

Normas do Tamanho da Categoria para 906 Palavras da Língua Portuguesa

Gerson Américo Janczura¹
Goiara Mendonça de Castilho
Universidade de Brasília
Nelson Rocha de Oliveira
Secretaria de Saúde do Distrito Federal

RESUMO - Normas do Tamanho da Categoria (TC) para 906 palavras da língua portuguesa são apresentadas. TC é igual ao número de associadas diferentes de uma palavra produzidas por, pelo menos, dois participantes. As associadas foram coletadas através da técnica de associação livre considerando-se, somente, a primeira resposta. Setecentos e cinquenta e seis universitários de diferentes instituições de ensino superior e cursos participaram do estudo, sendo que as associadas de cada pista foram produzidas por, pelo menos, 100 participantes. Os resultados mostraram que TC apresenta uma distribuição muito próxima da normalidade e é ortogonal às medidas de concretude, frequência de ocorrência da palavra, valência e alerta emocionais, cujos valores também são apresentados para cada palavra.

Palavras chave: tamanho da categoria de palavras, associação livre, normas

Norms for Word Set Size for 906 Portuguese Words

ABSTRACT - Norms are presented of Word Set Size (WSS) for 906 Portuguese words. WSS corresponds to the number of different associates produced by, at least, two subjects. The free association technique, which considered only the first response, was used to collect the associates. Seven hundred fifty six undergraduates from different schools and majors participated in this study. Associates were produced by at least 100 subjects per cue word. Results indicated that WSS distribution resembles a normal distribution and it is orthogonal to measures of concreteness, word frequency, valence and arousal whose values are presented for each word.

Keywords: word set size, free association, norms

Um dos materiais mais utilizados na pesquisa sobre a cognição humana são palavras. Estímulos verbais são úteis nas investigações sobre memória, categorização, representação mental do conhecimento, raciocínio lógico, percepção, tomada de decisão, atenção e, naturalmente, linguagem. No que tange à memória, por exemplo, a maioria dos mecanismos básicos sobre a consolidação de longo-prazo foi descoberta utilizando-se palavras escritas ou lidas em listas (Barsalou, 1992). Os motivos para essa escolha se associam aos fatos de as palavras exibirem validade ecológica, serem de fácil acesso e manipulação pelo pesquisador, e de a metodologia experimental ser uma ferramenta epistemológica poderosa.

Entre os atributos de diferentes naturezas das palavras — sintáticos, semânticos, perceptuais, emocionais — investigados em estudos experimentais, incluem-se concretude, frequência de ocorrência em materiais escritos, comprimento da palavra, valência emocional, alerta, familiaridade, similaridades fonológica e semântica, modalidade sensorial de apresentação, distintividade. A necessidade de manipular ou controlar as características desses estímulos verbais é amplamente reconhecida na literatura científica e pode ser ilustrada em diferentes paradigmas experimentais. Por exemplo, o desempenho linguístico é superior se as palavras forem mais concretas

(e.g., lápis, maçã) na leitura, compreensão de sentenças e definição de termos (Haberlandt & Graesser, 1985; Sadoski, Kealy, Goetz, & Paivio, 1997; Sadoski & Paivio, 2001). Palavras concretas também produzem melhor reconhecimento e recordação livre ou com pista (Li, Liu, & Du, 2003; Nelson & Schreiber, 1992). Outra característica que também afeta a memória para palavras é a sua frequência de ocorrência na língua. Uma dissociação entre os testes explícitos, reconhecimento e recordação, é observada: quanto maior for a frequência da palavra, maior será a probabilidade de ser recuperada (Hall, 1954; Mandler, Goodman, & Wilkes-Gibbs, 1982; Sumbly, 1963). O oposto é observado no reconhecimento: palavras mais raras na língua são mais provavelmente reconhecidas em uma lista que também contenha palavras com alta frequência (Glanzer & Bowles, 1976; Gorman, 1961).

O uso de palavras na pesquisa tem contribuído para a compreensão da interação da emoção e da cognição humana (e.g., Wilson, MacLeod, & Campbell, 2007). Duas variáveis emocionais que têm sido fatorialmente combinadas em delineamentos experimentais são o alerta (fator que reflete o nível de ativação/excitação produzido pela palavra) e a valência (fator que reflete o quanto uma palavra é percebida em função de sua agradabilidade). Ilustrando: as palavras *sosego* e *morango* estão associadas a um alto nível de agradabilidade e baixo nível de alerta, e as palavras *raiva* e *estupro* estão associadas a um alto nível de alerta e alto

¹ Endereço para correspondência: Universidade de Brasília, Instituto de Psicologia, PPB, Campus Universitário Darcy Ribeiro, Brasília, DF, Brasil. CEP. 70910-900. E-mail: janczura@unb.br

nível de desagradabilidade (Oliveira, Janczura, & Castilho, 2013). A valência e o alerta influenciam diferentes processos cognitivos, como memória (e.g., Kensinger & Corkin, 2003; Marx, Marshall, & Castro, 2008; Stein, 2010), atenção (Aquino & Arnell, 2007; Dresler, Mériaux, Heekeren, & Van der Meer, 2009; Kissler, Herbert, Winkler, & Junghofer, 2009; Pratto & John, 1991), busca visual (Mathewson, Arnell, & Mansfield, 2008; Ohman, Flykt, & Esteves, 2001; Ohman, Lundqvist & Esteves, 2001;), tomada de decisão (Hackley & Valle-Inclan, 1999), decisão lexical (Kousta, Vinson, & Vigliocco, 2009), entre outros.

