



---

# Verbal Fluency Test: A historical-critical review of the concept of fluency

## Teste de Fluência Verbal: Uma revisão histórico-crítica do conceito de fluência

### Test de Fluidez Verbal: una revisión histórico-crítica del concepto de fluência

Ana Paula de Oliveira Santana\*  
Karoline Pimentel dos Santos\*

#### Abstract

Verbal Fluency Test (VFT) is one of the most used tests in neuropsychology, although the concept of fluency that underlies the test has been rarely discussed. In this way, this study aims to present a historical-critical review of VFT, its genesis and modifications until now, in light of Discursive-Enunciation Neurolinguistic. The review points the concept of fluency in VFT shows to be more related to a verbal quantitative production perspective than to a linguistic process comprehension that allows this production. Also, there is a conceptual change of fluency concept in neuropsychology, so that fluency is no longer considered as an isolated intelligence factor, to be considered a task that mobilizes other cognitive activities such as memory, attention, vocabulary, among others. In this sense, VFT has not been focusing on linguistic fluency but on cognitive fluency through verbal production.

**Keywords:** *Language; Evaluation; Neuropsychology..*

#### Resumo

*O teste de fluência verbal (TFV) é um dos mais utilizados na área de neurolinguística e neuropsicologia, contudo, pouco se tem analisado o conceito de fluência que subjaz à sua elaboração. Desta forma, o objetivo desta retrospectiva é analisar histórica e criticamente o TFV, sua criação e modificações até os dias atuais, à luz dos pressupostos que norteiam a Neurolinguística Enunciativo-Discursiva. A análise aponta que o conceito de fluência parece sustentar-se mais pela perspectiva da produtividade linguística do que pela compreensão das vias que possibilitam tal produtividade. Observa-se, ainda, a ocorrência*

---

\* Universidade Federal de Santa Catarina.

**Conflito de interesses:** Não.

**Contribuição dos autores:** KPS - concepção, elaboração textual, revisão de literatura, análise crítica e discussão teórica; APOS - concepção, elaboração textual, revisão de literatura, análise crítica e discussão teórica.

**Endereço para correspondência:** Ana Paula de Oliveira Santana -

**E-mail:** anopasantana@hotmail.com

**Recebido:** 29/05/2015 **Aprovado:** 14/10/2015



*de uma mudança conceitual em que a fluência deixa de ser considerada como um fator de inteligência isolado para ser uma tarefa pela qual se mobilizam outras atividades cognitivas como memória, atenção, vocabulário, dentre outras. Nesse sentido, desloca-se a avaliação da fluência linguística para uma avaliação da fluência cognitiva por meio da produção verbal.*

**Palavras-chave:** Linguagem; Avaliação; Neuropsicologia.

### **.Resumen**

El test de fluencia verbal (TFV) es uno de los más utilizados en el área de neurolingüística y neurosiología, aunque poco se ha analizado el concepto de fluencia que subyace en su elaboración. De esta forma, el objetivo de esta retrospectiva es analizar histórica y críticamente el TFV, su creación y sus modificaciones hasta los días actuales, a la luz de los presupuestos que nortean la Neurolingüística Enunciativa-Discursiva. El análisis apunta que el concepto de fluencia parece sustentarse más por la perspectiva de la productividad lingüística, de que por la comprensión de las vías que posibilitan tal productividad. Se observa aún, que hay un cambio conceptual en que la fluencia deja de ser considerada como un factor de inteligencia aislado, para ser una tarea por la cual se mobilizan otras actividades cognitivas como memoria, atención, vocabulario, entre otras. En ese sentido, dislocase la evaluación de la fluencia lingüística para una evaluación de la fluencia cognitiva por medio de la producción verbal.

**.Palabras clave:** Lenguaje; Evaluación; Neuropsicología.

### **Introduction**

In Neuropsychology, the verbal fluency test is widely used in the evaluation of subjects suspected of having neurological damage, especially in patients with possible diagnosis of Alzheimer's dementia (AD), because task execution as proposed in the test requires the integrity of cognitive skills such as memory and executive functions<sup>1,2,3,4,5,6</sup>. The concept of fluency, however, is little discussed and more geared towards the idealization of the Verbal Fluency Test (VFT) than to the fluency phenomenon itself.

Back in 1986, there were writers who drew attention to the importance of testing fluency in assessment batteries<sup>7</sup>. Compared with other language tests, the weight of this test is very relevant and may even define, in extreme cases, the patient's pathology, although the psychometric properties of the test are not clear. However, the issue of asymmetry between the weights of the test becomes more worrisome when the fluency ability is not well defined. Therefore, one may wonder: "what is fluency?"

Fluency as measured by the VFT is fluency in speech that refers to neurological aspects. In this case, "fluency is an executive process characterized by the subject's ability to perform a series

of behaviors within a framework of established rules which may be verbal or nonverbal." [our translation<sup>8</sup>.

