

EFEITO DA LACUNA PREENCHIDA E PLAUSIBILIDADE SEMÂNTICA NO PROCESSAMENTO DE FRASES EM PORTUGUÊS BRASILEIRO¹

Marcus Maia

RESUMO

Investiga-se o efeito da lacuna preenchida (ELP) em português brasileiro através de experimentos de rastreamento ocular e de leitura automonitorada. Os resultados detectam a presença do ELP e indicam que o *parser* atua de modo estritamente sintático em sua fase inicial. As medidas das fases finais integrativas dos dois experimentos são divergentes e motivam uma discussão sobre efeitos do tipo *good-enough*.

PALAVRAS-CHAVE: Construções QU; Efeito da Lacuna Preenchida; Rastreamento Ocular.

1. Introdução

O chamado Efeito da Lacuna Preenchida (doravante ELP) ou *Filled Gap Effect* costuma ser referenciado na literatura da especialidade da Psicolinguística conhecida como Processamento de Frases (*Sentence Processing*), principalmente com base no trabalho de Stowe (1986) que,

¹ Uma primeira versão da pesquisa reportada neste artigo foi apresentada no III Workshop de Processamento Anafórico (III WPA), que teve lugar na Universidade Federal Fluminense, em Niterói, entre os dias 21 e 23 de maio de 2013. O autor agradece aos participantes do evento pelos comentários recebidos e também à bolsista de Iniciação Científica (PIBIC/CNPq) Vanessa Simões, aluna da Faculdade de Letras da UFRJ, que atuou como assistente de pesquisa na implementação e aplicação dos experimentos.

baseando-se em estudo original de Crain & Fodor (1985), compara frases como 1a e 1b, abaixo. Em experimento de leitura automonitorada, identifica-se que o pronome *us* na posição de objeto direto de *bring* é lido com latências significativamente mais elevadas em 1a do que em 1b.

- 1a. My brother wanted to know who Ruth will bring **us** home to at Christmas.
- 1b. My brother wanted to know if Ruth will bring **us** home to Mom at Christmas.

Na construção em 1a, o Sintagma QU *who*, segundo análise tradicional na teoria gerativa (cf. CHOMSKY, 1977), teria se deslocado para posição não argumental na periferia esquerda da oração subordinada, a partir de sua posição argumental de base, onde recebe caso e papel temático, podendo ser interpretado. Note-se que, em 1b, por outro lado, o operador *if* não teria a mesma história derivacional, sendo gerado diretamente na posição em que se encontra e não por movimento. O fato de que o mesmo pronome *us* requer maior tempo médio de leitura na construção de extração 1a do que na construção sem movimento seria, segundo Stowe, evidência de que a expectativa da lacuna é dependente do contexto sintático². O maior tempo de processamento requerido na leitura de 1a teria sido causado pela surpresa do processador ao encontrar preenchida a primeira posição sintática onde o QU poderia ser interpretado e/ou pela necessidade de se revisar a primeira análise, comprometida com a postulação da lacuna como primeiro recurso.

Frazier (1987), Frazier & Flores d'Arcais (1989) e Clifton & Frazier (1989), entre outros trabalhos, analisam o ELP como um caso de um princípio mais abrangente, incluído no âmbito das estratégias de economia da Teoria do *Garden Path* (TGP), o *Active Filler Principle*, traduzido em Maia & Finger (2005) como Princípio do Antecedente Ativo, que é formulado, nos termos de Clifton & Frazier (1989), da seguinte forma: “quando um antecedente de uma categoria XP tiver sido identificado em uma posição não argumental, tal como COMP, priorize-se a possibilidade de atribuí-lo a uma

² This evidence suggests that the development of a gap expectation is heavily dependent on syntactic context. (STOWE, 1986, p. 244).

lacuna correspondente na frase, em vez da opção de identificar um sintagma lexical da categoria XP”³. Desde então, o princípio do Antecedente Ativo tem sido observado em várias línguas⁴, através de diferentes técnicas experimentais, tendo se tornado mesmo, como sugerem Phillips & Wagers (2007), uma espécie de *baseline* para o estudo de diferentes questões no processamento sintático.

A importante questão do acesso semântico no *parsing* sintático também tem sido tratada na literatura de processamento de frases de modo relacionado ao ELP. Trata-se de se determinar **se e quando** uma avaliação da plausibilidade semântica entre o antecedente e o verbo poderia interagir com a probabilidade de postulação da lacuna. Por exemplo, usando a técnica de rastreamento ocular, Traxler & Pickering (1996) observaram um decréscimo significativo nos tempos de leitura de frases como 2b relativamente a 2a, argumentando que a implausibilidade de se utilizar uma garagem, em comparação a uma pistola, como instrumento do verbo *to shoot* (“atirar”) seria detectada durante (ou imediatamente após) a leitura do verbo.

- 2a. That’s the pistol with which the heartless killer shot the hapless man...
2b. That’s the garage with which the heartless killer shot the hapless man...

Naturalmente, essa questão se projeta no quadro mais amplo da controvérsia entre os modelos de processamento ditos de “dois estágios” e os modelos mais interativos. De um lado, por exemplo, um modelo como a TGP (e.g., FRAZIER, 1979, 1987; RAYNER, CARLSON & FRAZIER, 1983), que adota uma arquitetura serial e modular, propõe que o *parser* não teria acesso, na fase inicial do processo de compreensão de frases, a informações semânticas, cuja computação se daria em uma fase posterior à análise sintática.

³ When a filler of category XP has been identified in a non-argument position, such as COMP, rank the option of assigning its corresponding gap to the sentence over the option of identifying a lexical phrase of category XP.

⁴ Além do inglês, há estudos em holandês (FRAZIER, 1987), italiano (DE VICENZI, 1991), húngaro (RADÓ, 1999), japonês (AOSHIMA, PHILLIPS & WEINBERG, 2004), russo (SEKERINA, 2003), entre outras línguas. Em português, registre-se o estudo de França (2005).

De outro lado, os chamados modelos de “satisfação de condições” (*constraint-based*) consideram viável entreterem-se múltiplas análises em paralelo, de modo que informações de natureza semântica, por exemplo, seriam capazes de ser acessadas a tempo de influenciar as decisões de análise do processador sintático (e.g., MACDONALD, PEARLMUTTER & SEIDENBERG, 1994; TRUESWELL, TANENHAUS & GARNSEY, 1994).

