Available online at www.scholarsresearchlibrary.com

Scholars Research Library

Archives of Applied Science Research, 2017, 9 (2):16-29
(http://scholarsresearchlibrary.com/archive.html)

ISSN 0975-508X
CODEN (USA) AASRC9

An Algorithm for Noun and Verb Ranking in Linguistic Data (ALNOVE)

Ioannis Phinikettos* and Maria Kambanaros

Department of Rehabilitation Sciences, School of Health Sciences, Cyprus University of Technology, Limassol, Cyprus

ABSTRACT

The literature of the two similar methods, statistical classification and clustering analysis, is very broad. The classification procedure builds a model to separate and classify new data points. On the other hand, using clustering analysis, subgroups are created from a set of objects. In this research, we propose a new clustering method for classifying verbs and nouns as Easy, Medium or Hard for linguistic data using normal and language-impaired (LI) population responses on a verb/noun picture-naming test. One scope of the classification is to exclude the 'easy' and 'hard' items from the analysis as 'easy' items are just easy to answer and 'hard' items may be affected by other exogenous variables. The proposed algorithm first classifies the items for each LI group by applying the McNemar test using as reference the easiest and hardest items and then performs an overall ranking using the binomial test. An implementation showed that the difference in medium responses between the normal and LI populations is greater than the corresponding difference in easy responses. This implies that when trying to distinguish between normal and impaired individuals, it is more efficient to proceed with the medium responses. Finally, a classification model is built by creating cut-off points for the two word classes (verbs and nouns) to distinguish between the typical and atypical populations using the medium responses. The clinical outcome is a shortened version of the verb and noun picture-naming test for assessment and research purposes that is valid and reliable.

Keywords: Statistical classification, Cluster analysis, Binomial distribution

INTRODUCTION

There is robust cross-linguistic evidence that the ability to name verbs and nouns can be differentially affected across a wide range of cognitive-linguistic pathologies using picture-naming [1], although no reliable patterns link lesion site and/or brain disease with verb/noun processing differences [2]. Even so, disproportionate verb/noun naming deficits have guided hypotheses about the organization of grammatical word class in the brain [3].

The dissociation between verbs and nouns is claimed to be of fundamental importance, given that both word types are essentially universally available categories across all languages. In simple terms, verbs express states and events, that is, what happens to things, including actions, whereas nouns refer to entities such as people, animals, objects, and concepts. In Indo-European languages, verbs are usually marked overtly for tense, aspect, mood, and number, while nouns are marked for case, gender, and number, although languages vary in which features are marked overtly. There is still considerable debate in the literature-drawing, among others, from the fields of neuroscience, pathology, and linguistics—as to whether verb/noun naming deficits are a true breakdown of a specific grammatical category or whether dissociations can be attributed to lexical-semantic, or syntactic differences between the two word classes [4].

Verbs and nouns differ also on dimensions such as word frequency (i.e., nouns are usually more frequent) and image ability (verbs are considered less imagable), lexical-semantic variables that may affect the naming process for each word type [5]. The general lower image ability and frequency of verbs renders them more vulnerable than nouns to being impaired following brain damage or disease [6]. Moreover, age of acquisition (with nouns usually acquired earlier than verbs across languages) is known to be a significant predictor of naming performance [7].

A number of languages have produced verb and noun picture-naming tests as useful tools for the examination of naming deficits in acquired and developmental language impairments. In fact, picture-naming is most commonly used because it involves all stages of word production (i.e., from picture/concept to word form) and taps into the individual's knowledge about the target word, both quantitatively and qualitatively [8, 9].

The aim of the present study is to propose a new clustering method for classifying verbs and nouns as Easy, Medium or Hard for linguistic data by using the naming responses provided by different language-impaired (LI) populations, as compared to non-impaired or typical control groups, on a verb/noun picture-naming test used for the assessment of naming deficits in Greece and Cyprus [1]. This novel approach allows the exclusion of the 'easy' and 'hard' items from the analysis. The idea is that 'easy' simply implies that both typical and atypical populations having no difficulty naming the item, thus allowing no distinction between the groups. On the other hand, the low performance for both groups (language impaired *vs.* non-language impaired) on the 'hard' items could be a consequence of the extreme values of the corresponding psycholinguistic variables related to the target word.

The proposed algorithm starts by classifying the responses separately for each LI group using as references the responses with the lowest (hardest response) and highest (easiest response) scores. Then the score of all responses is compared with the scores of the easiest and hardest responses by applying several McNemar tests [10]. The responses that showed no significant difference with one of the reference responses are classified as 'easy' or 'hard' accordingly. Then, an overall ranking of the responses is performed by applying the Binomial test. An implementation using the statistical packages R [11] and SPSS [12] revealed it is more efficient to proceed with the medium responses when trying to distinguish between normal and language impaired individuals.

MATERIALS AND METHODS

Participants

A total of 87 language impaired (LI) individuals participated in this research: 73 adults and 14 children. For each LI participant per group, there was a chronologically age-matched non-impaired individual of the same gender, educational level and socioeconomic status. The demographic information of each LI group is presented in Table 1.

Table 1: Demographic characteristics of reported LI groups.

LI Groups	Age (mean) (range between brackets)	Gender	Education (years) (range between brackets)	Intelligence Level (range between brackets)
Broca aphasics	62.4 (30-81)	2 females 5 males	6.0 (4-8)	n/a
Alzheimer's Dementia	82.7 (74-92)	10 females 5 males	7.3 (0-14)	n/a
Schizophrenia	39 (25-62)	14 males 6 females	10.80 (6-16)	98.25 (88-105)
Relapsing Remitting Multiple Sclerosis	40.8 (17-56)	24 females 7 males	12.25 (9-18)	101.50 (80-110)
Children with SLI	6.9 (5.5-9.9)	4 females 10 males	primary school (grades1-4)	Non-verbal IQ>80

Broca's aphasia (BA) (N=7)

The participating individuals with aphasia had suffered a single, relatively localized lesion in the left hemisphere with no other neurological involvement. All were chronic aphasics and met the following criteria: no previous history of infarct, neurologically and physically stable (over 6 months post onset), no history of active or significant alcohol and/or drug abuse, no history of active psychiatric illness or other brain disorder (e.g., Parkinson's disease, Huntington's disease, Korsakoff's syndrome, Alzheimer's disease and other presentations of dementia, senility, and mental retardation), corrected-to-normal auditory and visual acuity for age. All participants were right-handed by self-report, showed a right hemiplegia and were native speakers of Greek. The diagnosis of Broca's aphasia was based on the results of the Greek version of the Boston Diagnostic Aphasia Examination [13], where the severity rating scale ranges from 0 to 5, with 0=no usable speech, 3=mild aphasia, and 5=minimal speech handicaps. The participants with Broca's aphasia presented with non-fluent speech and a mild–moderate to severe aphasia (BDAE severity rating 3-5). The reader is referred to Kambanaros & Grohmann [1] for a detailed description of all background assessment results.

