



A neurofisiologia do acesso lexical: palavras em português

Aniela Improta França (UFRJ)^a

Miriam Lemle (UFRJ)^b

Aline da Rocha Gesualdi (CEFET-RJ)^c

Mauricio Cagy (UFF)^d

Antonio Fernando Catelli Infantosi (UFRJ)^e

RESUMO: Este é um estudo de acesso lexical estimulado por *priming* e aferido por testes de extração de ERP (potenciais relacionados a eventos) em que foram comparados 60 pares de *prime* e alvo com semelhança fonológica (batata-barata) e 60 com semelhança morfológica (centro centralização) divididos em séries de palavras de diferentes tamanhos. Constatamos que o acesso a palavras pequenas é mais rápido do que a grandes em pares com semelhança fonológica. Porém a diferença de tamanho linear das palavras não cria uma gradação quando o *prime* e alvo compartilham a primeira instância de categorização: o acesso a *pureza* depois de *puro* se dá na mesma janela temporal do que o de *centralização* depois de *centro*, e estas duas são mais rápidas do que *barata* depois de *batata*. Ou seja, a morfologia é fator de aceleração no reconhecimento de alvos.

Palavras-chave: Sintaxe interna à palavra; Ponto de arbitrariedade saussuriana; Composicionalidade semântica; Computações em Forma Lógica; Eletrofisiologia da linguagem.

Introdução

Uma das operações lingüísticas inevitáveis é a do acesso lexical, que nos permite, com enorme facilidade e rapidez, entender e/ou produzir palavras soltas. Porém, por trás da aparente simplicidade da relação entre forma e significado no âmbito da palavra, há mecanismos que estão longe de serem completamente entendidos e que vêm gerando muita discussão na literatura lingüística dos últimos anos (FRANÇA, 2005, 2007; HAY, BAAYEN 2005; STOCKALL, MARANTZ 2006; STOCKALL *et al.* 2004; PYLKKÄNEN *et al.* 2004, 2003A; PHILLIPS *et al.* 2000; PINKER, 1999).

A controvérsia diz respeito à própria natureza da entidade acessada. Uma palavra como *globalização* é acessada inteira (HAY, BAAYEN 2005)? Ou analisamos suas unidades menores (morfemas): *glob-*, *-al*, *-iza*, *-ção* (STOCKALL, MARANTZ 2006)? E as unidades complexas, como *papo de anjo* ou *dar no pé*? São acessadas juntas ou separadas? (FIORENTINO, POEPEL, 2007; PINKER, 1999).

^a Programa de Pós-Graduação em Lingüística, UFRJ; anielaimprota@terra.com.br

^b Programa de Pós-Graduação em Lingüística, UFRJ; miriamlemle@gmail.com

^c Departamento de Engenharia Elétrica, CEFET-RJ; alinegesualdi@gmail.com

^d Departamento de Epidemiologia e Bioestatística, UFF; mauricio.cagy@gmail.com

^e Programa de Engenharia Biomédica, COPPE/UFRJ; afci@peb.ufrj.br

Há diferenças notáveis entre itens lexicais simples e complexos. O significado de palavras simples como *globo* deriva da concatenação da parte arbitrária do significado (raiz) com um morfema categorizador - um nominalizador, no caso de *globo*, enquanto palavras complexas como *globalização* trazem camadas morfológicas que podem ser interpretadas composicionalmente a partir do significado negociado em *globo*. No Modelo de Gramática que assumimos aqui, Morfologia Distribuída (MD- HALLE, MARANTZ 1993, 1994), existe a previsão de que haja uma diferença derivacional entre o acesso semântico à parte arbitrária e aquele que envolve uma parte composicional. Assim, o significado de *globo* é fixado por convenção, num módulo denominado Enciclopédia, mas os de *global*, *globalizar* e *globalização* seriam montados composicionalmente e calculados fase a fase na sintaxe, de acordo com a sucessão de concatenações de sufixos. Este processo é exatamente igual àquele com que se calcula o significado de sentenças concatenando os significados de suas partes componentes. Será possível capturar experimentalmente a dissociação do acesso lexical, flagrando a distinção entre a parte enciclopédica e a parte sintática do significado de palavras? Este é o principal objetivo deste trabalho.

Nos últimos cinco anos, uma série de experimentos (PYLKÄNNEN *et al.*, 2000; EMBICK *et al.*, 2001; PYLKÄNNEN *et al.*, 2002; PYLKÄNNEN *et al.*, 2003a; PYLKÄNNEN *et al.*, 2003b) propiciaram avanços consideráveis no campo do acesso lexical, utilizando o protocolo de decisão lexical. Em Pylkkänen *et al.* (2003b), que de certa forma sistematiza os achados dos outros estudos do grupo, os 2 estímulos foram estruturados através do protocolo de *priming*¹. Foram testados pares de palavras com semelhança fonológica pelo início, como *spin-spinach*, com semelhança fonológica pelo meio, como *teacher-reach*, e com semelhança morfológica, como *teacher-teach*. Depois da apresentação do *prime*, o acesso à palavra alvo era monitorado em duas dimensões: na ativação cortical, através de magnetoencefalógrafo (MEG), e na reação comportamental, a partir do aperto de botão pelo sujeito, indicando se era uma palavra ou uma não palavra.

A observação cuidadosa dos tempos de ativação cortical e reação comportamental aos alvos confirma a existência de dois mecanismos de pareamento fonológico: um para os fonemas que começam a palavra e outro para os fonemas que se localizam no meio da palavra. *Reach* do par *teacher-reach* é mais rápido do que *tea* do par *teacher-tea*, em termos de ativação cortical.

¹ O paradigma de *priming* no acesso lexical é um teste para desvendar aspectos da arquitetura do léxico mental, por exemplo, os critérios de agrupamento. Agrupam-se palavras inteiras ou fatias morfológicas? Por semelhança fonológica ou semântica? Para isso, os experimentadores constroem estímulos de muitos pares de palavras. Na metade dos casos, a primeira palavra do par, tecnicamente chamada de *prime*, terá algum relacionamento com a segunda palavra do par, tecnicamente chamada de *alvo*; por exemplo, no par *caro-carinho* há uma relação de semelhança fonológica entre *prime* e alvo. Nenhum relacionamento existirá entre *prime* e alvo nos pares restantes, por exemplo *caro-frio*. Para que possamos conhecer a reação do voluntário aos alvos temos que incumbi-lo de uma tarefa, como discriminar se o alvo é uma palavra ou uma não palavra. Com esse propósito, mesclamos aleatoriamente os pares de palavra-palavra com um número igual de pares de palavra-não palavra. Analisam-se o tempo e acuidade de resposta, podendo-se acessar indiretamente a influência que o *prime* exerceu em relação ao alvo. Os tipos de influências mais testados na literatura são semântica (*primo-tio*), fonológica (*cara-carinho*), morfológica (*cabeça-cabeçudo*), ortográfica (*pedra-vidro*) e sintática (*fazia-tinha*).