No presente artigo, investigam-se a atuação do ELP e o acesso semântico em seu processamento, na leitura de frases em português brasileiro, através de um experimento de rastreamento ocular e de um experimento de leitura automonitorada. No primeiro experimento, em um *design* simples, captura-se diretamente o ELP através da comparação dos padrões e dos tempos médios de fixação ocular nas áreas de interesse relevantes (medidas *on-line*) e também através do padrão de respostas às perguntas interpretativas, ao fim de cada frase lida (medida *off-line*). No segundo experimento, cruzam-se, em um *design* fatorial 2x2, as variáveis (i) preenchimento da lacuna (preenchida ou vazia) e (ii) plausibilidade semântica da relação entre o Sintagma QU extraído e o verbo (plausível ou implausível), colocando, portanto, em exame criterioso, além do ELP, também a questão dos tipos de informação disponíveis para o processador, no curso temporal da compreensão de frases. Os resultados das medidas *on-line* nos dois experimentos são bastante convergentes, no sentido de se estabelecer o ELP em português brasileiro, mas as medidas *off-line* divergem quanto à interferência do ELP, que é patente no experimento de rastreamento ocular e inexistente no experimento de leitura automonitorada. Como se verá, estas discrepâncias nos resultados *off-line* obtidos nos dois experimentos motivam ainda uma discussão no presente artigo, que é diretamente relevante para a chamada Hipótese *Good-Enough*. Essa importante questão da especificação e da profundidade do processamento foi explorada produtivamente por Christianson *et alii* (2001), Ferreira *et alii* (2002), estudos também replicados, em parte, através de testes de compreensão de frases equivalentes em português do Brasil por Ribeiro (2008), além de também discutidos em Maia (2013). Testando experimentalmente estruturas do tipo *Late Closure*, esses estudos concluem, de modo geral, que o processamento da linguagem muitas vezes é superficial (*shallow*), podendo resultar em representações semânticas imprecisas, incompatíveis com o valor de verdade do *input*.

O presente artigo se estrutura da seguinte forma: na seção 2.1, apresenta-se e discute-se o experimento de rastreamento ocular; na seção 2.2, apresenta-se e discute-se o experimento de leitura automonitorada; finalmente, na seção 3, apresentam-se as considerações finais do artigo, discutindo-se as possíveis razões para a discrepância nas medidas *off-line* dos dois experimentos.

2. Experimentos sobre o Efeito da Lacuna Preenchida

2.1. Experimento I – Rastreamento ocular

O experimento tem por objetivo verificar se o efeito da lacuna preenchida ocorre em PB, utilizando a técnica de rastreamento ocular. Estabeleceu-se uma única variável independente, o fator “preenchimento”, dividido em dois níveis, a saber, lacuna sem preenchimento (*Gap*) e lacuna preenchida (*Filled*). As condições experimentais foram, portanto, apenas duas, como exemplificado a seguir:

[G] Que livro o professor escreveu sem ler **a tese** antes?

O que o professor escreveu?

A) o livro

B) a tese

[F] Que livro o professor escreveu **a tese** sem ler antes?

O que o professor escreveu? A) o livro B) a tese

Quadro 1 – Exemplo das condições experimentais

Rastream-se as fixações oculares dos sujeitos, cuja tarefa consistia na leitura automonitorada de frases que apareciam completas em uma única linha na primeira tela, apresentando-se, na tela subsequente, uma questão interpretativa com duas opções de resposta, devendo-se responder fixando-se o olhar em uma delas. As variáveis dependentes foram estabelecidas como: (i) os tempos totais de fixação nas regiões de interesse, a saber, nas frases com a lacuna preenchida (F), o SN ocupando a posição de objeto direto do primeiro verbo e, nas frases sem a lacuna preenchida (G), o SN ocupando a posição

de objeto direto do segundo verbo (nos exemplos acima, o SN “a tese”); (ii) o padrão de fixação ocular nestas mesmas regiões críticas, nas duas condições, e (iii) os tempos totais de fixação nas alternativas de resposta às questões interpretativas finais. Caso o ELP fosse operativo em PB, esperava-se que o tempo de fixação total na região do SN preenchido da primeira posição possível de postulação da lacuna para o Sintagma QU movido (condição F) fosse significativamente mais elevado do que o tempo de fixação total do mesmo conjunto de SNs em posição de não preenchimento da primeira lacuna disponível (Condição G), e que o mesmo sintagma crítico demandasse mais fixações na condição F do que na condição G. Além destas duas medidas *on-line*, previa-se que, se o ELP persistisse ainda na fase interpretativa, na medida *off-line*, haveria mais erros na condição com a lacuna preenchida, indicados pelo padrão de duração de fixação total nos SNs de resposta da questão interpretativa. Como a questão incidia sempre sobre o SN crítico do Sintagma QU, esperava-se que, na condição G, em que a lacuna era encontrada como primeiro recurso, houvesse mais acertos do que na condição F, em que a lacuna estava preenchida.

2.1.1. Método

Participantes:

20 alunos do curso de Letras da UFRJ, com idade média de 21 anos e visão normal ou corrigida, participaram do estudo como voluntários, sendo 14 do sexo feminino.

Material:

Prepararam-se dez conjuntos de frases interrogativas QU, como o conjunto exemplificado na Figura 1, havendo-se distribuído sistematicamente as frases em duas versões, em quadrado latino, de modo que cada participante fosse exposto às duas condições, mas não às duas frases do mesmo conjunto. Acrescentou-se a cada versão o mesmo conjunto de 20 frases distrativas, apresentadas em randomização distinta a cada vez que o experimento era aplicado. Todas as frases, tanto as experimentais quanto as distrativas, eram apresentadas na tela do monitor acoplado ao rastreador ocular, em fonte Monaco 21, aparecendo em uma única linha. Na tela das questões interpretativas,

apresentava-se a pergunta em uma linha e as duas opções de resposta algumas linhas abaixo, situando-se a opção A à esquerda da tela e a opção B, à direita. O programa usado para apresentação, registro e análise dos estímulos foi o *TOBII Studio 2.3.2*.

