

Formas flexionais regulares de número em português brasileiro: acesso e representação lexical

Jefferson Alves da Rocha¹

José Ferrari-Neto²

Universidade Federal da Paraíba (UFPB)

Resumo: Este trabalho investiga o processamento de vocábulos flexionados em número em Português Brasileiro (PB), sugerindo um modelo de representação para eles. Observou-se como as formas flexionadas em número são armazenadas e recuperadas do léxico mental. Estudos recentes têm apontado que a frequência da base influi no processamento de formas flexionadas em gênero e em número (DOMINGUEZ; CUETOS; SEGUI, 1999). No tocante à flexão de gênero em PB, um efeito de frequência dominante foi registrado (CORRÊA; ALMEIDA; PORTO, 2004). A fim de prover evidências sobre se a frequência dominante afeta as formas flexionadas em número, elaborou-se um experimento, o qual consistiu de um teste de leitura de palavras composto por vocábulos com flexão regular de número. Procurou-se observar se a forma singular era acessada por inteiro, tendo em vista efeitos de frequência dominante, ou se a forma no plural era acessada por decomposição, tendo em vista efeitos de frequência não dominante, salvo os casos de alta frequência de formas no plural, nos quais estas formas também seriam acessadas por inteiro, em um modelo de dupla rota. Os resultados atestaram que as palavras com frequência dominante são processadas de forma mais rápida que as formas com frequência não dominante.

Palavras-chave: Flexão de Número. Léxico Mental. Português Brasileiro.

Introdução

Léxico é o componente da gramática que contém todas as informações - fonológicas, morfológicas, sintáticas e semânticas ó que os falantes sabem sobre palavras simples ou morfemas. Do ponto de vista psicolinguístico, é denominado *léxico mental* e corresponde a um repositório de conhecimentos declarativos sobre as palavras de uma língua. Estes conhecimentos podem ser de natureza fonológica, morfológica, sintática e semântica, podendo também haver conhecimentos pragmáticos e estilísticos sobre os itens lexicais. Do ponto de vista linguístico, o léxico é uma lista de elementos que são usados na formulação de sentenças. Consiste num conjunto de informações acerca dos itens lexicais que são acessados e manipulados pela gramática. Na teoria gerativa, tradicionalmente, o léxico é o repositório de

¹ Mestre em Linguística pelo Programa de Pós-Graduação em Linguística da Universidade Federal da Paraíba. E-mail: jefferson.rocha16@gmail.com.

² Doutor em Estudos da Linguagem pelo Departamento de Letras Clássicas e Vernáculas ó Centro de Ciências Humanas, Letras e Artes da Universidade Federal da Paraíba. E-mail: joseferrarin@ibest.com.br.

exceções, isto é, das propriedades idiossincráticas dos itens lexicais. O léxico é tudo aquilo que não pode ser gerado por regras.

A representação dos itens lexicais no léxico mental tem sido um tópico recorrente, devido a sua grande importância para a pesquisa psicolinguística. De acordo com Emmorey e Fromkin (1988), o estudo da natureza das representações armazenadas no léxico mental apresenta grande relevância porque determina, em boa medida, a natureza dos mecanismos de acesso. Do ponto de vista psicolinguístico, vários modelos de processamento lexical têm sido sugeridos, tratando de questões relacionadas ao acesso e à representação, além de mostrar como determinadas formas linguísticas estariam armazenadas e seriam recuperadas do léxico mental. Nesses modelos, é nítida a importância dada à questão das formas flexionadas: como essas formas seriam representadas no léxico mental? Como se daria o acesso e a representação de tais formas? Percebe-se que a problemática relacionada ao léxico mental gira em torno de se saber como as palavras são representadas, por um lado, e, por outro lado, saber como essas representações são acessadas. Para um e outro caso, foram sugeridas algumas hipóteses sobre a representação e o acesso das palavras.

Os modelos de processamento tratam de questões do acesso ao léxico mental, além de mostrar como determinadas formas linguísticas estariam armazenadas e seriam recuperadas do léxico mental. Assim, procura-se expor a configuração dos mecanismos no âmbito do léxico mental e como se dá o tratamento das formas linguísticas relacionadas aos modelos de processamento, como já foi apontado por Emmorey e Fromkin (1988). Ainda em relação aos modelos, pode-se observar que há três hipóteses diferentes em relação a cada um deles: *Full Parsing Hypothesis*, *Full Listing Hypothesis* e *Hybrid Model Hypothesis*. Tais hipóteses resultam em duas formas de conceber os itens linguísticos no momento do acesso, através de *morpheme-based access* ou *full-form access*.

Em relação aos modelos de acesso e representação lexical, propôs-se que as palavras seriam representadas em sua forma completa, ao que se denominou *Full Listing Hypothesis*, ou que seriam decompostas em seus morfemas constituintes, ao que se denominou *Full Parsing Hypothesis*. Uma via alternativa sugeriu que ambas as representações estariam disponíveis, surgindo assim a *Hybrid Model Hypothesis*. Por sua vez, no que se refere ao modo como são acessadas, as propostas caminharam na direção de haver duas formas de reconhecimento: por palavra completa (*full form access*) e por morfemas (*morpheme-based access*). Pode-se postular uma relação entre esses modelos teóricos de representação e acesso,

uma vez que o modo como as palavras são representadas determina, em grande parte, a maneira como serão acessadas e recuperadas do léxico mental. Assim, modelos *Full Listing* são mais compatíveis com acesso *full form*, ao passo que modelos *Full Parsing* são mais compatíveis com acesso *morpheme-based*. Modelos híbridos são compatíveis com as duas vias de acesso.

De acordo com a *Full Parsing Hypothesis*, portanto, radicais e morfemas presos são representados independentemente no léxico mental, sendo o morfema a unidade básica de armazenamento. Os modelos que assumem esta hipótese são os de Taft & Forster (1975), conhecido como *Modelo Serial de Busca*, e o de Taft (1994), chamado de *Modelo de Ativação Interativa*. Por sua vez, a *Full Listing Hypothesis* afirma que existem entradas lexicais para todas as palavras independentemente de sua estrutura morfológica. Assim, a unidade básica de armazenamento, segundo esta hipótese, seria a palavra. O modelo que assume esta visão é o de Butterworth (1983), que se classificaria entre os de acesso *full form*, ou seja, não compartilha a assunção da decomposição morfológica obrigatória. O reconhecimento por palavra completa (*full form access*) requer que o conjunto de itens que forma o *input* tenha uma representação estocada no léxico. Esta rota está então disponível apenas para uma palavra que esteja estocada como uma entidade simples, como, por exemplo, as palavras monomorfêmicas. Por outro lado, um típico candidato para um acesso baseado em morfemas seria um vocábulo polimorfêmico, que não é muito frequente, mas contém morfemas constituintes que aparecem frequentemente. Formas flexionadas, como as de plural em português ou inglês, seriam os melhores exemplos. Em um estudo feito com falantes de árabe argelino, Mimouni, Kehaya & Jarema (1998), apontaram evidências para a existência de uma dupla rota de acesso a palavras árabes flexionadas em número.