Thus, in Neuropsychology, fluency is not limited to linguistic aspects, but expands to other disciplines such as painting, logic, music, etc. All of these modalities, including language, relate to cognitive functioning, more specifically, the executive function, their integrity and/or capacity. Thus, fluency is thus understood both as an intelligence factor and as a creativity factor.

Therefore, the characteristics involving fluency measurement in the VFT are quite different from studies that measure fluency in stuttering, for example, as they seek to determine its severity and measure speech disruptions, which is the main complaint of the stuttering subject<sup>9</sup>. Moreover, fluency in the VFT seems to be regarded as a task triggers other cognitive activities such as memory, attention, vocabulary, etc. Therefore, some questions have arisen: Why is the test called VFT and how was it created? What is the relationship established between language and cognition, in that a low score on the fluency test indicates a cognitive deficit? While considering these questions, this paper is aimed at making a historical and critical analysis of the VFT, its creation and modifications to date, as well as its implications for language assessment.



## The First Thurstone World Fluency Test: a historical context

The verbal fluency test appeared in a scientific context where the usefulness of intelligence testing as a cognitive predictor was being questioned. In 1938, Thurstone considered IQ tests to be limited, to a large extent, since they did not measure the subject's expertise, his degree of mental abilities on specific tasks. According to the author's argument, two people could have the same intelligence quotient (IQ) - and yet perform the same task in qualitatively different levels<sup>10</sup>.

Thurstone believed in the need to understand the different mental abilities, which he referred to as factors. For the author, it was essential to measure these factors separately to understand the varied performance of the same patient on specific tests. The factors were divided into primary (superior and independent of one another) and secondary (general, activity parameters of primary factors), and language is to be characterized as the most important primary factor<sup>11</sup>. Based on this understanding of mental abilities, the author sought to answer several questions: Why did the same subject do well in a language test but not in another? What were language tests actually measuring? What do they require from subjects? Understanding these factors could not only contribute to the composition and administration of the tests, but also lead to a more detailed interpretation of particular injuries of subjects affected by language pathologies.

The author gave language a distinctive place over other cognitive functions, independently of other functions as well as of its own aspects: comprehension and expression. Language was understood as a complex skill characterized by two verbal factors: the "V Factor", relative to the ease of verbal comprehension, and the "W factor", referring to word fluency skills. Later, Thurstone also worked with another verbal factor, F, for ease of expression of ideas. The V and W factors are clearly distinguishable, since a type of verbal test, corresponding to the W factor, requires the subject to produce words without necessarily knowing their meanings while the other type of test, corresponding to the V Factor, requires the subject to understand the words that he sees. As these factors are independent, a person may have a high score in a comprehension test, which is indicative of high

verbal reasoning capacity, and yet, have a low score in word fluency. Likewise, another person may be very fluent (i.e., produce many words), even if they have a limited vocabulary.

Based on these statements, we may start to reflect on the concept of fluency contained in Thurstone's test. It can be seen that this concept was devoid of any commitment to the notion of semantic content of the words. This is because the intention was to measure a set of verbal skills alone; therefore, it did not matter much that the focus of analysis was on productivity, because when applied in combination, these factors (V and W) complemented each other with the flow and understanding of the verbal material.

Thus, the Thurstone Word Fluency Test, based on the skill concept of the W factor, does not intend to measure semantic association, verbal comprehension or speaker's vocabulary range. For Thurstone, fluency, represented by the W factor, was an independent verbal component, liable to evaluation apart from other factors, including motor ones. Thus, as previously mentioned, fluency in the Thurstone Fluency Test is understood through the concept of rate of word production, delimited by the rules inherent in the test itself and which includes, among other criteria, the presence of the letters S and C at the beginning of the word. However, it should be noted that, although the score is based only on the number of produced items, productions outside the preset criteria are not scored; thus to a limited extent, the idea of information flow is implicit in the concept of fluency in the test. In this context, the "information flow" is more related to the flow of mental ideas and the ability to express these ideas than to speech, articulation, fluency, pace (aspects which may be related to the analysis of fluency from a linguistic perspective).

The Thurstone Word Fluency Test consists of two phases. In the first phase, the subject is asked to write, in five minutes, as many words beginning with the letter S as possible. In the second, he is asked to write, in four minutes, as many four-letter words beginning with the letter C as possible. It is believed that the first letter, S, has greater lexical association, unlike letter C, whose lexical association is believed to be low<sup>12</sup>. Both phases are scored similarly; each word is worth one point, and the test score is the sum of all the right words, each worth 1 point, regardless of the parts, A or B. Repeated words are not considered in the score.



the examiner requested the subject to read some unclearly written words and there was no identification between the spoken word (intended) and the written word, yet the written word would score one point, because the intent of the subject would have met the criteria established in the test.

As regards the use of writing, it can be said that the hypothesis underlying this test, considering its assessment perspective, is that writing was conceived as a direct transcription of thought and speech. This idea leads to a conception from the nineteenth century, in which writing and internal language/thought were seen as a simulacrum. Even considering an intrinsic relationship between language and thought, this relationship is not one of representation. The reality of the inner psyche is the sign. However, if the internal language and the external language have the same symbolic nature, that does not mean that thought and language mirror each other, but rather, they maintain a non-direct, non-univocal, non-transparent relationship with each other<sup>13</sup>.

