

TÓPICOS E SUJEITOS NO PB: UMA ABORDAGEM EXPERIMENTAL

TOPICS AND SUBJECTS IN BRAZILIAN PORTUGUESE: AN EXPERIMENTAL APPROACH

*Eduardo Kenedy**

RESUMO: Este artigo apresenta os resultados parciais de uma pesquisa experimental sobre as estruturas de topicalização do português brasileiro (PB). Discutem-se os dados de um experimento *self-paced reading* que testou a hipótese de que o PB seja uma língua de proeminência de tópicos (e não de sujeitos), como é característico de línguas orientadas para o discurso, conforme a hipótese de Pontes (1987) e Kato (2006), dentre outros. Argumentos teóricos e empíricos contra tal hipótese são também examinados.

PALAVRAS-CHAVE: topicalização, tópicos, sujeitos, psicolinguística experimental.

ABSTRACT: This paper presents some preliminary results of an experimental research on Brazilian Portuguese's (BP) topicalization structures. We discuss the results of a self-paced reading experiment that tested the hypothesis according to which BP must be analyzed as a discourse-oriented language, as it is assumed by Pontes (1987) and Kato (2006) among others. We also provide a brief review of both theoretical and empirical arguments against Pontes and Kato's hypothesis.

KEYWORDS: topicalization, topics, subjects, experimental psycholinguistics.

* Universidade Federal Fluminense (UFF), Niterói/RJ. Professor Adjunto de Linguística. Email: eduardokenedy@hotmail.com – www.professores.uff.br/eduardo.

TÓPICOS E SUJEITOS NO PB: UMA ABORDAGEM EXPERIMENTAL

Introdução

Neste artigo serão apresentados e discutidos os resultados preliminares de uma pesquisa experimental sobre o processamento cognitivo das estruturas sintáticas “tópico > comentário” e “sujeito > predicado” no português brasileiro (PB).¹ O recurso à experimentação é uma tentativa de testar empiricamente, em situações laboratoriais controladas, a hipótese de que o PB possa ser caracterizado como língua orientada para o discurso, com proeminência de tópicos, por oposição às línguas orientadas para a sintaxe, com proeminência de sujeitos, conforme sustentam diversos estudos produzidos ao longo dos últimos 25 anos na linguística brasileira (cf., dentre outros, PONTES, 1987; NEGRÃO, 1990; GALVES, 1998, 2001; ORSINI, 2003; KATO, 1989, 2006). Tal como será aqui indicado, esses estudos validam suas hipóteses com base (i) na intuição linguística de seus autores, recorrendo, assim, à clássica metodologia dos julgamentos de gramaticalidade cara à linguística gerativa, ou (ii) em dados extraídos de *corpora*, como é característico em pesquisas socioparamétricas e sociolinguísticas. Julgamos que pesquisas experimentais que indiquem como brasileiros per-

¹ O projeto de pesquisa “A topicalização na interface sintaxe e discurso: uma abordagem psicolinguística” recebe financiamento do CNPq (processo 474941/2010-8) e da Faperj (processo E-26/110.531/2010). A bolsista de Iniciação Científica, com bolsa Faperj, Carla Mota Regis de Carvalho participou ativamente da elaboração, aplicação e análise do experimento apresentado neste artigo.

cebem e processam as estruturas “tópico > comentário” e “sujeito > predicado” no curso *on-line* da computação cognitiva da informação linguística podem fornecer evidências complementares ou contraditórias a (i) e (ii). Assumimos que o desempenho linguístico de brasileiros registrado em experimentos de tempo de reação possa indicar a natureza da língua-I, no sentido de Chomsky (1986), subjacente a tal comportamento, ou seja, acreditamos que uma língua-I com proeminência de tópicos produza reações comportamentais assimétricas aos estímulos “tópico > comentário” e “sujeito > predicado”.

O artigo está organizado em três seções. Na primeira seção, apresentaremos as diferentes interpretações teóricas acerca da topicalização em PB. Analisaremos os argumentos formulados em favor da hipótese de que o PB seja uma língua de proeminência de tópico e aqueles formulados contra essa hipótese, segundo os quais a topicalização no PB não assume *status* diferente do que ocorre, por exemplo, nas demais línguas românicas. Na segunda seção, faremos a exposição de nosso plano experimental para o tratamento do assunto. Descreveremos como estamos conduzindo nossa pesquisa por meio de diferentes paradigmas de experimentação. Os resultados da pesquisa já realizada com o método *self-paced reading* serão apresentados e discutidos na terceira seção, após a qual indicaremos as próximas etapas de nosso trabalho.

1. A topicalização no PB

A topicalização é um fenômeno sintático-discursivo muito produtivo nas línguas naturais. Pode ser caracterizada como o expediente computacional por meio do qual um determinado constituinte – o tópico, que geralmente se encontra isolado do restante da frase por uma pausa entoacional – é posicionado à periferia esquerda de uma sentença, que sobre ele apresenta alguma espécie de comentário. A topicalização configura a estrutura frasal “tópico > comentário”, que se distingue da estrutura sintática “sujeito > predicado” por marcar, no tópico, informações prosódicas e discursivas ausentes no constituinte quando não topicalizado. Note-se, por exemplo, a saliência discursiva do objeto direto alvo da topicalização em (1) abaixo, por contraste a seu *status* informacional menos marcado em (2), em que se encontra em sua posição canônica no interior da sentença.²

² A estrutura “sujeito > predicado” pode coincidir com a veiculação de informação do tipo “tópico > comentário”. Para os efeitos deste artigo, interessam-nos apenas as estruturas de

(1) Objeto direto topicalizado

Aquele livro, ainda não tive tempo de ler.