Há, ainda, modelos de processamento morfológico que têm proposto que ambas as formas de representação, *Full Parsing* e *Full Listing*, e ambas as formas de acesso, *morpheme-based* e *full-form*, estão disponíveis para o reconhecimento da palavra (LAINE; VAINIO; HYÖNÄ, 1999. COLÉ; SEGUI; TAFT, 1997). Dessa forma, a *Hybrid Model Hypothesis* propõe que morfemas, sejam livres ou presos, seriam as unidades básicas de representação lexical. Esta hipótese foi proposta para dar conta dos efeitos de frequência no reconhecimento de palavras complexas, uma vez que palavras morfológicamente complexas com alta frequência são reconhecidas e/ou recuperadas como uma unidade (com maior rapidez), ao passo que palavras complexas de baixa frequência sofreriam *parsing* morfológico

(sendo acessadas com menor rapidez). Os modelos que adotam o *Hybrid Model* são o Modelo de Endereçamento Morfológico Ampliado (*Augmented Addressed Morphology Model* - AAM) de Caramazza, Laudanna & Romani (1988) e o Modelo *MRM*, de Schreuder & Baayen (1995).

Alguns estudos experimentais objetivaram, do ponto de vista dos estudos sobre o léxico mental, prover evidências acerca dos diversos modelos de acesso e representação, conforme visto acima. Tais estudos levaram em conta, em maior ou menor grau, possíveis fatores intervenientes, tais como efeitos de frequência (número de vezes que uma dada forma lexical aparece em um *corpus* representativo do léxico de uma língua), frequência dominante e não dominante (usada em situações em que há um par de formas flexionadas, sobre as quais se deseja saber qual delas é a mais frequente), frequência de base e de superfície (relativa à frequência de uma dada forma em relação às suas respectivas flexões) e familiaridade. Esses fatores são tomados como relevantes na pesquisa sobre o léxico mental na medida em que podem evidenciar a pertinência dos modelos de acesso e representação.

Estudos experimentais

O estudo realizado no espanhol por Dominguez, Cuetos e Segui (1999) buscou investigar questões de processamento de itens lexicais flexionados em número e em gênero. Os autores realizaram três experimentos, o primeiro e o segundo sobre o gênero e o terceiro sobre o número.

Em relação ao primeiro experimento, utilizaram-se palavras masculinas e femininas, com mudança no morfema final de *o* para *a*. Utilizou-se um teste de decisão lexical. Neste experimento, a frequência³ de superfície das formas masculinas e femininas foi manipulada. Para os itens *masculino dominante*, a frequência de superfície da forma masculina foi maior que aquelas da forma feminina. Já para os itens *feminino dominante*, a frequência de superfície da forma feminina foi maior que aquelas da forma masculina. Esperava-se que houvesse entradas e representações lexicais para as duas formas (masculino e feminino). Desse modo, a raiz de cada uma das formas permaneceria separada no léxico. Com isso, o teste de decisão lexical apontou que os sujeitos decidiram mais rapidamente pelas formas de

³ No texto de Dominguez, Cuetos e Segui (1999) não fica evidente o procedimento para verificação da frequência das palavras experimentais.

superfície que são mais frequentes em relação aos casos citados anteriormente. Nos resultados do primeiro experimento referente ao gênero, foram encontrados efeitos significativos no fator *frequência relativa* e no fator *gênero*. Também se evidenciou efeito significativo de interação entre frequência e gênero. A diferença entre palavras masculinas e femininas dos itens de *masculino dominante* foi significativa. Já a diferença entre palavras masculinas e femininas dos itens de *feminino dominante* foi significativa apenas por sujeito e não por item. Concluiu-se que os tempos de decisão lexical foram mais rápidos nos membros dos pares mais frequentes independentemente do gênero.

Utilizando-se do mesmo procedimento experimental acima, os autores replicaram o experimento a partir dos resultados observados. A diferença nessa replicação de experimento consistiu no conjunto dos itens selecionados. Neste segundo experimento, os autores inseriram a frequência cumulativa para verificação dos efeitos dessa variável nos vocábulos masculinos e femininos. Compararam-se, assim, palavras masculinas com a mesma frequência de superfície, diferindo da frequência da forma feminina correspondente. Os resultados obtidos neste experimento mostraram a ausência de efeito significativo da frequência cumulativa. Quando uma dada palavra masculina apresentou a mesma frequência de superfície e diferiu da frequência da forma feminina correspondente, os tempos de respostas foram semelhantes. Os resultados encontrados foram na mesma direção dos pretendidos na hipótese, aferindo que a informação sobre gênero é incluída na entrada lexical de cada palavra, afetando o tempo do processamento.

Dominguez, Cuetos e Segui (1999) mencionaram o estudo realizado por Sereno e Jongman (1997). Nesse estudo, verificou-se que a frequência das formas no singular e no plural influenciava o tempo de resposta. Uma das questões verificadas foi se a frequência afetava os vocábulos analisados, isto é, as palavras com *frequência dominante* deveriam ter tempos médios de resposta mais rápidos, e as palavras com *frequência não dominante* deveriam apresentar tempos médios de repostas mais lentos. Os resultados do estudo de Sereno e Jongman (1997) levaram Dominguez *et al.* (1999) a postular, neste terceiro experimento, a representação individual da flexão regular de número. Verificaram que a frequência cumulativa contribuiu pouco para o tempo de resposta das formas flexionadas em número. Os resultados encontrados por Sereno e Jongman (1997) diferiram daqueles encontrados por Taft (1979) em que a frequência cumulativa influenciou mais no processamento, comparada com a forma de frequência de superfície. Os resultados

encontrados por Taft (1979) foram similares aos encontrados mais recentemente por Baayen, Dukstra e Schreuder (1997) no holandês. De modo geral, os resultados encontrados por Dominguez *et al.* (1999) direcionaram ao entendimento de que as formas flexionadas no plural são processadas com a entrada principal através da forma correspondente no singular. Por outro lado, a frequência da forma no singular aparece como fator decisivo no tempo de resposta.

