

## **O PROCESSAMENTO DOS CLÍTICOS *TE* E *LHE* NO PORTUGUÊS BRASILEIRO: A CONTRAPARTE COGNITIVA DA VARIAÇÃO**

PROCESAMIENTO DE LOS CLÍTICOS *TE* Y *LHE* EN EL PORTUGUÉS BRASILEÑO:  
LA CONTRAPARTE COGNITIVA DE LA VARIACIÓN

THE PROCESSING OF CLITICS *TE* AND *LHE* IN BRAZILIAN PORTUGUESE: THE  
COGNITIVE COUNTERPART OF THE VARIATION

Thiago Laurentino de Oliveira  
Universidade Federal do Rio de Janeiro  
thiago.laurentinodeoliveira@gmail.com  
0000-0002-9537-5264

### **Resumo**

Neste artigo, analisamos, a partir da metodologia experimental, a variação entre os clíticos *te* e *lhe* na função de objeto direto. Alguns estudos afirmam que *lhe*, originalmente um clítico dativo de 3SG, é reinterpretado no Português Brasileiro como um acusativo de 2SG, p. ex.: "Eu *lhe* vi", após a difusão da forma *você* no sistema pronominal (Ramos 1999; Almeida 2009; Souza 2014). Desse modo, *lhe* passa a variar com *te* no contexto sintático acusativo. Para analisar a contraparte cognitiva dessa variação, elaboramos um teste de leitura automonitorada (*self-paced reading*), aplicado a participantes nativos de dois estados brasileiros (Rio de Janeiro e Ceará). No experimento, testamos algumas hipóteses recorrentes na literatura sobre o tema: (i) a ambiguidade referencial do clítico *lhe* e (ii) o condicionamento diatópico acerca do uso de *lhe*.

**Palavras-chave:** Variação linguística; Abordagem experimental; Pronomes pessoais.

### **Resumen**

En este artículo se analiza, desde la metodología experimental, la variación entre los clíticos *te* e *lhe* en función de objeto directo. Algunos estudios afirman que, originalmente un clítico dativo de 3SG, se reinterpreta en portugués brasileño como un acusativo de 2SG, p. ej.: "Eu *lhe* vi", luego de la difusión de la forma *você* en el sistema pronominal (Ramos 1999; Almeida 2009; Souza 2014).

De esta manera, *lhe* comienza a variar con *te* en el contexto sintáctico acusativo. Para analizar la contraparte cognitiva de esta variación, desarrollamos una prueba de lectura autocontrolada, aplicada a participantes nativos de dos estados brasileños (Río de Janeiro y Ceará). En el experimento, probamos algunas hipótesis recurrentes en la literatura sobre el tema: (i) la ambigüedad referencial del clítico *lhe* y (ii) el condicionamiento diatópico sobre el uso de *lhe*.

**Palabras clave:** Variación lingüística; Enfoque experimental; Pronombres personales.

### Abstract

In this paper, I analyze, from the experimental methodology, the variation between *te* and *lhe* clitics as direct object. Some studies claim that *lhe*, originally a dative clitics of 3SG, is reinterpreted in Brazilian Portuguese as an accusative of 2SG, e.g.: "Eu *lhe* vi" after expansion of *você* form in the pronominal system (Ramos 1999; Almeida 2009; Souza 2014).

Thus, *lhe* vary with *te* at accusative context. To analyze the cognitive counterpart of this variation, I developed a self-paced reading test, applied to participants of two Brazilian States (Rio de Janeiro and Ceará). In the experiment, I tested some recurrent assumptions in literature about the topic: (i) the referential ambiguity of clitics *lhe* and (ii) the diatopic conditioning about the use of *lhe*.

**Keywords:** Linguistic variation; Experimental approach; Personal pronouns.

Recebido: 28/01/2019

Aceito: 27/05/2019

## 1. Considerações iniciais

No âmbito da expressão pronominal de segunda pessoa do singular (2SG), chama a atenção dos linguistas, há tempos, a complexa realização variável existente nessa parte da gramática do Português Brasileiro (PB). Além de envolver um vasto conjunto de condicionadores linguísticos e sociais, esse fenômeno se destaca também pela quantidade de variantes que recruta. Estas se distribuem entre as funções sintáticas da sentença nas quais pode haver um pronome pessoal.

Dentre essas variantes, analisaremos, neste artigo, as formas clíticas *te* e *lhe*, que podem ocorrer nas funções de objeto direto e objeto indireto no PB contemporâneo. A variação linguística envolvendo essas formas tem sido objeto de diferentes estudos. O ponto de partida para a exploração do tema é, quase sempre, diacrônico, uma vez que a origem desse fenômeno está relacionada ao processo de mudança linguística por que tem passado o quadro

de pronomes pessoais do PB nos últimos séculos, sobretudo após a emergência e difusão da forma gramaticalizada *você* (Lopes *et al.* 2018).

Na tentativa de mapear a dinâmica desse fenômeno de variação, trabalhos baseados em *corpora* têm revelado que, mesmo após a difusão de *você* no quadro pronominal, a forma *te* ainda figura como uma variante bastante produtiva. Brito (2001) propõe como hipótese a questão da ambiguidade referencial: o clítico *lhe*, teoricamente, também pode se referir à 3SG. Além disso, *lhe* não possuiria o traço semântico de [2SG], de modo que a interpretação de 2SG dependeria da presença de outra forma antecedente com esse traço (o pronome *você*, por exemplo).

Paralelamente a essas discussões mais gramaticais, há ainda questões de ordem extralinguística importantes de serem tratadas. Dentre elas, destacamos a hipótese de que a variante *lhe* seria um marcador linguístico (Labov 1972), associado aos dialetos nordestinos (Ramos 1999). Os resultados de pesquisas recentes têm sustentado essa hipótese, registrando um uso considerável dessa variante na fala de indivíduos nativos de localidades nordestinas (Araújo e Carvalho 2015; Almeida 2016).

Diante desses pontos, buscamos contribuir com o tema trazendo para a discussão aspectos relativos ao processamento linguístico de *te* e *lhe* por falantes do PB. Pretendemos, neste estudo, (i) mensurar o custo de processamento exigido pelas formas *te* e *lhe* para a ativação da referência à 2SG e (ii) observar se o fator diatópico exerce influência no processamento das variantes em análise. Para a concretização desses objetivos, adotamos a metodologia experimental da *leitura automonitorada* (*self-paced reading*), uma técnica *on-line* bastante conhecida e utilizada na psicolinguística. O experimento foi aplicado a falantes nativos do PB, oriundos dos estados do Rio de Janeiro e Ceará.

Organizamos este artigo da seguinte maneira: após esta introdução, fazemos uma breve revisão dos estudos acerca dos clíticos *te* e *lhe* na seção 2; delineamos, na seção 3, a técnica experimental escolhida como metodologia alternativa para análise; apresentamos e discutimos os resultados experimentais obtidos na seção 4; em 5, por fim, tecemos as considerações finais acerca deste trabalho.

