PSYCHOLOGY & NEUROSCIENCE

Psychology & Neuroscience, 2012, 5, 1, 77 - 81
DOI: 10.3922/j.psns.2012.1.10

Semantic word association: comparative data for Brazilian children and adults

Maxciel Zortea and Jerusa Fumagalli de Salles

Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

Abstract

The construction of associated word lists is important for the elaboration of psychological and neuropsychological tasks and experiments. It remains unknown whether differences exist in the semantic associations of words from childhood to adulthood, possibly indicating important lexico-semantic developmental changes that influence neuropsychological assessment. The present study compared semantic word associations in children and adults in terms of forward associative strength and set size. The participants included 247 children from the third grade of elementary school, aged 7 to 11 years (M = 9.17 years, SD = 0.83 years), and 108 adults, aged 16 to 49 years (M = 22.17 years, SD = 6.04 years) from the study of Salles et al. (2008). The task consisted of the participants responding to the first word that came to mind (associate) with a meaning related to each of 87 words presented aloud (target). The children's responses had significantly higher forward associative strength between the target and most frequent associate word and a smaller response diversity index. Although the meaning and total set size did not significantly differ between groups, 40.2% of the targets had a large meaning set size in the children compared with only 10.3% in the adults. Among the most strongly associated pairs, 56.3% were equal between the sample groups. These results suggest that the selection of stimuli for the construction of verbal cognitive tasks should consider specific word association norms for different ages. **Keywords:** word association; developmental age groups; lexicon, semantic memory; verbal stimuli.

Received 18 November 2011; received in revised form 2 March 2012; accepted 7 March 2012. Available online 29 June 2012.

Introduction

The semantic word association task is a method that is easy to administer and has a historical and consolidated use in psychology (de Deyne & Storms, 2008; Nelson, McEvoy, & Dennis, 2000) to study lexico-semantic knowledge and how people retrieve concepts and words from memory. Word association norms are used to create tasks to evaluate memory and language processes (e.g., false memory; Roedinger & McDermott, 1995; Stein, Feix, & Rohenkohl, 2006) in clinical neuropsychiatry (Bhatt, Laws, & McKenna, 2010) and non-clinical samples (Stein & Pergher, 2001). Due to changes over time in the cultural and linguistic meaning and use of words (Hirsh & Tree, 2001), continuously updating such norms is important. In the context of the development of word associations, this allows the identification of similarities and differences in associations between children and adults (Macizo, Gómez-Ariza, & Bajo, 2000), and between adults and elderly persons (Hirsh & Tree, 2001). Disparate results between these groups indicate that memory tasks (e.g., implicit memory; Holderbaum & Salles, 2011) and cued recall (Nelson & Schreiber, 1992) should consider specific normative data for each age group.

Depending on the instructions received by the participants, word association tasks may involve free association (Nelson, McEvoy, & Schreiber, 1999) in which participants must respond with the first word that comes to mind (*associate*) when a stimulus-word (*target*) is presented, or semantic association in which the response must have a meaning related to the target (Coronges, Stacy, & Valente, 2007; Salles, Holderbaum, & Machado, 2009). Nelson et al. (1999) proposed the psychological implication of such an evaluation, suggesting that the process of evocating new information is mediated by the preexisting and implicit association strength between a word to be recalled and a cue or contextual stimulus word. Therefore, assessing these measures that underlie the stimulus and are related to lexico-semantic knowledge becomes important.

Every word association task may provide *forward associative strength*, *meaning set size*, and *total set size* measures. Forward associative strength refers to the proportion of individuals who provide a response that is dominant or more frequently associated with a

Maxciel Zortea and Jerusa Fumagalli de Salles, Institute of Psychology, Federal University of Rio Grande do Sul, Brazil. Correspondence regarding this article should be directed to: Maxciel Zortea, Rua Ramiro Barcelos, 2600, sala 114, 1º andar, Porto Alegre, Rio Grande do Sul, 90035-003, Brazil. Phone: 55 51 3308 5341. E-mail: max.zortea@gmail.com

target in a group of respondents (Nelson et al., 2000). Associative strength is considered weak when the proportion is less than 10% in the group, medium when it ranges from 10% to 24%, and strong when it is 25% or higher (Salles et al., 2008). Meaning set size refers to the number of different responses given in response to a target by two or more people, and total set size includes all the responses (Nelson & Schreiber, 1992). Meaning set size is considered small when it has 1 to 8 different associations, average when it has 9 to 16 different associations, and large when it has 17 to 34 different associations (Nelson & Schreiber, 1992). To obtain a general measure of the agreement of responses for each target, Shannon's *H* statistic is used, which is a *response diversity index* for a group of people (Hirsh & Tree, 2001). This measure is obtained from the following formula:

$$H = -\sum_{i=1}^{k} pi \log_2 pi \quad (1)$$

where *i* represents the number of associations with the target and *pi* represents the probability that *i* is associated with the target.