Individuals with alzheimer's dementia (AD) (N=15)

All individuals were clinically diagnosed with first stage AD (1-5 years after diagnosis) by a specialist neurologist and a specialist psychiatrist based on case history information (e.g., deterioration of memory, reduced function in activities of daily living) and cognitive assessment (at least 2 deficits in cognitive functions/domains). Exclusion criteria from the naming study included: organic CNS pathology-neurological disorders, major psychopathology spectrum disorders, head trauma resulting in loss of consciousness for longer than five minutes, mental retardation, illicit substance dependencies including alcohol for the past six months. All individuals with AD lived in the community with a family member.

Individuals with schizophrenia (SCZ) (N=20)

All patients were clinically evaluated for psychiatric status according to DSM-IV-TR criteria (using SCID Axis I and SCID Axis II) and by a specialist neurologist to exclude neurological disorders. Schizophrenia type varied within the group: 10 individuals were reported as suffering from paranoid schizophrenia, six individuals with undifferentiated schizophrenia, and two individuals with catatonic and residual schizophrenia, respectively. In addition, 17 of the individuals with schizophrenia were on atypical antipsychotic medication, while the remaining three were on typical antipsychotic and mood stabilizer medication. Exclusion criteria from the naming study included: organic CNS pathology-neurological disorders, HIV/HCV infection, major psychopathology spectrum disorders (excluding schizophrenia), head trauma resulting in loss of consciousness for longer than five minutes, dementia, mental retardation, and current therapy with medications or medical conditions known to affect cognition, illicit substance dependencies including alcohol for the past six months prior to inclusion in the maintenance therapy, and non-native speakers of the Greek language. All had adequate hearing and vision for test purposes. The reader is referred to Kambanaros et al [14] for detailed demographic and clinical characteristics of the participants.

Individuals with relapsing remitting multiple sclerosis (RRMS) (N=31)

This group consisted of patients with RRMS, diagnosed according to the McDonald criteria [15]. The reader is referred to [16] for detailed demographic (age, education, gender distribution, intelligence level) and clinical characteristics of the RRMS patients (Expanded Disability Status Scale, disease duration, Beck Depression Inventory–Fast Screen).

All adult participants in the above-mentioned groups provided informed consent to participate in the study, and permission to conduct the research was obtained by the local ethics committee.

Children with specific language impairment (SLI) (N=14)

Children were diagnosed with SLI prior to the noun/verb naming study using a language battery of norm-referenced tests [17]. The language difficulties encountered by the children were predominantly in expressive language in the domains of (morpho) syntax and the lexicon. Hearing and vision were adequate for test purposes and the children with SLI exhibited normal performance on a screening measure of non-verbal intelligence [18]. Children also showed normal articulation, had no gross motor difficulties, and came from medium to high socio-economic status families. Participant selection criteria included a (Cypriot) Greek-speaking family background and no history of neurological, emotional, or behavioural problems. The parents of each participating child provided written consent.

Stimuli

The Greek Object and Action Test [19] is the tool administered to assess naming of nouns and verbs for assessment and/or research purposes for Greek-speakers. It contains 84 coloured photographs, 10-14 cm in size representing 42 actions (verbs) and 42 objects (nouns). The test in total (production and comprehension subtests) takes under an hour to administer. The GOAT was piloted on a group of twenty non-brain-injured, monolingual Greek speakers aged between 55 and 75 years (Table 2) [18]. Only items named with 80% accuracy or more are included in the test. Objects are concrete inanimate nouns and include manipulated instruments used for activities of daily living such as garage tools (e.g., hammer), garden equipment (e.g., rake), kitchen utensils (e.g., grater), and items from household (e.g., broom), office (e.g., pen), or personal use (e.g., comb). Verbs are mono transitive, though frequently allow their object dropped, and actions are restricted to past stereotypical roles, that is, a woman is shown performing household activities (e.g., mopping) and a man is performing more manly duties (e.g., hammering). These stereotypical roles depicted in the photographs are deemed to be appropriate for the ages and cultural groups tested. All target nouns in object naming are also items in the noun comprehension task. All target verbs in action naming are also targets in the verb comprehension task .

Table 2: Mean values (SD) of characteristics for noun and verb pictures on the GOAT.

Picture type	Word Frequency	Syllable length	AOA	Word image ability	Picture Complexity
Nouns	40.91	2.88 (0.803)	2.98 (0.76)	6.49 (0.49)	6.49 (0.28)
Verbs	40.11	2.95 (0.731)	2.82 (.58)	6.42 (0.16)	6.42 (0.67)
Mann-Whitney test					
z-values	-1.264	-1.296	-1.168	- 2.978	- 2.331
p-values	0.443	0.264	0.243	0.003*	0.020*

Note: GOAT: Greek Object and Action Test; AoA: Age of Acquisition; SD: Standard Deviation in Parentheses; *Significantly Different

Verb and noun word frequencies were calculated based on the printed word frequency count for Standard Greek [20]. A Mann-Whitney test revealed no significant difference between nouns and verbs (z=– 0.154, p=0.878). In addition, there was no significant difference in syllable length between nouns and verbs (z=– 0.610, p=0.542). Nouns and verbs were measured also for age of acquisition (AoA, estimated age ratings were based on first contact with the given noun/verb in either verbal or written form using a seven-point scale, with 1 representing 0-2 years of age, 2 being 3-4 years of age, up to 7 for 13 years of age and older), image ability (ratings were performed on an eight-point scale, with 0=impossible, 1=least image able, up to 7=most image able), and picture complexity (ratings were performed on a seven-point scale related to the ease with which the noun/verb picture was recognized, from 1=least ease to 7=most ease). A Mann-Whitney test revealed that the nouns and verbs were not significantly different on AoA (z=–1.168, p=0.243), but there was a significant difference for word image ability (z=–2.978, p=0.003) and picture complexity (z=– 2.331, p=0.20), with higher ratings for nouns compared to verbs, revealing nouns as more imagable and visually less complex than verbs upon picture identification. The reader is referred to Kambanaros & Grohmann, [1] for further details. All the same, such differences between verbs and nouns for image ability and picture complexity by virtue of depicting actions in a static fashion-are common phenomena reported in the literature across languages which investigate verbs and nouns with pictured stimuli [3] The adapted version of the GOAT was used to test the children with SLI and this had 35 nouns (instead of 42) as nouns with a mean age of acquisition greater than 6 years were removed [20].

PROCEDURE

Each participant was tested individually by a certified speech and language therapist (the first author, a certified bilingual speech pathologist, administered the GOAT on all LI populations). The same set of 84 items of the GOAT (42 verbs and 42 nouns, respectively 35 nouns and 39 verbs for the children with SLI) was used for the word comprehension and production tasks. The noun and verb tasks were counterbalanced across all participants.