## 2. Os clíticos *te* e *lhe* como acusativos de 2SG no PB

Nesta seção, apresentamos três aspectos concernentes ao uso dos clíticos *te* e *lhe* como objeto direto (acusativo) de 2SG no PB. Primeiramente, revisamos alguns trabalhos descritivos, de cunho histórico, que tratam das origens das variantes e dos processos de mudança linguística que as afetaram. Em seguida, comentamos os resultados gerais e mais significativos de pesquisas sociolinguísticas recentes sobre a variação entre *te* e *lhe* em dados de *corpora*.

Por fim, discutimos a hipótese diatópica levantada por alguns linguistas de que o clítico *lhe* – como acusativo de 2SG – seria um marcador linguístico no PB, associado aos falares nordestinos.

## 2.1. O percurso histórico das variantes

A variação entre as formas *te* e *lhe* no PB atual decorre de um processo mais amplo, ocorrido no sistema de tratamento do português, em que se verifica a associação de paradigmas pronominais distintos. Faria (1958) afirma que, no latim, os pronomes pessoais apresentavam flexões morfológicas para marcar o caso (nominativo, acusativo, dativo etc.) que a forma recebia na oração. A forma *Tū*, que indicava referência ao interlocutor (2SG), flexionava-se em *tē*, quando exprimia o acusativo (objeto direto), e em *tibī*, quando exprimia o dativo (objeto indireto).

Os ditos pronomes pessoais “do caso reto” do português vieram, segundo o autor, das antigas formas latinas de nominativo (por ex., *tu*); já os pronomes pessoais átonos “do caso oblíquo” (*me, te, nos, se*) são oriundos das formas acusativas latinas. Das flexões dativas (*mihi, tibi, sibi*), surgiu a série de “pronomes oblíquos tônicos” (*mim, ti, si*), que, em português, são obrigatoriamente encabeçados por preposição. Além disso, a língua portuguesa também herdou do latim a distinção *tu* (nominativo-sujeito) e *te* (acusativo/dativo-objeto). Esse resquício do sistema de casos morfológicos latinos sobrevive até hoje e, em grande parte da história do português, o emprego de *te* esteve associado ao paradigma de *tu*.

Apesar dessas correspondências, há diferenças importantes entre os sistemas pronominais latino e português. Câmara Jr. (1985: 96) destaca que o quadro pronominal português está “muito longe do sistema de casos latinos, em que o caso dependia da função do nome ou do pronome na frase”. Williams (1994: 148) também chama a atenção para o fato de que o quadro pronominal do português não deve ser visto como uma cópia fiel dos pronomes latinos, “já que algumas formas de acusativo são usadas como dativo e algumas de nominativo e dativo são usadas como objeto de preposições”.

Os autores ressaltam que as formas homônimas de acusativo e dativo existentes atualmente no português derivam de *tē*, a forma latina do acusativo: “*Me, te, se* (...) são reflexos do acusativo-ablativo *mē, tē, sē*. Mas em português os clíticos adverbiais indicam o objeto direto ou indireto, isto é, equivalem a um acusativo-dativo (...)” (Câmara Jr. 1985: 97); “Essas formas eram do acusativo em latim clássico, mas vieram a ser usadas como dativo e acusativo em português” (Williams 1994: 154).

A forma *lhe* inexistia no latim. Segundo Câmara Jr. (1985), esse pronome surge a partir do demonstrativo *ille* (“aquele”) flexionado no caso dativo, *illi*. Quando passa a funcionar como um pronome pessoal, *lhe* preserva essa marca casual, opondo-se à forma *o/a*, específica para o caso acusativo.

Tempos mais tarde, a gramaticalização e difusão de *você* (< *Vossa Mercê*) no sistema de tratamento do português criou a possibilidade de que certas formas pronominais de 3SG – como o pronome *lhe* – passassem a atuar na referência à 2SG.

Isso porque *Vossa Mercê* era, originalmente, uma construção nominal. Assim, certos pronomes de 3SG foram reinterpretados, em correlação com *você*, e inseridos no âmbito da 2SG. Com isso, o quadro pronominal sofre um rearranjo quanto ao inventário de pronomes pessoais.

Todavia, a implementação de *você* no sistema, juntamente com a migração de formas pronominais de 3SG para 2SG, não implicou no desaparecimento do pronome *tu* e suas formas correlatas (*te*, *teu*, *contigo* etc.).

Devido a isso, temos, no PB atual, um complexo fenômeno de variação linguística envolvendo a representação da 2SG (cf. Lopes *et al.* 2018). No que se refere às formas *te* e *lhe* em função de objeto direto, tema central deste artigo, surge a pergunta: como se dá a variação entre elas nas variedades do PB?

## 2.2. A alternância das variantes em dados de *corpora*

Souza (2014) explora, em sua dissertação, a variação na expressão de 2SG em posição de complemento na variedade fluminense. A autora mapeou todas as variantes de 2SG na função de acusativo/objeto direto levantadas em um *corpus* de 504 cartas pessoais escritas entre as décadas de 1880 e 1980.

De um total de 433 dados, ela contabilizou 337 ocorrências de *te* (77,8%), 40 de *o/a* (9,2%), 29 de *você* (6,7%), 17 de *lhe* (4%) e 10 de objeto nulo (2,3%). Esse resultado destaca, dentre outras coisas, a força da variante *te* na variedade do Rio de Janeiro frente a um uso reduzido da variante *lhe*.

Focalizando especificamente as variantes *te* e *lhe*, distribuímos, no Gráfico 1, as ocorrências dessas formas na diacronia analisada por Souza (2014):

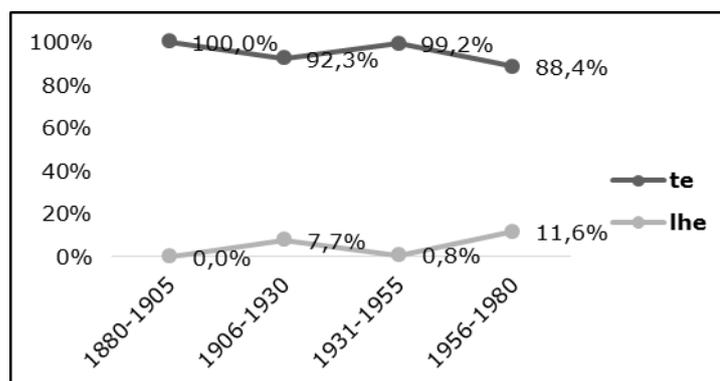


Gráfico 1: Distribuição diacrônica dos clíticos *te* e *lhe* na escrita epistolar fluminense (adaptado de Souza 2014)

Podemos observar facilmente que a variante *te* prevalece nos quatro recortes temporais estudados – 44 ocorrências entre 1880-1905, 132 entre 1906-1930, 123 entre 1931-1955, e 38 entre 1956-1980 –, com uma diferença bastante acentuada em relação aos índices de *lhe* nos mesmos recortes (nenhum dado entre 1880-1905, 11 dados entre 1906-1930, 1 ocorrência entre 1931-1955, e 5 entre 1956-1980).