(2) Objeto direto em posição canônica, não topicalizada

Ainda não tive tempo de ler *aquele livro*.

Li e Thompson (1976) caracterizaram a topicalização como um fenômeno sintático natural na linguagem humana, o qual apresenta, não obstante, diferentes níveis de profundidade e de produtividade entre as línguas. De acordo com os autores, há dentre as línguas aquelas orientadas para o discurso, tais como o chinês, que devem ser caracterizadas como línguas de proeminência de tópicos, já que sua configuração frasal básica é “tópico > comentário”. Por outro lado, línguas como o inglês apresentam a estrutura sintática “sujeito > predicado” como o modelo básico de frases, sendo, portanto, caracterizadas como línguas orientadas para a sentença, com proeminência de sujeitos. Nesses dois tipos de língua, a topicalização apresenta comportamento sintático diverso. Por exemplo, nas línguas de tópico não existem restrições para o constituinte que pode ser alvo da topicalização, ao passo que, nas línguas de sujeito, existe grande número dessas restrições. Nas línguas de sujeito, não há codificação morfológica especial para o item que sofre a topicalização, fenômeno existente nas línguas de tópico. Há, de fato, uma grande lista dos fenômenos linguísticos que separam as línguas orientadas para o discurso das orientadas para a sentença.

O português é tradicionalmente interpretado como uma língua orientada para a sentença, com proeminência de sujeitos, caracterizada pela ordem canônica “sujeito > verbo > objeto > adjuntos” (SVO) (cf., dentre outros, NEVES, 2000; MATEUS et al., 2003). Não obstante, a topicalização nessa língua apresenta-se como um fenômeno de extrema produtividade, sobretudo ao se considerar a língua oral e espontânea, de interação face a face em ambiente informal (VASCO, 1999, 2006). Segundo Tarallo (1990, p. 149), o português é uma língua SVO, mas manifesta grande flexibilidade de ordenação linear, conhecendo, diferentemente de outras línguas românicas, ordens não canônicas ou marcadas como OSV, VSO, VS etc., que resultam da topicalização. Essa flexibilidade na ordenação linear do português tem chamado a atenção dos estudiosos no curso dos últimos vinte e cinco anos. Em particular, o PB vem sendo interpretado como uma modalidade da língua

tópico que não estão sobrepostas à função sintática de sujeito.

portuguesa em que a topicalização ocorreria com maior produtividade. Diversos estudos, como os de Pontes (1987), Galves (1998, 2001), Decat (1989), Kato (1989, 2006), Callou, Moraes e Leite (1993), Vasco (1996, 2006) e Orsini (2003), têm encontrado, em diferentes *corpora* do PB, construções de tópico características de línguas orientadas para o discurso, assumidamente ausentes ou pouco produtivas em línguas com proeminência de sujeito, como é o caso do português europeu (PE). Tais evidências têm levado estudiosos brasileiros a levantar a hipótese de que o PB seja uma língua orientada para o discurso, com proeminência de tópico (PONTES, 1987; NEGRÃO, 1990; GALVES, 2001; ORSINI, 2003; KATO, 2006). De acordo com essa hipótese, a configuração frasal básica do PB seria “tópico > comentário”, por contraste à ordenação canônica “sujeito > predicado” do PE. Para Kato (2006), uma evidência a favor dessa hipótese é o fato de, em PB, ser possível realizar topicalizações recursivas sobre modificadores nominais, tal como se ilustra em (3), algo, segunda a autora, impossível no PE.

- (3) a. O pneu do carro da Maria furou-se.
- b. O carro da Maria furou o pneu.
- c. O carro da Maria, o pneu furou.
- d. A Maria, o carro furou o pneu.
- e. A Maria, o carro, o pneu furou.

Em (3a) temos a configuração básica “sujeito > predicado” à qual se seguem diversas possibilidades de topicalização, com sucessivas estruturas “tópico > comentário”. Em (3d) e (3e) ocorrem topicalizações radicais, que, segundo Kato (2006), são típicas de língua de proeminência de tópico. De acordo com a hipótese da autora, enquanto o PE conheceria apenas (3a) e ignoraria as demais construções, a situação do PB seria exatamente inversa: (3a) inexisteria no vernáculo brasileiro, que conheceria (3b-e).

Na interpretação do *status* da topicalização em PB, Duarte (1996) e Kenedy (2002) são vozes dissonantes. Para os autores, as topicalizações encontradas em certos *corpora* do PB são igualmente visíveis em dados do PE e também das demais línguas românicas. Defendem eles que os diferentes tipos de topicalização existentes em PB são, na verdade, um fenômeno característico do discurso oral e espontâneo em qualquer língua natural, diferentemente do que ocorre nas verdadeiras línguas de tópico, que, dentre outras coisas, possuem uma morfossintaxe especial para a topicalização. Para sustentar esse

