Ege Aphasia Test: Description of the Test and Performance in Normal Subjects

Ege Afazi Testi: Testin Tanımlanması ve Normal Kişilerdeki Performansı

Funda ATAMAZ, Arzu YAĞIZ ON, Berrin DURMAZ

Ege Üniversitesi Tıp Fakültesi Fiziksel Tıp ve Rehabilitasyon Anabilim Dalı, İzmir, Türkiye

Summary

Objective: Because of the fundamental differences between Turkish and Indo-European languages, translated forms of currently available aphasia tests would not be adequate for the Turkish language. This fact promoted us to develop a new aphasia evaluation test named Ege Aphasia Test (EAT). The present study describes the development of the test, its normative data in Turkish population, and its reliability. Also the effects of the demographic data on the various subtests are reported.
Materials and Methods: EAT consisted of nine subtests including spontaneous speech, praxis, understanding what is heard, production of language, repetition, naming, reading and understanding what is seen, writing and picturing, calculations. The test was administered to 133 healthy volunteers. Subjects were classified according to their age, gender and educational levels. Twenty-five were tested twice to collect test-retest reliability data.
Results: Educational level was the demographic variable mostly influenced the performance of the EAT. Detrimental impacts of aging have also been demonstrated, but coexisting effect of educational level could not be ruled out. EAT showed excellent test-retest reliability for the overall score and for the subtests.
Conlusion: EAT presented here allows evaluation of several aspects of language with high reliability. The test is influenced by demographic variables, thus adjustments in the test scores based on a particular patient's characteristics should be made and precise indicators of socio-cultural reality of the country should be considered. Diagnostic validity and acceptability of the test should be determined in further studies on the aphasic patients. *Turk J Phys Med Rehab 2007;53:5-10*
Key Words: Aphasia, aphasia evaluating test, normative data, Turkish

Özet

Amaç: Türk diliyle Batı dilleri arasındaki temel farklılıklar nedeniyle, bugün geçerli olan afazi testlerinin Türkçe'ye çevrilmiş halleri Türk dili için yeterli olmayacaktır. Bu gerçek bizi Ege Afazi Testi (EAT) adıyla yeni bir afazi değerlendirme testi geliştirmeye yöneltmiştir. Bu çalışma, testin gelişimini, Türk popülasyonundaki normal değerlerini ve tekrarlanabilirliğini tanımlamaktadır. Farklı alt testler üzerindeki demografik verilerin etkileri de sunulmaktadır.
Gereç ve Yöntem: EAT; spontan konuşma, praksi, gördüğünü anlama, dil üretimi, tekrarlama, isimlendirme, okuma/okuduğunu anlama, yazma/resim yapma ve hesaplamadan oluşan 9 alt testten oluşmaktaydı. Test 133 sağlıklı gönüllü tarafından dolduruldu. Gönüllüler yaş, cins ve eğitim düzeylerine göre sınıflandırıldılar. Yirmi beş gönüllü test-tekrar test tekrarlanabilirliği için 2 kez test edildi.
Bulgular: Eğitim düzeyi EAT performansını en fazla etkileyen demografik veriydi. Yaşlanmanın test üzerine olumsuz etkisi gösterilmekle birlikte, eşlik eden düşük eğitim düzeyinin etkisi dışlanamadı. EAT, ortalama skor ve alt testler için mükemmel bir test-tekrar test tekrarlanabilirliği gösterdi.
Sonuç: Burada sunulan EAT yüksek tekrarlanabilirliğiyle birlikte dilin bir kaç yönünün değerlendirilmesine olanak tanımaktadır. Test demografik verilerden etkilendiğinden, hastanın özelliklerine dayanarak test skorlarında düzeltmeler yapılmalı ve ülkenin sosyo-kültürel gerçeklerinin kesin göstergeleri göz önünde bulundurulmalıdır. Bundan sonraki çalışmalarda, afazik hastalar üzerinde testin tanısal geçerliliği ve kabul edilebilirliği çalışılmalıdır. *Türk Fiz Tıp Rehab Derg 2007;53:5-10*
Anahtar Kelimeler: Afazi, afazi değerlendirme testi, normal veri, Türkçe