As linhas do gráfico sugerem, pois, certa estabilidade na frequência de uso dessas variantes clíticas na documentação epistolar produzida ao longo de cem anos.

Ainda em relação à ocorrência da variante *lhe* nas cartas, Souza (2014) afirma que o uso dessa variante está relacionado com a variável *subgênero da carta*.

Na composição da amostra, havia missivas de natureza diversa – amorosas, familiares e pessoais (não-familiares); dentre eles, a pesquisadora constatou que o subgênero *carta pessoal* (no qual havia menor grau de intimidade entre os interlocutores) favorecia fortemente as ocorrências de *lhe*, que, segundo a autora, "assume um caráter [+formal] nesse tipo de missiva." (Souza 2014: 146).

O estudo de Camargo Jr. (2007) explorou a posição de complemento verbal acusativo de 2SG na variedade paulista. Como *corpus*, o autor utilizou produções textuais de estudantes do segundo segmento do ensino fundamental de um colégio particular da cidade de São Paulo. Os alunos foram solicitados a redigir uma carta, o que, naturalmente, propiciou a ocorrência de objetos diretos de 2SG. A tarefa foi dividida em duas modalidades – formal e informal.

Ao todo, Camargo Jr. (2007) contabilizou 1.524 dados de objeto direto de 2SG, sendo 595 deles produzidos na modalidade formal (39%) e 929, na modalidade informal (61%). Quanto às estratégias utilizadas, houve a ocorrência de clíticos, pronomes tônicos, sintagmas nominais e de objeto nulo. As formas clíticas foram majoritárias no corpus, correspondendo a 1019 dados (66,9% da amostra). Apresentamos, na Tabela 1, apenas os percentuais das variantes *te* e *lhe*, reportados pelo autor:

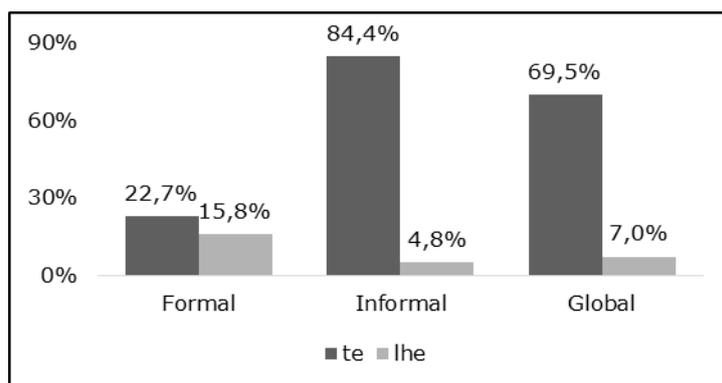


Tabela 1: A produtividade dos clíticos *te* e *lhe* em função acusativa na redação de estudantes paulistanos (adaptado de Camargo Jr. 2007: 130)

As barras percentuais ilustram que o pronome *te* foi a variante mais produzida pelos estudantes, com um percentual global próximo a 70% (708 dados, sendo 56 destes na modalidade formal e 652 na modalidade informal). Já o clítico *lhe* obteve uma frequência bem mais reduzida, correspondendo a 7% das formas clíticas produzidas pelos alunos. No contexto formal, essa variante teve um percentual de 15,8% (39 dados), 11% a mais do que foi verificado para o contexto informal (4,8%). Essa diferença vai ao encontro da hipótese levantada por Souza (2014) de que o clítico *lhe* exprime marca de formalidade.

Com base na variedade cearense, Araújo e Carvalho (2015) também abordam a variação entre *te* e *lhe* na função de objeto direto. Assim como Souza (2014), os pesquisadores também examinam uma amostra de cartas pessoais, composta por 186 documentos produzidos entre 1940 e 1990. Nesse *corpus*, os autores encontraram 149 ocorrências de clíticos de 2SG, sendo 90 dados do pronome *te* (60%) e 59 do pronome *lhe* (40%).

Ao analisarem a distribuição dos dados ao longo do período investigado, Araújo e Carvalho (2015) verificaram que *lhe* é predominante nas cartas escritas entre 1940 e 1950 (56%, 09/16 oco.), em comparação com as ocorrências de *te* (44%, 07/16 oco.); já entre as décadas de 1960 e 1970, o cenário se inverte em favor de *te* (54%, 34/63 oco.), que se sobrepõe ao índice de *lhe* (46%, 29/63 oco.); entre as décadas de 1980 e 1990, a distância entre os percentuais de *lhe* e *te* acentua-se, com 70% (49/70 oco.) de frequência para este, frente a 30% (21/70 oco.) para aquele. Vejamos os referidos percentuais projetados no Gráfico 2:

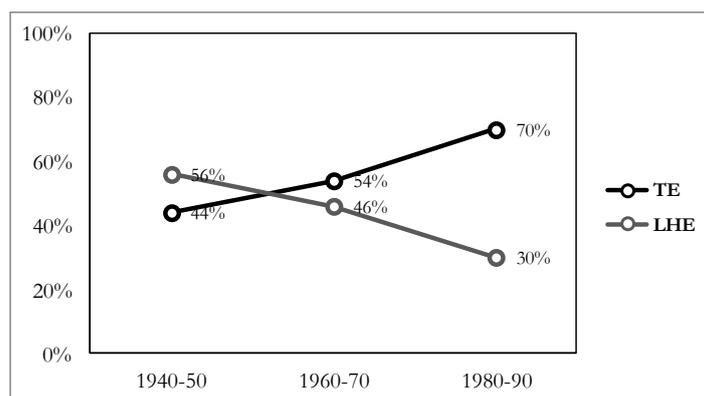


Gráfico 2: Distribuição diacrônica de *te* e *lhe* acusativos na escrita epistolar cearense (adaptado de Araújo e Carvalho 2015: 76)

O Gráfico 2 ilustra uma distribuição bastante interessante, pois vai de encontro à ideia de que *lhe* é a variante típica dos dialetos nordestinos. O índice de *lhe* nas cartas escritas entre 1980-1990 é superior àquele verificado, para o mesmo período, por Souza (2014), em cartas cariocas (pouco mais de 10%); entretanto, o gráfico sugere que a forma *lhe*, na variedade cearense, está caindo em desuso, enquanto *te* parece se estabelecer como a variante majoritária para a função de acusativo.

Outro estudo que também explora a variação *te/lhe* no objeto direto é o de Almeida (2016). A pesquisadora investigou o uso dos pronomes acusativos de 2SG nas cidades de Salvador e Santo Antônio de Jesus (Bahia). Para tanto, a autora examinou amostras de fala de 24 informantes (12 de cada cidade), estratificados segundo o sexo, faixa etária e escolaridade.