Os pesquisadores brasileiros dispõem de um número restrito de bases de dados que informam sobre as características das palavras. Atualmente, os seguintes atributos das palavras podem ser consultados em normas coletadas no Brasil: frequência de ocorrência em textos (Linguística Aplicada e Estudos da Linguagem [LAEL], 2006; Pinheiro, 1996; Sardinha, 2004), associadas a categorias naturais (Janczura, 1996, 2005), associadas semânticas em listas (Stein, Feix, & Rohenkohl, 2006; Stein & Gomes, 2009), associadas semânticas (Holderbaum & Salles, 2010; Salles et al., 2008; Salles, Holderbaum, & Machado, 2009; Salles, Machado, & Janczura, 2011); associadas a tríades de letras (Pompéia, Paes, & Bueno, 2003), emocionalidade (Oliveira et al., 2013; Santos, Silveira, Gomes, & Stein, 2009; Stein & Gomes, 2009), concretude (Janczura, Castilho, Rocha, & van Erven, 2007; Stein & Gomes, 2009).

Usualmente, os atributos das palavras são coletados utilizando diferentes procedimentos como a associação livre (e.g., produzir a primeira ou mais palavras que vêm à mente diante de uma pista verbal), julgamentos subjetivos (e.g., avaliar alguma dimensão das palavras utilizando escalas), quantificação de atributos (e.g., cálculo da frequência absoluta de ocorrência das palavras em textos), classificação de atributos (e.g., classificar as palavras em função de categorias gramaticais).

O procedimento de coleta de normas de palavras semanticamente relacionadas coletadas através do método da associação livre exibe alto grau de confiabilidade e informa qual é a probabilidade de um estímulo funcionar como uma pista para outra palavra acessar a consciência do indivíduo, ou seja, a medida corresponde a um índice de acessibilidade do conhecimento semântico relacionado na memória verbal e reflete a experiência lexical partilhada por muitos indivíduos (Nelson, McEvoy, & Schreiber, 2004). Nelson et al. (2004) apontaram vários usos para normas de associação livre na pesquisa psicológica: normas podem prever comportamentos; oferecem um mapa do conhecimento lexical partilhado linguisticamente e culturalmente; as probabilidades geradas pelas normas podem ser tratadas como uma das maneiras que refletem a recuperação do conhecimento na memória; os pesquisadores podem selecionar materiais para a pesquisa considerando o provável conhecimento dos participantes; normas podem ser comparadas com outras normas permitindo análises de proximidade e avaliação de modelos estatísticos; informações normatizadas podem ser usadas para o controle de variáveis, evitando efeitos de confusão; permitem avaliar se indivíduos ou grupos se desviam das normas; podem ser usadas na avaliação implícita dos efeitos do treinamento, psicoterapia e marketing. Algumas contribuições do uso

desse procedimento incluem a investigação da memória implícita (Nelson, McKinney, Gee, & Janczura, 1998), predição do uso de maconha e álcool (Stacy, 1997), percepção de pacientes com dor (Pincus, Pearce, McClelland, Farley, & Vogel, 1994), vieses da memória na depressão (Watkins, Vache, Verney, & Mathews, 1996), e no estudo do raciocínio condicional (Castilho & Janczura, 2012).

O objetivo do presente estudo é ampliar o conhecimento sobre as características das palavras que estão representadas na memória de longo-prazo e que podem ser controladas ou manipuladas experimentalmente, como no caso da frequência de ocorrência, a concretude e a emocionalidade. A característica sobre as palavras disponibilizada neste estudo se denomina Tamanho da Categoria (TC). Pesquisas evidenciaram que, em função da experiência linguística, conexões são estabelecidas entre as palavras (Anderson, 1983; Collins & Loftus, 1975). Nelson, Schreiber, McEvoy (1992) demonstraram que as conexões entre palavras não são idênticas porque variam em número, isto é, algumas palavras estão associadas a um grande número de palavras enquanto outras se associam a um número menor. Isso permite, por exemplo, investigar a influência do número de associadas ativadas implicitamente na memória sem que o indivíduo tenha consciência dessa informação. Underwood (1965, 1983) já havia sugerido essa possibilidade anteriormente. Entretanto, durante várias décadas, a falta de uma metodologia confiável retardou a investigação experimental dos efeitos produzidos pela rede associativas no desempenho humano.

A influência do TC de associadas na memória tem sido relatada experimentalmente desde os anos de 1970. A estratégia metodológica utilizada é comparar grupos de palavras que variam no tamanho da categoria em um paradigma de memória como a recordação com pista extra-lista ou o reconhecimento. Os resultados encontrados, em mais de vinte estudos publicados com vários experimentos, é que as palavras com um menor número de associadas produzem uma recuperação mais acurada e rápida do que aquelas associadas a um grande número de outras palavras. Esse efeito é obtido independentemente da concretude, frequência de ocorrência e ambiguidade das palavras, além de não interagir com processos explícitos de codificação como níveis de processamento, taxa de apresentação e tipo de instrução (para uma revisão, cf. Nelson et al., 1992).