Introduction

One of the greatest difficulties faced by Turkish health professionals in the fields of neurological rehabilitation is the absence of an accurate and validated test to evaluate aphasia, although its importance is well known. There are mainly two reasons causing this difficulty. First, the Turkish language that belongs to Ural-Altaic tongues considerably differs from the Indo-European language family. The fundamental differences between Turkish and Indo-European languages can be summarized as

follows; vowel harmony, the absence of gender, adjectives preceding nouns, verbs taking place at the end of the sentence and agglutination. For instance; in contrast to the English language, the verb is never expressed solely in Turkish, because all the verbs have a suffix that determines the subject. Thus the verb is expressed together with the subject just in one word. In other words, because of the agglutination, a sentence consisting of five to six words in English can be translated into Turkish just in one word. In the Turkish population, the degree of suffixes' loss determines the language abilities' loss. Because of these fundamental differences, translated forms of currently available aphasia evaluation tests would not be adequate for the Turkish language, since phonetic, morphological and syntactic structures of the used language have to be similar with the English language. Second, as the use of language is considerably influenced by demographic variables, particularly education, the test should be adequate for socio-cultural reality of Turkey (1,2).

These facts promoted us to develop a completely new aphasia test in Turkish language, compatible with the socio-cultural context of Turkey. The present study describes the development of this test, named Ege Aphasia Test (EAT). Performance of the EAT in Turkish population and impacts of the demographic influences such as age, gender and educational levels on the scores of the various subtests, and reliability of the test will also be reported.

Materials and Methods

Development of the EAT

The development of the EAT was held in three steps. In the first step, rehabilitation specialists, according to the expected goals of the EAT, defined the subsets. Then, a panel consisting of a Turkish linguist, a physicometrist and a rehabilitation specialist constituted the items of the subtests. Panel members gave special attention to the use of proper and understandable language for each item. In the third step, the test was presented to a different panel of 3 experts (2 blind specialist of rehabilitation, 1 blind specialist of public health) for relevance of both receptive and expressive language modalities.

The final version of the EAT consisted of nine subtests: Spontaneous speech, praxis, understanding what is heard, production of language, repetition, naming, reading and understanding what is seen, writing and picturing, calculations.

Spontaneous speech was evaluated by seven items that comprise simple to complex questions. An 8-point scoring system was used, with score 8 indicating that there was no answer, and score 0 indicating that there was normal and understandable speech. Praxis was evaluated in four categories: facial, upper limbs' functions, use of machine and complex. A 4-point scoring system was used in each command: Score 0 indicating the approximate performance of the command, score 1 indicating the failure to perform the command but approximate performance on imitation, score 2 indicating the failure to perform the command and imitation but approximate performance with object, score 3 indicating that there was no answer. A total score range for the praxis was 0 to 60. Understanding what is heard was evaluated in six categories. In these categories; test cards, yes/no questions, real objects and a simple paragraph were used as a stimuli, and different scoring was used depending on the response of verbal, gestural or eyes movement. In this subtest, one repetition of each command was allowed for all items, each repetition of command