O estudo realizado por Corrêa, Almeida e Porto (2004) critica, de certa forma, os experimentos feitos por Dominguez, Cuetos e Segui (1999), tendo em vista que utilizaram a frequência dos pares das palavras. Corrêa, Almeida e Porto (2004) argumentaram que a frequência de cada palavra, isolada da frequência do par, também poderia produzir efeitos que seriam levados em consideração no processamento lexical. Desenvolveram quatro experimentos. O trabalho desenvolvido por eles trata das formas com flexão em gênero. O primeiro experimento analisou se a frequência de dominância determina o reconhecimento de substantivos e adjetivos com flexão em gênero. Observou-se, para elaboração do experimento, que a frequência dominante é reconhecida mais rapidamente, independente do gênero do vocábulo. Sugerem, ainda, que a forma de superfície flexionada de dada palavra é reconhecida mais facilmente que os morfemas da mesma forma. Esse efeito favoreceu a hipótese de que as palavras são representadas pela forma completa, como sugeriram os dados obtidos no espanhol. Em relação ao experimento propriamente dito, Corrêa *et al.* (2004) utilizaram as mesmas variáveis do estudo realizado no espanhol por Dominguez *et al.* (1999). As variáveis foram as seguintes: a frequência de dominância nos pares de palavras com flexão de gênero (nos pares de palavras com masculino dominante e nos pares de palavras com feminino dominante); o gênero (masculino e feminino) e a categoria gramatical (substantivos e adjetivos). Os resultados não atestaram a hipótese da frequência dominante. Os efeitos principais de *categoria gramatical*, *interação entre gênero e categoria gramatical*, *interação entre frequência dominante e categoria gramatical* indicaram qual a categoria gramatical das palavras (substantivos ou adjetivos), além de indicar qual o tipo do processo de flexão envolvido (lexical ou sintático) afeta o acesso e a representação da flexão das referidas palavras. O efeito principal de categoria gramatical alcançou tempo de reconhecimento mais curto nos adjetivos do que nos substantivos. Os resultados foram favoráveis à ideia de que os substantivos femininos facilitaram o efeito significativo de frequência dominante, devido ao fato de ser um processo flexional lexical que se relaciona ao significado.

A fim de prover evidências empíricas que caracterizem e expliquem o comportamento de itens lexicais flexionados em número no léxico mental, da mesma forma que Corrêa *et al.* (2004) procederam para itens flexionados em gênero, elaborou-se o *design* experimental aqui conduzido. Inicialmente, formulou-se a hipótese de que a frequência dominante influencia o acesso às representações das palavras nominais flexionadas em número, evidenciando assim um modelo de acesso e representação em dupla rota. Desta forma, procurou-se analisar se a frequência de dadas formas lexicais influencia o tempo de processamento, isto é, se as formas mais frequentes eram acessadas por inteiro e se aquelas menos frequentes eram decompostas.

Experimento

Método e Participantes

Participaram deste experimento 80 (oitenta) sujeitos adultos, falantes nativos do PB. Os participantes foram alunos de cursos de graduação e da pós-graduação da UFPB (Universidade Federal da Paraíba). Esses participantes responderam o experimento de acordo com a categoria gramatical, isto é, 40 (quarenta) participantes realizaram o experimento com estímulos descritos como substantivos e 40 (quarenta) participantes realizaram o experimento com estímulos descritos como adjetivos.

Estímulos

As palavras utilizadas neste experimento apresentaram flexão regular em número, e o grupo de substantivos foi separado do grupo de adjetivos. A mudança flexional presente nos vocábulos consistiu apenas no acréscimo do morfema "s", indicativo de plural, mantendo, com isso, as diferenças entre forma singular e forma plural. A seguir, mostram-se exemplos do primeiro grupo (1) e do segundo grupo (2) de estímulos.

(1) *amigo* - *amigos* (substantivos).

(2) *grande* - *grandes* (adjetivos).

Os estímulos utilizados foram distribuídos por 4 (quatro) listas experimentais. Duas listas continham 24 (vinte e quatro) substantivos em cada uma delas e duas listas continham 24 (vinte e quatro) adjetivos também em cada uma delas. Em cada lista experimental estavam presentes 4 (quatro) condições experimentais, mais especificamente, 6 (seis) palavras por condição. Nas listas com substantivos, tivemos as seguintes condições: SD (Singular Dominante), SND (Singular Não Dominante), PD (Plural Dominante), PND (Plural Não dominante). Nas listas com adjetivos tivemos as mesmas condições experimentais. A seguir, podem-se visualizar, no quadro 1, as referidas condições, seguidas de exemplos dos estímulos experimentais.

Quadro 1: Condições experimentais e exemplos de vocábulos do experimento.

CATEGORIA GRAMATICAL	CONDIÇÕES EXPERIMENTAIS		EXEMPLOS DOS VOCÁBULOS UTILIZADOS NO EXPERIMENTO 1
substantivos	SD	Singular Dominante ⁴	presidente, filho, ministro, santa
	SND	Singular Não Dominante	aluna, cadela, pato, enfermeiro
	PD	Plural Dominante	crianças, santos, amigos, alunos
	PND	Plural Não Dominante	madrinhas, regentes, cadelas, gatas
adjetivos	SD	Singular Dominante	grande, novo, velha, pequeno
	SND	Singular Não Dominante	estudiosa, malvada, cheiroso, gostoso
	PD	Plural Dominante	belas, ricos, grandes, pobres
	PND	Plural Não Dominante	divertidos, educadas, malvadas, chatas

Cada estímulo foi apresentado apenas uma vez por lista, não se repetiram palavras com radicais de natureza semelhante, assim, um dado participante não foi exposto a um mesmo estímulo experimental. Além dos 24 (vinte e quatro) estímulos experimentais presentes em cada lista, também foram acrescentadas 50 (cinquenta) palavras distratoras. Todas as palavras foram randomizadas em cada lista. A seguir, exemplificamos, algumas das palavras distratoras utilizadas no experimento.

(3) *chefe, anjos* (distratoras - substantivos).

⁴ O que diferencia a forma dominante da não dominante é que as primeiras são mais frequentes que as segundas, ou seja, dado um par de formas como *aluna-alunas*, tem-se que *aluna* é singular dominante, pois é mais frequente que a sua contraparte plural *alunas*.

(4) *escuro, secos* (distratoras - adjetivos).

Frequência do experimento

Para adequar determinados vocábulos na composição do experimento, realizou-se uma busca pela frequência deles, com o intuito de selecionar quais vocábulos seriam utilizados e inseridos nas listas experimentais. Dessa forma, observamos a frequência dos vocábulos no *Corpus do Português*⁵. Realizamos uma busca pela frequência de 100 (cem) pares de substantivos (forma singular e forma plural) e 100 (cem) pares de adjetivos (forma singular e forma plural). Dentre essas palavras, dividimos aquelas que apresentaram frequência dominante e aquelas que apresentaram frequência não dominante. A partir da análise da frequência dos 200 (duzentos) pares de palavras (substantivos e adjetivos) no referido *corpus*, selecionaram-se as palavras mais frequentes e as palavras menos frequentes. Assim, realizou-se a distribuição dos substantivos na primeira e na segunda lista experimentais e também dos adjetivos na terceira e na quarta lista experimentais. Formaram-se, assim, as referidas listas experimentais, dividindo os vocábulos por categoria gramatical (substantivo e adjetivo), por número (singular e plural) e por dominância (dominante e não dominante).⁶

Teste de Familiaridade 1 e 2

Realizou-se, também um teste de familiaridade através do *Google Forms*⁷, elaborado para verificar a frequência de dominância das palavras. Os cem (100) pares de palavras (substantivos e adjetivos) utilizados para verificação da frequência no *Corpus do Português*, como foi citado na seção anterior, foram inseridos nos testes de familiaridade 1 e 2.

