

Luciana Garcia Ruiz

“Rebolar e requebrar, algo a mais do que dançar?”

Um estudo experimental sobre o acesso visual de verbos prefixados e pseudo-prefixados na Língua Portuguesa

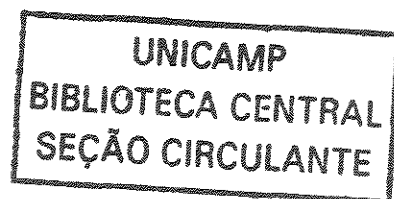
Dissertação apresentada ao Curso de Lingüística do Instituto de Estudos da Linguagem da Universidade Estadual de Campinas como requisito parcial para a obtenção do título de Mestre em Lingüística.

Orientador: Edson Françaço

Campinas

Instituto de Estudos da Linguagem

2001



200206410

UNIDADE	BC
N.º CHAMADA:	UNICAMP
	R859r
V	
Y	47544
P	837102
G	
D	<input checked="" type="checkbox"/>
PREÇO	R\$ 11,00
DATA	06-02-02
N.º CPD	

CM00163094-4

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA IEL - UNICAMP

R859r	<p>Ruiz, Luciana Garcia</p> <p>"Rebolar e requebrar, algo a mais do que dançar?": um estudo sobre o acesso visual de verbos prefixados e pseudo-prefixados na língua portuguesa / Luciana Garcia Ruiz. - - Campinas, SP: [s.n.], 2001.</p> <p>Orientador: Edson Françaço</p> <p>Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Estudos da Linguagem.</p> <p>1. Psicolingüística. 2. Processamento de palavras. 3. Morfologia. I. Françaço, Edson. II. Universidade Estadual de Campinas. Instituto de Estudos da Linguagem. III. Título.</p>
-------	--



Prof. Dr. Edson Françaço

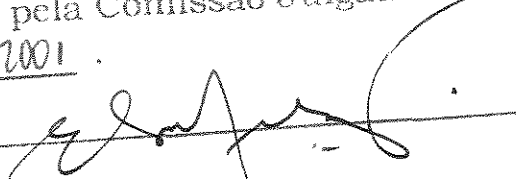
Prof.a. Dra. Aglael Juliana Aparecida Gama Rossi

Prof.a. Dra. Ester Míriam Scarpa

Este exemplar e a redação final da tese defendida por Luciana Garcia

Rui

e aprovada pela Comissão Julgadora em 28 / 11 / 2001.



Sem dúvida alguma, para Mamita e Tia Bê.

AGRADECIMENTOS

Pela presença, em espírito, iluminando meu caminho: ao meu pai.

Pelo estímulo para vencer os obstáculos: aos meus irmãos, principalmente ao Rogério.

Ao Edson, pelo conhecimento, paciência e perseverança.

A você, Nazaré, pelo amor que me dá segurança para seguir sempre em frente.

Pelos abraços nas horas fáceis e difíceis e por estarem sempre por aí: aos meus amigos, em especial à Paulinha.

Pelo colo e pela atenção, no momento mais complicado: à Eleonora.

AOS MEUS SUJEITOS, pela colaboração imprescindível.

Aos amigos e, principalmente, Amigas do LFAPE, pelo companheirismo e também por se sujeitarem aos meus testes.

Pelo trabalho lúdico, que me descansa das questões acadêmicas: ao Grupo Saia Rodada e a Tião Carvalho.

Pelas informações e ajuda a qualquer hora: aos funcionários da Secretaria, da Informática e da Biblioteca.

Pela disponibilidade de avaliar até em uma sexta de Carnaval: à Banca Examinadora.

E, por fim, a Deus e ao Bom Humor, que me permitiram continuar adiante quando parecia cientificamente impossível!!!

RESUMO

O propósito desta dissertação é estudar, por via de um enfoque experimental, o acesso lexical de palavras morfologicamente complexas, apresentadas visualmente. A questão de interesse central relaciona-se a como as palavras morfologicamente complexas – em especial, os verbos prefixados – estão representadas: as unidades de processamento correspondem à palavra integral, ou elas correspondem às unidades morfológicas constituintes (base e afixos)? Inicialmente procedeu-se a uma revisão bibliográfica, e em seguida uma série de experimentos foi planejada. Esse programa experimental foi parcialmente implementado, constituindo-se de dois experimentos (ao que se sabe, inéditos em Língua Portuguesa). No primeiro deles, replicou-se resultados tradicionais na literatura relativos à existência de efeito de *priming* semântico entre pares de palavras morfologicamente simples. Já no segundo experimento, trabalhamos o efeito de *priming* semântico com duas classes de verbos prefixados, os verbos prefixados transparentes e os pseudo-prefixados. No caso destas últimas, existe uma aparente prefixação formal, mas de fato esta palavra é uma outra unidade lexical semanticamente distante desta unidade analisada (por exemplo, *requebrar* nada tem a ver com *quebrar*). A comparação do comportamento dos sujeitos experimentais, numa tarefa de decisão lexical, frente a essas duas condições, permite estabelecer uma distinção no processamento de palavras que pode refletir diferenças de representação. Assim, conclui-se que as informações morfológicas têm um papel importante no acesso lexical.

SUMMARY

The aim of this dissertation is to experimentally study the lexical access of morphologically complex words, presented visually. The main question relates to how morphologically complex words – especially prefixed verbs – are represented: are processing units in correspondence to a whole-word representation, or are they in correspondence to morphological constituents (base and affixes)? Firstly, a literature review was conducted, and next a series of experiments was designed. Such an experimental program was partially implemented, and two experiments were run (to the best of our knowledge, none of them has ever been carried out in Portuguese). In the first, it was shown that semantic priming obtains between morphologically simple pairs; this is a common result in the psycholinguistic literature. In the second, semantic priming was employed to study the processing of two classes of morphologically complex verbs: those that are derivationally transparent, and those that are pseudo-prefixed. The latter are apparently the result of prefixation, but in fact form semantically different words (for instance, *reserve* cannot be put in correspondence with *serve*). The comparison of the subjects' responses on these two conditions, in a lexical decision task, allows one to determine processing differences related to differences in access representation. Thus, morphological information is shown to play a most important role in lexical access.

Este texto ambiciona ser o ponto de partida para um longo trabalho em uma área, até agora, pouco procurada pelos estudantes da Lingüística neste Instituto. Por se tratar, então, de uma pesquisa de base, vimos a necessidade de trazer aqui todo um panorama dos conceitos básicos que permeiam o estudo Psicolingüístico Experimental.

—

SUMÁRIO

RESUMO	9
SUMMARY	11
SUMÁRIO	15
CAPÍTULO I	17
INTRODUÇÃO	19
A. <i>BACKGROUND</i> DO PROBLEMA	19
B. ALGUNS CONCEITOS BÁSICOS	23
C. O QUE É O LÉXICO AFINAL?	23
D. MODELOS DE PROCESSAMENTO DA INFORMAÇÃO LEXICAL	26
CAPÍTULO II	57
O PROBLEMA E AS HIPÓTESES	59
A. AS PROPRIEDADES DISTRIBUCIONAIS DOS AFIÇOS, UMA PERSPECTIVA INOVADORA PARA A ORGANIZAÇÃO DAS INFORMAÇÕES MORFOLÓGICAS NO ACESSO LEXICAL	59
CAPÍTULO III	71
MATERIAIS E MÉTODOS	73
A. DIVERSAS MODALIDADES DE INVESTIGAÇÃO EXPERIMENTAL NA PESQUISA PSICOLINGÜÍSTICA	73
B. MORFO: UM PROGRAMA EXPERIMENTAL PARA O ESTUDO DA PREFIXAÇÃO NO PORTUGUÊS	82
BIBLIOGRAFIA	103
ANEXO I	109
EXPE	111

ANEXO II	115
EXPERIMENTO I: PRIMING SEMÂNTICO COM SUBSTANTIVOS E ADJETIVOS	117
EXPERIMENTO II - PRIMING SEMÂNTICO COM VERBOS PREFIXADOS	125
ANEXO III	135

CAPÍTULO I

INTRODUÇÃO

A. Background do Problema

A Psicolinguística é, basicamente, uma área de pesquisa científica que busca estudar a existência e o funcionamento dos mecanismos mentais envolvidos no processamento da linguagem humana. Essa busca se dá de forma interdisciplinar, já que, para chegar a proposições coerentes do ponto de vista da compreensão e da produção linguísticas, uma teoria psicolinguística deve atender às necessidades de correntes teóricas ligadas à Linguagem, à Psicologia e às Ciências da Cognição. Além disso, para dar mais confiabilidade às teorias que elaboram, os psicolinguistas são levados a testá-las por meio de experimentos; adicionando-se, então, a esse processo de pesquisa fatores estatísticos e computacionais para o controle e o manejo dos dados.

Em primeiro lugar, faz-se necessário evidenciar que, ao falarmos do estudo do “processamento de linguagem”, estamos abrangendo uma infinidade de questões sobre os processos cognitivos - listando apenas alguns dentre eles, temos a percepção acústica ou visual, a decodificação de impulsos nervosos em informação, a representação das informações linguísticas, a sua organização dentro do chamado dicionário mental etc. Portanto, é mister que se faça um recorte objetivo no estudo psicolinguístico, para que o pesquisador não se perca nas diversas possibilidades que este apresenta, tendo como consequência a produção de um trabalho teoricamente vago. Assim, certas áreas da pesquisa psicolinguística são privilegiadas em relação a outras pela facilidade de controle dos experimentos e de observação dos processos, como no caso da dicotomia **compreensão** x **produção** da linguagem. Pois quando tratamos da primeira, temos fatores mais controláveis do ponto de vista experimental, já que os dados utilizados no experimento podem ser previamente delimitados e levantados pelo pesquisador dentro de uma proposta teórica, na tentativa de controlar um possível erro experimental; enquanto que, no

tratamento da produção, temos que contar com fatores que fogem ao controle do pesquisador e, por vezes, até da teoria abordada por este; por exemplo, estratégias de processamento criadas pelos sujeitos do experimento, a capacidade de memória (ver Lima & França, no prelo, para saber mais sobre este tópico).

Em segundo lugar, mesmo tratando-se apenas da compreensão lingüística, há que se delimitar no mínimo três outros aspectos do estudo que suscitam questões psicolingüísticas muito complexas para o pesquisador: qual **área** do estudo lingüístico será abordada - Aquisição de Linguagem, Semântica, Morfologia, Sintaxe, Pragmática; qual o tipo de **unidade perceptual** será privilegiada nessa abordagem - se visual ou auditiva; como também, qual a **fase** do processamento lingüístico a ser focalizada - entre elas o **acesso**, a **representação** e a **organização** das informações lexicais, além dos **processos** extra e sub-lexicais. Detalhando mais abaixo cada uma dessas vertentes de pesquisa, o leitor poderá perceber, conforme dissemos anteriormente, quão diversas são as possibilidades da pesquisa psicolingüística. Por este motivo, voltamos a insistir na necessidade do pesquisador postular um caráter bastante objetivo e, de certa forma, pouco ambicioso para seu trabalho.

São diversas as possibilidades de uma abordagem psicolingüística em relação às áreas do estudo lingüístico, conforme mencionamos acima. Para especificar apenas algumas delas temos, por exemplo, na Aquisição de Linguagem, a aquisição dos padrões sonoros (fonético e fonológicos), escritos (alfabetização e letramento), em bebês e, respectivamente, numa fase mais tardia; na Sintaxe, a análise do processamento das anáforas, das frases encaixadas; na Semântica o processamento da metáfora, da ambigüidade e da homonímia; na Morfologia, a análise dos processos derivativos (seja na prefixação ou sufixação), flexionais (sejam eles nominais ou verbais) ou composicionais, além dos ajustes pragmático-discursivos.

O derradeiro¹, porém não menos essencial, recorte que faremos antes de iniciar a pesquisa propriamente dita, diz respeito à **natureza** do material com o qual iremos trabalhar

¹ As várias etapas compreendidas no processo de reconhecimento e na organização das informações lexicais serão descritas na explicação dos modelos de processamento, nos próximos sub-itens desse capítulo, por assim serem mais facilmente exemplificadas e definidas.

daqui para frente; ou seja, da mesma forma que temos que delimitar se nossa abordagem centrar-se-á na produção ou na compreensão da Linguagem, temos que privilegiar a modelidade de apresentação do estímulo que dará origem ao processo de reconhecimento. Se este impulso for **sonoro**, estaremos lidando com o chamado *input* auditivo e, por conseqüência, com todos os procedimentos relacionados aos processos de reconhecimento fônico, adequando nossa busca aos modelos, teorias e experimentos psicolinguísticos que compartilhem dos parâmetros necessários para a pesquisa do estímulo fonético-acústico. Por outro lado, se a natureza do impulso for **visual**, estaremos trabalhando com outra fonte de estímulo perceptual, o chamado *input* visual. Portanto, ao lidarmos com um material não mais sonoro, mas desta vez escrito, teremos outra fonte de dados para a qual será necessário buscar outro modelamento, principalmente no que concerne ao trabalho experimental. Pois, como poderemos ver logo mais, a maioria dos modelos teóricos de processamento que atentam para o acesso visual, contam também com um processador para o *input* auditivo, diferentemente dos modelos de reconhecimento de *input* sonoro, que na maioria das vezes não incluem o reconhecimento visual em seu modelamento.

Logo, o que pretendemos com todo esse mapeamento das possibilidades de pesquisa neste campo é explicar previamente alguns tópicos do nosso objeto de estudo - o **acesso visual das palavras prefixadas da Língua Portuguesa**² (LP) - como também, deixar claro que não ambicionamos dar conta de todo o processamento linguístico humano. Optamos, então, somente pelo processamento **morfológico** de alguns prefixos do Português e estamos propondo-nos a apresentar-lhes este assunto através de um panorama teórico que se inicie pela descrição de alguns modelos dos iniciais, como o *Autonomous Search Model* (Forster: 1976, 1979), o *Prefix-stripping Model* (Taft & Forster, 1975) e o *Morton's Logogen Model* (Morton: 1969, 1970, 1982), os quais evidenciam um caráter mais fortemente marcado

² Como objetivamos tratar apenas da língua Portuguesa, especialmente do Português brasileiro, não estamos contando aqui com alguns processos morfológicos concatenativos e não-concatenativos mais complexos como a *circunfixação* (no qual, o morfema é inserido “em torno” do radical; p. ex., *ge-wandel-d*, participio passado do verbo holandês *waldelen*, “andar”) e a *infixação* (processo no qual o morfema - infix - é inserido “dentro” do morfema “principal”, por exemplo, *sulat* “um escrito” - do Tagalog (Austronesiano)- tornar-se-á *s-in-ulat* “que foi escrito”). Para uma melhor explicação do modelamento destes e de outros exemplos de processos morfológicos não-concatenativos morfológicos, ver: Baayen, R. H. & Schreuder (1995) in: Dijkstra, T. & Smedt, K., *Computational Psycholinguistics: AI and Connectionist Models of Human Language Processing*.

pelos conceitos teóricos da Lingüística; chegando até outros modelamentos mais recentes, como o modelo *Augmented Addressed Morphology* (AAM) de Caramazza et alli (1988), cuja abordagem traz discussões sobre a ótica de um léxico organizado a partir de propriedades inerentes às palavras, as **propriedades distribucionais** (Laudanna & Burani, 1995). Traremos, em seguida, um capítulo intermediário com uma breve descrição de vários experimentos psicolingüísticos, com maior atenção para a modalidade que escolhemos para testar a teoria acima mencionada - os experimentos baseados nos efeitos de *priming semântico*. Este estudo experimental, que em sua totalidade constitui-se por dois experimentos, foi realizado no LFAPE/IEL (Laboratório de Fonética Acústica e Psicolingüística Experimental) durante os anos 1999 e 2000 e teve como sujeitos adultos com - no mínimo - um grau médio de escolaridade, todos graduandos ou já graduados em algum curso universitário. Incluiremos ainda a descrição do programa utilizado para a realização da tarefa experimental (Expe6), o material selecionado para compor as baterias de testes; além de uma extensa explicação da metodologia empregada para a elaboração do experimento em si; trata-se portanto de um capítulo sobre materiais e metodologia de pesquisa.

Finalizaremos o texto apresentando os dados colhidos a partir da abordagem experimental acima citada, juntamente com uma análise sobre sua significância estatística e o que isto pode representar quando se pensa no acesso e na organização lexicais.

B. Alguns conceitos básicos

Iniciaremos a nossa primeira seção procurando esclarecer ao leitor a noção de léxico na Psicolinguística. Em seguida, passaremos à exposição e conceitualização dos modelos de processamento lexical na seguinte ordem: primeiramente o **Modelo de Busca** (Search Model) incluindo-se neste tópico a distinção entre Modelo de Busca **Serial** e Modelo de Busca **Paralela** - ambos sob a premissa básica de que há um exame do léxico em busca da informação desejada. Posteriormente, apresentaremos uma vertente diferenciada pelo modo de acesso da informação lexical - os **Modelos de Ativação** - para os quais se propõe que a informação seja ativada a partir do momento em que o léxico é “acionado” para o processamento. Este modelamento tem uma certa afinidade com os chamados modelos **Conexionistas** de processamento, espécie um tanto mais complexa na sua conceitualização, porém não menos profícua teoricamente. Por fim, ao definir o Modelo **Augmented Addressed Morphology** - doravante AAM - (Chialant et Caramazza, op. cit.), apresentaremos para o leitor o que são as **propriedades distribucionais** e de que forma elas podem trazer uma nova perspectiva sobre a informação morfológica na organização e no acesso lexicais quando relacionadas a um modelo como o AAM.

C. O que é o Léxico afinal?

O termo *léxico* pode ser definido para algumas perspectivas do estudo lingüístico tal qual uma lista, um dicionário de palavras de que dispõe um falante/escritor para a sua comunicação, ou ainda, como um vocabulário específico a uma certa área de trabalho ou de estudo. Para algumas teorias morfológicas, o léxico pode ser caracterizado como uma “prisão” (ver Di Sciullo and Williams, 1987), pois nele estaria tudo aquilo que desobedecesse às regras morfológicas (inspirada talvez em Bloomfield (1930) que via o

léxico como um local das idiossincrasias da linguagem); já Aronoff (1976)³ trazia, sob perspectiva gerativista, um léxico como uma lista constituída pelas três classes maiores de palavras: os Verbos, os Adjetivos e os Nomes, juntamente com as chamadas WFRs (Word Formation Rules), que tornariam o falante capaz de fazer inferências na produção e interpretação de palavras morfologicamente complexas.⁴

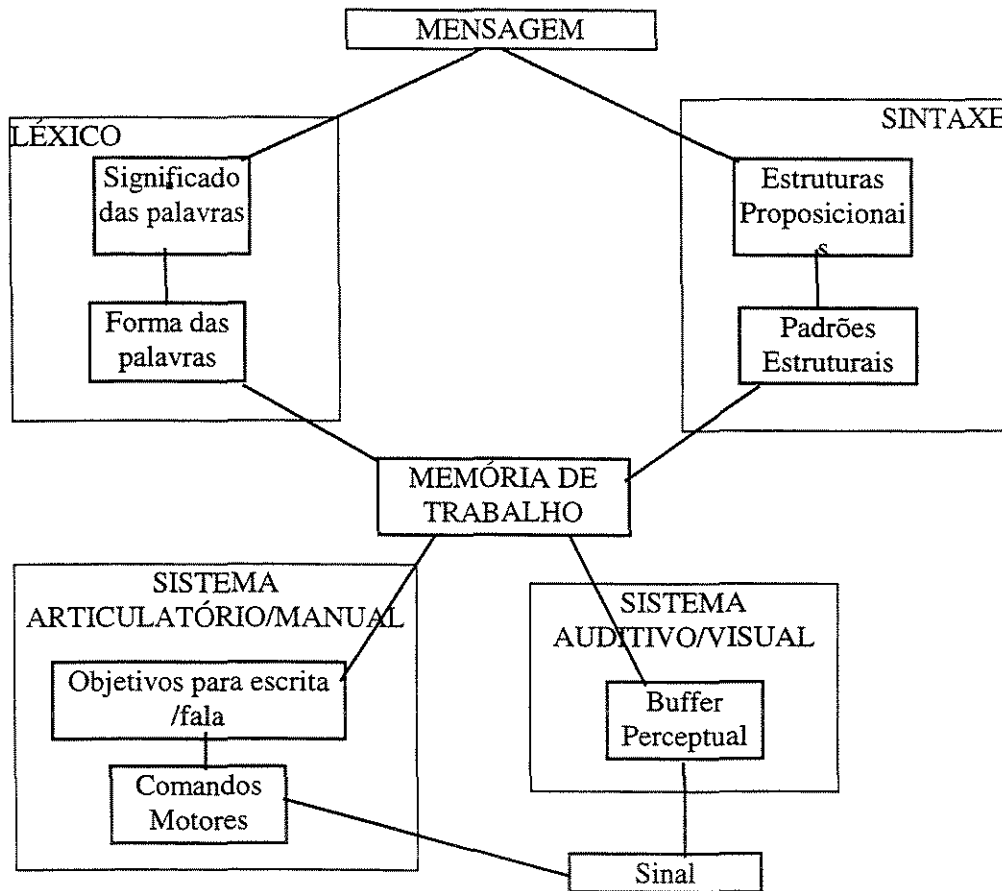
No entanto, a abordagem privilegiada pelo estudo psicolinguístico vê o Léxico como parte de um sistema de processamento mental da linguagem, seja na produção da fala ou da escrita, seja na decodificação de uma mensagem sonora ou visual. É através desta unidade do processamento, que funciona como uma espécie de dicionário mental, que podemos acessar⁵ e reconhecer as palavras de uma língua. Por meio desses processos de **acesso** e **reconhecimento lexical** é que se efetua a procura e seleção de itens no Léxico, com a finalidade de recuperar uma palavra que possa ser identificada com a informação dada pelo *input*. Logo, o conceito de **acesso lexical** está relacionado ao resgate, no Léxico, das informações perceptuais (dados provenientes da forma da palavra) e contextuais (dados provenientes no discurso ou texto), como também de objetos do mundo real e de formas de conhecimento partilhado de uma palavra; já o **reconhecimento lexical** pode ser entendido como o estágio que se alcança quando há somente um candidato restante, sendo que o *input* que deu origem ao processamento já foi identificado. Temos logo abaixo um exemplo didático das estruturas que podem ser componentes do processamento da linguagem; assim o leitor leigo poderá visualizar uma das possíveis localizações do léxico no processamento da mensagem linguística.

³ A noção do léxico, como um módulo independente no processamento linguístico, tem sua origem nas idéias propostas em *Remarks on Nominalization*, Chomsky (1970), obra na qual, pela primeira vez, sugeriu-se que o poder da Gramática Gerativa não está somente nas suas transformações; cada sistema pode ser gerido por regras interpretativas e princípios simples. Esta perspectiva do léxico influenciou claramente a teoria de formação de palavras proposta por Aronoff (1976).

⁴ As regras que nos permitiram entender em um certo momento, que um certo Ministro da República brasileira não era tão “imexível” quanto se julgava ser.

⁵ Apropriação do verbo inglês *access* (entrar, ter acesso), encontrada no jargão computacional, e que vem se disseminando na Língua Portuguesa.

Figura 1 - Componentes do processamento da linguagem.



(adaptado de Garman 1990: 182 - figura 4.1)

Será perceptível durante o texto que a concepção de léxico mental, assim como os conceitos de acesso e reconhecimento lexical, podem variar conforme a teoria que os descreve. Por conseguinte, os modelos teóricos tendem a se diferenciar pelo tipo de informação contida no léxico (morfológica, fonológica, semântica, sintática), pela forma de representá-las e organizá-las (palavras inteiras, constituintes morfo-fonológicos e de significação) e principalmente pela forma de acessar suas informações (modelos de busca x

modelos de ativação, modelos paralelos x modelos seriais).

Entretanto, antes de iniciarmos a descrição mais aprofundada de qualquer modelo de processamento lingüístico, se faz necessário ressaltar alguns fatores que não podem ser esquecidos durante um estudo desse cunho: primeiro, o léxico de um adulto com um grau médio de escolaridade é muito grande e nada estático, ou seja, tende a aumentar com a aquisição de novos termos. Segundo fato, todas as etapas deste processamento (percepção do estímulo nervoso, sua decodificação em informação lingüística e o acesso e reconhecimento desta última) ocorrem a uma velocidade muito alta. O leitor também deve ser avisado de que, ao direcionarmos nosso enfoque para o acesso lexical das palavras - portanto apenas à parte a etapa que antecede ao reconhecimento - estamos levando em conta que os fatos ligados ao processamento de informações semânticas e sintáticas podem constituir passos exteriores a esta fase.

D. Modelos de Processamento da Informação Lexical

Apresentaremos nesta seção, em uma ordem quase que cronológica da evolução do estudo, vários modelos de processamento da informação lexical. Iniciaremos por perspectivas, conforme o leitor poderá perceber, que se opõem em extremos quando se trata das linhas teóricas de estudo do acesso e representação das informações lexicais. É perceptível também, que a progressão das técnicas de pesquisa experimental e das teorias sobre a Linguagem e a Psicolingüística estão conduzindo o estudo dessa representação lexical para perspectivas teóricas dotadas de menos radicalismo.

1. Os Modelos de Busca

Como a própria denominação já indica, essa classe de modelos de processamento caracteriza-se por assumir que as palavras são recuperadas da memória através de processos

de *busca* pela informação requisitada. Esse resgate da informação ocorre de forma indireta sobre o Léxico, pois a busca pelo item lexical é intermediada por unidades de acesso lexical constituintes de um Léxico também intermediário e com características semelhantes a um índice. Ou seja, diferentemente de um dicionário, no qual, ao procurar uma palavra encontraremos também todas as suas informações de significado, classe gramatical, sinônimos, etc., temos, nos modelos de busca, um componente cuja função é remeter ao vocábulo propriamente dito, tal qual um índice, o **léxico de acesso**. Neste componente estão armazenadas apenas códigos de acesso às palavras, de maneira que sua função é viabilizar o acesso ao item lexical, tendo em vista um processo de busca que foi acionado a partir de um estímulo visual ou auditivo e que deve levar ao reconhecimento lexical. Para tornar este procedimento teoricamente mais eficiente, uma vez que um grande número de itens armazenados no Léxico de acesso está envolvido, os teóricos procuram propriedades organizacionais destes códigos que ajudem a acelerar o processo de busca.