For word production, participants were asked to name the noun or verb represented by the object or action depicted in the photograph, respectively, in a single word. The stimulus question was short and of equal length for both the noun and verb naming subtests (5 syllables): Ti ine afto? 'What's this?' for objects and Ti kani aftos/afti? 'What's he/she doing?' for actions. In the responses, nouns were supposed to be provided marked for nominative (which all participants did, though other case-markings would have been accepted too). Since Greek lacks a non-finite citation form (infinitive or gerund), verbs were required in the third person singular present tense (which all participants did, though other inflections would have been accepted too). Two examples were provided before testing. The stimulus was repeated once for participants who did not respond. If no response was given, the item was scored as incorrect. Again, no time limits were placed, hence no reaction times measured, and self-correction was allowed.

RESULTS AND DISCUSSION

A total of 7,168 spoken responses were produced by the LI groups: 3,556 responses for nouns and 3,612 responses for verbs. Correct responses scored 1 point, and were single words that named the object/noun and action/verb portrayed in the photograph.

Accuracy

The percentage of correct responses for nouns and verbs retrieved by each LI group on the GOAT is reported in Table 3. A Wilcoxon Signed Rank Test [23] which reports significance at p=0.05 for paired samples was performed to see whether there was a difference between noun and verb naming accuracies within each LI group. Overall, verbs were significantly more difficult to retrieve compared to nouns for all LI groups (χ^2=1.0819, df=5, p=0.956).

Table 3: Percentage (SD) of correct responses for noun and verbs by each LI group.

LI Group	Nouns	Verbs
AD	55.40 (20.68)	49.69 (25.62)
BA	52.04 (21.93)	43.20 (21.63)
MS	88.78 (8.34)	84.41 (10.56)
SCZ	85.24 (6.16)	75.83 (19.27)
SLI	68.71 (14.18)	62.93 (8.94)

Note: AD: Alzheimer's Dementia; BA: Broca's Aphasia; MS: Multiple Sclerosis; SCZ: Schizophrenia; SLI: Specific Language Impairment

Development of the algorithm

This section describes the newly proposed classification method. Let us denote by $P_1,\ldots,P_n$ the n individuals from a given LI group that were involved in the picture-naming study to answer k responses $R_1,\ldots,R_k$. Each response is marked as 0 for a wrong answer and as 1 for a correct one. The average score of a specific individual P_i is denoted as AP_i and the average of a response R_j is denoted as AR_j. This data is represented in Table 4.

Table 4: Score sheet of the k responses answered by the n individuals (LI groups).

Patient	R_1	R_2	R_k	-
P_1	1	0	1	AP_1
P_2	1	1	1	AP_2
P_n	0	1	0	AP_n
	AR_1	AR_2	AR_k	-

Since all LI individuals answered the same question ('What is this?' for the noun stimuli, 'What is he/she doing?' for the verb stimuli), the R_i are dependent random vectors. Therefore, to test if there exists a significance difference among the mean score of all responses (null hypothesis $AR_1=AR_2=\ldots=AR_k$), a Cochran's Q test was performed. If the Cochran's Q test showed a significant difference, we proceeded to classify the responses as Easy, Medium or Hard.

The responses from the LI group were first sorted in ascending order based on their average score AR_i. The responses with the highest average score (easiest responses) and the lowest average score (hardest responses) were chosen and denoted by R_{max} and R_{min} respectively. If there were two candidates for R_{max}, one was chosen at random (e.g. if R_1 and R_2 have the same highest average score, define $R_{max}=R_1$). The same applied for the hardest response R_{min}. Next, R_{max} was fixed as a reference and its average score was compared to the average score of every other response by applying several Mcnemar tests. The responses that showed no significant difference with R_{max} were classified as Easy. The procedure was repeated using as reference R_{min}. The responses that had no significant difference with R_{min} were classified as Hard. The remaining responses were classified as Medium.

Let us denote the sets of Easy, Medium and Hard responses as B_E, B_M and B_H respectively. If the intersection of B_E and B_H is not the empty set, we classify the responses that lie in the intersection as Medium and exclude them from both B_E and B_H.

The method described so far represents a classification for an impaired group. We repeat the procedure for all m impaired groups and use the classified data to create an overall classification for the responses. For the overall classification step, we used the Binomial test.

Let us define the functions N_i: Responses$\rightarrow\{0,1,2,\ldots,m\}$, where i denotes either Easy (E), Medium (M) or Hard (H), to denote the number of times a given response was classified as type i. Suppose we want to make inference about the probability of the number of times a response R_j was classified as Easy. The events of success are the Easy counts ($N_E(R_j)$) and the events of failure are the medium and hard counts ($N_M(R_j) + N_H(R_j)=m - N_E(R_j)$). This is like repeating an experiment m times and counting the number of successes. This is why we deal with a binomial distribution with m trials and some success probability p. Using the binomial test, we checked if the probability p is significantly greater than 0.5. If there was significant evidence, the response R_j was classified as Easy. Otherwise, the procedure was repeated for Medium (events of success $N_M(R_j)$) and then for Hard counts (events of success $N_H(R_j)$). If none of these counts showed significant evidence, the binomial test was applied twice by combining the Easy and Medium counts (events of success $N_E(R_j)+N_M(R_j)$) in the first case and the Medium and Hard counts (events of success $N_M(R_j)+N_H(R_j)$) in the second case. The response R_j was classified as Easy to Medium and Medium to Hard accordingly. If none of the binomial tests showed significant difference, the response R_j was classified as Medium.

Let $l \leq m$ denote the minimum number of successes that when applying the binomial test, the probability will be significantly greater than 0.5. For a given response R_j $N_E(R_j)$, $N_M(R_j)$ and $N_H(R_j)$ was computed. If any of these values was greater or equal to the minimum number of successes l, the response R_j was classified accordingly as Easy, Medium or Hard. Otherwise, the sums were tested $N_E(R_j)+N_M(R_j))$ and $N_M(R_j)+N_H(R_j)$ and if any of these values was greater or equal to l, then R_j was classified as Easy to Medium or Medium to Hard respectively. If none of the aforementioned restrictions was satisfied, then the response R_j was classified as Medium.

