

ISSO É UMA PALAVRA?¹

Kátia ABREU*

Instituto Superior Anísio Teixeira

Maria Carlota ROSA**

Universidade Federal do Rio de Janeiro

RESUMO

O presente artigo focaliza a reação de falantes a um conjunto de siglas. Eles deveriam decidir se estavam ou não diante de uma palavra. As respostas obtidas parecem evidenciar um grau considerável de incerteza. Os resultados autorizam o tratamento da formação de siglas como um tipo de morfologia especial, tal como as poucas descrições que abordam o tema vêm fazendo.

ABSTRACT

This paper focuses on speakers' reactions to a set of acronyms. The speakers would decide if a given acronym was a word or not. The answers they gave may be interpreted as evidence of a considerable amount of uncertainty in classifying the data. The results allow us to focus the formation of acronyms as a special type of morphology, as the few linguistic papers which deal with the subject do.

PALAVRAS-CHAVE

sigla, palavra, decisão lexical, léxico, morfologia improdutiva.

KEYWORDS

acronym, word, lexical decision, lexicon, unproductive morphology.

1. Introdução

O objetivo que os estudos morfológicos têm perseguido nos últimos 30 anos é o de representar o conhecimento do falante sobre o vocabulário, que lhe permite reconhecer as formações com as quais se depara como

próprias de sua língua ou não, rejeitar aquelas que não considere bem-formadas, estabelecer relações vocabulares e formar palavras novas, ciente não apenas das possibilidades de combinação, mas também das restrições de seleção (Basílio, 1980; Aronoff e Fuhrhop, 2002). Este artigo restringe-se a focalizar uma parte da expressão desse conhecimento, ou seja, que o falante sabe o que é uma palavra em sua língua – e o que não é. Na sua forma mais simples, a questão aqui focalizada diz respeito ao fato de que, caso apresentadas a um falante de português formas, por exemplo, como *clone* ['kloni] e *snrocs* [snrɔks] e a pergunta se seriam palavras do português, a resposta seria 'sim' para a primeira e 'não' para a segunda sem hesitação.

Podem, porém, incidir sobre as decisões de um falante real – a fonte dos dados das pesquisas – alguns complicadores. Um deles poderia resultar de uma falha na percepção de redundâncias numa dada estrutura. *Sindicado*, por exemplo, é termo de baixíssima frequência fora de um contexto jurídico relacionado a comissões de sindicância; o mesmo para os relacionados *sindicar*, *sindicante*. Em decorrência, um falante poderia considerar que *sindicado* não existe – e para ele isso seria verdadeiro – e que a forma correta (isto é, a que existe) é *sindicato*.

Um outro fator que poderia acrescentar algum grau de incerteza à resposta seria um falante ter consciência de que há muitas palavras que desconhece. A resposta à mesma pergunta apresentada anteriormente poderia não ser tão firme se esse informante tivesse de decidir se *elevamento* é palavra do português. Embora haja muitos nomes derivados de verbo que seguem o mesmo padrão de formação – como *deslocamento*, *cancelamento* –, para aquele que se utiliza apenas de *elevação*, pelo menos quatro alternativas quanto a *elevamento* seriam possíveis: *elevamento* poderia não ser aceito porque já existe *elevação* como sinônimo exato; poderia ser considerado uma palavra até possível, mas seria melhor confirmar sua existência num dicionário; poderia ser conhecida mas não empregada; ou ainda uma palavra com significado não exatamente igual ao de *elevação*².

Quaisquer que sejam as interferências, as respostas levantam questões acerca da relação entre os dados obtidos e o nível de abstração da descrição. Por um lado, seria possível argumentar que um falante ideal certamente reconheceria quaisquer redundâncias, e a descrição deveria desconsiderar as falhas; por outro, o não reconhecimento por vários informantes também poderia ser imputado a um falante ideal, que, nesse caso, representaria que no nível não mais de um indivíduo, mas do grupo, teria havido mudança. O segundo fator acima diz respeito ainda ao reconhecimento de uma interferência extralingüística na decisão. Ao recomendar a consulta a um dicionário para dirimir sua dúvida, indiretamente o falante está indicando: (a) que a palavra é formada de acordo com o esperado; (b) que, apesar disso, tal palavra não faz parte do seu conhecimento; (c) que fatores sociais, representados pelo dicionário, podem aprovar ou impedir o uso.