O procedimento adotado neste estudo para mapear o número de associadas consistiu em coletar normas de associação livre, seguindo as sugestões de Nelson, McEvoy e Dennis (2000) e Nelson et al. (1998), e envolveu solicitar a um número grande de participantes ($n = 100-200$) que produzisse a primeira palavra que ocorresse na mente em resposta a uma pista (i.e., palavra). A associação livre é semelhante a várias atividades que desempenhamos diariamente. Exemplos incluem buscar a melhor palavra para verbalizar uma opinião durante um diálogo, pensar em pistas verbais para executar uma busca no Google ou responder a uma pergunta inesperada.

A produção da primeira associada foi aplicada porque o tamanho relativo da categoria é proporcionalmente representado usando-se esse método ou a solicitação de múltiplas respostas (Herrmann, Chaffin, & Corbett, 1973). Além disso, a geração da primeira associação tem a vantagem

de evitar efeitos de encadeamento associativo e a inibição na recuperação (McEvoy & Nelson, 1982). Em contraste, no método de associação livre por múltiplas respostas é possível que a segunda ou posteriores associadas sejam enviadas pela resposta produzida anteriormente, que estaria desempenhando o papel de pista associativa no lugar do alvo pretendido. Segundo Nelson e Schreiber (1992), o TC é computado pelo número de associadas diferentes produzidas por, pelo menos, dois participantes, uma vez que as respostas idiossincrásicas se relacionam muito fracamente à pista e não exibem confiabilidade.

Às normas de TC coletadas para 905 palavras foram agregadas informações sobre outros atributos de cada uma, coletados em diferentes fontes: concretude (Janczura et al., 2007), valência e alerta (Oliveira et al., 2013), e frequência de ocorrência em materiais escritos (Linguística Aplicada e Estudos da Linguagem, 2006). A intenção foi disponibilizar para o pesquisador uma base de dados que viabilizasse a seleção de estímulos em função das respectivas variáveis.

Método

Participantes

Participaram voluntariamente da coleta de normas 756 estudantes universitários, cuja idade média era 24,5 anos ($DP = 7,79$, amplitude = 43 anos) sendo 75% mulheres e 25% homens. Os estudantes eram oriundos de uma instituição de ensino superior (IES) pública (34%) e duas privadas (47% e 19%, respectivamente) da cidade de Brasília. Os participantes tinham cursado entre 1 e 11 semestres letivos ($Média = 3,09$, $DP = 1,92$) sendo 54% da Psicologia, 18% da Enfermagem, 7,7% das Ciências Contábeis, 5% da Medicina e 15,3% distribuídos entre 32 cursos universitários. Metade dos participantes nasceu em Brasília (54,8%) e os demais têm origem em 93 locais distribuídos entre todas as regiões do país. O tempo médio de residência em Brasília era 18,97 anos ($DP = 9,36$, amplitude = 44 anos). Os participantes foram selecionados por conveniência¹.

Materiais

Um conjunto de 905 palavras, sendo 72,3% substantivos, 10,3% adjetivos, 2,6% verbos, 0,2% advérbios e 14,6% pertencentes a duas ou mais categorias gramaticais, foi utilizado. Essas palavras correspondem àquelas utilizadas anteriormente em normas de concretude (Janczura et al., 2007) e emocionalidade (Oliveira et al., 2013). As palavras foram organizadas, aleatoriamente, em nove blocos contendo um mínimo de 100 e um máximo de 101 palavras. Cada bloco tinha três páginas, sendo que cada página apresentava, em média, 34 palavras. A ordem de apresentação das páginas de cada bloco foi determinada randomicamente. Duas colunas foram impressas em cada página: a coluna da esquerda

apresentava as palavras a serem normatizadas e a coluna da direita, linhas pontilhadas onde os participantes escreviam as respostas. Cada bloco era acompanhado de uma folha, inserida antes da lista de palavras, que continha as instruções da tarefa.

Procedimentos

A coleta dos dados foi realizada em grupos nas salas de aula das respectivas instituições de ensino, durante o período de aulas. A tarefa dos participantes era escrever nas linhas pontilhadas, ao lado de cada palavra da lista, a primeira palavra relacionada à pista que viesse à mente. O tempo de execução da tarefa era livre. Previamente à distribuição dos blocos, os participantes foram esclarecidos sobre os objetivos da pesquisa e o caráter voluntário da participação. Após receberem todas as informações, os participantes assinaram o Termo de Consentimento Livre e Esclarecido e receberam o bloco de palavras. A folha de rosto de cada bloco continha as seguintes instruções: “Neste bloco você encontrará várias palavras. A sua tarefa é ler cada palavra silenciosamente e escrever, ao seu lado, na linha pontilhada, a primeira palavra que vem à mente após a leitura. Não pense muito, apenas escreva a primeira palavra que você lembrar. Não existem respostas certas ou erradas nesta tarefa. Escreva somente uma palavra ao lado de cada palavra. Por favor, escreva claramente, mas não se preocupe com a grafia. Muito obrigado pela colaboração!”.

