CoDAS

Este é um artigo publicado em acesso aberto (Open Access) sob a licença Creative Commons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado. Fonte: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S2317-17822016000500497&lng=pt&tlng=pt. Acesso em: 8 dez. 2017.

REFERÊNCIA

CALAIS, Lucila Leal et al. A study on the semantic association of Brazilian Portuguese words. **CoDAS**, São Paulo, v. 28, n. 5, p. 497-506, set./out. 2016. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S2317-17822016000500497&lng=pt&nrm=iso>. Acesso em: 8 dez. 2017. Epub 31-Out-2016. doi: http://dx.doi.org/10.1590/2317-1782/20162015268.

DOI: 10.1590/2317-1782/20162015268

Original Article

Artigo Original

A study on the semantic association of Brazilian Portuguese words

Um estudo sobre associação semântica de palavras do português brasileiro

Lucila Leal Calais[1]
Aveliny Mantovan Lima-Gregio[2]
Pablo Arantes[3]
Daniela Gil[4]
Alda Christina Lopes de Carvalho Borges[4]

Keywords

Linguistics
Language
Cognition
Speech Recognition Tests

ABSTRACT

Purpose: To collect the semantic association norms of 96 Brazilian Portuguese words for further application in the formulation of a speech recognition test using sentences with controlled word predictability. **Methods:** Study participants were 67 volunteers aged 18 to 27 years. A semantic association task with word class delimitation was used. **Results:** The mean sizes of the total and meaning sets were larger in the second recall to both verb and noun classes. The prevalent semantic association strength in the first recall was strong to the first word and moderate to the second word. In the second recall, the prevalent semantic association strength was moderate to both the first and second words. Significant negative correlation was observed between association strength and total and meaning set sizes. **Conclusion:** The semantic association norms of 96 Brazilian Portuguese words were collected as proposed and resulted in semantic associates for each target word in the verb and noun classes that were used in a speech recognition assessment considering sentences with controlled word predictability.

Descritores

Linguística
Linguagem
Cognição
Testes de Discriminação de Fala

RESUMO

Objetivo: coletar normas de associação semântica de 96 palavras do português brasileiro e aplicá-las, posteriormente, na elaboração de um teste de reconhecimento de fala com frases com controle da previsibilidade da palavra. **Método:** participaram 67 voluntários com idades entre 18 e 27 anos. Foi aplicada uma tarefa de associação semântica com delimitação de classe de palavras. **Resultados:** o tamanho médio do conjunto total e do significativo foi maior para a segunda evocação para as classes de verbos e substantivos. A força de associação semântica que prevaleceu na primeira recordação foi forte para a primeira e média para a segunda palavra. Na segunda recordação, a força de associação com maior prevalência foi média para a primeira e para a segunda palavra. Foi observada correlação negativa significativa entre a força de associação e os tamanhos total e significativo do conjunto. **Conclusão:** Normas de associação semântica de 96 palavras do português brasileiro foram coletadas conforme proposto e resultaram em palavras associadas semanticamente para cada alvo nas categorias de verbo e substantivo, as quais foram utilizadas em um teste de reconhecimento de fala com frases considerando a previsibilidade da palavra.

Correspondence address:
Lucila Leal Calais
Universidade Federal de São Paulo – UNIFESP
Rua Botucatu, 802, Vila Clementina, São Paulo (SP), Brazil,
CEP: 04023-062.
E-mail: calais@ig.com.br

Received: April 01, 2015

Accepted: December 25, 2015

Study carried out at Programa de Pós-graduação em Distúrbios da Comunicação Humana – PPGDCH, Departamento de Fonoaudiologia, Universidade Federal de São Paulo – UNIFESP - São Paulo (SP), Brazil.

[1]Prefeitura do Município de Piracicaba - Piracicaba (SP), Brazil.

[2]Universidade de Brasília – UnB - Brasília (DF), Brazil.

[3]Universidade Federal de São Carlos – UFSCar - São Carlos (SP), Brazil.

[4]Universidade Federal de São Paulo – UNIFESP - São Paulo (SP), Brazil.

Financial support: nothing to declare.

Conflict of interests: nothing to declare.

INTRODUCTION

The use of words in studies shows the need for a careful selection of these stimuli because their characteristics can generate undesirable effects of confusion. The necessity to avoid possible secondary influences demands knowledge on the properties of words, with collection of norms being frequently used to identify such attributes[1].

Semantics is among the main characteristics of words. Studies on semantic association norms involving research on language, learning, and memory have been conducted in other countries for over a century[1,2]. In Brazil, these studies are more recent, since the 1990s, and they are important to the extent that importing results from studies in other languages is not advisable due to lack of reliability[3].

The search for semantically associated words allows estimation of the set of semantic associates owing to some property or linguistic relationship, providing a way to assess the knowledge on words that individuals have acquired throughout their learning history[1].

Studies have focused on the semantic association of words and have correlated it with various aspects, such as semantic categories[3,4], syntagmatic and paradigmatic properties[5], development and/or aging processes[6-9], semantic context[1], semantic dementia[10], linguistic aspects as concreteness and frequency of words[2,11,12], memory[11,13], etc.

Free word association tasks have been widely used in the development of semantic association norms. They consist in presenting a target word to participants and asking them to produce a related word that comes to mind[2]. Two different tasks can be used: single response, in which individuals are required to produce only one response for each target word[1-3,6-9] and multiple response, in which individuals produce two or more responses in the sequence they come to mind[10,12,13].

Continuous association, or multiple response, has been criticized because it is subject to both response chaining, when the second word is generated as a function of the first word, not the target word, and retrieval inhibition, when the first word inhibits the production of other words[3,11]. The effect of retrieval inhibition can be mitigated when a limited number of responses are required. Another reduction of this effect can be obtained by increasing the variation of stimuli[5].

Although there are good reasons to restrict the collection of associates to one response per target word, there are cases in which a strong association may occur first, e.g., *sangue* and *vermelho* (blood and red), reducing the number of different associates and preventing sentences with more than one different semantic associate related to the same word from being produced. In this context, continuous responses enable greater variability of associations, in addition to better estimate the probability of weak associations[5].