O teste de familiaridade 1, composto por substantivos, foi respondido por 107 (cento e sete) participantes. Neste teste, foi perguntado o seguinte: "*Qual das palavras dos pares*

⁵ O *Corpus do Português* é um banco de dados de 45.000.000 (quarenta e cinco milhões) de palavras. Ao se digitar uma dada palavra no *corpus*, ocorre a busca pela frequência dessa palavra, por exemplo, aparecendo, em seguida, a quantidade relacionada à frequência.

⁶ As palavras distratoras também foram submetidas à análise da frequência no *corpus*, pois não podiam constar, no experimento, ocorrências de frequência similares as das palavras experimentais. Desse modo, as palavras distratoras foram selecionadas observando aquelas que não apresentaram frequência dominante nem frequência não dominante em relação às palavras experimentais. Tais palavras apresentaram frequência mediana em relação às dominantes e às não dominantes.

⁷ Os testes de familiaridade elaborados no *Google Forms* foram gerados e enviados através de um *link* para os participantes.

abaixo é mais frequente para você?", e, logo abaixo, constavam os 100 (cem) pares de palavras. Ao realizar o teste, o participante teria de optar pela forma no singular ou pela forma no plural. Deve-se ressaltar que os pares de palavras dos testes foram exibidos para os participantes de forma randomizada.

Em relação aos resultados, os sujeitos, que responderam ao teste 1, optaram por 72 (setenta e duas) das formas no singular como mais familiar e por 28 (vinte e oito) das formas no plural como mais familiar. Configurou-se, com isso, mais de 50% (cinquenta por cento) das respostas nas 72 formas citadas anteriormente em cada um dos pares de palavras apresentados no teste. Constatou-se, assim, a forma no singular como dominante em relação à forma no plural como não dominante.

O teste de familiaridade 2 seguiu o mesmo modelo do teste de familiaridade 1. A diferença presente nesse segundo teste se referiu à categoria gramatical, já que se utilizaram adjetivos. Esse teste foi respondido por 97 (noventa e sete) participantes. Faz-se importante salientar que algumas pessoas participaram dos dois testes. O procedimento de execução e de exibição do teste de familiaridade 2 foi idêntico ao do teste de familiaridade 1. Em relação aos resultados, os sujeitos que responderam ao teste 2 optaram por 100 (cem) das formas no singular como mais familiar, isto é, todas as formas de adjetivos no singular, neste teste, foram escolhidas como mais familiar em relação à forma no plural. Configurou-se, com isso, mais de 50% (cinquenta por cento) das respostas em cada um dos 100 (cem) pares de palavras apresentados no teste. Constatou-se a forma no singular como dominante em relação à forma no plural como não dominante, bem como aconteceu com o teste de familiaridade 1, envolvendo os substantivos.

Faz-se necessário salientar que, tanto substantivos quanto adjetivos nas suas respectivas formas no singular, motivaram a familiaridade do indivíduo. Os resultados apontaram ainda que os adjetivos no singular (100/100) facilitaram ainda mais a familiaridade em relação aos substantivos no singular (72/100).

Variáveis

As variáveis independentes foram: número (singular e plural), categoria gramatical (substantivos e adjetivos) e frequência dominante (singular dominante e plural dominante). A variável dependente foi o tempo de leitura de palavras.

Procedimento

O experimento foi elaborado no programa *Paradigm*. Os participantes do experimento responderam um teste de leitura de palavras, todos foram testados individualmente na sala do LAPROL (Laboratório de Processamento Linguístico) na UFPB (Universidade Federal da Paraíba). As instruções foram explicadas detalhadamente a cada participante antes da realização do experimento proposto. Antes da execução do experimento propriamente dito, o participante tinha contato novamente com as instruções experimentais, já que elas foram expostas na tela inicial do teste no computador.

Os participantes, assim, leram as instruções na tela do computador antes de iniciar a tarefa experimental. As instruções pediam que o participante lesse o mais rápido possível as palavras na tela do computador e, em seguida, clicasse com o botão esquerdo do *mouse*, imediatamente, aparecia na tela do computador uma fixação de cinco asteriscos por 500 (quinhentos) milésimos de segundos e, em seguida, a palavra seguinte já era mostrada, logo após o participante clicar. As palavras foram mostradas uma por vez na tela do computador. Esse processo se repetiu 24 (vinte e quatro) vezes com o mesmo participante, já que tivemos esse mesmo número de estímulos, além disso, também se fez o citado processo 50 (cinquenta) vezes com as palavras distratoras, ou seja, o participante foi exposto a 74 (setenta e quatro) palavras. Os vocábulos experimentais foram mostrados em fonte *arial*, tamanho 12 (doze), cor preta, em uma tela de fundo branco.

Resultados e discussões

Em relação aos resultados experimentais, foram realizadas duas análises. A primeira dessas análises foi feita não separando a categoria gramatical dos substantivos e a dos adjetivos, na mesma linha da análise realizada por Corrêa *et al.* (2004) no estudo sobre o gênero. A segunda análise consistiu em analisar de forma separada os substantivos e os adjetivos. Os resultados são os que vêm a seguir. Uma vez que os dados apresentaram distribuição normal, realizou-se uma análise de variância (ANOVA), bem como testes de comparação 2 x 2 (teste-t).

Substantivos e adjetivos

O gráfico 1 consiste no tempo médio de resposta de cada uma das condições experimentais. Aqui, nessas condições experimentais, apresentamos apenas as variáveis analisadas (gênero e dominância), já que substantivos e adjetivos foram observados de modo integrado.

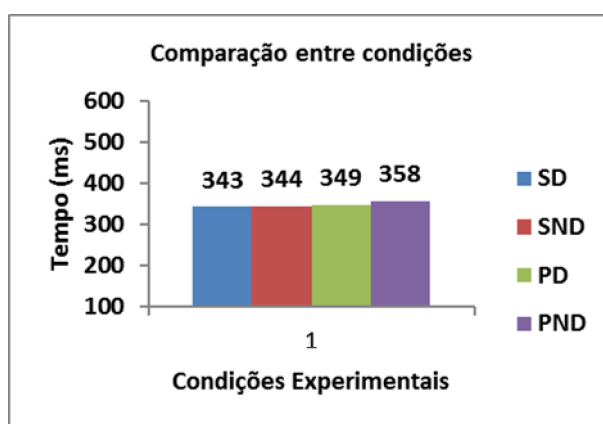


Gráfico 1: Teste-t (substantivos e adjetivos).