Este artigo discute um caso demonstrativo de algumas das dificuldades na tarefa de levar em conta os dados que explicitam o conhecimento do falante sobre palavras. Afinal, nem sempre é absoluto o nível de certeza dos falantes, como parece estar pressuposto na descrição das tarefas que são capazes de realizar. E isso acontece mesmo que os dados para essa tomada de decisão sejam como *clone* e *snrocs*. *Clone* tem muitas chances de ser conhecida de qualquer brasileiro que viva numa comunidade servida de energia elétrica, porque recentemente deixou os círculos científicos, serviu de título a uma novela de televisão de grande audiência e entrou no uso coloquial. A outra, por sua vez, é fonologicamente estranha: poderia ser uma palavra, mas não em português. Mas e se fosse apresentada por escrito, ao invés de ser apresentada oralmente? E como <SNROCS>, com maiúsculas, e não como <snrocs>? Poderia ser uma sigla (que referisse um sindicato nacional, por exemplo, já que começa pela seqüência <sn> – as possibilidades de significado seriam muitas). Nesse caso, a pronúncia estaria de acordo com os padrões do português (Barbosa *et al.*, 2003) e essa seqüência seria plenamente aceitável.

A discussão sobre a capacidade de o falante decidir se um termo é ou não uma palavra é aqui levada para um tipo de formação de palavras considerado *especial* (Sandmann, 1988, p. 146; Alves, 1990, p. 56), que começou a se

tornar freqüente a partir dos anos 1960 (Cunha, 1972, p. 130; Kehdi, 1992, p. 51; Nunberg, 2002). São as *siglas* ou *acrônimos*³. Aronoff e Anshen (1998, p. 246) classificam as siglas como *morfologia improdutiva*, porque, diferentemente da morfologia com base em radicais e afixos, as siglas são formações intencionais.

Embora já tenham sido classificadas como um tipo especial de composição (Alves, 1990, p. 56)⁴, as siglas também foram comparadas a palavras primitivas (Kehdi, 1992, p. 52), porque não se tem idéia do intitutivo a partir do qual uma sigla foi gerada, a não ser que seja explicitado. A opacidade estrutural da maior parte das siglas⁵, a que, ao fim das contas, Kehdi faz referência, impede o estabelecimento de redundâncias entre diferentes siglas, já que seus elementos formadores podem expressar diferentes tipos de formação⁶ e podem remeter para as seqüências mais variadas, reunidas nos intitutivos. A sigla é comparável a uma palavra primitiva porque sua estrutura é indecomponível em radicais e afixos.

As secções que se seguem estão organizadas do seguinte modo: primeiramente, apresenta-se a descrição do material e do método de análise. Em seguida, os resultados, a discussão dos resultados obtidos e a conclusão.

2. Materiais e método

Ao tratar as siglas como um tipo especial de morfologia, que em última análise parte de letras cujo número e seqüência decidirão o modo como serão pronunciadas – se soletradas ou não (Barbosa *et al.*, 2003) –, a literatura lingüística revela o quanto esse tipo de formação se afasta dos padrões produtivos descritos em secções sobre derivação ou sobre composição.

Uma vez que as siglas são apresentadas como formações especiais e como um tipo recente de formação de palavras⁷ resolveu-se confirmar se, apesar disso, eram sentidas como palavras. Para isso elaborou-se um teste (Abreu, 2005), utilizando-se a técnica do teste de decisão lexical (*lexical*

decision task). Os 29 informantes, 17 mulheres e 12 homens, entre 18 e 28 anos, todos alunos do primeiro período da graduação em Letras na UFRJ, participaram do experimento voluntariamente. Nenhum deles tinha problemas de visão. Cada um dos informantes tinha de responder “sim” ou “não”, se era ou não uma palavra, a cada sequência de quatro letras que lhes era apresentada na tela de um monitor, apertando para isso um botão verde (“sim”) ou vermelho (“não”) de um controle (*bottom box*). O tempo necessário para que uma nova sequência surgisse na tela era comandado pelo informante, mas, uma vez na tela, sua decisão, concretizada na pressão do botão verde ou do botão vermelho, deveria ser tomada o mais rapidamente possível. O tempo de resposta para cada item foi medido em milissegundos. Num teste de decisão lexical assume-se que a velocidade da reação ao estímulo indica a ativação de uma memória lexical e que alguns fatores podem acelerar essa reação. Frequência é um deles. Palavras frequentes provocam respostas mais rápidas nesse tipo de experimento.