Procedimento:

O experimento foi aplicado usando-se equipamento TOBII TX300, binocular, integrado a monitor de 23", na sala do LAPEX/UFRJ. Inicialmente, explicava-se a tarefa ao participante, solicitando-se que lesse cada frase rapidamente, automonitorando a sua passagem através da barra de espaço, no teclado do computador. Ao pressionar a barra de espaço, uma pergunta interpretativa, com duas opções de resposta, era chamada à tela, devendo o participante respondê-la, fixando o olhar por alguns segundos na opção que achasse correta. Como o sistema TOBII TX300 realiza correções de pequenos movimentos de cabeça, não se utilizou qualquer aparato de fixação da cabeça (*nasal clip e/ou chin rest*), o que permite maior naturalidade na leitura. Procedia-se, em seguida, à calibração de cada sujeito, que era sentado à distância de 60 a 65 cm da tela, devendo fixar o olhar e acompanhar o aparecimento e a movimentação de 12 pontos representados por círculos verdes. A calibração era repetida, caso não se obtivessem parâmetros aceitáveis, conforme indicado pelo programa. Em dois casos, substituíram-se sujeitos cuja calibração não obteve os níveis mínimos estabelecidos pelo programa. Após a fase da calibração, o sujeito era exposto a três frases de prática, sendo observado pelo experimentador, que podia indicar-lhe ajustes em relação à tarefa. Em seguida, o experimentador retirava-se da sala, deixando cada sujeito completar o experimento, que tinha duração média de 15 minutos.

2.1.2. Resultados

Registraram-se os tempos totais de fixação nas duas regiões de interesse, constituídas pelo mesmo conjunto de palavras das frases F e G, apresentadas em duas versões, em quadrado latino, obtendo-se os tempos médios de fixação indicados na Tabela 1 e ilustrados para uma das frases, no mapa de calor apresentado na Figura 1:

Condição G – SN objeto do segundo verbo	Condição F – SN objeto do primeiro verbo
0337ms	926ms

Tabela 1 – Tempos médios de leitura dos SNs críticos nas duas condições experimentais



Figura 1 – Mapa de calor da leitura da mesma frase nas duas condições por dez sujeitos

Uma análise de variância (ANOVA) por sujeitos indicou efeito principal altamente significativo do fator preenchimento $F(1,159) = 147$ $p < 0.000001$, o que também foi obtido em teste-t pareado [G]vs[F] $t(159)=12.12$ $p < 0.0001$.

Outra medida *on-line* obtida nesse experimento foi o número de fixações nos segmentos críticos. Na condição F, obtiveram-se, em média, 3 fixações no item preenchedor da lacuna, sendo, tipicamente, 1 progressiva, 1 regressiva e 1 no segundo passe da leitura da frase. Na condição G, obteve-se uma fixação no segmento crítico, na leitura da frase, sempre no primeiro passe, não se observando, em nenhum caso, fixações regressivas. O teste Chi-quadrado indica significância nessa diferença entre as duas condições experimentais ($X^2=20$, $p=0,0001^{***}$). O padrão de fixação obtido é, portanto, coerente com a medida do tempo total de fixação, identificando-se maior dificuldade de processamento, tanto pelo maior índice de fixações progressivas, quanto regressivas, no

caso da lacuna preenchida. Os *links* abaixo permitem visualizar animações de leitura real de frase de cada condição, identificando o padrão de fixação típico obtido no experimento:

https://drive.google.com/file/d/0Bx9dc1_fdJ1GSndTVUg0ZnFPUTA/edit?usp=sharing

https://docs.google.com/file/d/0Bx9dc1_fdJ1GcHdfeC0zLVYtUmM/edit

Figura 2 – Exemplos de leitura típica de frases com e sem a lacuna preenchida

Que livro o professor escreveu a tese sem ler antes?

Que livro o professor escreveu sem ler a tese antes?

Na medida *off-line*, constituída pelos tempos totais de fixação nas áreas das respostas às questões interpretativas, obtiveram-se os seguintes resultados, expressos no Gráfico 1 e na Tabela 2, além de também exemplificados em um caso específico, na Figura 3, abaixo.

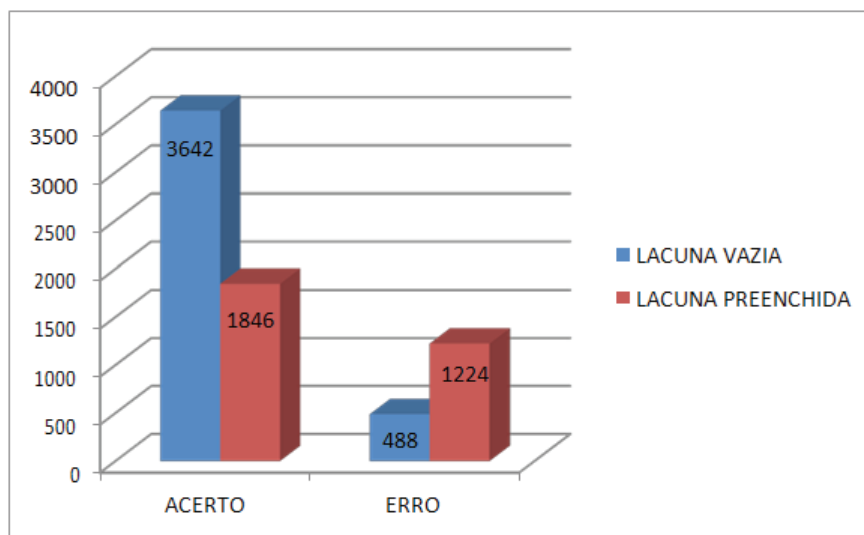


Gráfico 1 – Tempos Totais de Fixação nas respostas à questão interpretativa final

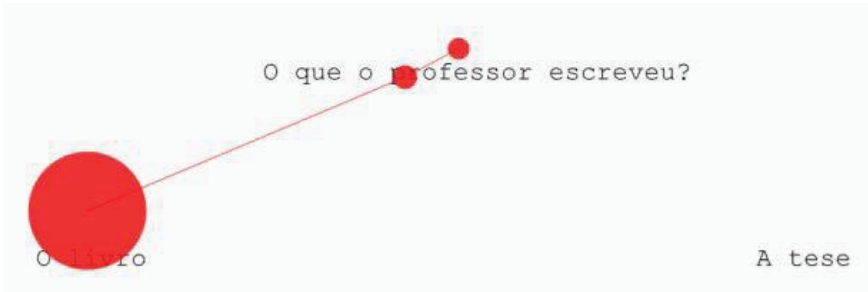


Figura 3 – Exemplo de padrão de fixação na resposta a uma pergunta interpretativa