The main point of this discussion is the assumption that internal language problems can be represented in writing. This seems to be the theoretical assumption that underlies the Thurstone test: the direct relationship between cognitive domain and writing production. Although, at the time, these discussions were not made, the test was criticized and its mode of expression was modified, particularly, to orality.

### Changes in the Thurstone Fluency Test

A The latest revision of the Thurstone Word Fluency Test was performed by Heaton et al in 1991. Ten age categories and six education and sex categories were added; in 2004, the race / ethnicity variable (black and white Americans) was added. However, the original version itself already had several issues that limited the application of the Thurstone Word Fluency Test, including low or lack of education in the population (also children at preschool age), patients with post-stroke movement disorders, chronic diseases, and impaired motor skills in normal aging. This scenario led Arthur Benton to prepare the oral version of the fluency test, in 1962, under the name of Controlled Verbal Fluency Task (CVFT), better known as FAS. In addition to tailoring the test to weaker and less

educated subjects, Benton reduced the length of the test from 9 to 3 minutes.

In this version of the test, the subject is asked to say as many words as possible, within a period of 60 seconds, that begin with each letter (F, A and S). The total score is given by the sum of all correct words beginning with the three letters. Slang and foreign words are considered to be correct. Proper names and repeated words are not taken into consideration.

In 1967, the CVFT became a component of the Neurosensory Center Examination for Aphasia, and, nine years later, it was included in the Multilingual Aphasia Examination Battery, under the name Controlled Oral Word Association Test (COWA) and replacement of the letters FAS with CFL or PRW.

The name change meant a better representation of the test procedure, which intends to shift from the centralization of productivity to the association between words, although production was still the only way to score in this evaluation<sup>14</sup>. Changing letters F, A and S, randomly defined, was justified by the frequency criterion of dictionary words in the English language, as adopted by the authors. At first, there was an attempt to establish an alphabetical hierarchy. Except for X and Z, which showed very low frequency of words in the dictionary, all the letters of the alphabet were evaluated according to their frequency in the beginning of dictionary words. Thus, as a resolution, the letters were divided into three levels of difficulty: easy (H, D, M, W, A, B, F, P, T, C, S), moderate (I, O, N, E, G, L, R) and difficult (Q, J, V, Y, K, L)<sup>15</sup>. Based on these definitions, the authors finally chose two alternative combinations of letters - CFL and PRW - that the examiner could choose for test application, wherein C, F, P and W have higher frequency, and are, therefore, at an easy level, whereas L and R are at a moderate level<sup>16</sup>. Once the criterion for letter choice had been established, the test could make a more precise distinction, from a linguistic point of view, between normal and pathological cases.

However, it should be noted that the COWA is based on graphemes instead of phonemes, which is contradictory to some extent, because it was also intended for application to uneducated or poorly educated subjects, who have lesser or no knowledge of the artificially acquired language code.

Despite the change, the test continued to be used under the same version in the Neurosensory

Center Examination for Aphasia, thus competing with the modified version, COWA. Following the proposed frequency orientation, the three letters of the CVFT, in this case F, A and S, are considered to be at an easy level, which possibly may involve different results compared with COWA, which uses a moderate frequency level letter in each of the letter combinations.

In 1996, with support from Benton himself, other researchers<sup>17</sup> reviewed the COWA and proposed new normative values that could better reflect the current profile of the population, quite different from the one considered when developing the latest versions of test. In these versions, the tested population belonged to rural areas and had a relatively low education level, different from the profile of the population that was contemporary to the authors. In this way, the aim was to update the normative data of the test, proposing a more heterogeneous sample of the population, with 360 normal subjects, divided by educational level, sex and age. The comparisons between the results of the original sample and the current sample, within a twenty years' interval between each other, indicated an increased score - cut off point - to the test, which the authors interpret as an expression of the directly proportional relationship between education and performance in the test.

From the 1980s onwards, test adaptations were performed in many countries, by replacing the required letters and/or changing the test score. These adaptations were designed to meet the specificity of the language and population profile where the test was to be applied. The frequency of occurrence of the letters was taken in consideration, in terms of being different according to the language of use.

In Brazil, no adjustments were made into Portuguese for the criterion of letter frequency; instead, the original version of the oral test was selected, with the letters F, A and S, widely used in neuropsychological evaluations in patients with cognitive damage and complaints. Regarding the population profile, FAS, also known as Phonological Verbal Fluency Test, in 345 normal elderly subjects aged above 60 years, with different education levels (1-3, 4-7, 8-11 and 12 or more years of schooling). The survey results indicated schooling as the most relevant factor in the test results. Although the authors have not submitted "cut off" point for the Brazilian population, they suggested the use of means, determined according

to age and education level of the research subjects, as a parameter for evaluation. It should be noted, however, that the proposed mean, as a statistical result of this research, is accompanied by a considerable standard deviation (SD) in each of the age groups, which casts some doubt on the reliability of using these means in the evaluation<sup>18</sup>.