Strength of association between pairs of associates is based on the number of concordant responses produced for the same target word. This strength is classified as weak (<10%), moderate (from 10 to 24%), or strong (>25%)[2,7,11]. It is directly related to set size, which can be total (all different responses are considered) or meaning (only responses generated by two or more participants are considered)[11,12].

Special care should be taken during the search for semantic associates regarding the acceptance of words derived from the target word, as in "*trabalho*" - "*trabalhador*" (work - worker), as well as with proper nouns, because of set size implications and structural similarity between the target word and other semantically associated words[2,7]. Considering the number of associates in the preparation of sentences to be recognized, the structural similarity between words could inappropriately advantage recognition. For instance, in the sentence "*O sorveteiro vendeu um sorvete*" (The ice cream man sold some ice cream/an ice cream stick), recognition of the word "*sorvete*" (ice cream) is favored by the word "*sorveteiro*" (ice cream man").

Caution should also be taken in the search for semantic associates with respect to counting. Generated words which vary in gender, e.g., "*avó*" and "*avó*" (grandfather and grandmother) or number, e.g., "*casa*" and "*casas*" (house and houses) and do not present clear semantic difference between items are grouped under the category of high frequency words[1-3,6,8,9].

The search for word association norms is an indispensable tool in the research on memory, representation and learning of concepts, influence of aging on learning and cognition, etc[1].

In this context, studies involving speech recognition tests using context sentences should consider the semantic association of words in the preparation of the material. Increased control over a particular property allows the researcher a more careful selection of the stimulus[1] and a less biased collection of the information desired.

In the national literature, there are several studies on the preparation of speech recognition tests using sentences, but only one survey addressing the property of word semantic association in the formulation of sentences was found[14]. In the present study[14], a speech recognition assessment considering sentences with controlled word predictability was prepared. Study participants were required to repeat only the last word of the sentence. The influence of type of sentence on this recognition was analyzed and high and low predictabilities were found.

Therefore, the objective of this study was to collect the semantic association norms of 96 Brazilian Portuguese words for further application in a speech recognition assessment using sentences with controlled word predictability[14].

METHODS

This study was conducted in the graduate study program of the discipline of Hearing Disorders, Department of Speech-language Pathology, Federal University of Sao Paulo - UNIFESP. The study project was approved by the Research Ethics Committee of the aforementioned institution under no. 0948/09. All participants signed an Informed Consent Form.

A list of 96 Brazilian Portuguese words belonging to the concrete category was used for the application of the semantic association task. The choice of words for the preparation of this list was based on two studies addressing word concreteness[14,15]. All selected words belonged to the noun class, were disyllabic, paroxytone, and presented high occurrence in Brazilian Portuguese

(more than 50 occurrences per million; a threshold established according to values commonly used in the literature[16]).

Study participants were 67 volunteers: 66 females (98.5%) and one male (1.5%) aged between 18 and 27 years (mean age of 20.56 years). All volunteers were enrolled at the Speech-language Pathology undergraduate course of a public university in 2009, distributed as follows: 28 freshmen (41.8%), 22 sophomores (32.8%), 6 juniors (8.9%), and 11 seniors (16.4%).

The words were arranged in a single version of the instrument which comprised 16 pages. On the first page, there was a header to be filled with the participant's identification data (age, gender, and course year), followed by the list of words. Each page was divided into six parts and each part contained a word in a caption. Under each caption, there were two numbered lines (1 and 2).

Participants were instructed to take the activity individually and in silence, with no time limit. A semantic association task with word class delimitation was used. Nouns were used with approximately half of the participants (33), whereas verbs associated with each of the words of the instrument were used with 34 participants. This option was adopted in order to collect words from these two classes for further use in the formulation of sentences for a speech recognition assessment. This way, the risk of generating words from only one grammatical class was avoided, which would hinder the development of sentences including the essential terms (subject, verb, and predicate). Two responses were requested for each target word.

The task was collectively applied in a single session for a total of four groups, depending on the availability of the participants. The following instruction was provided: "You have received paper sheets divided into six parts. Each of these parts contains a written word. You are requested to write the first two words with relational meaning that come to mind in the corresponding lines for each word." After that, an example for the noun class and one for the verb class were provided, e.g., the word "*praia*" (beach) associated with the noun "*mar*" (sea) and the verb "*nadar*" (swim). It was underscored that words derived from the target word, such as "*trabalho*" (work) and "*trabalhador*" (worker), as well as proper nouns would not be accepted.

The two words most frequently used for each target word were computed for both the first and second written words. The generated words that varied in gender, e.g., "*menino*" (boy) and "*menina*" (girl) or in number, e.g., "*asa*" (wing) and "*asas*" (wings), and that did not present clear semantic difference between the items were grouped under the category of high frequency words.

Set size was analyzed in two ways: first, the total set size (TSS) was calculated following a criterion similar to those proposed in other studies[11-13]. All different responses (words and abbreviations) were considered in the TSS, including idiosyncratic words and incorrect words (words of a class different from that requested, words derived from the target word, and nonexistent words). Regarding the idiosyncratic words, part of the studies[2,11] considered these words in the counting, but part of them[1,13] did not. Subsequently, the meaning set size (MSS) was calculated, in which only responses generated by two or more participants were considered. In the analysis of the meaning set size, the semantic associates were classified as small (1 to 8 associations), medium (9 to 16 associations), and large (17 to 34 associations). Meaning set analyses were performed similarly to the classification adopted in previous studies[2,11].

Semantic association strength between pairs was calculated by the percentage of occurrence of the semantic associate of highest frequency for each target word[2,7]. Words generated with agreement among the participants <10% were considered weak, from 10 to 24% were regarded as moderate, and ≥25% were considered strong.

Two other analyses were also conducted: correlation between semantic association strength and sizes of total and meaning sets and comparison between strength levels.

RESULTS

Participants completed the task in 28.5 minutes on average, with a range of 19 minutes between the shortest and longest times.

To calculate the total set size, the Student's t-test for paired samples was used to compare mean values between the first and second recall of words from the verb and noun classes. In the verb class, mean values of 13.2 (SD=5.5) and 20.7 (SD=4.5) were obtained for the first and second words, respectively, with significant difference between the values (p<0.001). In the noun class, mean values for the first and second words were 14.9 (SD=4.6) and 20.2 (SD=3.2), respectively, also with significant difference between the values (p<0.001).