In Algorithm 1 we present the classification algorithm (R_{ij} denotes the vector of scores for the jth response for group i and the significance level of all tests is arbitrary):

Algorithm 1 (ALNOVE). Classification algorithm.
1. **For** *i* **in 1 up to**
2. **Response overall difference:** Perform a Cochran's Q test to investigate if there are significant differences among the average score of all responses in group *i*.
3. **If** significant differences exist, then
4. **Group classification**: Compute R^i_{min} and R^i_{max}
5. **For** *J* **in 1 up to** *K*
6. **If** McNemar(R^i_{min}, R^i_j)=non-significant then $R_J \in B^i_H$
7. **Else if** McNemar(R^i_{max}, R^i_j)=non-significant then $R_j \in B^i_E$
8. **Else** $R_J \in B^i_M$
9. **If** $B^i_E \cap B^i_H \neq \varnothing$ then let $B^i_M := B^i_E \cap B^i_H$, $B^i_E := B^i_E / B^i_M$ and $B^i_H := B^i_H / B^i_M$.
10. **End for**
11. **End if**
12. **End for**
13. **Overall classification:** Compute the minimum number of successes *l*.
14. **If** $N_E(R_j) \geq l$ classify R_j as Easy.
15. **Else If** $N_M(R_j) \geq l$ classify R_j as Medium.
16. **Else If** $N_H(R_j) \geq l$ classify as Rj Hard.
17. **Else If** $N_E(R_j) + N_M(R_j) \geq l$ classify as Rj Easy to Medium.
18. **Else If** $N_M(R_j) + N_H(R_j) \geq l$ classify as Rj Medium to Hard.
19. **Else** classify Rj as Medium.
20. **End if**

Implementation of the classification method

In this section we apply the proposed classification method to the five impaired groups AD, BA, MS, SCZ and SLI (m=5) to classify the 42 verbs and 42 nouns that have been used in the GOAT study [1] as Easy, Medium or Hard to name. The classification procedure was applied separately for each grammatical category (nouns *vs.* verbs). The significance level of all tests was 0.05. Tables 4 and 5 show the group and over all naming classification for nouns and verbs respectively.

Table 5: Noun classification based on graded naming difficulty using individual LI groups correct/incorrect naming responses.

	LI Group	**AD**	**BA**	**MS**	**SCZ**	**SLI**	**Overall naming difficulty**
	Nouns						
1	ποτιστήρι (watering can)	E	M	E	E	E	Easy to medium
2	αναπτήρας (lighter)	M	M	E	E	H	Medium
3	τρίφτης (grater)	E	M	E	E	H	Medium
4	λίμα (file)	H	H	M	H	M	Medium to hard
5	ξυράφι (razor)	E	M	E	E	M	Easy to medium
6	ζυγαριά (scales)	H	H	E	E	H	Medium
7	χτένα (comb)	E	M	E	E	E	Easy to medium
8	σκούπα (broom)	M	M	E	E	E	Easy to medium
9	κόλλα (glue)	H	M	H	H	E	Medium
10	κλειδί (key)	E	M	E	E	E	Easy to medium
11	σίδερο (iron)	E	M	E	E	E	Easy to medium
12	τρυπάνι (drill)	H	M	H	H	M	Medium to hard
13	σφυρίχτρα (whistle)	H	M	E	E	E	Medium

14	σφουγγάρι (sponge)	H	H	E	E	M	Medium
15	ξύστρα (sharpener)	H	M	E	E	E	Medium
16	κόσκινο (sieve)	E	H	H	H	M	Medium
17	μολύβι (pencil)	E	M	E	E	E	Easy to medium
18	σφυρί (hammer)	E	M	E	E	M	Easy to medium
19	μικρόφωνο (microphone)	H	M	E	E	M	Medium
20	στυλό (pen)	H	M	E	E	E	Medium
21	κουτάλι (spoon)	E	M	E	E	E	Easy to medium
22	ψαλίδι (scissors)	E	M	E	E	E	Easy to medium
23	σφουγγάρι (sponge)	M	M	E	E	E	Easy to medium
24	τσουγκράνα (rake)	H	M	H	H	H	Medium to hard
25	καλάμι (fishing rod)	H	H	H	H	M	Medium to hard
26	κατσαρόλα (saucepan)	E	M	E	E	M	Easy to medium
27	μυστρί (trowel)	E	M	E	E	M	Easy to medium
28	πινέλο (paint brush)	M	M	E	E	M	Easy to medium
29	δίσκος (tray)	M	M	M	M	H	Medium to hard
30	βελόνα (needle)	M	M	E	M	H	Medium
31	σχοινί (rope)	M	M	M	E	E	Easy to medium
32	σκάλα (ladder)	E	M	E	E	E	Easy to medium
33	τηλεόραση (television)	E	M	E	E	E	Easy to medium
34	φάκελος (envelope)	E	M	E	E	E	Easy to medium
35	γραβάτα (tie)	E	M	E	E	E	Easy to medium
36	κουδούνι (bell)	E	M	E	M	E	Easy to medium
37	γάντι (glove)	E	M	E	E	E	Easy to medium
38	μπαλόνι (balloon)	M	M	E	E	E	Easy to medium
39	ρολόι (watch)	E	E	E	E	E	Easy
40	κρεβάτι (bed)	E	M	E	E	E	Easy to medium
41	καναπές (couch)	M	M	E	E	E	Easy to medium
42	εφημερίδα (newspaper)	E	M	E	E	E	Easy to medium

Note: AD: Alzheimer's Dementia; BA: Broca's Aphasia; MS: Multiple Sclerosis SCZ: Schizophrenia; SLI: Specific Language Impairment; E: Easy; M: Medium; H: Hard

Considering Tables 5 and 6, we notice that the noun 'ρολόι' (watch) was the easiest given that it was the only noun that has an overall classification as Easy. For the verbs, 'γράφει' (writing) and 'κάθεται' (sitting) were the easiest and 'μαζεύει' (ranking) was the hardest.

Table 6: Verb classification based on graded naming difficulty using individual LI groups correct/incorrect naming responses.

	LI group	AD	BA	MS	SCZ	SLI	Overall naming difficulty
	Verbs						
1	ζωγραφίζει (drawing)	E	M	H	H	M	Medium
2	καρφώνει (hammering)	E	M	E	M	H	Medium
3	τραγουδάει (singing)	H	M	E	E	E	Medium
4	γράφει (writing)	E	E	E	E	E	Easy
5	ανακατεύει (stirring)	M	H	M	M	M	Medium to hard
6	κόβει (cutting)	E	M	E	E	E	Easy to medium
7	πλένει (washing)	M	M	E	E	E	Easy to medium
8	μαζεύει (raking)	H	H	H	H	H	Hard
9	ψαρεύει (fishing)	H	M	E	E	E	Medium
10	μαγειρεύει (cooking)	E	M	M	E	E	Easy to medium
11	χτίζει (building)	M	H	E	E	H	Medium
12	βάφει (painting)	M	M	E	E	E	Easy to medium
13	ράβει (sewing)	E	H	E	E	M	Medium
14	σερβίρει (serving)	H	H	M	M	H	Medium to hard
15	ποτίζει (watering)	E	M	E	E	E	Easy to medium
16	ανάβει (lighting)	M	E	E	E	E	Easy to medium
17	τρίβει (grating)	E	M	M	E	H	Medium
18	λιμάρει (filing)	H	H	M	M	M	Medium to hard