A partir de questionários construídos para induzir à produção de formas pronominais de 2SG pelos informantes durante a entrevista, Almeida (2016) obteve 516 dados (218 dados em Salvador e 298 dados em Santo Antônio de Jesus). Das 218 ocorrências em Salvador, 91 foram do clítico *lhe* (42%) e 83, do clítico *te* (38%). Já em Santo Antônio de Jesus, dos 298 dados coletados, houve 122 ocorrências de *lhe* (41%) e 92 de *te* (31%). Esses números revelam uma presença notável de *lhe* na variedade baiana. Almeida (2016) correlacionou, ainda, a variável dependente com o fator extralinguístico *faixa etária*, o que forneceu outro dado interessante, conforme reproduzimos no Gráfico 3:

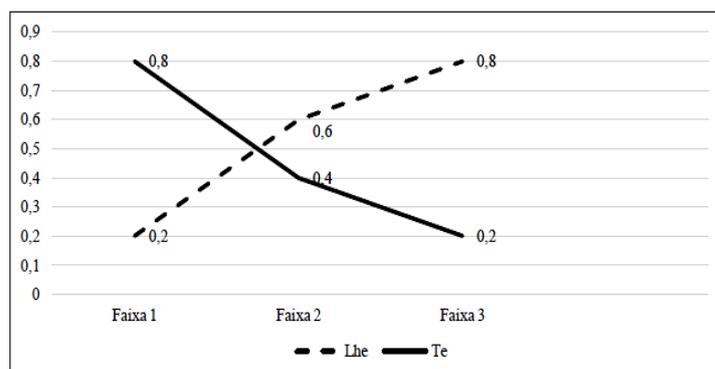


Gráfico 3: A atuação da variável *faixa etária* sobre os clíticos de 2SG em Salvador e Santo Antônio de Jesus (BA) (extraído de Almeida 2016: 130)

Lançando mão dos índices de peso relativo, Almeida (2016) afirma que os falantes mais idosos da amostra favorecem o uso de *lhe* (.800). Entre os indivíduos da faixa 2, esse favorecimento permanece, porém em menor índice (.600). Já os informantes mais jovens da amostra exibem um padrão oposto, visto que favorecem o uso da variante *te* (.800). Sob a ótica do modelo de mudança em curso na variedade baiana, segundo o qual o pronome *te* estaria se generalizando na posição de objeto direto.

Com base nos resultados das quatro pesquisas mencionadas, vimos que (i) as formas *te* e *lhe* constituem, de fato, variantes pronominais de 2SG na função de objeto direto, ocorrendo em diferentes variedades do PB e (ii) a distribuição dessas variantes parece oscilar significativamente de uma região para outra, tendo *lhe* uma frequência menos significativa em localidades da região sudeste (RJ e SP) do que em cidades da região nordeste (CE e BA).

Diante desse cenário, fica a questão: as diferenças observadas quanto à frequência da variante *lhe* estariam condicionadas à localidade de origem dos falantes do PB?

### 2.3. O uso de *lhe* como marcador linguístico: a hipótese diatópica

Diversos estudiosos têm defendido que o uso expressivo do pronome *lhe* como forma acusativa de 2SG seria um caso de variação diatópica, uma vez que esse pronome estaria relacionado aos falares nordestinos. Almeida (2009: 18), por exemplo, ao discutir sobre o estatuto do clítico *lhe* no PB atual, salienta a produtividade desse pronome na referência à 2SG e sua extensão para a função acusativa, destacando que “embora o fenômeno esteja presente em todo o Brasil (...), é em estados do Nordeste que ele se evidencia de maneira mais acentuada (...)”.

A autora toma por base o trabalho de Ramos (1999), que analisou o emprego dos clíticos de 3SG *le/lhe* como objeto direto no Espanhol Peninsular e no PB. Ramos (1999) relaciona o uso de *lhe* na 2SG à perda da distinção entre os pronomes *tu* e *você* em diversas regiões do Brasil, o que permitiu que *lhe* dativo de 3SG “invadisse” a função acusativa do pronome *te*. Diante do complexo quadro de reestruturação do sistema de tratamento, a pesquisadora postula a existência de três subsistemas no PB atual, tratados como três “gramáticas”, segundo a perspectiva gerativa adotada por ela:

- (i) **Gramática A:** *você* é a expressão universal de tratamento; *lhe* é a expressão de dativo de 2SG em relações de respeito/cortesia; *te* é a forma acusativa/dativa para o tratamento familiar/informal. Tal gramática corresponderia às variedades do eixo Rio-São Paulo.
- (ii) **Gramática B:** *você* é a forma generalizada de tratamento; *lhe* substituiu *te* nas funções acusativa e dativa e, por essa razão, não se limita às situações de formalidade/respeito/cortesia, estendendo-se ao tratamento familiar/informal. Essa gramática representaria grande parte dos estados nordestinos, dentre os quais se destacam Maceió, Recife, Salvador e João Pessoa.
- (iii) **Gramática C:** a distinção entre *tu* íntimo/familiar e *você* respeitoso/cortês se mantém; os usos de *te* e *lhe* obedecem a mesma distinção, sendo *te* utilizado como clítico de *tu* e *lhe* como clítico de *você*, *senhor/senhora*, tanto no dativo quanto no acusativo. Tal gramática compreenderia a região Norte e o estado do Maranhão.

Tal proposta consiste mais em uma hipótese do que em um mapeamento empiricamente verificado. A própria autora alerta para esse fato, advertindo que será necessária uma pesquisa mais ampla sobre a questão, a fim de que se possam fazer afirmações mais concretas.

Com efeito, se voltarmos aos estudos comentados em 2.2, perceberemos algumas incongruências entre a proposta de Ramos (1999) e o que já foi atestado em dados de *corpora*. Primeiramente, vemos nos dados de Souza (2014) – que investigou a variedade do RJ – e de Camargo Jr. (2007) – que analisou a variedade de SP – a ocorrência, ainda que pequena, do clítico *lhe* na função de objeto direto (relacionada ao tratamento mais formal); esses dados relativizariam a projeção da “Gramática A”, que não prevê o uso desse clítico como acusativo de 2SG.

Além disso, temos nos dados de Araújo e Carvalho (2015) e Almeida (2016) índices da forma *te* que invalidam a existência de uma “Gramática B” tal como fora postulada. Em vez de uma substituição completa de *te* por *lhe*, os resultados das pesquisas ilustram um acirrado quadro de variação linguística, inclusive com uma tendência à “revitalização” de uso de *te* na variedade baiana atual (Almeida 2016).

Apesar disso, é preciso reconhecer que a associação entre o uso de *lhe* como objeto direto de 2SG e as variedades nordestinas encontra respaldo nos resultados das investigações sociolinguísticas. A diferença de uso dessa variante nos parece ser mais quantitativa do que qualitativa. É mais coerente afirmar que *lhe* de 2SG tem um uso mais expressivo nas localidades nordestinas, em contraste com outras localidades do Brasil, do que dizer que esse pronome é a única forma de 2SG na função acusativa nessas regiões.