Other versions of the test were prepared by changing the criterion of initial letters to the formation of semantic categories, nomination of action verbs, and alternation between two categories given by the examiner.

In the semantic mode, the words to be spoken must belong to a specific category, predetermined by the examiner. The most widely used category is animals, because it is believed to be less vulnerable to the education variable. However, other categories are possible, for example: fruits, vegetables, professions, boys' names, girls' names, girls' clothes, people's names, musical instruments, food, furniture, types of transport, car parts, items found in supermarket, things that people drink, tools, clothes, a country's states, inanimate objects, people's names (same sex of the examined subject), things that make people happy, things that make people sad, things that people drive. Tests with similar tasks were also developed, for example, the Homophone Meaning Test (with homonyms), the Excluded Letter Fluency Task (generation of words that do not contain a specific vowel) and the Uses for Common Objects Task (thinking of alternative uses for everyday objects).

Even with the changes, the two modes, the phonological and the semantic ones, are indiscriminately called Verbal Fluency Test (VFT), or simply FAS, in the case of the phonological mode - and can be applied alone or as part of test batteries, as shown in the table below:

Table 1 shows that most tests use the list of animals as analytical categories. It also shows the amount of tests that use the VFT as a complement to cognitive assessment. It is part of the items that assess language. That is, eleven kinds of batteries using the fluency test; four of them are related to aphasia, indicating that the TFV falls within the scope of language assessment.

### **The semantic category of the Verbal Fluency Test**

In Brazil, the Verbal Fluency Test (TFV), animal category, was standardized in 1997<sup>19</sup>. The

**TABLE 1:** MAIN ASSESSMENTS USING THE VFT AND THEIR CATEGORIES OF ANALYSIS.

Fluência Semântica		Fluência Fonológica
Bateria	Categoria	Letras
Boston Diagnostic Aphasia Examination	Animais	
Woodcock Johnson III Tests of Cognitive Abilities	Comida, nomes de pessoas, animais	
Clinical Evaluation of Language Fundamentals	Animais, comida, profissões	
Consortium of the Establishment of a Register of Alzheimer's Disease	Animais	
Delis Kaplan Executive Function System	Tradicional: Animais e nomes de meninos VA*: roupas e nomes de meninas;	FAS e BRH
	Alternância de categorias: frutas e móveis VA*: vegetais e instrumentos music.	
Kaplan Baycrest Neurocognitive Assessment	Animais e nomes de pessoas.	C
Multilingual Aphasia Examination		CFL e PRW
NEPSY: A development Neuropsychological Assessment	Animais e comida	SF
Neurosensory Center Comprehensive Examination for Aphasia		FAZ
Repeatable Battery for the Assessment of Neuropsychological Status	Frutas e vegetais	
7 Minutes Screen	Animais	
* VA: Versão alternativa		

Table designed by the authors, based on the works of Strauss and Mitrushina, Boone, D'elia<sup>16</sup>.

survey involved 336 subjects without neurological or psychiatric damage and 66 subjects with neurological damage caused by degenerative diseases and various cognitive deficits. The profile of individuals without neurological damage followed the following exclusion criterion: under 15 years of age; a history of epilepsy, stroke, neurosurgery, psychiatric complaints; mother tongue other than Portuguese; uncorrected visual or hearing impairment; impairment of motor functions for orthopedic and rheumatologic causes, which could affect performance in the test. The subjects were divided into four education levels - illiterate, up to 4 years of schooling (incomplete), 4-8 years old (incomplete) and 8 or more years of education - age - <64 years (young) and > or equal to 65 years (elderly). Sex was also considered as a test variable.

The conclusion of the research is that there is interference of education level in the performance of the Brazilian population, whose social profile is materially heterogeneous. This would imply the need to reduce the score compared with the one previously proposed in 1994 (18 points for normal population). Thus, these are the cutoff scores in Brazil (minimum grades expected of a subject without damage), with reference to education level: 9 for illiterates and up to 8 years of schooling (incomplete), and 13 for 8 or more years of schooling.

It should be noted that in the case of the “animals” category, each word is also worth one point, and the total score is the sum of all words. The names of endangered, imaginary or magic animals, e.g., unicorns, mermaids and dragons, are admitted, but proper names of animals are not taken into account. Likewise, variations of the same name, such as dog and bitch or lion and lioness, as well as repetitions, were not considered in the score.

In this semantic version, the meaning of words (contrary to what was proposed in the Thrustone Word Fluency test) is not only relevant but also necessary to carry out the task. Although the test score remains as the sum of all words produced, in view of the semantic-cognitive issues involved in the word list, there is a resulting tendency to automatic formation of categories<sup>20</sup>; therefore, qualitative criteria deemed as alternative are proposed in the test, influencing the interpretation of the results<sup>21</sup>.