argumento, Duarte (1996, p. 353) retirou de diversos *corpora* de fala casual do PE, do francês e do italiano inúmeros casos de topicalizações de modificadores nominais e de topicalizações recursivas idênticos aos encontrados em PB. Para a autora, a chave para distinguir línguas de tópicos e línguas de sujeito não deve ser a quantidade ou a profundidade da ocorrência de topicalizações em um dado *corpus*, mas, sim, o inventário das propriedades morfossintáticas da língua. Segundo Duarte (1996, p. 355), as línguas de tópico (i) são verbo-finais; (ii) codificam morfossintaticamente o tópico (mas não o sujeito); (iii) não dispõem de expletivos; (iv) não possuem ou possuem escassamente passivas; (v) possuem anáforas pronominais e nulas orientadas para o tópico (mas não para o sujeito); (vi) não fazem restrições gramaticais para o constituinte que pode figurar como tópico; (vii) dispõem de construções com tópicos múltiplos; e (viii) não apresentam morfologia verbal de concordância. Como o PB parece não possuir quase nenhuma dessas propriedades, a autora chega à conclusão de que interpretá-lo como língua de tópico é um equívoco.

Como se vê, a interpretação tipológica do PB com relação às estruturas de topicalização é ainda controversa. Isso ocorre, segundo entendemos, porque os estudos sobre o tema baseiam-se todos em argumentações puramente conceituais motivadas pela intuição do próprio linguista, como é o caso de Galves (2001) e Kenedy (2002), ou pela análise e interpretação de diferentes *corpora*, como é o caso de Pontes (1987) e Duarte (1996). Entendemos que estudos baseados na intuição do linguista ou em dados retirados de *corpus* podem ser usados infinitamente para sustentar ou rejeitar qualquer hipótese teórica sobre uma língua. Com efeito, intuições linguísticas são indefinidamente variáveis de indivíduo a indivíduo e só podem ser utilizadas como argumento sério quando submetidas a rigorosos controles de variáveis e a confiáveis testes estatísticos de variância. Análises de *corpora*, por sua vez, possuem limitado poder explanatório, uma vez que seus resultados podem ser fortemente enviesados pelo perfil sociocultural dos sujeitos falantes e/ou pelo gênero textual característico do *corpus* em análise. Isso quer dizer que *corpora* com perfis socioculturais e gêneros textuais muito diferentes normalmente apresentarão dados distintos, mesmo em relação a um único fenômeno gramatical como a topicalização.

Diante das limitações de análises intuitivas ou baseadas somente em *corpora*, Kenedy (2009) apontou a necessidade de abordar a problemática da topicalização no PB sob a perspectiva da psicolinguística experimental. Segundo o autor, experimentos formulados para capturar a realidade psicológica da topicalização poderão verificar se no processamento cognitivo da in-

formação linguística brasileiros apresentam reações diferentes às estruturas “tópico > comentário” e “sujeito > predicado”, manifestando preferência natural por uma delas. Assumimos que o recurso à pesquisa experimental poderá trazer importantes contribuições sobre a pesquisa tipológica do PB. A metodologia experimental permite (i) o controle objetivo das variáveis gramaticais atuantes no estímulo linguístico, (ii) o controle do perfil sociocultural dos sujeitos participantes do experimento, (iii) o controle do gênero textual utilizado nos estímulos e (iv) o tratamento estatístico dos dados obtidos a partir do comportamento dos sujeitos. Com a conjugação desses fatores controlados, torna-se possível formular generalizações mais confiáveis sobre a língua-I subjacente ao comportamento linguístico dos seres humanos.

2. Plano de pesquisa experimental

Nosso projeto de pesquisa experimental sobre a topicalização em PB, presentemente em desenvolvimento no Laboratório de Psicolinguística da Universidade Federal Fluminense (Lapsi/UFF), compreende diversos paradigmas de experimentação *on-line* e *off-line* úteis no estudo da compreensão da linguagem, tais como a leitura automonitorada (*self-paced reading*), a audição automonitorada (*self-paced listening*), a ativação e a reativação de itens (*priming*), o rastreamento ocular (*eye-tracking*) e o julgamento imediato de gramaticalidade (*speeded judgment task*).³ Essas técnicas são capazes de registrar, por exemplo, medidas como o tempo de reação a estímulos linguísticos, a aceitabilidade de estruturas sintáticas e o índice de erros/acertos em tarefas de compreensão elicitadas durante ou imediatamente após o curso do processamento cognitivo de estruturas linguísticas. Assumimos que testes dessa natureza possam contribuir na investigação sobre a tipologia sintática do PB na medida em que são capazes de verificar se as estruturas “tópico > comentário” e “sujeito > predicado” constituem diferentes objetos para o processamento da informação linguística em brasileiros. Em outras palavras, as medidas *on-line* e *off-line* obtidas nesses experimentos permitirão o cotejo de eventuais assimetrias no processamento de tópicos e de sujeitos, as quais servirão de base para generalizações acerca da competência linguísticas dos falantes do PB.

Parece natural assumir que as diferenças entre as línguas de tópico e as línguas de sujeito sejam psicologicamente reais, isto é, correspondam a diferentes expedientes cognitivos que derivam e representam, no interior

³ Para uma boa revisão sobre métodos em psicolinguística experimental, ver Mitchell (2004) e Maia e Finger (2005).

das mentes dos falantes/ouvintes, cada uma dessas estruturas frasais. Com essa assunção, torna-se possível testar experimentalmente a hipótese de que os falantes brasileiros sejam naturalmente predispostos à estrutura frasal “tópico > comentário”, por contraste à estrutura “sujeito > predicado” – algo que deve ser verdadeiro caso a hipótese defendida em, dentre outros, Pontes (1987) e Kato (2006) esteja correta.