Empirical studies have attempted to compare forward associative strength and set size between non-clinical samples of children and adults, but no studies have used a response diversity index like the *H* statistic to compare these two groups. Macizo et al. (2000) compared performance in a free association task in four groups of Spanish participants. Three of the groups consisted of children aged 8 to 9 years, 10 to 11 years, and 12 to 13 years. The fourth group consisted of adults (ages not presented). No significant differences in forward associative strength were found between children and adults. However, the youngest group (aged 8 to 9 years) presented higher forward associative strength between pairs compared with the other groups of children. The authors suggested that such a decrease with age may be attributable to a process of learning through restructuring in which words acquire new associations as a function of new meanings and interpretations. Macizo et al. (2000) also found an increase in meaning set size in the 8- to 9-year-old group and 10- to 13-year-old group with a decrease in adulthood. This pattern was similar when only the number of idiosyncratic responses given by a single participant was considered. According to the authors, learning occurs through a tuning process between the onset of adolescence and adulthood in which associations are refined so that the number of words to explain a phenomenon is reduced.

Coronges et al. (2007) investigated total and meaning set sizes in a sample of 1,097 seventh graders aged 12 to 13 years compared with the adult sample reported by Nelson et al. (1999) that consisted of more than 6,000 adults. Although the total set size was larger in the seventh graders, the meaning set size was smaller compared with the adults. According to Coronges et al. (2007), the younger participants may filter their responses less and possess more divergent thought. However, one methodological issue must be considered. The adult sample in the study by Nelson et al. (1999) completed a free association task, whereas the children responded to a semantic association task.

Finally, de la Haye (2003) investigated the *agreement on the most frequent response* to each target from a list of 200 words in a free association task in French children (9- to 11-years of age) and adults (18- to 34-years of age). The responses were similar among the children. However, only 62.5%, 63%, and 68.5% of the most frequent associations at 9, 10, and 11 years of age, respectively, were the same as those given by the adults.

To overcome some of the limitations of these studies, the present study used the same task (i.e., semantic word association) for both children and adults, thus facilitating between-group comparisons. Moreover, the use of various measures of word association such as the ones proposed here for the same samples may expand investigations in the area. Conducting research in another language such as Brazilian Portuguese may indicate the degree of generalization of age differences in word associations. The present study compared semantic word associations between children and adults by assessing forward associative strength, meaning set size, total set size, the response diversity index, and agreement between groups in terms of the most frequent associate of each target.

Methods

Participants

The participants were 247 third-grade children, aged 7 to 11 years ($M = 9.17$ years, $SD = 0.83$ years), from five public primary schools in the city of Porto Alegre, RS, Brazil. Boys comprised 46.6% of the sample, and girls comprised 46.2% of the sample. Gender information was not available for 13 cases. A portion of the sample (154 children) was obtained from the database of the study by Salles et al. (2009). A sample of 108 young adults, undergraduate students enrolled in the first 3 years of Psychology and Speech Therapy courses, were the participants from the study by Salles et al. (2008). They were aged 16 to 49 years ($M = 22.17$ years, $SD = 6.04$ years). Women comprised 69.4% of the sample, and men comprised 23.1% of the sample (7.4% did not provide such information). Data were reanalyzed using the following selection criteria of valid responses. The present study was approved by the Research Ethics Committee of the Universidade Federal do Rio Grande do Sul.

Apparatus

The target list consisted of 87 target words in Brazilian Portuguese with various levels of occurrence frequency (from nine to 48,037 occurrences according to the norms of Sardinha [2003]), degree of concreteness (from 2.35 to 6.85 according to the norms of Janczura, Castilho, Rocha, Van Erven, & Huang [2007]), and word

length (60 target words had six or fewer letters, and 27 had seven or more letters). The Appendix presents this information for each target word.