Ao ouvir *teacher* [ti...], todas as representações estocadas na mente com este começo, como *tea*, *teen*, *team*, *teach* e obviamente, *teacher* são ativadas (ativação múltipla). Mas à medida que o pareamento continua, no ponto [...tʃ...], *tea* não pode mais ser pareado com o *input* e é então suprimido.

Esta supressão imediata leva *tea* a um nível mais baixo de ativação do que o nível das palavras em repouso na mente, ou seja, *tea* se torna ativamente desativado. Assim, quando em seguida se apresenta *tea* como estímulo-alvo, para que o voluntário consiga acessar a representação da palavra, o voluntário terá que reativá-lo a partir de um ponto negativo. Por isso esta ativação é mais lenta do que a ativação das representações que se assemelham ao meio da palavra *prime*.

Vejamos como isso se dá, analisando o acesso lexical a *reach* do par *teacher-reach*. Durante o escaneamento de *teacher*, à medida que mais informações do meio da palavra [...itʃr] chegam ao cérebro, através do escaneamento contínuo, palavras que têm estes fonemas mediais começam a ser ativadas. Note que estas ativações não se originaram de uma rotina de pareamento pelo início, portanto não serão inibidas. Por exemplo, *each* e *reach*, são ativados porque eles têm fonemas que podem ser pareados com a parte medial da palavra *teacher*. Porém, já que o ponto de pareamento medial não proporciona uma visão retroativa do início da palavra-alvo já escaneada, estas palavras com pareamentos não iniciais, mas com semelhança interna, continuam no páreo com níveis diferentes de ativação e competem por reconhecimento. Isto acontece por uma questão de segurança do sistema de reconhecimento, já que é comum termos que entender palavras mal pronunciadas, cujos inícios são pouco audíveis. Assim, quando uma dessas palavras que permaneceram por algum tempo na competição aparece como alvo, o início da ativação de sua representação se dá mais rapidamente do que *tea* no par *teacher-tea*.

Contrastivamente, se do ponto de vista da ativação lexical, o alvo *reach* é ativado mais rapidamente, em termos comportamentais ocorre o oposto. Como não há inibição, a fase final do processo, onde se dá a escolha da representação que é igual ao estímulo, fica mais lenta porque a palavra com a ativação mais alta tem que ser *avistada* entre muitos competidores que não foram suprimidos, o que torna a tarefa de escolha mais difícil, atrasando o *aperto do botão pelo voluntário*, que é o dado acessível para as atuais metodologias de Psicolingüística. Por isso, *reach* em *teacher-reach* provoca uma reação comportamental mais lenta do que a reação a *tea* em *teacher-tea*. Este descompasso entre achados da Psicolingüística demonstra que a avaliação neurofisiológica pode *desempacotar* ações cognitivas que não poderiam ser apreciadas somente a partir dos tempos de reação.

Outro fato de grande interesse que aflorou na pesquisa é que *teach* no par *teacher-teach* é ativado mais rapidamente do que *tea* no par *teacher-tea* e mais até do que *reach* no par *teacher-reach*. Ou seja, *teacher* influencia fortemente a ativação de *teach*. Note que este achado é paradoxal, pois, de acordo com as expectativas, o fato de *teacher-teach* terem inícios semelhantes deveria atrapalhar a ativação de *teach*. Logo após a parte [...tʃ...], *teach* deveria ter o mesmo destino de *tea*: seria inibido como representação do *prime* e por conseguinte demoraria para ser ativado como alvo. Porém, o tempo de ativação de *teach* está nos níveis encontrados quando o *prime* e o alvo são a mesma palavra, isto é, em níveis de facilitação máxima.

A razão deste aparente contra-senso está no fato de a relação entre *tea* e *teacher* ser fonológica, enquanto entre *teach* e *teacher* existe uma raiz em comum que se junta a um verbalizador presente tanto em *teach* como no interior da palavra *teacher*. Esta relação morfológica torna operante o efeito positivo do *priming*, e o acesso lexical se dá mais rapidamente. Com este teste, os autores estabeleceram parâmetros para discriminação entre relações de *semelhança* e relações de *identidade*, e empregam o termo *semelhança* quando ela é fonológica e *identidade* para aquela que é exclusivamente morfológica. Também foi possível captar com maior detalhe as etapas internas do processo de acesso no teste neurofisiológico em relação ao comportamental, que traz empacotado no tempo de resposta uma série de computações que só puderam ser destrinchadas através do acesso direto à eletrofisiologia do acesso lexical.

No presente teste, usamos palavras do português e analisamos a eletrofisiologia do acesso lexical, enfocando comparações entre pares com relacionamento fonológico pelo início, organizados em um Grupo Fonológico (GF), e pares com relação morfológica, organizados em um Grupo Morfológico (GM). Enquanto os pares no GF compartilhavam segmentos fonológicos iniciais, os pares no GM compartilhavam a primeira concatenação [Raiz + primeiro morfema categorizador]. Escolhemos trabalhar só com estes dois tipos de pares, fonológicos com relacionamento pelo início e morfológicos com compartilhamento de raiz, pois ambos, tendo início semelhante, poderiam oferecer uma comparação direta entre as computações que envolvem processos fonológicos e morfológicos. Contudo, a maior contribuição do nosso desenho foi termos elaborado condições contendo palavras de três tamanhos com semelhança fonológica e palavras com duas, três ou quatro camadas morfológicas após a camada saussuriana. Este desenho propiciou um avanço em relação aos testes precursores (PYLKÄNNEN *et al.*, 2000; EMBICK *et al.*, 2001; PYLKÄNNEN *et al.*, 2002; PYLKÄNNEN *et al.*, 2003A; PYLKÄNNEN *et al.*, 2003B) na medida em que, ao utilizarmos palavras com diferentes tamanhos fonológicos e diferentes profundidades morfológicas, pudemos variar o ponto da arbitrariedade saussuriana (AS) no GF e mantê-lo igual no GM, como se pode ver no Quadro 1:

Ponto de Arbitrariedade Saussureana

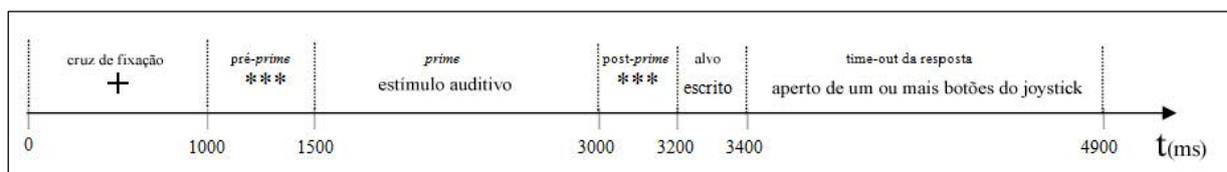
GF	F5/6 segmentos	barat	a	[raiz + n]
	F7/8 segmentos	problem	a	[raiz + n]
	F7/8 segmentos	jagatiric	a	[raiz + n]
GM	M2 (2 camadas)	pur	eza	[[raiz + a] + n]
	M3 (3 camadas)	arm	amento	[[[raiz + n] + v] + n]
	M4 (4 camadas)	inici	alização	[[[[raiz + n] + adj] + v] + n]

Quadro 1: Diferentes pontos de arbitrariedade saussuriana nos alvos dos subgrupos de GF e GM

Notem que no nosso desenho experimental há uma correspondência em termos de número de fonemas entre as gradações estabelecidas para os alvos fonológicos e morfológicos. Os alvos nos pares fonológicos com menor tamanho, ou seja, aqueles com cinco e seis segmentos (*barata*), equivalem em tamanho aos menores alvos do grupo morfológico com duas camadas (*pureza*). Da mesma forma, alvos com tamanho de sete e oito segmentos (*problema*) equivalem aos menores alvos morfológicos com três camadas (*armamento*). E, finalmente, aqueles com nove a dez segmentos (*jaguarica*) equivalem aos menores alvos morfológicos com quatro camadas (*globalizado*). No caso dos pares morfológicos, entretanto, contamos o número de camadas e não de segmentos. Sendo assim, havia palavras do grupo morfológico com número de segmentos bastante superior ao dos alvos fonológicos correspondentes (*jaguarica*, do GF com dez segmentos, e *inicialização*, do GM com doze segmentos). Ou seja, nos subgrupos morfológicos com duas, três e quatro camadas, havia sempre palavras do mesmo tamanho ou maiores do que aquelas nos subgrupos fonológicos correspondentes. Quando havia discrepância de tamanho, a palavra com maior número de segmentos era sempre a do GM.

2 Metodologia

Aplicamos um teste de *priming* transmodal (*prime* auditivo / alvo visual) que proponha uma tarefa de decisão lexical do tipo palavra ou não-palavra após a leitura do alvo. O Quadro 2 mostra o cenário do experimento com eventos alinhados à linha do tempo, como foi construído em plataforma *Presentation 12.0* da *Neurobehavioral Systems Trial Version*.



Quadro 2: Linha de tempo com eventos do experimento em milésimos de segundos

Havia 150 pares de palavras e 150 de não palavras, distribuídos pelas condições, que serão descritas a seguir, automaticamente aleatorizados no início do teste de cada sujeito, com a manutenção da integridade dos pares.

2.1. As séries de estímulos

No Gráfico 1 podemos ver a distribuição dos 300 estímulos por grupo. O número de itens pode ser visto em cada fatia do gráfico pizza.

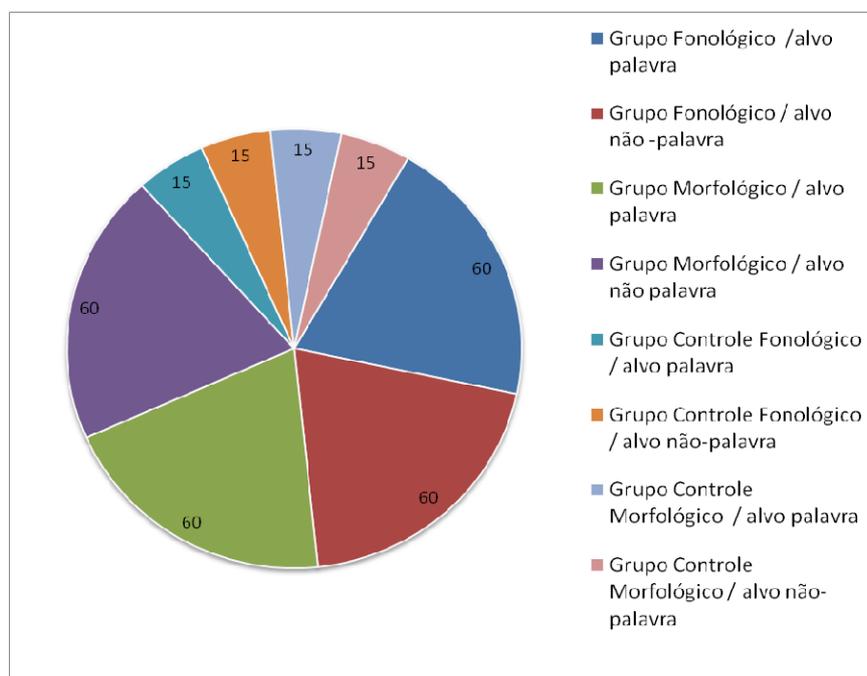


Gráfico 1: Distribuição dos 300 estímulos (pares *prime*-alvo) por grupo

Neste trabalho, fizemos uma comparação entre pares de palavras (*prime*-alvo) com relação fonológica e pares de palavra da mesma família morfológica e, para tal, construímos estímulos *prime*-alvo com relacionamento fonológico e morfológico. No GF, constando de palavras com início fonologicamente semelhantes, foram comparados 60 pares de palavras e 60 de não palavras: *prime* e *alvo*. Estes estímulos foram divididos em três subgrupos com 20 pares cada, com *prime*-alvo de tamanhos variados: F5-6, com cinco ou seis fonemas; F7-8, com sete ou oito fonemas; e F9-10, com nove ou dez fonemas; como aparecem exemplificados nos quadros abaixo:

	F5-6	F7-8	F9-10
1.	BATATA-barata	SARCASMO-sargento	TRAMBOLHO-trampolim
2.	FRANCA-frango	PROGRAMA-problema	JACARANDÁ-jaguatirica
3.	COMPRA-contra	PROJETO-profeta	SEMENTE-semelhança
4.	ÂNGULO-ancora	RECADO-retrato	TARÂNTULA-tartaruga
5.	ALERTA-alegre	TORTURA-torpedo	LABAREDA-labirinto
6.	TROMBA-tronco	ESTRADA-estrela	TORNEIRA-tornozelo

Quadro 3: Grupo F – palavras: *PRIME* e alvo com semelhança fonológica (20 pares em cada coluna/subgrupo)

	F5-6	F7-8	F9-10
1.	RESMA-remba	FARPA-farmécio	VASSOURA-vassemeiro
2.	PALMA-palfo	AVESSO-avelano	CARIE-carianjeno
3.	PERTO-presco	PRAIA-prazura	PIANO-pianquelia

Quadro 4: Grupo F – não palavras: PRIME e alvo com semelhança fonológica
(20 pares em cada coluna/subgrupo)

Em relação às palavras morfológicamente semelhantes, foram comparados 60 pares de palavras e 120 de não palavras *prime* e alvo. No GM, os estímulos foram divididos em três subgrupos, com 20 pares cada, com duas, três ou quatro camadas morfológicas além da primeira camada saussuriana: M2, M3 e M4, como estão exemplificados nos Quadros 5 e 6.

	M2	M3	M4
1.	PURO-pureza	IDEIA-idealismo	GLOBO-globalizado
2.	FRANÇA-francês	ARMA-armamento	CENTRO-centralização
3.	JOVEM-jovial	BRILHO-brilhante	MARGEM-marginalizado
4.	PONTA-pontudo	CANTO-cantador	INÍCIO-inicialização
5.	BOM-bondade	PRÊMIO-premiado	PAI-paternalista
6.	POBRE-pobreza	VACINA-vacinação	NAÇÃO-nacionalização

Quadro 5: Grupo M – palavras: PRIME e alvo com relação morfológica
(20 pares em cada coluna/subgrupo)

	M2	M3	M4
1.	VIDA-vidoso	TERRA-terrense	ROCHA-rochante
2.	BELO-belura	ARTÉRIA-arterioso	MENTE-mentalagem
3.	PENSÃO-pensioneiro	FLAUTA-fláutico	ÚTIL-utilizamento

Quadro 6: Grupo M – não palavras: PRIME e alvo com relação morfológica
(20 pares em cada coluna/subgrupo)

Adicionalmente, tínhamos o Grupo Controle (GC) com 60 pares *prime*-alvo que nos permitiram comparar a situação em que *prime*-alvo são exatamente a mesma palavra. Fizemos quatro subgrupos: o CF (Controle Fonológico) palavra, o CF não palavra, o CM (Controle Morfológico) palavra e o CM não palavra, como aparecem exemplificados nos Quadros 7, 8, 9 e 10.

1.	GAROTO- garoto
2.	PETISCO-petisco
3.	ANTIMÔNIO-antimônio

Quadro 7: Grupo CF – palavras: PRIME e alvo iguais (15 pares em cada coluna)

1.	PESANCO-pesanco
2.	LÂMOLA-lâmola
3.	JOTABA-jotaba

Quadro 8: Grupo CF – não palavras (15 pares em cada coluna)

1.	CORTADA-cortada
2.	APLICÁVEL-aplicável
3.	RENTABILIDADE-rentabilidade

Quadro 9: Grupo CM – palavras: PRIME e alvo iguais (15 pares em cada coluna)

1.	VORADA –vorada
2.	DROZÁVEL – drozável
3.	SENTIAGILIDADE - sentiagilidade

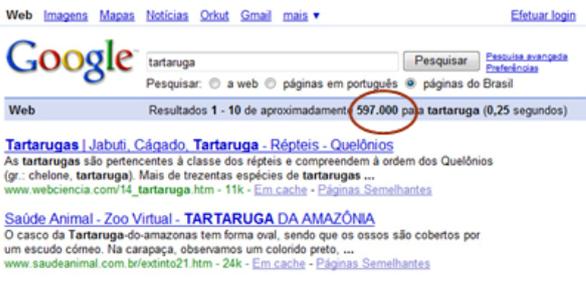
Quadro 10: Grupo CM – não palavras (15 pares em cada coluna)

2.2. A frequência dos estímulos

Nos últimos anos, a frequência tem sido encarada como um fenômeno complexo, promovendo um tipo de impacto diferente, dependendo da fase do acesso lexical e do tipo de palavra (EMBICK *et al.*, 2001; PYLKKÄNEN, 2002; STOCKALL *et al.*, 2004). Por exemplo, na fase inicial de ativação múltipla, uma palavra-alvo freqüente sob os aspectos semântico, fonológico, fonotático, e de extensão (número de fonemas), isto é, com características semelhantes àquelas da maioria das palavras de uma dada língua, vai ocasionar uma ativação rápida de um grande número de palavras. Nesta fase, freqüência equivale a rapidez. Porém, na fase seguinte, a tarefa é a supressão do excesso de palavras ativadas com o objetivo de preservar a única representação correta para um dado estímulo, ou seja, a tarefa se torna mais difícil e mais demorada se houver muitas palavras ativadas.

Nesta nova fase, a literatura aponta para a hipótese de que quanto mais freqüente for a palavra, maior será o adensamento das ativações múltiplas e mais demorada será a escolha da representação certa (EMBICK *et al.*, 2001; PYLKKÄNEN, 2002; STOCKALL *et al.*, 2004).