Uma ANOVA bivariada por sujeito indicou efeito principal significativo do fator ACERTO ($F(1,19) = 280$ $p < 0.000001$) e efeito principal do fator PREENCHIMENTO ($F(1,19) = 28.8$ $p < 0.000036$), além de interação entre os dois fatores ($F(1,19) = 114$ $p < 0.000001$). Testes-t pareados indicam diferenças significativas entre as condições nas direções previstas pelas hipóteses, ou seja, o tempo total de fixação na resposta certa foi significativamente maior na condição G, com lacuna disponível (3642ms) do que na condição F, com lacuna preenchida (1846ms) ($t(19) = 8.55$ $p < 0.0001$). O tempo total de fixação na resposta incorreta foi significativamente maior na condição F, com lacuna preenchida (1224ms), do que na condição G com lacuna disponível (488ms), evidenciando que a lacuna preenchida provocou maior dificuldade de interpretação do que a lacuna disponível. Na condição com lacuna disponível, o tempo total de fixação na resposta correta (3642ms) foi significativamente maior do que o tempo total de fixação na resposta incorreta (488ms) ($t(19) = 16.91$ $p < 0.0001$). Embora na condição com lacuna preenchida o tempo total de fixação na resposta correta (1846ms) também tenha sido significativamente maior do que na resposta incorreta (1224ms) ($t(19) = 4.53$ $p < 0.0002$), observe-se que a diferença entre os tempos de fixação corretos e incorretos no caso da lacuna disponível é de 3154ms, enquanto que, no caso da lacuna preenchida, a diferença cai para 622ms, diferença igualmente significativa, confirmando maior facilidade na interpretação correta da condição com a lacuna disponível ($t(38) = 15,5$, $p = 0,0001$).

2.1.3. Discussão

Os resultados obtidos no estudo de rastreamento ocular permitem estabelecer a ocorrência do ELP em português brasileiro. Postulando a lacuna como primeiro recurso, o leitor, conseqüentemente, encontra maior dificuldade, conforme se infere tanto pela comparação dos tempos médios de fixação quanto pelo padrão de fixação ocular nos segmentos críticos, para processar a primeira posição pós-verbal quando a mesma está preenchida do que quando ela se encontra disponível. Os resultados da medida *off-line* indicam a persistência do ELP, mesmo na fase interpretativa. É importante notar que o registro do ELP nesta fase final, integrativa, do processo de compreensão de frases estaria a indicar novas evidências para a hipótese *Good-enough* (CHRISTIANSON *et alii*, 2001; FERREIRA *et alii*, 2002). Estudando frases cuja resolução se dá pelo Princípio *Late Closure*, esses autores identificaram que a análise inicial do *input* persiste na memória de trabalho do leitor e interfere na interpretação final da frase, resultando em representações que não refletem o valor de verdade do *input*, imprecisas, ou, no dizer de Ferreira *et alii*, simplesmente *good enough*.

2.2 Experimento II – Leitura automonitorada

Este experimento tem como objetivos: (i) confirmar através da metodologia de leitura automonitorada se o efeito da lacuna preenchida é, de fato, operativo em português brasileiro, e (ii) se a plausibilidade da relação semântica entre o sintagma QU e o verbo, de cuja posição de argumento interno o sintagma foi deslocado, interfere no efeito em medidas *on-line* e *off-line*. As variáveis independentes foram, portanto, as seguintes: (1) preenchimento da lacuna (*Filled* ou *Gap*) e (2) plausibilidade (**Plausível** ou **Implausível**). O cruzamento das variáveis gera um *design* 2x2, produzindo as condições experimentais exemplificadas abaixo, seguidas das suas respectivas perguntas interpretativas:

[GP] Que livro / o professor / escreveu / sem / ler / **a tese** / antes? O professor escreveu o livro?

[FP] Que livro / o professor / escreveu / **a tese** / sem / ler / antes? O professor escreveu a tese?

[GI] Que livro / o professor / **cozinhou** / sem / ler / **a tese** / antes? O professor cozinhou o livro?

[FI] Que livro / o professor / cozinhou / **a tese** / sem / ler / antes? O professor cozinhou a tese?

Quadro 2 – Exemplo de conjunto experimental

A tarefa experimental foi a leitura automonitorada seguida de questão interpretativa, havendo-se dividido cada frase em sete segmentos, conforme indicado pelas barras oblíquas nos exemplos acima. As variáveis dependentes foram três, a saber, uma medida *on-line*, constituída pelos tempos médios de leitura dos segmentos críticos (em negrito, nos exemplos), e duas medidas *off-line*, os índices e os tempos médios de respostas corretas às questões interpretativas⁵. As hipóteses explicitadas acima permitem as seguintes previsões de resultados: no que se refere à variável dependente *on-line*, se o ELP se instancia em PB, prevê-se que os tempos médios de leitura do segmento crítico na condição GP seja significativamente menor do que o segmento crítico na condição FP, em que a lacuna está preenchida. Caso o processador não acesse a informação sobre a plausibilidade semântica da relação verbo/objeto a tempo de influir na decisão do *parser*, seriam igualmente esperadas diferenças significativas entre os segmentos críticos em GI, cujo segmento crítico deveria exibir latências médias de leitura significativamente mais baixas do que sua contraparte em FI, em função do ELP, que não seria, portanto, afetado por razões relacionadas à plausibilidade semântica. Quanto às variáveis dependentes *off-line*, preveem-se para o percentual de respostas às questões interpretativas maiores índices de erros nos casos de preenchimento da lacuna e também nos casos de implausibilidade do que em suas contrapartes não preenchidas e plausíveis. Preveem-se ainda para os tempos de resposta às questões interpretativas maiores latências para as respostas corretas nas condições com lacuna preenchida e implausível do que para suas contrapartes com lacuna disponível e plausível.

⁵ Os tempos médios obtidos para as respostas incorretas às questões interpretativas não produziram diferenças significativas entre si ou em comparação com as respostas corretas e não serão reportados.

2.2.1. Método

Participantes:

Participaram do experimento, como voluntários, 28 alunos do curso de graduação em Fonoaudiologia da UFRJ, com idade média de 23 anos e visão normal ou corrigida, sendo 25 do sexo feminino.

Material:

Foram construídos 16 conjuntos de frases como o exemplificado no Quadro 2. Esses conjuntos incluíam os dez conjuntos testados no experimento 1, de rastreamento ocular, havendo-se produzido mais seis conjuntos para o presente teste. Distribuíram-se as frases experimentais em quadrado latino, constituindo-se quatro versões do experimento, cada uma composta, portanto, por todas as condições experimentais, apresentadas quatro vezes, sem repetição de material lexical. Em cada versão, acrescentou-se o conjunto de 32 frases distrativas que foram apresentadas em randomização com as frases experimentais, através do programa *Psyscope, build 57*. Nas condições em que a lacuna estava preenchida, o segmento crítico era, justamente, o preenchedor da lacuna, constituído por palavra de duas (tese, vinho, taco etc) ou de três sílabas (pipoca, dinheiro, mochila etc) no segmento 4, imediatamente posterior ao verbo. Nas condições em que a lacuna não estava preenchida, o mesmo conjunto de palavras ocorria sistematicamente na posição do segmento 6, imediatamente posterior a um segundo verbo na frase, protegido por um segmento final (segmento 7), que encerrava cada frase. Nas condições em que se manipulava a plausibilidade semântica, exploraram-se inadequações de natureza semântica ou pragmática entre o verbo e seu objeto, tais como a estranheza de se cozinhar um livro, ferver uma carta, comer uma roupa, etc.