19	ξυρίζει (shaving)	E	M	E	E	M	Easy to medium
20	ζυγίζει (weighing)	E	H	E	E	H	Medium
21	χτενίζει (combing)	E	M	E	E	E	Easy to medium
22	σκουπίζει (sweeping)	E	M	E	E	E	Easy to medium
23	κολλάει (glueing)	H	M	E	M	E	Medium
24	κλειδώνει (locking)	H	M	E	E	E	Medium
25	σιδερώνει (ironing)	E	M	E	E	E	Easy to medium
26	τρυπάει (drilling)	H	M	M	E	H	Medium
27	σφυρίζει (whistling)	H	M	E	M	M	Medium
28	σφουγγαρίζει (mopping)	E	M	E	E	E	Easy to medium
29	ξύνει (sharpening)	H	H	M	E	E	Medium
30	κοσκινίζει (sifting)	E	M	M	H	M	Medium
31	τραβάει (pulling)	H	M	M	E	M	Medium
32	ανεβαίνει (climbing)	E	M	E	E	E	Easy to medium
33	βλέπει (watching)	E	M	E	E	E	Easy to medium
34	στέλνει (sending)	H	H	E	M	H	Medium
35	δένει (tying)	E	M	M	M	H	Medium
36	χτυπάει (ringing)	M	H	E	E	E	Medium
37	φοράει (wearing)	M	M	H	H	E	Medium
38	φουσκώνει (blowing)	E	M	E	E	E	Easy to medium
39	κουρδίζει (winding)	H	M	H	H	M	Medium to hard
40	κοιμάται (sleeping)	E	M	E	E	E	Easy to medium
41	κάθεται (sitting)	E	E	E	E	E	Easy
42	διαβάζει (reading)	E	M	E	E	M	Easy to medium

key: ad=Alzheimer's Dementia; ba=Broca's Aphasia; ms: Multiple Sclerosis; scz: Schizophrenia; sli: Specific Language Impairment; e: Easy; m: Medium; h: Hard

For the next sections, all responses that were classified as Easy to Medium or Medium to Hard were treated as Easy and Hard respectively. First, the hard responses were analysed with the observation that all contained at least one psycholinguistic variable with an extreme value. Further, the difference between the average score of 'easy' and 'medium' responses was larger for the impaired individuals compared to the normal population. These findings suggest that it could be considered reasonable to exclude the 'easy' and 'hard' responses from the analysis.

Hard responses – which psycholinguistic variables might be responsible for the low average score

In this section we analysed the 'hard' verbs and the 'hard' nouns. For this type of items, the LI individuals have scored significantly lower and a possible connection with the corresponding psycholinguistic variables was investigated. To address this issue, the percentiles for all psycholinguistic variables was constructed. A psycholinguistic variable for a specific item was considered a possible factor if it fell in the lower 10% (or very close) of the corresponding distribution. For age-of-acquisition (AoA), since it is proportionally reversed, we considered the corresponding upper tail. Tables 7 and 8 report on the percentiles for 'hard' verbs and 'hard' nouns.

Table 7: Percentiles of 'hard' verbs on all psycholinguistic variables.

Hard verbs	AoA	Picture Complexity	Image ability	Name Agreement	Familiarity	Frequency
μαζεύει (ranking)	.3452	.0476	.3810	.0238	.6190	.6786
κουρδίζει (winding)	.8929	.1786	.2143	.0714	.0476	.1429
σερβίρει (serving)	.8929	.3571	.3810	.1786	.2857	.4405
ανακατεύει (mixing)	.3452	.1190	.3810	.4286	.7262	.2738
λιμάρει (filing)	.9762	.4286	.5595	.4286	.3095	.0714

We observed that for all 'hard' verbs and nouns, there exists at least one psycholinguistic variable with the given property. For the hard verb 'μαζεύει' (ranking) for example, we notice that low picture complexity and low name agreement can be considered as possible factors for the LI groups low naming performance. For the hard noun 'λίμα' (nail file) the possible factors are high AoA and low picture complexity. Also we notice that none of the 'hard' items contained a psycholinguistic variable in the upper (lower for AoA) 25% of the corresponding distribution.

Table 8: Percentiles of hard nouns on all psycholinguistic variables.

Hard nouns	AoA	Picture Complexity	Image Ability	Name Agreement	Familiarity	Frequency
δίσκος (serving tray)	.7143	.0238	.5357	.6786	.3690	.0714
τσουγκράνα (rake)	.3810	.2143	.1071	.0952	.1429	.5119
τρυπάνι (drill)	.9524	.3095	.0714	.1429	.0952	.5952
καλάμι (fishing rod)	.8810	.0714	.1071	.1905	.0714	.5119
λίμα (nail file)	.9286	.0476	.2262	.2857	.5833	.5714

Easy responses

Let us first consider the statistic of the McNemar test and see why it might be wise to skip the easy responses. Suppose we consider an individual (P_1) with lexical impairment versus a normal individual (P_2). We expect that both individuals will perform similarly (answer correctly) when they are given the easy verbs.

Consider Tables 9 and 10. Table 9 contains the score of the two individuals for all classified responses and Table 10 the corresponding contingency table.

Table 9: Score of two individuals (P_1 impaired and P_2 normal).

Response	Class	P_1	P_2
R_1	E	1	1
R_2	E	1	1
R_i	M	0	1
R_{i+1}	M	0	1
R_{k-1}	H	0	0
R_k	H	0	0

Table 10: Contingency table of the score of two individuals.

	P_2 correct	P_2 incorrect
P_1 correct	*a*	*b*
P_1 incorrect	*c*	*d*

The individuals can be compared using the McNemar test with the corresponding statistic χ^2= (b-c)^2 / (b+c). As can be seen, this statistic does not include the value a nor d which correspond to correct and incorrect answers by both individuals respectively. We expect that the value α will be dominated by the easy responses and also d by the hard responses. The numbers b and c are thus dominated by the medium responses.

To investigate this issue, consider Figures 1 and 2 that the following two scatter plots from our data.

