxford: Wiley-Blackwell, v. IV, p. 118-179.
- DUARTE, M. E. L. (1995). *A perda do Princípio 'Evite Pronome' no Português Brasileiro*. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem da Universidade Estadual de Campinas, Campinas.
- GALVES, C. M. C. (2001). *Ensaio sobre as gramáticas do português*. Campinas: Ed. Da UNICAMP, 280p.
- GIVÓN, T. (1991). Isomorphism in the grammatical code: cognitive and biological considerations. *Studies in Language*, Amsterdam, v. 15, n.1, p.85-114.

- HAIMAN, J. (1980). The iconicity of grammar: isomorphism and motivation. *Language*, Washington, v. 56, n.4, p.515-540.
- HASPELMATH, M. (1993). More on the typology of inchoative/causative verb alternations. In: COMRIE, B.; POLINSKY, M. (Eds). *Causatives and transitivity*. Amsterdam: John Benjamins, p. 87-111.
- KEMMER, S. (1993). *The middle voice*. Amsterdam: John Benjamins, 299p.
- LANGACKER, R. W. (1991). *Foundations of Cognitive Grammar: descriptive applications*. Stanford: Stanford University Press, 589p.
- MALDONADO, R. (2006). *A media voz: problemas conceptuales del clítico se*. México: Universidad Nacional Autónoma, p. 480.
- NARO, A. J. (1976). The genesis of reflexive impersonal in Portuguese. *Language*, Washington, v. 52, n.4, p.779-810.
- NEGRÃO, E. V.; VIOTTI, E. (2008). Estratégias de impessoalização no português brasileiro. In: FIORIN, J.F.; PETTER, M.M.T. (Orgs.). *África no Brasil: a formação da língua portuguesa*. São Paulo: Contexto, p.179-203.
- NEGRÃO, E. V.; VIOTTI, E. (2010). A estrutura sintática das sentenças absolutas no português brasileiro. *Linguística*, Santiago, v. 23, p. 61-82.
- NEGRÃO, E. V.; VIOTTI, E. (2014). Contato entre quimbundo e português clássico: impactos na gramática de impessoalização do português brasileiro e angolano. *Linguística*, Santiago, v. 30, n.2, p. 289-330.
- PERLMUTTER, P. M. (1978). Impersonal passives and the unaccusative hypothesis. *Proceedings of the 4th Annual Meeting of the Berkeley Linguistics Society*, p. 157-190.
- REINHART, T.; SILONI, T. (2004). Against the unaccusative analysis of reflexives. In: ALEXIADOU, A. et al. (Eds). *The unaccusativity puzzle: explorations of the Syntax-Lexicon interface*. Oxford: Oxford University Press, p.159-180.