Procedimento:

Os participantes foram testados individualmente na sala do LAPEX (Laboratório de Psicolinguística Experimental – UFRJ), em equipamento iMac CORE I5, 4GB RAM, 500GB HD, com tela de 21 polegadas, rodando o programa *Psyscope (build 57)*, no sistema operacional X (*Leopard*). O participante era orientado sobre a tarefa oralmente pelo experimentador, que lhe instruía a ler com atenção e rapidamente cada segmento de frase, pressionan-

do a barra de espaço do teclado até a sua conclusão, quando aparecia na tela uma pergunta em cor diferente (azul) da cor das frases (preta), devendo, nesse momento, o participante pressionar, rapidamente, a tecla identificada com a palavra SIM ou a tecla identificada com a palavra NÃO, de acordo com a resposta que julgasse mais adequada. Antes de iniciar o experimento em si, cada participante realizava uma sessão de prática com 4 frases, sendo observado pelo experimentador, que podia lhe dar indicações sobre o procedimento, garantindo a realização correta da tarefa. Em seguida, o experimentador deixava a sala e o experimento se iniciava. A duração de cada sessão experimental era de cerca de vinte minutos.

2.2.2. Resultados

Os tempos médios de leitura de cada um dos segmentos em que se dividiram as frases nas quatro condições experimentais são apresentados no Gráfico 2 e na Tabela 3, abaixo. Esses resultados foram submetidos a uma análise de variância por sujeitos, tomando-se como medidas repetidas o preenchimento da lacuna (*Filled x Gap*) e a plausibilidade (Plausível x Implausível), em um *design* fatorial 2x2. O resultado da ANOVA indicou um efeito principal altamente significativo do fator “preenchimento da lacuna” ($F(1,111) = 79.2$ $p < 0.000001$), mas não se observou efeito principal significativo do fator “plausibilidade” ($F(1,111) = 1.64$ $p < 0.202332$).

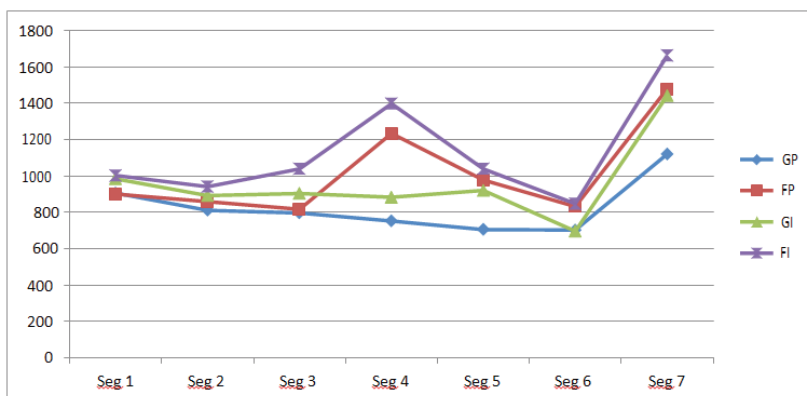


Gráfico 2 - Tempos médios de leitura dos segmentos em cada condição experimental

	Seg 1	Seg 2	Seg 3	Seg 4	Seg 5	Seg 6	Seg 7
GP	906	812	797	7053	706	703	1121
FP	901	860	817	1235	978	832	1479
GI	985	893	904	883	921	697	1442
FI	1003	942	1040	1398	1040	848	1663

Tabela 3 – Tempos médios de leitura dos segmentos em cada condição experimental

Também não se registrou interação entre os dois fatores ($F(1,111) = 2.31$ $p < 0.131059$). Foram também realizados testes-t pareados entre as condições relevantes, obtendo-se os seguintes resultados: cruzaram-se o tempo médio de leitura do segmento 4 na condição FP (1234ms) com o tempo médio de leitura do segmento 6 da condição GP (703ms), obtendo-se resultado altamente significativo, na direção esperada ($t(111)=6.70$ $p < 0.0001^{***}$), tendo em vista a latência mais elevada do segmento crítico da condição em que a lacuna estava preenchida (FP), em comparação com o menor tempo médio de leitura do mesmo conjunto de palavras em posição não problemática. Da mesma forma, observa-se o ELP na comparação entre os segmentos relevantes das condições implausíveis: o mesmo conjunto de palavras lido no segmento 4 crítico de FI em 1398ms apresentou tempos médios significativamente menores quando medido no segmento 6 de GI em que foi lido em 697ms, em posição em que não se preenchia lacuna sintática ($t(111)=7.25$ $p < 0.0001^{***}$). Por outro lado, não se observa diferença significativa entre os tempos médios de leitura dos segmentos relevantes das condições com lacuna preenchida plausível FP (1235ms) e com lacuna preenchida implausível FI (1398ms), conforme indicado no teste-t ($t(111)=1.54$ $p < 0.1272$). Da mesma forma, os tempos médios de leitura dos segmentos críticos das condições com lacuna também não diferem significativamente, seja em condição de plausibilidade (703ms), seja em condição de implausibilidade (697ms), segundo o teste-t ($t(111)=0.10$ $p < 0.9225$).

No que diz respeito às medidas *off-line*, aferiram-se os percentuais de acerto às perguntas interpretativas em cada condição, obtendo-se os resultados expressos no Gráfico 3. Observam-se índices de acerto significativamente mais

elevados para as condições plausíveis do que para as condições implausíveis, seja com lacuna preenchida ou não, indicando que o efeito de plausibilidade não capturado nas medidas *on-line* se fez sentir nesta fase integrativa final da compreensão. O teste Chi quadrado indica diferenças significativas entre os índices de acerto de GP e de GI ($X^2=4,2$, $p=0,04$) e entre os índices de acerto de FP *versus* FI ($X^2=6,1$, $p=0,01$), confirmando o efeito de plausibilidade sobre a interpretação das sentenças. Entretanto, não se registram diferenças significativas nem entre GP e FP ($X^2=0,04$, $p=0,83$) nem entre GI e FI ($X^2=0,05$, $p=0,81$), o que indica que, ao contrário do que se obteve nas medidas *on-line*, o efeito da lacuna preenchida já não se instancia na fase interpretativa, capturada na medida *off-line*.