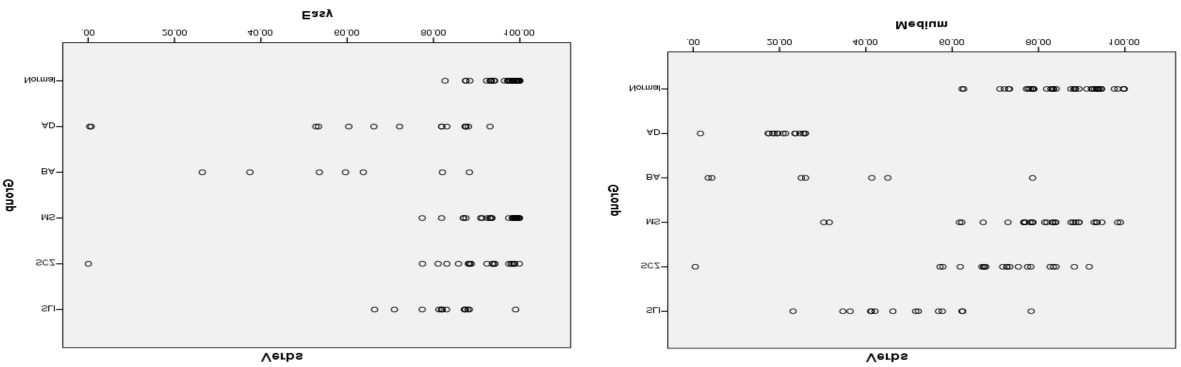

Figure 1: Scatterplots for easy and medium verbs for all groups

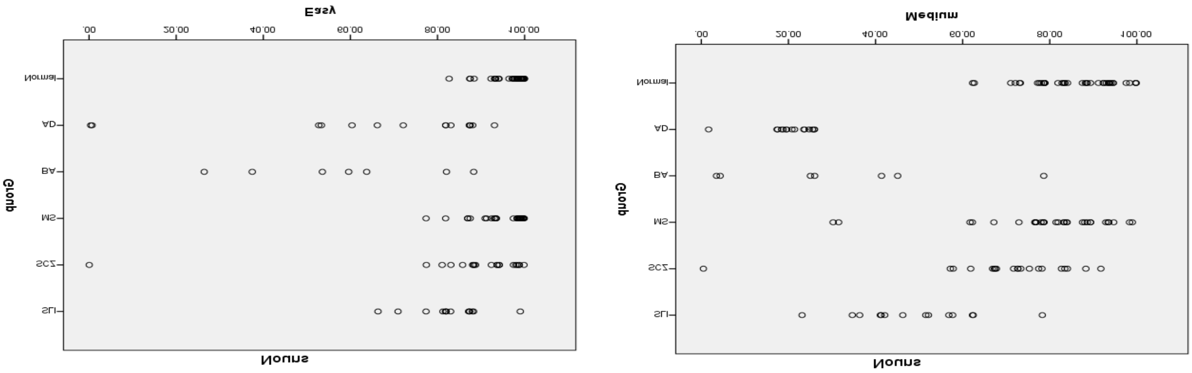

Figure 2: Scatterplots for easy and medium nouns for all groups

These scatter plots contain the performance of the individuals for all impaired groups for the easy and medium responses for both noun and verbs. Let us now try to distinguish the normal from the impaired individuals using the verbs. Considering the first scatterplot of the easy verbs, we observe that the normal individuals perform very similar to SCZ and MS, close to SLI and that there were some individuals in the BA and AD groups that had similar performance to the normal population.

On the other scatterplot of medium verbs, we notice that all clinical groups have spread more from easy verbs than the normal. Specifically, there was no AD patient that overlapped with the normal individuals and there was just one BA that scored in the normal range. The SLI group also seem to perform much worse in the medium verbs but the MS and SCZ scored more similar to the normal group. The same image can be seen when considering the scatterplots of easy and medium nouns.

To provide evidence to our observation, we will run a two-way mixed ANOVA with factors the groups of people and the class of verb (easy and medium verbs). The results of these analyses are given below (consider Figures 3-8)

Ad (Figure 3)

We notice that the decrease from easy to medium responses for both verbs and nouns is steeper for the AD individuals and this can be seen by the significant interaction (verbs: f $(1,53)=41.82$, p-value<.001, $\eta^2=0.441$ and nouns: f$(1,53)=167.31$, p-value<.001, $\eta^{2=}0.759$).

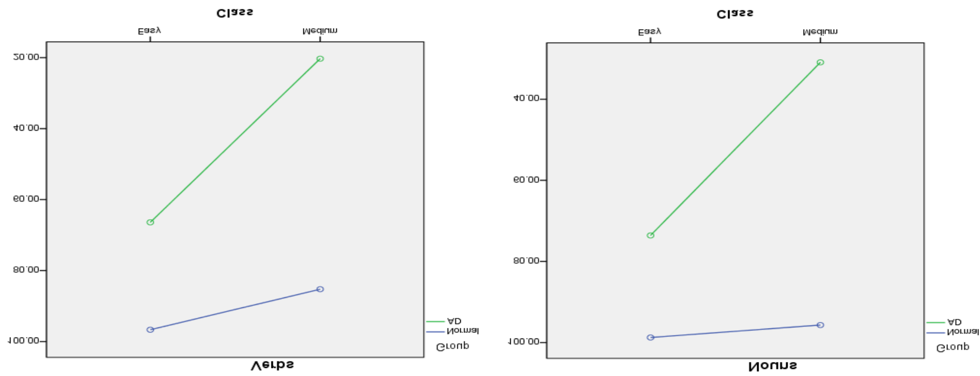

Figure 3: Average score of normal and AD individuals for easy and medium classes for both noun and verbs

Ba (Figure 4)

The decrease from easy to medium responses for both verbs and nouns is steeper for the BA individuals given by the significant interactions (Verbs: F $(1,45)=10.779$, p-value=.002, $\eta^{2=}0.193$ and Nouns: F$(1.45)=73.269$, p-value<0.001, $\eta^{2=}0.619$).

Ms (Figure 5)

We notice that the behaviour of MS individuals is similar to the normal people for verbs. (F (1.69), p-value=0.142, $\eta^2=0.031$) but the decrease is steeper from easy to medium nouns (F $(1,69)=13.477$, p-value<0.001, $\eta^2=0.163$).

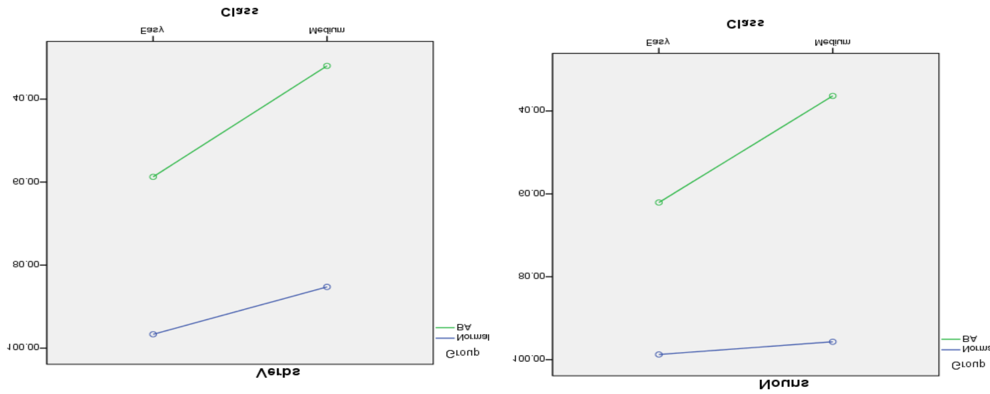

Figure 4: Average score of normal and BA individuals for easy and medium classes for both noun and verbs