Within this approach, the main components of the test that were analyzed are the cluster (forming

the subcategories) and switching (switching between these categories.) In the semantic category “animals”, for example, bee, fly and butterfly are a subcategory, while cat, dog and guinea pig are another. If this answer (1) were, in this order, the subject’s answer, there would be 2 clusters (two sub-categories) and 1 switching (an exchange between categories).

(1) Bee, fly, butterfly, cat, dog and guinea pig

If the subject had inserted an animal that did not belong to any of these categories, such as “lion”, the result would have been different, depending on where the animal was inserted. For example, considering the possible occurrences of “lion”

(2) **a**, bee, **b** fly, **c**, butterfly, **d**, cat, **e**, dog, **f**, guinea pig, **g**.

Traditionally, it is believed that the formation of a cluster requires three elements that characterize a category, usually from a taxonomic perspective. Thus, if the word “lion” occurred in positions a, d or g, the results would still be the same, with two clusters and one switching; however, if the occurrence were in positions b, c, e or f, the results would be different, forming one cluster and zero switching.

It should be noted, however, that the test score does not change; regardless of the position of occurrence of “lion”, the result is the sum of all the valid names - in this case, seven.

Some authors<sup>22</sup> suggested that the cluster should be formed from two animals instead of three, whereas “two” is enough to form one category. Still, others suggested that switching is considered as an exchange not only between subcategories - whose occurrence the author names as cluster switching - but also between individual words and between individual words and subcategories - occurrence named as hard switching - considering that switching would be evidence of mental shifting ability, which may occur even between individual words<sup>23</sup>.

Following these evaluation criteria, the analysis of the hypothetical subject’s answers can be considered again:

(1) bee, fly, butterfly, cat, dog and guinea pig

(2) **a**, bee, **b**, fly, **c**, butterfly, **d**, cat, **e**, dog, **f** guinea pig, **g**

The results of the qualitative analysis in (1) would be the same obtained according to traditional criteria; however in (2), the results are shown to be quite different: with the occurrence of a lion in a, d, and g, there are two clusters and two switchings; in the other instances, b, c, e or f, there are two clusters and three switchings.

Another possibility of forming categories but which, however, is not valid in cluster formation, is the phonological one, for example “goat, horse, dog.” Goat and horse belong to the same category, while “dog” is evoked by phonological association (/ k / in the Portuguese equivalents *cabra, cavalo, cachorro*) instead of semantic association, which is required for the formation of a valid cluster. The inverse criterion is valid in fluency by phonological category, such as FAS and COWA, since the relationship between words is not relevant in the test assessment.

The qualitative criteria correlated to the test results of the Brazilian population showed that subjects with higher education usually perform better, producing more answers and larger clusters<sup>24</sup>. Using the concept of prototype<sup>25</sup> for the category “animals”, it was found that prototypicality is established distinctly between more and less educated subjects: dog, for the most educated, horse for the illiterate, demonstrating - according to the authors - that the educated and uneducated have different cognitive strategies. The study in question is very important in that it corroborates the notion that the qualitative assessment of the test may contribute cues on the performance of the subjects.

Although it can be noted that the studies tend to deepen the effectiveness of the VFT and even understand the variables that can interfere with its application, the concept of fluency is not usually focused on language; the relationship established is restricted to the performance of cognitive functions activated during execution of the task proposed in the test. For example, the cluster is seen as an execution strategy, while switching is considered responsible for the mental search and flexibility. Thus, poor performance in phonological fluency is related to frontal lobe damage, while poor performance in semantic fluency is associated with temporal lobe damage. Studies using a spectroscope and research reviews with magnetic resonance<sup>26</sup> showed that the category task activates both the left and the right

prefrontal cortex. Other studies claimed<sup>27</sup> that the VFT involves the left prefrontal cortex and temporal lobes, both for the phonological and semantic categories. It is worth noting, however, that these relations are still under discussion.

Discussions on the relationship between language and cognition are still not held in depth in the tests. However, it is known that this relation, although not direct, is constitutive. The question lies, however, in the ratio attributed to the relationship between verbalization and functioning of language. The VFT seems to be more related to the assessment of cognitive integrity in general, than to language assessment. In this sense, language seems to have a more instrumental nature, that of representation of thought, as if “thought” could be seen through it.

Moreover, there are discussions in the literature on the linguistic and cognitive assessment performed by standard tests<sup>28</sup> that were often devised to measure verbal behavior related to brain injuries. This association is, to some extent, “problematic” primarily because it assumes a homogenizing character. Although the statistical standard of the tests allows a correct classification of diseases, via taxonomic evaluation, it does not allow the understanding of language effectively and does not affect the interpretation of the patient’s linguistic production. Based on the evaluation of this type of test, some inadequacies can be mentioned:

- Decontextualization of the proposed language tasks, simulating artificial situations for an alleged linguistic activity;

- Predominance of metalinguistic tasks that, while necessary for diagnosis, cannot substitute linguistic activities and the consideration of epilinguistic processes involved in the reconstruction of language by aphasic subjects;

- The fact that the nature of the tasks proposed matches exercises founded on written language, with strong educational commitment (in the worst sense of “school”), when they are not restricted to techniques to approach the phenomenon in order to raise facts required for the academic description of aphasia;

- Failure in empirical results: the reductionist theoretical perspective of the language phenomenon ultimately restricts the facts, precisely those that are not the most significant or the most relevant to characterize the language difficulties of aphasic



subjects and to provide support for follow-up (p: 06).