Aplicando estes achados para os grupos de estímulos utilizados neste teste, uma palavra alvo do Grupo F5-6 como BALA, por exemplo, ativa *pata, fala, pala, bala, sala, para, rala, paca, passa, gala, pano, papa* e muitas outras. O resultado então é a ativação múltipla e imediata de representações, quando só uma corresponde perfeitamente ao estímulo BALA. Na fase de supressão, o processo se torna mais lento pois BALA é apenas uma dentre muitas representações ativadas. Por outro lado, se a palavra-alvo tiver alguma singularidade, por exemplo, uma fonotática infreqüente como XEROX, ou for muito grande, como PINDAMONHANGABA, a etapa inicial da ativação cortical pode demorar um pouco mais, entretanto o resultado final constará de poucas palavras ativadas e, assim, na fase final do processo, a da supressão e escolha, o acerto no alvo se dará com facilidade.

Para F5/6 e F7/8 (palavras curtas + freq.)	Para F9/10 (palavras longas + infreq.)
 <p>batata</p> <p>Resultados 1 - 10 de aproximadamente 1.630.000 para batata (0,12 segundos)</p> <p>BATATA Rica em carboidratos, a batata é grande fonte de energia. Contém ainda sais minerais, vitamina C e, em pequenas quantidades, vitaminas do Complexo B. ... www.vitaminasecia.hpg.ig.com.br/batataorientacao.htm - 8k - Em cache - Páginas Semelhantes</p> <p>ABBA - Associação da Batata Brasileira by FPazzini III Seminário Brasileiro da Batata Para baixar as palestras, clique aqui ... Cadeia Brasileira da Batata. Resumos de trabalhos científicos. Leis nacionais ... www.abbabatatabrasileira.com.br/ - 18k - Em cache - Páginas Semelhantes</p>	 <p>tartaruga</p> <p>Resultados 1 - 10 de aproximadamente 597.000 para tartaruga (0,25 segundos)</p> <p>Tartarugas Jabuti, Cágado, Tartaruga - Répteis - Quelônios As tartarugas são pertencentes à classe dos répteis e compreendem à ordem dos Quelônios (gr.: chelone, tartaruga). Mais de trezentas espécies de tartarugas ... www.webciencia.com/14_tartaruga.htm - 11k - Em cache - Páginas Semelhantes</p> <p>Saúde Animal - Zoo Virtual - TARTARUGA DA AMAZÔNIA O casco da Tartaruga-do-amazonas tem forma oval, sendo que os ossos são cobertos por um escudo córneo. Na carapaça, observamos um colorido preto, ... www.saudeanimal.com.br/extinto21.htm - 24k - Em cache - Páginas Semelhantes</p>
Número mínimo de aparições no Google: 1.500.000	Número mínimo de aparições no Google: 500.000
Número máximo de aparições no Google: 3.000.000	Número máximo de aparições no Google: 3.000.000

Quadro 11: Faixas aceitáveis de freqüência no *Google*

Haveria, portanto, uma compensação entre as palavras curtas de média freqüência e as longas, no que diz respeito ao curso temporal do acesso lexical. Assumindo estas generalizações, os estímulos utilizados neste teste foram classificados em dois grupos: um de palavras menores (F5-6 e F7-8) e outro de palavras maiores (F9-10). Foi verificado o número de consultas *Google* (*hits*) que haviam sido feitas aos itens lexicais escolhidos. Usando o número médio de hits entre as palavras menores (F5-6 e F7-8) e as maiores (F9-10), foram estabelecidas duas faixas *Google* de freqüência como mostra o Quadro 11: uma mais estreita para as palavras dos grupos F5-6 e F7-8 (entre 1.500.000 e 3.000.000 toques) e outra mais larga para o grupo 9/10 (entre 500.000 e 3.000.000 toques).

2.3 As comparações

Foram estabelecidos quatro tipos de comparações entre os subgrupos: (i) entre os três tamanhos de alvo (F5-6, F7-8 e F9-10) dentro do Grupo F; (ii) entre os três tipos de alvo do

Grupo M, com duas, três e quatro camadas (M2, M3 e M4) depois da saussuriana; (iii) entre os alvos das séries fonológica e morfológica. Em cada uma destas comparações, dois parâmetros foram verificados: o nível de ativação, informado pela amplitude da onda, e a demora de ativação, informada pela latência da onda. Para a construção do Gráfico 3, utilizamos um parâmetro-controle, que traz a latência e a amplitude em casos em que *prime* e alvo são exatamente iguais. O valor do controle foi levado a zero, de forma a sobrepor x no gráfico cartesiano, e os outros valores foram recalculados relativamente a este valor-controle. Depois de recalculados, os valores plotados acima de x correspondem ao efeito positivo do *priming*, ou seja, os casos em que a existência do *prime* facilita o acesso ao alvo. Os valores plotados abaixo de x correspondem aos efeitos do *prime* negativo, isto é, os casos em que o *prime* atrapalha o alvo.

2.4 Os voluntários

Trinta e seis voluntários (18 homens, idade média 28,3 anos), destros e com visão normal ou corrigida para o normal, todos alunos de diferentes cursos de Graduação da Universidade Federal do Rio de Janeiro, aceitaram participar do teste e para isto assinaram Termo de Consentimento Livre e Esclarecido, conforme estabelece a Resolução número 196, de 10 de outubro de 1996, do Conselho Nacional de Saúde. Os voluntários declararam não ter ingerido bebidas alcoólicas a partir do dia anterior ao teste, estar com boa saúde e não estar fazendo uso de barbitúricos, antidepressivos nem sendo tratados de nenhuma doença sistêmica. Os testes relativos a quatro voluntários, todos homens, foram eliminados por não terem atingido o mínimo de 85% de acerto na tarefa comportamental de aperto de um entre dois botões com o objetivo de discriminar palavras de não palavras.

3 Resultados

3.1 Comparações entre subgrupos do Grupo Fonológico

F5-6	F7-8	F9-10
BATATA-barata	SARCASMO-sargento	TRAMBOLHO-trapolim

Em termos das comparações dentro do Grupo Fonológico (GF), os alvos dos pares F5-6 não se mostraram estatisticamente diferentes daqueles dos pares F7-8, assim com os alvos dos pares F7-8 não se mostraram estatisticamente diferentes daqueles dos pares F9-10. Porém, houve resultado significativo ao compararmos os alvos dos pares F5-6 com aqueles dos pares F9-10. Isto significa que tivemos latência e amplitude de onda significativamente diferentes quando havia diferença linear de quatro segmentos (por exemplo, entre *barata* primada por *batata* em comparação com *trapolim* primado por *trambolho*).