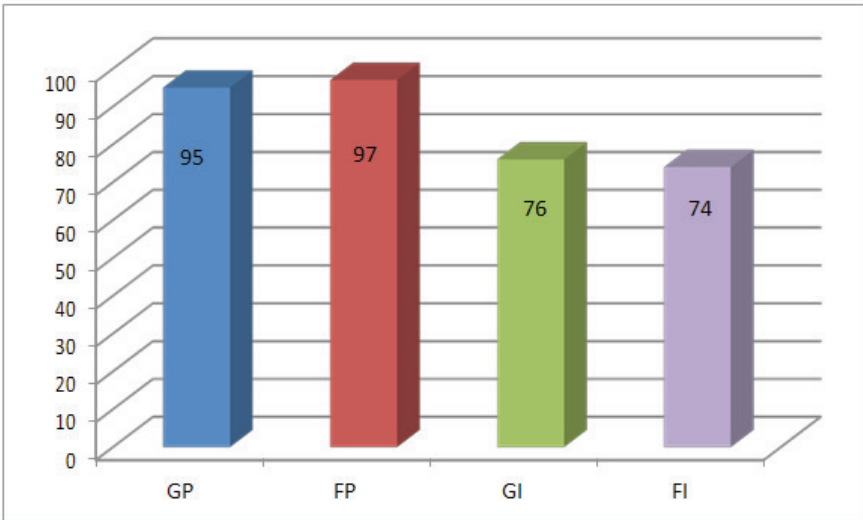


Gráfico 3 - Percentuais de acerto das questões interpretativas em cada condição

Aferiram-se também os tempos médios de resposta para as questões interpretativas assinaladas corretamente, obtendo-se os valores representados no Gráfico 4. Aplicou-se aos dados uma ANOVA bifatorial por sujeitos, não se observando efeito principal do fator PREENCHIMENTO da lacuna ($F(1,111) = 0.815$ $p<0.37$), mas se observando efeito principal significativo do fator PLAUSIBILIDADE ($F(1,111) = 9.12$ $p<0.003$). Não se registra, igualmente, interação significativa entre os dois fatores ($F(1,111) = 0.780$ $p<0.379$).

Testes-t pareados indicaram significância na comparação entre as condições plausível e implausível, seja com lacuna disponível ([GP x GI] $t(111)=1,19$, $p<0,03$), seja com lacuna preenchida ([FP x FI] $t(111)=3,07$, $p<0,0027$), mas não se obtiveram evidências para o ELP, seja comparando-se as condições plausíveis entre si ([GP x FP] $t(111)=0,01$, $p<0,99$), seja comparando-se as condições implausíveis ([GI x FI] $t(111)=0,96$, $p<0,33$).

Os resultados das duas medidas *off-line*, a saber, os percentuais e os tempos médios de acerto nas respostas às questões interpretativas foram, portanto, coerentes entre si, havendo ambos capturado efeitos de plausibilidade, mas não efeitos relacionados ao ELP.

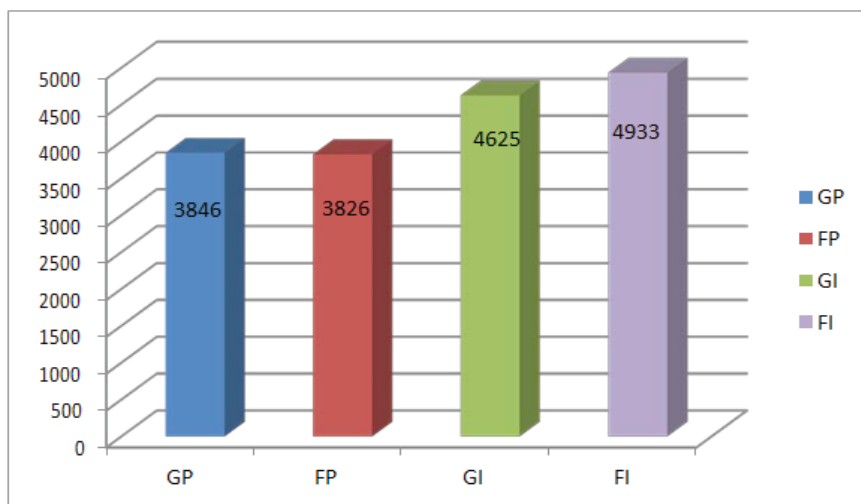


Gráfico 4 - Tempos médios de acerto das questões interpretativas em cada condição

2.2.3. Discussão

Por um lado, os resultados das medidas *on-line*, obtidos no experimento de leitura automonitorada reportado nesta seção, confirmam o ELP, já encontrado no Experimento 1. Não se observou, no entanto, nas medidas *on-line*, qualquer efeito de plausibilidade, que só é instanciado na fase interpretativa final, conforme indicado pelos resultados das medidas *off-line* apresentados na seção 2.2.2. Esse contraste entre os resultados *on-line* e *off-line* aduz evidências

em favor dos modelos mais estruturais do tipo *syntax-first*, tais como a TGP, em oposição às previsões de modelos mais interativos que preveem o acesso a informações de natureza semântica já nas fases iniciais do processamento. Por outro lado, neste segundo experimento, não se observou na medida *off-line* nenhuma interferência do ELP. Ou seja, embora este princípio seja claramente instanciado durante o processo de leitura, conforme indicado pelos resultados dos tempos médios de leitura dos segmentos críticos, na fase final, interpretativa, encontra-se evidência em favor de uma reanálise efetiva da disrupção sintática ocasionada durante a leitura pelo ELP. A interpretação, neste caso, não teria sido *good enough*, mas efetiva. Embora tenham entrado em *garden-path* provocado pelo ELP, conforme claramente evidenciado pelos resultados obtidos nas medidas *on-line*, os leitores se recuperaram da disrupção causada pelo ELP, apresentando padrão de respostas interpretativas não influenciadas pelo princípio estrutural. A questão que, naturalmente, se coloca nesse ponto é a de entender a discrepância nos resultados *off-line* obtidos nos dois experimentos: por que o experimento de rastreamento ocular indicou o ELP atuando na fase interpretativa e o experimento de leitura automonitorada não? Na seção final deste artigo, especula-se sobre as possíveis razões para esta discrepância.

