

Conexões conceituais: Um estudo psicolingüístico de *priming* encoberto

Aniela Improta França*

Miriam Lemle**

Isabella Lopes Pederneira***

Juliana Novo Gomes***

Resumo – Exploramos a hipótese de que há relações semânticas primordiais entre palavras (do tipo *qualia*), de que outras são mais tênues e que estas relações podem ser definidas com mais precisão através de uma sintaxe interna à semântica. Para verificar esta hipótese, formatamos o presente estudo de *priming* semântico encoberto que compara quatro séries de pares de palavras com níveis de relacionamento progressivamente mais tênues: Série 1, pares do tipo escola-aluno, Série 2, praia-calção; Série 3, cadeira-revólver, Série 4, batata- parobo, esta última envolvendo uma não palavra que justifica a tarefa requisitada ao voluntário de distinguir entre alvos palavra e alvos não palavra.

Palavras-chave – *Priming* semântico. *Qualia*. Sintaxe da semântica.

* Professora Adjunta do Departamento de Lingüística da UFRJ (Rio de Janeiro, Brasil); e-mail: anielaimprota@terra.com.br. ** Professora Titular do Departamento de Lingüística da UFRJ; e-mail: miriamlemle@gmail.com. Miriam Lemle e Aniela França são, respectivamente, Coordenadora e Pesquisadora de Eletrofisiologia da Linguagem do Laboratório CLIPSEN (<http://www.lettras.ufjf.br/clipsen>). *** Bolsistas de Iniciação Científica e alunas da graduação de Letras da UFRJ.

1. Introdução

De acordo com Modelo da Morfologia Distribuída, uma versão não lexicalista do Programa Minimalista, as palavras são formadas através de mobilizações sintáticas feitas em traços abstratos, que a princípio não são dotados de conteúdo fonológico. Ao final de cada fase de derivação sintática, unidades morfofonológicas, estas dotadas de conteúdo fonológico, começam a competir por inserção nos nós terminais que foram criados pela sintaxe. Desta forma, este modelo de gramática é separacionista já que a computação sintática é separada da implementação fonológica (MARANTZ, 1997, 2001; HARLEY, 1998; MARVIN, 2002).

Com estas previsões teóricas, no campo experimental, muitos protocolos se colocam como objetivo demonstrar o fato de que, assim como na derivação de sentença, internamente à palavra também há processamento sintático (PYLKKANEN et alii, 2002 e 2003).

Uma das metodologias empregadas para demonstrar o processamento sintático interno à palavra é o protocolo experimental de “priming”. O método baseia-se na suposição de que uma palavra possa ser acessada mais rapidamente se precedida em curto prazo por outra com a qual compartilhe propriedades semânticas, segmentos fonológicos ou unidades da morfologia.

Para estudar a influência que uma palavra pode ter no acesso lexical de outra, apresenta-se uma palavra (*prime*) antes de outra (*alvo*) que se quer estudar. Por exemplo, poderíamos comparar a influência que “martelo” (*prime*) tem sobre “martelada” (*alvo*), comparativamente a “telefone” (*prime*) / “machadada” (*alvo*).

A estimulação é feita com vários pares *prime*/ *alvo* (“martelo” / “martelada”) misturados aleatoriamente a um igual número de pares não-*prime*/ *alvo* (“telefone” / “machadada”) e

também pares *prime* / não-palavra (“cadeira”/“parobo”). Esta metodologia pode utilizar dois tipos de *prime*: um que fica na tela por um tempo de 200ms, assim como o seu alvo, e outro, que é apresentado subliminarmente por 38 ms e seu alvo por 200 ms.

Através do tempo de resposta, nos testes psicolingüísticos podemos então avaliar a influência das representações segundo noções bastante claras na teoria: semelhança morfológica e semelhança fonológica. As relações puramente nocionais, como, por exemplo, idéia/noção vêm sendo utilizadas somente como grupo controle, porque essa faceta das palavras é, por enquanto, um campo bem menos compreendido.