Compunham o teste 64 itens, apresentados por escrito, em maiúsculas, todos eles formados por quatro letras (*vide* Anexo 2). A frequência dos itens, que determinou sua escolha, teve por base o *corpus* CETEN-FOLHA. Metade dos 64 itens eram siglas. Das siglas, 16 eram silábicas, como *OTAN*, sendo oito de baixa frequência, como *IPEA*, e oito de alta frequência, como *AIDS*. As outras 16 siglas eram siglas soletradas, como *FGTS*. Também as siglas soletradas foram divididas em siglas de alta frequência, como *IPTU*, e siglas de baixa frequência, como *OCDE*⁸. Os outros 32 itens serviram de controle: 16 palavras primitivas, de alta e de baixa frequência (como *CASA* e *NOJO*, respectivamente), e 16 não-palavras, oito silábicas (como *GROL*) e oito soletradas (como *NJHT*). Para cada item houve, por conseguinte, 29 decisões. Cada grupo de oito itens, submetidos a 29 alunos, totaliza 232 decisões, divididas entre “sim” e “não”. Somados os oito grupos, obteve-se o total de 1856 decisões. A ordem em que os itens foram apresentados a cada informante foi aleatória.

As instruções do teste foram apresentadas a cada informante por escrito, na tela do monitor, ao mesmo tempo em que eram lidas em voz alta pelo

aplicador. Após a leitura, o aplicador simulava o experimento, familiarizando o informante com o equipamento e certificando-se de que as instruções haviam sido compreendidas. Após essa prática, o aplicador retirava-se do laboratório e o informante realizava o teste a sós.

A expectativa inicial era de que: (a) as palavras primitivas de alta frequência, como CASA, seriam mais rapidamente reconhecidas como palavras; (b) as siglas soletradas de baixa frequência e as não-palavras soletradas teriam tempo equivalente de decisão; (c) as não-palavras silábicas teriam o maior tempo de decisão, porque poderiam ser palavras desconhecidas.

3. Resultados

A Tabela 1 a seguir resume os resultados obtidos.

TABELA 1
Resumo dos resultados

Categoria	dec sim ¹	%	tr sim ² (ms)	dec não	%	tr não (ms)	Exemplo
PA ³	231	99,5	1086	1	0,4	1780	CASA
PB ⁴	224	96,5	1301	8	3,4	4490	SELA
SLA ⁵	111	47,8	1437	121	52,1	1598	INSS
SSB ⁶	101	43,5	2668	131	56,4	2471	IPEA
SSA ⁷	87	37,5	2291	145	62,5	2109	OTAN
NS ⁸	55	23,7	2503	177	76,2	2366	TILA
SLB ⁹	49	21,1	2357	183	78,8	1593	PDBG
NL ¹⁰	9	3,8	2225	223	96,1	1658	MRIT

¹dec = decisão. – ²tr = tempo de resposta (em milissegundos). – ³PA = palavra de alta frequência. – ⁴PB = palavra de baixa frequência. ⁵SLA = sigla soletrada de alta frequência. – ⁶SSB = sigla silábica de baixa frequência. – ⁷SSA = sigla silábica de alta frequência. – ⁸NS = não-palavra silábica. ⁹SLB = sigla silábica de baixa frequência. – ¹⁰NL = não-palavra soletrada.

Os resultados indicam que os alunos participantes reconhecem os itens dos grupos *palavra de alta frequência* (PA) e *palavra de baixa frequência* (PB) como palavras. Não há aqui evidências de qualquer efeito de frequência, mesmo para os itens considerados de baixa frequência no experimento.

Da mesma forma, os alunos participantes reconhecem como *não-palavras* os itens dos grupos *não-palavra silábica* (NS) e *não-palavra soletrada* (NL). A formação silábica, porém, é capaz de influenciar algumas decisões para “sim”. O falante pode considerar um item desse tipo como sendo uma palavra existente da qual ele não tem conhecimento. Tal fato não se verifica no grupo *não-palavra soletrada* (NL) em que cada item apresenta uma estrutura que não é possível no português. A quantidade de decisões “não”, por isso, quase alcança a totalidade.

O número de decisões “não” e o número de decisões “sim” do grupo *sigla soletrada de alta frequência* (SLA) estão muito próximos, o que pode indicar a existência de dúvida no julgamento. Nesse caso, pode ser efeito da frequência o fato de os alunos terem percebido que se tratava de siglas.