O “Modelo Autônomo de Busca” elaborado por Kenneth Forster (1976, 1979) pode ajudar-nos a ilustrar melhor os modelos que adotam esta perspectiva. Constituindo-se basicamente de um sistema de processamento lingüístico e um subsistema que trata do reconhecimento visual das palavras, este modelo evidencia um caráter modular, pois assume-se que cada uma das etapas do processamento lingüístico independe da outra. Ou seja, para Forster (1976, 1979, 1995), o módulo dedicado ao acesso das informações relativas à forma da palavra não recebe qualquer influência relacionada aos processamentos de ordem semântica ou sintática; logo, cada tipo de informação lingüística é legada ao seu processador específico. Em resumo, esta proposta é compatível com o conceito de modularidade da mente de Fodor (1983).

Desta forma, o sistema de reconhecimento de palavras se subdivide, primeiramente, em componentes ou arquivos de acesso periféricos (*access files*), dentre os quais temos: um componente de **acesso ortográfico**, que lida com as palavras escritas, portanto com o *input* visual, um segundo componente responsável pelo **acesso fonológico**, o qual lida com o *input* fonético-acústico, portanto com as informações provenientes da fala, e um terceiro componente responsável pelo tratamento da informação semântica. A união destes três

componentes forma o chamado léxico de acesso, sobre o qual se dá, através de um mecanismo de busca, a primeira etapa do reconhecimento lexical. Segundo Forster (1995), cada arquivo de acesso consiste de um conjunto de índices e de **códigos de acesso**, que levam às suas respectivas entradas lexicais. A forma de representar esse código de acesso difere para cada um dos componentes de acesso; no arquivo de acesso ortográfico, a representação seria algo relacionado à ortografia da palavra, por exemplo, suas primeiras letras; já para o código de acesso fonológico, a especificação corresponderia a um aspecto fonológico da palavra. Seguindo esta “linha” representacional, teríamos para o código de acesso semântico algo que especificasse a categoria semântica à qual a palavra pertence (p.113). O componente que contém a “verdadeira” entrada lexical, ou seja, onde o item está completamente representado, com todas as informações sobre sua ortografia, fonologia, semântica e sintaxe é chamado **Léxico Principal**.

Cada um destes componentes de acesso é subdividido em unidades de armazenamento cuja função é estocar a informação lexical de acesso; estas unidades podem ser visualizadas como espécies de pequenos escaninhos (*bins*). Propõe-se, pela teoria, uma organização interna a estes escaninhos com a finalidade de tornar a busca pelo item lexical mais rápida e eficiente. O principal critério organizacional adotado por Forster nesse modelo é a chamada **freqüência vocabular**⁶ - termo que pode ser entendido como um efeito psicolinguístico condicionado ao número de vezes que uma palavra é utilizada em um determinado contexto ou corpus linguístico. Por exemplo, palavras como *bola, carro, escola, jornal, telefone, criança, dinheiro, saúde* entre diversas outras são, provavelmente, muito mais freqüentes em um Léxico-padrão, ou seja, de um brasileiro adulto com um nível médio de escolaridade, do que termos como *noética, procumbir, friável, sobrepeliz, ventrudo* entre outros⁷. Portanto, Forster propõe que os termos estejam estocados nos escaninhos por ordem decrescente de freqüência de ocorrência de cada palavra; ou seja, sob uma ordenação tal que as palavras mais utilizadas pelos falantes possam estar mais

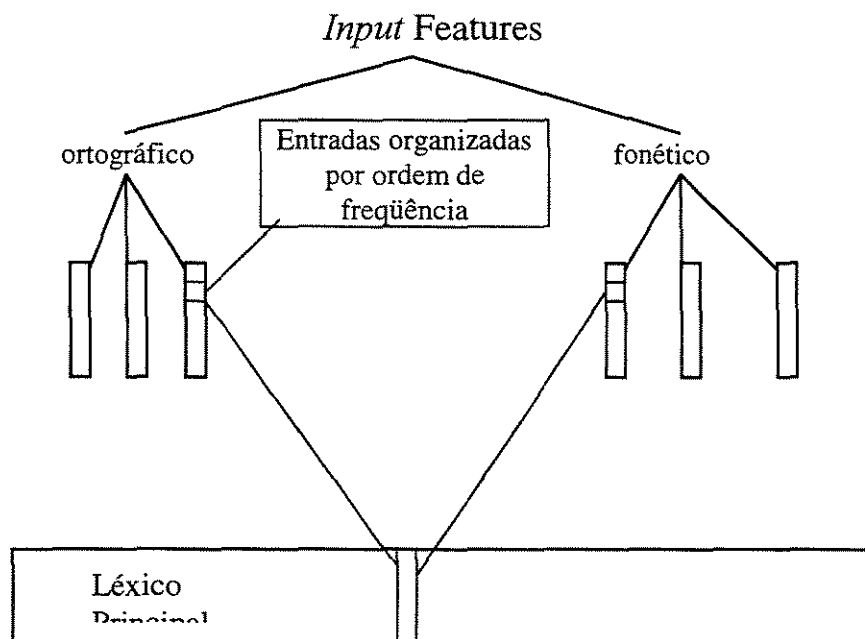
⁶ A freqüência é um conceito básico para toda a pesquisa psicolinguística, pois faz parte da organização de vários modelos de Léxico e serve como critério de elaboração e análise para muitos experimentos da área.

⁷ Nenhum destes termos caracteriza-se como uma não-palavra, apenas são raros no nosso cotidiano, mas nossos dicionários estão repletos deles. Estes aqui apresentados foram retirados do *Dicionário Básico da*

prontamente acessíveis quando forem requisitadas. Para decidir o que é um termo de alta ou de baixa frequência, o psicolinguísta costuma recorrer à bases de dados que computem informações sobre a quantidade de ocorrências de um termo em um determinado contexto (textos de jornais, livros, revistas, etc.).

Apresentaremos agora uma analogia interessante do modelo de Forster (1976), para que, desta forma, o leitor se situe um pouco melhor na compreensão deste processo de organização e busca pelas informações lexicais. Garman (1990) imaginou o *Modelo de Busca Autônoma* como uma biblioteca. Vejamos: ao chegar em uma biblioteca “ideal” à procura de um determinado livro, não é comum que se vá direto às suas prateleiras, pois estas além de numerosas costumam ter um número muito grande de livros, dificultando a pesquisa. O método mais eficaz de encontrarmos a obra procurada é ir ao catálogo, que usualmente tem uma organização baseada em um parâmetro específico - por exemplo, a ordem alfabética do título ou do sobrenome do autor - mas não encontraremos a obra realmente ali; o que teremos à disposição é um código que nos levará ao livro que procuramos, uma “marca de localização abstrata”. Nesta biblioteca ideal, pode-se imaginar ainda, que as prateleiras com os livros mais consultados estejam dispostas de forma mais acessível para o leitor. Temos em vista, portanto, um processo de busca que se dá de forma sucessiva e só termina a partir do momento em que se encontra o livro procurado.

Figura 2 - Modelo Autônomo de Busca de Forster (1976, 1979)



(adaptado de Forster, 1979)

Assim é o processo de busca proposto no modelo de Forster (1976); uma busca que, por possuir um caráter seqüencial, examinando apenas uma alternativa de resposta por vez, pode ser classificada por **serial**. Torna-se conseqüência de um procedimento de busca como este, que a organização das informações lexicais com base no conceito de freqüência vocabular venha a ser um ponto crucial para a estruturação do modelo. Desta forma, temos que nos modelos os quais compartilham do conceito de organização por ordem de freqüência de uso da informação nos arquivos de acesso, as palavras mais freqüentes são buscadas antes das palavras de baixa freqüência, em um procedimento que se inicia a partir do momento em que o mecanismo de busca é acionado durante o estágio do acesso lexical. Esta busca é classificada como *auto-terminante* (*self-terminating*), se o procedimento se encerrar assim que o termo buscado for encontrado, e como *exaustiva* se houver a necessidade de se percorrer a lista toda mesmo após a obtenção do termo.

No *Prefix Stripping Model*, uma versão do modelo elaborada no ano de 1975⁸, Forster, juntamente com Marcus Taft, apresentou uma hipótese de decomposição morfológica, na qual as entradas do léxico de acesso estariam estocadas não apenas por ordem de frequência, como também representadas na forma de morfemas, fossem eles raízes ou prefixos. Assim, a entrada lexical para uma palavra morfológicamente complexa (com mais de um morfema) seria uma descrição ortográfica do radical da palavra. Por consequência, todo *input* formado por mais de um constituinte morfológico seria segmentado em um processo pré-lexical e analisado em unidades de representação separadas no léxico de acesso. Deste modo, uma palavra como *incansável* seria analisada como *in-*, *cans-*, *-vel*, em um pré-processamento, *amoral*, como *a-* e *moral*.

Até este ponto o modelo não geraria tantas questões, mas o que aconteceria caso uma palavra como *referir* tivesse que ser acessada? O processo de extração de morfemas faria com que a busca se desse sobre o verbo *#ferir#* como um constituinte da palavra, ao invés da forma presa *-ferir*? Lembre-se que a mesma base pode ser encontrada nas palavras *inferir*, *interferir*, *conferir*, por exemplo. Logo, o que se observa é que o modelo não conseguiria dar conta de uma classe especial de palavras prefixadas, aquelas que causam uma espécie de “ambigüidade” no processamento morfológico, pois, não há nenhuma informação sobre qual seria a solução de um caso como esse, ou de como ambas as formas estariam alocadas para que não houvesse um problema desta ordem.

Apesar disto, este modelo empolgou a comunidade psicolinguística particularmente porque foi o primeiro a incluir em sua estruturação fatores de natureza morfológica, além de atribuir-lhes um caráter pré-lexical, de maneira que o processo de extração dos morfemas sempre ocorreria antes de as palavras morfológicamente complexas serem acessadas, acarretando um acesso exclusivo aos constituintes morfológicos. Além disso, acreditava-se que o modelo de *Prefix Stripping* teria um caráter mais econômico no que diz respeito ao seu poder representacional, o que supostamente significaria uma economia da carga de memória utilizada para o processamento linguístico. Contudo, seria necessária a elaboração de regras que coibissem certas análises, como no caso da palavra *referir*, acima citada, e

⁸ Apresentada no artigo “Lexical storage and retrieval of prefixed words”. *Journal of Verbal Learning Behavior*, 14, 638-

que promovessem todas as outras concatenações legais do ponto de vista do processo de formação de palavras, causando talvez, um aumento no tempo total de processamento. Ou seja, neste caso, a economia representacional não significa uma economia no tempo de processamento; pelo contrário, o procedimento para o acesso e reconhecimento, não apenas de palavras morfologicamente “ambíguas” como aquelas mencionadas acima, mas também para o reconhecimento de novas formas, seria bastante complicado causando um aumento no tempo final de processamento. Comprometer-se ia até a modularidade do modelo; posto que, para resolver as ambigüidades e a necessidade de compreensão das novas formas, seria necessário recorrer aos outros módulos de acesso, como o semântico e o sintático e talvez até a outros mecanismos cognitivos, para verificar a aceitabilidade das formas envolvidas no processamento.

Um outro problema decorrente desta mesma questão representacional vem à tona quando tratamos das bases “presas⁹” como *-gredir*, constituinte das palavras *regredir*, *progredir* e *transgredir*; que, apesar de não trazerem nenhum caráter ambíguo, morfologicamente falando, exigiriam uma representação separada no léxico, ou de regras especiais de concatenação de morfemas. Apresentando uma faceta ainda mais complicada, se pensarmos em um vocábulo derivado de *progredir*, como *progressão*. Como estaria representada esta forma e de que maneira se daria a busca deste termo - já que no modelo aqui descrito há uma premissa fundamental de análise segmentada da palavra? Teríamos uma representação particular para formas como *-gress-*? Estes dados também não se encontram detalhados na descrição do modelo, mas há uma opção de representação existente no modelo que tenta dar conta de algumas destas questões, não abandonando porém a idéia de percepção dos constituintes das palavras. Os autores atentaram para a possibilidade de uma representação total da palavra na memória na qual os prefixos estivessem apenas indicados, ou seja, teríamos uma representação como *re(gredir)*. Mas será que uma representação como essa resolve as questões pendentes sobre o modelo? Afinal, esta forma de se representar o item lexical não parece descartar a prévia análise

647.

⁹ Termo usado em oposição ao termo *forma livre*, que tem como exemplo qualquer forma que possa representar uma unidade de significação por si só, ex.: *gostar*, *querer*.

segmentada do termo morfológicamente complexo. Uma hipótese mais plausível para o acesso e a representação de palavras morfológicamente complexas e de bases presas como *progresso*, *regressão*, entre outras, propõe que termos como estes possam ser acessados e representados de forma integral no Léxico. Esta hipótese será discutida posteriormente, na apresentação do Modelo AAM (Caramazza, Laudanna & Romani, 1988).

Desta forma apresentamos várias evidências de uma capacidade limitada para os processos de busca no reconhecimento de palavra, principalmente pela possibilidade de um acesso único às alternativas de resposta. Há, porém, modelos de processamento que admitem, na possibilidade de um acesso **paralelo**, um processamento que analise várias alternativas de uma única vez, optando por aquela que melhor se “encaixe” no padrão da informação. Imaginemos, então, um modelo de busca paralela que tivesse como base de busca os morfemas¹⁰ de uma língua e a palavra a ser acessada fosse *abrindo*. O modelo poderia, num primeiro momento, buscar todas as palavras que fossem formadas pelo morfema *abr-*, como *abrir*, *abridor*, *abrirá*, *abriste*, *abrindo*, *abre* e assim por diante. O sistema de reconhecimento terminaria sua tarefa após examinar todas as possibilidades e encontrar aquela que melhor se adequasse ao termo. Normalmente, encontraremos este procedimento de acesso à informação como parte do processamento lexical nos modelos que serão o assunto da próxima seção: os modelos de **acesso direto** ou de **ativação**.

¹⁰ Levando em conta que neste momento, a título de exemplificação, *morfema* signifique menor unidade de significação, cuja a forma não possa ser dividida em mais constituintes.

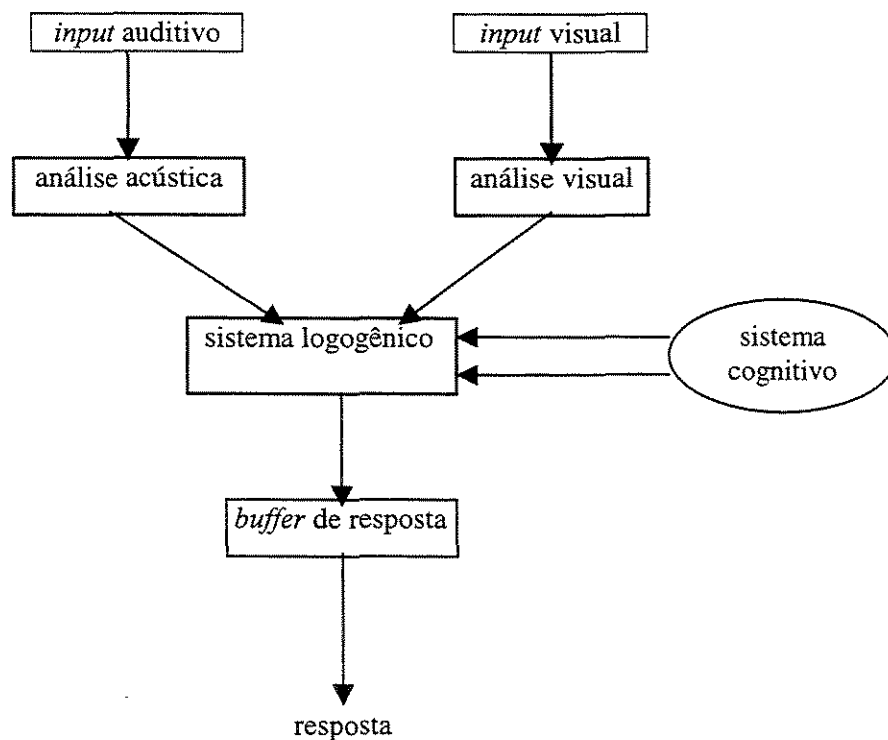
2. O Modelo de Ativação de Morton (1970)

Sob esta perspectiva dos **modelos de ativação**, acessar as palavras durante um processo de compreensão, seja de natureza visual ou auditiva, constitui-se em um trabalho de mapeamento dos sinais lingüísticos em representações de significado; dando-se, desta forma, o início da compreensão. Trata-se, portanto, de um modelamento de acesso direto às entradas lexicais que se baseiam em “unidades detectoras de palavras”. Veja que nesse caso não há um código abstrato de acesso que leve à representação da palavra propriamente dita; propõe-se que cada letra ou fonema provoque uma ativação nas palavras que a contém. Por exemplo, quando uma palavra como *bolacha* nos é visualmente apresentada, temos sob esta perspectiva que ao lermos a letra *b* serão ativadas *bolo*, *bolacha*, *boneca*, *botão*, *bala*, *batata*, *biscoito*, *batalha* e assim por diante; mas quando *b+o* são lidos, apenas *boneca*, *bolacha*, *botão* e *bolo* continuam recebendo mais energia para se manterem ativadas. O processo segue-se assim até que a identificação completa se dê, com apenas uma palavra permanecendo com o nível mais alto de ativação entre as concorrentes, e portanto, tornando-se a única possível de ser recuperada na memória.

Um modelamento pioneiro que leva em conta a ativação para o reconhecimento das palavras é o *Modelo Logogênico* de Morton (1970). Neste caso, cada palavra no léxico possui o seu detector, ou seja, uma unidade de ativação independente que monitora o sinal de *input* (seja auditivo, seja visual). Nesta unidade, chamada por Morton de **logogene**, estão representadas as informações de ordem ortográfica, fonológica, sintática e semântica de cada item lexical. Conforme exemplificamos no parágrafo acima, mais de um logogene pode ser ativado pelo mesmo *input*; entretanto, apenas aquele que alcançar um certo limite de ativação para o reconhecimento será liberado. A informação do logogene segue para uma unidade que é dedicada a transformá-la em um *output* - o **response buffer** - e os limiares de ativação voltam a zero. Porém, por alguns momentos, os limites de ativação permanecem maiores que zero para uma possível reutilização do termo. A informação enviada ao *buffer* de resposta só estará disponível em forma de um *output* após um processo de checagem

elaborado pelo **Sistema Cognitivo** - componente no qual se encontram todas as informações que não estão relacionadas ao conhecimento lexical como as sintáticas, discursivas, pragmáticas, etc. Veja a esquematização do modelo logogênico apresentada abaixo.

Figura 3 - A arquitetura do Modelo de logogênico adaptada de Morton (1970)



Os pontos mais atraentes do Modelo Logogênico de Morton (1970) estão ligados à sensibilidade aos efeitos de frequência vocabular e à sua capacidade interativa. Ou seja, diferentemente de alguns modelos de acesso direto, os limites de ativação das unidades logogênicas para que o reconhecimento ocorra são sensíveis aos efeitos de frequência de ocorrência da palavra; dado que os logogenes de palavras mais frequentes têm limiares de ativação mais baixos, logo ficam “pré-dispostos” a serem acionados. Além disso, dizer que

o modelo é interativo significa atribuir-lhe uma capacidade de lidar concomitantemente com as diferentes áreas do conhecimento lingüístico; ou de outra forma, pode-se dizer que os níveis de ativação dos logogenes são afetados por “expectativas” geradas pela estruturas sintático-semânticas do *input*. Desta forma, se uma palavra for sintaticamente consistente com o *input*, uma quantidade menor de informação fonético-acústica será exigida para o seu reconhecimento do que se esta mesma palavra fosse apresentada isoladamente (ver Lively, S. E. et al.: 1994; p.282). Isso indica que, quando relacionadas semântica, sintática ou morfológicamente, as unidades logogênicas têm seus limiares de ativação diminuídos para a tarefa de reconhecimento. Por exemplo, ao apresentarmos a palavra *escola*, o logogene correspondente a *professor* fica mais propenso a ser ativado, ou ainda que a palavra *dançar* se torne mais disponível quando apresentada sob o contexto sintático de *Uma bailarina leva a vida a...*, possibilitando um reconhecimento mais rápido do que se fossem apresentadas isoladamente. Esta última característica lega ao modelo de Morton (1970) um caráter semelhante aos modelos ligados ao Conexionismo, que têm como premissa básica - como o próprio nome já evidencia - a conexão entre as unidades de processamento. Contudo, se o leitor puder se lembrar, o modelo em questão diverge bastante dos modelos de busca como aqueles elaborados por Forster, que privilegiam um caráter independente dos módulos de processamento tornando-se mais eficientes para o processamento da palavra isolada.

Apesar de uma “arquitetura” despojada e simples, o modelo Logogênico de Morton apresenta alguns problemas na sua projeção, por exemplo, não há nada no modelo que especifique quais são as unidades perceptuais utilizadas para o mapeamento do *input* até os logogenes da memória; também não há nenhum tipo de explicação da forma como as informações lingüísticas são integradas nos logogenes. Além do mais, o modelo não dá conta dos chamados efeitos de **similaridade lexical**, efeito psicolingüístico que reflete as relações entre palavras foneticamente parecidas - é o caso, por exemplo, das palavras *sessão* e *seção* que são ditas vizinhas auditivas mas não visuais, ou de [‘sɛdə] e [‘sedə], vizinhas visuais mas não auditivas, ou ainda de *cedo* (forma verbal de ceder) e *cedo* (advérbio) vizinhas visuais e auditivas; pois, em vista desse efeito, ao ativar de forma parcial vários logogenes independentemente, não há mecanismo algum que iniba altos níveis de ativação

dessas unidades “vizinhas” àquela que serviu como *output*.

Podemos questionar também o que aconteceria no caso da produção, se em uma concorrência entre duas palavras pela recuperação de um termo na memória, uma delas fosse muito mais freqüente do que a outra, portanto com menor limiar de ativação? Por exemplo, em relações de eufemismo como *morte* e *passamento*; supondo que a palavra a ser recuperada fosse aquela de menor freqüência, será que o logogene da palavra *morte* estaria disponível mais rapidamente mesmo não sendo esta a opção requisitada? Estas informações não estão especificadas no modelo. Voltando ao acesso e reconhecimento, especula-se que mesmo após a identificação da informação requerida ao logogenes, as opções que foram ativadas no processamento paralelo continuam com seus níveis de ativação salientes, podendo representar um gasto desnecessário de energia.

Seguiremos agora com a apresentação de um modelo Conexionista elaborado por McClelland, Rumelhart & PDP Group (1981).

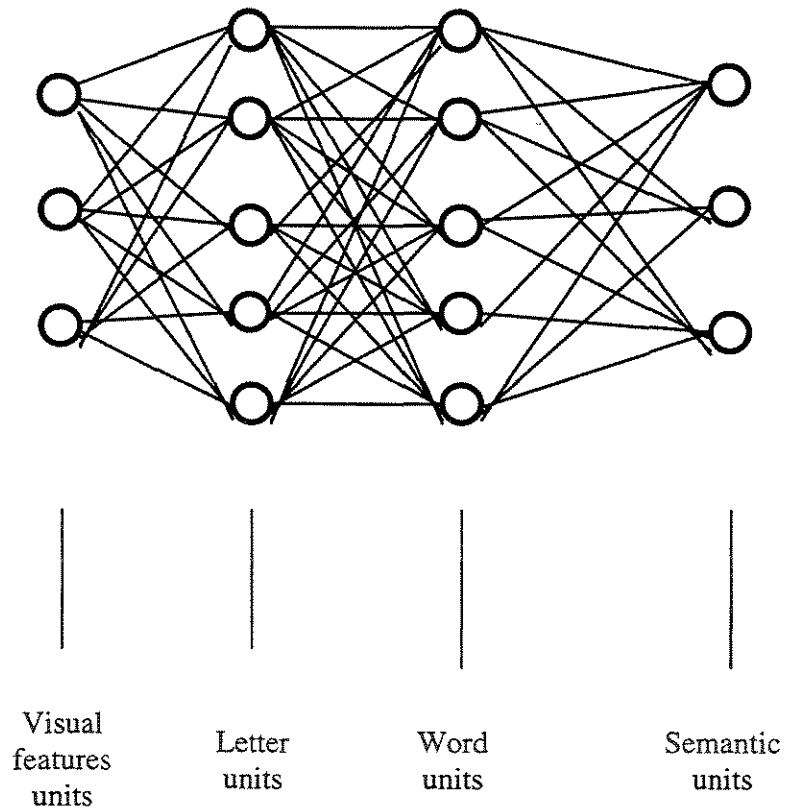
3. Um Modelo Conexionista para o Reconhecimento

Inserir uma abordagem conexionista na pesquisa lingüística ainda é um trabalho muito delicado, mesmo que estejamos apenas descrevendo um modelo delineado para o reconhecimento visual de palavras. Isto se deve ao fato de que o estudo conexionista tem uma abrangência que vai além das fronteiras da pesquisa psicolingüística, usando-se de conceitos relacionados a áreas como a Neurociência, as Ciências Cognitivas, a Computação, a Física, pois pela via conexionista postula-se que o processamento da informação ocorre por meio da interação entre um grande número de unidades de processamento em um mecanismo semelhante aos mecanismos de sinapses que ocorrem entre os neurônios. Além disso, sua forma de lidar com a informação espelha-se em conceitos computacionais, trazendo uma visão um pouco diferente no que concerne às etapas de um processamento lingüístico. Assim, o trabalho de um lingüista, ao entrar em contato com esta área, tem de ser mais cauteloso pelo fato de que, para recuperar o que

particularmente lhe interessa, necessita ter conhecimento suficiente sobre aquilo que está deixando de fora do seu recorte. Não ambicionamos, desse modo, fazer uma descrição detalhada dos diversos modelos existentes; o que traremos aqui é, novamente, um enfoque muito específico inserido em um delineamento das características mais gerais do modelo de McClelland, Rumelhart & PDP Group (1981).