Descreveremos a seguir os resultados de um dos experimentos correntemente em aplicação no Lapsi. Trata-se de um teste de reação, no paradigma *self-paced reading*, que confronta o processamento cognitivo de estímulos com tópicos e de estímulos com sujeitos.

3. Experimento *self-paced reading*

Nesse paradigma experimental, os sujeitos são levados a ler frases na tela de um computador. Cada frase do experimento é apresentada em segmentos, que podem ser constituídos de palavras ou de sintagmas e são dispostos na tela de maneira serial e não cumulativa. O tempo de exibição de cada segmento é registrado por um dispositivo interno ao computador. Detalhamos o paradigma *self-paced reading* na tabela a seguir.

1º)	O sujeito participante do experimento pressiona uma tecla no computador;
2º)	Com o acionamento da tecla, o primeiro segmento X de uma frase é exibido na tela;
3º)	O sujeito lê X conforme sua velocidade natural de leitura;
4º)	Ao concluir a leitura de X, o sujeito pressiona novamente uma tecla no computador;
5º)	Com esse novo acionar da tecla, X deixa de ser exibido e o segundo segmento Y é então trazido à tela;
6º)	O sujeito passa à leitura de Y e repete os procedimentos citados até que todos os segmentos da frase sejam lidos;
7º)	Após a leitura do último segmento da frase, o sujeito responde a uma pergunta de caráter interpretativo sobre o conteúdo do que acabou de ler;
8º)	Após a emissão de sua resposta, quase sempre uma opção entre “sim” ou “não”, o sujeito pressiona mais uma vez uma tecla no computador para dar início à exibição do primeiro segmento de uma nova frase;
9º)	Todo o procedimento citado é repetido até que todas as frases do experimento sejam lidas pelo sujeito.
10º)	Após a resposta à última pergunta do experimento, uma frase de agradecimento é apresentada na tela do computador e o teste é encerrado.

Tabela 1: O paradigma *self-paced reading*.

A lógica por detrás do *self-paced reading* é a de que o tempo de leitura dos segmentos exibidos na tela, medido pelo computador a partir do acionamento de uma tecla conforme o esquema acima, é fortemente influenciado pelo processamento cognitivo demandado pelo estímulo linguístico presente no segmento. Como os estímulos são apresentados de maneira serial e não cumulativa, isto é, como o segundo segmento só é exibido após o desaparecimento do primeiro, o sujeito deverá manter o primeiro segmento ativo em sua memória de trabalho e integrá-lo sintaticamente ao segundo segmento quando este for exibido, de modo a produzir mentalmente uma representação sintática coerente entre ambos.⁴ A maior ou a menor facilidade de integração sintática entre um segmento e outro poderá ser capturada pelos tempos relativos despendidos na leitura de cada segmento. Ou seja, um tempo de leitura maior no estímulo Y_1 em comparação a Y_2 deve ser considerado indício de que o processamento cognitivo de Y_1 é mais custoso do que o de Y_2 . Se Y_1 parece ser mais custoso do que Y_2 , então é possível dizer que a integração sintática de estímulos como Y_1 demanda maior carga de processamento cognitivo, o que será interpretado pelo pesquisador como evidência de que a computação mental de Y_1 e a de Y_2 são assimétricas na competência linguística dos sujeitos submetidos ao experimento. Ora, se fizermos com que, em um dado experimento, a integração de Y_1 corresponda à estrutura “tópico > comentário” e a integração de Y_2 corresponda à estrutura “sujeito > predicado”, poderemos então verificar se essas configurações frasais são computadas assimetricamente pelos indivíduos.

3.1 Design experimental

Criamos um desenho experimental de modo a apresentar um sintagma nominal (NP) como o primeiro segmento a ser lido pelo sujeito, o qual pode, a princípio, ser associado na sentença a uma posição de tópico ou a uma posição de sujeito. As posições hierárquicas de tópicos e de sujeitos assumidas na teoria sintática contemporânea são ilustradas simplificada-mente em (4).

⁴ Essa exposição concentra-se na representação mental correspondente à integração sintática de apenas dois segmentos. Naturalmente, as integrações demandadas no paradigma *self-paced* são proporcionais ao número de segmentos utilizados no experimento. Assim, um experimento composto por três segmentos demanda a manutenção do primeiro segmento na memória, a integração do segundo segmento ao primeiro, a manutenção do composto desses dois segmentos na memória e a integração desse composto com o terceiro segmento.

- (4) a. posição sintática do tópico [_{CP} TÓPICO [_{IP} [_{VP}]]]
 b. posição sintática do sujeito [_{CP} [_{IP} SUJEITO [_{VP}]]]

No experimento, a definição do NP como tópico ou como sujeito na sentença dá-se apenas na leitura do segundo segmento, que introduz o sintagma verbal (VP) da frase, conforme se ilustra em (5) abaixo.

- (5) [_{NP} 1º segmento / _{VP} 2º segmento / 3º segmento]
 a. Essa janela / **venta muito** / no verão.
 b. Essa janela / **fica aberta** / no verão.