Resultados

O TC foi estimado contando-se o número de respostas diferentes produzidas por, no mínimo, dois participantes. Cada uma das respostas foi avaliada por, pelos menos dois juízes, que observaram os seguintes critérios de codificação: (a) foram tratadas como respostas semelhantes variações de número que não implicassem diferenças semânticas significativas, sendo tabulada a palavra com maior frequência absoluta de ocorrência (e.g., *casa* e *casas* foram codificadas como *casa*, mas *pai* e *pais* foram tratados como respostas diferentes), (b) foram tratadas como respostas semelhantes sinônimos (e.g., *cão* e *cachorro*), sendo tabulada a palavra com maior frequência absoluta de ocorrência; (c) foram tratadas como respostas semelhantes variações de gênero que não implicassem em diferenças semânticas significativas (e.g., *branco* e *branca*; *macio* e *macia*; *fofo* e *fofa*), sendo tabulada a palavra com maior frequência absoluta de ocorrência; (d) erros de grafia foram corrigidos considerando o significado da pista (e.g., a resposta *cansasso* para a pista “cochilar” foi recodificada como *cansaço*); (e) foram codificadas como respostas diferentes variações do tempo verbal (e.g., *dormi* e *dormindo* como associadas da pista “dormir”); (f) foram tratadas como respostas diferentes conjuntos de palavras que denotassem um sintagma diferente (e.g., as respostas *fogo* e *de fogo* para a pista “arma”); (g) respostas que incluíam artigo ou adjetivo foram recodificadas tratando como núcleo do sintagma a outra palavra (e.g., *muito feliz* foi recodificado como *feliz*, *muito*; *o amigo* foi

1 Projeto aprovado pelo Parecer nº 988.985 do Comitê de Ética em Pesquisa/ICH/UnB

recodificado como *amigo*, o); (h) foram considerados erros respostas que correspondiam à própria pista (e.g., produzir *arma* como associada da pista “arma”) ou respostas que não eram palavras. Registrou-se, ainda, como “em branco” a ausência de resposta e como “ilegíveis” respostas cuja grafia não permitia o seu reconhecimento.

As normas podem ser acessadas no repositório de dados de pesquisas do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT/MCT), no endereço <http://repositoriopesquisas.ibict.br/dvn/dv/Janczura>. O número total de informações numéricas disponibilizadas pela base é 13.579, sendo que o pesquisador poderá selecionar estímulos em função dos critérios desejados utilizando filtros que estão associados a cada uma das colunas. Além disso, é possível descarregar a base de dados que se encontra no formato .xlsx (Excel).

A base de dados apresenta as seguintes medidas para cada uma das palavras usadas como pista: TC, frequência absoluta de respostas idiossincrásicas, frequência absoluta de respostas diferentes, frequência absoluta de respostas em branco, frequência absoluta de erros, frequência absoluta de respostas ilegíveis, tamanho da amostra de respondentes para cada palavra (N), número de respostas válidas (i.e., excluindo as respostas em branco, ilegíveis e erros), Concretude e respectivo Desvio Padrão (DP/C) extraídos de Janczura et al. (2007), Valência e respectivo Desvio Padrão (DP/V) e Alerta e respectivo Desvio Padrão (DP/A) extraídos das normas de Oliveira et al. (2013) e Frequência Absoluta de Ocorrência extraída de LAEL (2006).

A Figura 1 mostra que o TC tem uma distribuição semelhante à normalidade (assimetria = -0,59, curtose = 0,037), com uma amplitude de 24 (valor mínimo = 2 e valor máximo = 26) e média igual a 14,22 (DP = 3,786).

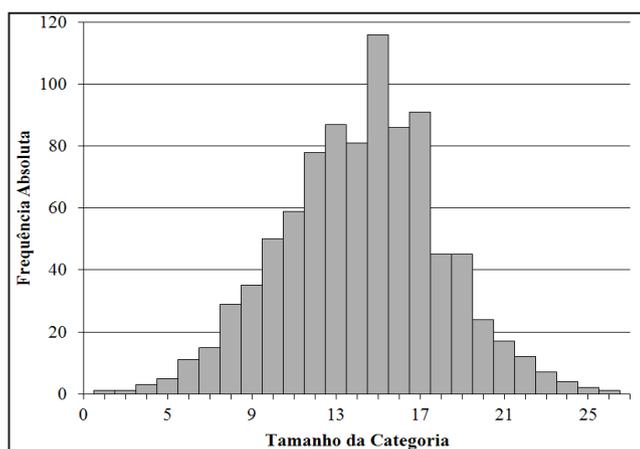


Figura 1. Distribuição de frequência do TC

O número total de respostas coletadas foi 100.295, assim distribuídas: 69,5% correspondentes ao TC, 29% a respostas idiossincrásicas, 0,2% a erros, 1% a respostas em branco e 0,31% das respostas estavam ilegíveis. Quatro palavras foram eliminadas das normas (hecatombe, zombador, paparicar e acalento) porque um número significativo de respostas em branco ($\geq 60\%$) foi registrado.