Essa notável diferença quantitativa no uso de *lhe*, registrada entre diferentes localidades do Brasil, parece impactar a percepção dos falantes, que avaliariam essa variante, nos termos de Labov (1972), como um *marcador sociolinguístico*.

Embora sua utilização não gere estigma social sobre os usuários da variante, *lhe* é condicionado por fatores estilísticos – como grau de formalidade

– e sociais – como cidade de origem e faixa etária. Isso explica o fato de que diferentes estudiosos identifiquem o Nordeste como a área onde a realização de *lhe* 2SG se dá com maior intensidade, sendo tratada como uma marca dialetal.

Conduzindo a discussão para o viés experimental, levantamos a seguinte questão: como os falantes de dois dialetos distintos (um que utilize expressivamente a forma *lhe* como 2SG e outro que não a utilize) processam os clíticos *te* e *lhe* na função acusativa? Na tentativa de explorar essa questão, reportamos, a seguir, o projeto e os resultados de um experimento linguístico envolvendo essas variantes.

### 3. Metodologia experimental

Na presente seção, descrevemos o teste de leitura automonitorada elaborado por Oliveira (2018) e aplicado em dois estados brasileiros – Rio de Janeiro e Ceará. O intuito principal era verificar a eficiência dos clíticos *te*, *lhe* e *o/a* para ativar a informação de 2SG na mente dos falantes. Embora o clítico *o/a* também tenha sido investigado por Oliveira (2018), nos restringiremos às análises dos clíticos *te* e *lhe*, tendo em vista os propósitos deste artigo e as limitações desta exposição.

#### 3.1. Desenho do experimento

A técnica da *leitura automonitorada* compreende a exibição de um enunciado dividido em partes menores (palavras, sintagmas ou mesmo sentenças completas). Os participantes leem os enunciados na tela do computador segundo uma rotina previamente estabelecida: para iniciar a tarefa, apertam uma tecla que traz à tela o primeiro segmento do enunciado; após a leitura deste, os sujeitos vão acionando repetidamente a mesma tecla para ler os demais segmentos, até que tenham lido o enunciado inteiro. Essa rotina justifica a identificação de “automonitorada” (ing. *self-paced reading*; Mitchell 2004; Garrod 2006), já que é o próprio participante que aciona o aparecimento dos estímulos na tela do computador.

Para o pesquisador, interessa analisar os intervalos de tempo entre a exibição de cada segmento, a partir do ritmo de acionamento da tecla de comando pelos participantes. A lógica da tarefa é que tempos de leitura maiores em segmentos específicos do texto podem indicar maior custo de processamento linguístico dos itens presentes nestes segmentos.

#### 3.2. Método

A fim de testar a eficiência dos clíticos *te* e *lhe* na ativação de um referente de 2SG, foram criados itens experimentais que envolviam a menção de três nomes próprios dentro de uma estrutura de subordinação completiva com verbos *dicendi*. Nessa configuração, um dos nomes próprios admitia a interpretação de referente de 2SG. As frases experimentais apresentavam o seguinte formato:

01. Sônia prometeu para **Aldo** na companhia de **Davi**: Eu **te** ajudo com a prova.  
PERGUNTA: Sônia ajuda Davi com a prova?

Na oração matriz, com o verbo *dicendi*, aparecem os três nomes próprios. Além de introduzir referentes com traço [+pessoa], essa sentença forja uma situação dialógica, necessária para a inserção de pronomes de 2SG. Na oração completiva, aparecem dois pronomes pessoais: “eu”, ligado ao sujeito da primeira sentença – em (01), [Sônia] –, e “te”, vinculado ao constituinte da primeira sentença ([para Aldo], em (01)).

Depois de ler as sentenças com os nomes próprios e com o clítico, os participantes tinham de responder a uma pergunta relacionada com o clítico. No exemplo em (01), a resposta esperada seria “não”, uma vez que o clítico *te* se refere a “Aldo”. “Davi”, nesse contexto, é um referente de 3SG, fora do eixo falante-ouvinte, que presencia a conversa entre “Sônia” e “Aldo”.

Dentro desse formato, torna-se possível verificar se o clítico *lhe* gera ambiguidade referencial (Brito 2001). Em (02), temos o mesmo enunciado na versão com *lhe*:

02. Sônia prometeu para **Aldo** na companhia de **Davi**: Eu **lhe** ajudo com a prova.  
PERGUNTA: Sônia ajuda Davi com a prova?

Nesta versão, “Davi” pode ser interpretado como o referente de *lhe*, caso os participantes optem por uma interpretação de 3SG desse clítico. Essa possibilidade é vedada para *te*, que funciona somente como pronome de 2SG. Dessa forma, é possível examinar, no experimento, qual leitura será atribuída a *lhe* em contextos ambíguos.

### 3.3. Participantes

Participaram do experimento 60 sujeitos, sendo 30 naturais do Rio de Janeiro e 30 do Ceará. Destes, 24 eram do sexo masculino. Os participantes tinham, em média, 25 anos de idade. Todos possuíam nível superior (completo ou em andamento).

### 3.4. Materiais

Foram criados 30 enunciados experimentais, sendo 10 enunciados para cada tipo de clítico (*te*, *lhe* e *o/a*<sup>1</sup>). Os participantes foram organizados em 3 grupos distintos e cada grupo visualizava 10 itens experimentais mais 20 enunciados distratores. A divisão dos enunciados por grupos adotou a organização descrita em (03):

03.  
Grupo 1: 5 enunciados com CLÍTICO TE + 5 enunciados com CLÍTICO LHE + 20 distratores  
Grupo 2: 5 enunciados com CLÍTICO LHE + 5 enunciados com CLÍTICO O/A + 20 distratores  
Grupo 3: 5 enunciados com CLÍTICO TE + 5 enunciados com CLÍTICO O/A + 20 distratores

<sup>1</sup> Conforme já esclarecemos anteriormente, não discutiremos os resultados relativos ao clítico *o/a*.

Os enunciados distratores apresentavam a mesma estrutura sintática das frases experimentais, porém não continham nenhuma forma pronominal de 2SG. Quanto à distribuição dos itens experimentais, adotou-se o formato conhecido na abordagem experimental como *within subjects* (“dentro sujeitos”), segundo o qual todos os participantes são expostos a todas as condições experimentais. Cada grupo era formado por 20 participantes, sendo 10 nativos do Rio de Janeiro e 10 do Ceará. A inserção dos sujeitos nesses grupos foi aleatória.

Os enunciados experimentais e distratores foram inseridos no software *PsychoPy v1.8* (Peirce 2009). Com essa ferramenta, foi possível programar a apresentação dos estímulos no computador e registrar as respostas e os tempos de leitura dos participantes. O *PsychoPy* também permitiu a randomização dos enunciados, de modo que cada participante lia as frases experimentais em sequências diferentes.