Two of these criticisms are particularly worth of notice, in the case of the VFT: decontextualization of the proposed task and comparability to school exercises. Except in special situations, it is not common, although possible, that one is prompted to say all the animals that come to mind or all clothes, or any other category in a given time, in everyday interaction.

It should also be noted that, in a dialog situation between examiner and patient, there is an asymmetrical relationship, in which the examiner has the answers, the language skills, the evaluation and judgment, and the patient, given a decontextualized - and therefore meaningless - task, has reduced face-saving skills, which affects his performance in the test. In addition, the established relationship eventually dismisses the speaker as the interlocutor; since this type of test is standard; it also ends up excluding the subject, his history and subjectivity, and aligns him to the single class of patient. Thus, interlocution ceases to exist to the extent that the interaction is replaced by responses elicited by artificial stimuli.

It should be noted, however, that although the subject does an atypical activity in terms of his everyday life, he attempts to signify and contextualize such activity, based on his interaction with the environment. For example, during an assessment that involves a list of animals, a patient can perform the task based on his memory from the zoo, and list his visit to the cages of the animals in a sequence. In this case, the patient uses a strategy to deal with the decontextualization of the proposed task, and evokes a mental scheme, necessary for executing the task, which confirms the importance of context in language tasks.

Another point to note is the relationship between the VFT and formal learning. The criterion of forming categories refers to the concept of a classic categorization that does not address the specificity of relationships established cognitively and consensually over the world. For example, "sardine, salmon and whale" cannot be considered as a category, since a whale is not a type of fish, taxonomically. In our culture, however, a whale can be categorized as such, perceptually and consensually. This view is based on the scientific concept proposed by Vygostky<sup>2</sup>, which is definitely

learned at school, as opposed to the spontaneous concept, categorization linked to the subject's social experiences.

Although Vygotsky has not established a hierarchy between these two concepts - the scientific concept and the spontaneous concept - common sense has it that the scientific concept, acquired at school, is the most favored, and the spontaneous concept turns out to be stigmatized. The criterion for formation of categories in the VFT reflects this hierarchy and disregards the situated social sphere where these speakers are constituted and through which they interact symbolically. As a result, there is a concern about the control of the "schooling" variable in the test, whose evaluation turns out to choose the concept of school as the sole criterion for categorization.

According to Jakobson<sup>30</sup>, the Verbal Fluency Test, as it is currently applied, ultimately mobilizes only one of the linguistic operations, that of selection, wherein the items should maintain a similarity relationship with one another, while the operation of combination is not considered (it is noteworthy that the language in operation involves, for the author, operations, selection and combination). In this sense, fluency in the VFT appears to be limited when compared to the language in actual use, where these two operations of language are present at all linguistic levels.

Moreover, the effort in producing individual words is different from the production effort in conversation shifts; thus, there is a need for case studies, even if merely descriptive, with more comprehensive linguistic observation patterns with spontaneous speech, so that fluency can be evaluated.

## Conclusion

This paper has reported that the first change made by Benton has modified not only the shape and the name of the test, but especially its essence, as it shifts the evaluation from written language to spoken language - two modes that have their own characteristics and require different cognitive and linguistic mechanisms for task execution. Both for the CVFT and later versions of the test, including the current ones, the same evaluation criteria of the original test was maintained, based on productivity. However, it can be seen that there is a change in the notion of language underlying the test, because

originally, it was believed that it could be measured alone, and therefore regardless of other cognitive mechanisms. The current application of the test envisions the deployment of a set of cognitive activities such as memory, attention, vocabulary, which are required for executing the task.

What emerges is that the concept of fluency in the VFT is increasingly different from the concept used in language studies, as it is geared towards the cognitive domain exclusively. In this current perspective, verbal fluency is no longer seen as a language skill assessment and is now interpreted as the verbalization of the flow of cognitive processes; therefore, it is considered as a direct consequence of the cognitive state of the subject. In other words, it shifts from speech fluency assessment to cognitive fluency assessment through verbal production, i.e., it does not evaluate fluency itself, or even the previous original concept of fluency, but it seeks to assess the cognitive integrity of the subjects through the flow of the processes involved in the verbal search and selection.

From the linguistic point of view, the concept of fluency in the VFT differs from the social use of language and takes on a metalinguistic character, where lists of words relevant to a single verb class or a single letter are required. In this scenario, the articulation required for the enunciation context - selection (paradigmatic axis) and combination (syntagmatic axis) - is disregarded; only one of these axes - the paradigmatic one - is evaluated, and language fluency assessment is simplified. This is due to the fact that when the test disregards the combination of words, including the grammatical and discursive complexity that are established between words and through words, it does not address the possible interpretation of traces that indicate the speech production process, such as disfluencies, rhythm, speed, physical behaviors. The combination of linguistic material into syntagmatic traditionally allows the analysis, discussion and judgment of fluency in both typical subjects and subjects with fluency disorders. It is through enunciation complexity that the subject shows fluency disruptions and the materialized traces that act as clues for researching the subject.