The meaning set size was calculated likewise, by comparing the mean values obtained for the first and second recall of word classes. In the verb class, mean values of 4.5 (SD=1.6) and 5.9 (SD=1.8) were obtained for the first and second words, respectively, with significant difference between the values (p<0.001). In the noun class, mean values for the first and second words were 4.8 (SD=1.8) and 6.0 (SD=1.4), respectively, also with significant difference between the values (p<0.001).

Regarding the analysis of the meaning set of the 96 target words in the verb class, 100% of the sets of semantic associates were classified as small for the first recall – Verb 1 (V1), whereas 94.8% of the sets were classified as small and 5.2% as medium for the second recall – Verb 2 (V2). As for analysis of the meaning sets in the noun class, 97.9% of the sets of associates were classified as small and 2.1% as medium for the first recall – Noun 1 (N1), whereas 94.8% of the sets were classified as small and 5.2% as medium for the second recall – Noun 2 (N2).

Table 1 shows the quantitative analysis of semantic association strength between the target word and the two most frequently associated words for the first and second recalls, considering all target words for the verb and noun classes.

It is possible to observe that, for the first recall (V1 and N1), the prevalent semantic association strength was strong, followed by moderate and weak to the first word, whereas it was mainly moderate, followed by weak and strong to the second word. For the second recall (V2 and N2), the prevalent semantic association strength was moderate, followed by strong and weak to the first word, whereas it was seldom strong to the second word.

Table 2 shows the correlation (Pearson's correlation coefficient) between semantic association strength and total and meaning

set sizes for the verb and noun classes. Significant negative correlations ($p<0.05$) can be observed for both classes. For verbs, there was strong ($r>0.5$) correlation[17] with the first word of V1 and with both words of V2 in the total set size and only with the first word of V1 in the meaning set size. Correlation with the second word of V1 in the total set size was considered moderate ($0.3>r<0.49$) and there was weak ($0.10>r<0.29$) correlation with the other words of the meaning set size. For nouns, there was strong correlation with the first word of N1 and with both words of N2 in the total set size and only with the first word of S1 in the meaning set size. Correlation with the second word of N1 in the total set size was considered weak, as well as with the second word of N1 and N2 in the meaning set size. Correlation with the first word of N2 in the meaning set size was considered moderate.

The following results were obtained for the verb class in the comparison (ANOVA) between semantic association strength and set sizes: significant difference ($p<0.05$) was observed for the sizes of the total and meaning sets for the two words in the first and second recalls, except for the second recall to the second word in the meaning set. The set sizes decreased as the association strength increased, but this pattern was constant only for the total set size. The following results were obtained for the noun class in the comparison (ANOVA) between semantic association strength and set sizes: significant difference ($p<0.005$) was observed for the sizes of the total and meaning sets for the two words in the first and second recalls, except for the first and second recalls to the second word in the meaning set. Similarly to what occurred in the verb class, the set sizes decreased as the association strength increased, and this pattern was constant for the total set size.

All of the words in this assessment generated at least two associates for the verb and noun groups. It is worth noting that, in several cases, more than two associates were generated because they presented similar values at the highest level of association strength. Table 3 shows the high-frequency semantic associates of the 96 target words for verbs and nouns.

DISCUSSION

Semantic association norms of 96 Brazilian Portuguese words were collected in the present study. Word class delimitation for the semantic associate and the writing of two semantically associated words for each target word were the procedures employed. The study aim was to generate a varied number of associates that would enable the formulation of distinct sentences from the same target word.

Among the results obtained with the semantic associates, it was possible to observe that the total set size presented larger mean values for the second recall compared with the first recall. The greater variation between the responses presented by participants to the second word was equally observed in both classes (verbs and nouns). Regarding the meaning set size, a behavior similar to that previously described was observed, with increased number of associates to the second evoked word, but with smaller mean values. A possible explanation for this increase in the sets in the second recall would be the effect of the creation of a semantic chain, or response chaining[3,11,18]. According to this hypothesis, it is possible that the second word be a semantic associate of the first recalled word, not of the target word. Thus the set of associates tends to increase considering that all individuals recall the same word in the first recall.

The size of the total and/or meaning sets has been presented in previous studies[2,3,5,7,9,12]; however, the means of evoked words for both the total set [2,7,9] and the mean set[2,5,7,9,12] were considerably larger. Only one study[3] reported a mean value for the total set

Table 1. Percentage of semantic association strength levels for the first and second recalls to verbs and nouns

Semantic association strength	V 1		V 2		N 1		N 2	
	1st word	2nd word	1st word	2nd word	1st word	2nd word	1st word	2nd word
Weak	1.0	28.1	10.4	45.8	3.1	26.0	9.4	47.9
Moderate	19.8	62.5	75	54.2	39.6	69.8	78.1	51.0
Strong	79.2	9.4	14.6	0.0	57.3	4.2	12.5	1.0

V1 = 1st verb recall; V2 = 2nd verb recall; N1 =1st noun recall; N2 = 2nd noun recall

Table 2. Correlation between semantic association strength and set sizes for verbs and nouns

Semantic association strength	Verb			
	TSS		MSS	
	r	p	r	p
Verb 1 - 1st word	–0.61	<0.001*	–0.53	<0.001*
Verb 1 - 2nd word	–0.31	0.002*	–0.15	0.152
Verb 2 - 1st word	–0.61	<0.001*	–0.22	0.032*
Verb 2 - 2nd word	–0.57	<0.001*	–0.12	0.227
Noun 1 - 1st word	–0.71	<0.001*	–0.60	<0.001*
Noun 1 - 2nd word	–0.29	0.004*	–0.05	0.610
Noun 2 - 1st word	–0.60	<0.001*	–0.33	0.001*
Noun 2 - 2nd word	–0.58	<0.001*	–0.15	0.135

*Statistically significant result (p<0.05); r = Pearson's correlation coefficient
Caption: TSS = Total set size; MSS = Meaning set size

Table 3. Target words with high-frequency semantic associates and set sizes for verbs and nouns