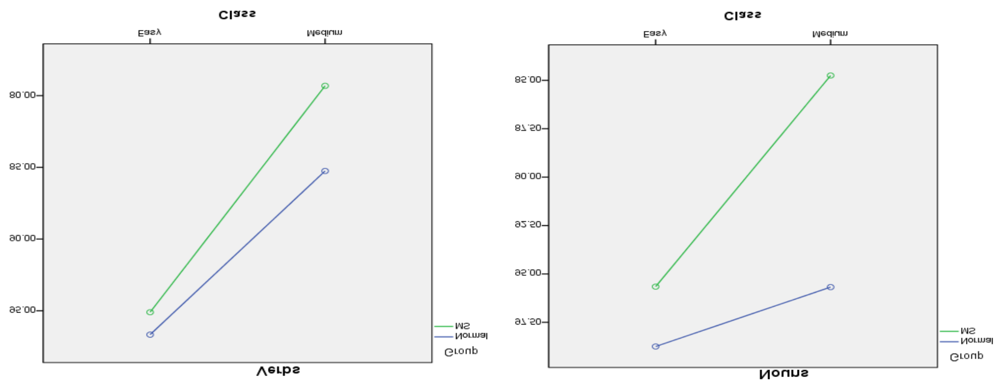

Figure 5: Average score of normal and MS individuals for easy and medium classes for both noun and verbs

Scz (Figure 6)

We notice that the behaviour of SCZ individuals is similar to the normal people for the verbs but the p-value of the interaction is marginally higher than the significance level of 0.05 (F (1.58)=3.48, p-value=.067, η^2=0.057). On the other hand there exists a steeper decrease from easy to medium nouns (F (1.58)=11.846, p-value<0.001, $\eta^{2=}$0.434).

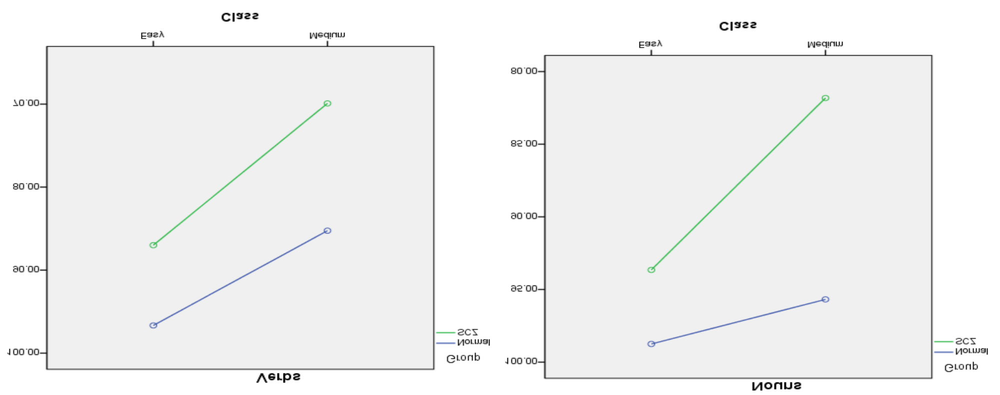

Figure 6: Average score of normal and SCZ individuals for easy and medium classes for both noun and verbs

SLI (Figure 7)

The decrease from easy to medium responses for both verbs and nouns is steeper for the SLI individuals given by the significant interactions (Verbs: F (1.52)=40.072, p – value<0.001, $\eta^{2=}$0.435 and Nouns: F (1,52)=130.105, p – value<0.001, $\eta^{2=}$0.714).

All patients (Figure 8)

Overall, the decrease from easy to medium responses for both verbs and nouns is steeper for the impaired population given by the significant interactions (Verbs: F (1.125)=15.119, p-value<0.001, $\eta^{2=}0.108$ and Nouns: F (1.125)=36.574, p-value<0.001, $\eta^{2=}0.224$).

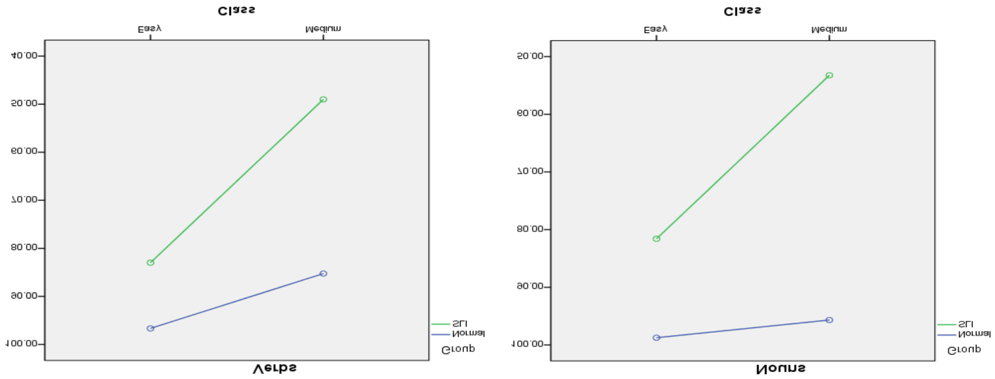

Figure 7: Average score of normal and SLI individuals for easy and medium classes for both noun and verbs

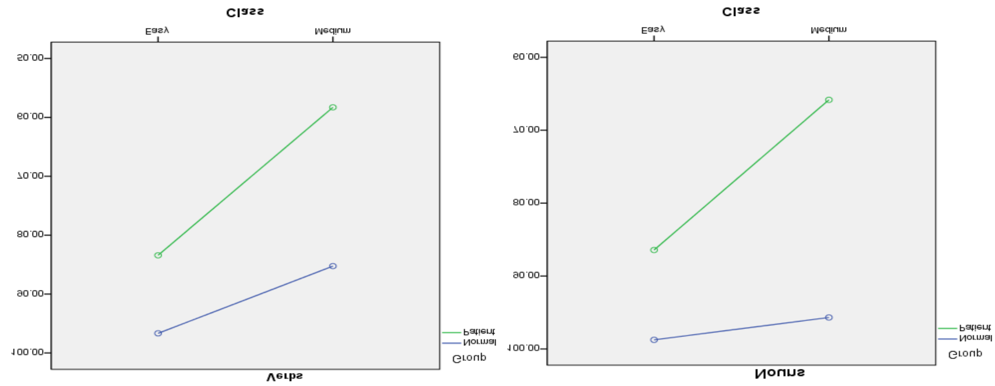

Figure 8: Average score of normal and all LI individuals for easy and medium classes for both noun and verbs

Medium responses – warning and cut-off points

Here we will exclude the easy and hard responses and perform the analysis using the medium ones. Let us assume that all individuals are tested using only on verbs. Since we are dealing only with one class, we will assume that, under the normal population, all the medium verbs have the same success probability p. Given that there are n_m medium verbs, the number of correct answers of a normal individual will follow a binomial distribution with n_m trials and probability of success P i.e Bin(n_m, p). Equivalently, the number of incorrect responses will follow Bin (n_m, q) where q=1- p First we estimate the success probability P by the empirical estimator $\hat{p}$ and the failure probability q by $\hat{q} = 1 - \hat{p}$

Let $X \sim \mathrm{Bin}(n_m, \hat{q})$ The random variable X counts the number of incorrect answers an individual given in n_m trials with probability of success$_{\hat{p}=0.}$. Consider the probabilities P ($X \geq x$). We seek for the minimum positive integers k_w and k_c such that P ($X \geq k_w$) ≤ 0.05 and P($X \geq k_c$) ≤ 0.001. This means that there is less than 0.05 or 0.001 chance that a normal individual will answer incorrectly in more than k_w or k_c verbs respectively. The number k_w will serve as a warning point and k_c as a cut-off point.