Procedures

In the semantic word association task, the participant had to respond as quickly as possible to the first word (forward association) that came to mind with a meaning related to a word said by the evaluator. Examples like "church" and "football" were given to verify that the participant understood the instructions. The children were assessed individually and provided the associated word orally to increase the speed of their response. The adults were assessed collectively, and their responses were in written form. Responses such as "I don't know" (in the case of children) or that were blank, illegible, words identical to the target word, or words that did not fit the word criterion (in the case of adults) were considered invalid. To assess forward associative strength, only the most frequent associates of each target were used.

Statistical analysis

Because of the asymmetric distribution of the data found with the Kolmogorov-Smirnov test, nonparametric Mann-Whitney tests for differences between average ranks were conducted for forward associative strength, meaning set size, total set size, and the response diversity index. The χ^2 test was used for frequency variables such as the categories of forward associative strength and meaning set size.

Results

Table 1 presents the mean and standard deviation of each analyzed variable by group. The analysis showed significant differences between groups for forward associative strength (U [86] = 2958.5, z = -2.49, p = .013) and the response diversity index (U [86] = 2713.0, z = -3.22, p = .001). The analysis of effect sizes showed a Cliff's Delta (δ) = 0.22 for forward associative strength and δ = 0.28 for the response diversity index. No significant difference was found between children and adults in meaning set size and total set size. Categorical analyses (Table 2) were also conducted for forward associative strength and meaning set size to determine which category was more frequent in each group. The χ^2 test revealed significantly more pairs with a strong association between the target and most frequent associate in the children (χ^2 [2] = 44.9, p < .001). The phi coefficient for this contrast was φ = .72. The adults had more pairs in the medium and strong categories of association between the target and most frequent associate (χ^2 [2] = 35.2, p < .001, φ = .64) and a higher frequency of targets that generated a medium meaning set size (χ^2 [2] = 78.7, p < .001, φ = .95). In terms of the percentage of agreement between groups with regard to the most frequently associated pair of each target, the frequency analysis showed that 56.3% of the most frequent responses were the same in the two groups.

Table 1. Mean and standard deviations for forward associative strength, total set size, meaning set size, and response diversity index in each group

	Children	**Adults**
Forward associative strength	.37 (*.21*)[a]	.28 (*.14*)[b]
Total set size	36.64 (*17.45*)[a]	37.48 (*8.99*)[a]
Meaning set size	13.64 (*7.30*)[a]	12.33 (*3.29*)[a]
Response diversity index	3.59 (*1.17*)[a]	4.12 (*.76*)[b]

[a,b] Same letters indicate no significant differences at the .05 level between the average ranks of the groups (Mann-Whitney).

Table 2. Target-associate pair percentages for forward associative strength categories and associate percentages relative to each target for meaning set size categories in each group

	Association strength			**Significant set size**		
Group	*Weak*	*Medium*	*Strong*	*Small*	*Medium*	*Large*
Children	3.4	34.5	62.1	29.9	29.9	40.2
Adults	3.4	46.0	50.6	11.5	78.2	10.3

Discussion

The present results showed that the semantic word association measures were different between children and adults. With regard to forward associative strength between each target and its most frequent associate, higher strength was observed in the children than in the adults. These results differ from Macizo et al. (2000) in which no differences were found between groups of children aged 8 to 13 years and adults. Notably, the authors used a free association task that tends not to restrict the response as much as the semantic word association task used in the present study. Such a task requires that a response is semantically associated with the target. Moreover, Macizo et al. (2000) found differences between the 8- to 9-year-old children and 10- to 13-year-old children. The authors attributed these differences to the acquisition of new concepts and therefore new word associations. Sauzéon, Lestage, Raboutet, N'Kaoua, & Claverie (2004) used a semantic fluency task and found that the number of conceptual categories and elements in each category tended to increase at about 11 to 13 years of age.

Although the present study did not conduct statistical analyses for each age range of children (7 to 11 years old), the comparison between children and adults suggests the occurrence of an acquisition process of new concepts or conceptual categories. The targets did not generate associations for the adults that were as strong as the ones for the children. Additionally, the adults presented a higher response diversity index

compared with the children, indicating that their responses were more heterogeneous. The categorical analysis of forward associative strength also suggests this heterogeneity in which 62.1% of the targets in the group of children were strongly associated with their most frequent associate, whereas the percentage was smaller (50.6%) in the adults. Meaning set size also exhibited a between-group difference with regard to the percentage of targets that had a medium number of associates (29.9% for the children *vs*. 78.2% for the adults).