Conforme foi dito na etapa anterior, o modelo escolhido para exemplificar o reconhecimento visual das palavras pela perspectiva conexionista tem em seu delineamento alguns elementos semelhantes àqueles encontrados no Modelo Logogênico que acabamos de lhes descrever. A primeira semelhança está na adoção de uma perspectiva que se utiliza igualmente de detectores de palavras em sua elaboração, porém estas unidades no presente modelo têm estrutura bastante diferente daquela idealizada para um logogene. Assim, o **Modelo de Ativação Interativa** proposto para o reconhecimento visual por McClelland e Rumelhart (1986) consiste em uma **rede** de associações subdividida em duas **camadas** (*layers*) - constituídos por unidades detectoras de palavras ou traços - cuja função é a integração da informação a ser processada nestas unidades. Ou seja, neste modelamento os autores partiram da idéia de que existem unidades detectoras não apenas para palavras, como também para as letras e os traços visuais e de significado. Estas unidades encontram-se separadas em seus respectivos níveis de uma forma tal que é a maneira como a informação “transita” por estes níveis que dá a este modelo o aspecto de “rede”. Isto é, cada unidade está conectada a outras unidades não apenas em seu próprio nível de processamento, como também nos níveis adjacentes, criando um aspecto de malha. Além desse aspecto, o mecanismo de processamento da informação é pensado como um conjunto de processos simultâneos e contínuos, no qual o “produto” de um estágio é sempre parte integrante do estágio adjacente - este *modus operandis* conhecido como **processamento em cascata**, foi postulado no artigo de McClelland (1979). Temos a seguir uma representação do modelo:

Figura 5.a - Estrutura Básica do *Interactive Activation Model* adaptada de McClelland & Rumelhart (1981)



Estas conexões entre as unidades da rede podem variar quanto ao tipo de energia que liberam: se estão em seu próprio nível, inibem-se umas às outras; porém, se as conexões se dão entre os níveis, tendem a estimular a ativação entre as camadas. Outro ponto de variação entre as camadas está relacionado à força de suas conexões, que, durante a concorrência pela identificação entre duas unidades, pode privilegiar relações como frequência e contexto.

Na projeção de McClelland & Rumelhart (1986), como pôde ser visto na figura 5a apresentada acima, são quatro os níveis apropriados para o reconhecimento; sendo que cada um deles é capacitado para detectar uma parte de informação. Esses detectores são específicos para cada nível do processamento, sendo dispostos da seguinte maneira: na primeira camada temos as **unidades para traços** (*features*) **visuais** como uma espécie de identificador intermediário para a ativação das **unidades de letras**, as quais se encontram no nível superior e que, por conseguinte, ativarão as **unidades de palavras e unidades semânticas** que forem consistentes com a informação que vem passando de nível a nível. As conexões entre essas unidades possuem a capacidade de modificar a força de inter-relação conforme a “rede” se expande, “aprendendo” gradualmente a reconhecer as palavras impressas, posto que a rede foi delimitada pensando-se também na capacidade de aquisição de linguagem.

Conforme indicamos acima, o modelo conta ainda com processos inibidores de ativação em qualquer uma de suas camadas para que, deste modo, os padrões inconsistentes com a informação a ser reconhecida não sejam ativados. Estes processos ocorrem dentro de um mesmo nível e entre suas unidades detectoras, enquanto que os processos de ativação, se dão entre os níveis. A existência de fortes conexões de inibição entre duas unidades dentro de um mesmo nível pode, em caso de ativação de uma dessas unidades, implicar em uma carga inibidora de mesma proporção na sua contraparte. Por exemplo, supondo que as palavras *fogo* e *água* tivessem uma conexão inibidora como esta da qual estamos falando, então quanto mais a unidade de palavra *fogo* recebesse energia para ativação, mais energia da unidade para *água* seria retirada. Podemos identificar os processos internos a cada camada tal qual uma competição; pois de forma semelhante ao modelo Logogênico, neste

modelo conexionista as unidades têm que alcançar um limiar de ativação suficiente para que apenas uma seja ativada e o processo de reconhecimento progride - posto que outras unidades dentro de cada nível podem ser parcialmente ativadas. Assim, aquelas que “vencerem” a competição, conseguindo energia suficiente para ultrapassar o limiar de ativação, começarão a ativar as unidades do nível seguinte e inibir aquelas com as quais estiveram competindo dentro de seu próprio nível até esta etapa. Este procedimento segue até que o acesso ao termo desejado se dê como *output* final da operação.

Portanto, o reconhecimento visual das palavras, neste modelo, se inicia com ativação das unidades detectoras de traços através da apresentação de um estímulo visual de uma palavra, por exemplo, *gato*. Então, no nível das representações das letras, todas as unidades relativas a essas letras serão inicialmente ativadas, porém de um modo difuso, ou seja, G é ativado para caber em qualquer uma das quatro posições das letras de *gato*, assim como as letras A, T e O. Conforme o limite de ativação de cada letra para cada posição vai recebendo mais energia, a força de conexão entre as unidades referentes às letras G-A-T-O em outras posições, que não àquelas correspondentes à da palavra, será inibida. Isto é, G1 - identificável como “G na primeira posição” - permanece recebendo energia de ativação, enquanto G2, G3 e G4 vão sendo inibidas, repetindo o procedimento para A2, T3 e O4 e suas respectivas ativações para as outras posições. Sequencialmente, as conexões entre o nível das letras e das unidades de palavras passarão a receber energia de excitação nas unidades mais “aptas” ao reconhecimento do estímulo apresentado. Mas esta ativação “seletiva” da camada de unidades de letras não impede que palavras como *fato*, *mato*, *pato*, por exemplo, recebam também energia de ativação. Caso houvesse na Língua Portuguesa um substantivo como *cato*, por exemplo, este receberia mais ativação que as outras palavras semelhantes a *gato* pelo fato de que as consoantes *g* e *c* têm um padrão de articulação muito parecido.

Estes fatos se devem também à ocorrência de uma **ativação cruzada** (*cross-activation*) nos estágios iniciais do processo de ativação; de forma que todas as unidades, seja qual for o nível onde se encontrem, são parcialmente ativadas e conforme o processo avança essa ativação passa a ser mais seletiva. Forster (1995) nos aponta que “os efeitos

dessa ativação cruzada serão passados para o nível superior mais próximo”, afetando também a camada das unidades semânticas (lugar onde se encontram os lemas, para Levelt, 1989) e fazendo com que, por exemplo, supostos traços de *mato* sejam ativados juntamente com os de *gato*. O mecanismo de inibição das unidades detectoras do modelo pode ser visto, sob este contexto, como uma forma de se equilibrar os efeitos deste tipo de ativação.

Essa modalidade de ativação pode ser comparada com a **ativação espalhada** (*spreading activation*), mecanismo que serve de base para modelos conexionistas elaborados para a produção como o de Gary Dell (1986, 1988). Em ambos os mecanismos temos um mecanismo de ativação que se espalha através das unidades (ou nós); porém no caso desta última, este mecanismo se dá de maneira associativa; ou seja, se retomarmos ao nosso exemplo de estímulo visual *gato*, teremos que a palavra *cachorro* será mais facilmente ativada do que a palavra *açúcar*, pois as representações dos dois primeiros termos estariam conectadas através de um “atalho” associativo em termos de significação, ou um atalho semântico.

Essa hipótese de que um processo associativo pode ajudar na ativação das unidades de processamento está presente na descrição do tratamento do tempo perfeito do Inglês por Rumelhart & McClelland (1986). Este delineamento conta com uma carga maior de teorias linguísticas do que no modelo conexionista anteriormente apresentado, dado que se direciona para o tratamento de informações morfológicas na aquisição de padrões verbais, fazendo-o por caminhos associativos, não apenas no nível semântico, mas como um mecanismo que poder ser pensado para o nível lexical desde de o processamento da parte formal do vocábulo.

3.1 O tratamento do Tempo Perfeito do Inglês sob a perspectiva de Rumelhart & McClelland (1986)

Um aspecto ainda não abordado em nossa descrição refere-se à maneira como são tratadas as formas morfológicamente complexas no modelo PDP. No artigo *On learning the*

past tenses of English verbs (Rumelhart & McClelland, op. cit.), os autores se concentraram na aquisição do padrão do tempo passado de verbos regulares e irregulares da Língua Inglesa, tentando adaptar o modelo descrito acima a esse objetivo. Para tal finalidade, Rumelhart & McClelland trabalharam observando crianças durante os estágios de aquisição do padrão verbal de passado.

No artigo, são descritos três estágios básicos na aquisição do tempo perfeito em relação a quatro “tipos” de verbos. No primeiro estágio, as crianças ainda usam um número muito pequeno de verbos no passado, sendo que todos eles tendem a ser muito freqüentes e de maioria irregular; já no segundo estágio de aquisição, “emerge a evidência do conhecimento implícito das regras lingüísticas” (p.219), pois as crianças passam a usar um número muito maior de verbos no passado, porém sua maior parte é de verbos regulares. Além disso, elas tendem a “regularizar” todos verbos irregulares - mesmo aqueles já aprendidos - e estendem a aplicação da regra de passado regular à palavras desconhecidas adicionando-lhes um *-ed* final, se acreditarem que estas palavras servem para descrever ações. No terceiro estágio, as crianças começam a dar conta tanto dos verbos regulares como dos irregulares, com a tendência a aplicar a forma regular quando entram em contato com formas novas; porém esta é uma tendência que permanece até a idade adulta. Desta forma, a classificação dos verbos acaba por ficar atrelada à divisão desses estágios, ou seja, temos como ponto de partida os verbos que primeiro são aprendidos - que são em sua maioria verbos cujo o passado é irregular como *get, come, take, give, go*; os regulares; os irregulares adquiridos posteriormente e as novas formas verbais¹¹. Temos abaixo uma tabela ilustrando as características dos três estágios da aquisição do padrão do tempo perfeito do inglês:

¹¹ Rumelhart & McClelland alertam que a caracterização do processo de aquisição deste tempo verbal como um processo trifásico pode, erroneamente, levar a crer que este processo é estanque. Todavia, o processo de aquisição é gradual, há momentos de transição entre cada um dos estágios. Além disso, essas mudanças são diferentes de um estágio para o outro: entre a primeira e a segunda fase são breves, mas entre a segunda e a terceira fase costumam perdurar vários anos; particularizando o segundo estágio como uma etapa bastante instável do processo de aquisição.

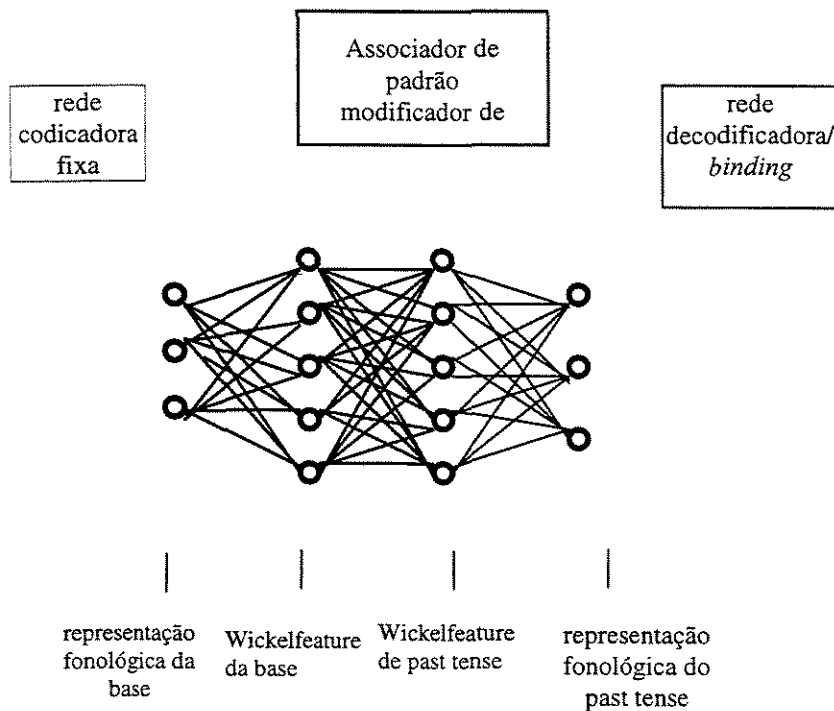
Tabela 1: CARACTERÍSTICAS DOS TRÊS ESTÁGIOS DE AQUISIÇÃO DO TEMPO PERFEITO NO INGLÊS

Tipo de verbo	Estágio 1	Estágio 2	Estágio 3
Verbos iniciais	Correto	Regularizados	Correto
Regulares	—	Correto	Correto
Outros irregulares	—	Regularizado	Correto ou Regularizado
Novos	—	Regularizado	Regularizado

(adaptado de Rumelhart & McClelland, op. cit., p.221)

O fato de a inserção da marca de tempo passado ser adotada como regra para palavras novas leva os autores a tomá-la como uma unidade possível de ser representada no modelo McClelland & Rumelhart (1981). Tomemos como base a figura 5a, na qual ilustramos a estrutura do modelo através de unidades organizadas em níveis de processamento, levando em conta porém, que estamos tratando de um caso específico, no qual as unidades de conexão da rede são correspondentes a apenas uma parte do processamento. Vejamos a ilustração:

Figura 6 - Estrutura básica do modelo para a aprendizagem do tempo perfeito



(adaptado de Rumelhart & McClelland, op. cit., p.222)

Nesta representação mais localizada do modelo, podemos notar que a estrutura constitui-se de duas partes básicas: “uma **rede associadora de padrões**, que aprende as relações entre a forma da base e a forma de *perfeito* e uma **rede decodificadora**, que converte as representações de traços do tempo passado em representações fonológicas” (p. 223). As unidades podem ser subdivididas em dois grupos, um utilizado para a representação do padrão do *input* correspondente à forma da raiz do verbo em processo de aquisição. O outro conjunto de unidades tem a função de representar o padrão do *output* gerado pelo modelo como sua aposta atual sobre o passado correspondente à forma de base representada nas

unidades de *input*.

Segundo McClelland & Rumelhart (1981), as formas relacionadas às bases verbais e aos passados destas têm que estar representadas no modelo, “de maneira que estes traços forneçam uma base conveniente para a apreensão das regularidades incorporadas nas formas de perfeito dos verbos da língua inglesa.” (p.233). Portanto, para a apreensão de tais regularidades são postuladas estruturas - representadas na figura 6 acima - chamadas de *Wickelfeatures*, um traço conjuntivo (ou sensível ao contexto) que têm a função de capturar um traço do fonema central, um traço do fonema anterior e um traço do fonema posterior. Esta estrutura é uma representação compactada dos chamados *Wickelfones*, uma representação anteriormente proposta por Wickelgren (1969, apud Rumelhart & McClelland, op. cit.) a qual era composta de seqüências de três fonemas (um antecessor à esquerda de um fonema central, o próprio fonema central e seu sucessor à direita) com o intuito de “capturar o contexto” no qual se inseria o fonema. Caso o fonema a ser representado fosse o primeiro fonema da palavra, este viria precedido de um símbolo # correspondente à fronteira de palavra; caso fosse o último, esse mesmo símbolo estaria à sua direita - assim as representações sempre se constituíam de trincas. Esta “estruturção” do Wickelfone garante a delimitação suficiente do contexto no qual um fonema ocorre, fornecendo dados suficientes para “distinguir entre diferentes casos de regra de *tempo perfeito* e para a caracterização das variáveis contextuais que determinam as sub-regularidades entre os verbos irregulares no tempo passado” (p. 234). A representação à base de Wickelfones traz, no entanto, inconvenientes que inviabilizam sua utilização em um modelo como este aqui apresentado: são muito numerosos - cerca de 43.000¹² - e deveras específicos.

Devido a este fator, os autores postularam os *Wickelfeatures*, uma representação do traço, ao invés do fonema em si. Então, para poder especificar o traço pelo qual cada fonema foi caracterizado, Rumelhart & McClelland (1986) elaboraram um quadro procurando manter um número restrito de diferentes traços, ao mesmo tempo em que buscaram dar um código distintivo ao máximo de fonemas possíveis; de forma que cada

¹² Esse número é a potência cúbica - já que os Wickelfones são trincas de fonemas - dos 35 fonemas conhecidos do

fonema passou a ser categorizado em quatro dimensões:

- a primeira categoria classifica os fonemas em três tipos diferentes: consoantes contínuas, consoantes interrompidas e vogais;
- a segunda categoria é uma especificação da anterior, pois divide a categoria acima da seguinte maneira: as consoantes contínuas entre fricativas, líquidas e semivogais, as consoantes interrompidas entre oclusivas e nasais e as vogais entre altas e baixas;
- a terceira categorização diz respeito ao local aproximado de articulação, ou seja, anterior, médio e posterior;
- a quarta dimensão subcategoriza vogais e consoantes respectivamente entre curtas e longas e entre surdas e sonoras.

Sob essa categorização Rumelhart & McClelland procuram dar conta do processo de aprendizado do tempo perfeito e de sua generalização, pois argumentam que com dois conjuntos de 460 *Wickelfeatures* - um para as unidades de *input*, que representam as formas de cada base verbal, e outro conjunto para as formas de *output*, que representam a forma de cada verbo no tempo passado - o modelo pode lidar bem até com contextos especiais de verbos com padrão de regularidade parcial.

As simulações para o treinamento do modelo resultaram em três pontos básicos de comportamento da rede: primeiro, o modelo captura os três estágios básicos do padrão de aquisição explicados acima; segundo, o modelo consegue capturar os aspectos mais diferentes no desempenho dos diversos tipos de verbos regulares e irregulares e, por último, o modelo é capaz de responder apropriadamente tanto aos verbos regulares e irregulares trabalhados durante o treinamento da rede, quanto a verbos que nunca lhe haviam sido apresentados. Estes resultados foram obtidos em apenas um longa simulação que contou com três etapas de treinamento, cada uma delas correspondendo respectivamente a uma fase do processo de aprendizagem do tempo perfeito do Inglês em crianças.

A simulação da primeira fase de aquisição consistiu em um treinamento com dez verbos de alta frequência, apresentados em dez ciclos - este pequeno treinamento já foi

Inglês; sem incluir ainda os casos em que há fronteira de palavra.

suficiente para que a rede produzisse bons resultados e uma performance equivalente àquela de uma criança na primeira fase do processo de aquisição. Já para os estágios subsequentes foram acrescentados 410 verbos de média frequência (334 regulares e 76 irregulares) e realizadas mais 190 baterias de testes, cada teste consistindo na apresentação de cada um dos 420 verbos. Comparativamente ao comportamento das crianças no processo de aquisição, podemos dizer que o início desta fase de treinamento corresponde à segunda fase na aquisição e o desempenho, ao final das apresentações das 190 baterias de treinamento, corresponde ao terceiro estágio do processo de aquisição. O modelo apresenta um desempenho praticamente livre de erros neste ponto do treinamento; então, são apresentados 86 verbos de baixa frequência (72 irregulares e 14 regulares) e suas respostas são coletadas. Durante esta fase, a força das conexões não é ajustada¹³ e o comportamento de rede é a simulação do comportamento dos falantes perante a conjugação de novos verbos. O resultado é bem otimista, pois 90 % dos *Wickelfeatures* são gerados corretamente; obviamente há um desempenho mais fraco perante os verbos irregulares, da ordem de 85%; ou seja, nada que comprometa a confiabilidade do modelo. Esta última fase do treinamento é comparável às aquisições de novos verbos realizadas após ter-se estabilizado o padrão da língua já adquirido.

Os autores ressaltam, entretanto, que a experiência do treinamento da rede não captura exatamente a experiência de aprendizado de uma criança, pois há aumento abrupto do corpus de uma fase para outra, além de uma diferença exagerada no comportamento da rede da fase inicial para as fases finais (p.241).

¹³ Assim que uma forma base de um verbo e um alvo são apresentados ao modelo em uma bateria de testes, o padrão de associação da rede computa o resultado gerado a partir do *input*. O modelo compara cada unidade de saída com o alvo, a força das conexões são então ajustadas para que o modelo se estabilize e novos *inputs* possam ser apresentados.

4. Modelo Augmented Addressed Morphology (AAM)

O Modelo AAM elaborado por Caramazza, Laudanna & Romani (1988) apresenta-nos uma perspectiva intermediária em relação àquelas já descritas sobre a representação das palavras morfologicamente complexas no léxico de acesso; pois, conforme pudemos constatar nas páginas anteriores, há uma certa radicalidade bipartida no modo de se representar a informação relativa à forma da palavra neste estágio do processamento lexical. Se por um lado, temos no Modelo de Taft & Forster (1975) a premissa de que toda palavra formada por mais de um constituinte morfológico passa, incondicionalmente, por um processo pré-lexical de segmentação, causando um acesso exclusivo aos seus morfemas, no outro extremo, temos os modelos de Forster (1976) e Morton (1970) entre outros, que postulam a palavra como uma unidade plena no que diz respeito ao acesso lexical, seja ele realizado de forma direta ou indireta. Temos ainda que o modelo de McClelland & Rumelhart (1981) consiste em um caso à parte ao passo que inclui, conforme já dissemos, níveis relacionados também aos traços visuais e letras; mas para tratar da representação das palavras com estrutura morfologicamente complexa, como as formas verbais no tempo passado, são postulados wickelfones separados para forma e base verbais e também para as marcas de tempo.

Pudemos notar também que ambas as vertentes, ao optar por um radicalismo na forma de representar itens morfologicamente complexos, geram diversas questões das quais não conseguem dar conta dentro de suas principais projeções para o acesso lexical. É o caso, por exemplo, do acesso às palavras cujo processamento analisado poderia detectar uma prefixação aparente - como *referir*, do reconhecimento de novos termos afixados, da existência do efeito de frequência como fator de ativação para palavras homógrafas nos modelos de ativação, entre outros.

De maneira que o modelo AAM vem como a solução a algumas destas questões, à medida em que propõe ambas as possibilidades de representação, aproveitando o que cada uma dessas hipóteses têm de melhor. Sua elaboração se deu, portanto, baseada nessas

lacunas deixadas pelas abordagens de representação integral ou de decomposição das palavras no acesso lexical. Apresentaremos abaixo as principais características deste trabalho e, um pouco mais à frente, a descrição de alguns experimentos nos quais os autores se basearam para reforçar sua argumentação.

O modelo adotado por Caramazza et alli (1988) pode ser definido inicialmente pela exposição de algumas de suas hipóteses, consistindo a primordial dentre elas, que o sistema de acesso lexical funcione de modo mais transparente possível. De maneira tal que se assume que o processamento está baseado na informação carregada explicitamente na forma superficial do estímulo; e, como consequência disso, nos estágios iniciais do processamento de um estímulo (*input*) ortográfico, o sistema pode fazer uso apenas da informação ortográfica de superfície fornecida por um estímulo isolado. Decorre, portanto, desta hipótese que todas as formas complexas do ponto de vista morfológico, parcial ou totalmente irregulares, devem ter uma representação independente, já que em ambos os casos não há uma forma comum na representação de constituintes que possa ser recuperada a partir da forma de superfície do *input*. O leitor pode ter como um exemplo de formas parcial e totalmente irregulares para a flexão na Língua Portuguesa, respectivamente, os pares *ler X lido* e *ser X fui*. Veja que no primeiro par o processo morfológico é considerado parcialmente irregular porque parte da forma base de *ler* foi suprimida antes do acréscimo da forma de particípio *-ido*; enquanto, no segundo exemplo, o processo flexional ocorreu gerando uma forma supletiva com a supressão de total de forma da base verbal. Apresentaremos abaixo as outras hipóteses lançadas pelo modelo AAM:

- a similaridade lexical desempenha um papel de restrição tão importante no modelo que leva o mecanismo de acesso a ser caracterizado como um sistema paralelo de ativação, no qual o grau de ativação de uma representação ortográfica armazenada (unidade lexical) é uma função de similaridade entre os grafemas constituintes da seqüência de letras de *input* - denominada similaridade **grafêmica** - e a representação armazenada;
- pressupõe-se que as palavras morfológicamente complexas já conhecidas sejam acessadas através de **unidades de acesso** contendo uma representação total da palavra.

Enquanto que em caso de palavras de formação regular, porém ainda desconhecidas para o falante, e casos nos quais apenas a base da palavra lhe seja familiar propõe-se um acesso através de unidades de acesso decompostas em morfemas. Estas últimas unidades também podem ser acessadas se a frequência da base do estímulo for significativamente maior do que a frequência total da ocorrência em processamento;

- por ser um modelo que opta pela ativação paralela e assumindo a hipótese acima exposta, temos por decorrência que, assim que uma forma ortograficamente transparente é apresentada como *input*, haverá uma ativação de ambas as unidades de acesso - decomposta e totalmente representada, diretamente proporcional à frequência dessas unidades de acesso;
- temos ainda que bases e afixos são unidades independentes desde que elas sejam unidades produtivas na língua; o que leva a postular que palavras que se encaixem em um paradigma regular de flexão ou derivação permitem representações decompostas.

Assumindo tudo o que foi dito acima, podemos dizer que o modelo baseia os procedimentos de acesso em princípios de similaridade ortográfica, principalmente, no que diz respeito à sua transparência ortográfica. Ou seja, as informações contidas na estrutura de superfície da palavra vão “dizer” ao sistema se aquele estímulo pode ser acessado via uma unidade de representação integral ou analisada da palavra. De forma que devemos entender por esse princípio de similaridade lexical que o *input* deva satisfazer um conjunto de condições suficientes para que ele seja válido como fonte de ativação para qualquer unidade de acesso lexical (cf. Chialant & Caramazza, op. cit., p.64).