Em (5a), o segmento crítico destacado em negrito define o NP “Essa janela” como tópico da sentença, enquanto em (5b) o segmento crítico seleciona esse NP como sujeito do verbo. Assumimos que, no processamento cognitivo desses estímulos, os sujeitos deverão atribuir a (5a) uma representação mental como a de (4a), ao passo que, diante de (5b), atribuirão a representação de (4b). O interessante é que a decisão pela representação (4a) ou (4b) só será possível durante a leitura do segundo segmento, que introduz o VP.

Na prática do experimento, o sujeito participante acionará uma tecla no computador e o primeiro segmento de uma frase, algo como o NP “Essa janela”, será exibido. Após a leitura desse segmento, o sujeito pressionará novamente uma tecla e, assim, o primeiro segmento desaparecerá da tela e dará lugar ao segundo, algo como o VP “venta muito” ou o VP “fica aberta”. Por fim, após a leitura do segundo segmento, o sujeito pressionará mais uma vez uma tecla e o segundo segmento dará lugar ao terceiro e último segmento. Conforme dissemos, essa técnica fará com que o sujeito tenha de integrar mentalmente o NP lido no primeiro segmento ao VP exibido no segundo segmento. VPs do tipo em (5a) levarão o sujeito a computar mentalmente a estrutura “tópico > comentário”, enquanto VPs do tipo em (5b) levarão o sujeito a computar mentalmente a estrutura “sujeito > predicado”.

3.2 Hipóteses

Conforme se assume em qualquer paradigma de tempo de reação, latências de leitura maiores no segmento crítico indicarão maior complexidade em seu processamento cognitivo. Assim, a hipótese de que o PB seja uma língua de proeminência de tópicos prevê que o processamento de (5b) demandará maior tempo de leitura em relação a (5a). Tal hipótese é coerente com

a ideia de que a representação em (4a) esteja automaticamente disponível como *default* na competência linguística de um indivíduo cuja língua-I seja orientada para o tópico. Inversamente, o processamento de (5b) envolve a quebra da expectativa pela estrutura (4a) e a consequente elaboração de outra representação mental, no caso, aquela indicada em (4b). A quebra da expectativa e a reestruturação sintática são os fatores cognitivos que, segundo a hipótese em favor do PB como língua de tópico, deverão provocar maiores médias na leitura do segmento crítico de (5b).

Pelo exposto, a hipótese de que o PB seja uma língua de proeminência de tópicos será confirmada no experimento caso os tempos médios de leitura de VPs que introduzem a estrutura “sujeito > predicado” sejam superiores aos tempos médios de leitura de VPs que introduzem a estrutura “tópico > comentário”. Caso não haja assimetrias entre os tempos de leitura dos dois tipos de VPs e/ou caso os tempos médios de leitura de VPs que introduzem a estrutura “sujeito > predicado” sejam inferiores aos de VPs que introduzem a estrutura “tópico > comentário”, então o experimento não confirmará a hipótese.

3.3 Variáveis e condições

A variável independente do experimento é o tipo de estrutura sintática da sentença, estabelecida na conjugação do NP do primeiro segmento com o VP do segundo segmento. Trata-se, portanto, de um experimento simples, com somente duas condições: (i) “tópico > comentário”, conforme (6a), e (ii) “sujeito > predicado”, conforme (6b).

(6) Exemplos das duas condições do experimento

a. condição “tópico > comentário”: Essa janela / **venta muito** / no verão.

b. condição “sujeito > predicado”: Essa janela / **fica aberta** / no verão.

A variável dependente é o tempo de leitura do segmento crítico. Para os propósitos deste artigo, as respostas apresentadas pelos sujeitos à pergunta interpretativa feita após a leitura de cada frase não serão consideradas, posto serem irrelevantes para a análise.

3.4 Sujeitos

Participaram do experimento sessenta sujeitos, divididos em três grupos conforme sua escolarização máxima: nível fundamental, médio ou supe-

rior. Os sujeitos de nível fundamental e médio foram selecionados aleatoriamente dentre estudantes da rede pública e privada de ensino do município de Niterói, região metropolitana do Rio de Janeiro, em diversas escolas do centro e da região sul da cidade. Os sujeitos com nível superior foram selecionados aleatoriamente dentre os recém-formados da UFF no campus do Gragoatá, em diversas habilitações – exceto Letras.⁵ A média de idade dos sujeitos foi 19 anos. Ao todo, trinta e dois sujeitos eram do sexo feminino e vinte e oito do sexo masculino.

3.5 Materiais

O experimento foi composto por vinte frases experimentais. Com esse número, foi possível que cada sujeito de cada nível testado fosse exposto dez vezes a cada condição do experimento. Adotou-se a distribuição *between subjects*, também conhecida como *quadrado latino*, de forma que todos os sujeitos fossem expostos a todas as condições sem que lessem as versões “tópico > comentário” e “sujeito > predicado” de um mesmo estímulo. Tal distribuição é importante para evitar que o sujeito participante tome consciência do tipo de estrutura linguística que lhe está sendo apresentada.

Além das vinte frases experimentais, quarenta frases distratoras foram adicionadas aos estímulos. Essas frases não possuem qualquer relação com a estrutura sujeito *versus* tópico e, assim, cumprem a função de distrair a atenção do sujeito, de modo a evitar a tomada de consciência das estruturas sob teste.