O gráfico 1 apresenta os tempos médios de resposta à tarefa experimental, observando substantivos e adjetivos simultaneamente. Os resultados foram submetidos à análise de variância (ANOVA, *design* fatorial 2(número) x 2(dominância)). Os resultados mostraram um efeito significativo de *número* ($F(1,479) = 6,44$ $p < 0,01$). Quanto às médias obtidas nas palavras singulares, comparadas aos plurais, verificou-se que aquelas foram lidas mais rapidamente que estas, conforme ilustra o gráfico 2:

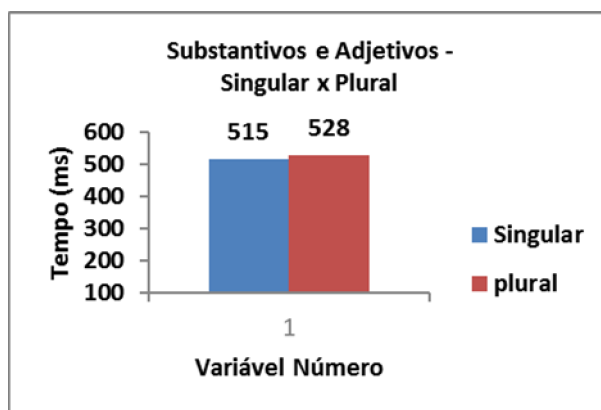


Gráfico 2: Número dos substantivos e adjetivos.

O efeito significativo de *dominância* não foi registrado ($F(1,479) = 1,36$ $p < 0,24$), sugerindo que a frequência dominante não afeta os tempos médios de leitura, de acordo com o ilustrado no gráfico a seguir:

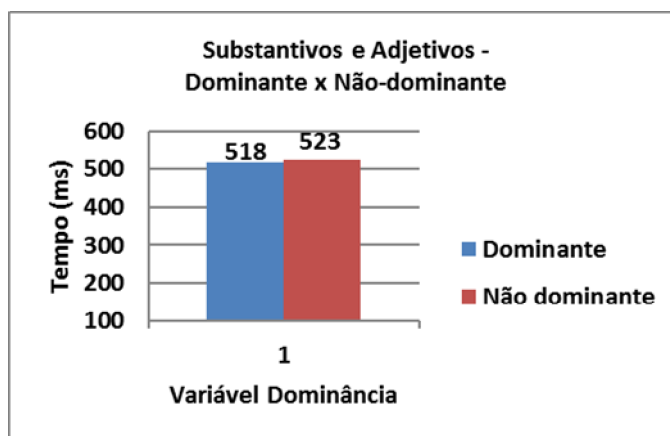


Gráfico 3: Dominância dos substantivos e adjetivos.

Um teste comparativo de médias (teste-t), com correção de Bonferroni, foi aplicado para a comparação das condições entre si, tomando-as duas a duas. Os resultados são mostrados no quadro 2:

Quadro 2: Comparações entre médias

PAIRWISE COMPARISONS (comparações duas a duas)	
[Singular Dominante] vs [Singular Não Dominante]	$t(479)=0,08$ $p < 0,9343$
[Singular Dominante] vs [Plural Dominante]	$t(479)=0,94$ $p < 0,3485$
[Singular Dominante] vs [Plural_Não Dominante]	$t(479)=1,94$ $p < 0,0527$
[Singular_Não Dominante] vs [Plural_Dominante]	$t(479)=2,31$ $p < 0,0213$
[Singular_Não Dominante] vs [Plural_Não Dominante]	$t(479)=4,17$ $p < 0,0001$
[Plural_Dominante] vs [Plural_Não Dominante]	$t(479)=2,97$ $p < 0,0031$

Os gráficos expostos anteriormente se referem aos tempos de leitura de palavras na variável *número* e na variável *dominância*, observados nos gráficos 2 e 3, respectivamente. Essas variáveis se referem ao tempo de leitura de palavras experimentais caracterizadas como substantivos e adjetivos de modo integrado. Assim, observou-se o que levou mais ou menos tempo para ser processado, isto é, ao se verificar o gráfico 2, percebe-se que as palavras no singular foram processadas mais rapidamente, obtendo, com isso, tempo de leitura de palavras menor (515ms) ao se comparar com o tempo de leitura das palavras no plural (528ms), que

levaram mais tempo para serem processadas⁸. Em relação à variável *dominância*, isolaram-se os tempos de leitura de palavras das formas dominantes e das formas não dominantes. O tempo de leitura de palavra das formas dominantes (518ms) foi menor que o tempo de leitura de palavras das formas não dominantes (523ms), nesse caso, também se obteve resultado significativo. Em relação à interação das variáveis *número* e *dominância*, não se obteve resultado significativo (Número*Dominância $F(1,479) = 1,56$ $p < 0,21$).

Em seguida, aborda-se a análise voltada especificamente para os substantivos e para os adjetivos, separadamente.

Substantivos

A seguir será mostrado o gráfico 4, que consiste no tempo médio de resposta de cada uma das condições utilizadas no experimento proposto e analisado em questão.

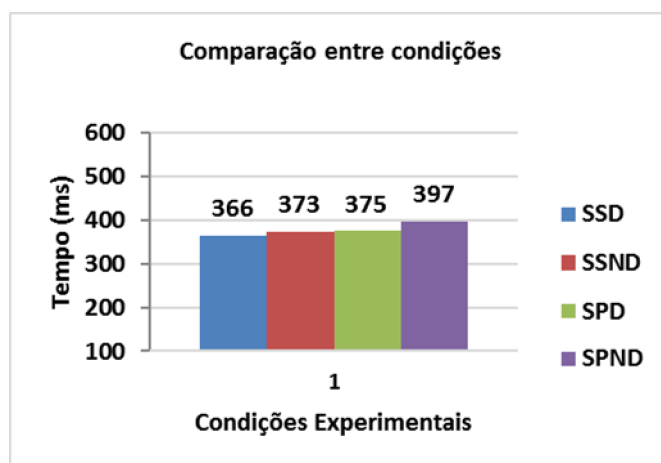


Gráfico 4: Teste-t (substantivos).

O gráfico 4 apresenta os tempos médios de resposta à tarefa experimental. Os resultados também foram submetidos à análise de variância ANOVA, *design* fatorial 2(número) x 2(dominância), a fim de se comprovarem as hipóteses aqui assumidas. Os resultados mostraram um efeito significativo de *número* ($F(1,239) = 4,63$ $p < 0,03$), indicando que a flexão da palavra afetou os tempos médios de leitura. Observando-se as médias obtidas

⁸ Uma possível observação para essa diferença observada seria a de afirmar que ela decorre do aumento do tamanho linear da palavra plural em relação à sua respectiva forma singular. No entanto, o acréscimo de uma letra à forma singular nem sempre implica em maior custo de processamento, pois fatores como a frequência e a estrutura morfológica subjacente parecem ser mais decisivos nesse caso, conforme mostram os estudos de Dominguez, Cueto e Seguí (1999) e os de Corrêa, Almeida e Porto (2004).

nas palavras singulares, comparadas às plurais, verificou-se que aquelas foram lidas mais rapidamente que estas, conforme ilustra o gráfico a seguir:

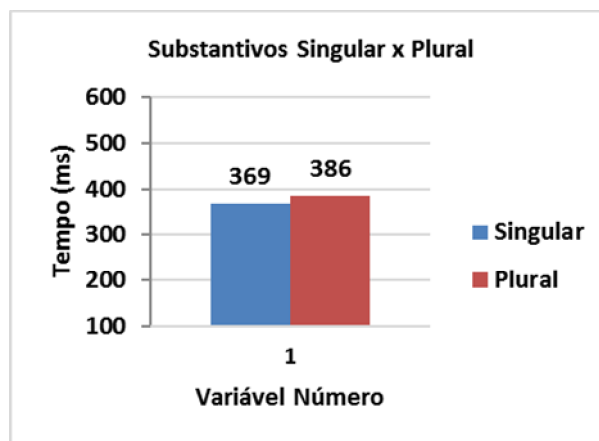


Gráfico 5: Número dos substantivos.