O leitor já deve ter notado que há uma diferença entre os conceitos **unidade de acesso lexical** e **representação lexical**. A representação lexical é postulada como uma forma invariável, compartilhada por vários itens lexicais, que carrega toda a informação sobre as propriedades fonológicas e ortográficas desta forma lexical: é a parte do item lexical relacionada ao conhecimento formal - diferente em essência de sua contraparte relacionada ao conhecimento semântico, conhecida por **lema** (ver Levelt 1989, para saber mais sobre este conceito). Já as unidades de acesso lexical são apresentadas como unidades intermediárias entre o estímulo e a representação lexical, podendo, por este motivo, assumir

a forma inteira da palavra ou de seus constituintes, pois são dependentes apenas dos aspectos superficiais da estrutura da palavra. Um exemplo da diferença entre ambos os conceitos pode ser exposto através da forma morfológicamente complexa *prosperidade*, dado que por essa definição sua representação lexical será sempre *próspero* e *-idade*, ao passo que as unidades de acesso podem ser *prosperidade*, *prosper* e *-idade*. Conforme vimos dizendo desde o princípio do trabalho, a diferença na representação da unidade de acesso dependerá da hipótese de organização lexical abordada.

No entanto, é importante ficar claro que, para a ativação uma unidade de acesso, um estímulo não lhe deve ser apenas idêntico ortograficamente, senão, no exemplo que acabamos de utilizar, teríamos também a ativação da unidade de acesso lexical do substantivo *idade*. Logo, a ativação de uma unidade lexical deve satisfazer também às condições relacionadas à distribuição das unidades lexicais na língua. No caso apresentado, por exemplo, poderíamos ter como uma restrição à ativação do termo *idade*, o fato de que a forma *-idade#* sempre aparece ao final de uma palavra, sendo uma forma presa sempre precedida de uma fronteira de morfema e sucedida por uma fronteira de palavra¹⁴; enquanto que *#idade#* está sempre entre fronteiras de palavras. Portanto, a ativação do substantivo *idade* não satisfaz **as restrições distribucionais** da unidade de acesso de *-idade*, uma vez que a entrada lexical contém todas as informações lexicais e gramaticais daquela unidade, inclusive as propriedades de combinação de seus constituintes.

Podemos perceber, através de uma perspectiva de acesso lexical como esta rapidamente apresentada no modelo AAM, que é inegável a importância da Morfologia em um estudo que pretende tratar do processamento das palavras morfológicamente complexas - incluindo-se aqui o acesso lexical e a organização dos itens no léxico - estejam elas sob os processos de flexão ou de derivação. O intuito de Chialant & Caramazza (1995) é buscar a natureza da representação lexical para formas regularmente flexionadas, questionando o processamento através da relação dessas formas com suas bases. Para tal finalidade, os

¹⁴ Temos uma convenção terminológica a ser explicitada aqui: para fronteira de palavra, ou seja, a delimitação por espaços em branco que determina a palavra para o seu reconhecimento visual, usaremos o símbolo "#", enquanto para delimitação interna à palavra, a fronteira de morfema usaremos o símbolo "-". Não pretendemos, entretanto, fazer qualquer discussão interna à esta teoria neste trabalho.

autores procuraram na literatura por resultados experimentais que fornecessem evidências da importância da estrutura morfológica por meio de critérios como a relevância da **freqüência da base** (*root*), a existência do efeito de **priming morfológico**¹⁵ e o papel da estrutura morfológica no reconhecimento de **não-palavras morfológicamente decomponíveis**.

O conceito de freqüência de base, elaborado primeiramente por Taft (1979), é uma extensão do conceito de freqüência de palavra, para a qual leva-se em conta que no caso do acesso lexical das palavras flexionadas há uma influência da freqüência da base. O valor de uma freqüência de base é obtido pela soma da freqüência de todas as palavras que compartilham daquela base; por exemplo, para saber a freqüência da base de *gelar*, temos que somar a freqüência de superfície (ou de palavra) de palavras como *gelou*, *gelo*, *congelar*, *congelador* e todas as outras palavras formadas por esta base. Segundo o estudo de Marcus Taft acima mencionado, constatou-se que, em um teste de decisão lexical¹⁶, uma palavra com baixa freqüência de superfície e alta freqüência de base como *sized* demorou menos tempo para ser reconhecida do que uma palavra como *raked* cujas características são opostas, alta freqüência de ocorrência da palavra e baixa freqüência de base. Fatos como estes indicam que, pelo menos em algum nível do processamento lexical, houve uma análise decomposta dos constituintes morfológicos.

O fato de que o efeito de *priming*¹⁷, um efeito bastante abordado no estudo psicolinguístico - que se realiza através da aceleração do acesso de itens lexicais em virtude da apresentação prévia de outros itens lexicais a eles relacionados - pode ser levado em conta para evidenciar o acesso às estruturas morfológicas durante o processamento lexical. Já na década de 70, estudos como os de Morton & Murrell (1974) foram ponto de

¹⁵ Esta modalidade experimental está descrita de forma mais detalhada no Capítulo II desta dissertação.

¹⁶ Ver também no Capítulo II a definição de um experimento de Decisão Lexical.

¹⁷ Estes termos estão definidos no capítulo relacionado aos materiais e métodos nos experimentos psicolinguísticos. Por ora, podemos dizer resumidamente que, por tempo de reação, pode-se entender o período de tempo que um sujeito experimental leva para cumprir uma tarefa, por exemplo, de reconhecimento visual de palavras. Supondo que nesta tarefa o sujeito tivesse que responder se uma determinada seqüência corresponde ou não a uma palavra de sua língua. Ou seja, o tempo de reação (TR) do sujeito é o intervalo de tempo que compreende a apresentação do estímulo visual e o ato de responder a ele. Já o efeito de *priming* está relacionado a um efeito associativo que pode abranger várias áreas do conhecimento linguístico, portanto podemos ter efeitos de *priming* morfológico, fonológico, sintático, ortográfico e principalmente, semântico. (Ver explicação no segundo capítulo).

demonstração de que uma simples apresentação de palavras morfologicamente relacionadas afeta o tempo de reconhecimento de sua contraparte relacionada, como as palavras *cars* (*prime*) e *car* (alvo). Ou seja, houve uma influência da palavra-*prime* *cars* sobre o tempo de decisão lexical da palavra-alvo *car* apresentada posteriormente. Estes resultados apontam para a evidência de que as unidades de ativação e/ou de representação incluem subcomponentes morfêmicos de uma palavra, que podem por sua vez contribuir para uma recuperação mais rápida das unidades de acesso lexical. (cf. Chialant & Caramazza, op cit).

Porém, os experimentos de decisão lexical relacionados às não-palavras apontam para uma variação no tempo de decisão diretamente ligada à possibilidade de análise de uma não-palavra¹⁸ - ou seja, uma seqüência de letras que não faz parte da língua do falante - em constituintes morfológicos durante o processamento. Por exemplo, nos experimentos de Taft & Forster (1975), foi constatado que os tempos de decisão lexical para se rejeitar uma não-palavra constituída de uma base e um prefixo reais (**dejuvenescer*, adaptando de **dejuvenate*) são muito menores do que os tempos de decisão para uma não-palavra constituída de um prefixo real e uma “pseudo-base” (**depertório*, cuja a base seria **pertório*, de *repertório*). No estudo posterior de Caramazza et al. (1988), demonstrou-se que os tempos de reação e/ou as taxas de erros em tarefas de decisão lexical são afetados pela complexidade morfológica das palavras mesmo quando os efeitos de similaridade ortográfica com outras palavras são controlados; pois observou-se que a reação dos sujeitos tende a melhorar, em ordem crescente, nos fatores tempo e exatidão em relação a não-palavras que tenham base e afixo reais, como no exemplo *tecla-deiro*; em oposição a uma base real e um pseudo-afixo, como em *tecla-daro* ou *ticla-deiro*; sendo que o desempenho é mais rápido e revela uma baixa taxa de erro se a não-palavra for formada sem seqüências que se pareçam com morfemas, como *ticla-diro*. Este tipo de resultado experimental pode ser melhor explicado via uma hipótese de representação analisada, já que para uma hipótese de representação total da palavra, a diferença na constituição interna de uma não-palavra não pode afetar o tempo de reação.

Todos esses dados levam à confirmação da hipótese de que a representação das

¹⁸ O conceito detalhado de não-palavra é abordado na segunda nota do Capítulo II.

unidades morfológicas de acesso lexical não pode ser vista sob o aspecto de um radicalismo teórico até então predominante na literatura da área; radicalismo esse representado na obrigatoriedade de representação e acesso analisados ou na obrigatoriedade da integralidade destes. O texto de Chialant & Caramazza (1995), assim como o modelo de *Augmented Addressed Morphology*, nos apresentam, além de uma perspectiva que valoriza a importância de informações morfológicas no acesso lexical às palavras morfológicamente complexas, a idéia de que existam fatores relacionados a propriedades combinatórias dos próprios morfemas que devam ser levados em conta quando se pensa no acesso a essas informações morfológicas. São essas condições de distribuição das unidades lexicais na língua, chamadas de **propriedades distribucionais**, que serão abordadas em capítulos posteriores.

CAPÍTULO II

O PROBLEMA E AS HIPÓTESES

A. As Propriedades Distribucionais dos afixos, uma perspectiva inovadora para a organização das informações morfológicas no acesso lexical

O artigo *Distributional Properties of Derivational Affixes: Implications for Processing*, de Laudanna & Burani (1995), identifica-se por abandonar a tradicional discussão sobre a decomposição e a não-decomposição das palavras morfológicamente complexas durante o acesso lexical, propondo, dessa forma, uma nova maneira de se olhar para os mecanismos de acesso e representação lexical. Sob esta perspectiva apresentada pelos autores, a pesquisa sobre o processamento visual das palavras morfológicamente complexas necessita, em primeiro lugar, de uma investigação que classifique os tipos de palavras com base em “outras dimensões de relevância psicolinguística” que não apenas aqueles já definidos pelos moldes linguísticos. Ou seja, parâmetros que levem em conta “onde” e “quantas vezes” palavras e morfemas de uma língua ocorrem; portanto, parâmetros quantitativos e distribucionais tal qual a **freqüência** e o **tamanho**¹⁹ dos constituintes morfológicos e das palavras, como também as relações quantitativas entre formas realmente afixadas e as formas **pseudo-afixadas** (*pressentir x predar*, por exemplo).

Laudanna & Burani apresentam sua argumentação tendo em vista que os fatores de distribuição dos prefixos ali discutidos podem variar de uma língua para outra; logo, assumem uma posição pouco comum entre os pesquisadores da área. Há, na pesquisa linguística, uma tendência à generalização para as teorias dos fatos localmente encontrados nas línguas investigadas, posto que é um objetivo já tradicional nesta área de estudo a busca por universais teóricos. Em *Distributional Properties*, a tentativa é demonstrar que se os

parâmetros distribucionais e lingüísticos da morfologia derivacional são relevantes no modelamento da organização do léxico mental e, em particular, dos mecanismos que permitem o acesso às representações lexicais; então, segue-se que quanto maior a variação dos valores destes parâmetros de uma língua para outra, maior poderá ser a variação na organização dos sistemas de acesso lexical entre estas línguas. Assim, os autores apontam para o fato de que os procedimentos devotados ao acesso das representações lexicais correspondentes às palavras devem ser sensíveis às regularidades inerentes às palavras e às unidades sub-lexicais (cf. p.347).

Apesar de adotarmos os conceitos apresentados no modelo *Augmented Addressed Morphology*, encontramos uma lacuna em um dos pontos argumentativos nos quais os autores se fixam; segundo Caramazza et al.(1988), e por extensão Laudanna & Burani (1995), o acesso lexical se dá com base em estímulos ortográficos, das formas de superfície da palavra escrita, influenciado pela similaridade e transparência ortográfica, de maneira que acabam por privilegiar somente a modalidade escrita do estudo lexical. Até este ponto, não haveria maiores problemas na opção por uma modalidade específica de estímulo, afinal, defendemos logo de início a necessidade de recortes na pesquisa experimental. Contudo, um problema se insere a partir do momento em que as conclusões sobre o reconhecimento da palavra escrita servem de base para toda uma estruturação do processamento lexical humano - um processo muito mais genérico - ao negligenciarem que dados como frequência e produtividade também são influenciados pela fala, os autores acabam por favorecer o estímulo ortográfico como estímulo básico para a organização, acesso e reconhecimento lexicais. Não pretendemos, neste momento, aprofundarmo-nos em modelos de reconhecimento da palavra falada²⁰; mas também não podíamos deixar de observar que um modelo de reconhecimento da palavra escrita não deve ser tomado como base de todo o processamento lexical, já que é de domínio dos profissionais que trabalham

¹⁹ Apesar da “deselegância” do termo para uma categorização, que na versão de língua inglesa é apresentado através do termo *length*, preferimos traduzí-lo por *tamanho*, pois é justamente esta a idéia relacionada a este parâmetro quantitativo: o número de letras ou de sílabas que constituem o prefixo. Enquanto que nas opções previamente cogitadas *dimensão* e *extensão*, sempre apresentavam uma noção de espaço que não gostaríamos de imprimir ao termo.

²⁰ Como o leitor pode perceber, já apresentamos alguns modelos que levam em conta ambas as formas de *input* como Forster (1979), Morton (1970). Para saber mais sobre o assunto veja Lively et al. (1994).

com a linguagem a existência de comunidades que não se utilizam da escrita. Portanto, neste ponto, nossa adoção do modelo AAM é parcial, pois pensamos que as características do estímulo de superfície da palavra podem influenciar apenas no léxico de acesso, ou seja, no primeiro estágio do processamento lexical. O artigo de Laudanna & Burani (op. cit.) é praticamente uma extensão das hipóteses lançadas pelo modelo acima descrito, e aos seus conceitos, como aos do modelo AAM, adotamos parcialmente e com as mesmas ressalvas.

Em *Distributional Properties* vemos o mecanismo de acesso lexical caracterizado como um sistema de ativação sensível à estrutura de superfície das palavras, no qual o grau de ativação de uma representação para o acesso ortográfico depende de uma função de similaridade grafêmica entre uma seqüência de letras, que serviu como *input*, e uma representação estocada. Estas representações podem estar armazenadas tanto em forma de constituintes morfológicos ou na forma integral; sendo que esta última só pode existir para palavras previamente conhecidas.

Assim, a representação das palavras derivadas pode ser vista em Laudanna & Burani (op. cit.) sob a perspectiva de propriedades distribucionais dos afixos e de **propriedades lingüísticas redefinidas em termos distribucionais** - na seção acima mencionamos tais propriedades ao falarmos das “restrições distribucionais” que impediam que termos homógrafos, como o sufixo *-idade#* e o substantivo *#idade#*, fossem concomitantemente ativados e acessados durante o processamento do estímulo visual da palavra *prosperidade*. A conceitualização de tais propriedades distribucionais tem como base a existência de parâmetros quantitativos definidos tanto em termos das características individuais dos morfemas e de suas ocorrências, por exemplo, o **tamanho** ou a **freqüência** dos sufixos derivacionais - quanto nas bases de suas co-ocorrências condicionadas com outros morfemas na língua, por exemplo, os números ou tipos de prefixos derivacionais que ocorrem em combinação com radicais/raízes ligados.

Apresentamos abaixo cada um dos fatores ou propriedades que podem ser vistos como distribucionais na Língua Italiana²¹, além das propriedades redefinidas como tal:

²¹ Os exemplos foram elaborados em português para que o leitor possa compreendê-los melhor.

- **freqüência da palavra** - conforme dito anteriormente, costuma ser definida pelo número de vezes em que um certo termo aparece em um corpus lingüístico determinado e analisado. Por exemplo, “quantas vezes foi utilizada a palavra **necessão** nos textos jornalísticos de um certo período de tempo”. Neste caso, conforme apresentamos na seção anterior, há também a quantificação de freqüência de base para afixos e bases de palavras; portanto, a freqüência total de uma palavra como *reavaliar* seria a soma das freqüências de ocorrência das formas de superfície, do prefixo *re-* e da base *avaliar*;
- **produtividade dos afixos** - já conhecida na literatura de áreas como Morfologia e outras áreas relacionadas com o estudo do Léxico, a produtividade de um afixo está relacionada com sua “capacidade” de produzir novas palavras na língua. Por exemplo, em termos de desinências verbais, podemos dizer que, na Língua Portuguesa, os verbos de primeira conjugação são mais produtivos que os de segunda e terceira conjugação; fato que pode ser visto pelos neologismos criados na área de Informática como *deletar*, *butar*, *resetar*;
- **tamanho do afixo**- relaciona-se com a quantidade de letras ou sílabas que formam um afixo. Os autores afirmam que na Língua Italiana há uma forte tendência de que o número de palavras prefixadas decresça em uma ordem inversamente proporcional ao tamanho do prefixo.²² Ou seja, geralmente quanto maior um afixo menos freqüentemente ele é utilizado para formar novas palavras; por exemplo, um prefixo como *re-* é muito mais utilizado para a produção de novos vocábulos do que um prefixo como *intra-*;
- **ambigüidade ortográfica** (“*ortographic ambiguity or confusability*”) - conceito de ambigüidade formal apresentado em termos de seqüências ortográfico/fonológicas²³ que podem corresponder a prefixos em palavras prefixadas como *re-* em *refazer* e a seqüências não-morfêmicas em palavras não prefixadas como *relação*. Tanto no Italiano

²² Por exemplo, a partir de uma análise preliminar, não publicada, do *Fonêmico* (base de descrição lexical disponível no LAFAPE), constatamos um dado que vêm favorecer esta hipótese, é a indicação de que os prefixos de duas e três letras como *re-*, *in-*, *co-*, *des-*, *pre-* e *sub-* apresentam-se mais produtivos na Língua Portuguesa do que prefixos de quatro letras ou mais tal como *sobre-*, *anti-*, *para-*, *meta-*, *intra-*, *contra-*, *supra-*, posto que os primeiros são formadores de uma grande parte das palavras prefixadas registradas na fonte acima citada.

²³ Temos uma observação quanto a este parâmetro no que diz respeito à pressuposição no texto de Laudanna & Burani (1995) de que as seqüências ortográficas e fonológicas coincidam. Por levarmos em conta que as fronteiras ortográficas são muito mais estanques do que as fronteiras fonológicas, assumiremos neste trabalho que esse

como no Português, essas palavras são facilmente encontradas, principalmente no que concerne aos prefixos de duas e três letras. A questão relativa a este critério se resume, então, a como essas palavras seriam representadas e acessadas se as hipóteses de representação morfêmica e, conseqüentemente, do acesso analisado durante o processamento lexical fossem assumidas. Já encontramos em Mattoso Câmara (1975) um questionamento semelhante sobre os afixos através da palavra **fogão** - que pode representar tanto o utensílio doméstico quanto o aumentativo da palavra **fogo** - levantando a dúvida se a adição do sufixo **-ão** criaria duas novas entradas no léxico para a mesma forma ortográfica; e para a nossa pesquisa a questão de como se daria seu processamento se a forma fosse obrigatoriamente analisada em constituintes. Este conceito é muito importante quando lembramo-nos de que o modelo AAM privilegia análises decompostas para novas palavras no léxico e que, para estas análises, levar-se-ão em conta as características superficiais do estímulo.

Temos que atentar, porém, para uma determinação teórica de que esses parâmetros só podem ser considerados distribucionais se puderem ser aplicados “em larga escala” dentro dos processos morfológicos. Este “detalhe” é parte constitutiva da hipótese básica de “*Distributional Properties of Derivational Affixes*”, pois os autores assumem que em uma pesquisa sobre os processos morfológicos derivacionais, sob uma abordagem na qual são buscadas variáveis psicolinguísticas relevantes ao sistema de acesso lexical, deve-se levar em conta tanto as características do *input* como as *restrições cognitivas* impostas pelo processamento lexical (p.348). As características do *input* são significativas a partir do momento em que se assume que um sistema de processamento lexical é modelado e afetado pela sua exposição a uma determinada língua; portanto a organização do léxico de um “usuário” de uma determinada língua provavelmente será afetada pelas características inerentes à ela. Já as restrições cognitivas são vistas como um fator que independe da língua, ou das propriedades dos processamentos e das representações no que diz respeito ao espaço computacional disponível para o processamento da informação lexical e à

parâmetro só é válido para seqüências ortográficas.

capacidade de estocagem dos *buffers* (cf. p.348). Inserem-se, nesta última classe, as restrições que impedem a ativação para o processamento de decomposições incorretas ou inconsistentes, como nos casos das palavras pseudo-afixadas - por exemplo, um processamento que, erroneamente, analisasse as palavras *refeição* e *relação* como termos morfológicamente complexos, levando à uma conseqüente ativação de formas incorretas como o prefixo *re-* e os substantivos *feição* e *lação*. A título de comparação, é válido lembrar que uma análise como esta não seria possível em termos como *rejeição* e *redação*, pois, dada à inexistência de pseudo-fronteiras fonológicas, estes termos, diferentemente daqueles, não suscitam restrições cognitivas, posto que as restrições fonológicas são prontamente “acionadas”.

Podemos perceber, após a apresentação das hipóteses e dos conceitos básicos envolvidos no trabalho de Laudanna & Burani (1995), que uma das noções que emergem do texto está relacionada à idéia de que os afixos derivacionais podem ser tomados como unidades de processamento pelo léxico de acesso conforme as seqüências em que aparecem, tal qual uma base verbal ou nominal. Essa “propriedade” é denominada **saliência do afixo** e sua “legitimidade” está baseada em características inerentes aos próprios afixos, em parâmetros distribucionais relativos à sua ocorrência em uma dada língua - como a freqüência e a produtividade- e também, nas co-ocorrências com outras unidades sub-lexicais de relevância. Por este motivo é que as propriedades distribucionais dos afixos são vistas como um fator de determinação do conteúdo e da organização do léxico de acesso.

Estas propriedades de distribuição podem exercer suas influências na organização lexical de duas formas, ou independentemente, ou em conjunto a outro parâmetro. Por exemplo, a relação entre a produtividade e o tamanho dos prefixos: há uma tendência na Língua Italiana - e que pode ser também observada na Língua Portuguesa - de que os prefixos mais produtivos sejam os de menor tamanho, ou seja, formados por uma, duas, ou três letras, que geralmente constituem-se apenas de uma sílaba, como *a-*, *in-*, *de(s)-*, *co-*, *pre-*, *re-*, *en-*, *ex-*, *sub-*, entre outros.

Porém, um parâmetro quantitativo tal qual a freqüência já é considerado como fator de influência independente nos processo lexicais, desde modelos como o de Morton (1970)

- no qual reduz os limiares de ativação dos logogenes - mas também pode ser encontrado como modificador dos pesos de conexão em modelos conexionistas com o de McClelland & Rumelhart (1981), e fator de reorganização do léxico de acesso no *Prefix Stripping Model*. No entanto, a frequência pode ser considerada como fator de influência no acesso lexical, também em conjunto com o efeito de ambigüidade ortográfica ou com o nível de produtividade dos afixos - por este fato é que a frequência é apresentada como uma propriedade redefinida em termos distribucionais. Laudanna & Burani nos apresentam um estudo quantitativo (p. 351) levando em conta o grau de ambigüidade que um prefixo pode causar e seu tamanho, e concluem, com base nesses dados, que prefixos muito curtos, com apenas uma letra, tendem a ser altamente produtivos e ambíguos; mas, conforme cai o índice de produtividade, diminui também o índice de confusabilidade. Efeito inverso pôde ser constatado entre frequência e extensão dos prefixos; ou seja, quanto maior o prefixo, menor sua frequência de ocorrência em palavras da língua.

Entretanto, quando a comparação se dá em três dimensões: a frequência com o grau de confusabilidade e o tamanho do afixo, temos algo parecido com o primeiro efeito, porém em menor escala. O que Laudanna & Burani constatam são índices mais altos de frequência para prefixos que causam menor ambigüidade, mas estes prefixos tendem a ter até três letras; ou seja, são de pequena dimensão, sendo que os prefixos de duas letras alcançam os mais altos níveis de frequência²⁴.

Enfim, o leitor pode agora compreender o porquê dessas propriedades estarem relacionadas com a distribuição lexical e diretamente baseadas nos estímulos de superfície, posto que elas próprias se relacionam e se influenciam.

1. Pseudo-prefixo, um caso de ambigüidade em termos morfológicos

Suponhamos o contato de um leitor com as seguintes formas verbais: a) *reconstruir*,

²⁴ Não encontramos oficialmente ainda algum estudo quantitativo semelhante no Português, mas faz parte de nossos planos realizá-lo. Todavia, ao elaborarmos a lista de nosso segundo experimento nos deparamos com o fato de que o prefixo *re-*, no Português, é de pequena extensão, alta produtividade e frequência; quanto à ambigüidade, o leitor

refazer, religar e b) *relevar, retrair, revoltar*. Conforme explicamos anteriormente, pela proposta do modelo AAM e, por conseqüência, de *Distributional Properties of Derivational Affixes*, temos em a) palavras claramente prefixadas representadas no léxico de duas maneiras, uma forma analisada que compreende os constituintes da palavra (prefixo e base) e outra forma cuja representação é integral, ambas passíveis de ativação - embora os autores afirmem que o procedimento mais rápido para palavras já conhecidas seja a ativação da segunda opção. Não obstante, isto não quer dizer que as unidades de representação correspondentes ao prefixo e às bases verbais não recebam energia para ativação; há apenas uma diferença no limiar de ativação entre elas.

Já em b) temos uma classe muito especial de palavras que representa um problema significativo para modelos de processamento que, como o AAM, levam em conta a existência de unidades de representação para os constituintes morfológicos; pois pela sua forma ortográfica de superfície poderiam ser erroneamente analisadas como formadas pelo prefixo *re-* e, respectivamente, pelas bases verbais *levar, trair* e *voltar*. Estas palavras serão denominadas de **pseudo-prefixadas**; elas podem interferir no processamento lexical causando um efeito denominado por Laudanna & Burani de **ambigüidade ortográfica** ou **confusabilidade**. A principal conseqüência de um efeito como este para o processamento lexical está no fato de que mecanismos inibidores da ativação de tais constituintes têm de ser levados em conta nos modelos de reconhecimento das palavras prefixadas.