has gained 1 point. A total score range was 0 to 150. Production of language was evaluated in six categories, which comprise vowels, syllable with two letters, syllable with three letters, idioms, numbers and days of week. A total score range of this subset was 0 to 91. Repetition was evaluated by asking the respondent repeat the words and sentences. One command repetition of each item was allowed for all, each repetition of item has gained 1 point. The score range was 0 to 74. Naming was evaluated in eight categories. In these categories; real objects, incomplete sentences, test cards and verbal questions were used as stimuli. A total score range for naming subtest was 0 to 190. Reading and understanding what is seen was evaluated in six categories. In these categories; geometric figure cards, letter cards, picture cards, word cards, yes/no questions and a simple paragraph were used as stimuli. The reading was assessed by six-point scoring, with score 0 indicating that there was normal and understandable reading and score 6 indicating that there was no reading, giving a total score of 0 to 208. The subtest of writing and picturing was evaluated in seven categories. In these categories; verbal and visual stimuli were used as stimuli, in which the respondents were asked to write the letters, words and sentences and to picture the geometric figures. The score range for this subtest was 0 to 86. The subtest of calculation was evaluated in 4 categories that include calculation of hour, money and simple arithmetic and complex arithmetic process. The score range for this subtest was 0 to 30.

Subjects

EAT was applied to 133 volunteers (80 women and 53 men), aged between 20 and 80 years, after they gave their written consent. They were recruited among hospital employees and inpatients' and outpatients' relatives or caregivers. Individuals were excluded if they had taken any drugs known to affect mental status, if they had neurological illness, major or minor brain injury and other health problems (psychiatric disease, disease known to affect vision, alcohol abuse). All the subjects were native Turkish speakers living in the Aegean Region, although they were from different parts of Turkey. The participants completed the EAT at home or at our department by the help of a physician.

In order to evaluate influence of education on the performance of the test, the respondents were divided into four groups according to their educational level: Illiterate: no schooling, no reading-writing or incompletion of elementary school; Low: completion of elementary school (5 to 8 years of schooling); Mid: completion of high school (11 years of schooling); High: university degree (at least 15 years of schooling). To evaluate the effect of age on the performance of the test, three age groups were formed: young: between 20 and 39 years; middle-aged: between 40 and 59 years and old: between 60 and 80 years.

Of the subjects, twenty-five were re-tested after 1 to 3 months interval by the same physician and at the same place, in order to collect test-retest reliability data. Their characteristics were similar to those of the overall sample in terms of age, gender and educational level.

Statistics

Data were statistically analyzed by using SPSS/PC V10.0. The subgroups regarding age and educational level were compared for the performance in the various subtests by using Kruskal Wallis test. A level of 0.05 was accepted as significant. If a difference was determined among the subgroups, then Post-Hoc multiple comparisons were performed to determine which means

differ. The comparison of the performance of the subgroups regarding gender was made by independent samples t test. Correlations between age, gender, educational level and test performance were computed by Pearson's correlation analysis. Reliability of the test was evaluated by determining the intraclass correlation coefficient (ICC) between the two measurement points.

Results

Figure 1 and 2 represent the distribution of age and gender of the respondents regarding their educational levels. Comparison of the age groups showed that, young respondents had significantly higher education level than the older group (p<0.05), while no significant differences existed between the middle-aged and older respondents. Educational level of men was significantly higher comparing to that of women (p<0.05). No differences were found for the distribution of gender between the age groups.

Statistical analyses were not performed for the subtests of spontaneous speech, production of language and repetition, because all the scores were zero for these subtests in all respondents. Thus, data of the remaining subtests will be given here (praxis, understanding what is heard, naming, reading and understanding what is seen, writing and picturing and calculations). Comparison of performance of the remaining subtests among the subgroups of education revealed significant differences between the subgroups in all subtests and overall test (p<0.05, Table 1). Group comparisons revealed that, illiterate respondents had significantly lower performance than the respondents with more education in these subtests. The respondents with low educational level had significantly lower performance in these subtests as well, comparing to those with mid and high educational level (p<0.05). Significant differences were found between the respondents with mid and high educational level for the performance of reading and understanding what is seen and calculation; giving significant differences in the overall test. When the items of these subtests were analyzed, it has been found that these differences were due to items involving paragraph reading and complex calculations. Correlation analysis revealed that an increase in the educational

level corresponded with an increase of the performance of all subtests (r values are between 0.47 and 0.61, p<0.05, Table 3).