Thus, in the TVFT, the concept of fluency still seems to be based more on the perspective of linguistic productivity (quantitative) than on understanding the pathways that enable such productivity (qualitative), disregarding the linguistic

context in which this productivity occurs and the textual organization of language levels for verbal interaction.

Regarding the cognitive functions executed by the VFT, we believe that fluency, under the dialogical perspective, requires not only the integrity of semantic memory (word inventory) and executive function (selection and search), but also the articulation of linguistic levels within a legitimate social event. In addition, for the purpose of comparison, it should be noted that, unlike the VFT, the two operations, selection by similarities and combination by contiguity, are recruited: one seeks the selection of words (among many other similar ones) of a particular class that plays the role of theme in the sentence, and seeks the selection of words that are within the contextual relationship with the other elements of the sentence.

Within the general framework, we realize that the subject's preservation of the meaningful use of language is ultimately left out of language assessments. The methodology used in the test points to a static notion of language, which is repeated at all times and is not subject to variations that depend on context, the interlocutor and the dynamic relationship between the subject and his own language.

There is no doubt about the relevance of the test and its effectiveness in the evaluation of patients with suspected neurological damage. This is because the test measures the combination of a number of cognitive skills whose integrity must be maintained for execution of the proposed task. However, it is understood that fluency, while a language skill, must be evaluated based on the elements that constitute it within a perspective of language functioning. In other words, these concepts underlying the assessment have to be clearly understood, so that one can discuss not only what the test actually measures, but mostly its role and relevance in the assessment of language. Consequently, the proposed discussion is not only relevant to neuropsychology, but especially to speech therapy, in consideration of the test for the interpretation of the subject's language loss and reminiscences in therapy.

## References

1. STRAUSS E. A compendium of neuropsychological tests: Administration, Norms, and Commentary. Oxford: Oxford University Press, 2006.

2. MCDOWD J et al. Understanding verbal fluency in healthy aging, Alzheimer's disease, and Parkinson's disease. *Neuropsychology*. 2011; 25(2): 210-25.
3. ALLAIN P, ETCHARRY-BOUYX F, VERNY C. Executive functions in clinical and preclinical Alzheimer's disease. *Rev Neurol*. 2013; 169(10): 695-708.
4. EASTMAN JA et al. Cortical Thickness and Semantic Fluency in Alzheimer's Disease and Mild Cognitive Impairment. *Am J Alzheimers Dis*. 2013; 1(2): 81-92.
5. MÁRDH S, NÄGGA K, SAMUELSSON S. A longitudinal study of semantic memory impairment in patients with Alzheimer's disease. *Cortex*. 2013; 49(2): 528-33.
6. NITRINIR. et al. Testes neuropsicológicos de aplicação simples para o diagnóstico de demência. *ArqNeuropsiquiatr*, 1994; 52(4): 457-65.
7. MARSHALL JC. The description and interpretation of aphasic language disorder. *Neuropsychologia*, 1986; 24(1): 5-24.
8. BECKER N et al. Estratégias de evocação lexical com critério semântico em adultos após acidente vascular cerebral no hemisfério direito. *Letrônica: Revista Digital do PPGL*, 2014; 7(1): 325-47.
9. CELESTE LC, ALMEIDAA, MARTINS-REIS VO. A auto-avaliação de pessoas com gagueira em relação à expressão de atitudes. *Distúrbios da Comunicação*. 2014; (26)(1): 168-75. Disponível em: <http://revistas.pucsp.br/index.php/dic/article/view/15153/14190>. Acesso em: 28/05/2015.
10. THURSTONE LL. Primary mental abilities. *Science*, 1948; 26: 585.
11. THURSTONE, LL. Psychological implications of factor analysis. *Am Psychol*, 1948; 3(9): 402-8.
12. COHEN MJ, STANCZAK DIE. On the reliability, validity, and cognitive structure of the Thurstone Word Fluency Test. *ArchClinNeuropsychol*. 2000; 15(3): 267-279.
13. BAKHTIN M. *Marxismo e filosofia da linguagem*. São Paulo: Hucitec, 1929/1981.
14. RUFF RM et al. The psychological construct of word fluency. *Brain Lang*. 1997; 57(3): 394-405.
15. BORKOWSKI JG, BENTON A L, SPREEN, O. Word fluency and brain damage. *Neuropsychologia*. 1967; 5(2): 135-40.
16. MITRUSHINA MM, BOONE KL, D'ELIAL. *Handbook of normative data for neuropsychological assessment*. New York: Oxford University Press, 2005.
17. RUFF RM, LIGHT RH, PARKER SB, LEVIN, HS. Benton Controlled Oral Word Association Test: Reliability and updated norms. *Arch ClinNeuropsychol*, 1996; 11(4): 329-38.
18. MACHADO TH et al. Normative data for healthy elderly on the phonemic verbal fluency task-FAS. *DementNeuropsychol*, 2009; 3(1): 55-60.
19. BRUCKI SMD. et al. Dados normativos para o teste de fluência verbal categoria animais em nosso meio; Normative data for the animals category verbal fluency test in our environment. *ArquivosNeuropsiquiatr*, 1997; 55(1): 56-61.
20. GRUENEWALD PJ, LOCKHEAD GR. The free recall of category examples. *J ExpPsychol Learn Mem Cogn*, 1980; 6(3): 225-40.
21. ABWENDER DA. et al. Qualitative analysis of verbal fluency output: Review and comparison of several scoring methods. *Assessment*, 2001; 8(3): 323-38.
22. TROYER, A. Normative data for clustering and switching on verbal fluency tasks. *J ClinExpNeuropsychol*. 2009; 22(3): 370-8.
23. TROYER AK, MOSCOVITCH M, WINOCUR G, LEACH L, FREEDMAN, M. Clustering and switching on verbal fluency tests in Alzheimer's and Parkinson's disease. *J IntNeuropsychol Soc*. 2001; 4(2): 137-43.
24. BRUCKI SMD, ROCHA MSG. Category fluency test: effects of age, gender and education on total scores, clustering and switching in Brazilian Portuguese-speaking subjects. *Braz J Med Biol Res*, 2004; 37(12): 1771-7.
25. ROSCH E. et al. Basic objects in natural categories. *Cogn Psychol*. 1976; 8(3) 382-439.
26. VANNORSDALL TD. et al. Altering automatic verbal processes with transcranial direct current stimulation. *Front Psychiatry*, 2012; 3:1-6.
27. CAVACO S et al. Semantic fluency and phonemic fluency: Regression-based norms for the Portuguese population. *Arch Clin Neuropsychol*. 2013; 28(3): 262-71.
28. VIGOTSKY, L. S. *Pensamento e linguagem*. 2. ed. São Paulo: Martins Fontes, 1989.
29. COUDRY MI. *Diário de Narciso*. São Paulo: Martins Fontes, 2001/1986.
30. JAKOBSON R. *Linguística e Comunicação*. 22a edição. São Paulo: Cultrix, 1969/2010.