Este problema pode se tornar maior ainda se observarmos que, em muitos casos de palavras pseudo-afixadas, as “pseudo-bases” são formas verbais de alta freqüência; portanto, teoricamente haveria uma competição pela ativação das unidades referentes às duas formas de representação e, concomitantemente, um mecanismo de inibição agindo sobre as formas ambíguas. Estes mecanismos de inibição por vezes são tão fortes que se traduzem inclusive em “*primings* negativos”, ou seja, nas palavras experimentais os sujeitos apresentam tempos maiores que nas palavras-controle (ver Bushell, C. & A Martin; 1997, p. 1062). De maneira que uma proposta de processamento lexical como esta apresentada em *Distributional Properties*, a qual inclui efeitos de ordem psicolinguística - ou assim

pode ver ao final do terceiro capítulo os resultados do experimento envolvendo palavras pseudo-prefixadas.

redefinidos - como princípios de organização lexical, torna-se mais coerente do que propostas como as de *Prefix Stripping*, por exemplo, no sentido em que tenta abarcar situações de limite - mas nem por isso raras - para os mecanismos de acesso lexical. E, apesar de avocar os problemas que uma análise segmentada pode trazer ao processamento lexical, ao levar em conta a importância da informação morfológica, o modelo também assume a palavra como uma unidade pertinente para o processamento.

2. Pseudo-prefixadas vs. prefixadas transparentes: existem semelhanças no processamento morfológico de *rebolar* e *refazer*?

A partir do momento em que nos propusemos a investigar o acesso lexical das palavras morfológicamente complexas na Língua Portuguesa e, por conseqüência, buscar alguns parâmetros lingüísticos relevantes para a organização lexical no que concerne os termos prefixados, decidimos que a elaboração de um experimento era ponto de crucial importância para nossa pesquisa. Na etapa de delineamento do nosso estudo experimental, buscamos nos experimentos já realizados em Chialant & Caramazza (op. cit.), Stanners, Neiser & Paiton (1979) e Laudanna & Burani (op. cit.), entre outros, a melhor maneira de se investigar o acesso a palavras morfológicamente complexas. Optamos, então, inicialmente por um teste de *priming* morfológico, no qual a tarefa delegada aos sujeitos é a decisão lexical. Nosso teste seria constituído por palavras prefixadas, palavras morfológicamente simples, além das não-palavras.

Porém, conforme coletávamos material, nos deparamos com um problema na elaboração da questão base para nosso experimento: como poderíamos verificar a ativação das pseudo-bases em através de um efeito de *priming* morfológico, sem que obtivéssemos também um efeito de ordem semântica? A partir desta questão, uma outra mais interessante despontou: o que aconteceria com as palavras pseudo-prefixadas quando apresentadas em contexto de *priming* semântico? Refletindo mais sobre o efeito psicolingüístico escolhido para o desenvolvimento experimental de nosso trabalho, poderíamos nos questionar

também se, em havendo diminuição no tempo de reação das pseudo-prefixadas, este se daria nas mesmas proporções do efeito que esperávamos para a condição de apresentação das prefixadas transparentes e o que isto poderia, então, nos indicar sobre o processamento lexical humano?

Ao idealizarmos nosso cronograma experimental - que está descrito no próximo capítulo e ao qual batizamos “Morfo”, procuramos inserir, de uma maneira menos caótica que aquela apresentada acima, a verificação de algumas hipóteses sobre o processamento de cada uma das classes de palavras morfologicamente complexas, à luz da hipótese de representação proposta no modelo AAM - a qual vimos descrevendo desde o início deste tópico. De forma que, supondo uma dupla representação para as palavras de prefixação transparente, conforme indicado no modelo, esperamos que, à luz de uma facilitação semântica, deva haver uma redução no tempo de reação da palavra-alvo relacionada, caso a informação morfológica seja realmente acessada.

Quanto às palavras pseudo-prefixadas temos duas possibilidades: se há uma análise restrita apenas ao sentido da representação integral, então, o tempo de decisão lexical para o alvo relacionado será semelhante ao tempo de decisão lexical de um alvo não-relacionado. Porém, se houver análise dos dois sentidos, o verdadeiro e o “pseudo”, a informação morfológica terá sido acessada; o *prime* terá ocorrido, resultando na facilitação do acesso ao termo relacionado e na conseqüente diminuição do tempo de decisão lexical. De forma que, para que esta hipótese se confirme, o tempo de decisão lexical da palavra relacionada terá que ser menor do que o tempo da não-relacionada; pois para esta última não terá havido previamente nenhuma espécie de ativação parcial.

Enumerando as hipóteses elaboradas para a realização do experimento temos então:

1. Palavras de prefixação transparente, apresentadas previamente a bases relacionadas em significação - seja de oposição ou de sinonímia, devem acelerar o acesso lexical desta segunda palavra. Este efeito, conhecido na literatura como efeito de *priming* semântico, é traduzido por um tempo de reação menor dos informantes perante este contexto, do que se apresentássemos duas palavras não-relacionadas entre si.

2. Palavras pseudo-prefixadas poderão apresentar aceleração nos tempos de reação, quando comparadas aos tempos dos pares de palavras não relacionadas entre si, caso prevaleça a análise morfológica perante o acesso integral da palavra, levando a ativação das palavras relacionadas à pseudo base. Ou poderão apresentar tempos de reação semelhantes ao contexto de não-relação se não houver acesso analisado, prevalecendo o acesso à entrada integral da palavra, sem ativação das palavras relacionadas à pseudo-base.

CAPÍTULO III

MATERIAIS E MÉTODOS

No capítulo anterior, mencionamos em alguns momentos a necessidade do uso de experimentos de caráter psicolingüístico para atestar a confiabilidade das hipóteses propostas pelo pesquisador. Nossa tarefa neste capítulo é, inicialmente, descrever alguns tipos de experimentos psicolingüísticos, de forma que o leitor possa se situar quanto às modalidades de investigação experimental. Em seguida, apresentaremos o experimento através do qual testamos a nossa questão, descrevendo-o desde seu delineamento até a execução.

É válido evidenciar, primeiramente, que esses experimentos não provam ou invalidam qualquer teoria, eles apenas nos dão indícios de que os fatores teóricos são válidos ou não naquele determinado contexto experimental. Essa “capacidade limitada” dos testes deve-se a um recorte necessário para a elaboração do que chamamos de *design* ou **delineamento experimental**. Pois um experimento é realizado, geralmente, em um ambiente de laboratório e com uma finalidade específica; para que assim, o pesquisador possa controlar razoavelmente diversos fatores que influenciem as respostas dos **sujeitos experimentais**. Portanto, no momento em que se opta por uma pesquisa deste cunho, temos que delimitar não apenas o **material** lingüístico a ser utilizado, como também, quais os **sujeitos** dos testes e qual a **modalidade experimental** será empregada para a questão que foi determinada previamente dentro da teoria abordada; seguindo esta **metodologia**, o pesquisador procura garantir a confiabilidade de seus dados.

A. Diversas modalidades de investigação experimental na pesquisa psicolingüística

Em se tratando do tipo de experimento psicolingüístico a ser utilizado, temos primeiramente uma divisão básica que atende à forma de captação do estímulo pesquisado,

isto é, o sujeito pode ser submetido a um estímulo exclusivamente acústico ou visual, ou ainda a um estímulo que alie as duas formas perceptuais. É válido reforçar que a escolha do material estudado é determinante para a adequação da modalidade de experimento a ser empregada - por exemplo, ao optarmos pelo estudo do reconhecimento visual da palavra isolada, tivemos que buscar pelas modalidades que fossem delineadas com o objetivo de captar respostas a estímulos visuais; porém, podemos encontrar na literatura vários estudos de grande importância sobre o reconhecimento da palavra falada, ou de segmentos sonoros constituintes de uma palavra, ou ainda de unidades maiores que esta quando se relacionam com o processamento de informações sintáticas.

Com a finalidade de não nos distanciarmos muito de nossa questão investigativa, serão descritos abaixo apenas alguns dos métodos utilizados na pesquisa do reconhecimento da palavra em forma acústica ou visual; pois, desta forma, poderemos mostrar ao leitor como se insere a pesquisa psicolinguística na prática experimental²⁵. Estamos nos limitando ainda a somente descrever estes experimentos, de forma que não apresentaremos nenhuma discussão sobre as vantagens e desvantagens de cada um deles

Temos logo abaixo, portanto, modalidades experimentais apresentadas em uma ordem que tenta obedecer à evolução da própria investigação experimental na Psicolinguística.

- **Identificação perceptual** - caracteriza-se por ser um dos procedimentos mais simples no estudo do reconhecimento da palavra, dado que a tarefa dos sujeitos consiste apenas em tentar identificar palavras isoladas de forma precisa. Para dificultar relativamente esta tarefa de identificação, pode-se adicionar um mascaramento na apresentação das palavras, ou seja, alguma forma de distorção (p. ex. menos contraste, diferença nas letras entre caixa alta e baixa, padrão de letra diferente, etc.). A tarefa pode ser classificada entre direcionada, se for apresentado aos sujeitos um conjunto alternativas de respostas anteriormente a cada palavra-alvo; ou espontânea, quando o sujeito fica livre para responder ao estímulo. Esse tipo de experimento é considerado uma boa

²⁵ A descrição aqui apresentada está baseada nos artigos de Lively et al. (1994) e em Forster, K. (1995).

forma de se examinar os efeitos de frequência vocabular, de similaridade lexical²⁶ e para se observar relações estruturais entre as palavras estocadas no léxico mental.

- **Gating** - é uma versão aprimorada dos experimentos acima descritos, pois somente uma porção do estímulo é apresentado a cada tentativa. Por exemplo, na primeira tentativa de identificação somente 30 ms. da palavra são apresentados; em seguida, um tempo para a resposta é dado, caso esta não ocorra, apresenta-se 60 ms. na tentativa seguinte., e assim por diante até que o sujeito responda à tarefa. Essa porção de tempo de apresentação do estímulo pode ser tanto do início da palavra, como do seu final; além disso, o estímulo pode ser totalmente mascarado e a cada tentativa ir se apresentando porções “desmascaradas” do estímulo. O que se pretende com *Gating* é dar conta da distribuição temporal da fala, ou seja, saber quanto da informação fonético-acústica é necessária para que o reconhecimento da palavra se dê.
- **Decisão lexical** - diferentemente dos anteriores, os experimentos de decisão lexical não mascaram seus estímulos. Tanto para estímulos auditivos quanto para estímulos visuais, a tarefa dos sujeitos é relativamente simples: após um estímulo ser rapidamente apresentado, pede-se ao sujeito que o classifique entre as opções de *palavra* e *não-palavra* da sua língua. Geralmente, como forma de abolir o aprendizado por parte dos sujeitos, ou ainda como maneira de disfarçar o objetivo do experimento, as seqüências são apresentadas aleatoriamente dentro de sua classificação. Nos experimentos de reconhecimento visual, é um procedimento normal que a seqüência de letras a ser classificada seja apresentada isoladamente na tela do computador por um período de tempo suficiente para que o sujeito possa reconhecê-la. Em seguida, ele deve acionar a tecla em uma caixa de respostas correspondente à resposta certa - o tempo decorrido desde a apresentação do estímulo até que a resposta seja dada é chamado de **tempo de reação**; é justamente esta medida que o pesquisador terá como dado para a sua análise.

²⁶ Ou seja, se o reconhecimento de uma determinada palavra é afetado pela presença de uma palavra semelhante na mesma língua.

Para chegarmos a tal medida de tempo temos levar em conta o tempo que o sujeito leva para se conscientizar de que o estímulo apresentado tem ou não relação com alguma forma representada previamente no seu léxico; mais o tempo necessário para que ele tome uma decisão sobre a resposta a ser dada, somando-se ainda o tempo requerido para se iniciar a resposta²⁷. Este experimento é caracterizado por adotar um método *on line* de investigação, pois seu delineamento experimental direciona e treina o sujeito - através da apresentação muito acelerada das seqüências de letras na tela - para que a resposta ao estímulo seja dada de forma muito rápida, anterior a qualquer possibilidade de análise pós-perceptual. Este procedimento é diferenciado dos exemplos acima citados, os quais, por não exigirem uma resposta “imediate” de seus sujeitos, acabam tendo como resultado tempos influenciados por análises pós-lexicais.

- ***Naming*** - neste experimento, a tarefa dos sujeitos consiste em repetir, o mais rápido possível e em voz alta, as palavras que lhe são apresentadas. Normalmente o estímulo não está mascarado, e os sujeitos não têm que tomar qualquer tipo de decisão sobre sua natureza - como classificá-lo entre palavra e não-palavra, apenas têm de repeti-las corretamente. Pois, a verificação do experimento está relacionada com a hipótese de que os sujeitos têm que verificar se existe uma representação no léxico do termo a ser repetido e recuperar a informação estocada sobre esta representação, para que a pronúncia das palavras se dê de forma correta. A diferença de tempo entre a apresentação do estímulo e o início da vocalização do sujeito é medida por um instrumento, o *voiced-operated relay*. O número de erros é considerado uma variável-dependente e, normalmente, esta tarefa experimental não costuma levar em conta que fatores como a freqüência de ocorrência da palavra influenciem no tempo de resposta do sujeitos.

- ***Shadowing*** - este experimento é uma versão ampliada da tarefa de “naming” apresentada

²⁷ O leitor deve estar atento para o fato de que os termos “tempo de reação” e “tempo de associação” não equivalem conceitualmente. Tempo de associação é, aproximadamente, o intervalo temporal que compreende a tomada da decisão lexical.

logo acima. Porém, neste caso os sujeitos têm que repetir enunciações mais longas em sincronia com seu interlocutor. A intenção deste experimento é avaliar a interação de diferentes níveis do processamento da linguagem. Marslen-Wilson (1985) encontrou, em sua pesquisa, evidências de que os sujeitos realizam análises sintática e semântica do *input* durante a tarefa. Os padrões de erros aumentaram e, conseqüentemente o desempenho dos sujeitos decaiu, conforme o material apresentado como estímulo foi se tornando sintática, semântica e lexicalmente anômalo. Isto representa um indício de que os sujeitos (*shadows*) não repetem mecanicamente os estímulos a que são submetidos - eles se utilizam aparentemente das estruturas lingüísticas de que dispõem para o processamento cognitivo. Como os experimentos de Decisão Lexical, o *Shadowing* é considerado um experimento relativamente livre das influências pós-perceptuais, por exigir uma resposta imediata do sujeito.

- **Monitoramento de fonemas** - diferentemente dos experimentos anteriores (com exceção ao experimento de *Shadowing*), que originalmente foram elaborados para o reconhecimento visual das palavras, o monitoramento de fonemas centra-se na natureza das representações segmentais da palavra falada - portanto, busca indícios sobre a contribuição de informações lexicais e **sub-lexicais** (por exemplo, fonêmicas) no processamento de palavras isoladas. Nesta modalidade, os sujeitos têm que dar respostas muito rápidas sempre que um “fonema-alvo” é apresentado em uma palavra-estímulo. Ou seja, sua resposta não é espontânea como nos experimentos de *Naming* e *Shadowing*; mas um “procedimento de escolha forçada” (*forced-choice procedure*) como nos testes de Decisão Lexical.

Temos ainda uma classe de experimentos que, por basearem-se na chamada **relação de priming**, serão apresentados separadamente dos anteriores. Conceituar o termo *prime* para o leitor leigo, constitui-se uma tarefa complicada já que não há equivalente na língua portuguesa para o verbo inglês. O costume de alguns pesquisadores é adotar para o português o francesismo *amorçage*, o qual significa “detonar”, e que não denota muito bem

o efeito obtido pelo experimento. Essa dificuldade terminológica reflete-se na descrição dos experimentos que se utilizam da relação de *prime* no seu delineamento.

Podemos dizer, então, que na apresentação de duas seqüências de palavras, a existência de uma relação entre elas ocasiona o *priming*: uma aceleração no acesso lexical da segunda seqüência apresentada. Através da manipulação e observação dos tipos de relação existentes entre as palavras *primes* e seus respectivos alvos, pode-se obter evidências sobre arquitetura do sistema de reconhecimento das palavras. Esta definição nos leva à identificação do *priming* como um efeito lingüístico suscitado pelo contexto em que as seqüências são dispostas.

Entretanto, não podemos confundir o *priming* como um tipo de experimento, pois ele é, como dissemos ainda agora, uma relação decorrente do contexto em que um estímulo é apresentado. De forma que a tarefa dos sujeitos em um experimento de decisão lexical que conta com a relação de *prime* continua sendo a classificação do estímulo entre as alternativas de *palavra* e *não-palavra*; ou seja, diante de estímulos como *casa*, *gato*, *pato*, **porape*, **codada*, **dodera*, responder à pergunta: “Essa seqüência é uma palavra da Língua Portuguesa? Sim ou não?”. Portanto continuamos tendo em vista um experimento de decisão lexical, a diferença está na apresentação dos estímulos, os quais são apresentados unitariamente em experimentos de decisão lexical sem *priming*, e aos pares de seqüências relacionadas ou não-relacionadas entre si nos experimentos de decisão lexical que contém uma relação de *priming*.

Ressaltamos que esta **relação de *priming*** pode ocorrer entre diversas áreas do conhecimento lexical, ficando a critério do pesquisador selecionar a modalidade que mais lhe aprouver, ou que melhor se encaixar na questão abordada em sua pesquisa. Desta forma, os experimentos com *priming* acabam por se distinguirem conforme a relação privilegiada; ou seja, denomina-se *prime ortográfico*, se a relação de interesse estiver centrada nos padrões ortográficos das unidades de reconhecimento; **morfológico**, se a relação estiver entre os constituintes de formação do vocábulo; **fonológico**, quando envolve os padrões sonoros (fonéticos e fonológicos) que afetam o reconhecimento das palavras, **sintático**, quando se questiona a influência dos contextos sintáticos no processamento lexical, e

semântico, se a relação entre as seqüências apresentadas envolver questões de associação de sentido. O que tanto intriga os pesquisadores nesse tipo de estudo é que se pode olhar para as influências dos contextos lingüísticos, na relação entre os *primes* e suas seqüências-alvo, no intuito de se observar a organização do sistema de reconhecimento da palavra. Tentaremos a seguir nos fazer entender através da exemplificação de possíveis experimentos em cada uma das modalidades de design experimental:

- **Priming ortográfico** - conforme dissemos anteriormente, os sujeitos são submetidos a rápidas apresentações de duas seqüências de letras na tela do computador, estas podem ser palavras seguidas de outras palavras - subdivididas entre *primes*, controles e distratores - ou de não-palavras. Por exemplo, *porco/porto*, *lixo/vema*, *porco/criga*, *lâmpada/espada*, *grama/porto*, *porco/carro*, *lixo/rádio* e assim por diante (uma seqüência é apresentada após a outra, jamais os pares estarão simultaneamente na tela). As seqüências são apresentadas isoladamente na tela, mesmo que a relação se estabeleça aos pares. Aqui, como estamos apenas descrevendo e simulando um experimento, a palavra sob o efeito de *prime* (*porto*) é apenas uma; no entanto, em um teste real, ela e os outros *primes* estariam suficientemente mascarados entre várias outras palavras igualmente repetidas, de forma que o sujeito não a percebesse como tal. A medida que importa neste tipo de teste é, novamente, o tempo que o sujeito leva para classificar o estímulo entre as opções de palavra e não-palavra; tempo que pode ser diminuído na apresentação da palavra *porto*, no contexto de semelhança ortográfica com a palavra *porco*. Assim, o tempo de resposta do par *porco/porto* teria de ser menor que o tempo do par *porco/carro*, *porco/criga* e *grama/porto*, dentre os quais não há relação de semelhança ortográfica. Nesses experimentos, costuma-se apresentar uma seqüência na tela, uma **tela de intervalo** em branco, depois de alguns milissegundos o seu par e, em seguida, uma tela de intervalo um pouco mais longa para que o sujeito possa responder se a **segunda** seqüência é palavra ou não de sua língua; somente após a resposta ser dada pelo sujeito é que o par seguinte será apresentado. Entretanto, se faz necessário que durante a elaboração do experimento tempos limite de resposta sejam postulados, a fim

de que tempos de resposta muito longos não acabem sendo influenciados por fatores pós-lexicais, já que o que se busca em experimentos dessa modalidade são os efeitos de semelhança da estrutura ortográfica das palavras nos estágios iniciais do acesso lexical.

- **Priming morfológico** - se na relação de *priming* ortográfico, focalizamos um contexto de estrutura de escrita em busca de influências de unidades ortográficas de acesso, no *priming* morfológico buscamos, através da repetição de determinados morfemas como bases ou afixos, uma relação entre unidade visual e unidade morfológica de acesso lexical. Por exemplo, em uma série de tarefas de decisão lexical, nas quais o grau de similaridade ortográfica foi altamente controlado, Laudanna et al. (1989) encontraram que, em curtos intervalos de tempo, ocorria a facilitação no reconhecimento da **palavra-alvo** somente entre palavras cuja relação de *prime* era morfológica, e não apenas ortográfica. De modo que, entre pares de palavras da Língua Italiana como *posto* (lugar) X *posti* (lugares) e *collo* (pescoço) X *colpo* (golpe), somente o primeiro par de estímulos, formado por palavras que compartilham da mesma base (*post-*), apresenta uma resposta mais rápida ao teste de decisão lexical. Este indício sugere, portanto, que essas palavras compartilhem de uma mesma representação lexical. Os autores observaram ainda que o *prime* de uma base homográfica sem relação morfológica pode ter um efeito inibidor sobre o reconhecimento da palavra alvo, sugerindo que, mesmo quando idênticas na grafia, estas bases constituem diferentes representações lexicais. Assim, segundo Laudanna et al. (1989, op. cit.), quando duas palavras como *portare* (carregar) e *porte* (portas) são apresentadas como estímulos para uma tarefa de decisão lexical, na qual a segunda é o alvo, os sujeitos demoram mais para reconhecê-la do que quando precedida por uma palavra sem base homógrafa, levando à evidência de que uma facilitação no acesso só é obtida quando o *prime* e o alvo compartilham tanto a forma quanto o significado.
- **Priming fonológico** - nesta modalidade - a exemplo do estudo de Meyer, Schvaneveldt e Ruddy (1974) - apresenta-se aos sujeitos pares de palavras que abordam diferentes tipos

de relação, por exemplo: (a) uma ortografia semelhante, mas que sejam constituídos de formas fonológicas diferentes (*sede/rede*), (b) ortografia e forma fonológica semelhantes (*vela/tela*) e paradoxalmente (c) formas fonológicas e ortográficas diferentes (*casa/boto*). No caso do estudo citado, os autores conseguiram encontrar sujeitos que respondiam mais rápido ao teste de decisão lexical quando expostos a pares da relação de tipo (b) se comparados a pares não-relacionados de tipo (c) e produziam respostas mais lentas quando expostos a pares de tipo (a) em relação a pares não relacionados tipo (c). O objetivo de estudos como esse é provar a importância do código fonológico no reconhecimento visual das palavras.

- **Priming semântico (associativo)** - é a modalidade de *priming* mais utilizada dentro desta área de pesquisa. Sua “descoberta” se deve a Meyer & Schvaneveldt que, em 1971, encontraram indícios de que os sujeitos são mais rápidos para tomar decisões lexicais quando apresentados a pares de palavras semanticamente relacionadas do que a pares não-relacionados. Segundo Balota (1994), os experimentos que utilizam o efeito de *priming* semântico tornaram-se muito bem vindos na comunidade psicolinguística; pois a medida dependente utilizada - latência de resposta - era a base para os experimentos cognitivistas da época; um outro motivo de interesse é que o efeito do *priming* semântico é facilmente replicável e muito evidente. Além disso, a tarefa parece perfeita para a pesquisa linguística no que se refere ao mapeamento da arquitetura do nível das representações semânticas, como também das operações de recuperação que agem sobre estas representações. No entanto, os efeitos do *priming* semântico são muito questionados quando se envolve estudos que buscam a relação do contexto semântico e das relações de associação no reconhecimento da palavra. Por exemplo, palavras como *caminhão* e *caminhonete* são semanticamente relacionadas porque ambas pertencem a uma mesma categoria de automóveis. Contudo, se pensarmos em *caminhão* e *frete*, por exemplo, encontraremos uma relação apenas associativa, pois não compartilham de uma mesma categoria semântica e sequer compartilham muitos traços semânticos; mas não há como se negar que ambas têm uma relação de associação. Alguns estudos como os de

Palermo & Jenkins (1964) apontam, porém, para ocorrências em que palavras como *luva* e *chapéu* - que compartilham da mesma categoria - não costumam produzir uma forte relação de associação ao se realizarem testes cuja tarefa dos sujeitos seja gerar respostas associadas a partir de uma palavra apresentada como estímulo.

B. Morfo: um programa experimental para o estudo da prefixação no Português

1 Delineando um programa experimental: a criação de “Morfo”

“Morfo” é o nome escolhido para a série de estudos experimentais de decisão lexical que pretendemos realizar no intuito de testar, via *priming* semântico, as hipóteses sobre os mecanismos de acesso lexical das palavras morfologicamente complexas. Os experimentos estão, inicialmente, programados para a observação das duas categorias de prefixação que vimos apresentamos desde o início: transparente e pseudo-prefixada. Entretanto, é válido ressaltar que não se pretende dar conta de todas as perguntas em um único experimento. O estudo experimental, como vimos dizendo desde o início deste trabalho, requer um cuidado nunca excessivo com o que se pergunta aos informantes e com a maneira pela qual isto é perguntado; em geral não devemos nos precipitar ao ponto de incluir, por exemplo, mais do que duas variáveis de interesse em um único teste. Ainda assim, às vezes, há necessidade de restringirmo-nos apenas a uma variável de interesse em cada experimento.