Analysis of performance of the test among the age groups revealed that the younger group had significantly higher performance in all of the subtests comparing to the old group (p<0.05), and in the subtests of reading and understanding what is seen, calculation and overall test comparing to the middle-aged group (p<0.05) (Table 2). The performance of the middle-aged group was significantly higher than the old group for the subtests of praxis, reading and understanding what is seen and overall test (p<0.05) (Table 2). Correlation analyses showed statistically significant negative correlations between the age and performance for all subtests and overall test (r values are between 0.15 and 0.29, p<0.05, Table 3).

No significant differences were found between men and women in all subtests except praxis, the performance of which were found to be significantly higher in men comparing to women (p<0.05).

EAT showed excellent test-retest reliability for the overall score (ICC= 0.99). Test-retest correlation coefficient ranged from 0.84 for praxis to 0.99 for naming, reading and understanding what is seen, writing and picturing.

Discussion

Here we reported a new Turkish aphasia test, which we have developed as a guide to assess all aspects of language abilities, including spontaneous speech, speech fluency and speech output, auditory comprehension, repetition, naming, written output, reading comprehension, picturing and praxis.

Several standardized tests are commonly used for evaluating aphasia (3-5). EAT, as far as the structure and objectives of the testing are concerned may be likened to these tests.

EAT application begins with an assessment of fluency, vocal quality and loudness, and the pronunciation and clarity of speech. At the same time, examiner also assesses strength and coordination of the speech muscles before evaluating comprehensive and expressive abilities. Comprehension is evaluated by sampling different types of language skills by using
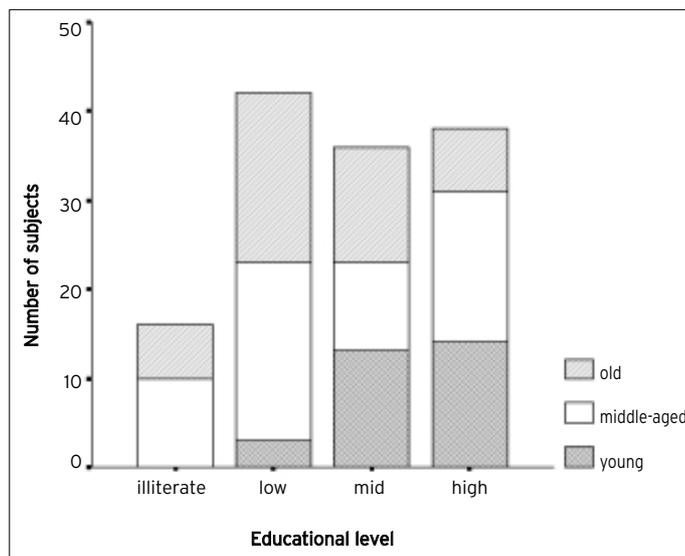

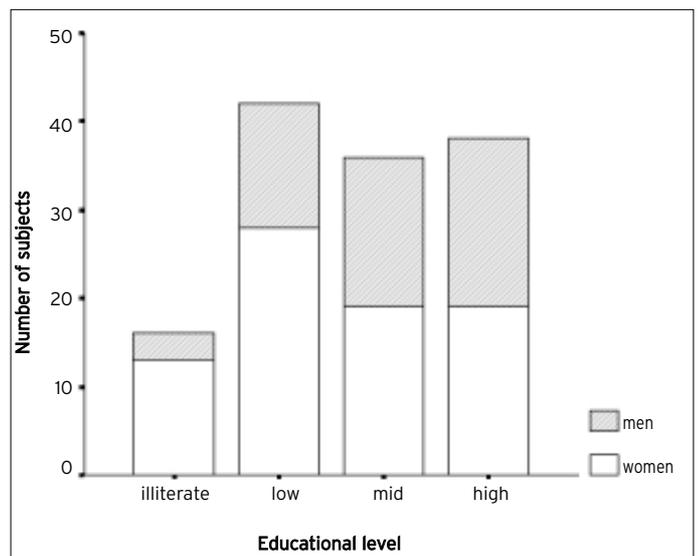Figure 1. The distribution of subjects in groups by age and education level.