Neste teste, nos aventuramos em explorar exatamente este campo das relações nocionais: a idéia de que há relações semânticas primordiais entre palavras e de que outras, porém, são mais tênues. Estamos aqui chamando de relação primordial aquela em que uma palavra está envolvida na definição de outra, a partir de uma estrutura sintática mínima do tipo Xis-Barra, que compreenda um núcleo, o seu complemento e o seu especificador. Em uma perspectiva semântica, estas são relações do tipo *qualia* (PUSTEJOVSKY, 1993, 1995). Usamos estas relações para elaborar os estímulos (pares *prime* / alvo) da Série 1 (cf. Figura 1), como nos exemplos seguintes:

- (i) cinto / fivela - cinto com fivela → relação todo-parte
- (ii) golfe / bola - golfe com bola → relação constitutiva
- (iii) escola / aluno - escola para aluno → relação télica

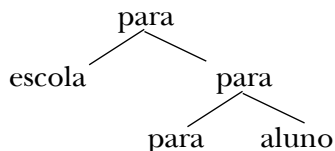


Figura 1 - Estrutura de pares da Série 1

Para elaborar os estímulos da Série 2, nos ativemos a pares de palavras cujas relações dependem de acesso a conhecimento de mundo além da sintaxe estreita. Observem os exemplos abaixo:

- (i) praia / calção → praia é um lugar onde os freqüentadores do sexo masculino podem ir usando calção.
 (ii) balão / festa → balão é um tipo de objeto de recreação que pode ser usado em algumas festas.

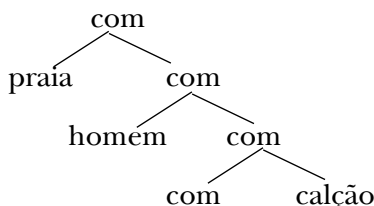


Figura 2 – Estrutura de pares da Série 2

Levando em conta o espírito da Morfologia Distribuída, decidimos explorar a hipótese de que haja isomorfismo entre a sintaxe e a “sintaxe” interior aos conceitos. Se assim fosse, nos estímulos da Série 1 as relações de prime iriam operar mais fortemente do que na Série 2. Isto porque na Série 1 a informação requerida no alvo já estaria contida no prime. Em contraste, na Série 2, supusemos que a informação contida no prime fosse mais indiretamente relacionada às informações que estão no alvo nesta “sintaxe” da semântica.

Para compor o desenho do experimento, elaboramos a Série 3 com palavras não relacionadas semanticamente. Para conectá-las é necessário criar muitas proposições no discurso que se sustentem na experiência individual. Por exemplo: cadeira / revólver? na minha casa tem uma cadeira onde meu pai sobe para esconder o revólver em cima do armário. Nossa expectativa é que a visão prévia de “cadeira” não vá facilitá-

tar o acesso à palavra alvo, revólver. Esta série é, portanto o grupo controle para o nocaute do prime. A Série 4 contém pares de palavras - não palavra destinadas exclusivamente a permitir que haja uma tarefa de decisão lexical para o voluntário: apertar um botão de *joystick* para decidir se o alvo é uma palavra ou uma não palavra. Esta tarefa proporciona que o voluntário se mantenha atento à palavra alvo.

2. Materiais e métodos

2.1 Protocolo de priming encoberto

A técnica experimental utilizada foi a de prime encoberto. Nesta técnica, o prime, que é a primeira palavra do par, é sempre mostrado subliminarmente, no nosso caso por 38ms. Seguindo ao prime, exibíamos o alvo por 200ms. O voluntário era instruído a opinar se o alvo era palavra ou não-palavra através do apertado de botão de um *joystick*. O botão superior deveria ser acionado para o julgamento 'palavra' e o inferior para 'não palavra'.

Como o teste envolvia o manuseio correto do *joystick*, fazíamos um pré-teste para habituar o voluntário ao uso automatizado deste instrumento. Este treinamento permitia que o teste só começasse depois que os sujeitos tivessem automatizado o acionamento dos três botões que seriam empregados durante o experimento lingüístico para julgar o estímulo como palavra ou não-palavra. O treinamento para o uso do *joystick* durava em média três minutos e podia ser repetido se o condicionamento efetivo não acontecesse dentro do período estipulado, entretanto raramente houve necessidade desta repetição.