O grupo *sigla soletrada de baixa frequência* (SLB) apresentou alto índice de rejeição. Nesse caso, a formação pode ter sido a responsável pelo julgamento, já que as siglas exibidas tinham seqüências de consoantes em um padrão não possível no português. O número de decisões “não” foi superior ao demonstrado no grupo NS, que apresenta itens que não são palavras do português, mas que têm estrutura silábica.

4. Discussão

Uma das maneiras de representar o conhecimento que permite ao falante tomar decisões acerca de palavras apela para módulos distintos de conhecimento, o *léxico* e a *morfologia* (Aronoff, 1976; Aronoff e Anshen, 1998). Todas as palavras complexas cujo significado seja previsível com base na depreensão de sua estrutura são *palavras potenciais* e ficam a cargo da morfologia. As palavras cujo significado não seja previsível pela análise de sua estrutura são parte do léxico. Segundo Aronoff e Anshen (1998,

p. 238), as *palavras existentes* são aquelas que compõem o léxico mental do falante, que ambos os autores fazem equivaler a uma lista de itens irregulares. Comporiam essa lista as palavras primitivas que, por não terem estrutura interna, representariam a arbitrariedade saussureana da relação entre som e significado. Também comporiam essa lista palavras morfológicamente complexas cujos componentes não sejam, todos eles, reconhecidos pelo falante-ouvinte; e, por fim, também estariam no léxico as palavras cujo significado não fosse depreensível da soma de suas partes componentes.

[w]e will say that any word that is stored in a single speaker/hearer's mental lexicon or list of irregular items is an existing word, and that nothing else is. In particular, a word that meets all the criteria for being a word of the language but that is not in an individual's mental lexicon does not exist for that person, though it may exist for another speaker/hearer.

[diremos que qualquer palavra que está armazenada no léxico mental de um falante-ouvinte ou lista de itens irregulares é uma palavra existente e nada mais que isso. Em particular, uma palavra que preenche todos os critérios para ser uma palavra da língua mas que não está no léxico mental de um indivíduo não existe para ele, embora possa existir para um outro falante-ouvinte.]

Por seu turno, não fariam parte da lista ou léxico nem as palavras derivadas nem os compostos bem-formados:

The unlisted word is a potential word, and we will say that morphologically well-formed complex potential words are provided by the morphology, not by the lexicon. Thus, the conventional idea that the existing words of a language [...] comprise all the words in [...] some [...] comprehensive dictionary does not apply in this model of the lexicon and the morphology. The difference between which words exist and which are potential is defined solely in terms of the individual's lexicon and morphology.

[A palavra não listada é uma palavra potencial e diremos que as palavras potenciais complexas bem-formadas são providas pela morfologia, não

pelo léxico. Assim, a idéia convencional de que as palavras existentes de uma língua [...] compreendem todas as palavras em [...] algum [...] grande dicionário não se aplica neste modelo do léxico e da morfologia. A diferença entre que palavras existem e que palavras são potenciais define-se tão-somente em termos do léxico e da morfologia individuais.]

As siglas, como já afirmado anteriormente, são formações que se assemelham a palavras primitivas, na medida em que sua estrutura é opaca. Além disso, tal como outras palavras da língua, as siglas podem servir de base a derivados: *petista*, *pefelista*, *aidético*, *cepeguinho* são alguns exemplos. Nos termos propostos por Aronoff e por Aronoff e Anshen, os itens criados pela *morfologia improdutiva* estariam armazenados no léxico. O acesso a esses itens poderia ser facilitado pela sua freqüência. O que remete ao experimento apresentado nas secções anteriores.

Os resultados obtidos confirmaram algumas das expectativas. Como esperado, as palavras primitivas freqüentes (PA), como CASA, OBRA ou LOJA, ficaram no topo das decisões no reconhecimento como palavras, embora tenha causado estranheza que um informante não tenha reconhecido em CASA uma palavra. Um pouco abaixo ficaram as palavras de baixa freqüência (PB), como ANTA, BULE ou SELA.