### 3.5. Variáveis e condições

No experimento em questão, foram controladas duas variáveis independentes: o *tipo de clítico* e o *estado de origem*. A primeira continha três níveis, enquanto a segunda apresentava dois níveis:

04.

- (A) TIPO DE CLÍTICO: *te*, *lhe* ou *o/a*
- (B) ESTADO DE ORIGEM: *Rio de Janeiro* e *Ceará*

O experimento lidou, portanto, com seis condições experimentais: *te-RJ*, *lhe-RJ*, *o/a-RJ*, *te-CE*, *lhe-CE* e *o/a-CE*. Dessas, discutiremos aqui – vale frisar novamente – os resultados de quatro condições: *te-RJ*, *lhe-RJ*, *te-CE* e *lhe-CE*. No que tange às variáveis dependentes, obtiveram-se, a partir do teste de leitura automonitorada, duas medidas: a *resposta das perguntas interpretativas* e o *tempo de leitura das frases*:

05.

- (a) RESPOSTA DAS PERGUNTAS INTERPRETATIVAS: *sim* ou *não*
- (b) TEMPO DE LEITURA DAS FRASES: *em milissegundos (ms)*

### 3.6. Procedimentos

O experimento foi aplicado aos participantes em um “*Macbook Pro*, monitor de 15” (*Apple*, Macintosh). Cada sujeito recebia, individualmente, do pesquisador responsável, as orientações para realizar o teste. Todos os participantes cumpriam uma etapa de treino antes de iniciar o teste, na presença do pesquisador, a fim de receber orientações específicas e/ou de sanar dúvidas acerca da tarefa. Os sujeitos eram instruídos a manusear apenas a tecla “espaço” (para avançar para as telas seguintes) e, na tela da pergunta, as teclas “a” (para “sim”) e “l” (para “não”). Essas teclas foram sinalizadas com um adesivo colorido.

Nas instruções, o pesquisador esclarecia a rotina do teste: clicando “espaço”, apareceria um asterisco no centro da tela, o que indicaria a exibição de um novo enunciado na tela seguinte; clicando “espaço” novamente, surgiria

a primeira frase, que deveria ser lida no menor tempo possível; acionando “espaço” pela terceira vez, apareceria a segunda frase, que também deveria ser lida o mais rápido possível; em mais um “espaço”, surgiria uma pergunta relativa ao enunciado lido anteriormente. A pergunta só desapareceria da tela quando o participante clicasse em “a” ou “l”, o que também deveria ser feito rapidamente. Respondida a pergunta, apareceria novamente um asterisco no centro da tela. Tal processo se repetiria até o fim, quando uma mensagem de agradecimento seria mostrada.

Após a etapa de treino, os participantes iniciavam o experimento propriamente dito. Durante a tarefa, os sujeitos ficavam sozinhos ou em um espaço suficientemente isolado no local de aplicação. Os participantes levaram, em média, 5 minutos para concluir a tarefa.

Os estímulos eram exibidos em meio escrito, de modo não-cumulativo, na cor preta, em fundo branco. Foi utilizada a fonte *Times New Roman*, tamanho 14, centralizado. Cada segmento do enunciado ocupava uma única linha. Quanto ao tamanho dos segmentos, a primeira sentença possuía 15 sílabas, a segunda sentença, 9 sílabas, e a pergunta interpretativa, 11 sílabas.

### 3.7. Previsões

Como previsões para o experimento, destacamos que:

(i) No que se refere à variável dependente *respostas das perguntas interpretativas*, podemos esperar que, se o pronome *lhe* for, de fato, uma variante ambígua, que admite tanto a leitura de 2SG quanto a de 3SG, então os índices de respostas para a interpretação de 2SG desse clítico deverão ser significativamente menores do que os índices verificados para o clítico *te*, que nunca provoca ambiguidade;

(ii) Quanto à variável dependente *tempo de leitura das frases*, podemos prever um tempo de leitura significativamente maior para as frases relacionadas ao pronome *lhe*, em comparação com o tempo de leitura das frases relacionadas a *te*, caso aquele pronome tenha uma interpretação ambígua, o que demandará maior esforço cognitivo e tempo de processamento durante a leitura dos enunciados pelos participantes;

(iii) Relacionando a *hipótese diatópica* à tarefa experimental em questão, podemos projetar que, se o clítico *lhe* funciona de maneira distinta na “gramática” dos falantes nordestinos e dos falantes da região sudeste (Ramos 1999), então os participantes cearenses – que utilizam frequentemente essa variante em seu vernáculo – deverão exibir uma *performance* diferenciada dos participantes fluminenses, isto é, os cearenses deverão optar, com frequência significativamente maior, pela leitura de 2SG de *lhe* e gastar um tempo significativamente menor para ler os enunciados com esse pronome.

## 4. Resultados e análises

### 4.1. Resultados

Organizamos a exposição dos resultados segundo as variáveis dependentes analisadas. Primeiramente, expomos os resultados relativos à *resposta da pergunta interpretativa*, traduzida no Gráfico 4 como “interpretação da informação de 2SG”:

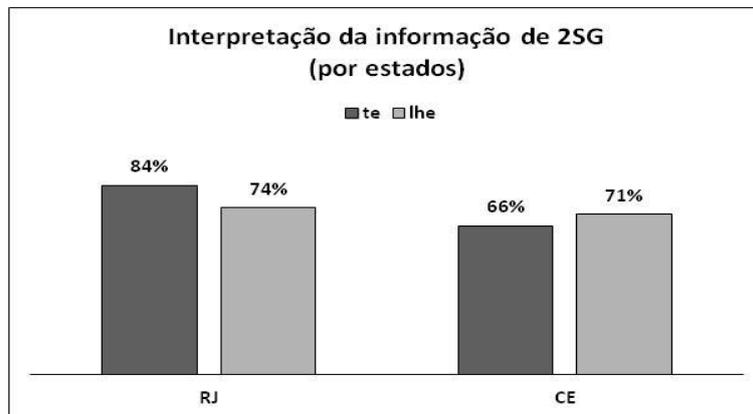


Gráfico 4: Interpretação da informação de 2SG segundo o tipo de pronome (em %).

Verificamos, no Gráfico 4, que *te* favoreceu a interpretação de 2SG entre os participantes do RJ mais frequentemente (84%) do que o clítico *lhe* (74%). Apesar da distância de 10% entre os índices de respostas, a aplicação do teste estatístico revelou que a diferença entre esses percentuais não é significativa ( $\chi^2=2,44$  (1),  $p = 0,11$ ). Entre os participantes cearenses, a forma que mais propiciou a leitura de 2SG foi *lhe* (71%), em detrimento da forma *te* (66%). Entretanto, essa diferença também não foi estatisticamente significativa ( $\chi^2=0,37$  (1),  $p = 0,54$ ).