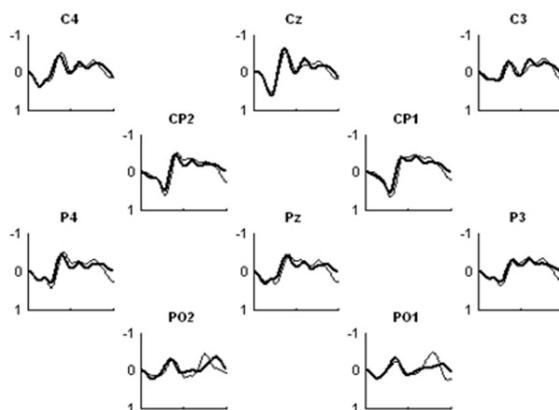


Gráfico 2: ERPs – linha grossa F5-6; linha fina F9-10

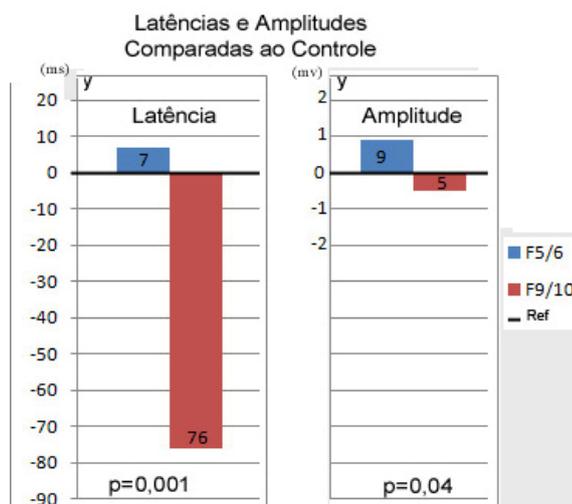


Gráfico 3: Comparações do Grupo Fonológico: latência em ms e amplitude em mv

O Gráfico 2 mostra derivações de interesse para a linguagem (C4, Cz, C3, CP2, CP1, P4, Pz, P3, PO2, PO1) com diferenças nos parâmetros de latência e amplitude das ondas. No Gráfico 3 os alvos no subgrupo F5-6 tiveram latências menores, medidas em milissegundo, e amplitudes maiores, medidas em microvolts, do que F9-10. Através dele podemos estabelecer uma comparação entre os valores médios de latência e amplitude a valores do Grupo controle. Em relação à latência, a coluna abaixo de $y = 0$ mostra valor de *priming* negativo, ou seja, latência maior do que a relativa a pares em que o *prime* e o alvo são a mesma palavra. A coluna acima de

y= 0 mostra latência menor do que a relativa a pares em que o *prime* e o alvo são a mesma palavra.

O Gráfico 4 mostra os valores absolutos para os subgrupos de GF. A configuração das ondas segue o formato *default* de ondas com ascendência mais íngreme, isto é, uma onda rápida que tem a tendência de chegar a uma amplitude maior antes de descer rumo à mudança de polaridade.

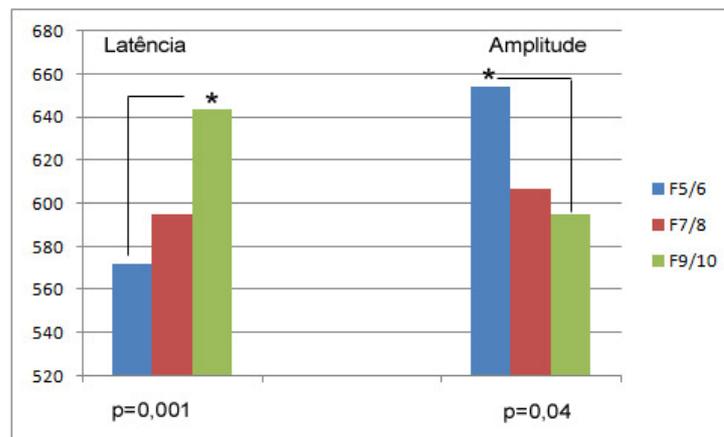


Gráfico 4: Valores absolutos dos sub-grupos de GF

3.2 Comparações entre subgrupos do Grupo Morfológico

M2
PURO-pureza

M3
IDÉIA-idealismo

M4
INÍCIO-inicialização

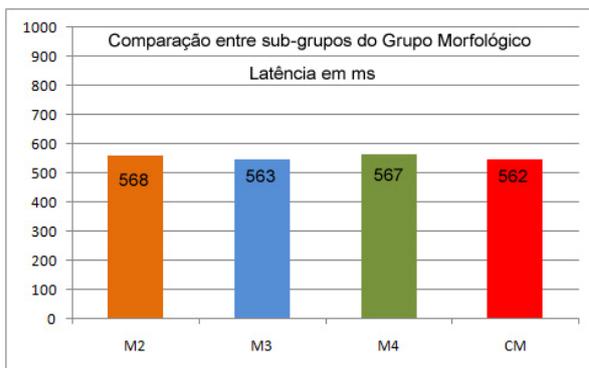


Gráfico 6: Valores absolutos de latência dos ERPs nos subgrupos de GM

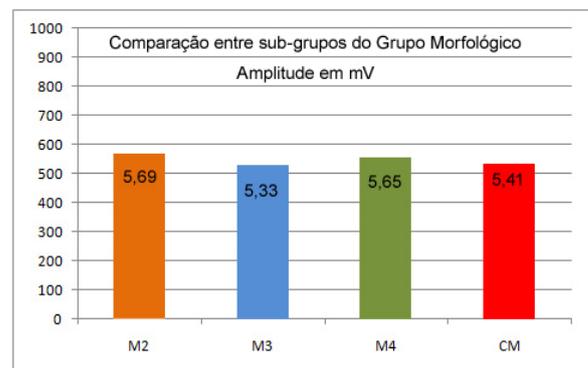


Gráfico 7: Valores absolutos de amplitude dos ERPs nos subgrupos de GM

O Gráfico 6 mostra valores estatisticamente iguais para as latências nos três subgrupos de M e também para o Grupo Controle relativo a M (CM). Nos subgrupos M2, M3 e M4, a diferença do tamanho linear dos alvos não causa nenhum impacto na amplitude e latência dos ERPs. De forma semelhante, o Gráfico 7 também mostra valores estatisticamente iguais, mas, desta vez, para as amplitudes nos três subgrupos de M e também para o Grupo Controle relativo a M (CM).