Esse resultado numericamente próximo levou a rever os itens dos grupos *palavra* (PA e PB) e a questionar os critérios empregados para considerar uma palavra muito ou pouco freqüente. Em princípio, a classificação como muito freqüente estaria informando que um item lexical é muito empregado, a ponto de não só ser conhecido, mas de poder ser recuperado com facilidade. A classificação das palavras numa escala de freqüência foi colhida numa base de dados eletrônica, o CETEN-Folha, *corpus* de cerca de 24 milhões de palavras do português do Brasil, construído com base em extratos de notícias de todas as edições do ano de 1994 do jornal *Folha de S.Paulo*. Freqüência, nesse *corpus*, não é necessariamente um indicador de que um termo é ou não freqüente, em abstrato, mas que ele é freqüente ou não no português escrito, culto e em áreas específicas como política, economia ou divertimento. Nesse

sentido, seria realmente pouco usual o emprego de BAFO ou de ANTA em boa parte da cobertura de fatos por esse jornal. Por outro lado, essa base de dados é rica em siglas, que não são comuns em outro tipo de publicação, como as revistas voltadas para o público adolescente ou para o público feminino, por exemplo.

As não-palavras também confirmaram as expectativas iniciais. A quase totalidade das decisões para as não-palavras soletradas (NL) foi “não”. Comparado esse resultado com o obtido para as não-palavras silábicas (NS), confirmou-se a previsão de que não-palavras “ilegais”, como *ONRS*, seriam mais fáceis de classificar do que as não-palavras “legais”, como *FESA* (*vide* Wagenmakers *et al.*, 2004): o tempo de resposta foi bem maior para aquelas do que para estas. As siglas soletradas de baixa frequência (SLB) foram interpretadas como seqüências ilegais e também foram classificados como não-palavras num número expressivo de decisões.

As demais siglas (SLA, SSB e SSA) alcançaram resultados muito próximos para “sim” e para “não”, com leve vantagem para serem consideradas não-palavras. Um aspecto deverá ser revisto: haveria uma diferença significativa nos resultados caso os itens que compunham o teste fossem apresentados em minúsculas? A questão surgiu em razão de as siglas, especialmente as soletradas, serem usualmente grafadas em maiúsculas, e isto poderia ter gerado um viés nos resultados obtidos. Por outro lado, caso a apresentação de todos os itens tivesse sido feita em minúsculas não haveria como evitar o viés oposto. Some-se a isso que algumas das siglas são homônimas de não-siglas: *INCA* ou *PISA*, por exemplo. Se apresentada a seqüência como <*inca*> ou <*pisa*>, muito possivelmente não se poria para o leitor a possibilidade de estar diante de siglas. Esta questão deriva do tempo de resposta em todos os grupos. Com dados para o inglês e com base em experimentos não com seqüências escritas, mas sim com seqüências sonoras, Aitchison (1994, p. 8) refere cerca de meio segundo (450 ms) como o tempo para rejeitar uma não-palavra. No experimento aqui relatado, o tempo de decisão para os itens dos diferentes grupos vai de pouco mais de um segundo (1086 ms) a quase quatro segundos e meio (4490 ms). No caso das siglas SLA, SSB e

SSA, tanto nas respostas para “sim” como para “não”, o tempo de decisão foi aproximado e longo.

Uma outra questão que deriva da apresentação escrita é que a possível dúvida quanto a estar diante de seqüências “ilegais” ou de siglas não teria lugar se a apresentação dos dados fosse também feita oralmente.

5. Conclusão

Numa proposta que busca representar a competência do falante no que respeita à formação e estrutura das palavras, dados como os apresentados aqui evidenciam a dificuldade de um falante real em reconhecer esse tipo de morfologia improdutiva que são as siglas como um processo que forme palavras.

Os resultados obtidos podem ser interpretados como evidência de que as siglas, embora cada vez mais freqüentes na atualidade, têm situação ímpar em relação ao restante do léxico. As razões que levam diferentes autores a dar tratamento diferenciado a esse tipo de formação parece refletir-se no desempenho dos falantes. Com base no desempenho no teste, mesmo se examinado o resultado para cada indivíduo, independentemente de qual tenha sido a decisão tomada, o tempo de decisão evidencia alguma dificuldade em classificar as seqüências como palavras ou não-palavras.