Cada VP em cada segmento crítico de cada condição experimental foi composto por duas palavras que perfazem um total de quatro ou cinco sílabas, de tal forma que assimetrias no desempenho dos sujeitos não possam ser devidas ao tempo de leitura de um estímulo mais extenso em relação a outro menos extenso. Os NPs usados como sujeitos e tópicos foram balanceados para o traço [animacidade]. Os verbos usados nos estímulos eram tais que selecionavam um NP sujeito referencial e, assim, eram usados na condição “sujeito > predicado”, ou eram tais que não selecionavam sujeito referencial, de modo a serem usados na condição “tópico > comentário”.

⁵ Acreditamos que linguistas, professores de português e estudantes de Letras poderão ter o seu desempenho linguístico alterado em experimentos psicolinguísticos em razão de sua maior consciência metalinguística e/ou de sua maior atenção à norma culta, razão pela qual não os selecionamos no presente experimento.

3.6 Procedimentos

Cada sujeito recebeu, individualmente, instruções para a realização do experimento, que eram apresentadas oralmente, na interação direta com o experimentador, e também por escrito, ao início da tarefa, na tela do computador. Todos participavam, também, de um pré-teste (treinamento), realizado diante do experimentador, em que ocorriam apenas frases distratoras e que possuía o mesmo *design* do experimento real. Esse pré-teste tinha intenção de confirmar o perfeito entendimento, por parte do sujeito, da tarefa a ser desempenhada. O experimento real só tinha início quando o sujeito demonstrasse ter compreendido completamente a tarefa a que se submeteria. Ao iniciar o experimento real, o sujeito encontrava-se sozinho, sem a interferência do experimentador ou de qualquer outro indivíduo, em uma sala com isolamento acústico necessário para a perfeita concentração na tarefa.

Ao pressionar uma barra branca destacada no teclado, surgia na tela do computador o primeiro dos três segmentos de cada frase a ser lida pelo sujeito. Com a conclusão da leitura de um segmento, o sujeito deveria pressionar novamente a barra branca para autorizar a apresentação do próximo e assim deveria proceder até que todos os segmentos tivessem sido lidos. Após a leitura do último segmento, uma pergunta interpretativa era apresentada na tela. O sujeito deveria pressionar, no teclado do computador, uma tecla verde, caso considerasse que a resposta para a pergunta fosse *sim*, e uma tecla vermelha caso considerasse *não* como a resposta correta.

O experimento foi elaborado e aplicado no *software* Psycope versão X – B46, suportado pelo Sistema Operacional X do computador PowerBook G4 (*laptop* da Apple, Macintosh), tela LCD de 15". As palavras foram apresentadas em fonte 25, na cor preta, com fundo branco. O relógio interno do computador registrou os tempos de leitura de cada segmento das frases-estímulo.

3.7 Resultados

No conjunto dos sessenta sujeitos participantes, os tempos médios de leitura da condição “tópico > comentário” foram superiores às latências da condição “sujeito > predicado”. Quando inserido em um contexto como [Essa janela / **venta muito** / no verão], o segmento crítico era lido em médios 2066 milissegundos, enquanto demandava 1591 milissegundos de leitura ao se encontrar na estrutura [Essa janela / **fica aberta** / no verão].

A diferença obtida, 475 milissegundos, quase meio segundo, foi considerada significativa no teste estatístico ($p < .05$).⁶

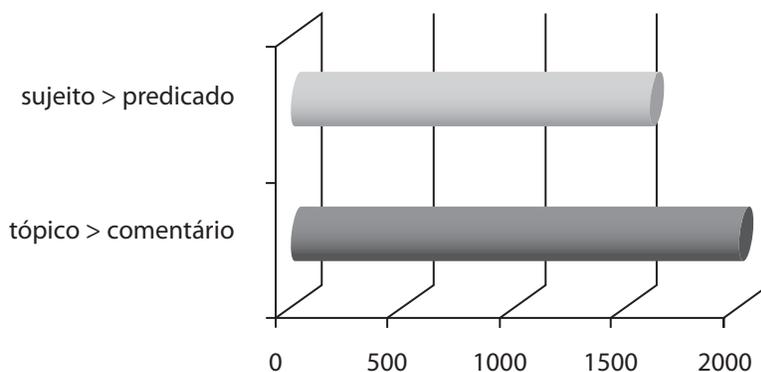


Gráfico 1: Tempos médios de leitura em cada condição experimental.
Tempos de reação em milésimos de segundo

Ao analisar os três grupos escolares em separado, a tendência no atraso da condição “tópico > comentário” permaneceu. No nível fundamental, as latências dessa condição chegaram a 2256 milissegundos, por contraste aos 1689 milissegundos da condição “sujeito > predicado” – diferença de 567 milissegundos. No nível médio, estruturas de tópicos demandaram 372 milissegundos a mais na leitura – 2021 para os VPs que selecionam tópicos e 1649 para os VPs que selecionam sujeitos. O mesmo ocorreu no nível superior, que registrou médias de 1855 milissegundos para os tópicos e 1364 para os sujeitos. Todas as diferenças são significativas ($p < .05$).⁷

⁶ Anova: [F1 (1,60) = 1,21; $p < .05$]. [F2 (1,5) = 20,36; $p < .05$].

⁷ Anova: Nível superior [F1 (1,20) = 12,63; $p < .05$], [F2 (1,5) = 179,66; $p < .05$]. Nível médio: [F1 (1,20) = 11,64; $p > .05$], [F2 (1,5) = 118,64; $p < .05$]. Nível fundamental: [F1 (1,20) = 9,84; $p < .05$], [F2 (1,5) = 137,22; $p < .05$].