O número de respondentes por palavras variou de 100 a 177 sujeitos (*Média* = 110,82, *DP* = 6,5). A menor Valência das palavras foi 1,35 (1 = palavras com máxima desagradabilidade) e a maior 8,58 (9 = palavras com máxima agradabilidade), com amplitude = 7,23; o Alerta variou de 1,75 (1 = condição máxima de relaxamento) a 8,15 (9 = condição máxima de estimulação), com amplitude = 6,40; a concretude variou de 1,61 (1 = palavra altamente concreta) a 6,93 (7 = palavra altamente abstrata), com amplitude = 5,32. A frequência variou de 1 a 78.264 (amplitude = 78.263). Esses valores indicam que a amostra incluiu palavras relaxantes, neutras ou alertadoras (fator alerta); palavras negativas, neutras ou positivas emocionalmente (fator valência); palavras julgadas como mais concretas até mais abstratas; e palavras de diferentes frequências de ocorrência na língua portuguesa, ou seja, desde raras até muito frequentes.

Uma análise correlacional entre os atributos das palavras foi realizada a fim de verificar a ortogonalidade entre as medidas. As seguintes correlações de Pearson ($n = 905$) foram observadas entre o TC e os atributos Concretude, Valência, Alerta e Frequência, respectivamente: -0,11 ($p = 0,001$, bicaudal), 0,10 ($p = 0,001$, bicaudal), -0,03 ($p = 0,384$), 0,08 ($p = 0,014$, bicaudal). Apesar de algumas correlações serem significativas, a sua magnitude é muito fraca, explicando, no máximo, 1% da variância. Isso sugere que os fatores são independentes. Quando as palavras são agrupadas nos níveis do TC pequeno ($TC \leq 8$), médio ($9 \leq TC \leq 16$) e grande ($17 \leq TC \leq 34$), seguindo as sugestões de Nelson e Schreiber (1992), Nelson et al. (1992) e Nelson et al. (1998), pode-se comparar os valores médios dos fatores concretude, alerta, valência e frequência para cada nível, conforme se observa na Tabela 1.

A tabela revela que 65,4% das palavras possuem o TC médio, 27,4% são grandes e 7,2% são pequenas. Apesar dessa desproporção, os valores médios dos demais atributos são muito semelhantes em função desse fator. As médias gerais para o total de palavras são: Concretude = 4,47 ($DP = 0,25$), Valência = 5,30 ($DP = 0,24$), Alerta = 4,81 ($DP = 0,09$), Frequência (sem outliers) = 783,08 ($DP = 160,78$). Também são semelhantes as médias referentes às repostas idiossincrásicas ($M = 30,32$, $DP = 5,49$), número de repostas diferentes ($M = 43,23$, $DP = 11,39$), repostas em branco ($M = 0,99$, $DP = 0,26$), erros ($M = 0,26$, $DP = 0,17$) e repostas ilegíveis ($M = 0,29$, $DP = 0,10$).

Tabela 1. Médias dos atributos das palavras em função do TC

Tamanho da Categoria	Atributos				
	Concretude	Valência	Alerta	Frequência	Frequência
Pequeno (n = 65)	4,73	5,11	4,78	786,72	604,69
Médio (n = 592)	4,46	5,24	4,92	1789,31	827,74
Grande (n = 248)	4,22	5,56	4,74	2352,05	916,8

Nota. Frequênciaa = médias sem outliers (n = 94 casos)

A tabela 2 apresenta as correlações bivariadas de Pearson entre o TC e os demais atributos das palavras. Os resultados permitem avaliar se os respectivos fatores estão associados ou são independentes.

Nenhuma das correlações entre os fatores foi significativa, sugerindo que os fatores são ortogonais.

Tabela 2. *Correlações entre o TC e os demais atributos*

Atributo	TC		
	Pequeno	Médio	Grande
Concretude	-0,02 (p = 0,90)	-0,08 (p = 0,06)	-0,06 (p = 0,31)
Valência	0,17 (p = 0,18)	0,07 (p = 0,09)	0,09 (p = 0,18)
Alerta	-0,21 (p = 0,87)	0,004 (p = 0,92)	-0,03 (p = 0,65)
Frequência	0,04 (p = 0,78) ^a	0,06 (p = 0,13)	-0,02 (p = 0,71)

Nota. ^acorrelação sem outliers (2 casos excluídos)

Discussão

As normas associativas relatadas neste estudo expressam um índice relativo do TC, mas não um valor absoluto do número de associadas na memória de longo prazo. Isso significa que os valores assinalados a cada palavra não refletem a base de conhecimento semântico absoluto dos indivíduos sobre os estímulos, mas uma probabilidade (Nelson et al., 2000). Assim, a verificação de que a palavra barco possui 9 associadas e água possui 20 não significa que o conhecimento sobre a primeira palavra seja mais restrito do que o conhecimento sobre a segunda. Esses valores indicam que, quando uma das palavras é pensada, o número de associadas mais provavelmente acessadas na memória é mais amplo para água do que para barco. Mais importante é observar, ainda, que a acessibilidade a esse conhecimento semântico não está associada à frequência de ocorrência das palavras na língua portuguesa, sua valência emocional, o grau de alerta ou a concretude, sugerindo que o TC emerge como um atributo independente das palavras, dirigido por processos subjacentes diferentes dos demais atributos. A ortogonalidade dessas medidas é semelhante às normas de associação semântica coletadas em outros países, assim como as características descritivas de sua distribuição (e.g., Nelson et al., 2004).