Notas

- * Kátia Abreu é doutoranda em Lingüística na Universidade Federal do Rio de Janeiro e professora do Instituto Superior Anísio Teixeira. *E-mail*: katia.nazareth@terra.com.br.
- ** Maria Carlota Rosa é doutora em Lingüística pela Universidade Federal do Rio de Janeiro e professora associada da Universidade Federal do Rio de Janeiro. *E-mail*: carlota@ufrj.br.
- 1 Agradecemos ao Prof. Marcus Maia (UFRJ), por ceder seu laboratório para o teste, e aos 29 alunos do primeiro período da Faculdade de Letras da UFRJ em 2005/1 que aceitaram participar do teste.
- 2 Apresentada a pergunta a três professores universitários, em meio a conversas informais em ocasiões distintas, e não como parte de pesquisa, surgiram respostas diferentes: que uma existe e a outra, não; que ambas existem, mas uma delas é mais comum; que ambas existem mas são um pouco diferentes, porque *elevamento* pressupõe um artefato mecânico que gera o movimento e por essa razão se diria *elevação de preços* e não *elevamento de preços*. As respostas foram acompanhadas da modalização “acho que” e do conselho de procurar um dicionário.
- 3 Os termos são aqui empregados como sinônimos. Para uma revisão da literatura, *vide* Abreu (2004).
- 4 Alves (1990, p. 50) considera as siglas um tipo de *composição sintagmática*, porque nelas “os membros integrantes de um segmento frasal encontram-se numa íntima relação sintática, tanto morfológica quanto semanticamente, de modo a constituírem uma única unidade léxica”.
- 5 Não serão levadas em conta siglas do tipo *Embrafilme* ou *Portobras*.
- 6 O tipo de sigla mais comum reúne as iniciais de cada palavra do intitutivo, mas as opções de formação são variadas (*vide* Sandmann, 1988; Alves, 1990; Abreu, 2004).
- 7 Não se faz referência aqui a abreviaturas empregadas na escrita que, se oralizadas, são pronunciadas por extenso, como *Ilmo*, *Sr*.
- 8 Para a tradução das siglas empregadas, *vide* Anexo 1.

Referências

- ABREU, Kátia Nazareth Moura de. *Um caso de morfologia improdutiva no português do Brasil: a formação de siglas e de acrônimos*. Dissertação (Mestrado), UFRJ, Rio de Janeiro, 2004.
- _____. *O reconhecimento das siglas como palavras da língua portuguesa: uma análise preliminar*. Monografia final apresentada à disciplina de doutorado Introdução à Psicolinguística/ Tópicos Avançados em Processamento Sintático, ministrada pelo Prof. Marcus Maia, no Programa de Pós-graduação em Linguística da UFRJ. 2005.
- AITCHISON, Jean. *Words in the mind: An introduction to the mental lexicon*. Oxford: Blackwell, 1994.
- ALVES, Ieda Maria. *Neologismos: criação lexical*. São Paulo: Ática, 1990.
- ARONOFF, Mark. *Word formation in generative grammar*. Cambridge, Mass.: The MIT Press, 1976.
- ARONOFF, Mark; ANSHEN, Frank. Morphology and the lexicon: lexicalization and productivity. In: SPENCER, Andrew; ZWICKY, Arnold (Ed.). *The handbook of morphology*. Oxford: Blackwell, 1998. p. 236-247.
- ARONOFF, Mark; FUHRHOP, Nanna. Restricting suffix combinations in German and English: Closing suffixes and the monosuffix constraint. *NLLT*, n. 20, p. 451-490, 2002.
- BARBOSA, Filipe; ROSA, Maria Carlota; GONÇALVES, Carlos Alexandre; RESENDE Jr., Fernando Gil. Algoritmo para a leitura de siglas em um sintetizador de voz. In: *Anais do XX Simpósio Brasileiro de Telecomunicações*. Rio de Janeiro: IME e PUC-RJ, 2003. p. 672-675.
- BASÍLIO, Margarida. *Estruturas lexicais do português: uma abordagem gerativa*. Petrópolis: Vozes, 1980.
- CETEN-Folha (Corpus de Extractos de Textos Electrónicos NILC/Folha de S.Paulo). Disponível em: <http://acdc.linguatca.pt/cetenfolha/>. Acesso em 26/07/2006.

CUNHA, Celso Ferreira da. *Gramática da língua portuguesa*. Rio de Janeiro: MEC/FENAME, 1972.

KEHDI, Valter. *Formação de palavras em português*. São Paulo: Ática, 1992.

NUNBERG, Geoffrey. Letter perfect, 2003. Disponível em: <http://www-csli.stanford.edu/~nunberg/acronyms.html>. Acesso em 26/07/2006.

SANDMANN, Antônio José. *Formação de palavras no português brasileiro contemporâneo*. Curitiba: Scientia & Labor e Ícone, 1988.