Após o teste do *joystick*, começava o treinamento para o teste lingüístico (*warm-up*). Durante este período de aproxi-

madamente cinco minutos, todas as instruções para o teste apareciam na tela do computador. Além disso, e ainda antes do início do teste, as instrutoras ofereciam sua assistência para o esclarecimento de qualquer dúvida que ainda pudesse existir.

Como a aparição do *prime* na tela pode ser perceptível, antes de iniciar o teste, instruíamos os voluntários para que focalizassem o julgamento na segunda palavra.

Após o julgamento, uma cruz de fixação ocular aparecia na tela por 1,500 ms, para que houvesse uma cessação dos recursos cognitivos recrutados para o processamento da palavra apresentada, em prol de um redirecionamento de atenção para o próximo estímulo.

2.2 Estímulos e voluntários

O tempo de resposta para cada estímulo apresentado é gravado pela plataforma utilizada e comparado ao restante dos tempos da amostra. Isto permite que se testem hipóteses sobre o modo e o curso do processamento de cada par de palavra testada. O método baseia-se na suposição de que uma palavra possa ser acessada mais rapidamente se precedida em curto prazo por outra com a qual compartilhe propriedades semânticas, segmentos fonológicos ou unidades da morfologia.

Os estímulos, que consistem em pares de palavras distribuídas na 4 séries já descritas, foram selecionados entre palavras de alta frequência através do julgamento dos experimentadores.

No experimento contamos com a participação de 38 sujeitos normais (21 mulheres e 17 homens), destros, falantes nativos de português. Tinham idade média de 19,9 anos, todos alunos de graduação da Universidade Federal do Rio de Janeiro.

A apresentação dos estímulos se dava em um computador, no Laboratório CLIPSEN, localizado na Sala H 308 do

prédio de Letras da UFRJ, sendo criteriosamente observadas as condições ideais do ambiente de teste: silêncio, boa iluminação, temperatura condicionada.

2.3 Protocolo de estimulação cinética

O *script* computacional para a apresentação dos estímulos de leitura auto-monitorada foi criado na plataforma *Presentation*, da *Neurobehavioral Systems*, Albany, Estados Unidos. Os tempos de leitura gravados pelo *Presentation* foram agrupados por grupos.

2.4 Protocolo estatístico

Os tempos de reação em cada Série passaram por duas análises estatísticas de comparação: primeiro voluntário por voluntário, série a série e depois usando a médias de todos os voluntários por série. O teste estatístico utilizado foi o ANOVA 1 (Analysis of Variance 1), que revela se existe relação de diferença ou igualdade entre as médias das amostras do tempo de reação aos segmentos em questão. Se a diferença fosse significativa, se acharia $p \leq 0,05$. Uma rotina feita em MATLAB – produzido por *The MathWorks, Inc.*, em Massachusetts – permitia ainda que se retirassem os *ouliers* (dados discrepantes) iguais a pelo menos três vezes o valor do desvio padrão.

3. Resultados

No Gráfico 1, podemos ver os tempos médios obtidos para os 50 pares de palavras da Série 1. Cada ponto representa o tempo de reação a um alvo de um dado par. Como pode-se observar, 90 % dos tempos de resposta nesta série ficaram na faixa entre 500 e 600 ms.

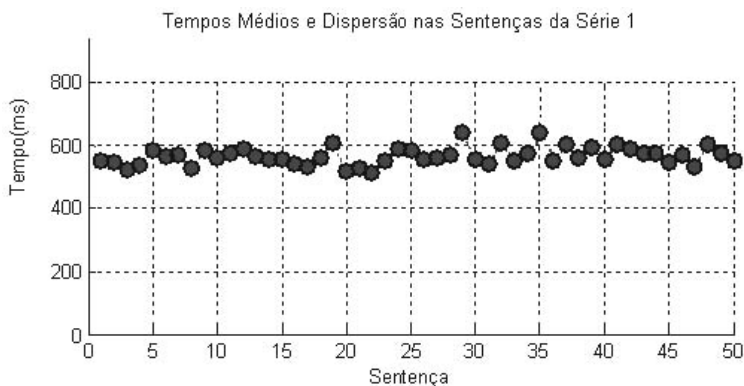


Figura 3 - Série 1: escola/aluno

No gráfico dos tempos médios de resposta da Série 2 (cf. Gráfico 2), temos 66% dos tempos de respostas entre 550 e 650 ms.