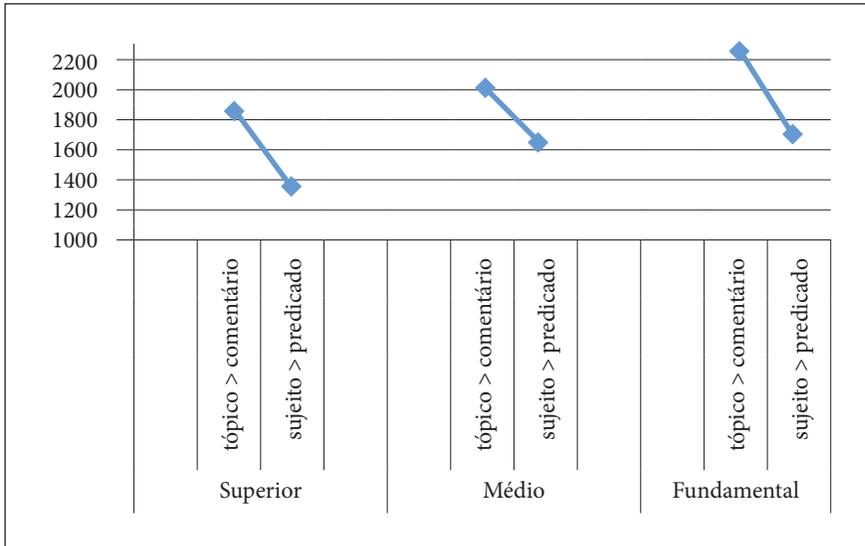


Gráfico 2: Dispersão dos tempos de reação por nível de escolarização.
Tempos em milésimos de segundos

Os sujeitos de escolarização mais alta despenderam menos tempos médios de leitura em ambas as condições. Interessantemente, as reações aos estímulos tornam-se mais rápidas conforme os sujeitos adquirem maior escolarização. Não obstante, a tendência de leituras mais lentas na condição “tópico > comentário” é constante nos três grupos de escolarização.

3.8 Discussão

Os resultados do experimento não confirmam a hipótese de que o PB seja uma língua de proeminência de tópicos. Com efeito, os tempos de leitura da condição “sujeito > predicado” foram significativamente menores em todos os grupos de sujeitos participantes, o que pode ser tomado como evidência de que a estrutura “sujeito > predicado” fazia parte da expectativa dos sujeitos participantes. É possível dizer que os resultados são perfeitamente inversos à hipótese do PB como língua de tópico.

A inversão da hipótese do experimento pode ser compreendida da seguinte maneira. Com um NP no *buffer* de sua memória de trabalho, os sujeitos do experimento criavam a expectativa de um VP que desse continuidade à estrutura “sujeito > predicado” que haviam começado a representar

mentalmente. Com a introdução de um VP coerente com essa expectativa, o processamento cognitivo do estímulo era facilitado, algo que se tornou visível com os tempos de leitura mais rápidos na condição “sujeito > predicado”. Já se um VP incoerente com tal expectativa é introduzido, então os sujeitos precisavam refazer sua representação mental, reanalisando o NP do primeiro segmento como tópico seguido de comentário. O custo cognitivo dessa reanálise é capturado pelas maiores latências na leitura da condição “tópico > predicado”.

Os resultados do experimento indicam, portanto, que os sujeitos participantes processaram perfeitamente a estrutura “tópico > comentário”, tanto quanto a estrutura “sujeito > predicado”, afinal ambas constituem derivações sintáticas perfeitamente gramaticais em PB. Não obstante, os resultados atestaram que era cognitivamente menos custoso para esses sujeitos processar estruturas “sujeito > predicado”, algo que não é esperado para uma língua-I orientada justamente para tópicos, com a configuração frasal típica “tópico > comentário”.

Para concluir

O experimento de compreensão de leitura aqui apresentado é apenas um dentre diversos outros tipos que estamos aplicando correntemente a respeito da problemática da topicalização no PB. Nosso próximo estudo apresentará uma versão oral do experimento descrito neste artigo: o *self-paced listening*. Por se tratar de *input* oral, tal experimento permitirá maior naturalidade na elaboração dos estímulos psicolinguísticos. Também realizaremos experimentos de aceitabilidade de diferentes tipos de topicalização, verificaremos assimetrias na orientação de anáforas de sujeito e de tópico e buscaremos sujeitos nativos do PE e de outras línguas classicamente descritas como orientadas para o sujeito ou para o tópico, de modo a formar um ponto de referências para análises dos resultados a serem obtidos com falantes do PB.

Por enquanto, nossos dados não podem confirmar ou refutar com segurança a hipótese de Pontes (1987) e de todos os estudos que a esse se seguiram. Generalizações mais confiáveis só se tornarão possíveis quando tivermos realizado um número significativo de experimentos com resultados conclusivos. Por agora, entretanto, é possível levantar a hipótese de que a topicalização do PB possa representar uma instância daquilo que Naro e Scherre (1993) e Varejão (2006) indicaram acerca da falta de concordância

verbo-nominal e Arim, Ramiro e Freita (2005) e Kenedy (2007) apontaram sobre as relativas cortadoras da língua portuguesa. Segundos esses linguistas, tais fenômenos provavelmente não são inovações do PB, mas, sim, usos regulares da língua entre a população iletrada, no caso da concordância, e entre qualquer falante em situação oral informal, no caso das cortadoras, tanto no Brasil quanto em Portugal. Seriam as topicalizações do PB também uma ilusão de ótica?