## Introdução

Na Neuropsicologia, o teste de Fluência Verbal tem sido largamente utilizado na avaliação de sujeitos com suspeita de danos neurológicos, especialmente de pacientes com possível diagnóstico de Demência de Alzheimer (DA), já que a execução

da tarefa proposta no teste exige a integridade de habilidades cognitivas como memória e funções executivas<sup>1,2,3,4,5,6</sup>. O conceito de fluência, contudo, apresenta-se pouco discutido e mais inclinado à idealização do Teste de Fluência Verbal (TFV) do que ao evento de fluência em si.

Em 1986, já havia autores que chamavam a atenção para a importância dos testes de fluência nas baterias de avaliação<sup>7</sup>. O peso deste teste em relação aos outros testes de linguagem é bastante relevante, podendo mesmo definir, em casos extremos, o quadro patológico do paciente, muito embora as propriedades psicométricas do teste não sejam claras. Mas a questão da assimetria entre os pesos dos testes se torna mais preocupante quando a habilidade de fluência não é bem delimitada, o que leva ao questionamento: “mas o que é a fluência?”.

A fluência de que trata o TFV é uma fluência na fala que remete a aspectos neurológicos. Nesse caso, “a fluência é um processo executivo caracterizado pela capacidade de um indivíduo realizar uma série de comportamentos dentro de uma estrutura de regras estabelecidas, podendo estes ser verbais ou não verbais”<sup>8</sup>.

Assim, na Neuropsicologia, a fluência não se restringe aos aspectos linguísticos, mas expande-se para outras modalidades como pintura, lógica, música, etc. Todas estas modalidades, incluindo a linguística, dizem respeito ao funcionamento cognitivo; mais especificamente, à função executiva, sua integridade e/ou capacidade. Desta forma, a fluência é compreendida tanto como um fator de inteligência quanto como um fator da criatividade.

Logo, as características que envolvem a mensuração da fluência no TFV são bastante diferentes dos estudos de mensuração da fluência na gagueira, por exemplo, em que se busca medir sua severidade e mensurar as rupturas de fala, que é a principal queixa do sujeito<sup>9</sup>. Por outro lado, a fluência no TFV parece ser considerada como uma tarefa pela qual se mobilizam outras atividades cognitivas como memória, atenção, vocabulário, dentre outras. Diante disso, levantam-se algumas questões: Por que o teste recebe este nome e como ele foi criado? Qual a relação estabelecida aqui entre linguagem e cognição para que um escore baixo no teste de fluência indique um déficit cognitivo? Diante desses questionamentos, o objetivo deste artigo é analisar histórica e criticamente o TFV, sua criação e modificações até os dias atuais, bem como suas implicações na avaliação da linguagem.