Figure 2. The distribution of subjects in groups by gender education level.

auditory and visual inputs. Stimuli that require a definite action as a response are used in order to assure the analysis of oral and written comprehension. Production of language is evaluated by repetition, oral agility, naming and oral reading. Expressive functions are also measured by some subtests regarding vocabulary (semantics) and grammar (syntax). Semantic categories include nouns related to objects, shapes, numbers and letters. The test considers not only nouns but also transitive and intransitive action verbs. EAT additionally provides evaluation of pathologies frequently associated with aphasia, such as apraxia and dysarthria. The scoring system allows the examiner to differentiate responses elicited with various types of errors. Production of language is scored in a more rigorous way, considering qualitative data, answer delay and sensitivity to phonemic and semantic clues.

EAT is similar to Minnesota Test from standpoint of providing a wide sampling of potentially adequate types of stimuli, of situations in which responses can be elicited, and of the kinds of circumstances in which a patient can use language successfully. EAT seems similar to the Boston Diagnostic Aphasia Examination (BDAE) as well, due to its objective basis for the identification and differentiation of aphasic syndromes.

The main purpose of an aphasia evaluation test is to identify the presence of language impairment. However, as the use of language is known to be considerably influenced by demographic variables such as age, gender and particularly education, the performance of some individuals may overlap that of the patients with cerebral disorders (6,7). Thus providing normative data for the test target population is necessary before applying the test to the patients, to be able to provide normal-pathologic boundaries. We therefore attempted to apply the test to the normal population to provide normative data, before examining its validity and acceptability in aphasic patients. Indeed, the distribution our respondents regarding their age, gender and educational levels reflects the socio-cultural reality of Turkey (Figure 1 and 2). Low education or illiteracy is very common in the elderly individuals, especially in the elderly women. Because of the educational policies in the last couple of decades, people, particularly women are getting more educated, thus difference between men and women is getting decreased. As a result, young individuals regardless of their gender have relatively more education comparing to the older individuals. Considering these distinct features, examining the possible influences of demographic variables on the test performance is even more important in Turkish population.

We have found that, the demographic variable most strikingly influenced the performance of the EAT was the educational level. Illiterate individuals have the poorest performance in almost most of the subtests, supporting the previous studies demonstrating the compromising effect of illiteracy on different aspects of language, including oral comprehension, short-term memory, grammatical judgment and meta-phonological analysis (8,9). The influence of low educational level was to a lesser degree comparing to that of illiteracy, but still significant comparing to that of higher educational level. The impacts of low educational level have been investigated in a few studies involving different versions BDAE. The scores have been found to be impaired especially among subjects with less than 9 years of education (10,11). A more recent study compared the performance of the subjects with less and more than 8 years of schooling. The authors verified that such a division could not exert the greatest impact on subject performance and emphasized the necessity of studying the subjects with less than 8 years of education in more detail. The authors also noted that the correlation between educational level and test performance does not occur linearly; the interval from 0 to literacy obviously has much more impact than the interval from 6 to 8 years and so on (12). Our

Table 1. Subtests of Ege Aphasia Test: mean scores for educational subgroups.