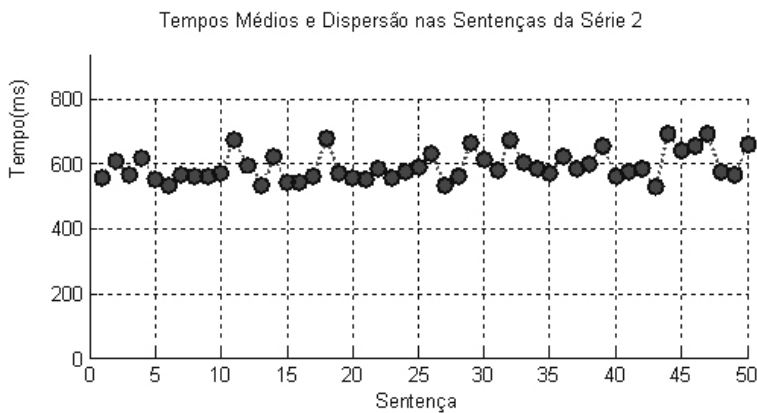


Figura 4 - Série 2: praia/calção

Na Série 3 (cf. Gráfico 3), 66% dos tempos médios de respostas obtidos ficaram entre 650 e 750ms.

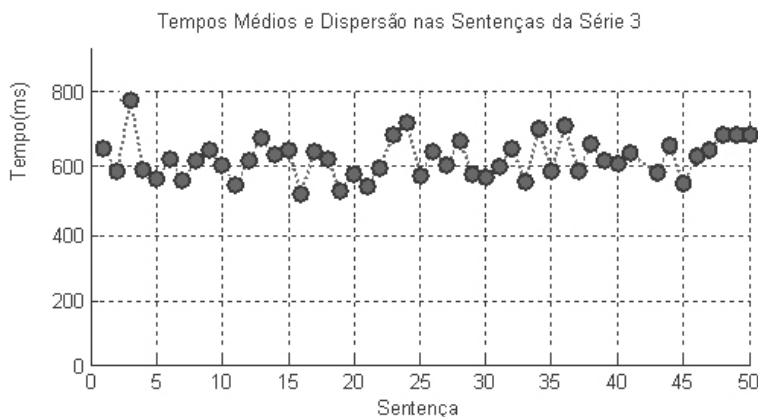


Figura 5 – Série 3: cadeira / revólver

Na série 4, os tempos de resposta, como podemos observar no Gráfico 4, ficaram $\approx 80\%$ entre 700 e 800 ms.

Podemos, ainda, notar que há um maior aglomerado de pontos, isto porque como esta série correspondia à tarefa que o voluntário deveria efetuar (como explicado em voluntários e estímulos), esta série possui 150 pares de *prime*/não-palavra, número equivalente à soma de todos os pares das Séries 1, 2 e 3. Isto é necessário para que o voluntário veja o mesmo número de alvos palavras do que de alvos não palavras.

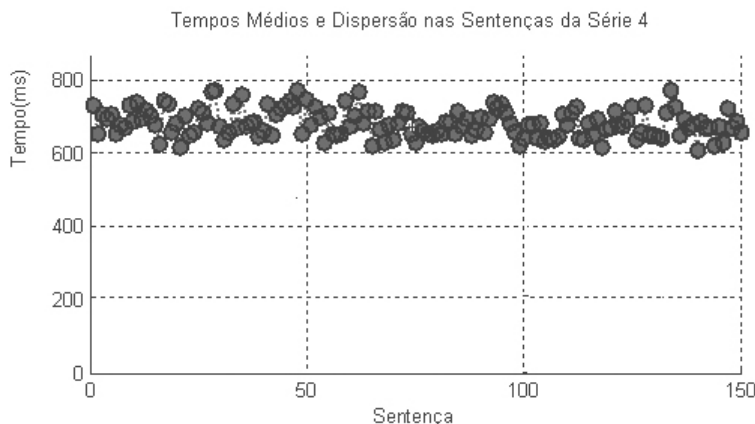


Figura 6 – Série 4: banaba / parobo

No Gráfico 5 podemos ver os tempos de reação médio em cada série, ainda sem o tratamento estatístico.