Um efeito significativo de *dominância* também foi registrado ($F(1,239) = 3,98$ $p < 0,04$), sugerindo que a frequência dominante também afeta os tempos médios de leitura. Formas flexionadas dominantes são lidas mais rapidamente do que as suas contrapartes não-dominantes, de acordo com o ilustrado no gráfico a seguir:

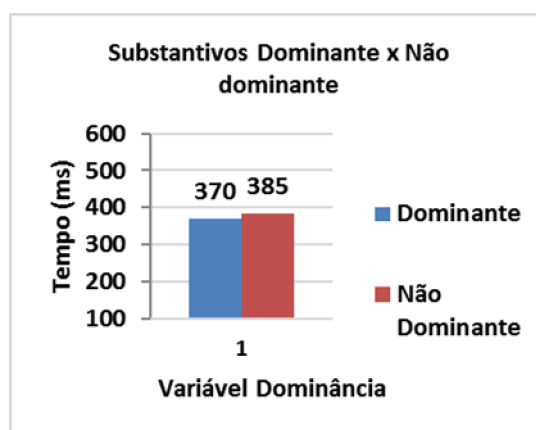


Gráfico 6: Dominância dos substantivos.

Um teste comparativo de médias (teste-t), com correção de Bonferroni, também foi aplicado para a comparação das condições entre si, tomando-as duas a duas. Os resultados são mostrados no quadro a seguir:

Quadro 3: Comparações entre médias

PAIRWISE COMPARISONS (comparações duas a duas)	
[Singular Dominante] vs [Singular Não Dominante]	$t(239)=0,58$ $p<0,5630$
[Singular Dominante] vs [Plural Dominante]	$t(239)=0,75$ $p<0,4518$
[Singular Dominante] vs [Plural Não Dominante]	$t(239)=2,15$ $p<0,0327$
[Singular Não Dominante] vs [Plural Dominante]	$t(239)=0,55$ $p<0,5860$
[Singular Não Dominante] vs [Plural Não Dominante]	$t(239)=3,76$ $p<0,0002$
[Plural Dominante] vs [Plural Não Dominante]	$t(239)=4,48$ $p<0,0001$

Pode-se observar que as condições com efeito significativo foram constatadas apenas ao se comparar cada uma das condições com a condição SPND. Ao se comparar as condições SSD, SSND e SPD com a condição SPND, averiguou-se efeito significativo. Ao se verificar a comparação entre quaisquer outras das condições não se constatou efeito significativo. Observa-se que, mesmo com diferenças pequenas, as condições com forma no plural foram mais lentas que as condições com forma no singular.

Os gráficos expostos acima se referem aos tempos de leitura de palavras na variável *número* e na variável *dominância*, observados nos gráficos 5 e 6, respectivamente. Essas variáveis se referem ao tempo de leitura de palavras dos estímulos experimentais caracterizados como substantivos. Assim, observou-se o que levou mais ou menos tempo para ser processado, isto é, ao se verificar no gráfico 5, percebe-se que as palavras no singular foram processadas mais rapidamente, obtendo, com isso, tempo de leitura de palavras menor (369ms) ao se comparar com o tempo de leitura das palavras no plural (386ms), que levaram mais tempo para serem processadas. Salienta-se que o plural foi mais lento, analisando-se, assim, que a forma flexionada é acessada por decomposição. Aqui, faz-se importante salientar uma comparação com o teste de familiaridade 1 feito através do *Google Forms*, os resultados apontaram preferência pela forma singular como mais familiar, coincidindo com os dados obtidos aqui na medida *on-line*. Em relação à variável *dominância*, isolaram-se os tempos de leitura de palavras das formas dominantes e das formas não dominantes. Como o esperado, o tempo de leitura de palavra das formas dominantes (370ms) foi mais rápido que o tempo de leitura de palavra das formas não dominantes (385ms). Com isso, verifica-se que a frequência de dominância é um fator que influencia no processamento, contribuindo para um tempo de leitura de palavras mais rápido. Observando-se os dados da ANOVA, não foi constatado efeito significativo de interação entre *número* e *dominância* ($F(1,239) = 1,29$ $p<0,25$).

Em resumo, os resultados obtidos aqui, no que se refere aos substantivos, direcionam-se para aqueles encontrados no estudo do gênero realizado por Corrêa, Almeida e Porto

(2004). Os resultados indicaram que a forma singular é acessada por inteiro, como o são as formas flexionadas em gênero, e a forma plural, por sua vez, é acessada por decomposição, ressalvados os casos de formas plurais com frequência dominante, que também se comportam como as formas no singular. Assim, substantivos flexionados em número no plural tendem a ser representadas *Full Parsing* e acessados pelo modelo *morpheme-based*, isto é, por decomposição. Porém, uma vez que suas formas plurais se tornam dominantes, podem receber representação *Full Listening*, sendo então acessadas pelo modelo *full-form*, já as formas singulares são representadas em *Full Listening*.

Adjetivos

Uma mesma análise foi conduzida com os adjetivos. O gráfico 7 traz o tempo de leitura de palavras das condições experimentais utilizadas.

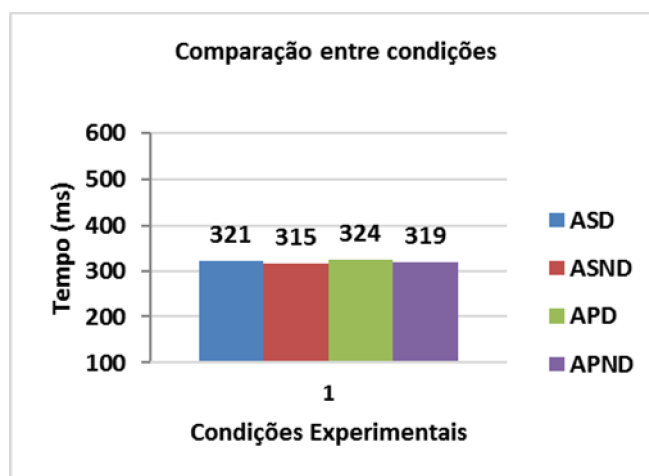


Gráfico 7: Teste-t (adjetivos).