O PB é uma das raras línguas cujas entranhas foram despidoradamente devassadas em público. Poucas línguas europeias possuem tanta descrição sociolinguística em detalhes como o PB, fato que a nós linguistas brasileiros só nos causa orgulho. O preço por esse avanço, porém, pode ser certa exotificação da língua. É preciso verificar se o que se passa no PB representa de fato inovação em relação à sua origem românica ou se não se trata, na verdade, de apenas um retrato vivo da língua oral de uma comunidade subletrada em plena atividade em seu nicho mais natural: a interação comunicativa oral e face a face.

Referências Bibliográficas

ARIM, E.; RAMILO, M.; FREITAS, T. Mudança em curso e os média: o caso das relativas. In: MATEUS, M. do; NASCIMENTO, F. (Org.). *A língua portuguesa em mudança*. Lisboa: Caminho, 2005.

CALLOU, D.; MORAES, J.; LEITE, Y. A topicalização no português do Brasil: sintaxe e prosódia. In: II CONGRESSO DA ASSEL, 2009, Rio de Janeiro. *Anais do II Congresso da Assel*. Rio de Janeiro: Ed. UFRJ, 1993.

CHOMSKY, N. *Knowledge of Language: Its Nature, Origin and Use*. NY: Praeger, 1986.

DECAT, M. B. N. Construções de tópico em português: uma abordagem diacrônica à luz do encaixamento no sistema pronominal. In: TARALLO, Fernando (Org.). *Fotografias sociolinguísticas*. Campinas, SP: Pontes, 1989.

DUARTE, I. A topicalização no português europeu: uma análise comparativa. In: DUARTE, I.; LEIRIA, I. (Org.) *Actas do Congresso Internacional sobre o Português*. Lisboa: APL/Colibri, 1996.

GALVES, C. *Ensaio sobre as gramáticas do português*. Campinas: Editora da Unicamp, 2001.

GALVES, C. Tópicos, sujeitos, pronomes e concordância no português brasileiro. *Cadernos de estudos lingüísticos*, Unicamp, Campinas, n. 34, 1998.

KATO, M. Comparando o Português da América com o Português de Portugal e com outras línguas. Museu da Língua Portuguesa, 2006. Disponível em: <http://www.museudalinguaportuguesa.org.br/colunas_interna.php?id_coluna=13>.

KATO, M. A. Tópico e sujeito: duas categorias na sintaxe? *Cadernos de estudos lingüísticos*, Unicamp-IEL, Campinas, n.17, 1989.

KENEDY, E. Por uma pesquisa em “sintaxe experimental” sobre a topicalização no português do Brasil. In: XV CONGRESSO DA ASSEL, 2009, Rio de Janeiro. *Anais do XV Congresso da Assel – Linguagens em diálogo: pesquisa e ensino na área de Letras*. Rio de Janeiro: Ed. UFRJ, 2009.

KENEDY, E. *A hipótese da antinaturalidade de pied-piping em orações relativas*. Tese (Doutorado) – Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, 2007.

KENEDY, E. *Aspectos estruturais da relativização em português: uma análise baseada no modelo raising*. Dissertação (Mestrado) – Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, 2002.

LI, C. N.; THOMPSON, S. A. Subject and topic: a new typology of language. In: LI, C. N. (Org.). *Subject and topic*. New York: Academic Press Inc., 1976.

MAIA, M.; FINGER, I. *Processamento da Linguagem*. Pelotas: Educat, 2005.

MATEUS, M. M. et al. *Gramática da língua portuguesa*. Lisboa: Caminho, 2003.

MITCHELL, D. C. On-line Methods in Language Processing. In: CARREIRAS, M.; CLIFTON JR., C. (Ed.). *The on-line study of sentence comprehension: Eyetracking, ERPs and beyond* New York. NY: Psychology Press, 2004. p. 15-32.

NARO, A.; SCHERRE, M. Sobre as origens do português popular do Brasil. *D.E.L.T.A.*, v. 9, p. 437-54, 1993.

NEGRÃO, E. V. *O português brasileiro: uma língua voltada para o discurso*. Tese (Livre Docência) – Universidade de São Paulo (USP), São Paulo, 1990.

NEVES, M. H. M. *Gramática de usos do português*. São Paulo: Editora Unesp, 2000.

ORSINI, M. T. *As construções de tópico no português do Brasil: uma análise sintático-discursiva e prosódica*. Tese (Doutorado em Língua Portuguesa) – Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, 2003.

PONTES, E. *O tópico no português do Brasil*. Campinas: Pontes, 1987.

TARALLO, F. *Tempos lingüísticos: itinerário histórico da língua portuguesa*. São Paulo: Ática, 1990.

VAREJÃO, F. *Variação em estruturas de concordância verbal e em estratégias de relativização no português europeu popular*. Tese (Doutorado em Língua Portuguesa) – Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, 2006.

VASCO, S. L. *Construções de tópico na fala popular*. Tese (Doutorado em Língua Portuguesa) – Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, 2006.

VASCO, S. L. *Construções de tópico no português: as falas brasileira e portuguesa*. Dissertação (Mestrado) – Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, 1999.

Recebido em 18 de fevereiro de 2011

Aceito em 17 de junho de 2011