Children's vocabulary gradually increases over the years (Sauzéon et al., 2004), and this increase may be related to the present findings. Notably, however, the effect sizes found in the independent group analysis (Mann-Whitney test) for both forward associative strength and the response diversity index were fairly small, indicating that this difference was not apparent. A longitudinal or cross-sectional study with a larger number of age groups may clarify whether the semantic association of words is able to illustrate the acquisition process of new concepts or enhancement of existing concepts. Additionally, the ages of the adults were not presented in previous studies (e.g., Macizo et al., 2000; Coronges et al., 2007), making rigorous comparisons difficult.

With regard to total and meaning set sizes, Mann-Whitney tests did not reveal differences between groups. Moreover, adults exhibited a higher frequency of medium set sizes, whereas children did not exhibit a preference for any of the categories. These results appear to conflict with other studies. For example, Coronges et al. (2007) verified a larger meaning set size in adults compared with 12- to 13-year-old children. When all of the provided responses were considered (i.e., total set size), the measure was larger in the children. In contrast, Macizo et al. (2000) found that the meaning set size on average increased from 8 to 10 years of age but decreased in adulthood. These opposing results may be attributable to the different instructions given, such as for the free association task in the study by Macizo et al. (2000) and the semantic association task in the study by Coronges et al. (2007), different stimulus number and quality, and the diverse sample sizes.

However, the fact that no significant difference was found between groups in the present study with regard to meaning set size and total set size deserves more in-depth consideration. First, the semantic word association task, in contrast with the free association task, includes restricted responses that may naturally decrease the number of different associations for each target. Second, the word list of the present study had 38 targets (43.7%) with the goal of generating associates with high forward associative strength and categorical relationships such as antonyms (e.g., *noite* [night] and *dia* [day]). We would thus expect that the set size would be smaller in the case of categorical relationships, considering the negative correlation between forward associative strength measures and set size (Salles et al., 2008, 2009). When we analyzed separately the responses to the 49 and 38 targets between children and adults, we found that the children had larger meaning and total set sizes in the list of 49 targets and smaller meaning and total set sizes in the list of 38 targets. In contrast, adults maintained the same means. Salles et al. (2009) found that third-grade children exhibited a preference for establishing thematic relationships between the target and most frequent associate. Therefore, list characteristics such as targets with a trend toward generating associates with either thematic or categorical relationships may have influenced the results. To overcome this limitation, future studies should control list characteristics to clarify this effect. Moreover, the children's responses were oral, whereas adult responses were written. We determined that requiring written responses from children at 7 or 8 years of age would be inadequate because fast and automatic responses were necessary. With the procedure used herein, we believe that both the children and adults could respond quickly and automatically.

Overall, considering word association norms according to the characteristics of the study samples—children and adults—enables the more careful assessment of the mnemonic and linguistic processes that consider the semantic issue. This is a function of the differences relative to association strength and response diversity. Similarly, the low agreement (56.3%) of responses that were most often provided for the targets between groups, consistent with results from other studies (de la Haye, 2003), supports this need. Furthermore, Nelson et al. (1999) suggested that preexisting associative relationships between words influences any task that involves single words. This underscores the importance of word association experiments in cognitive assessment.

References

Bhatt, R., Laws, K. R., & McKenna, P. J. (2010). False memory in schizophrenia patients with and without delusions. *Psychiatry Research*, *178*, 260-265.

Coronges, K. A., Stacy, A. W., & Valente, T. W. (2007). Structural comparison of cognitive associative networks in two populations. *Journal of Applied Social Psychology*, *37*, 2097-2129.

de Deyne, S., & Storms, G. (2008). Word associations: Norms for 1,424 Dutch words in a continuous task. *Behavior Research Methods*, *40*, 198-205.

de la Haye, F. (2003). Normes d'associations verbales chez des enfants de 9, 10 et 11 ans et des adultes. *L'Année Psychologique*, *103*, 109-130.

Hirsh, K. W., & Tree, J. J. (2001). Word association norms for two cohorts of British adults. *Journal of Neurolinguistics*, *14*, 1-44.

Holderbaum, C. S., & Salles, J. F. (2011). Semantic priming effects in a lexical decision task: comparing third graders and college students in two different stimulus onset asynchronies. *Spanish Journal of Psychology*, *14*, 589-599.

Janczura, G. A., Castilho, G. M., Rocha, N. O., van Erven, T. J. C., & Huang, T. P. (2007). Normas de concretude para 909 palavras da língua portuguesa. *Psicologia: Teoria e Pesquisa, 23*, 195-204.