Subtests	Education Level				
	Illiterate mean±SD (min-max)	Low mean±SD (min-max)	Mid mean±SD (min-max)	High mean±SD (min-max)	p
Praxis*	1.1±0.6 (0-2)	0.42±0.5 (0-2)	0.1±0.3 (0-1)	0.0±0.0 (0-0)	0.00
Understanding what is heard*	6.5±3.8 (1-13)	2.7±1.8 (0-8)	1.2±1.3 (0-6)	1.3±1.3 (0-6)	0.00
Naming*	16.3±10.7 (4-34)	2.9±3.6 (0-18)	1.4±1.4 (0-6)	0.6±0.9 (0-2)	0.00
Reading and understanding what is seen*	92.6±84.1 (2-196)	3.9±2.5 (0-9)	2.1±1.5 (0-8)	1.5±1.3 (0-6)	0.00
Writing and picturing*	35.6±27.3 (1-89)	4.1±5.0 (1-27)	1.0±1.0 (0-5)	0.6±0.4 (0-1)	0.00
Calculations*	10.8±6.6 (2-22)	2.9±3.6 (0-15)	0.9±1.9 (0-7)	0.05±0.22 (0-1)	0.00
Total*	163.2±119.7 (22-307)	17.1±12.5 (2-48)	6.7±5.0 (1-22)	4.0±2.9 (0-11)	0.00
*p <0.05, comparisons were made with Kruskal Wallis					

findings strongly support this study, since the performance of the subjects with mid and high educational level did not considerably differ from each other except for the items that required high level of education. It should be noted that, educational level may not have such an important impact in the aphasic patients, likely due to masking effect of the cerebral lesions (13), thus controlled clinical trials to demonstrate this relationship in the Turkish aphasic patients are needed.

Detrimental impacts of aging on language, especially on word discrimination, naming, oral sentence reading and repetition of high and low-probability phrases have been reported previously (6,7,11,14,15). Supporting these studies, our study showed that, the young individuals have the best performance on the EAT and detrimental impacts occur on different aspects of EAT with increasing age. Considering the highest educational level of the young subjects in our population, coexisting effect of educational level could not be ruled out as has been emphasized previously (10,12,16).

Women had poorer results on the subtest of praxis comparing to men. This difference was likely due to low acceptability of the item about driving in the women –particularly in elderly women who had low education- who do not drive for socio-cultural reasons. No differences were found between men and women for the remaining subtests. Although influence of gender on language remains controversial in the literature, our findings are contrary to the previous studies demonstrating higher performance of the women comparing to men (13,17). It should be noted however that, educational level of the women was considerably lower than men in our population, thus coexisting effect of low educational level factor should be considered. Larger sample sizes would permit the use of statistical regression techniques to ensure that the effect of the education level was independent from aging and gender.

According to the American Psychological Association (APA) guidelines, aphasia test manuals would include test-retest reliability

Table 2. Subtests of Ege Aphasia Test: mean scores for age groups.

Subtests	Age Groups			
	Young mean±SD (min-max)	Middle mean±SD (min-max)	Old mean±SD (min-max)	p
Praxis*	0.06±0.25 (0-1)	0.2±0.5 (0-2)	0.5±0.6 (0-2)	0.00
Understanding what is heard*	1.1±1.5 (0-6)	2.6±2.9 (0-13)	2.8±2.4 (0-11)	0.01
Naming*	1.2±2.4 (0-12)	3.9±7.6 (0-34)	4.2±6.5 (0-33)	0.09
Reading and understanding what is seen*	2.0±2.0 (0-8)	14.8±45.6 (0-196)	19.4±47.3 (0-176)	0.18
Writing and picturing*	1.2±2.2 (0-11)	5.7±11.0 (0-53)	9.8±21.3 (0-89)	0.04
Calculations*	1.5±3.4 (0-15)	2.1±4.6 (0-21)	3.5±5.1 (0-22)	0.12
Total*	7.1±9.7 (0-42)	29.6±68.5 (0-307)	40.5±78.6 (2-286)	0.09

*p<0.05, comparisons were made with Kruskal Wallis

Table 3. Correlation results (Pearson's r) between Ege Aphasia Test subtests and demographic variables (age and education level).