3.3 Comparações entre GF e GM

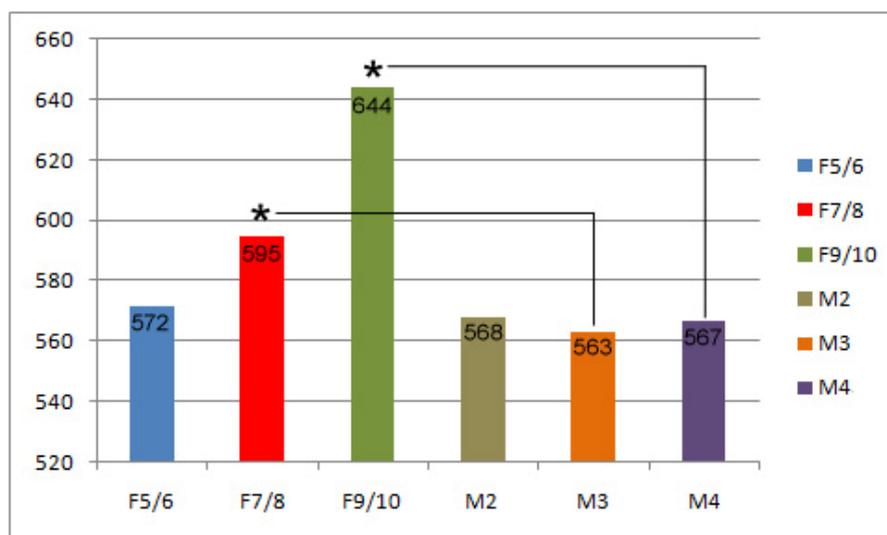


Gráfico 8: Comparações entre as latências de GF e GM

Conclusão

Constatamos que o acesso a palavras pequenas é mais rápido do que a grandes em pares com semelhança fonológica. Entretanto, esta diferença só se tornou uma gradação estatisticamente significativa quando os alvos tinham quatro segmentos fonológicos de distância (F5-6 comparado com F9-10).

Considerando os grupos categorizados por profundidade morfológica, constatamos que a diferença do número de camadas e, conseqüentemente, do número de segmentos fonológicos não cria uma gradação quando o *prime* e o alvo compartilham a primeira instância de categorização: o acesso a *pureza* depois de *puro* se dá em janela temporal igual à do acesso a *centralização* depois de *centro*.

Este achado fala a favor de que a latência encontrada é sempre relacionada ao ponto da negociação semântica, ou seja, na primeira concatenação entre raiz e morfema categorizador.

Este é o ponto que pode ser legitimamente chamado de acesso lexical, onde uma onda de polaridade negativa com amplitude mais pronunciada e menor latência pode ser detectada. Além disso, temos que ressaltar que esta onda relativa a palavras de até quatro camadas morfológicas é mais rápida do que a onda que marca o acesso lexical a *barata* depois de *batata*. Isso equivale a dizer que o reconhecimento lexical seria determinado pela morfologia e incidiria no ponto em que a raiz ganha categorização, o ponto do *merge* da raiz com a primeira peça vocabular categorizadora.

Este achado aparentemente paradoxal tem interpretação à luz da Teoria da Morfologia Distribuída (MD) da seguinte maneira. De acordo com a MD, as palavras não são pré-montadas, mas sim formadas dinamicamente a partir da distribuição de tarefas entre três listas diferentes que participam de três estágios diferentes na derivação da estrutura, sendo a interpretação a última dessas tarefas.

Na Enciclopédia (Lista 3) se estabelece o pareamento entre forma e sentido, isto é, é lá que um significado arbitrário é atribuído ao composto [Raiz + primeiro morfema categorizador]. É lá portanto que o verdadeiro acesso lexical acontece. Nos três pares de M2, M3 e M4, respectivamente PONTA-pontudo, IDÉIA-idealismo e GLOBO-globalizado, o *prime* e o alvo compartilham uma mesma Raiz concatenada a um mesmo morfema categorizador *n* (*enezinho*). Nos três casos, há negociação semântica neste mesmo ponto. Então, independentemente de camadas morfológicas mais externas, por exemplo, [-udo]_a que adjetiviza o nome *ponta*; [-al]_a e [-ismo]_n que respectivamente adjetiviza e nominaliza de novo a idéia; e [-al]_a, [-iz]_v, e [-ado]_a que adjetiviza, verbaliza e adjetiviza o nome *globo*, as duas palavras nos três pares de exemplo são idênticas, já que em uma certa fase da derivação, *pontudo* foi *ponta*, *idealismo* foi *idéia* e *globalizado* foi *globo*. A diferença entre as palavras destes três pares são as camadas externas, cujos valores semânticos são obtidos composicionalmente partindo do valor estabelecido pelo acesso lexical, ou seja, daqueles valores obtidos pela primeira concatenação (*merge*) entre raiz e morfema categorizador.

Nossos resultados experimentais, que mostram tempos iguais para os três subgrupos morfológicos e tempos menores para os subgrupos morfológicos em relação aos correspondentes subgrupos fonológicos, refletem exatamente que, para o acesso lexical, o que depende recursos cognitivos detectáveis pelo processo de extração de ERPs é a computação da negociação semântica, que faz a relação arbitrária entre forma e significado.

Se por um lado podemos num primeiro momento estranhar e quase lamentar a total invisibilidade das computações composicionais que se estabelecem a partir do ponto da arbitrariedade saussuriana, por outro lado este resultado nos força a levar a sério os achados descritos nos Gráficos 6, 7 e 8. Esses gráficos mostram que a metodologia aqui empregada para a extração de ERPs não está capturando a computação na interface da sintaxe com a semântica nas camadas sintáticas posteriores àquela em que se dá a fixação da parte arbitrária do significado das palavras.

Na verdade, estes gráficos confirmam de forma inequívoca as previsões da Morfologia Distribuída de que existe uma separação entre Listas e Operações. O modelo MD difere dos modelos lexicalistas justamente por advogar que as leituras semânticas composicionais provêm de uma computação que não tem início a partir de um Léxico pré-montado em uma outra espécie

de sistema computacional diferente e anterior à sintaxe. Pelo contrário, a mesma sintaxe já atua desde o interior das palavras complexas, a partir do ponto da arbitrariedade saussuriana: "*It's Syntax all the way down*" (MARANTZ, 1997). Os nossos achados eletrofisiológicos estão documentando exatamente isto: depois da detecção da arbitrariedade saussuriana entramos em outro módulo da gramática.