Os dois experimentos que aplicamos, por exemplo, foram delineados sob essa perspectiva; o primeiro tendo como questão a existência, na Língua Portuguesa, um efeito de *priming* semântico; o segundo, se esse *priming* pode ocorrer em palavras morfologicamente complexas. No Experimento I, delineamos uma estrutura pouco complexa, trabalhando apenas com pares de substantivos e adjetivos, que tendem a ser morfologicamente mais simples do que os verbos - que antes da derivação ou flexão já apresentam uma desinência de infinitivo. Já o Experimento II foi projetado tendo em vista os resultados positivos do processo experimental anterior. Pudemos, portanto, acrescentar,

nesta segunda etapa, uma variável de interesse a mais, além de mudarmos a classe gramatical. No Experimento II, nosso interesse esteve centrado na existência ou não do efeito de *priming* semântico nas classes de palavras prefixadas assumidas no capítulo anterior. De forma que nele, utilizamo-nos, então, das palavras prefixadas, classificadas entre transparentes e pseudo-prefixadas.

Outro elemento a ser cuidadosamente trabalhado em um delineamento experimental para testes de decisão lexical são as não-palavras. No que concerne a esses estímulos é recomendável, em primeiro plano, que sejam formados por seqüências de letras que obedecem ao padrão de formação da língua em questão; ou seja, para o Português, por exemplo, os logatomas obedecem em sua maioria ao padrão silábico CVCV²⁸. No entanto, mesmo obedecendo a este padrão, poderemos ter, segundo Forster (1999, p.8), um efeito de *priming* de forma ou uma competição pela identificação, relacionados à dificuldade de discriminação da identificação entre os estímulos de palavra/não-palavra, podendo este efeito inclusive “transbordar” para os estímulos subseqüentes. A idéia de que há diferença nos critérios de identificação de não-palavras por parte dos sujeitos, em experimentos desta modalidade, é apontada por Forster como um dos pontos de um modelo de processamento proposto por Jacobs & Grainger (1992), e, basicamente, consiste na suposição de que quando uma discriminação é facilmente realizável, as decisões são baseadas na soma da ativação de todas as unidades de palavras. Simulações mostram que o *priming* nestes casos é equivalente para palavras e não-palavras; mas, quando uma discriminação mais apurada é requerida, os padrões de ativação são baseados localmente e, sob estas condições, Forster (op. cit.) afirma que novas simulações apontaram para um efeito de *priming* mais fraco para as palavras do que para as não-palavras. Uma opção para “forçar” os sujeitos a adotar análise mais criteriosa é mudar apenas uma letra das palavras para as não-palavras, por exemplo, adotamos como não-palavra a corruptela *LODAR, a partir do verbo RODAR; e não utilizamos não-palavras do tipo *PODOPAR.

²⁸ Em alguns deles colocamos encontros consonantais passíveis de serem encontrados do Português, como a letra *s* antes de *p* em **muspir*, **rospar* ambas derivadas das palavras *cuspir* e *raspar*, respectivamente. Ou ainda o encontro *pr* em **produvir*, *tansar* corrupção de *dansar*, cujo o encontro em questão é o *ns*. Estes são apenas alguns dentre os vários utilizados.

Além de trabalhar as questões lingüísticas em um experimento, o pesquisador jamais pode se esquecer da faceta perceptual de seu experimento, i. e., tem que estar igualmente atento às questões que dizem respeito à forma e ao tempo de apresentação dos estímulos visuais. Em Perea & Gotor (1997) e Bushell & Martin (1997), encontramos dados que nos mostram que os efeitos de *priming* semântico e associativo dissipam-se muito rapidamente. Portanto, se o intuito é estabelecer se há ou não efeito associativo, a apresentação de uma palavra detonadora do efeito de *priming*, nos testes de decisão lexical, tem que ocorrer muito rapidamente; de uma forma quase subliminar. Esta propriedade de distinção entre as apresentações mais longas ou mais curtas de um estímulo é conhecida como *Stimulus Onset Asynchrony*²⁹ (daqui por diante, SOA). Um SOA curto pode ter entre 30 ms e 250 ms, e sua contagem abrange desde o início da apresentação da palavra *prime* até o final da tela de intervalo que precede a palavra alvo. No primeiro experimento utilizamos um SOA de 240 ms e no Experimento II um SOA um pouco maior, de 250 ms.³⁰

Por fim, no Experimento I, a tarefa dos informantes é aquela apresentada neste capítulo na definição de um teste de decisão lexical: ao serem apresentados a um estímulo visual, os sujeitos têm de lê-lo e responder rapidamente se este é uma palavra de sua língua materna ou não. Nesta versão, no entanto, a primeira seqüência é apresentada por apenas 190 ms. Temos, então, um intervalo de 50 ms e a subsequente apresentação por 600 ms da palavra alvo da resposta dos informantes³¹ - esta **pode** ser um substantivo da nossa língua ou apenas um conjunto de letras que nada significam. Os sujeitos têm até 3000 ms para responder ao estímulo, pressionando os botões em uma caixa de respostas.

Para saber sobre a programação dos experimentos de Morfo, sua sintaxe e

²⁹ Observações sobre *Stimulus Onset Asynchronies* em pequenas e longas durações vêm sendo realizadas desde de o final da década de 70, e têm como referência inicial o trabalho de Fischler e Goodman (1978), no qual foi constatado efeito de *priming* associativo com SOA de 40ms, mas não com 90ms de SOA. Ainda como referência importante para este estudo temos o artigo de Neely (1991), no qual, além de tratar da diferença de resultados conforme o SOA, o autor também traça um panorama detalhado do estudo do reconhecimento visual das palavras isoladas.

³⁰ Não pudemos trabalhar com SOA's mais curtos ainda, porque o modo gráfico da ferramenta utilizada não comporta tempos mais baixos que estes. Futuramente, usando um programa como DMASTER, planejamos trabalhar com tempos de apresentação mais curtos e com experimento que utilizem máscara visual.

³¹ O programa EXPE6 não consegue lidar com a apresentação de tempos menores que este. Por este motivo, não

comandos, o leitor pode se dirigir aos textos anexos a essa dissertação. No Anexo I, encontra-se o artigo do programa computacional utilizado para o feitiço destes experimentos, o programa francês **Expe6**, cedido pelo “Laboratoire de Sciences Cognitives et Psycholinguistique” - EHESS - CNRS, Paris; além de um arquivo impresso - o qual foi parte integrante dos testes - com o material lingüístico de pesquisa aliado à programação computacional, de maneira que o leitor possa visualizar melhor a programação e a elaboração de um experimento psicolingüístico.

Traçados nossos limites básicos, descreveremos detalhadamente os experimentos I e II. A descrição destes dar-se-á em três etapas: primeiro apresentaremos as palavras com as quais trabalhamos e como balanceamos o experimento; em seguida o perfil dos nossos sujeitos e, por último, os dados obtidos e sua interpretação.

2. Experimento I

2.1 As palavras

Conforme foi dito acima, neste experimento, nosso material constituiu-se de pares de substantivos e adjetivos; estes pares são relacionados semanticamente e também são constituídos de palavras bastante freqüentes da Língua Portuguesa. Além dos pares experimentais e controles, temos de ter em um teste de decisão lexical, devidamente balanceado, pares de distratores que se dividem entre pares de palavras e pares de palavra e não-palavra. Esta exigência deve-se ao fato de que os sujeitos experimentais não podem perceber qual a real finalidade do teste; ou seja, para eles a pergunta a ser respondida é se o segundo estímulo apresentado é uma palavra do Português ou não. Neste experimento, de um total de 160 pares de estímulos, temos 40 pares testes e 120 pares de distratores divididos da seguinte maneira:

A. Dos 40 pares testes:

pudemos trabalhar com tempos de 60 a 90 ms, conforme a literatura sugere serem os ideais.

- 20 pares experimentais são formados por itens relacionados subdivididos entre *primes* e alvos. Ex: certo/errado;
- 20 pares são controles dos pares experimentais, formados pelos mesmos alvos, porém com *primes* diferentes. Ex: estrela/errado. É a comparação dos tempos de reação entre esses dois alvos em contextos associativos diferentes que nos interessam como ponto de análise.

B. Dos pares distratores³²:

- 20 pares são formados por palavras semanticamente relacionadas entre si, tais quais os experimentais, porém estes não têm seus respectivos controles e, portanto, não terão termos de comparação dos tempos de reação. Estão na lista apenas para distrair os informantes, e, assim como todos os outros pares descritos a seguir, comungam da função de não deixar que os sujeitos percebam quais são os pares que estão sendo testados.
- 20 pares são formados por pares não-relacionados de substantivos. Ex: janela/borracha;
- 80 pares distratores são formados por uma palavra *prime* e uma não-palavra - formada a partir de uma palavra da língua portuguesa na qual uma letra foi acrescentada, subtraída ou substituída. Ex: pedra/COPRO.

Temos abaixo as respectivas listas com os pares experimentais e controles.

³² Na literatura este termo é amplamente empregado em sua tradução inglesa *fillers*, mas optamos pelo uso do correspondente em português.

Tabela 2: Pares experimentais e controles apresentados no Experimento I

<i>Prime</i>	<i>Alvo</i>	<i>Controle</i>	<i>Alvo</i>
CERTO	ERRADO	UNHA	ERRADO
GATO	CACHORRO	SAPATO	CACHORRO
OURO	PRATA	LEITE	PRATA
CASA	COMIDA	ANEL	COMIDA
GORDO	MAGRO	JORNAL	MAGRO
DOCE	SALGADO	SORRISO	SALGADO
BATALHA	GUERRA	COPO	GUERRA
LIVRO	ESCOLA	MORNO	ESCOLA
BOM	RUIM	ESPADA	RUIM
PORTA	JANELA	PERNA	JANELA
CHEIO	VAZIO	PIMENTA	VAZIO
PALHAÇO	CIRCO	BOLSA	CIRCO
FLOR	FRUTO	TREM	FRUTO
CALÇA	CAMISA	POSTE	CAMISA
PAPEL	CANETA	BANANA	CANETA
PRETO	BRANCO	MENOR	BRANCO
HOMEM	MULHER	TIJOLO	MULHER
CLARO	ESCURO	ORELHA	ESCURO
GARFO	FACA	CHUVA	FACA
ALTO	BAIXO	FICHA	BAIXO

É válido lembrar, entretanto, que os pares de ambas as tabelas e os 120 estímulos restantes tiveram suas apresentações mescladas e aleatorizadas para cada um dos sujeitos. Além disso, pela possibilidade de criarmos um efeito de *repetition prime* escolhemos dentre várias aleatorizações aquelas em que os estímulos repetidos tivessem uma maior distância possível entre as duas apresentações. Este efeito de facilitação, traduzido pela aceleração do acesso lexical de um determinado item, é provocado pela reincidência próxima da apresentação de uma forma em uma mesma lista; segundo a literatura, podemos obter efeito de *priming* de repetição com a distância aproximada de 20 palavras entre as duas apresentações de um mesmo vocábulo. Se o leitor puder recordar, alguns modelos contam com “resíduos” de ativação pós-acesso, por exemplo, nos modelos Logogênicos os itens mantêm-se parcialmente ativados por um pequeno período de tempo, para que uma possível

reativação não implique na repetição de todo o processo. Já nos modelos conexionistas, é a mudança no peso das conexões o equivalente desse “residual” que pode facilitar a reativação de uma determinada unidade.

2.2 Os sujeitos

O perfil de nossos sujeitos já foi parcialmente delimitado a partir do momento em que determinamos que nossa pesquisa se daria sobre o léxico de adultos, com grau de escolaridade médio e, obrigatoriamente, falantes nativos da Língua Portuguesa. Na aplicação do primeiro experimento, contamos com apenas 5 informantes voluntários entre graduandos e pós-graduandos dos cursos de Letras e Linguística, deste Instituto, e alguns voluntários cujo grau de escolaridade constituía-se pelo menos da graduação incompleta.

2.3 O procedimento

A aplicação do teste é individual e inclui um rápido de treinamento, com 15 pares, sobre os quais o sujeito está autorizado a perguntar suas dúvidas ao assistente. O tempo total do teste está por volta de dez minutos, incluindo-se aqui o treinamento e o tempo (indeterminado) que os informantes levam para ler as instruções.

Apresenta-se, muito rapidamente, um par de seqüências de estímulos visuais aos sujeitos, conforme os parâmetros de tempo foram explicados acima. O primeiro estímulo (*prime*) sempre é uma palavra da Língua Portuguesa escrita em letras minúsculas, e o segundo estímulo (*alvo*) subsequente está escrito em letras maiúsculas.

2.4 Os resultados

Os dados referentes aos sujeitos são automaticamente armazenados pelo programa

em arquivos separados - de extensão *.res - com nomes correspondentes aos arquivos originais. Estes testes foram agrupados em um único arquivo, separados pelos códigos de cada informante e foram analisados estatisticamente através de uma ANOVA simples pelo programa *SigmaStat*. Neste teste, opusemos *primes* contra controles e obtivemos significância estatística ($P = 0,0017$). Esta significância só não foi maior porque a amostra, além de pequena, variou bastante em comportamento. Além disso, em análise estatística descritiva, a média de diferença de tempo de reação entre *primes* e alvos foi de 39 ms; respectivamente os tempos de reação médios foram de 512 ms e 551 ms.

Tabela 3: Média dos tempos de reação para substantivos e adjetivos em experimento de decisão lexical com SOA curto (ms)

Condição	Tempo
Relacionado	512
Não-relacionado	551
Diferença total	39

($P = 0,00178$)

Portanto, com base nestes resultados, pudemos constatar um efeito de aceleração nos substantivos e adjetivos quando precedidos por um termo associado semanticamente. Dessa forma, a obtenção positiva de um efeito de *priming* semântico nos “autorizou” a seguirmos para a nossa verdadeira questão de interesse: a existência de *priming* semântico nos verbos prefixados.

3. Experimento II

3.1 As palavras

O motivo inicial pelo qual escolhemos a prefixação como alvo de um estudo morfológico experimental é aparentemente simples: do ponto de vista de um processo

morfo-fonológico, a prefixação é um processo muito mais regular, partindo-se da perspectiva de que esta não traz muitas mudanças nas regiões de fronteiras de morfemas se comparada com a derivação; além de ser mais regular também do que a derivação e a composição no aspecto semântico da formação de novos vocábulos. Porém, dentro do próprio processo derivativo, procuramos por prefixos mais “regulares”, isto é, por prefixos que não variassem (ou pelo menos que o fizessem em menor grau) justamente nesses dois pontos: forma e significado. Para isso, procedemos em uma busca nas palavras prefixadas da Língua Portuguesa listadas no programa **Listas**, desenvolvido pela professora Eleonora Albano juntamente com uma equipe de pesquisadores do LAFAPE (Albano et alli, 2000), o qual tem como base de dados o *Minidicionário Aurélio*, listado e analisado sob vários aspectos gramaticais (categorias gramaticais), fonéticos (padrão acentual, transcrição fonética), formais (padrão silábico, número de sílabas) e que permite a geração de listas específicas do interesse do usuário.

É necessário esclarecer, no entanto, que, como buscamos informações com um fim muito específico para os prefixos, não pudemos encontrar na literatura da área de morfologia no Brasil nada tão abrangente que pudesse satisfazer a esses detalhes pelos quais pesquisávamos; portanto, obras como *Formação do Português Brasileiro e Competência Lexical*, de José Sandman (1989; 1991), foram consultadas mais como fonte de confirmação de alguns dados - como significados e ocorrências possíveis de alguns prefixos - do que como base para a pesquisa propriamente dita.

Em nosso levantamento prévio, que teve intuito seletivo e jamais ambicionou algo como descrever toda a prefixação da Língua Portuguesa, pudemos notar que os prefixos mais produtivos de nossa língua - e por conseqüência mais freqüentes - tendem a ser aqueles formados por duas e três letras como *pre-*, *re-*, *des-*, *de-*, *co-*, *sub-*. Essa produtividade decresce conforme aumenta a extensão do prefixo; qualquer leitor não-especialista pode observar que prefixos como *anti-*, *contra-*, *intra-*, *mega-*, entre outros, são menos freqüentemente encontrados e utilizados no dia-a-dia. Um segundo critério seletivo foi observar quais desses prefixos, dentre aqueles mais produtivos e freqüentes, tinham um processo de formação mais regular do ponto de vista do seu significado. Por exemplo, um

prefixo como *des-* pode assumir diferentes significados tais como “reverter uma ação” nas palavras *desfazer* e *desanimar*, “negação” em palavras como *desapego*, *desafeto* e afastamento ou retirada em *destampar*, *desvelar* e *desterrar*³³; já no caso dos prefixos *co-* e *sub-* o processo de formação, além de menos produtivo, gera palavras com forte fronteira morfo-fonológica, traduzida por vezes no sinal gráfico de hífen, como *subliminar*, *sub-mundo*, *co-articular*, *co-habitar* - o que para o nosso experimento não seria de muita valia, pois tornariam as palavras desta classe formal e visualmente diferentes das não-palavras e dos *primes*, podendo causar algum tipo de desvio indesejado.

Chegamos, por fim, à conclusão de que os prefixos *pre-* e *re-* eram os que mais se enquadravam nos padrões que delineamos para “Morfo” por serem, como procurávamos, muito freqüentes e produtivos em nossa língua, além de mais regulares quanto ao aspecto semântico. Tivemos ainda, como conseqüência de uma ocorrência maior de palavras pseudo-prefixadas, que optar pelo prefixo *re-* para nosso teste. Lembre-se que, neste momento, já tínhamos em vista um experimento com duas variáveis de interesse (palavras prefixadas transparentes e pseudo-prefixadas), portanto, não seria desejável introduzir um outro prefixo como uma variável a mais no experimento.

Tendo eleito o prefixo a ser pesquisado, partimos em busca de palavras que apresentassem pseudo-prefixação, ou seja, aquelas que formalmente parecem prefixadas mas não são, como *retratar*, *retrair*, *remediar*; verificando sempre nos dicionários a possibilidade de existência de uma acepção reconhecida como prefixada. Na etapa posterior, a busca se deu por verbos morfologicamente simples relacionados por meio de sinonímia ou de antonímia com a base e pseudo-base, contrariamente à fase seguinte, na qual o inventário se deu sobre verbos não-relacionados com as bases e pseudo bases. Realizamos, então, um pré-teste com os alunos do primeiro ano do curso de Lingüística deste Instituto para sabermos quais seriam os “melhores” pares associados dentre os que listamos. Este pré-teste foi constituído apenas pelas bases verbais das palavras experimentais a serem testadas pareadas com verbos relacionados semanticamente. Pedimos, então, aos alunos que atribuísem notas de zero a cinco às relações, tendo na nota

³³ Além de ter também variação formal, como os alomorfes *de-*, *dis-* e *di-*. Veja os respectivos exemplos:

três um valor mediano; ver as respectivas médias no Anexo III ao final deste texto.

Seguem abaixo as palavras prefixadas e seus respectivos pares relacionados, não relacionados e não-palavras utilizados em nosso experimento. As palavras prefixadas estão agrupadas conforme as duas classificações elaboradas em nossa hipótese: de prefixação transparente e pseudo-prefixadas.

Tabela 3: Pares de estímulos prefixados transparentes utilizados como testes no Experimento II e seus respectivos alvos

Verbos prefixados	Verbos relacionados com a base
REAGIR	ACIONAR
READAPTAR	ADEQUAR
REACENDER	APAGAR
REAJUSTAR	APERTAR
REAVALIAR	CALCULAR
RECONSTRUIR	DESTRUIR
RELEMBRAR	ESQUECER
REDEFINIR	EXPLICAR
REABRIR	FECHAR
RECAIR	LEVANTAR
REAFIRMAR	NEGAR
REALINHAR	ORGANIZAR
REDIVIDIR	PARTIR
RECONQUISTAR	VENCER
RECOMPRAR	VENDER

Tabela 4: Pares de estímulos das palavras prefixadas transparentes utilizados como controle no Experimento II

Verbos controles	Verbos relacionados com a base
SONHAR	ACIONAR
ESFOLAR	ADEQUAR
PEDALAR	APAGAR
BRINDAR	APERTAR
MATAR	CALCULAR
PEDIR	DESTRUIR
BOCEJAR	ESQUECER
CANTAR	EXPLICAR
ESCORRER	FECHAR
ENTREGAR	LEVANTAR
SORRIR	NEGAR
APALPAR	ORGANIZAR
BABAR	PARTIR
TRANCAR	VENCER
SENTIR	VENDER

Tabela 5: Pares de estímulos pseudo-prefixados utilizados como testes no Experimento II e seus respectivos alvos

Verbos pseudo-prefixados	Verbos relacionados c/ a pseudo-base
REPOUSAR	ATERRISSAR
REPORTAR	CARREGAR
REVELAR	CUIDAR
REVERTER	DERRAMAR
REQUERER	DESEJAR
REFLETIR	DOBRAR
RETRAIR	ENGANAR
RETALHAR	ESCULPIR
REQUEBRAR	ESTRAGAR
REMEDIAR	INTERVIR
REFERIR	MACHUCAR
REMOVER	MEXER
RECOBRAR	PAGAR
RECORRER	PARAR
RELEVAR	TRAZER

Tabela 6: Pares de estímulos utilizados como controle no Experimento II e seus respectivos

alvos

Verbos controle	Verbos relacionados c/ a pseudo-base
DIGITAR	ATERRISSAR
FOFOCAR	CARREGAR
GRAVAR	CUIDAR
APROVAR	DERRAMAR
CRITICAR	DESEJAR
CORRIGIR	DOBRAR
COBRIR	ENGANAR
CHEIRAR	ESCULPIR
ENCOSTAR	ESTRAGAR
DESTACAR	INTERVIR
ESCOLHER	MACHUCAR
BUSCAR	MEXER
DEFENDER	PAGAR
GANHAR	PARAR
ACERTAR	TRAZER

Após a seleção dos melhores pares, partimos para a “invenção” das não-palavras, processo delimitado pela perspectiva de que os logotomas poderiam ter apenas uma letra de diferença de palavras da Língua Portuguesa.

O Experimento II totalizou um número maior de pares experimentais e, por consequência do balanceamento, teve também mais pares de controle e de distratores, constituindo um total de 240 pares de estímulos divididos da seguinte maneira:

A. Dos 60 pares testes:

- 15 pares experimentais são formados por palavras prefixadas transparentes e os verbos relacionados às suas respectivas bases subdivididos entre *primes* e alvos exemplificados na tabela 3;
- 15 pares experimentais são formados por palavras pseudo-prefixadas e os verbos relacionados à suas respectivas pseudo-bases subdivididos entre *primes* e alvos exemplificados na tabela 5;
- 15 pares são controles dos pares experimentais de prefixação transparente, formados pelos mesmos alvos, porém com *primes* diferentes, exemplificados na tabela 4;

- 15 pares são controles dos pares experimentais pseudo-prefixados formados pelos mesmos alvos, porém com *primes* diferentes, exemplificados na tabela 6. É a comparação dos tempos controle versus *prime* que nos interessam como medida.

B. Dos pares distratores:

- 60 pares dos quais 30 são formados por verbos relacionados entre si, tais quais os experimentais, e 30 são formados por pares de verbos sem qualquer relação de significado, porém ambos não têm seus respectivos controles e, portanto, não terão termos de comparação dos tempos de reação;
- 120 pares distratores são formados por um verbo *prime* e uma não-palavra. Ex: nadar/CODER.

3.2 Os sujeitos

O perfil de nossos sujeitos é o mesmo do Experimento I: adultos, com grau de escolaridade médio (i.e., com pelo menos o atual ensino médio completo) e, obrigatoriamente, falantes nativos da Língua Portuguesa. Na aplicação deste experimento, entretanto, contamos com uma mostra maior: 11 informantes voluntários cujo grau de escolaridade constituía-se pelo menos da graduação incompleta.

3.3 O procedimento

A aplicação do teste é individual e inclui um rápido de treinamento, com 15 pares, sobre os quais o sujeito está autorizado a perguntar suas dúvidas ao assistente. O tempo médio para a execução do teste está por volta de treze minutos, incluindo-se aqui o treinamento e o tempo (indeterminado) que os informantes levam para ler as instruções.

Os sujeitos são muito rapidamente apresentados a um par de seqüências de

estímulos visuais, conforme os parâmetros de tempo explicados acima. Novamente, o primeiro estímulo (*prime*) sempre é uma palavra da Língua Portuguesa escrita em letras minúsculas, e o segundo estímulo (alvo), como é de praxe nos experimentos de *prime*, é apresentado em letras maiúsculas.

4.4 Os resultados

Os dados referentes aos sujeitos foram automaticamente armazenados pelo programa em arquivos separados - de extensão *.res - com nomes correspondentes aos arquivos originais, conforme no Experimento I, já que este é um procedimento rotineiro do próprio EXPE. Estes testes foram agrupados em um único arquivo, separados pelos códigos de cada informante e foram analisados estatisticamente através de uma ANOVA de duplo fator (*Two way Anova*). Neste teste, foram opostos os tempos *primes* contra controles dentro de cada grupo, ou seja, para a verificação do efeito de *priming* semântico com duas variáveis, os dados são comparados dentro de sua própria condição de controle e alvo, mas também entre as condições de transparência e não-transparência. Obtivemos significância estatística marginal ($p=0,0636$), conforme suspeitávamos, apenas nas palavras prefixadas transparentes, ou seja, nas palavras realmente prefixadas. Na pseudo-prefixação não obtivemos significância estatística alguma ($p=0,4454$). Estes resultados apontam para um efeito de *priming* semântico apenas para as palavras cujo o processo morfológico de prefixação é transparente, ou seja, cujos constituintes podem ser analisados em diferentes morfemas durante o acesso lexical.

Tabela 4: Índice de significância estatística na comparação entre as médias (em ms.) das duas condições de apresentação em experimento de decisão lexical com SOA curto (ms.) e *priming*.