Target words	Semantic associates* V1 and V2	Semantic associates* N1 and N2
Costa (shore/back)	V1-nadar/andar, viajar, navegar (swim/walk, travel, navigate) V2-coçar/nadar, admirar, mergulhar (scratch/swim, admire, dive)	N1-praia/corpo (beach/body) N2-mar/massagem (sea/massage)
Alvo (target)	V1-atirar/acertar (shoot/hit) V2-acertar/mirar (hit/aim)	N1-tiro, flecha (shot, arrow) N2-dardo/tiro (dart/shot)
Grupo (group)	V1-conversar/juntar (talk/gather) V2-trabalhar/estudar (work/study)	N1-pessoa/amigos (person/friend) N2-trabalho/amigo, família, pessoas, união (work/friend, family, people, union)
Ponto (dot/stop)	V1-marcar/finalizar (mark/score/end) V2-terminar/escrever, desenhar, parar, acabar, chegar (finish/write, draw, end, arrive)	N1-final/vírgula (period/comma) N2-ônibus/final (bus/end)
Peça (piece/play)	V1-montar/quebrar, assistir (assemble/break, watch) V2-assistir/montar (watch/assemble)	N1-teatro/quebra-cabeça (theater/break, puzzle) N2-carro, quebra-cabeça, teatro (car, puzzle, theater)
Luta (fight)	V1-brigar/bater (fight/beat) V2-ganhar, machucar/bater (win, hurt/beat)	N1-briga/luva (fight/glove) N2-briga/machucado (fight/injured)
Mata (woods)	V1-correr/plantar, acampar, caminhar (run/plant, camp, walk) V2-cortar/plantar (cut/plant)	N1-árvore, floresta (tree, forest) N2-árvore/animais (tree/animals)
Pena (feather/fountain pen/pity)	V1-voar/sentir (fly/feel) V2-cair/escrever (fall/write)	N1-galinha/pássaro (chicken/bird) N2-dó/ave (pity/bird)
Reino (kingdom)	V1-governar/morar, mandar (rule/live, order) V2-mandar/morar, batalhar (order/live, battle)	N1-animal, rei (animal, king) N2-rei/castelo, rainha (king/castle, queen)
Obra (construction)	V1-construir/trabalhar (build/work) V2-construir/criar (build/create)	N1-pedreiro/arte (mason/art) N2-quadro/construção, tijolo (picture/construction, brick)
Cerca (fence)	V1-pular/fazer (jump/make) V2-separar/proteger, construir (divide/protect, build)	N1-madeira/arame, fazenda (wood/barbwire, farm) N2-fazenda/arame, madeira (farm/barbwire, wood)
Posto (gas station/clinic/position)	V1-abastecer/parar (fuel/stop) V2-parar/cuidar, abastecer, pagar, vacinar (stop/care, fuel, pay, vaccinate)	N1-gasolina (gasoline) N2-carro/álcool (car/ethanol)
Jogo (game)	V1-brincar/ganhar (play/win) V2-ganhar/divertir, torcer, participar, competir (win/amuse, cheer for, play, compete)	N1-futebol/bola (soccer/ball) N2-bola, tabuleiro (ball, board)
Pista (track/floor)	V1-correr/dançar, andar (race/dance, walk) V2-andar/dirigir (walk/drive)	N1-carro/corrida (car/race) N2-carro/corrida (car/race)
Povo (people)	V1-juntar, unir, falar (gather, join, talk) V2-unir, lutar, trabalhar, criticar (gather, argue, work, criticize)	N1-pessoas/gente (persons/people) N2-multidão/pessoas (crowd/people)
Vôlei (volleyball)	V1-jogar/sacar (play/serve) V2-sacar/ganhar (serve/win)	N1-bola/esporte (ball/sport) N2-bola, rede (ball, net)
Língua (tongue/language)	V1-falar/beijar (speak/kiss) V2-comer/falar (eat/speak)	N1-boca/português (mouth/Portuguese) N2-dente/boca, fala (tooth/mouth, speech)
Mundo (world)	V1-girar/viajar, viver (revolve/travel, live) V2-conhecer/girar (know/revolve)	N1-água, globo, terra (water, globe, Earth) N2-pessoas, terra (people, Earth)
Chefe (boss)	V1-mandar/comandar (demand/command) V2-trabalhar/obedecer, mandar (work/obey, demand/order)	N1-trabalho/emprego (work/job) N2-trabalho/cargo, cozinha, dinheiro, empregado, escritório, salário (work/position/function, kitchen, money, employee, office, salary/wages)
Cara (face/head/person)	V1-bater, pintar/limpar, olhar, lavar, maquiar (hit/slap, paint/clean, look, wash, make up) V2-olhar, pintar, expressar, cuidar, lavar (look, paint, express, care, wash)	N1-rosto/coroa (face/head) N2-face/homem, moeda, rosto (face/man, coin, face)
Órgão (organ)	V1-doar/funcionar (donate/function) V2-tocar/doar (touch/donate)	N1-corpo/coração (body/hear) N2-pulmão/vida (ling/life)
Papa (Pope/pap)	V1-rezar/comer (pray/eat) V2-rezar/falar, comer (pray/talk, eat)	N1-igreja/comida (church/food) N2-religião/igreja, comida (religion/church, food)

*In case of a draw in the semantic association strength between the first and second recalls, all of the words were mentioned
Caption: V1 = 1st verb recall; V2 = 2nd verb recall; N1 = 1st noun recall; N2 = 2nd noun recall

Table 3. Continued...