WAGENMAKERS, Eric-Jan M.; ZEELENBERG, René; STEYVERS, Mark; SHIFFRIN, Richard; RAAIJMAKERS, Jeroen G. W. Nonword repetition in lexical decision: Support for two opposing processes. *The Quarterly Journal of Experimental Psychology*, n. 57, p. 1-20, 2004.

Anexo 1

Siglas empregadas no teste e seu significado

INCA	Instituto Nacional de Câncer
AIDS	<i>Acquired Immune Deficiency Syndrome</i>
OTAN	Organização do Tratado do Atlântico Norte
INPE	Instituto Nacional de Pesquisas Espaciais
ECAD	Escritório Central de Arrecadação e Distribuição
UNIP	Universidade Paulista
CIMI	Conselho Indigenista Missionário
AIEA	Agência Internacional de Energia Atômica
IPEA	Instituto de Pesquisa Econômica Aplicada
CECA	Comunidade Européia do Carvão e do Aço
DARF	Documento de Arrecadação de Receitas Federais
ALCA	Área de Livre Comércio das Américas
OVNI	Objeto Voador Não Identificado
SAEB	Sistema de Avaliação da Educação Básica
PISA	Programa Internacional de Avaliação de Alunos
FLIP	Festa Literária Internacional de Parati
INSS	Instituto Nacional de Seguridade Social
IBGE	Instituto Brasileiro de Geografia e Estatística
PSDB	Partido da Social Democracia Brasileira
PMDB	Partido do Movimento Democrático Brasileiro
ICMS	Imposto sobre Circulação de Mercadorias e Serviços
FGTS	Fundo de Garantia por Tempo de Serviço
CNBB	Conselho Nacional dos Bispos do Brasil
IPTU	Imposto Predial, Territorial e Urbano

ISSO É UMA PALAVRA?

IBPT Instituto Brasileiro de Planejamento Tributário
ABLH Associação Brasileira de Lan Houses
CPMF Contribuição Provisória sobre Movimentação Financeira
SMTU Secretaria Municipal de Transportes Urbanos
PDBG Programa de Despoluição da Baía de Guanabara
IGPM Índice Geral de Preços de Mercado
CNPq Conselho Nacional de Desenvolvimento Científico e Tecnológico
OCDE Organização para Cooperação e Desenvolvimento Econômico

Anexo 2

Grupos de dados do Experimento

Siglas silábicas de alta frequência (SSA)	Siglas soletradas de alta frequência (SLA)
SSA 1 INCA	SLA 1 INSS
SSA 2 AIDS	SLA 2 IBGE
SSA 3 OTAN	SLA 3 PSDB
SSA 4 INPE	SLA 4 PMDB
SSA 5 ECAD	SLA 5 ICMS
SSA 6 UNIP	SLA 6 FGTS
SSA 7 CIMI	SLA 7 CNBB
SSA 8 AIEA	SLA 8 IPTU

Siglas silábicas de baixa frequência (SSB)	Siglas soletradas de baixa frequência (SLB)
SSB 1 IPEA	SLB 1 IBPT
SSB 2 CECA	SLB 2 ABLH
SSB 3 DARF	SLB 3 CPMF
SSB 4 ALCA	SLB 4 SMTU
SSB 5 OVNI	SLB 5 PDBG
SSB 6 SAEB	SLB 6 IGPM
SSB 7 PISA	SLB 7 CNPQ
SSB 8 FLIP	SLB 8 OCDE

ISSO É UMA PALAVRA?

Palavras de alta frequência (PA)	Palavras de baixa frequência (PB)
PA 1 CASA	PB 1 ANTA
PA 2 HOJE	PB 2 TORA
PA 3 SETE	PB 3 BULE
PA 4 OBRA	PB 4 TATU
PA 5 REDE	PB 5 NOJO
PA 6 FORA	PB 6 BAFO
PA 7 LOJA	PB 7 GULA
PA 8 PEÇA	PB 8 SELA

Não-palavras silábicas (NS)	Não-palavras soletradas (NL)
NS 1 TILA	NL 1 MRIT
NS 2 FESA	NL 2 NJHT
NS 3 GROL	NL 3 ONRS
NS 4 LOUT	NL 4 CRVL
NS 5 MIOT	NL 5 GRLZ
NS 6 DESA	NL 6 IBCF
NS 7 FIJE	NL 7 DJMP
NS 8 CISE	NL 8 RTVS