Algumas cautelas devem ser tomadas ao interpretarmos normas, segundo Murphy e Davidshofer (1988). A primeira se relaciona à amostra de participantes que, espera-se, representem o tipo de população alvo. Universitários de IES privadas e pública compuseram a respectiva amostra, que apresentou significativa diversidade nos fatores demográficos faixa etária, local de nascimento, curso universitário e respectivo semestre letivo. Apesar deste estudo não incluir outros níveis de escolaridade ou ter controlado o nível socioeconômico e a origem dos participantes, acreditamos que as normas oferecem uma amostra representativa do TC para a população de uma maneira geral, porque se trata de termos utilizados no cotidiano da maioria dos brasileiros; as palavras não expressam conhecimento especializado; respostas idiossincrásicas foram excluídas do cálculo do TC; cada pista foi avaliada por, em média, 100 indivíduos; a amostra de palavras incluiu termos com diferentes

frequências de ocorrência na língua portuguesa; quase metade da amostra de participantes era oriundo de diferentes regiões do país. Ressalte-se, também, que a participação de universitários na coleta de normas tem sido validada em vários estudos em diferentes línguas, como exemplificam as normas na língua inglesa de Battig e Montague (1969), McEvoy e Nelson (1982), Nelson, McEvoy, Walling e

Wheeler (1980), as normas na língua italiana de Barca, Burani e Arduino (2002), as normas em alemão de Ruts et al. (2004), as normas espanholas de Dasí, Soler e Ruiz (2004) e as normas em português lusitano de Marques (1997), além das normas brasileiras citadas anteriormente.

A segunda cautela se associa ao tamanho da amostra de participantes. Assume-se que, se as normas fossem coletadas novamente com os mesmos participantes, os mesmos valores do TC seriam encontrados. Isso significa que os valores são fixos e não sujeitos a erro amostral. Cuidados para aumentar a confiabilidade das normas foram tomados: cada palavra normatizada foi avaliada por um número elevado de participantes, no mínimo 100, e solicitou-se, apenas, a primeira associada. Conforme discutido anteriormente, esses cuidados têm se mostrado eficientes e eficazes na produção de medidas confiáveis de associação semântica.

Por último, cautela deve ser tomada em relação à atualidade e durabilidade das normas. Por quanto tempo essas normas poderão ser utilizadas, ou seja, qual é a estabilidade da medida TC? Um exame na literatura revela que normas não são frequentemente renormatizadas e as razões estão relacionadas ao custo alto desse procedimento e o tempo consumido para executá-lo. A renormatização de parâmetros é urgente quando mudanças aparentes são observadas no fator descrito. Por exemplo, normas que descrevem habilidades cognitivas como raciocínio analítico e tomada de decisões que foram coletadas antes das novas tecnologias (e.g., smartphones, tablets) e do advento da internet. Não há razões para considerar que o conhecimento de termos cotidianos, como as palavras utilizadas nas normas presentes, venha a sofrer mudanças substanciais impelidas por algum evento que tenha impacto social no conhecimento lexical dos indivíduos. Caso aconteça, o efeito será, muito provavelmente, associado a alguma palavra em particular que o pesquisador poderá prontamente eliminar.