Target words	Semantic associates* V1 and V2	Semantic associates* N1 and N2
Mato (bush)	V1-cortar/cheirar (cut/smell) V2-plantar/cortar (plant/cut)	N1-grama/ bicho, animais, cachorro, capim, floresta, moita, planta, terra, verde (lawn/animal/ bug, animals, dog, grass, forest/woods, plant, earth/land, green) N2-verde/capim, grama (green/grass, lawn)
Faixa (strip/band/crosswalk)	V1-passar, ultrapassar/atravessar, pintar, limitar (pass/transfer, overtake, paint, limit) V2-colocar/cortar, usar (put on/cut, wear/use)	N1-pedestre/cabelo (pedestrian/hair) N2-pedestre/cabelo (pedestrian/hair)
Serra (saw/mountain range)	V1-subir/cortar (climb/go up/cut) V2-descer/subir (go down/go up/climb)	N1-praia/montanha (beach/hill/mountain) N2-neblina/estrada, ferramenta, viagem (fog/road, tool, trip)
Vídeo (video)	V1-assistir/ver, gravar (watch/see, record) V2-gravar, fazer, assistir/relembrar, alugar, divertir, filmar (record, make, watch/recall, rent/lease, amuse, video record/shoot)	N1-filme/televisão (movie/television) N2-filme/televisão (movie/television)
Cabo (cable)	V1-conectar, puxar (connect, pull) V2-ligar/conectar (plug/connect)	N1-televisão/aço (television/steel) N2-televisão/fio (television/wire)
Campo (field/countryside)	V1-jogar/correr (play/run) V2-jogar/descansar (play/rest)	N1-futebol/grama (soccer/lawn) N2-flor, grama (flower, grass)
Neto (grandson)	V1-amar, brincar (love, play) V2-amar/brincar (love/play)	N1-avô/família (grandfather/family) N2-avô/criança (grandfather/child)
Time (team)	V1-torcer/jogar, unir (cheer/play, join) V2-jogar/ganhar (play/win)	N1-futebol/jogo, torcida, união (soccer/game, fans, union) N2-torcida/equipe, união (fans/team, union)
Vila (town/village)	V1-morar/viver (dwell/live)V2-morar/visitar (dwell/ visit)	N1-casa/bairro (house/home/neighborhood) N2-cidade, rua (city/town, street)
Gente (people/folk)	V1-conversar/conhecer (talk/meet) V2-conhecer/chorar, aprender, correr, juntar, respeitar (meet/cry, learn, run, gather, respect)	N1-pessoa/povo (person/people) N2-mundo, pessoas (world, persons)
Porto (port/harbor)	V1-navegar/parar, embarcar (navigate/stop, board) V2-navegar, chegar (navigate, arrive)	N1-navio/barco (ship/boat) N2-barco/mar (boat/sea)
Banda (band)	V1-tocar/cantar (play/sing) V2-cantar/tocar (sing/play)	N1-música/show (music/song/concert) N2-música/bateria, guitarra (music/song/drums, guitar)
Texto (text)	V1-ler/escrever (read/write) V2-escrever/ler (write/read)	N1-palavra/redação (word/composition) N2-leitura/palavra (reading/word)
Filme (movie)	V1-assistir/ver (watch/see) V2-assistir, gostar (watch, like)	N1-cinema/pipoca (cinema/popcorn) N2-pipoca/cinema (popcorn/cinema)
Globo (globe/Globo TV)	V1-girar, assistir (revolve, watch) V2-assistir, rodar (watch, go round)	N1-televisão/mundo (television/world) N2-mundo/bola (world/ball)
Feira (marketplace)	V1-comprar/gritar, comer (buy/shout, eat) V2-comprar/comer (buy/eat)	N1-frutras/pastel (fruits/fried pie) N2-frutas/legumes, verdura (fruits/vegetables, greens)
Lista (list)	V1-fazer/assinar (make/sign) V2-ler, escrever, lembrar (read, write, remember)	N1-compras/mercado (shopping/market) N2-compras, mercado, supermercado (shopping, market, supermarket)
Palco (stage)	V1-dançar/apresentar, cantar (dance/perform, sing) V2-cantar, apresentar (sing, perform)	N1-teatro/apresentação, show (theater/ performance, concert) N2-show/teatro (concert/theater)
Clube (club)	V1-nadar/divertir (swim/have fun/enjoy) V2-nadar/jogar (swim/play)	N1-piscina/esporte, sol (swimming pool/sport, sun) N2-sol/amigos (sun/friends)
Praça (square)	V1-descansar/passear (rest/walk/wander) V2-passear/brincar, conversar (walk/wander, talk)	N1-banco/árvore (bench/tree) N2- banco/árvore (bench/tree)
Praia (beach)	V1-nadar/bronzear, divertir, queimar (swim/suntan, enjoy/have fun, sunburn) V2-nadar/bronzear (swim/sun tan)	N1-sol/areia (sun/sand) N2- areia/sol (sand/sun)
Dólar (Dollar)	V1-comprar/gastar (buy/spend) V2-vender/trocar, gastar, comprar (sell/exchange, spend, buy)	N1-dinheiro/verde, viagem (money/cash, trip) N2- real/dinheiro, viagem (Real currency/money, trip)

*In case of a draw in the semantic association strength between the first and second recalls, all of the words were mentioned
Caption: V1 = 1st verb recall; V2 = 2nd verb recall; N1 = 1st noun recall; N2 = 2nd noun recall

Table 3. Continued...