Macizo, P., Gómez-Ariza, C. J., & Bajo, M. T. (2000). Associative norms of 58 Spanish words for children from 8 to 13 years old. *Psicológica*, *21*, 287-300.

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.

Nelson, D. L., McEvoy, C. L., & Dennis, S. (2000). What is free association and what does it measure? *Memory & Cognition*, *28*, 887-899.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1999). *The University of South Florida word association, rhyme and fragment norms*. Web site: http://w3.usf.edu/FreeAssociation/; accessed at March 27, 2012.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 803-814.

Salles, J. F., 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*. *Revista Interamericana de Psicologia*, *43*, 57-67.

Salles, J. F., Holderbaum, C. S., Becker, N., Rodrigues, J. C., Liedtke, F. V., Zibetti, M. R., & Piccoli, L. F. (2008). Normas de associação semântica para 88 palavras do Português Brasileiro. *Revista PSICO*, *39*, 362-370.

Sardinha, T. B. (2003). *The bank of Portuguese.* Web site: http://www2.lael.pucsp.br/~tony/tony/pesquisa.html; accessed at November 12, 2005.

Sauzéon, H., Lestage, P., Raboutet, C., N'Kaoua, B., & Claverie, B. (2004). Verbal fluency output in children aged 7-16 as a function of the production criterion: Qualitative analysis of clustering, switching processes, and semantic network exploitation. *Brain and Language*, *89*, 192-202.

Stein, L. M., & Pergher, G. K. (2001). Criando falsas memórias em adultos por meio de palavras associadas. *Psicologia: Reflexão e Crítica*, *14*, 353-366.

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*, 166-176.

Appendix. Target words and their occurrence frequency, degree of concreteness, and word length

Target words	OF	DC	WL	Target words	OF	DC	WL
aberto	2568	-	1	longe	2919	2.83	1
agua	6805	6.42	1	mãe	5508	5.95	1
alegria	1367	2.44	2	magro	237	5.27	1
amarelo	1034	5.23	2	maionese	92	-	2
antes	18249	-	1	medo	3014	2.52	1
areia	737	6.25	1	meia	5829	6.18	1
astronauta	89	-	2	mês	21079	-	1
atmosfera	508	-	2	montanha	496	6.14	2
bandeira	1863	6.42	2	natal	3380	-	1
biblioteca	858	-	2	noite	12472	4.77	1
bola	4994	-	1	ontem	48037	-	1
borracha	350	6.42	2	orquestra	1195	-	2
boxe	609	-	1	papel	6983	6.64	1
brasa	100	-	1	porta	4887	6.62	1
capim	121	-	1	preciosa	81	-	2
carro	10668	6.85	1	queijo	540	6.76	1
casa	21801	6.63	1	rádio	3213	6.72	1
cemitério	966	-	2	raiva	538	2.72	1
copo	372	6.72	1	rei	2204	5.00	1
crime	5510	3.56	1	resposta	3680	-	2
dente	257	6.75	1	resultado	7657	-	2
dentro	9021	-	1	rico	1000	-	1
doce	1136	4.83	1	sal	479	-	1
droga	1512	4.90	1	sapo	188	-	1
elefante	150	6.73	2	satisfação	635	2.35	2
erva	114	-	1	seco	480	3.88	1
escorpião	77	-	2	sede	2955	-	1
exército	4883	5.78	2	silenciosa	127	-	2
fácil	3339	-	1	sorte	1739	-	1
fantoche	9	-	2	sujo	248	-	1
febre	646	-	1	sul	3579	-	1
feio	376	2.74	1	tarde	7226	3.80	1
final	19027	-	1	táxi	674	-	1
floresta	716	-	2	terra	6001	5.94	1
fogão	179	6.65	1	toalha	189	-	2
forte	6159	3.53	1	tosse	100	5.35	1
fralda	67	-	1	trabalho	16858	4.47	2
frio	1814	3.42	1	transporte	2335	5.01	2
fruta	355	6.57	1	vazio	560	-	1
importante	6414	-	2	velho	3393	-	1
indústria	7940	-	2	vento	989	4.36	1
isca	68	-	1	vermelho	1861	4.96	2
letra	767	5.25	1	zero	1751	-	1
leve	1293	-	1				

OF, occurrence frequency from Sardinha (2003); DC, degree of concreteness from Janczura et al. (2007); WL, word length (1 = WL ≤ 6 letters; 2 = WL ≥ 7 letters).