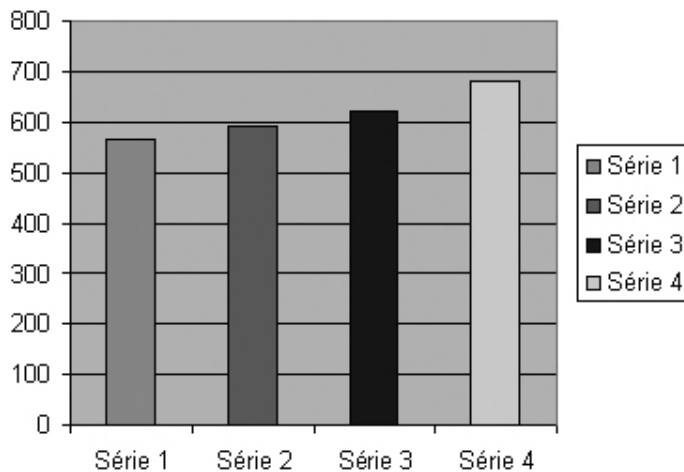


Figura 7 – Médias brutas do tempo de reação aos alvos das séries experimentais

Depois de contabilizados linearmente, nossos resultados brutos passaram por uma normalização e por uma análise comparativa de variância, cujo índice de significância (teste ANOVA) é dado pelo coeficiente estatístico $p = 0,05$. Temos, então, no Gráfico 6 à luz da estatística, o boxplot da comparação entre as Séries 1 e 2:

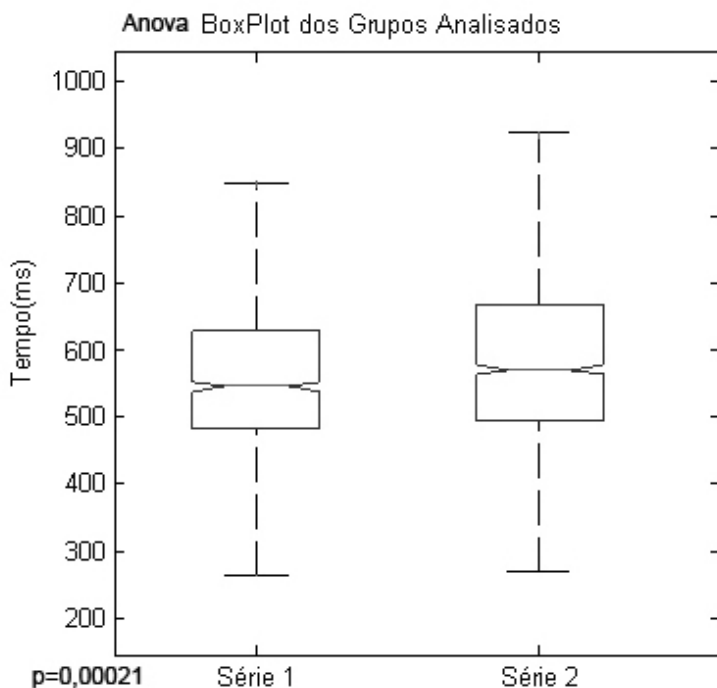


Figura 8 - ANOVA 1 dos tempos de reação entre as Série 1 e a Série 2

Podemos ver os tempos médios de resposta da Série 1 entre 500 e 600 ms e da Série 2 entre 550 e 650 ms. Note o coeficiente, $p=0$, que indica que as diferenças entre os tempos de reação entre estas séries é estatisticamente significativo, ou seja os tempos de reação em relação à Série 1 são significativamente menores do que os em relação à Série 2.