Target words	Semantic associates* V1 and V2	Semantic associates* N1 and N2
Parque (park)	V1-brincar/passear (paly/walk/wander) V2-andar/correr, passear, divertir (walk/run, wander, have fun)	N1-árvore/diversão (tree/fun) N2-árvore/criança (tree/child)
Ilha (island)	V1-nadar/isolar (swim/isolate) V2-nadar/isolar (swim/isolate)	N1-mar/água, coqueiro (sea/water, coconut tree) N2- coqueiro/árvore, mar (coconut tree/tree, sea)
Tela (screen/canvas)	V1-pintar/assistir (paint/watch) V2-pintar/proteger (paint/protect)	N1-cinema/televisão (cinema/television) N2- cinema/computador, filme, mosquito (cinema/computer, movie, mosquito)
Loja (store)	V1-comprar/vender (buy/sell) V2-vender, comprar (sell, buy)	N1-roupa/compras (clothes/shopping/purchases) N2- roupa/compras, sapato (clothes/shopping/purchases, shoe)
Teto (ceiling/roof)	V1-cair/proteger (fall/protect) V2-cair, construir (fall, build)	N1-casa/cobertura, lustre, telhado, vidro (house/cover, chandelier, roof, glass) N2- casa, chão, telhado, telha (house, floor, roof, tile)
Bomba (bomb/candy)	V1-explodir/estourar (explode/burst) V2- matar/destruir (kill/destroy)	N1-guerra/chocolate (war/chocolate) N2- chocolate/explosão, guerra (chocolate/explosion, war)
Bolsa (handbag/purse)	V1-guardar/carregar (keep/carry) V2-guardar/comprar (keep/buy)	N1-dinheiro/couro, mulher (money/leather, woman) N2- celular, mulher (mobile phone, woman)
Carne (meat)	V1-comer/comprar (eat/buy) V2-assar/cozinhar, fazer (roast/cook, make)	N1-boi/vaca N2- churrasco/porco, vaca (barbecue/pork/pig, beef/cow)
Folha (sheet/leaf)	V1-escrever/cair (write/fall) V2-escrever/cair (write/fall)	N1-árvore/papel (tree/paper) N2- árvore, papel (tree, paper)
Quadro (picture/frame)	V1-pintar/pendurar (paint/hang) V2-pintar/comprar (paint/buy)	N1-pintura/parede (painting/wall) N2- arte, parede (art, wall)
Caixa (box/teller)	V1-guardar/fechar (keep/close) V2-guardar/presentear, abrir, organizar (keep/gift, open, organize)	N1-banco, papelão, sapato (bank, pasteboard, shoe) N2- dinheiro, papelão, presente (money, pasteboard, gift)
Dado (die)	V1-jogar/brincar (throw/play/toy) V2-ganhar/brincar, contar (win/play, count)	N1-jogo/números (game/numbers) N2-tabuleiro/sorte (board/luck)
Padre (priest)	V1-rezar/orar (pray/orate) V2-abençoar/falar (bless/speak)	N1-igreja/missa (church/mass) N2- igreja/religião (church/religion)
Boca (mouth)	V1-falar/beijar (speak/kiss) V2-comer/beijar (eat/kiss)	N1-dente/beijo (tooth/kiss) N2-língua/beijo, comida, dente (tongue/kiss, food, tooth)
Ouro (gold)	V1-comprar/ganhar, enriquecer (buy/win, become rich) V2-enriquecer/brilhar, ganhar (become rich/shine, win)	N1-prata/joia, riqueza (silver/jewelry, wealth) N2-joia/bronze (jewelry/bronze)
Bola (ball)	V1-jogar/rolar, chutar (play/roll, kick) V2-brincar/chutar (play/kick)	N1-futebol/jogo (soccer/game) N2-criança, futebol (child, soccer)
Foto (photo)	V1-tirar/revelar (take/develop) V2-revelar/recordar (develop/remember)	N1-lembrança/câmera, família, paisagem (memory/camera, family, landscape) N2-família/recordação (family/remembrance)
Rio (river)	V1-nadar/correr, navegar (swim/flow, navigate) V2-nadar/pescar (swim/fish)	N1-água/barco, peixe (water/boat, fish) N2-peixe/barco (fish/boat)
Casa (house)	V1-morar/dormir (dwell/sleep) V2-descansar/limpar, morar (rest/clean, dwell)	N1-família/lar, teto (family/home, ceiling/roof) N2-família/quarto (family/bedroom)
Banco (bank/bench)	V1-sentar/pagar (sit/pay) V2-sentar/depositar (sit/deposit)	N1-dinheiro/praça (money/square) N2-dinheiro/praça (money/square)
Noite (night)	V1-dormir/escurecer (sleep/grow dark) V2-dormir/admirar, sair (sleep/admire, go out)	N1-lua/estrela (moon/star) N2-estrela/lua, sono (star/moon, sleep)
Livro (book)	V1-ler/comprar (read/buy) V2-escrever/estudar (write/study)	N1-leitura/história, página, papel (reading/history, page, paper) N2-escola/estudo (school/study)
Filho (son)	V1-amar/cuidar (love/care) V2-amar/cuidar, educar (love/care, educate)	N1-mãe/pai (mother/father) N2-pai/mãe (mother/father)

*In case of a draw in the semantic association strength between the first and second recalls, all of the words were mentioned
Caption: V1 = 1st verb recall; V2 = 2nd verb recall; N1 = 1st noun recall; N2 = 2nd noun recall

Table 3. Continued...

Target words	Semantic associates* V1 and V2	Semantic associates* N1 and N2
Rede (net/hammock)	V1-pescar/balançar (fish/swing) V2-dormir/descansar, navegar (sleep/rest, navigate)	N1-descanso, peixe (rest, fish) N2-descanso, peixe (rest, fish)
Carro (car)	V1-dirigir/andar (car/ride) V2-bater/andar (crash/ride)	N1-pneu/passeio, roda, volante (tire/ride, wheel, steering wheel) N2-rua, trânsito (street, traffic)
Terra (earth/land)	V1-plantar/arar (plant/plow) V2-plantar/colher (plant/sow/harvest)	N1-plantação/água, árvore, barro, marrom, planeta, planta (crop/water, tree, mud, brown, planet, plant) N2-mato/água, flor, interior, marrom, minhoca, planta (grass/water, flower, countryside, brown, worm, plant)
Homem (man)	V1-casar/amar (marry/love) V2-trabalhar/conversar, casar (work/talk, marry)	N1-mulher/cabelo, pai, trabalho (woman/hair, father, work) N2-pessoa/mulher, pai (person/woman, father)
Rua (street)	V1-andar/atravessar (walk/cross) V2-dirigir/caminhar, andar, olhar, correr, passear (drive/wander, walk, look, run, stroll)	N1-estrada/asfalto (road/asphalt) N2-avenida, calçada, carro (avenue, sidewalk, car)
Meia (sock/stocking/half/half price)	V1-vestir/calçar (wear/put on) V2-lavar/dividir (wash/divide/split)	N1-tênis/pé, sapato (sneaker/foot, shoe) N2-pé/cinema, frio, metade, sapato, tênis (foot/cinema, cold, half, shoe, sneaker)
Água (water)	V1-beber (drink) V2-banhar, nadar (bathe, swim)	N1-sede/copo (thirst, glass, cup) N2-rio, sede (river, thirst)
Linha (line/thread)	V1-costurar/desenhar (sew/draw)V2-costurar, seguir (sew, follow)	N1-agulha/costura (needle/seam) N2-costura/roupa (seam/clothes)
Corpo (body)	V1-cuidar/movimentar (care/move) V2-vestir/tocar, equilibrar, malhar, cuidar (dress/change, balance, work out, care)	N1-órgão/boca, pessoa, saúde (organ/mouth, person, health) N2-pessoa, roupa, saúde, tronco, vida (person, clothes, health, torso, life)
Prova (test/exam)	V1-estudar/fazer (study/take) V2-estudar, sofrer, fazer (study, suffer, take)	N1-nota/teste (grade/test) N2-escola, estudo (school, study)
Disco (record/disco)	V1-tocar/ouvir (play/listen) V2-escutar, dançar (listen, dance)	N1-música/antiguidade, festa, visual (music/song/antique, party, scenery) N2-música/vitrola (music/record player/turntable)
Porta (door)	V1-abrir/fechar (open/close) V2-fechar/abrir (close/open)	N1-madeira/casa (wood/house) N2-casa/madeira (house/wood)
Festa (party/feast)	V1-dançar/divertir (dance/have fun) V2-beber/dançar (drink/dance)	N1-bexiga/música (balloon/music) N2-bolo/música, amigos (cake/music, friends)
Sala (living room)	V1-sentar/estudar (sit/study) V2-descansar/assistir (rest/watch)	N1-sala, sofá (room, sofa) N2-televisão/sofá (television/sofa)
Carta (letter)	V1-escrever/ler (write/read) V2-escrever/enviar (write/mail)	N1-envelope/correio (envelope/post office) N2-selo/correio, saudade (stamp/post office, longing)
Quarto (bedroom)	V1-dormir/arrumar (sleep/clean up) V2-descansar/estudar (rest/study)	N1-cama/cozinha (bed/kitchen) N2-cama/sono (bed/sleep)
Leite (milk)	V1-beber/tomar (drink/have) V2-tirar/beber, misturar (milk/drink, blend)	N1-vaca/copo (cow/glass) N2-caixa/café, chocolate, geladeira (bottle/coffee, chocolate, refrigerator)
Fogo (fire)	V1-queimar/apagar (burn/put out) V2-esquentar/queimar, apagar (warm/burn, put out)	N1-madeira/água, incêndio (wood/water, fire) N2-água/bombeiro (water/firefighter)
Sangue (blood)	V1-doar/machucar (donate/hurt) V2-morrer, examinar, machucar, jorrar (die, examine, hurt, gush)	N1-corpo/hospital (body/hospital) N2-machucado/exame, hemoglobina, morte, vermelho (injury/examination, hemoglobin, death, red)
Mesa (table)	V1-comer/colocar (eat/set) V2-comer/sentar, estudar (eat/sit, study)	N1-cadeira/comida (chair/food) N2-cadeira/prato (chair/plate)
Ponte (bridge)	V1-atravessar/passar (cross/pass) V2-construir/ligar (build/connect)	N1-rio/carro (river/car) N2-rio/carro, madeira (river/car, wood)