O gráfico acima apresenta os tempos médios de resposta à tarefa experimental. Os resultados também foram submetidos à análise de variância, *design* fatorial 2 (número) x 2 (dominância), a fim de se comprovarem as hipóteses aqui assumidas. Os resultados mostraram um efeito significativo de *número* ($F(1,239) = 5,01$ $p < 0,02$), indicando que a flexão do adjetivo afetou os tempos médios de leitura. Observando-se as médias obtidas nos adjetivos singulares, comparadas aos adjetivos na forma no plural, verificou-se que aquelas foram lidas mais rapidamente que estas, conforme ilustra o gráfico abaixo:

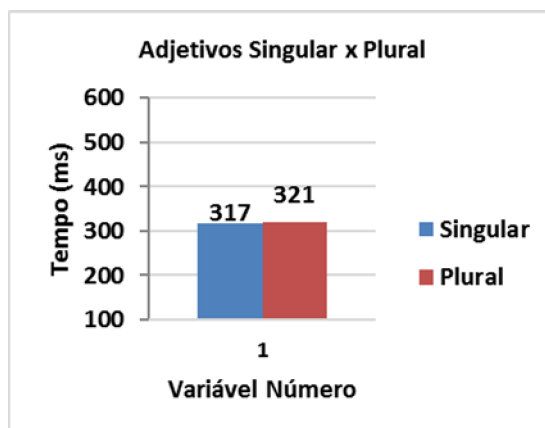


Gráfico 8: Número dos adjetivos.

Um efeito significativo de *dominância* também foi registrado (Dominância $F(1,239) = 3,59$ $p < 0,05$), sugerindo que a frequência dominante também afeta os tempos médios de leitura. Formas flexionadas dominantes são lidas mais rapidamente do que as suas contrapartes não dominantes, de acordo com o ilustrado no gráfico a seguir:

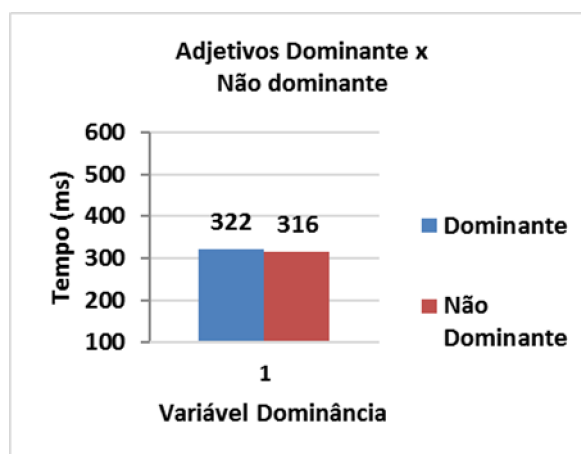


Gráfico 9: Dominância dos adjetivos.

Assim, como se aplicou nos resultados referentes aos substantivos, também se realizou aqui um teste comparativo de médias (teste-t), com correção de Bonferroni, aplicado para a comparação das condições entre si, tomando-as duas a duas. Os resultados são mostrados na tabela a seguir:

Quadro 4: Comparações entre médias

PAIRWISE COMPARISONS (comparações duas a duas)	
[Singular Dominante]vs[Singular_Não Dominante]	$t(239)=1,87$ $p < 0,0622$
[Singular Dominante]vs[Plural_Dominante]	$t(239)=1,17$ $p < 0,2435$
[Singular Dominante]vs[Plural_Não Dominante]	$t(239)=0,53$ $p < 0,5939$
[Singular Não Dominante] vs [Plural_Dominante]	$t(239)=3,33$ $p < 0,0010$
[Singular Não Dominante] vs [Plural_Não Dominante]	$t(239)=2,05$ $p < 0,0414$
[Plural Dominante]vs[Plural_Não Dominante]	$t(239)=1,48$ $p < 0,1403$

Pode-se observar que as condições com efeito significativo foram constatadas apenas ao se comparar as condições APD e APND com a condição ASND. Ao se verificar a comparação entre quaisquer outras das condições não se constatou efeito significativo. As condições com forma no plural foram mais lentas que as condições com forma no singular.

Os gráficos expostos anteriormente se referem aos tempos de leitura de palavras na variável *número* e na variável *dominância*, observados nos gráficos 8 e 9, respectivamente. Essas variáveis se referem ao tempo de leitura de palavras dos estímulos experimentais caracterizados como adjetivos. Assim, observou-se o que levou mais ou menos tempo para ser processado, isto é, ao se verificar no gráfico 8, percebe-se que as palavras no singular foram processadas mais rapidamente, obtendo, com isso, tempo de leitura menor (317ms) ao se comparar com o tempo de leitura das palavras no plural (321ms), que levaram mais tempo para serem processadas. Salienta-se que o plural foi mais lento, analisando-se, assim, que a forma flexionada é acessada por decomposição. Aqui, faz-se importante salientar uma comparação com o teste de familiaridade 2 feito através do *Google Forms*, os resultados apontaram preferência exclusiva pela forma no singular como mais familiar, coincidindo com os dados obtidos aqui na medida *on-line*. Em relação à variável *dominância*, isolaram-se os tempos de leitura das palavras das formas dominantes e das formas não dominantes. O tempo de leitura das formas dominantes (322ms) foi mais lento que o tempo de leitura das formas não dominantes (316ms). Observando-se os dados da ANOVA, não foi constatado efeito significativo de interação entre *número* e *dominância* (Número*Dominância $F(1,239) = 0,431$ $p < 0,5$).

Em resumo, os resultados obtidos aqui, no que se referem aos adjetivos, também se direcionam para aqueles encontrados no estudo do gênero realizado por Corrêa, Almeida e Porto (2004). Salientando-se que, o estudo citado anteriormente não analisou separadamente substantivos e adjetivos. Ao se estabelecer a divisão entre substantivos e adjetivos para análise dos dados, notam-se os mesmos direcionamentos nos resultados obtidos naquele estudo, como já foi dito acima.

Abaixo, mostra-se o quadro com os efeitos das variáveis independentes e os efeitos de interação obtidos através da análise de variância do experimento sobre o número regular.

Tabela 5: ANOVA das variáveis independentes.

Classe gramatical	$F(1,478) = 33,8 \text{ } p < 0,000001$
Número	$F(1,478) = 6,46 \text{ } p < 0,011347$
Dominância	$F(1,478) = 1,37 \text{ } p < 0,241747$
Classe gramatical*Número	$F(1,478) = 2,83 \text{ } p < 0,093314$
Classe gramatical*Dominância	$F(1,478) = 6,49 \text{ } p < 0,011187$
Número*Dominância	$F(1,478) = 1,58 \text{ } p < 0,209687$
Classe gramatical*Número*Dominância	$F(1,478) = 0,933 \text{ } p < 0,337543$

Os efeitos significativos foram evidenciados em: *classe gramatical*, *número* e *interação entre classe gramatical e dominância*. Os dados com os referidos efeitos significativos, além de expostos na tabela acima, seguem a seguir: classe gramatical ($F(1,478) = 33,8 \text{ } p < 0,000001$); número ($F(1,478) = 6,46 \text{ } p < 0,011347$), interação classe gramatical*dominância ($F(1,478) = 6,49 \text{ } p < 0,011187$).