Agora, no Gráfico (7) apresentamos, o boxplot da comparação entre as Séries 3 e 4:

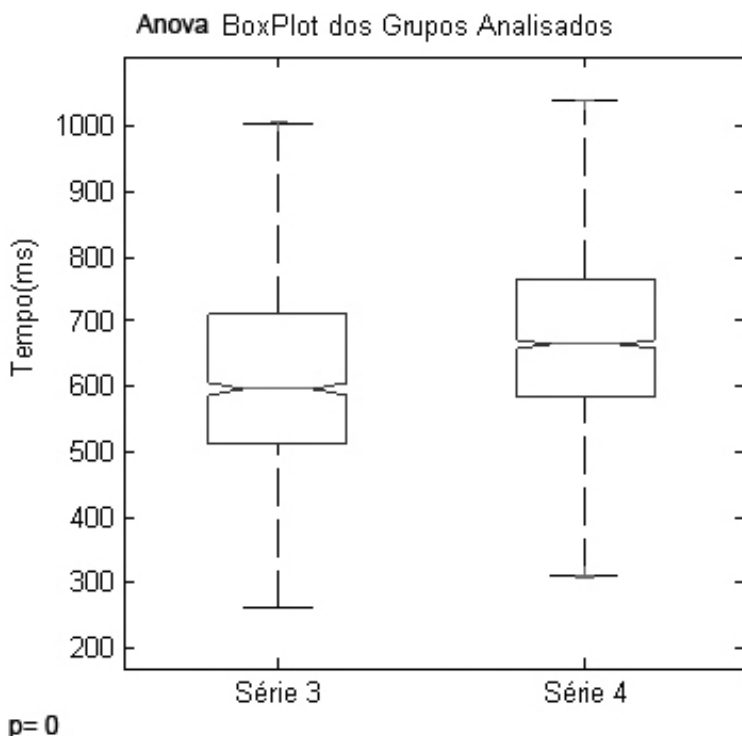


Figura 9 - ANOVA 1 dos tempos de reação entre as Série 3 e a Série 4

Os tempo médio de resposta da Série 3 ficou entre 650 e 750 ms e da Série 4, entre 700 e 800 ms. O coeficiente, $p=0$, indica que as diferenças entre os tempos de reação entre as duas séries é estatisticamente significativo.

4. Discussão

No protocolo de priming semântico o significado do prime tem alguma relação com o significado do alvo e por isso favorece o reconhecimento do alvo em termos de tempo. Este trabalho procurou captar a relação prime-alvo em termos sintáticos, buscando pares de palavras que pudessem ser

juntadas de uma maneira que a segunda consistisse em uma noção que já está latente na própria definição da primeira.

Sabe-se que nocionalmente cada palavra tem diversos aspectos. Por exemplo, livro: quem o produz, para que serve, qual a constituição (matéria, tamanho, peso, etc), forma geral de diferenciação de outros objetos ao redor. Estas relações, conhecidas como *qualia* (plural de *qualis*), remontam aos gregos, e vem sendo retomadas em Pustejovski (1993, 1995). Um *qualis* funciona como uma possível definição latente de uma palavra. Por exemplo, um livro pode ser definido por seu *qualis* télico: o que se faz com um livro? Um livro serve para ser lido? Vou ler o livro agora. Ou por seu *qualis* constitutivo: de que é feito um livro? De folhas de papel presas por uma capa? Vou subir no livro para alcançar a prateleira. Ou por seu *qualis* agentivo: quem fez este livro? Seu autor. ? Os livros do Chomsky são importantes. Ou por seu *qualis* formal: Qual a fronteira entre livro e não livro? O livro é geralmente mais extenso do que uma revista e não é periódico. ? Não quero uma revista sobre música; quero um livro.

De uma maneira geral nas línguas do mundo, preposições são relacionadores que, de forma diferente em cada língua, fazem ressaltar um ou outro *qualis* de uma palavra. Por exemplo: bule para leite, *qualis* télico; colher de pau, *qualis* constitutivo; miado de gato, *qualis* agentivo; trono: cadeira de rei, *qualis* formal. Construímos os estímulos de forma que uma destas preposições surgisse implicitamente, na tentativa inescapável e inadiável de se estabelecer um relacionamento sintático entre as duas palavras do par.

Na Série 1, utilizamos pares de palavras cuja relação se reduzia sempre a um desses *qualia*, expresso por uma preposição básica (de, com, para). A nossa previsão de que este conjunto teria o tempo de reação mais rápido se baseou na idéia de que esta relação básica poderia ser expressa por

uma estrutura Xis-barra, tendo uma dessas preposições como núcleo. Ou seja, procuramos compatibilizar a noção de *qualia* com uma peça do vocabulário: a preposição, que conectaria as duas palavras. De fato, como prevíamos, os tempos de reação a estes pares foi o menor das quatro séries.