*In case of a draw in the semantic association strength between the first and second recalls, all of the words were mentioned
Caption: V1 = 1st verb recall; V2 = 2nd verb recall; N1 = 1st noun recall; N2 = 2nd noun recall

Table 3. Continued...

Target words	Semantic associates* V1 and V2	Semantic associates* N1 and N2
Rocha (rock)	V1-quebrar/endurecer (break/toughen) V2-quebrar/jogar, destruir, bater (break/throw, destroy, hit)	N1-pedra/duro, gruta, praia (stone/toughness, grotto, beach) N2-montanha/cachoeira (hill/mountain/waterfall)
Rosa (rose)	V1-cheirar/ganhar (smell/receive) V2-cheirar/plantar (smell/plant)	N1-flor/amor (flower/love) N2-cor, flor (color, flower)
Dança (dance)	V1-divertir/mexer, movimentar, aproveitar, animar, festejar (have fun/swing, move, enjoy, celebrate) V2-divertir/pular (have fun/jump)	N1-música/bailarina (music/ballet dancer) N2-música/salão (music/ballroom)
Chuva (rain)	V1-molhar/cair (wte/fall) V2-cair/esfriar, lavar (fall/cool, wash)	N1-água/guarda-chuva (water/umbrella) N2-água/frio, guarda-chuva, nuvem, rua, trovão (water/cold, umbrella, cloud, street, thunder)
Câncer (cancer)	V1-morrer/adoecer (die/sicken) V2-curar, adoecer/morrer (cure, sicken/die)	N1-doença/hospital (illness/hospital) N2-morte/doença (death/illness)
Arma (arm/weapon)	V1-matar, atirar (kill, shoot) V2-atirar, machucar, matar (shoot, wound, kill)	N1-fogo/tiro (fire/shot) N2-bala, tiro (bullet, shot)
Massa (pasta/dough)	V1-comer/cozinhar, fazer, preparar (eat/cook, make, prepare) V2-fazer/comer (make/eat)	N1-macarrão/comida (noodle/food) N2-pão/restaurante (bread/restaurant)
Branca (white)	V1-pintar/limpar, sujar (paint/clean, dirty) V2-lavar/limpar, desenhar (wash/clean, draw)	N1-roupa/neve, pomba, preta (clothes/snow, dove, blackness) N2-paz/neve, parede (peace/snow, wall)

*In case of a draw in the semantic association strength between the first and second recalls, all of the words were mentioned
Caption: V1 = 1st verb recall; V2 = 2nd verb recall; N1 = 1st noun recall; N2 = 2nd noun recall

closer to that (mean=19.06) found in this survey. With respect to the increased number of associates for the second recall, only one research[5] conducted a similar procedure (requested three recalls to each target word) and observed an increase in the meaning set size for the second and third recalls. The differences between the results obtained can be explained by the different procedures used, the reduced number of words in common (between two and eight), and the differences in the aspects of concreteness, extension, and grammar class. Interestingly, the study that presented similar data[3] requested only one response to participants and used categories (e.g., farm animals) as targets, not single words. Another study[1] described the category size considering the different responses and the words were presented in sentences of different linguistic contexts, which invalidated the comparison of results.

Regarding the meaning set sizes, most of them were classified as small, without occurrence of large sets in any of the categories considered, for both the first and second written words. Previous studies showed that most sets were either large[3] or medium[2,7]. A national study[8] comparing different age groups reported a majority of large sets for the group of children, medium sets for adults, and small sets for the elderly, without occurrence of any large set only in the group of elderlies. It is worth noting that, in the mentioned studies, only one word was requested and there was no restriction of category.

Association strength between the pair of semantic associates was another aspect analyzed in this research. The semantic association strength tended to decrease from the first (mostly strong) to the second (mostly moderate) word in the first recall. As for the second recall, most of the associations presented moderate strength for both considered words (Table 1). The hypothesis of response chaining[3,11] could explain this decrease in association strength, that is, a certain response is more likely to occur in the first words, whereas one of these responses is less likely to occur, due to greater variability, at a second moment.