Confrontando os índices de interpretação da informação de 2SG segundo o tipo de clítico nos dois estados, observamos uma diferença percentual maior nas condições com *te*: 84% nas respostas dos participantes fluminenses frente a 66% nas respostas dos cearenses. Nas condições com *lhe*, os índices são bem próximos: 74% (RJ) e 71% (CE). Analisando esses percentuais estatisticamente, temos que as diferenças diatópicas não são significativas (*te*:  $\chi^2=2,16$  (1),  $p = 0,14$ ; *lhe*:  $\chi^2=0,06$  (1),  $p = 0,80$ ).

Prosseguindo com a análise, examinamos, no Gráfico 5, os resultados relativos às *médias de tempo de leitura da frase com o clítico*:

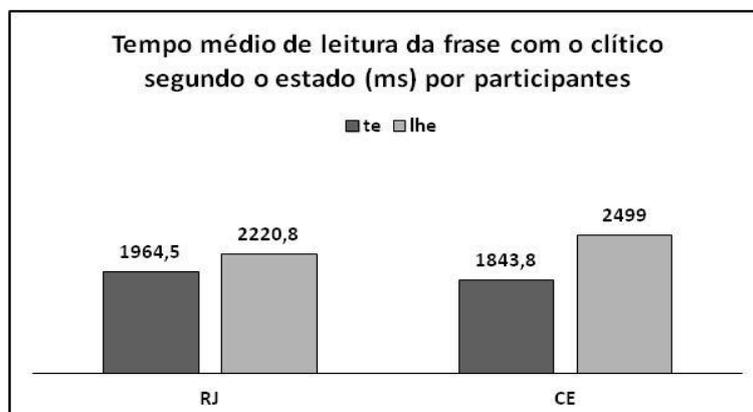


Gráfico 5: Média de tempo de leitura (ms) das frases com o clítico pelos participantes.

Percebemos que tanto os participantes do RJ quanto os participantes do CE gastaram menos tempo para ler as frases que continham o clítico *te*: registrou-se um tempo médio de 1964,5 milissegundos para os primeiros e de 1843,8 milissegundos para os últimos. As frases que continham o clítico *lhe* contabilizaram médias maiores, sendo 2220,8 milissegundos entre os participantes do RJ e 2499 milissegundos entre os participantes do CE. A análise estatística indicou que a média de leitura para as frases com *te* foi significativamente menor em relação à média da frase com *lhe* nos dados do Ceará ( $t(19)=3,82$   $p < 0,01$ ); o mesmo não se verifica para a diferença entre as médias de tempo dos participantes do Rio de Janeiro ( $t(19)=1,52$   $p = 0,14$ ).

Dando continuidade à apresentação dos resultados, passamos então às médias de tempo de leitura da pergunta interpretativa:

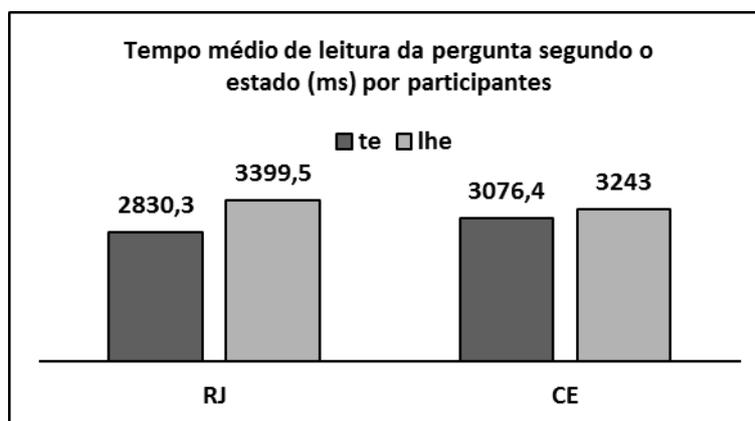


Gráfico 6: Média de tempo de leitura (ms) da pergunta interpretativa pelos participantes.

Novamente, vemos que o tempo médio de leitura das perguntas que sucediam as frases com o clítico *te* pelos indivíduos do RJ foi de 2830,3 milissegundos, mais baixo do que a média registrada para a leitura das perguntas que sucediam as frases com o clítico *lhe* (3399,5 milissegundos). Essa diferença também ocorreu entre as médias dos indivíduos do CE: 3076,4 milissegundos para as perguntas que sucediam as frases com *te* e 3242 milissegundos para as perguntas subsequentes às frases com *lhe*. Na verificação estatística, detectou-se diferença significativa apenas entre as médias de tempo dos participantes do RJ ( $t(19)=2,69$   $p < 0,05$ ), não sendo relevante a diferença observada entre as médias dos participantes do CE ( $t(19)=0,83$   $p = 0,41$ ).

## 4.2. Discussão

Os resultados do experimento de leitura automonitorada sinalizaram que os falantes fluminenses e cearenses parecem processar de modo semelhante as variantes *te* e *lhe* em função acusativa. Ambos optaram, majoritariamente, pela leitura de 2SG para os enunciados que traziam o clítico *lhe*. No entanto, os participantes das duas localidades leram mais rapidamente os estímulos experimentais que traziam o clítico *te*.

Com relação à variável dependente *resposta da pergunta interpretativa*, verificamos que as duas variantes ativaram, na maioria dos casos, a leitura de

2SG na mente dos participantes dos dois estados. Embora os percentuais indiquem maior ativação da informação de 2SG em enunciados com *te* pelos sujeitos do RJ (84%) e em enunciados com *lhe* pelos sujeitos do CE (71%), a checagem estatística da significância das diferenças registradas não nos permite falar em comportamentos distintos entre os falantes das duas variedades. Parece mais acertado considerar que tais diferenças nos percentuais se encontram dentro da faixa de aleatoriedade, previsível em testes com respostas *sim/não*.

No que diz respeito à variável *tempo médio de leitura* (da frase com o clítico e da pergunta interpretativa), constatamos que houve um efeito da variável *tipo de clítico* nas diferenças entre os tempos de leitura dos enunciados experimentais, em favor da variante *te*. Tanto os participantes fluminenses quanto os participantes cearenses leram mais lentamente as frases com o clítico *lhe*. Essa diferença foi estatisticamente significativa em momentos diferentes da tarefa: durante a leitura das frases com os pronomes para os participantes do CE e durante a leitura das perguntas interpretativas para os participantes do RJ.

Segundo os resultados desse experimento, não é possível sustentar a hipótese segundo a qual a forma *lhe* é ambígua para os falantes do PB, pela possibilidade de poder se referir à 3SG. Os participantes dos dois estados efetuaram predominantemente a leitura de 2SG para esse clítico. Além disso, também não atestamos um efeito significativo da hipótese diatópica no desempenho dos participantes cearenses: estes tiveram um comportamento similar ao dos participantes fluminenses, registrando um tempo de processamento maior para efetuar a leitura dos enunciados com o clítico *lhe*.

Não podemos dizer, entretanto, que a hipótese diatópica deva ser descartada. Acreditamos que outras variáveis de natureza cognitiva (ou melhor, sociocognitiva) não controladas neste experimento possam atuar no favorecimento da variante *lhe*, mais frequentemente encontrada nos dialetos nordestinos. Somente após a realização de mais pesquisas experimentais sobre o tema é que poderemos alcançar um entendimento mais claro dessa questão.