Contudo, ainda tomando esta mesma metodologia, tentaremos desenvolver uma investigação para testar o caso das palavras que não podemos afirmar que vão suscitar análise unânime entre falantes. É possível ver *refrigerar* em *refrigerante*; *restaurar* em *restaurante*? Temos feito uma observação factual com muitos falantes e numerosas palavras complexas. As opiniões parecem divergir.

No nosso próximo estudo, já em curso, faremos a promediação não por participantes e sim por palavras, para estudar casos como os exemplos acima. Em teste de percepção informal, algumas pessoas afirmam enxergar *prometer* dentro de *promissória*. Outras, levando em conta as diferenças semânticas, acreditam que não produzem *promissória* a partir de *prometer*, embora possam até se dar conta de que há *uma promessa de pagar* na palavra *promissória*. Para esta nova e instigante investigação, com análises item a item, a base de comparação serão as latências e amplitudes já evidenciadas em testes como este, em que todas as palavras utilizadas não poderiam ser segmentadas de modos diferentes por diferentes falantes.

ABSTRACT: This is a lexical access study in Brazilian Portuguese, stimulated by a priming protocol, and assessed by the extraction of event-related brain potentials. Sixty prime-target pairs bearing onset similarity (*batata-barata*) were compared with 60 pairs which hold morphological relationship (*centro-centralização*). Each of the two groups of 60 pairs of words was subdivided into three series of 20 pairs of words of different lengths. It was verified that, in relation to the pairs with onset similarity, access to small targets is faster than access to large ones. Contrastingly, the differences in length did not result in different activation times when prime and target belonged to the same morphological family and shared the first merge between root and a categorizing morpheme. So, cortical access to *pureza* after *puro* takes exactly the same time as access to *centralização* after *centro*, and both targets are faster than *barata* after *batata*. Thus, morphology seems to be an acceleration factor in relation to recognition of targets.

Keywords: syntax all the way down; saussurean arbitrariness, semantic compositionality, computations in Logic Form, electrophysiology of language

Referências

- EMBICK, D; HACKL M.; SCHAEFFER, J.; KELEPIR, M; MARANTZ, A. (2001) A magneto-encephalographic component whose latency reflects lexical frequency. **Cognitive Brain Research**, 10(3): 345–8.
- FIORENTINO, R. & POEPEL, D. (2007). Compound Words and Structure in the Lexicon. *Language and Cognitive Processes*, v12, p. 953-1000.
- FRANÇA, A. I. (2007) A Interface Lingüística-Neurociência da Linguagem. **Cadernos de Estudos Lingüísticos**, UNICAMP, v. 49, p. 151-166.



- FRANÇA, A. I. (2005) O léxico mental em ação: muitas tarefas em poucos milissegundos **Revista Linguística**, n.1 v. 1
- HALLE, M.; MARANTZ, A. (1994) Some key features of distributed morphology. In: **MITWPL Papers on Phonology and Morphology**, p. 275-288. Cambridge, MA: MIT Press.
- HALLE, M.; MARANTZ, A. (1993) Distributed morphology and the pieces of inflection. In: HALE, K.; KEYSER, S. J. (Eds.) **The View from the Building 20: Essays in Honor of Sylvain Bromberger**, p. 111-176. Cambridge, MA: MIT Press.
- HAY, J.; BAAYEN, R. H. (2005), Shifting paradigms: gradient structure in morphology. In: *Trends Cogn Sci.* v. 9, p. 342-348.
- MARANTZ, A. (1997) No escape from syntax: don't try morphological analysis in the privacy of your own lexicon. In: DIMITRIADIS, Alexis *et al.* (Eds) *Upenn Working Papers in Linguistics 4.2. Proceedings of the 21st Annual Penn Linguistics Colloquium*, Philadelphia, pp. 201-225.
- PHILLIPS, C; PELLATHY, T; MARANTZ, A; YELLIN, E; WEXLER, K; POEPPEL, D; MCGINNIS, M; ROBERTS, T. (2000) Auditory cortex accesses phonological categories: an MEG mismatch study. **J Cogn Neurosci**, nov, 12(6):1038-55.
- PINKER, S. (1999) **Words and Rules: The Ingredients of Language**. NY: Basic Books. HarperCollins.
- PYLKKANEN, L; FEINTUCH, S; HOPKINS, E; MARANTZ, A. (2004) Neural correlates of the effects of morphological family frequency and family size: an MEG study. **Cognition**. Apr;91(3):B35-45.
- PYLKKÄNEN, L; MARANTZ, A. (2003a) Tracking the time course of word recognition with MEG. **Trends in Cognitive Sciences**. 2003; 7:187-189.
- PYLKKÄNEN, L.; GONNERMAN, L.; STRINGFELLOW, A.; MARANTZ, A. (2003b) Disambiguating the source of phonological inhibition effects in lexical decision: an MEG study. (ms). Disponível em: <http://www.psych.nyu.edu/pylkkanen/papers/Inhibition_ms_CBR_w_figs_submitted.pdf>
- PYLKKÄNEN, L.; STRINGFELLOW, A.; MARANTZ, A. (2002) Neuromagnetic evidence for the timing of lexical activation: an MEG component sensitive to phonotactic probability but not to neighborhood density. **Brain and Language**. 81(1-3): 666-78.
- PYLKKÄNEN, L.; STRINGFELLOW, A; FLAGG, E; MARANTZ, A. (2000) A neural response sensitive to repetition and phonotactic probability: MEG investigations of lexical access. **Proc. 12th Int. Conf. Biomagnetism**, Helsinki University of Technology, pp. 363-367.
- STOCKALL, L; MARANTZ, A. (2006) A single route, full decomposition model of morphological complexity: MEG evidence. *The Mental Lexicon* 1:1
- STOCKALL, L.; STRINGFELLOW, A.; MARANTZ, A. (2004). The precise time course of lexical activation: MEG measurements of the effects of frequency, probability, and density in lexical decision. **Brain and Language**. 90 (1), 88.