Referências

- Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Aquino, J. M., & Arnell, K. M. (2007). Attention and the processing of emotional words: Dissociating effects of arousal. *Psychonomic Bulletin & Review*, *14*(3), 430-435. doi: 10.3758/BF03194084
- Barca, L., Burani, C., & Arduino, L. S. (2002). Word naming times and psycholinguistic norms for Italian nouns. *Behavior Research Methods, Instruments, & Computers*, *34*(3), 424-434. doi: 10.3758/BF03195471
- Barsalou, L. W. (1992). *Cognitive Psychology: An overview for cognitive scientists*. Hillsdale, New Jersey: LEA.
- Battig, W. F., & Montague, W. E. (1969). Category norms of verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology*, *80*(3), 1-46.
- Castilho, G. M., & Janczura, G. A. (2012). Raciocínio condicional: A conclusão depende do conhecimento armazenado na memória. *Estudos de Psicologia*, *17*(1), 53-61.
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407-428. <http://dx.doi.org/10.1037/0033-295X.82.6.407>
- Dasí, C., Soler, M. J., & Ruiz, J. C. (2004). Normative data on the familiarity and difficulty of 196 Spanish word fragments. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 559-563. doi: 10.3758/BF03195602
- Dresler, T., Mériaux, K., Heekeren, H. R., & Van der Meer, E. (2009). Emotional Stroop task: Effect of word arousal and subject anxiety on emotional interference. *Psychological Research PRPF*, *73*(3), 364-371. doi: 10.1007/s00426-008-0154-6
- Glanzer, M., & Bowles, N. (1976). Analysis of the word-frequency effect in recognition memory. *Journal of Experimental Psychology: Human Learning and Memory*, *2*(1), 21-31. <http://dx.doi.org.ez54.periodicos.capes.gov.br/10.1037/0278-7393.2.1.21>
- Gorman, A. M. (1961). Recognition memory for nouns as a function of abstractness and frequency. *Journal of Experimental Psychology*, *61*(1), 23-29. <http://dx.doi.org/10.1037/h0040561>
- Haberlandt, K. F., & Graesser, A. C. (1985). Component processes in text comprehension and some of their interactions. *Journal of Experimental Psychology: General*, *114*(3), 357-375. <http://dx.doi.org/10.1037/0096-3445.114.3.357>
- Hackley, S. A., & Valle-Inclan, F. (1999). Accessory stimulus effects on selection: Does arousal speed decision making? *Journal of Cognitive Neuroscience*, *11*(3), 321-329. doi:10.1162/089892999563427
- Hall, J. F. (1954). Learning as a function of word frequency. *The American Journal of Psychology*, *67*(1), 138-140. doi: 10.2307/1418080
- Herrmann, D. J., Chaffin, R. J. S., & Corbett, A. T. (1973). A factor analysis of six measures extracted from the Battig and Montague (1969) category norms: Evidence for three properties of categories in semantic memory. *Journal of Verbal Learning and Verbal Behavior*, *12*, 666-674. doi:10.1016/S0022-5371(73)80046-5
- Holderbaum, C. S., & Salles, J. F. D. (2010). Priming semântico em crianças: Efeitos da força de associação semântica e frequência do alvo. *Aletheia*, *33*, 95-108.
- Janczura, G. A. (1996). Normas associativas para 69 categorias semânticas. *Psicologia: Teoria e Pesquisa*, *12*(3), 237-244.
- Janczura, G. A. (2005). Contexto e normas de associação para palavras: A redução do campo semântico. *Paidéia*, *15*(32), 417-425.
- Janczura, G. A., Castilho, G. M., Rocha, N. O., & van Erven, T. J. C. (2007). Normas de concretude para 909 palavras da língua portuguesa. *Psicologia: Teoria e Pesquisa*, *23*(2), 195-204.
- Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition*, *31*(8), 1169-1180. doi: 10.3758/BF03195800
- Kissler, J., Herbert, C., Winkler, I., & Junghofer, M. (2009). Emotion and attention in visual word processing—An ERP study. *Biological Psychology*, *80*(1), 75-83. doi:10.1016/j.biopsycho.2008.03.004
- Kousta, S. T., Vinson, D. P., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition*, *112*(3), 473-481. doi:10.1016/j.cognition.2009.06.007
- Linguística Aplicada e Estudos da Linguagem. (2006). *Banco de Português, sub-corpora escrito*. Retrieved from <http://ww2.lael.pucsp.br/corpora/bp/index.htm>.
- Li, Y., Liu, Y., & Du, X. (2003). The effect of word concreteness on implicit and explicit memory. *Psychological Science*, *26*(5), 823-828.
- Mandler, G., Goodman, G. O., & Wilkes-Gibbs, D. L. (1982). The word-frequency paradox in recognition. *Memory & Cognition*, *10*(1), 33-42. doi: 10.3758/BF03197623
- Marques, J. F. (1997). Normas de tipicidade e familiaridade para diferentes categorias de itens verbais. *Revista Portuguesa de Psicologia*, *32*, 35-55.
- Marx, B. P., Marshall, P. J., & Castro, F. (2008). The moderating effects of stimulus valence and arousal on memory suppression. *Emotion*, *8*(2), 199-207. doi: 10.1037/1528-3542.8.2.199
- Mathewson, K. J., Arnell, K. M., & Mansfield, C. A. (2008). Capturing and holding attention: The impact of emotional words in rapid serial visual presentation. *Memory & Cognition*, *36*(1), 182-200. doi: 10.3758/MC.36.1.182
- McEvoy, C. L., & Nelson, D. L. (1982). Category norms and instance norms for 106 categories of various sizes. *American Journal of Psychology*, *95*(4), 581-634.
- Murphy, K. R., & Davidshofer, C. O. (1988). *Psychological testing: Principles, and applications*. New Jersey: Prentice Hall.
- Nelson, D. L., Schreiber, T. A., & McEvoy, C. L. (1992). Processing implicit and explicit representations. *Psychological Review*, *99*(2), 322-348. <http://dx.doi.org/10.1037/0033-295X.99.2.322>
- Nelson, D. L., McEvoy, C. L., & Dennis, S. (2000). What is and what does free association measure. *Memory and Cognition*, *28*(6), 887-899.
- Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 402-407. doi: 10.3758/BF03195588
- Nelson, D. L., McEvoy, C. L., Walling, J. R., & Wheeler, J. W. (1980). The university of South Florida homograph norms. *Behavior Research Methods & Instrumentation*, *12*(1), 16-37. doi: 10.3758/BF03208320