## 5. Considerações finais

Neste artigo, abordamos a variação *te/lhe* em função acusativa, fenômeno diretamente relacionado com o processo de reestruturação do sistema de tratamento por que passa o PB desde a emergência e difusão da forma gramaticalizada *você*. Apesar das origens distintas dessas formas, pesquisas sincrônicas e diacrônicas a partir de *corpora* do PB revelam que *te* e *lhe* atuam como variantes da expressão de 2SG na função de objeto direto em diferentes dialetos do país.

A partir dos resultados de estudos anteriores, inserimos o tema na abordagem experimental, a fim de trazer dados relativos à contraparte cognitiva da variação linguística. Dentre as hipóteses existentes acerca do tema, revisitamos, em especial, duas delas: (i) a ambiguidade referencial do clítico *lhe*, que, devido às suas origens, pode se referir à 3SG ou à 2SG; (ii) o

condicionamento diatópico de *lhe*, que é visto como uma marca dos dialetos nordestinos.

Os resultados obtidos através do teste de leitura automonitorada, no entanto, forneceram evidências contrárias a tais hipóteses. Os participantes dos dois estados investigados (RJ e CE) efetuaram predominantemente a leitura de 2SG para os enunciados ambíguos em que havia o clítico *lhe*. Além disso, os mesmos sujeitos exibiram um comportamento bastante similar quanto ao processamento das variantes analisadas, mesmo pertencendo a variedades do PB distintas quanto ao uso de *lhe*.

A divergência verificada entre as hipóteses construídas a partir de *corpora* linguísticos e os resultados do teste experimental revela a necessidade de que mais pesquisas sejam promovidas. Longe de trazer resultados conclusivos, esperamos, com este estudo, instigar a realização de novas análises, que visem a ampliar a compreensão dos aspectos cognitivos atinentes aos fenômenos variáveis do PB e estabelecer um diálogo profícuo entre os estudos baseados no uso e no processamento linguístico.

## Referências bibliográficas

- Almeida, Gilce de Souza. 2009. *Quem te viu quem lhe vê: a expressão do objeto acusativo de referência à segunda pessoa na fala de Salvador*. Tese de Mestrado, Universidade Federal da Bahia. Salvador. [em linha] Disponível em: <http://repositorio.ufba.br/ri/handle/ri/10995>
- Almeida, Gilce de Souza. 2016. Uso dos pronomes-objeto de segunda pessoa na fala de Salvador e de Santo Antônio de Jesus, *Veredas*, 20: 122-135.
- Araújo, Francisco Jardes Nobre de e Hebe Macedo de Carvalho. 2015. TE e LHE como clíticos acusativos de 2ª pessoa em cartas pessoais cearenses, *LaborHistórico*, 1: 62-80.
- Brito, Onilda Regina Marchioni de. 2001. *"Faça o mundo te ouvir": a uniformidade de tratamento na história do português brasileiro*. Tese de mestrado, Universidade Estadual de Londrina. Londrina. Inédita.
- Câmara Jr., Joaquim Mattoso. 1985. *História e estrutura da língua portuguesa*, 4ª ed., Rio de Janeiro, Padrão.
- Camargo Jr., Arnaldo Rebello. 2007. *A realização do objeto direto em referência ao interlocutor*. Tese de mestrado, Universidade de São Paulo. São Paulo. [em linha] Disponível em: <https://teses.usp.br/teses/disponiveis/8/8142/tde-03032008-114747/pt-br.php>
- Faria, Ernesto. 1958. *Gramática superior da língua latina*, Rio de Janeiro, Livraria Acadêmica.
- Garrod, Simon. 2006. Psycholinguistic research methods, em Keith Brown (ed.), *Encyclopedia of Language & Linguistics*, 2ª ed., Amsterdam, Elsevier Science: 251-257.
- Labov, William. 1972. *Sociolinguistic patterns*, Philadelphia, University of Pennsylvania Press.
- Labov, William. 1994. *Principles of linguistic change: internal factors*, Oxford, Blackwell.

- Lopes, Célia Regina, Leonardo Lennertz Marcotulio, Rachel Lucena, Thiago Laurentino de Oliveira e Camila Duarte de Souza. 2018. A reorganização do sistema pronominal de 2ª pessoa na história do português brasileiro: outras relações gramaticais, em Célia Regina Lopes (coord.), *Mudança sintática das classes de palavra: perspectiva funcionalista, história do português brasileiro*, Vol. 4, São Paulo, Contexto: 142-185.
- Mitchell, Don. 2004. On-line methods in language processing: introduction and historical review, em Manuel Carreiras e Charles Clifton Junior (eds.), *The on-line study of sentence comprehension: eyetracking, ERPs and beyond*, New York, Psychology Press: 15-32.
- Oliveira, Thiago Laurentino de. 2018. *Língua e percepção: o processamento dos clíticos com referência ao interlocutor no português brasileiro*. Tese de doutorado, Universidade Federal do Rio de Janeiro. Rio de Janeiro. [em linha] Disponível em: [https://www.academia.edu/36304117/L%C3%8DNGUA\\_E\\_PERCEP%C3%87%C3%83O\\_O\\_Processamento\\_dos\\_cl%C3%ADticos\\_com\\_refer%C3%A2ncia\\_ao\\_interlocutor\\_no\\_Portugu%C3%AAs\\_Brasileiro](https://www.academia.edu/36304117/L%C3%8DNGUA_E_PERCEP%C3%87%C3%83O_O_Processamento_dos_cl%C3%ADticos_com_refer%C3%A2ncia_ao_interlocutor_no_Portugu%C3%AAs_Brasileiro)
- Peirce, Jonathan. 2009. Generating stimuli for neuroscience using PsychoPy. *Front. Neuroinform*, 2: 10.
- Ramos, Conceição de Maria Araújo. 1999. *O clítico de 3ª pessoa: um estudo comparativo português brasileiro/espanhol peninsular*. Tese de doutorado, Universidade Federal de Alagoas. Maceió. Inédita.
- Souza, Camila Duarte de. 2014. *Eu te amo, eu lhe adoro, eu quero você: a variação das formas de acusativo de 2ª pessoa em cartas pessoais (1880-1980)*. Tese de mestrado, Universidade Federal do Rio de Janeiro. Rio de Janeiro. [em linha] Disponível em: <http://www.posvernaculas.letras.ufrj.br/images/Posvernaculas/3-mestrado/dissertacoes/2014/3-SouzaCD.pdf>
- Williams, Edwin Bucher. 1994. *Do latim ao português: fonologia e morfologia históricas da língua portuguesa*, 6ª ed., (Trad. Antônio Houaiss), Rio de Janeiro, Tempo Brasileiro.

NOTA:

O Autor é o único responsável do conteúdo e da forma deste artigo.