In our example we have n_m=19, the success probability $\hat{p} = 0.863$ and $\hat{q} = 1 - 0.863 = 0.137$ which gives X ~ Bin (19, 0.137). The warning and cut-off points are given by k_w=5 and k_c=7.

A similar analysis may be applied using only the 11 medium nouns. For this case we have n_m=11$_{,}$ the success probability $\hat{p} = 0.942$ and $\hat{q} = 1 - 0.942 = 0.058$ which gives X ~ Bin(19, 0.058). The warning and cut-off points are given by k_w=3 and $k_{c=}$5.

Also we can combine both medium verbs and medium nouns to apply the GOAT test using the 30 medium responses. Now since the success probability of noun and verbs differ, we cannot apply the binomial distribution to create the warning and cut-off values. Instead the warning and cut-off points can be constructed using the Poisson binomial distribution [23]. Their corresponding values are given by k_w=7 and $k_{c=}$10.

Shortened version of the GOAT

Clinical studies can be very time consuming especially when language assessment measures have a large number of items to be administered. One scope of the classification method presented in this paper, is to dismiss the 'easy' and 'difficult' items from the GOAT, a verb/noun picture-naming assessment measure, as not being informative. By this we mean, that the 'easy' and 'difficult' verb/noun target items/responses were not able to distinguish language impaired (i.e., those suffering from a naming/lexical impairment) from non-impaired groups. As a consequence, the GOAT tool is significantly shortened from its original 84 items (42 verbs and 42 nouns) down to 30 items (19 verbs and 11 nouns). Moreover, in the clinical setting, for the assessment of naming abilities, when administering the shortened-version of the GOAT, clinicians will be informed that greater than 7 errors will suggest a naming/lexical deficit that requires further investigation across a wider range of tasks.

CONCLUSION

A new algorithm (ALNOVE) was proposed to dismiss redundant or non-informative items from a picture-based naming measure used for the assessment of lexical impairments across language-impaired populations for Greek. This has allowed a shortened version of the tool to be developed that is fast to administer and reliable for clinicians.

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

REFERENCES

[1] Kambanaros, M., and Grohmann, K. K., Grammatical class effects across impaired child and adult populations. *Frontiers in Psychology*, **2015.** p. 6.

[2] Arévalo, A.L., et al., Action and object processing in brain-injured speakers of Chinese. *Neuropsychology,* **2011.** 25(6): p. 792.

[3] Crepaldi, D., et al., A place for nouns and a place for verbs? A critical review of neurocognitive data on grammatical-class effects. *Brain and Language,* **2011.** 116(1): p. 33–49.

[4] Kemmerer, D. Word classes in the brain: implications of linguistic typology for cognitive neuroscience. Cortex, **2014.** 58: p. 27–51.

[5] Bastiaanse, R., Wieling, M., and Wolthuis, N., The role of frequency in the retrieval of nouns and verbs in aphasia. *Aphasiology,* **2016.** 30(11): p. 1221–1239.

[6] Crepaldi, D., et al., Noun--verb dissociation in aphasia: The role of imageability and functional locus of the lesion. *Neuropsychologia,* **2006.** 44(1): p. 73–89.

[7] Brysbaert, M., and Ellis, A.W., Aphasia and age of acquisition: are early-learned words more resilient? *Aphasiology,* **2016.** 30(11): p. 1240–1263.

[8] Kambanaros, M., Group effects of instrumentality and name relation on action naming in bilingual anomic aphasia. *Brain and Language,* **2009.** 110(1): p. 29–37.

[9] Kambanaros, M., and Grohmann, K.K., Profiling (specific) language impairment in bilingual children: Preliminary evidence from Cyprus. *Solutions in the Assessment of Bilinguals,* **2013.** p. 146–174.

[10] McNemar, Q., Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika,* **1947.** 12(2): p. 153–157.

[11] R Development Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria. **2016.**

[12] Corp, I.B.M., IBM SPSS statistics for windows, version 22.0. IBM Corp Armonk, NY. **2011.**

[13] Tsolaki, M., The Neuropsychological Evaluation of the Elderly [in Greek]. *Thessaloniki.* **1997.**

[14] Kambanaros, M., et al., Action and object naming in schizophrenia. *Journal of Clinical and Experimental Neuropsychology,* **2010.** 32(10): p. 1083–1094.

[15] McDonald, W. I., et al., Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis.*Annals of Neurology,* **2001.** 50(1): p. 121–127.

[16] Kambanaros, et al., Verb--noun dissociations in relapsing-remitting multiple sclerosis: verb effects of semantic complexity and phonological relatedness. *Aphasiology,* **2017.** 31(1): p. 49–66.

[17] Kambanaros, M., et al., On the nature of verb--noun dissociations in bilectal SLI: A psycholinguistic perspective from Greek. Bilingualism: *Language and Cognition,* **2014.** 17(1): p. 169–188.

[18] Raven, J. C., et al., Manual for Raven's Progressive Matrices and Vocabulary Scales (Section 3)-Standard Progressive Matrices. San Antonio, TX: Harcourt Assessment. **2000.**

[19] Kambanaros, M., Verb and noun processing in late bilingual individuals with anomic aphasia. Flinders University of South Australia, School of Medicine. **2004.**

[20] Hatzigeorgiu, N., et al., Design and Implementation of the Online ILSP Greek Corpus. In LREC. **2000.**

[21] Kambanaros, M., Grohmann, K.K., and Michaelides, M., Lexical retrieval for nouns and verbs in typically developing bilectal children. *First Language,* **2013.** 33(2): p. 182–199.

[22] Wilcoxon, F., Individual comparisons by ranking methods. *Biometrics Bulletin,* **1945.** 1(6): p. 80–83.

[23] Hong, Y., On computing the distribution function for the Poisson binomial distribution. *Computational Statistics & Data Analysis,* **2013.** 59: p. 41–51.