Na Série 2, procuramos fazer uma composição entre *prime* e alvo que exigisse uma outra palavra entre elas funcionando como ponte semântica. A relação entre as três seria mediada por duas preposições como núcleos. Ou seja, o par envolveria dois *qualia*. Os resultados mostraram que esta relação já longínqua, que necessitava de uma “ponte”, ainda funcionou na facilitação do alvo.

Surpreendentemente, a Série 3 ainda foi um pouco mais rápida do que a Série 4, que envolveu não palavras, e o resultado da comparação entre estas duas séries foi estatisticamente significativo ($p=0$).

5. Conclusão

A metodologia utilizada teve sucesso em captar a diferença do efeito de *prime* nas Séries 1 e 2 e foi sensível às diferenças entre a intermediação por um *qualis* e por dois *qualia*. Assim, pudemos comprovar estatisticamente a hipótese aventada de que existe isomorfismo entre a sintaxe e a “sintaxe” interior aos conceitos – a sintaxe da semântica.

Declaramos na Discussão que o protocolo psicolingüístico de priming semântico foi capaz de detectar que o significado do *prime* tem alguma relação com o significado do alvo através da análise dos tempos de resposta em relação a outros alvos não relacionados ou menos relacionados. Através da utilização de protocolo neurolingüístico pode-se ir além desta análise ao se utilizar também um outro

parâmetro investigativo: a amplitude do sinal elétrico cortical relacionado à aparição de alvos que foram seguidos por primes semanticamente relacionados. Planejamos agora dar um passo adiante testando nossos estímulos através de técnica neurolingüística de extração de potenciais relacionados a eventos (ERP).

Conceptual Connections : A masked priming behavioral study

Abstract – We explore the hypothesis that there are primordial semantic relationships between words, of the *qualia* type, and that these relationships may be more precisely defined syntactically. To verify this hypothesis we have formatted the present masked priming semantic study that compares four series of word pairs with semantic relationships increasingly weaker. Series 1 stimuli are of the type *escola-aluno* (school-student), Series 2, *praia-calção* (beach-trunks); Series 3, *cadeira-revólver* (chair-gun), Series 4, *bataba-parobo* (potato-blicket), the latter involving a non-word as target, to justify the task that is required from volunteers to discriminate between word and non word targets.

Key words – Semantic priming. Qualia. Syntax of semantics.

Referências bibliográficas

HARLEY, H.; NOYER, R. *Licensing in the Non-Lexicalist Lexicon: Nominalizations, vocabulary Items and the Encyclopedia*. Cambridge, MA: MIT, 1998. (Working Papers in Linguistics n. 32, p. 119-137).

MARANTZ, A. No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon. In: Dimitriadis, A.; Siegel, L.; Surek-Clark, C.; Williams, A. (eds.) *Proceedings*

of the 21st Annual Penn Linguistic Colloquium, U Penn Working Papers in Linguistics 4.2: 201-225. Philadelphia: Penn Linguistics Club.1997

MARANTZ, A. Words. Ms. MIT, 2001.

MARVIN, T. *Topics in the Stress and Syntax of Words*. 2002. PhD Thesis, MIT.

PUSTEJOVSKY, J. Type Coercion and Lexical Selection. In: Pustejovsky, J. (ed.) *Semantics and the Lexicon*. Dordrecht: Kluwer Academic Publishers, The Netherlands, 1993.

_____. *The Generative Lexicon*. Cambridge, MA: MIT Press, 1995.

PYLKKÄNEN, L.; GONNERMAN, L.; STRINGFELLOW, A.; MARANTZ, A. *Disambiguating the source of phonological inhibition effects in lexical decision: an MEG study*. 2003. Ms. Disponível em: www.psych.nyu.edu/pylkkanen/papers/Inhibition_ms_CBR_w_figs_submitted.pdf.

PYLKKÄNEN, L.; STRINGFELLOW, A.; MARANTZ, A. *Neuromagnetic evidence for the timing of lexical activation: an MEG component sensitive to phonotactic probability but not to neighborhood density*. *Brain and Language*, v. 81, p. 666–678, 2002

Apresentado para publicação em novembro de 2005.

Aprovado para publicação em novembro de 2005.