Condição	Médias experimentais	Médias controles	Diferenças
Prefixadas Transparentes P=0.0636	629	656	27
Pseudo-prefixadas P=0,4454	691	704	13

*p só é marginalmente significante se $p < 0,09999$

CAPÍTULO IV

Discussão Final

As primeiras respostas para as nossas perguntas iniciais emergiram deste estudo experimental. Obtivemos efeito de *priming* semântico com itens com SOA's curtos (aproximadamente 250 ms.); este é um experimento inédito no Português, posto que não havia nenhum estudo conhecido desta ordem aqui no Brasil. Em segundo lugar, e mais importante ainda, obtivemos efeito de *priming* semântico com SOA curto com um material morfológicamente complexo.

Relembrando nossas hipóteses, no início do trabalho, tínhamos que, segundo o modelo *Augmented Addressed Morphology*, a informação morfológica pode ser representada de duas maneiras diferentes no léxico do acesso: uma analisada entre unidades de acesso lexical, como afixos e bases, e outra com a forma integral da palavra, decorrendo daí a possibilidade de um acesso em duas vias também para as palavras morfológicamente complexas. Posto que nosso objeto de interesse é o processamento morfológico da Língua Portuguesa, as questões restringiram-se, então, ao acesso das palavras prefixadas, classificadas segundo a transparência morfológica e semântica do processo de prefixação. Até que ponto a forma do estímulo ortográfico poderia influenciar na ativação de unidades semanticamente relacionadas à forma da base verbal daquelas palavras prefixadas? Segundo a nossa "taxonomia", quando a prefixação fosse semântica e morfológicamente transparente, os verbos prefixados, apresentados previamente a bases verbais relacionadas em significação, deveriam acelerar o acesso lexical destas últimas. Este efeito seria traduzido por um tempo de reação menor dos informantes perante este contexto do que em contextos de não relação. Obtivemos resultados estatísticos marginalmente significativos de que o acesso de uma palavra prefixada acelera o acesso de um outro termo relacionado. Esta marginalidade traduzida em um índice de significância $p = 0,0636$, posto que este índice é considerado incontestavelmente significativo quando $p < 0,05$, pode ser consequência da falta de um controle rigoroso da frequência das palavras durante a elaboração da lista experimental. Pois dados de um recente experimento apresentado por Schreuder, Burani & Baayen (no prelo), no qual a frequência foi controlada em todas as

suas dimensões - ou seja em relação à ocorrência da palavra, à ocorrência do lema-base (palavra e suas flexões) e a frequência de família, na qual são levadas em conta também as derivações e as composições - os autores obtiveram resultados significantes do ponto-de-vista estatísticos para ambas as prefixações.

Já as palavras pseudo-prefixadas poderiam apresentar aceleração nos tempos de reação, caso prevalecesse a análise morfológica no lugar do acesso integral da palavra, o que levaria à ativação das palavras relacionadas à pseudo base. Ou, ainda, poderiam apresentar tempos de reação semelhantes ao contexto de não-relação, caso não houvesse acesso analisado, prevalecendo o acesso à entrada integral da palavra, sem ativação das palavras relacionadas à pseudo-base. Os dados experimentais apontam para conclusões mais parecidas com a segunda opção; aparentemente não há acesso analisado destas palavras, mas um acesso apenas à unidade plena da palavra.

Por fim, temos em vista a continuidade de nosso estudo experimental e teórico no doutoramento com a busca por resultados, os quais nos dão indícios sobre o “comportamento” dos verbos de prefixação opaca em contexto de associação semântica. Nossa intuição, baseada em modelos como AAM e artigos como *Distributional Properties*, nos diz que serão os fatores vinculados, por exemplo, à frequência de uso das palavras, os determinantes das diferenças representacionais. Palavras como *prever*, *pressentir* continuam sendo uma questão de interesse para nossa pesquisa; não a perdemos de vista, apenas adiamos a implementação de experimentos e de reflexões, pois não tínhamos as respostas às primeiras questões. Aquele que trabalha com a ciência experimental tem sempre de refrear a sua ânsia por respostas às questões mais complexas, pois ao pular etapas experimentais e desdenhar perguntas aparentemente menores, acaba por perder os parâmetros necessários para a escolha e análise dos seus dados. Nosso próximo passo nesta pesquisa é justamente verificar se o tempo de reação de verbos como aqueles acima citados será semelhante aos tempos dos verbos de prefixação transparente.

Bibliografia

- ALBANO, E., MOREIRA, E., SILVA, A., AQUINO, P, KAKINOHANA, R. Um conversor ortográfico-fônico e uma notação prosódica mínima para a síntese de fala em língua portuguesa. In: Scarpa, E., org., (2000).
- ARONOFF, M. (1976) Word Formation in Generative Grammar . *Linguistic Inquiry Monograph, vol. 1*. MIT Press, Cambridge.
- BAAYEN, R. H. & SCHREUDER, R. (1995). Modelling the processing of morphologically complex words. In: Dijkstra, T. & Smedt, K., eds., (1996).
- BALOTA, D. A. (1994). Visual word recognition. In: In: Gernsbacher, M. A., ed.,(1994).
- BESNER, D. & HUMPHREYS, G., eds., (1991). *Basic processes in reading? Visual word recognition*. Hillsdale, NJ: Erlbaum.
- BUSHELL, C. & MARTIN, A. (1997). Automatic semantic priming os nouns and verbs in patients with Alzhiemer's disease. *Neuropsychologia*, 35 (8), pp 1059-1067.
- BLOOMFIELD, L.(1930) *Language*. Henry Holt and Company, New York.
- CAMARA JR., J. M. (1975) *Estrutura da Língua Portuguesa*. Rio de Janeiro: Vozes.
- CARAMAZZA, A., LAUDANNNA, A. & ROMANI, C. (1988) Lexical access and inflectional morphology. In: *Cognition*, 28, pp. 297-332.

- CHIALANT, D. and CARAMAZZA, A. Where is morphology and how is it processed?
The case of written word recognition. In: Feldman, L. B., ed., (1995), pp. 55-76
- COOPER, W. E. & WALKER, E. C. T., ed., (1979) *Sentence Processing: Psycholinguistic studies presented to Merrill Garret* Hillsdale, NJ: Erlbaum.
- DIJKSTRA, T. & SMEDT, K. (eds.), (1996). *Computational Psycholinguistics*. London: Taylor and Francis.
- FELDMAN, L. B., ed., (1995). *Morphological Aspects of Language Processing*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- FODOR, J. A. (1983) *The Modularity of Mind*. Cambridge, MA: MIT Press.
- FORSTER, K. I. Accessing the mental lexicon. In: Wales, R. J. and Walker, E., eds., (1976).
- _____. Levels of processing and the structure of the language processor. In: Cooper, W. E. & Walker, E. C. T., eds., (1979).
- _____. Lexical Processing. In: Osherson, D. and Lasnik, H., eds., (1995), pp. 95-131.
- _____. (1999) The microgenesis of priming effects in lexical access. *Brain and Language*. 68, pp. 05-15.
- GARMAN, M. (1990) *Psycholinguistics*. Cambridge: Cambridge University Press
- GARNHAM, A. (1989) *Psycholinguistics: Central Topics*. London: Routledge.

- GERNSBACHER, M. A. (1994) *Handbook of Psycholinguistics*. London: Academic Press.
- JACOBS, A. M. & GRAINGER, J. (1992) Testing a semistochastic variant of the Interactive Activation Model in different word recognition experiments. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 1174-1188.
- HENDERSON, L. On mental representation of morphology and its diagnosis by measures of visual access speed. In: MARSLEN-WILSON, W., ed., (1989), pp. 357-389.
- LAUDANNA, A. BADECKER, W. & CARAMMAZZA, A.(1989) *Priming homographic stems*. *Journal of Memory and Language*, 28, pp. 531-546.
- LAUDANNA, A. and BURANI, C. Distributional properties of derivational affixes: implications for processing. In: FELDMAN, L. B., ed., (1995), pp. 345-364.
- LAUDANNA, A.; BURANI, C. & CERMELE, A. Prefixes as processing units. *Language and Cognitive Processes*, 9 (3), (1994), pp. 295-316.
- LEVELT, W. J. M. (1989) *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press
- LIMA, P.L.C. & FRANÇOZO, E. Metáfora e memória de trabalho. *Revista Palavra*. (no prelo)
- LIVELY, S. E., PISONI, D. B. and GOLDINGER, S. D. Spoken Word Recognition: research and Theory. In: Gernsbacher, M. A., ed.,(1994), pp 265-301.
- MARSLEN-WILSON, W. (1989) *Lexical Representation and Process*. Cambridge, MA: MIT Press

- MCCLELLAND, J. L (1979) On the time relation of mental processes: an examination of systems of processes in cascade. *Psychological Review*, 86, pp. 287-330.
- MCCLELLAND, J. L. & RUMELHART, D. E. (1981) An interactive model of context effects in letter perception. Part 1: an account of basic findings. *Psychological Review*, 88, pp. 375-407.
- MCCLELLAND, J. L. & RUMELHART, D. E., eds., (1986) *Parallel distributed processes: Exploration in the microstructure of Cognition*, Vol. 2 Cambridge, MA: MIT Press
- MCCLELLAND, J. & L RUMELHART, D. E. (1986) The appeal of parallel distributed processing. In: McClelland, J. L. & Rumelhart, D. E., eds., (1986), pp. 03-44.
- MEYER, D. E, SCHVANEVELDT, R. W. & RUDY, M. G. (1974). Functions of graphemic and phonemic codes in visual word recognition. *Memory & Cognition*, 2, pp. 309-321.
- MORTON, J. Word recognition. In: Morton, J. & Marshall, J. eds., (1970), pp 107-156.
- _____ & MARSHALL, J., eds., (1970) *Psycholinguistics 2: Structures and processes*. Cambridge, MA: MIT Press.
- MURREL, G. A. & MORTON, J. (1974). Word recognition and morphemic structure. *Journal of Experimental Psychology*, 102, pp. 963-968.
- NEELY, J. H. Semantic priming effect in visual word recognition. In: Besner, D. & Humphreys, G., eds., (1991), pp. 236-264.

- OSHERSON, D. and LASNIK, H. (1995) *Language: An Invitation to Cognitive Science*, vol. 1. Cambridge, MA: MIT Press.
- PALERMO, D. S. & JENKINS, J. J. (1964). Word association norms: Grade school through college. In Balota, D (1994)
- PALLIER, C., DUPOUX, E. & JEANNIN, X. EXPE: an expandable programming language for on-line psychological experiments. *Behaviour Research Methods, Instruments and Computer*, 1997, 29 (3), 322-327.
- PEREA, M. & GOTOR, A. (1997). Associative and semantic priming effects occur at a very short stimulus-onset asynchronies in lexical decision and naming. *Cognition*, 62, pp. 223-240.
- RUMELHART, D. E. & MCCLELLAND, J. L. On learning the past perfect of English verbs. In J. L. McClelland & D. E. Rumelhart , ed., (1986).
- SANDMANN, A. J. (1989) *Formação das Palavras no Português Brasileiro Contemporâneo*. Curitiba: Editora UFPR
- SANDMANN, A. J. (1991). *A Competência Lexical*. Curitiba: Editora UFPR
- SANDRA, D.(1994). The morphology of the mental lexicon: internal word structure viewed from a psycholinguistic perspective. *Language and Cognitive Processes*, 9 (3), 227-269.
- SWINNEY, D. A. (1979). Lexical access during sentence comprehension: (re)consideration of context effects. *Journal of Verbal Learning and Verbal Behaviour* 18, 545-569.

SCARPA, E. (org.) (2000). *Estudos de Prosódia no Brasil*. Campinas: Editora da UNICAMP.

SCHREUDER, R. (1994). Prefix stripping revisited. *Journal of Memory and Language*, 33, pp. 357-375.

STANNERS, R. F.; NEISSER, J.J AND PAINTON, S. (1979). Memory representation for prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 18, pp. 733-743.

TAFT, M (1979). Recognition of affixed words and the word frequency effect. In: *Memory and Cognition*, 7, 263-272.

TAFT, M & FORSTER, K.I. Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 14, 638-647.

WALES, R. J. AND WALKER, E., eds., (1976). *New approaches to language mechanisms*. Amsterdam: North-Holland.

WICKELGREN, W. A. (1969) Context-sensitive coding, associative memory and serial order in (speech) behaviour .Apud: J. L. McClelland & D. E. Rumelhart, eds., (1986).

ANEXO I

EXPE³⁴

O programa EXPE utilizado em nossa pesquisa é uma ferramenta computacional - delineada por Christophe Pallier, Emmanuel Dupoux e Xavier Jeannin, no “Laboratoire de Sciences Cognitives et Psycholinguistique ”/ EHESS - CNRS, Paris - para preparar e realizar experimentos psicolinguísticos que necessitem coletar respostas *on-line* ou *off-line* dos sujeitos experimentais, ou seja, dados colhidos no momento do processamento lexical ou posteriores a ele. EXPE pode ser usado para elaborar uma série de experimentos psicolinguísticos como decisão lexical aliados ou não a um efeito de *priming*, monitoramento de fonemas, que envolvam a coleta dos tempos de resposta dos sujeitos, tendo também a opção de trabalho com estímulos visuais ou auditivos.

Normalmente, quando se trabalha com pesquisa para a qual seja necessário usar experimentos psicológicos com equipamentos computadorizados temos que optar entre o uso de linguagens de programação como Pascal, C - as quais, por não serem criadas exatamente para a descrever experimentos, acabam requerendo, portanto, um nível muito elevado de conhecimento em programação para um pesquisador não-especialista em computação, além de muitos detalhes a serem observados quando há necessidade de sincronia e precisão quanto ao tempo dos resultados. Pode-se ainda optar por softwares específicos para a pesquisa experimental já existentes no mercado; estes seriam perfeitos se não fossem tão específicos, limitando por vezes o delineamento experimental ou a fonte de estímulo pesquisada, tornando-se exclusivos a apenas um tipo de experimento. A ferramenta Expe apresenta-nos algumas diferenças em relação a esses pacotes de programas atuais e também em relação às linguagens de programação, pois foi desenvolvido no intuito de manter as características de uma linguagem computacional; porém, com adaptações que tornam a sua programação uma tarefa mais acessível e de fácil manejo - principalmente para aqueles que trabalham com a pesquisa experimental, mas não têm o hábito da

³⁴ Este anexo está baseado no artigo “EXPE: an expandable programming language for on-line psychological experiments” de Pallier, C. et al. (1997), com o acréscimo de algumas informações nossas adquiridas através da experiência no trabalho com o programa.

programação computacional.

O programa vem acompanhado de arquivos tutores e de um manual, os quais ajudam aos leigos em programação na compreensão dos comandos e de alguns dos mecanismos necessários para se elaborar o experimento. Além disso, EXPE pode ser usado, ainda que precariamente, em equipamentos de tecnologia defasada como PC's antigos com 640 kb de memória. Fator que o torna mais acessível se comparado as ferramentas de mercado mais modernas, principalmente aquelas que - igualmente ao EXPE - possuam interface gráfica para a elaboração dos experimentos. Apesar de que a máquina ideal para que se “rodem” experimentos com funções de tempo real, seja no mínimo um modelo 386SX20 - pois a lentidão no processamento dos dados pode invalidar os resultados experimentais.

A sintaxe do programa³⁵ é “minimalista”, segundo os autores, pois o programa consiste em séries de linhas que podem ser linhas de comando, ou de dados ou de comentários; com as duas últimas caracterizadas por convenções específicas. Por exemplo, se estivermos iniciando uma base de dados, há de se acrescentar antes o termo “data” entre aspas, e se estivermos introduzindo um comentário como: *a primeira coluna traz somente substantivos de alta frequência, enquanto a segunda, os de baixa frequência*, para que este não seja levado em conta pelo programa temos que iniciá-lo pelo sinal de porcentagem (%). Existem comandos que permitem que mecanismos cíclicos (“*loopdata*”) e de aleatorização (“*shuffle*”), entre outros sejam inseridos na programação do experimento. Esta última tarefa pode ser importante dentro de uma estratégia experimental, se devidamente planejada pelo pesquisador, para que se evitem desvios na apresentação dos dados. Quando não se trata de um valor atribuído a uma variável, como “Verdadeiro” ou “falso”, não há padrão para a escrita dos comandos em caixa alta ou baixa; ou seja, podemos escrever *dAtA*, *Data*, *DATA*, *data*, que se o termo não estiver entre aspas a máquina poderá processá-lo da mesma maneira.

Na apresentação dos estímulos, o programa permite que experimentos como os de *Gating* sejam realizados, pois pode apresentar estímulos sonoros, como também vários

³⁵ Assumimos aqui a acepção técnica do termo da programação em computação.

eventos de ordem acústica em sincronia a estes estímulos. Entretanto, conforme dissemos anteriormente, é aconselhável que uma máquina com maior capacidade de memória seja utilizada, de forma que a mistura dos sons não ocorra *off-line* devido à lentidão desta. No entanto, a função de vídeo do programa ainda é consideravelmente primitiva, somente as funções básicas do Borland Pascal são disponíveis. Isto pode criar alguns problemas inesperados, por exemplo, para a elaboração de “Morfo” usamos algumas palavras que, por serem acentuadas ou terem cedilha - como *alçar* e *pôr*, necessitaram caracteres gráficos especiais; pois, no momento em que eram apresentadas na interface gráfica apresentavam-se com caracteres ininteligíveis. Também ainda não é possível aliar apresentar estímulos sonoros e visuais em conjunto, esta é uma expectativa dos autores para versões futuras do programa.

Os comandos relativos aos tempos de resposta e de intervalo são facilmente manipuláveis; com eles, no caso do programa experimental “Morfo” - que se utilizou de efeitos de *priming* - pudemos estipular tempos de apresentação das palavras-alvo, dos *primes*, um tempo de intervalo entre os dois estímulos e o tempo de resposta para os sujeitos em cada apresentação. Assim, conforme descrevemos no capítulo II, conseguimos apresentar os estímulos em pares cuja duração era sempre a mesma na tela por meio de comandos como “wait” e “waittill” e de variáveis como “inter_trial interval” (para a duração do intervalo entre as duas seqüências), “blank duration” (para a duração da tela de intervalo), etc. Em relação à captação do tempo de reação do sujeitos - que no nosso caso foi uma medida de uma reação em processo (on-line) e não de dados analisados em processamento pós-lexical (off-line), para isso estipulamos um tempo máximo para que o sujeito respondesse ao estímulo - pode ser melhor comensurado se o pesquisador usar uma caixa de resposta em uma porta paralela do computador, pois o teclado pode medir também os tempos de reação, porém com menor exatidão do que uma caixa de respostas. Os resultados são gravados em arquivos de formato ASCII e em arquivos de extensão “.res” de especificação própria, de forma que há duas fontes de resultados, cabe ao pesquisador a escolha de consulta, como também a escolha de gravação, já que cada sujeito pode gerar um arquivo de respostas próprio ou todos os sujeitos podem ter seus TR (tempos de reação)

registrados em um mesmo arquivo.

O programa EXPE está facilmente disponível, pode ser conseguido gratuitamente via rede (*internet*) na URL <http://ruccs.rutgers.edu/pallier> ou pelo site do LSCP, <http://www.ehess.fr/centres/lscp/expe/expe.html>, acompanhando a documentação que se constitui de um manual, alguns arquivos tutores e artigo. Os autores apenas pedem que os pesquisadores citem o artigo-fonte deste anexo sempre que se reportarem ao seu trabalho.

ANEXO II

Experimento I: Priming semântico com substantivos e adjetivos

Exemplo de programação dos dados em EXPE6.

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%% Tutor05.pro -- created by Dpx, Oct 1997% %%%%%%%%%
%%%%%%%%% novo: Short.pro -- modificado por LGR, %%%%%%%%%
%%%%%%%%% Setembro/2000 %%%%%%%%%
%%%%%%%%% Instruções por Luciana Garcia Ruiz %%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
%%%%%%%%% UM PRIME SEMÂNTICO ENVOLVENDO UM TAREFA DE DECISÃO
%%%%%%%%% LEXICAL COM SOA CURTO
```

```
%%%%%%%%% Conteúdo: Este arquivo apresentará uma palavra
%%%%%%%%% durante 190 ms., seguida de um 50 msec. de
%%%%%%%%% intervalo e de uma palavra-alvo (por mais 700 ms.,
%%%%%%%%% no máximo).
```

```
%%%%%%%%% A tarefa consiste em desempenhar uma decisão lexical
%%%%%%%%% acelerada no segundo item.
```

```
%%%%%%%%% O tempo final de resposta não ultrapassa 3000 ms.
%%%%%%%%% RTS (tempos de reação) são coletados e armazenados.
```

```
%%%%%%%%% Keywords: call, procedure, endproc, waittill,
%%%%%%%%% variables, goto
%%%%%%%%% To run: EXPE6 short subject_name
```

```
defvar rightkey leftkey resp_certa resp_sujeito
defvar prime_duration blank_duration target_duration
rt_deadline inter_trial_interval
defvar a
defvar b
```

```
prime_duration:=190
blank_duration:=50
target_duration:=700
rt_deadline:=3000
inter_trial_interval:=1500
```

```
call "instruções"
```

```
graphmode
```

```
%%%%%%%%% Esta , a sessão de treinamento %%%%%%%%%
```

```

WithData "treinamento"
    wait inter_trial_interval

    onpage 1 cls
    onpage 1 writexy centerx centery #1
    showpage 1
    wait prime_duration
    a:=getonset showpage 0
    onpage 1 cls
    onpage 1 writexy centerx centery #2
    enablebutton
    waittill a+blank_duration
    rtonset:=getonset showpage 1
    waitwhile (not buttonpressed) and
(clock<rtonset+target_duration)
    showpage 0
    waitwhile (not buttonpressed) and
(clock<rtonset+rt_deadline)

    disablebutton

```

```

LoopData

```

```

%%%%%%%%%% Esta é a Sessão Experimental %%%%%%%%%%

```

```

cls
writexy centerx centery - 3 line "E agora vamos ao
experimento real!"
writexy centerx maxy - 3 line "...Aperte uma tecla para
continuar..."

```

```

readbutton

```

```

WithData "dados"

```

```

cls

```

```

    wait inter_trial_interval
    onpage 1 cls
    onpage 1 Writexy centerx centery #1

    showpage 1
    wait prime_duration
    a:=getonset showpage 0
    onpage 1 cls

```

```

onpage 1 writexy centerx centery #2
enablebutton

waittill a+blank_duration

rt onset:=getonset showpage 1

waitwhile (not ButtonPressed) and
(clock<RtOnset+target_duration)

showpage 0

WaitWhile (not ButtonPressed) and
(clock<RtOnset+rt_deadline)