Some studies[2,7] have also reported increased values (between 44% and 50%) for pairs with strong and moderate association strength and decreased values (1.1% and 6%) for pairs with weak strength. One study[9], which considered different groups of participants (young adults and elderly), described few pairs of associates with weak association strength (between 3.4% and 5.7%) and a higher percentage (between 46% and 50%) for pairs with moderate and strong strength. Only one response was requested in these studies, thus it was not possible to analyze the reduction of association strength.

With respect to the correlation between semantic association strength and sizes of the total and meaning sets (Table 2), the results showed that the variables are inversely proportional. The stronger the association strength, the smaller the set size, that is, the smaller the number of associates generated per target word, the stronger the association between the most frequent associate and the target word. It is worth emphasizing that significant correlations occurred in all associations in the total set and in half of the associations in the meaning set. Most correlations were considered strong for the total set and moderate for the meaning set.

Similar analyses were conducted in other studies[2,7,11,12], in which significant negative correlation was observed between the variables association strength and total and meaning set size. Considering the correlation strength and the set sizes, some studies found strong correlation for the total size[2,7], whereas others reported moderate correlation[11,12]. Regarding the meaning size, all studies described strong correlations[2,7,11].

The results obtained in the comparison between association strength and total and meaning set sizes corroborate the correlation between both. The association is more evident and

shows a consistent pattern of decreased strength as the total set size increases. Similar analysis has not been mentioned in the studies consulted for comparison of results.

Regarding the semantic associates, the two words generated with greater association strength for both the first and second recalls were considered equally for each group (verbs and nouns). In many cases, there was a tie in the values of words, so that they were all classified to be possibly considered in the formulation of sentences.

Finally, the lists of associates generated allowed control of the semantic association aspect in the formulation of a speech recognition assessment with predictability control of the final word.

CONCLUSIONS

The semantic association norms of 96 Brazilian Portuguese words were collected as proposed and resulted in semantic associates for each target word in the verb and noun classes that were used in a speech recognition assessment considering sentences with controlled word predictability.

REFERENCES

1. Janczura GA. Contexto e normas de associação para palavras: a redução do campo semântico. Paidéia (Ribeirão Preto). 2005;15(32):417-25. http://dx.doi.org/10.1590/S0103-863X2005000300011.
2. Salles JF, Holderbaum CS, Becker N, Rodrigues JC, Liedtke FV, Zibetti MR, et al. Normas de Associação semântica para 88 palavras do português brasileiro. Psico (Porto Alegre). 2008;39(3):260-8.
3. Janczura GA. Normas associativas para 69 categorias semânticas. Psicol, Teor Pesqui. 1996;12(3):237-44.
4. De Deyne S, Storms G. Word associations: network and semantic properties. Behav Res Methods. 2008a;40(1):213-31. PMid:18411545. http://dx.doi.org/10.3758/BRM.40.1.213.
5. De Deyne S, Storms G. Word associations: norms for 1424 Dutch words in a continuous task. Behav Res Methods. 2008b;40(1):198-205. PMid:18411543. http://dx.doi.org/10.3758/BRM.40.1.198.
6. Macizo P, Gómez-Ariza CJ, Bajo MT. Associative norms of 58 Spanish words for children from 8 to 13 years old. Psicologica (Valencia). 2000;21(3):287-300.
7. Salles JF, Holderbaum CS, Machado CS. Normas de associação semântica de 50 palavras do português brasileiro para crianças: tipo, força de associação e set size. Interam J Psychol. 2009;43(1):57-67.
8. Zortea M. Estudo sobre associação semântica de palavras em crianças, adultos jovens e idosos [dissertação]. Porto Alegre: Universidade Federal do Rio Grande do Sul; 2010.
9. Zortea M, Salles JF. Estudo comparativo das associações semânticas de palavras entre adultos jovens e idosos. Psicol, Teor Pesqui. 2012;28(3):259-66. http://dx.doi.org/10.1590/S0102-37722012000300001.
10. Vesely L, Bonner MF, Reilly J, Grossman M. Free association in semantic dementia: the importance of being abstract. [abstract]. Brain Lang. 2007;103(1-2):154-5. http://dx.doi.org/10.1016/j.bandl.2007.07.092.
11. Nelson DL, Schreiber TM. Word concreteness and word structure as independent determinants of recall. J Mem Lang. 1992;31(2):237-60. http://dx.doi.org/10.1016/0749-596X(92)90013-N.
12. Stein LM, Gomes CFA. Normas brasileiras para listas de palavras associadas: associação semântica, concretude, frequência e emocionalidade. Psicol, Teor Pesqui. 2009;25(4):537-46. http://dx.doi.org/10.1590/S0102-37722009000400009.
13. Stein LM, Feix LF, Rohenkohl G. Avanços metodológicos no estudo das falsas memórias: construção e normatização do procedimento de palavras associadas. Psicol Reflex Crit. 2006;19(2):166-76. http://dx.doi.org/10.1590/S0102-79722006000200002.
14. Calais LL. Reconhecimento de fala em idosos: elaboração e aplicação de um teste considerando a previsibilidade da palavra [tese]. São Paulo: Universidade Federal de São Paulo; 2011.
15. Janczura GA, Castilho GM, Rocha NO, Van Erven TJC, Huang TP. Normas de concretude para 909 palavras da língua portuguesa. Psicol, Teor Pesqui. 2007;23(2):195-204.
16. Bybee JL. Phonology and language use. Cambridge: Cambridge University; 2001.
17. Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale: Erlbaum; 1988.
18. Nelson DL, McEvoy CL, Bajo MT. Lexical and semantic search in cued recall, fragment completion, perceptual identification, and recognition. Am J Psychol. 1988;101(4):465-80. PMid:3232723. http://dx.doi.org/10.2307/1423225.

Author contributions

LLC is the main author of the article, DG is the study co-advisor, and ACLCB is the study advisor; LLC, AMLG, PA, DG and ACLCB contributed to the study from its design to the final writing of the manuscript. This study is part of a doctoral dissertation; PA and DG also contributed to the considerations and decisions involving the methodology of the study, analysis of results, and writing of the manuscript.