3. Considerações finais

O estudo de rastreamento ocular e o de leitura automonitorada reportados no presente artigo convergem para um resultado comum, no que diz respeito ao ELP em português brasileiro, em sua aferição *on-line*. Em ambos os estudos, obtiveram-se latências médias significativamente mais elevadas nos casos do SN preenchendo a primeira lacuna disponível para a interpretação do Sintagma QU deslocado do que nos casos em que o SN era apresentado posteriormente à resolução do QU em lacuna disponível, fossem as regiões críticas apresentadas na frase completa na tela, sem segmentação (teste de rastreamento), fossem elas apresentadas em segmentação não cumulativa (teste de leitura automonitorada). Tais resultados confirmam achados estabelecidos desde Stowe (1986) para o inglês, além de vários estudos realizados desde então em diferentes línguas, que vêm permitindo concluir que, nos casos de extração de Sintagma QU, o processador sintático postula a lacuna para a interpretação adequada do sintagma, como primeiro recurso.

Deslocado para posição não argumental (A-barra), em que não é possível a satisfação de requisitos gramaticais tais como caso e papel temático, o sintagma torna-se “ativo” na busca pela primeira posição sintática que preencha as condições gramaticais adequadas para que possa ser interpretado. Ao encontrar tal posição preenchida por outro sintagma, o processador entra em *garden-path*, surpreendendo-se e precisando reanalisar sua decisão *default* que se revelou incorreta, de modo a chegar à análise sintática que permita, finalmente, interpretar a frase. Conforme demonstrado nos dois experimentos reportados no presente estudo, este efeito *garden-path* aparece cedo, no curso temporal do processamento dessas frases, sendo claramente indicado pelos tempos médios de leitura do SN crítico preenchido da lacuna, que são significativamente mais elevados do que nas condições *baseline*, em que o mesmo conjunto de SNs não ocupa a posição da lacuna como primeiro recurso. Por outro lado, conforme demonstrado pelos resultados do segundo experimento, ao contrário do efeito estrutural do ELP, capturado cedo no curso temporal do processamento, o efeito de plausibilidade semântica só é instanciado na fase integrativa final da compreensão, aduzindo evidências em favor dos modelos do tipo *syntax first*.

Como observamos na seção anterior, no entanto, a convergência obtida nas medidas *on-line* nos dois testes não se instancia de modo paralelo nas medidas *off-line*. No experimento de rastreamento ocular, a ANOVA que testou os tempos totais de fixação nas respostas à questão interpretativa final indicou efeito principal do fator PREENCHIMENTO, com teste-t pareado confirmando que o tempo total de fixação na resposta certa foi significativamente maior na condição G, com lacuna disponível do que na condição F, com lacuna preenchida.

Em contraste, as duas medidas *off-line* do experimento de leitura automonitorada, a saber, os índices e os tempos de resposta às questões interpretativas finais, indicaram apenas efeitos do fator PLAUSIBILIDADE, não se registrando, nesta fase final do teste, quaisquer efeitos do fator PREENCHIMENTO.

Essa falta de paralelismo nas medidas *off-line* dos dois testes exige algumas considerações, no sentido de se entenderem melhor essas diferenças para se estabelecer se a conclusão do primeiro teste de que o ELP também interfere na fase de interpretação das frases pode ainda ser mantida como con-

clusão do estudo ou se, de fato, a melhor conclusão seria a de que o ELP só se instancia durante o processamento *on-line*, mas não atua na fase integrativa final da compreensão, conforme se obteve no segundo teste.

Desde logo, queremos afirmar que a melhor conclusão nos parece ser a de que o ELP também produz efeitos na fase interpretativa, pós-*parsing*, da compreensão. Afinal, o princípio metodológico da evidência positiva garante esta conclusão: a ausência de evidência não deve ser tomada como evidência da ausência. Neste sentido, a existência de evidência positiva de que o ELP ocorre na fase final, obtida no primeiro experimento, não pode ser sobrepujada pela evidência negativa para o ELP, no segundo experimento.

De qualquer forma, impõe-se uma reflexão sobre as diferenças encontradas nas medidas finais, nos dois experimentos. Observe-se que, no primeiro experimento, a tarefa final proposta consistia de uma pergunta informacional QU, em que duas alternativas de SNs eram apresentadas, devendo-se responder fixando-se o olhar em um dos SNs, forçando-se, portanto, uma escolha entre a alternativa correta e a incorreta. Já no segundo experimento, a tarefa consistia na resposta a uma pergunta do tipo sim/não, que devia ser respondida, pressionando-se a tecla marcada com “sim” ou com “não”, no teclado do computador. Há portanto diferenças de método e de *input* nas duas tarefas. Nossa avaliação seria a de que a metodologia direta de captura do olhar, aliada à escolha forçada, impediu o erro do tipo II, ou seja, o chamado “falso negativo”, em que se aceita uma hipótese nula inválida, que não se pôde impedir no segundo experimento, com metodologia indireta, sem escolha forçada. Além disso, há uma terceira diferença importante entre as medidas *off-line* nos dois experimentos: apenas no experimento de leitura automonitorada se incluíram as condições de implausibilidade, tornando a tarefa, além de mais indireta e menos explícita, ainda mais complexa em termos de sua avaliação consciente, a despeito dos 2/3 de frases distrativas presentes em ambos os experimentos. Todos esses fatores reunidos podem ter impedido que o ELP fosse observado na resposta interpretativa no segundo experimento. Já no experimento de rastreamento, com método mais direto, escolha forçada, centrada apenas no fator PREENCHIMENTO, sem as condições de implausibilidade, o ELP persistiu na fase interpretativa, não tendo ocorrido a reanálise que se verificou no segundo experimento.