DisableButton

wait 100
resp_sujeito:="b"
if button==rightkey resp_sujeito:="word"
if button==leftkey resp_sujeito:="nonword"
resp_certa:=(resp_sujeito == #5)

save #0 button rt resp_sujeito resp_certa
showpage 0

```

LoopData

TextMode

%%MATERIAIS%%

Data "treinamento"

pato	LAGO	Word	Related
cheiro	NERATU	Nonword	N/A
macaco	BOLA	Word	Unrelated
barco	COLTA	Nonword	N/A
banana	RORVE	Nonword	N/A
folha	RODA	Word	Unrelated
carro	FREIO	Word	Related
olho	ESBUGA	Nonword	N/A
bolo	CRIANÇA	Word	Related
teclado	CANETA	Word	Unrelated
caneta	PORTA	Word	Unrelated
chave	COTUGA	Nonword	N/A
sacola	TREDUTA	Nonword	N/A
tela	PAPEL	Word	Unrelated
colcha	CAMA	Word	Related

EndData

Data "dados"

alegria	CORTINA	distrator2	word	word
perto	LONGE	distrator1	word	word
conjunto	ATULTO	distrator3	word	nonword
fumaça	FOGO	distrator1	word	word
toalha	BANHO	distrator1	word	word
forte	OGUA	distrator3	word	nonword
ponto	PAZUL	distrator3	word	nonword
pijama	CERRENTE	distrator3	word	nonword
doença	ESPOTA	distrator3	word	nonword
jogo	BOLA	distrator1	word	word
cabelo	PENTE	distrator1	word	word
água	GELO	distrator1	word	word
livro	ESCOLA	experim	word	word
órgão	TERVOSO	distrator3	word	nonword
lista	GRALHO	distrator3	word	nonword
espada	RUIM	controle	word	word
ciência	FORRO	distrator2	word	word
choro	CELCHA	distrator3	word	nonword
lápiz	BORRACHA	experim	word	word
desenho	PRODA	distrator3	word	nonword
prova	PANTE	distrator3	word	nonword
certo	ERRADO	experim	word	word
envelope	SEMANA	distrator2	word	word
cheque	BEJO	distrator3	word	nonword
curto	LONGO	distrator1	word	word
roxo	SUCA	distrator3	word	nonword
viagem	PARRO	distrator3	word	nonword
bom	RUIM	experim	word	word
grama	MOSICA	distrator3	word	nonword
tarde	CONETA	distrator3	word	nonword
ponte	PETADE	distrator3	word	nonword
prisão	CARMORE	distrator3	word	nonword
sorriso	SALGADO	controle	word	word
redondo	QUADRADO	distrator1	word	word
amor	PORATA	distrator3	word	nonword
tontura	CONTRALE	distrator3	word	nonword
porta	JANELA	experim	word	word
bermuda	ENFERMEIRA	controle	word	word
morena	CESTU	distrator3	word	nonword
ficha	BAIXO	controle	word	word
ciúme	PORVETE	distrator3	word	nonword
corajem	MEDO	distrator1	word	word
árvore	BRANCO	controle	word	word
sonho	PAROL	distrator3	word	nonword

mapa	CROVO	distrator3	word	nonword
pincel	TINTA	distrator1	word	word
rosa	FIPRA	distrator3	word	nonword
loja	PIPOCA	distrator2	word	word
café	LEITE	experim	word	word
sábado	CACHORRO	controle	word	word
onda	TEPINO	distrator3	word	nonword
boca	FURVA	distrator3	word	nonword
pesca	BOSE	distrator3	word	nonword
antena	LUVEM	distrator3	word	nonword
claro	ESCURO	experim	word	word
bolacha	CORRIDA	distrator2	word	word
chave	GAVETA	distrator1	word	word
disco	FRUTA	distrator3	word	nonword
tênis	MECACO	distrator3	word	nonword
esperto	GERRADO	distrator3	word	nonword
orelha	ESCURO	controle	word	word
chão	PASTA	distrator2	word	word
cartão	TALOR	distrator3	word	nonword
plástico	CALHER	distrator3	word	nonword
tampa	CALO	distrator2	word	word
palhaço	CIRCO	experim	word	word
médico	ENFERMEIRA	experim	word	word
vaca	MOXTA	distrator3	word	nonword
pano	FELHA	distrator3	word	nonword
estoujo	MOCHILA	distrator1	word	word
anel	COMIDA	controle	word	word
teatro	PRAVA	distrator3	word	nonword
mola	FIGUGA	distrator3	word	nonword
preto	BRANCO	experim	word	word
nota	CENETA	distrator3	word	nonword
começo	BAITO	distrator3	word	nonword
página	CANEMA	distrator3	word	nonword
garfo	FACA	experim	word	word
quarto	MAPEL	distrator3	word	nonword
banana	CANETA	controle	word	word
forno	BISTA	distrator3	word	nonword
perfume	COMETA	distrator2	word	word
bebida	FLOR	distrator2	word	word
doce	SALGADO	experim	word	word
tomate	FILHO	distrator2	word	word
trabalho	LAVANCA	distrator3	word	nonword
alto	BAIXO	experim	word	word
perna	JANELA	controle	word	word
tecido	PAZELA	distrator3	word	nonword
cidade	ESTRELA	distrator2	word	word
alarme	GETO	distrator3	word	nonword
criança	VROSSO	distrator3	word	nonword

batata	NERTO	distrator3	word	nonword
homem	MULHER	experim	word	word
grosso	RELO	distrator3	word	nonword
santo	MARTO	distrator3	word	nonword
velho	PAVALHO	distrator3	word	nonword
cheio	VAZIO	experim	word	word
linha	TAMA	distrator3	word	nonword
cigarro	BORRACHA	controle	word	word
fígado	BORATA	distrator3	word	nonword
tijolo	MULHER	controle	word	word
vontade	PEIJE	distrator3	word	nonword
rótulo	JARFIM	distrator3	word	nonword
conta	FROGA	distrator3	word	nonword
multa	POEIRA	distrator2	word	word
estado	ETIQUETA	distrator2	word	word
laranja	ESCOLA	controle	word	word
barco	ILHA	distrator1	word	word
poste	CAMISA	controle	word	word
coelho	PAVALO	distrator3	word	nonword
areia	CIMENTO	distrator1	word	word
cola	GIROFA	distrator3	word	nonword
chinelo	GOIABA	distrator2	word	word
pastel	ESCADA	distrator2	word	word
bolha	QUEIJO	distrator2	word	word
barraca	LERDE	distrator3	word	nonword
quadro	CEBOLA	distrator2	word	word
gato	CACHORRO	experim	word	word
calça	CAMISA	experim	word	word
gordo	MAGRO	experim	word	word
sombra	PAPOCA	distrator3	word	nonword
jornal	MAGRO	controle	word	word
firme	GISTA	distrator3	word	nonword
azeite	PRAFUSO	distrator3	word	nonword
bolso	CRAZO	distrator3	word	nonword
música	LETRA	distrator1	word	word
fada	PIDRA	distrator3	word	nonword
caminho	FLUTA	distrator3	word	nonword
tela	BARACO	distrator3	word	nonword
creme	FALME	distrator3	word	nonword
unha	ERRADO	controle	word	word
dente	ELHO	distrator3	word	nonword
bonito	FRICO	distrator3	word	nonword
frio	CASACO	distrator1	word	word
preço	PRASENTE	distrator3	word	nonword
sapato	LEITE	controle	word	word
papel	CANETA	experim	word	word
pimenta	VAZIO	controle	word	word
recado	BILHETE	distrator1	word	word

fome	BRUXA	distrator2	word	word
bolsa	CIRCO	controle	word	word
poema	COVO	distrator3	word	nonword
botão	TEMPA	distrator3	word	nonword
porto	NAVIO	distrator1	word	word
ponta	LACO	distrator3	word	nonword
torto	RETO	distrator1	word	word
chuva	FACA	controle	word	word
casa	COMIDA	experim	word	word
cópia	MELHA	distrator3	word	nonword
pedra	FOFOCA	distrator2	word	word
carne	PEREDE	distrator3	word	nonword
espelho	CAXA	distrator3	word	nonword
ave	PEIXE	distrator1	word	word
chato	PROMA	distrator3	word	nonword
tambor	PREDE	distrator3	word	nonword
bota	FAXA	distrator3	word	nonword
garrafa	NEMERO	distrator3	word	nonword
telha	CARTA	distrator2	word	word
torre	DASTO	distrator3	word	nonword
EndData				

%%%%%%%%%%%%%% `instruções %%%%%%%%%%%%%%%

data "instruções"

Neste experimento você verá na tela palavras da Língua Portuguesa seguidas por seqüências de letras - Maiúsculas - que podem ser palavras ou não-palavras. A primeira palavra será apresentada muito rapidamente. Sua tarefa consiste em ler as DUAS SEQUÊNCIAS e responder apenas se a SEGUNDA seqüência , uma palavra da Língua portuguesa ou não.

Caso você perca alguma palavra, não de desespere, siga em frente.

Para responder ao teste use a caixa de resposta a sua frente, escolhendo os botões para a respostas SIM e NÃO.

Por favor, comunique ao assistente a sua escolha.

Haverá uma fase de treinamento, anterior ao teste propriamente dito.

Lembre-se: , necessário que você leia as DUAS seqüências.

Muito obrigada pela contribuição.

Aperte qualquer tecla para começar o treinamento...

```
enddata
```

```
procedure "instruções"
```

```
    hidecursor
```

```
    writedata "instruções"
```

```
    readbutton
```

```
    cls
```

```
    echo "Por favor aperte a tecla marcada SIM"
```

```
    rightkey:=readbutton
```

```
    wait 1000
```

```
    echo "Por favor aperte agora a tecla marcada NÃO"
```

```
    leftkey:=readbutton
```

```
    wait 1000
```

```
endproc
```

OBS: A ordem de apresentação da lista dos itens experimentais foi aleatorizada a cada sujeito.

Experimento II - Priming semântico com verbos prefixados

Exemplo de programação dos dados em EXPE6.

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%Tutor05.pro -- created by Dpx, Oct 1997%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%novo:Short.pro - modificado por LGR, dezembro/2000%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%Instruções por Luciana Garcia Ruiz      %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%UM PRIME SEMÂNTICO ENVOLVENDO UM TAREFA DE DECISÃO LEXICAL
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%COM SOA CURTO
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Conteúdo: Este arquivo apresentará uma palavra durante
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%190ms, seguida de um 60ms de intervalo e de uma palavra-
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%alvo (por mais 700ms, no máximo).
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% A tarefa consiste em desempenhar uma decisão lexical
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%acelerada no segundo item. O tempo final de resposta não
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%ultrapassa 3000ms. RTS (tempos de reação) são coletados e
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%armazenados.

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Keywords: call, procedure, endproc, waittill, begif,
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%endif, variables, goto, label
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% To run:    EXPE6 short subject_name

defvar rightkey leftkey resp_certa resp_sujeito
defvar prime_duration blank_duration target_duration rt_deadline
inter_trial_interval
defvar a
defvar b

prime_duration:=190
blank_duration:=60
target_duration:=700
rt_deadline:=3000
inter_trial_interval:=1500

call "instruções"

graphmode

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Esta é a sessão de treinamento %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

WithData "treinamento"

    wait inter_trial_interval

onpage 1 cls
```

```

onpage 1 writexy centerx centery #1
showpage 1
wait prime_duration
a:=getonset showpage 0
onpage 1 cls
onpage 1 writexy centerx centery #2
enablebutton
waittill a+blank_duration
rtonset:=getonset showpage 1
waitwhile (not buttonpressed) and (clock<rtonset+target_duration)
showpage 0
waitwhile (not buttonpressed) and (clock<rtonset+rt_deadline)

disablebutton

```

LoopData

```

%%%%%%%%%% Esta é a sessão Experimental %%%%%%%%%%

```

```

cls
writexy centerx centery - 3 line "E agora vamos ao experimento
real!"
writexy centerx maxy - 3 line "...Aperte uma tecla para
continuar..."

```

```

readbutton
WithData "dados"
cls

```

```

wait inter_trial_interval
onpage 1 cls
onpage 1 Writexy centerx centery #1

showpage 1
wait prime_duration

a:=getonset showpage 0
onpage 1 cls
onpage 1 writexy centerx centery #2
enablebutton

waittill a+blank_duration
rtonset:=getonset showpage 1

waitwhile (not ButtonPressed) and (clock<RtOnset+target_duration)

showpage 0

```

```
WaitWhile (not ButtonPressed) and (clock<RtOnset+rt_deadline)
  DisableButton
```

```
wait 100
resp_sujeito:="b"
if button==rightkey resp_sujeito:="word"
if button==leftkey resp_sujeito:="nonword"
resp_certa:=(resp_sujeito == #4)
```

```
save #0 button rt resp_sujeito resp_certa
showpage 0
```

LoopData

TextMode

%%MATERIALS%%

Data "treinamento"

pato	LAGO	Word	Related
cheiro	NERATU	Nonword	N/A
macaco	BOLA	Word	Unrelated
barco	COLTA	Nonword	N/A
banana	RORVE	Nonword	N/A
folha	RODA	Word	Unrelated
carro	FREIO	Word	Related
olho	ESBUGA	Nonword	N/A
bolo	CRIANCA	Word	Related
teclado	CANETA	Word	Unrelated
caneta	PORTA	Word	Unrelated
chave	COTUGA	Nonword	N/A
sacola	TREDUTA	Nonword	N/A
tela	PAPEL	Word	Unrelated
colcha	CAMA	Word	Related

EndData

Data "dados"

reconquistar	VENCER	Transp	word
recompor	ESCORRER	distrator1	word
defender	AFANCAR	distrator3	nonword
copiar	IMITAR	distrator2	word
agarrar	TODAR	distrator3	nonword
espurrar	PARTENCER	distrator3	nonword
apalpar	ORGANIZAR	Ctrl	word
encostar	ESTRAGAR	Ctrl	word
votar	DROER	distrator3	nonword
trocar	DANTAR	distrator3	nonword
tolerar	TEMRERAR	distrator3	nonword
redividir	PARTIR	Transp	word
importar	JERVER	distrator3	nonword
sorrir	NEGAR	Ctrl	word

cursar	GELIR	distrator3	nonword
haver	GIPAR	distrator3	nonword
rejeitar	ESTICAR	distrator1	word
secar	FOLTAR	distrator3	nonword
pecar	POLHAR	distrator3	nonword
costurar	GOLAR	distrator3	nonword
passear	AVIRAR	distrator3	nonword
abordar	GORCER	distrator3	nonword
escolher	MACHUCAR	Ctrl	word
criticar	DESEJAR	Ctrl	word
aprontar	PENDURAR	distrator1	word
complicar	POMITAR	distrator3	nonword
pressupor	DISPERSAR	distrator1	word
honrar	OSUPAR	distrator3	nonword
largar	PROCAR	distrator3	nonword
pedalar	APAGAR	Ctrl	word
guardar	ESTORRER	distrator3	nonword
digitar	ATERRISSAR	Ctrl	word
esvaziar	NASCER	distrator1	word
cruzar	GLOUVAR	distrator3	nonword
guerrear	BATALHAR	distrator2	word
dever	AFOGER	distrator3	nonword
calar	CURLIR	distrator3	nonword
cumprir	PESTAR	distrator3	nonword
faltar	ENROLHER	distrator3	nonword
desmamar	DIGERIR	distrator1	word
apurar	LOUPAR	distrator3	nonword
cercar	BULTAR	distrator3	nonword
propor	CEITAR	distrator3	nonword
baleiar	PUZAR	distrator3	nonword
requerer	DESEJAR	Pseudo	word
enfurecer	ARRATAR	distrator3	nonword
queimar	TEGER	distrator3	nonword
suspirar	ENTORDAR	distrator3	nonword
reavaliar	CALCULAR	Transp	word
provar	FEDAR	distrator3	nonword
morrer	PEVAR	distrator3	nonword
namorar	ENCERAR	distrator1	word
chatear	ABORRECER	distrator2	word
batizar	DEILAR	distrator3	nonword
apontar	MOSTRAR	distrator2	word
roncar	SAJAR	distrator3	nonword
cortar	FATJAR	distrator2	word
optar	JOFREAR	distrator3	nonword
pedalar	EXPLICAR	Ctrl	word
testar	SURFIR	distrator3	nonword
amar	AVISAR	distrator1	word
matar	CALCULAR	Ctrl	word
possuir	TERMINAR	distrator1	word
cansar	DIMINUIR	distrator1	word
destacar	INTERVIR	Ctrl	word
readaptar	ADEQUAR	Transp	word

vestir	PUNDIR	distrator3	nonword
informar	ESTORRER	distrator3	nonword
retrair	ENGANAR	Pseudo	word
comportar	ENTROLAR	distrator3	nonword
recomprar	VENDER	Transp	word
ganhar	PARAR	Ctrl	word
reaparecer	EMPURRAR	distrator1	word
exceder	PAVEGAR	distrator3	nonword
localizar	PEMAR	distrator3	nonword
reportar	CARREGAR	Pseudo	word
pescar	BLATIR	distrator3	nonword
reverter	DERRAMAR	Pseudo	word
combinar	MISTURAR	distrator2	word
arriscar	VRECISAR	distrator3	nonword
despregar	FALAR	distrator1	word
arranhar	RASPAR	distrator2	word
começar	ACABAR	distrator2	word
requebrar	ESTRAGAR	Pseudo	word
desativar	VIAJAR	distrator1	word
adquirir	CONSEGUIR	distrator2	word
compensar	INDENIZAR	distrator2	word
custar	SACER	distrator3	nonword
vestir	ESAR	distrator3	nonword
brindar	APERTAR	Ctrl	word
retornar	PAPOIAR	distrator3	nonword
jogar	TRODAR	distrator3	nonword
recobrar	PAGAR	Pseudo	word
remover	MEXER	Pseudo	word
girar	RODAR	distrator2	word
abaixar	ESPETAR	distrator1	word
reafirmar	NEGAR	Transp	word
relevar	TRAZER	Pseudo	word
comparar	ABIMAR	distrator3	nonword
chover	SOBRIR	distrator3	nonword
olhar	ENXERGAR	distrator2	word
precisar	DROMETER	distrator3	nonword
sentir	VENDER	Ctrl	word
pedir	DESTRUIR	Ctrl	word
poluir	TANSAR	distrator3	nonword
cuidar	ETRAR	distrator3	nonword
cheirar	ESCULPIR	Ctrl	word
acelerar	ANGOLIR	distrator3	nonword
escolher	RESFROAR	distrator3	nonword
marcar	ASSINALAR	distrator2	word
fumar	LECRAR	distrator3	nonword
amanhecer	ABOTOAR	distrator1	word
existir	CEMITIR	distrator3	nonword
bater	LUTAR	distrator2	word
conhecer	PRODUVIR	distrator3	nonword
aumentar	AMASSAR	distrator1	word
dançar	REBOLAR	distrator2	word
substituir	COMOTER	distrator3	nonword

esfolar	ADEQUAR	Ctrl	word
desunir	SEPARAR	distrator2	word
pretender	SONEGER	distrator3	nonword
viver	TOUVIR	distrator3	nonword
visitar	ESPROMER	distrator3	nonword
adorar	CHONRAR	distrator3	nonword
assustar	SALTAR	distrator1	word
regressar	DEDENDER	distrator3	nonword
dirigir	DEPEDIR	distrator3	nonword
chegar	ESPALHAR	distrator1	word
branquear	PERTEBER	distrator3	nonword
aprender	ESTUDAR	distrator2	word
servir	LODAR	distrator3	nonword
resolver	ADODAR	distrator3	nonword
relembrar	ESQUECER	Transp	word
sublinhar	PEGAR	distrator1	word
elaborar	AFASTER	distrator3	nonword
acertar	TRAZER	Ctrl	word
buscar	MEXER	Ctrl	word
brindar	CADAR	distrator3	nonword
sentar	DEIBAR	distrator3	nonword
mastigar	PERGUNGAR	distrator3	nonword
ponderar	ACACAR	distrator3	nonword
remediar	INTERVIR	Pseudo	word
atirar	PUDAR	distrator3	nonword
estacionar	ENXUGAR	distrator1	word
roubar	MATER	distrator3	nonword
limpar	PROEGER	distrator3	nonword
redefinir	EXPLICAR	Transp	word
realinhar	ORGANIZAR	Transp	word
entregar	LEVANTAR	Ctrl	word
seduzir	LEDIR	distrator3	nonword
reacender	APAGAR	Transp	word
insistir	DEPORAR	distrator3	nonword
lixar	PINTAR	distrator2	word
pegar	PRITAR	distrator3	nonword
esfregar	ACABAR	distrator1	word
participar	CORRIGIR	distrator1	word
gastar	POSCAR	distrator3	nonword
entrar	PISER	distrator3	nonword
aprovar	DERRAMAR	Ctrl	word
seguir	BINGAR	distrator3	nonword
pressentir	SOLTAR	distrator1	word
fofocar	CARREGAR	Ctrl	word
acalmar	CROLAR	distrator3	nonword
correr	DALVAR	distrator3	nonword
desmentir	PINTAR	distrator1	word
referir	MACHUCAR	Pseudo	word
observar	ISOTAR	distrator3	nonword
recorrer	PARAR	Pseudo	word
isolar	ILHAR	distrator2	word
pedir	BOBRAR	distrator3	nonword

arrancar	POUTAR	distrator3	nonword
elevantar	SUBIR	distrator2	word
desatar	MANEIRAR	distrator1	word
mudar	SALAR	distrator3	nonword
reajustar	APERTAR	Transp	word
recusar	MOLHAR	distrator1	word
retalhar	ESCULPIR	Pseudo	word
atingir	LIGERIR	distrator3	nonword
exportar	ATUMAR	distrator3	nonword
salvar	PADAR	distrator3	nonword
falhar	CHUDAR	distrator3	nonword
segurar	TREIVAR	distrator3	nonword
perdoar	PAXAR	distrator3	nonword
reagir	ACIONAR	Transp	word
adivinhar	DORELAR	distrator3	nonword
reconstruir	DESTRUIR	Transp	word
cobrir	ENGANAR	Ctrl	word
arrepiar	TEMAR	distrator3	nonword
lavar	PORAR	distrator3	nonword
babar	PARTIR	Ctrl	word
chutar	BREZAR	distrator3	nonword
habitar	PANCHAR	distrator3	nonword
discutir	OTUPAR	distrator3	nonword
desarmar	COLAR	distrator1	word
investir	COMORIR	distrator3	nonword
explodir	LEXAR	distrator3	nonword
duvidar	SORCEAR	distrator3	nonword
ordenar	ARRUMAR	distrator2	word
frear	BRECAR	distrator2	word
colher	PEGAR	distrator2	word
fazer	CROAR	distrator3	nonword
recair	LEVANTAR	Transp	word
revelar	CUIDAR	Pseudo	word
reabrir	FECHAR	Transp	word
chicotear	AGRONTAR	distrator3	nonword
zombar	TIGAR	distrator3	nonword
clamar	GRITAR	distrator2	word
escurecer	CLAREAR	distrator2	word
cozinhar	ASSAR	distrator2	word
supor	IMAGINAR	distrator2	word
acomodar	ALOJAR	distrator2	word
desenhar	PIMETAR	distrator3	nonword
desmaiar	ESBONDER	distrator3	nonword
ganhar	PERDER	distrator2	word
saber	CODAR	distrator3	nonword
conduzir	INVENCAR	distrator3	nonword
julgar	CONDENAR	distrator2	word
refletir	DOBRAR	Pseudo	word
causar	ROSPAR	distrator3	nonword
repousar	ATERRISSAR	Pseudo	word
desculpar	MEMECER	distrator3	nonword
bocejar	ESQUECER	Ctrl	word

impor	COMPRAR	distrator1	word
gravar	CUIDAR	Ctrl	word
arrancar	TRALAR	distrator3	nonword
trancar	VENCER	Ctrl	word
perceber	DEGONAR	distrator3	nonword
trabalhar	EMGRESTAR	distrator3	nonword
combinar	ESTOURAR	distrator1	word
adoecer	MUSPIR	distrator3	nonword
passar	SOLCAR	distrator3	nonword
sonhar	ACIONAR	Ctrl	word
corrigir	DOBRAR	Ctrl	word
melhorar	PRODESSAR	distrator3	nonword
desconfiar	FRICAR	distrator3	nonword
prender	PESCER	distrator3	nonword
acontecer	AFANCAR	distrator3	nonword
escorrer	FECHAR	Ctrl	word
defender	PAGAR	Ctrl	word
EndData			

%%Instruções%%

data "instruções"

Neste experimento você verá na tela palavras da Língua Portuguesa seguidas por seqüências de letras - Maiúsculas - que podem ser palavras ou não-palavras. A primeira palavra será apresentada muito rapidamente. Sua tarefa consiste em ler as DUAS SEQÜÊNCIAS e responder apenas se a SEGUNDA seqüência, uma palavra da Língua portuguesa ou não.

Caso você perca alguma palavra, não de desespere, siga em frente.

Para responder ao teste use a caixa de resposta a sua frente, escolhendo os botões para a respostas SIM e NÃO.

Por favor, comunique ao assistente a sua escolha.

Haverá uma fase de treinamento, anterior ao teste propriamente dito.

Lembre-se: , necessário que você leia as DUAS seqüências.

Muito obrigada pela contribuição.

Aperte qualquer tecla para começar o treinamento...

enddata

procedure "instruções"

```

hidecursor
writedata "instruções"
readbutton

```

```
cls
echo "Por favor aperte a tecla marcada SIM"
rightkey:=readbutton
wait 1000
echo "Por favor aperte agora a tecla marcada NÃO"
leftkey:=readbutton
wait 1000
endproc
```

O

BS: A ordem de apresentação da lista dos itens experimentais foi aleatorizada a cada sujeito.

ANEXO III

1	copiar	imitar	4,73
2	ferir	machucar	4,70
3	frear	brecar	4,53
4	querer	desejar	4,53
5	velar	cuidar	4,46
6	ativar	acionar	4,33
7	constituir	formar	4,26
8	mover	mexer	4,26
9	pousar	aterrissar	4,26
10	adaptar	adequar	4,20
11	adquirir	conseguir	4,20
12	tardar	demorar	4,20
13	trair	enganar	4,20
14	conquistar	vencer	4,13
15	marcar	assinalar	4,13
16	aprender	estudar	4,06
17	dividir	partir	4,00
18	portar	carregar	4,00
19	talhar	esculpir	4,00
20	tocar	encostar	4,00
21	contar	enumerar	3,93
22	acomodar	alojar	3,80
23	cortar	fatiar	3,80
24	fletir	dobrar	3,80
25	lutar	brigar	3,80
26	ligar	unir	3,73
27	verter	derramar	3,73
28	ajustar	apertar	3,66
29	colher	pegar	3,66
30	armar	montar	3,60
31	mediar	intervir	3,60
32	bolar	imaginar	3,46
33	compensar	indenizar	3,46
34	definir	explicar	3,46
35	sentir	perceber	3,46
36	cobrir	tampar	3,40
37	condicionar	habituar	3,40

38	mover	andar	3,40
39	alinhar	organizar	3,33
40	formular	articular	3,33
41	agir	atuar	3,26
42	quebrar	estragar	3,26
43	bater	lutar	3,13
44	clamar	gritar	3,13
45	avaliar	calcular	3,06
46	ordenar	arranjar	2,86
47	lembrar	esquecer	2,60
48	afirmar	teimar	2,36
49	cobrar	receber	2,26
50	agrupar	separar	2,00
51	levar	trazer	2,00
52	abrir	fechar	1,93
53	acender	apagar	1,93
54	buscar	levar	1,93
55	começar	acabar	1,93
56	comprar	vender	1,93
57	negar	afirmar	1,93
58	cair	levantar	1,86
59	construir	destruir	1,86
60	correr	parar	1,86
61	criar	destruir	1,86
62	iniciar	terminar	1,86
63	parar	continuar	1,80
64	voltar	permitir	1,73
65	alçar	pousar	1,66
66	cobrar	pagar	1,66
67	conhecer	ignorar	1,66
68	tirar	colocar	1,66
69	negar	confirmar	1,60
70	quebrar	concertar	1,60
71	baixar	subir	1,46
72	povoar	destruir	1,33

OBS: Os valores da quarta coluna são as médias aritméticas de cada par sobre as respostas de cada aluno.

Para que o leitor compreenda melhor o pré-teste, apresentamos abaixo as instruções fornecidas aos alunos para a sua realização.

“A lista de pares de verbos abaixo foi construída com base na existência de algum tipo de relacionamento (semântico ou meramente associativo) entre eles. Por exemplo, entre *roubar* e *tomar* há – intuitivamente – uma relação que se assenta na idéia de que quando alguém rouba ou toma um objeto de outra pessoa, então a posse desse objeto muda da segunda para a primeira pessoa. Da mesma forma, entre *calar* e *tagarelar* há uma relação que gira em torno da idéia de “falar”. Estamos interessados em saber como nossa intuição de falantes do português permite julgar a ‘força’ ou a ‘proximidade’ dessa relação. Às vezes é possível dizer que a relação entre dois verbos é mais ‘fraca’ que outra. Para o par *calar* e *tagarelar*, acima, pode-se dizer que a relação é fraca ou distante, pelo menos se comparada à relação entre *calar* e *emudecer*. De modo a traduzir esses julgamentos em números, vamos convencionar que se um par de verbos tem uma relação ‘forte’ de significado, então daremos a nota 5 para ela; se a relação for fraca, a nota será 1; e, finalmente, se a intuição for de que o relacionamento é intermediário, a nota será 3.”