Subtests	Age	Education
Praxis	0.29**	-0.58**
Understanding what is heard	0.22**	-0.57**
Naming	0.16*	-0.57**
Reading and understanding what is seen	0.15*	-0.47**
Writing and picturing	0.21*	-0.54**
Calculations	0.17*	-0.61**
Total	0.18*	-0.55**

*p<0.05, ** p<0.01

data (17). EAT has high reliability with ICC of the test-retest subtests scores greater than 0.84.

Here we presented EAT developed to evaluate Turkish patients with aphasia. The test allows evaluation of several aspects of language related to comprehension and production processing, quantification of the intensity of impairment and a qualitative analysis of functional interactions among different language system components. As the main purpose of the test is to detect the presence or absence of aphasia, its diagnostic validity and acceptability must be determined in further studies on the patients with different kinds of aphasia. Normative data demonstrated that EAT is influenced by demographic variables, particularly education, thus adjustments in the test scores based on a particular patient's age, education or gender should be made and precise indicators of socio-cultural reality of the country should be considered.

Acknowledgments

References

1. Özkan M. "Türkçemiz ve Uydurmacılık". Türk Edebiyatı 1992;2:18-9.
2. Öner M. "Ortak Türk Alfabesi" Hakkında Notlar, Yeni Türkiye-Türk Dünyası Özel Sayısı 1997;3:207-11.
3. Goodglass H, Kaplan E. The assessment of Aphasia and Related Disorders. Philadelphia: Lea & Febiger; 1972.
4. Goodglass H, Kaplan E. The assessment of Aphasia and Related Disorders, 2nd ed. Philadelphia: Lea & Febiger; 1983.
5. Wertz RT. Neuropathologies of speech and language: An introduction to patient management. In: Johns DF, editor. Clinical management of neurogenic communitive disorders, 2nd Ed. Boston: Little Brown and Co.: 1985. p. 23.
6. Ross KB, Wertz RT. Possible demographic influences on differentiating normal from aphasic performance. J Commun Disord 2001;34:115-30.
7. Le Dorze G, Bédard C. Effects of age and education on the Lexico-semantic content of connected speech in adults. J Commun Disord 1998;31:53-71.
8. Roselli M. Neuropsychology of illiteracy. Behav Neurol 1993;6:107-12.
9. Roselli M, Ardila A, Rosas P. Neuropsychological assessment in illiterates. II. Language and praxic abilities. Brain Cogn 1990;12:281-96.
10. Ardila A, Rosselli M. Spontaneous language production and aging: sex and educational effects. Int J Neurosci 1996;87:71-8.
11. Pineda DA, Rosselli M, Ardila A, Mejia SE, Romero MG, Pérez C. The Boston Diagnostic Aphasia Examination-Spanish version: Influence of demographic variables. J Int Neuropsychol Soc 2000;6:802-14.
12. Radanovic M, Mansur LL. Performance of a Brazilian population sample in the Boston Diagnostic Aphasia Examination: A pilot study. Braz J Med Biol Res 2002;35:305-17.
13. Duffy JR, Keith JC. Performance of non-brain injured adults on the PICA: descriptive data and a comparison to patients with aphasia. Aphasia-Apraxia-Agnosia 1980;2:1-30 (abstract).
14. Rosselli M, Ardila A, Florez A, Castro C. Normative data on the Boston Diagnostic Aphasia Examination in a Spanish-speaking population. J Clin Exp Neuropsychol 1990;12:313-22.
15. Shewan CM, Henderson VL. Analysis of spontaneous language in the older normal population. J Commun Disord 1988;21:139-54.
16. Deloche G, Hannequin D, Dordain M, Perrier D, Pichard B, Quint S, Metz-Lutz MN, Kremin H, Cardebat D. Picture confrontation oral naming: performance differences between aphasics and normals. Brain Lang 1996;53:105-20.
17. Flanagan JL, Jackson ST. Test-Retest Reliability of Three Aphasia Tests: Performance of Non-Brain-Damaged Older Adults. J Commun Disord 1997;30:33-40.