Esses resultados não fornecem evidências favoráveis à hipótese aqui assumida, de que substantivos flexionados em número em PB são acessados e representados por mecanismos de dupla rota. Pode-se aferir que a flexão de número da forma plural apresenta, pelo menos, mais uma computação no momento do processamento em relação à forma no singular que, via de regra, apresenta apenas uma computação. Pode-se observar ainda que a frequência exerceu um papel importante na determinação da forma como os substantivos plurais foram acessados e representados.

Considerações finais

O estudo concernente ao léxico mental do PB aqui conduzido mostrou evidências dos efeitos de frequências. Com isso, formulou-se a hipótese de que efeitos de frequência de dominância podem influenciar o tempo de leitura de palavras dos sujeitos em um teste experimental sobre os vocábulos com flexão regular e irregular em número. Os estudos realizados no espanhol por Dominguez, Cuetos e Segui (1999) atestaram efeitos significativos na forma de frequência de superfície, já os estudos feitos por Corrêa, Almeida e Porto (2004) encontraram efeitos significativos de frequência dominante e cumulativa. Observando-se esses dois estudos, formulou-se o teste experimental deste trabalho. O *design* experimental foi

criando para verificar e analisar se os efeitos encontrados para gênero no estudo realizado por Corrêa, Almeida e Porto (2004), também podem ser encontrados para o número em PB.

Os resultados encontrados aqui no experimento sobre número regular foram na mesma direção daqueles encontrados para o gênero em PB. Quando a forma de conceber os itens linguísticos, no momento do acesso, é dada por inteiro ó *full form* ó utiliza-se a *Full Parsing Hypothesis*. Quando a forma de conceber os itens linguísticos, no momento do acesso, é dada por decomposição - *morpheme-based* - utiliza-se a *Full Listing Hypothesis*. Pode-se afirmar que a frequência das formas é um fator a ser considerado no reconhecimento das palavras. Em relação ao número (singular e plural), a forma no singular é acessada mais rapidamente, tendo em vista que, a forma no singular é acessada por inteiro e a forma no plural por decomposição, já que é acessada mais lentamente. É importante mencionar os casos das formas no plural com frequência dominante, o acesso às formas também é dado por inteiro e não por decomposição. Esses resultados não corroboram explicitamente para um modelo de dupla rota, podem, apenas, apontar que as formas de plural, que sofreram decomposição, sofreram mais uma decomposição para além daquelas das formas de singular. Por fim, faz-se necessário pensar que os dados obtidos no experimento sobre o número regular tendem a fornecer importantes informações sobre o modo de se conceber o léxico mental. Assim, constata-se que podem indicar uma possível interação entre linguística e psicolinguística, precisando-se, desse modo, de mais dados experimentais para assegurar de fato uma relação mais evidente.

Referências

- BAAYEN, R. H.; DUKSTRA, T; SCHEREUDER, R. Singulars and plurals in dutch: Evidence for a parallel dual route model. *Journal of Memory and Language*, Nova York, v. 37, p. 94-117, jul. 1997.
- BUTTERWORTH, B. Lexical Representation. In: _____. (Org.) *Language production: development, writing and other language processes*. London: Academic Press, v.2, 1983.
- CARAMAZZA, A., LAUDANI, A., ROMANI, C. Lexical acess and inflection morphology. *Cognition*. v. 28, 2. ed, p. 297-332, abr. 1988.
- COLÉ, P.; SEGUI, J.; TAFT, M. Words and Morphemes as Units for Lexical Access. *Journal of Memory and Language*, 37, 1997.

CORPUS DO PORTUGUÊS: banco de dados. Disponível em: <<http://www.corpusdoportugues.org>>. Acesso em: 17 nov 2015.

CORRÊA, L. M. S.; ALMEIDA, D. A. de A.; PORTO, R. S. On the representation of Portuguese gender-inflected words in the mental lexicon. *Brain and Language*, Rio de Janeiro, v. 90, p. 63-73, jul.-set. 2004.

CUNHA, Antônio Sérgio Cavalcanti da. *Palavras derivadas no léxico mental: abordagens gerativas e psicolinguísticas*. 2000. 105 f. Tese (Doutorado em Letras) - PUC: Rio de Janeiro, 2000.

DOMINGUEZ, A.; CUETOS, F.; SEGUI, J. The Processing of Grammatical Gender and Number in Spanish. *Journal of Psycholinguistic Research*, Estados Unidos, v. 28, n. 5, 1999.

EMMOREY, K. D.; FROMKIN, V. A. The mental lexicon. In: NEWMAYER, F. J. *Language: psychological and biological aspects*. Cambridge: Cambridge University Press, 1988.

LAINE, M; VAINIO, S; HYONA, J. Lexical access routes in morphologically rich language. *Journal of Memory and Language*, 40, 1999.

MIMOUNI, Z; KEHAYIA, E; JAREMA, G. The mental representation of singular and plural nouns in Algerian as revealed through auditory priming in agrammatic aphasic patients. *Brain and Language*, 61, 1998.

SERENO, J. A; JONGMANN, A. *Processing of English Inflectional Morphology Memory and Cognition*, 25, 1997, p. 426 - 437.

TAFT, M; FORSTER, K. I. Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, v. 14, 1975.

TAFT, M; FANNY, M. Lexical representation of gender: a quasiregular domain. *Journal of Psycholinguistic Research*. v. 27, n. 1, 1988.

TAFT, M. Interactive activation as framework for understanding morphological processing. In: *Language and Cognitive Processes*, Special Issue, ago. 1994, p. 271-294.

Number inflected forms in Brazilian Portuguese: access and lexical representation

Abstract: This article investigates the processing of number inflected words in Brazilian Portuguese (BP), and it suggests a model of representation for them. It was observed how number inflected words are stored in the mental lexicon and retrieved from it. Recent studies have pointed out that the frequency of the base affects the processing of word inflected both in gender and in number (DOMINGUEZ; CUETOS; SEGUI, 1999). Regarding gender inflection in BP, a frequency effect was

observed (CORRÊA; ALMEIDA; PORTO, 2004). In order to provide evidences concerning the dominant frequency of forms inflected in number, an experiment was carried out, which consisted of a word reading test, formed by words with regular number inflection. It was intended to observe if the singular form was accessed in the whole, taking into account dominant frequency effects, or if the regular plural form was accessed by decomposition, taking into account non-dominant frequency effects, excepting the cases of high frequency of plural forms, in which this forms would also be accessed in the whole, in a double-route model. The results suggest that the frequency dominant words are processed faster than the ones with non-dominant frequency.

Keywords: Number Inflection. Mental Lexicon, Brazilian Portuguese.

Recebido em: 30 de junho de 2017.

Aprovado em: 09 de julho de 2017.