No primeiro experimento, ao chegar ao final da frase com lacuna preenchida, em que foi conduzido a *garden-path* e a reparo da estrutura por reanálise, o leitor, ainda ressentido pelo “tropeço” na leitura da frase, é colocado explicitamente diante de duas alternativas, com medida direta de aferição do olhar, sem a variável da PLAUSIBILIDADE. Sua escolha reflete, então, o *garden-path* que acaba de sofrer na leitura das frases com lacuna preenchida que são, então, interpretadas com menor índice de acerto do que na condição com a lacuna disponível, que não ocasionou *garden-path*. Neste caso, estaria ocorrendo o chamado efeito *good-enough*, conforme proposta de Christianson *et alii* (2001), também evidenciado em Ribeiro (2008) para o português com estruturas *Late Closure*. Seguindo o Princípio do Antecedente Ativo, o leitor postula a lacuna após o primeiro verbo da frase como primeiro recurso. Ao encontrar esta lacuna preenchida, entrando em *garden-path*, o leitor precisará reanalisar o Sintagma QU como tendo sido extraído da lacuna após o segundo verbo, estabelecendo a análise correta da frase. Quando, no entanto, perguntado, ao final da leitura, por exemplo, “o que o professor escreveu?”, devendo optar por uma de duas alternativas, os sujeitos apresentam índice de respostas indicando que sua análise inicial da estrutura, ainda que inadequada e corrigida, pode ser persistente, influenciando a interpretação final da frase, que seria, então, no dizer de Ferreira, simplesmente *good enough*, ou seja, apenas boa o bastante para alguns propósitos. Em outras palavras, a interpretação seria rasa ou subespecificada ou ainda, no dizer de Phillips *et alii* (2011), falível. Por outro lado, se nosso raciocínio está correto, a tarefa de leitura dessas mesmas frases, aferida através de metodologia indireta, não explícita, contendo condições extras de plausibilidade, teria criado condições para a reanálise adequada do ELP nesta fase final integrativa da compreensão.

Se verdadeira, esta análise oferece novas evidências para a questão discutida em Maia (2013), sobre a aferição do curso temporal da computação dos subprocessos gramaticais e a avaliação da sua profundidade, em duas dimensões do processamento, a saber, os métodos e o *input* linguístico.

Referências

- AOSHIMA, S.; PHILLIPS, C. & WEINBERG, A.S. Processing filler-gap dependencies in a head- final language. *Journal of Memory and Language* 51: 23-54, 2004.
- CHOMSKY, Noam. On Wh-Movement. In: CULICOVER, Peter; WASOW, Thomas, AKMAJIAN, Adrian (Eds.). *Formal Syntax*. New York: Academic Press, 1977. pp. 71-132.
- CLIFTON, C. & FRAZIER, L. Comprehending sentences with long-distance dependencies. In: TANENHAUS, M.K. & CARLSON, G. (Eds.), *Linguistic structure in language Processing*. Dordrecht: Kluwer Academic Press, 1989.
- CHRISTIANSON, K.; HOLLINGWORTH, A.; HALLIWELL, J.; FERREIRA, F. Thematic roles assigned along the garden-path linger. *Cognitive Psychology*, 42: 368-407, 2001.
- CRAIN, S. & FODOR, J.D. How can grammars help parsers? In: DOWTY, D.R.; KARTTUNEN, L. & ZWICKY, A.M. (Eds.). *Natural language parsing*. Cambridge: Cambridge University Press, 1985.
- DE VINCENZI, M. *Syntactic Parsing Strategies in Italian*. Dordrecht: Kluwer Academic Publishers, 1991.
- FERREIRA, F.; BAYLEY, Karl G. D.; FERRARO, V. Good-Enough representations in language comprehension. *Current Directions in Psychological Science*, 11:11-5, 2002.
- FRANÇA, A. I. Neurofisiologia da linguagem: aspectos micromodulares. In: MAIA, Marcus; FINGER, Ingrid. (Org.). *Processamento da Linguagem*, 1ed. Pelotas: Editora Educat, 2005, v. 1, p. 459-479.
- FRAZIER, L. On comprehending sentences: Syntactic parsing strategies. Tese de Doutorado. University of Connecticut (reproduzida por: Indiana University Linguistics Club), 1979.
- _____. Syntactic processing: Evidence from Dutch. *Natural Language and Linguistic Theory*, 5: 519-560, 1987.
- _____. FLORES D'ARCAIS GB. Filler-driven parsing: a study of gap filling in Dutch. *Journal of Memory of Language*, 28: 331-44, 1989.
- MACDONALD, M.C.; PEARLMUTTER, N.J. & SEIDENBERG, M.S.

The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101, 676-703, 1994.

MAIA, Marcus; FINGER, Ingrid (Orgs.). *Processamento da Linguagem*. 1. ed. Pelotas, RS: EDUCAT, 2005. 535p .

_____. Linguística Experimental: aferindo o curso temporal e a profundidade do processamento. *Revista de Estudos da Linguagem (UFMG)*, 21: 9-42, Belo Horizonte, 2013.

PHILLIPS, C. & WAGERS, M. Relating Structure and Time in Linguistics and Psycholinguistics. In GASKELL, G. (Ed.), *Oxford Handbook of Psycholinguistics*. Oxford University Press, 2007. pp. 739-756.

_____, _____, & LAU, E. F. Grammatical illusions and selective fallibility in real-time language comprehension. In: J. Runner (Ed.), *Experiments at the Interfaces, Syntax & Semantics*, 37: 153-186. Bingley, UK: Emerald Publications, 2011.

RADÓ, J. Discourse effects in gap-filling. Poster presented at AMLaP-99, University of Edinburgh, 1999.

RAYNER, K., CARLSON, M., & FRAZIER, L. The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior*, 22, 358-374, 1983.

RIBEIRO, Antonio J. C. A abordagem *Good-Enough* e o processamento de frases do português do Brasil. *Veredas (UFJF Online)*, 12 (2): 62-75, 2008.

SEKERINA, I. A. Scrambling and processing: dependencies, complexity, and constraints. In: S. Karimi (Ed.). *Word order and scrambling*. Malden, MA: Blackwell Publishing, 2003.

STOWE, Laurie. Parsing wh-constructions: evidence for on-line gap location. *Language and Cognitive Processes*, 1:227-463, 1986.

TRAXLER, MJ & PICKERING MJ. Plausibility and the processing of unbounded dependencies: An eyetracking study. *Journal of Memory and Language*, 35: 454-475, 1996.

TRUESWELL, J. C.; TANENHAUS, M. K. & GARNSEY, S.M. Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, 33:285-318, 1994.

FILLED GAP EFFECT AND SEMANTIC PLAUSIBILITY IN BRAZILIAN PORTUGUESE SENTENCE PROCESSING

ABSTRACT

The Filled Gap Effect (FGE) is investigated in Brazilian Portuguese through eye-tracking and self-paced reading experiments. Results detect the presence of FGE, indicating that the parser is strictly syntactic in the early stage of processing. The final measures in the two experiments present discrepant results, motivating a discussion on possible *good-enough* effects.

KEY WORDS: Language Processing; Eye-tracking; WH-constructions; Filled Gap Effect; Good-Enough Hypothesis.

Recebido em: 06/03/2014

Aprovado em: 25/07/2014