- Nelson, D. L., McKinney, V. M., Gee, N., & Janczura, G. A. (1998). Interpreting the influence of implicitly activated memories on recall and recognition. *Psychological Review*, *105*(2), 299-324. <http://dx.doi.org/10.1037/0033-295X.105.2.299>
- Nelson, D.L., & Schreiber, T.A. (1992). Word concreteness and word structure as independent determinants of recall. *Journal of Memory and Language*, *31*, 237-260. doi:10.1016/0749-596X(92)90013-N
- Ohman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General*, *130*, 466-478. doi: 10.1037/0096-3445.130.3.466
- Ohman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: A threat advantage with schematic stimuli. *Journal of Personality and Social Psychology*, *80*, 381-396. doi:10.1037//0022-3514.80.3.381
- Oliveira, N. R., Janczura, G. A., & Castilho, G. M. (2013). Normas de alerta e valência para 908 palavras da língua portuguesa. *Psicologia: Teoria e Pesquisa*, *29*(2), 185-200.
- Pincus, T., Pearce, S., McClelland, A., Farley, S., & Vogel, S. (1994). Interpretation bias in responses to ambiguous cues in pain patients. *Journal of Psychosomatic Research*, *38* (4), 347-353. doi:10.1016/0022-3999(94)90039-6
- Pinheiro, A.M.V. (1996). *Contagem de frequência de ocorrência de palavras expostas a crianças na faixa pré-escolar e séries iniciais do 1o. grau*. São Paulo: Associação Brasileira de Dislexia.
- Pompéia, S., Paes, A.T., & Bueno, O.F.A. (2003). Teste de complementação de letras: Estímulos para uso no Brasil. *Psicologia: Teoria e Pesquisa*, *19*(1), 65-73.
- Pratto, F., & John, O. P. (1991). Automatic vigilance: The attention grabbing power of negative social information. *Journal of Personality and Social Psychology*, *61*(3), 380-391. <http://dx.doi.org/10.1037/0022-3514.78.6.1024>
- Ruts, W., De Deyne, S., Ameel, E., Vanpaemel, W., Verbeemen, T., & Storms, G. (2004). Dutch norm data for 13 semantic categories and 338 exemplars. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 506-515. doi: 10.3758/BF03195597
- Sadoski, M. & Paivio, A. (2001). *Imagery and text: A dual coding theory of reading and writing*. Mahwah, New Jersey: Lawrence Erlbaum.
- Sadoski, M., Kealy, W. A., Goetz, E. T., & Paivio, A. (1997). Concreteness and imagery effects in the written composition of definitions. *Journal of Educational Psychology*, *89*, 518-526. doi: 10.1037/0022-0663.89.3.518
- Salles, J. F., Holderbaum, C. S., Becker, N., Rodrigues, J. D., Liedtke, F. V., Zibetti, M. R., & Piccoli, L. F. (2008). Normas de associação semântica para 88 palavras do português brasileiro. *Psico*, *39*(3), 362-370.
- Salles, J. F. D., Holderbaum, C. S., & Machado, L. L. (2009). Normas de associação semântica de 50 palavras do português brasileiro para crianças: Tipo, força de associação e set size. *Interamerican Journal of Psychology*, *43*(1), 57-67.
- Salles, J. F., Machado, L. L., & Janczura, G. A. (2011). Efeitos de priming semântico em tarefa de decisão lexical em crianças de 3ª. série. *Psicologia: Reflexão e Crítica*, *24*(3), 597-608.
- Santos, R. F., Silveira, R. A. T., Gomes, C. F. A., & Stein, L. M. (2009). Normas de emocionalidade para a versão brasileira do paradigma Desse-Roediger-McDermott (DRM). *Psicologia: Teoria e Pesquisa*, *25*(3), 387-394.
- Sardinha, T.B. (2004). *Linguística de Corpus*. Barueri: Manole.
- Stacy, A. W. (1997). Memory activation and expectancy as prospective predictors of alcohol and marijuana use. *Journal of Abnormal Psychology*, *106*(1), 61-73. doi: 10.1037/0021-843X.106.1.61
- Stein, L. M. (2010). *Falsas memórias*. Porto Alegre: Artmed.
- Stein, L. M., & Gomes, C. F. A. (2009). Normas brasileiras para listas de palavras associadas: Associação semântica, concretude, frequência e emocionalidade. *Psicologia: Teoria e Pesquisa*, *25*(4), 537-546.
- Stein, L. M., Feix, L. F., & Rohenkohl, G. (2006). Avanços metodológicos no estudo das falsas memórias: Construção e normatização do procedimento de palavras associadas. *Psicologia: Reflexão e Crítica*, *19*(2), 166-176.
- Sumby, W. H. (1963). Word frequency and serial position effect. *Journal of Verbal Learning and Verbal Behavior*, *1*(6), 443-450. doi:10.1016/S0022-5371(63)80030-4
- Underwood, B. J. (1965). False recognition produced by implicit verbal responses. *Journal of Experimental Psychology*, *70*(1), 122-129. doi: org/10.1037/h0022014.
- Underwood, B. J. (1983). *Attributes of memory*. Glenview, IL: Scott, Foresman.
- Watkins, P. C., Vache, K., Verney, S. P., & Mathews, A. (1996). Unconscious mood-congruent memory bias in depression. *Journal of Abnormal Psychology*, *105*(1), 34-41. <http://dx.doi.org/10.1037/0021-843X.105.1.34>
- Wilson, E., MacLeod, C., & Campbell, L. (2007). The information-processing approach to emotional research. In J. A. Coan & J. J. B. Allen (Ed.), *Handbook of emotion elicitation and assessment* (pp.184-202). New York: Oxford University Press.

Recebido em 14.07.2015

Primeira decisão editorial em 13.06.2016

Versão final em 03.07.2016

Aceito em 19.07.2016 ■