European Journal of Physical and Rehabilitation Medicine
EDIZIONI MINERVA MEDICA

CROSS-CULTURAL ADAPTATION INTO ITALIAN AND VALIDATION OF THE FRENCHAY DYSARTHRIA ASSESSMENT – 2

Valentina RIOLO, Nicole PIZZORNI, Eleonora GUANZIROLI, Bruna AGOSTINIS, Megghi CONFORTOLA, Nicolò SCHETTINO, Giulia GILARDONE, Antonio SCHINDLER, Franco MOLTENI

European Journal of Physical and Rehabilitation Medicine 2021 Sep 09
DOI: 10.23736/S1973-9087.21.07029-5

Article type: Original Article

Article first published online: September 9, 2021
Manuscript accepted: August 31, 2021
Manuscript revised: July 20, 2021
Manuscript received: April 20, 2021

CROSS-CULTURAL ADAPTATION INTO ITALIAN AND VALIDATION OF THE FRENCHAY DYSARTHRIA ASSESSMENT – 2

Valentina Riolo[1], Nicole Pizzorni[2], Eleonora Guanziroli[1], Bruna Agostinis[3], Megghi Confortola[4], Nicolò Schettino[5], Giulia Gilardone[6], Antonio Schindler[2], Franco Molteni[1]

[1] Villa Beretta Rehabilitation Center, Valduce Hospital, Costa Masnaga, Italy.

[2] Department of Biomedical and Clinical Sciences "Luigi Sacco", University of Milan, Milan, Italy

[3] Fondazione Renato Piatti, Varese, Italy

[4] UOC Neuropsichiatria dell'Infanzia e Adolescenza, ASST Valtellina e alto Lario, Sondrio, Italy

[5] ASST Lecco, Lecco, Italy

[6] Department of Neurorehabilitation Sciences, Casa di Cura del Policlinico, Milan, Italy

Corresponding author:

Nicole Pizzorni
Ospedale L. Sacco,
Via GB Grassi 74,
20157 Milan, Italy
Tel. +39 0239043207
e-mail: nicole.pizzorni@unimi.it

Running title: Validation of the Italian FDA-2

Abstract

Background A comprehensive evaluation of dysarthria is required to make an accurate differential diagnosis with other communication disorders and plan effective rehabilitation programs. The Frenchay Dysarthria Assessment-2 (FDA-2) is a valid, reliable and widely-used protocol for the assessment of dysarthria. An Italian version of the FDA-2 is currently lacking.

Aim To perform a cross-cultural adaptation of the FDA-2 in Italian and to validate the Italian version of the FDA-2

Design Validation study

Setting Inpatient rehabilitation center

Population 69 patients with dysarthria and 112 healthy controls.

Methods The FDA-2 was translated and cross-culturally adapted to Italian. The validation study was carried out in 4 steps: (1) 42 audio-recorded samples of FDA-2 items from 11 patients with dysarthria were independently assessed by 7 speech and language pathologists for interrater reliability and re-assessed after 6 weeks for intrarater reliability; (2) 11 patients were simultaneously assessed by 3 speech and language therapists for interrater reliability of the whole Italian version of the FDA-2 and re-assessed within 24 hours for test-retest reliability; (3) the Italian version of the FDA-2 was administered to 112 healthy volunteers to gain normative data; (4) 49 patients with different types of dysarthria were assessed using the Italian version of the FDA-2, the Therapy Outcome Measure impairment scale and the Robertson Profile for the validity analysis.

Results Interrater and intrarater reliability ranged from good to excellent (ICC >0.75) except for 3 audio-recorded items. The overall protocol demonstrated excellent (ICC >0.9) inter-rater and test-retest reliability for all the sections and the total score. Normative data were gained for 6 age groups. For the validity analysis, a statistically significant difference was found between dysarthric patients and healthy subjects for all sections and the total score. The FDA-2 significantly correlated to the Therapy Outcome Measure (r=0.75) and the Robertson Profile (r=0.81).

Conclusion The Italian version of the FDA-2 yield satisfactory reliability and validity, comparable to the psychometric properties of the original version.

Clinical Rehabilitation Impact Speech and language pathologists can rely on a valid and reliable tool in Italian for the assessment of dysarthria in both clinical and research practice.

Keywords: dysarthria – speech-language pathology – validation study

Introduction

Dysarthria is a motor speech disorder occurring after a central or peripheral nervous system damage and resulting in respiration, phonation, resonance, articulation, and prosody disturbances[1,2]. The main symptoms are weakness, slowing, incoordination, altered muscle tone, and inaccuracy of oral and vocal movements, attributed to an abnormal muscular activation of the structures involved in speech production[3]. Dysarthria can be congenital, caused by prenatal or perinatal brain damages, or acquired, due to cerebrovascular accidents, traumatic brain injuries, tumors or progressive neurological diseases[2]. Several classification systems of motor speech disorders have been proposed in the last years. The most commonly used was provided by the Mayo Clinic group, who developed a cogent organization and classification of speech symptoms[4]. The Mayo Clinic classification system is based on the correspondence between the underlying neurological impairment and the perceptual features of speech and distinguishes between six different types of dysarthria: spastic, ataxic, hypokinetic, hyperkinetic, flaccid and mixed. Another dysarthria type named unilateral upper motor neuron dysarthria, occurring after unilateral upper motor neuron lesions, was added in 1975 by the same authors[1].

Although data regarding dysarthria prevalence within the general population are not available, it is not a rare condition with a severe impact on people quality of life (QOL). Even mild dysarthria may have consistent social and psychological effects reducing speech intelligibility and leading to isolation, depression and loss of independence[5].

A comprehensive evaluation of dysarthria is required to make an accurate differential diagnosis with other communication disorders (i.e. apraxia of speech, aphasia) and plan effective rehabilitation programs. A thorough documentation of the patient's medical history should be gathered and followed by detailed perceptual analyses of all the motor system components involved in speech production (respiration, phonation, resonance,

articulation, and prosody). Muscle strength, speed, range, stability, tone, and accuracy of muscular movements should be considered too[1].

Using valid and reliable instrument is a crucial component of research quality and clinical practice. Reliability allows evaluating the stability of measures administered at different times to the same subjects (test–retest reliability), estimating the equivalence of sets of items from the same test (internal consistency) and establishing the equivalence of ratings obtained by different observers (interrater reliability)[6]. Validity refers to the degree to which an instrument measures the construct it purports to measure[6].

To date, only few validated assessment tools for the evaluation of dysarthria are available worldwide. In Italy, the Robertson Dysarthria Profile is the only available test for dysarthria assessment[7]. Normative data were obtained by testing 60 healthy subjects[7] and reliability was analyzed in 50 patients with dysarthria[8]; however, validity data are lacking. Another validated dysarthria assessment tool is the Frenchay Dysarthria Assessment (FDA), developed in English by Enderby[9]. The FDA was originally designed based on the Mayo clinic classification system[4] to identify the nature and the patterns of oro-motor movements associated with different neurological diseases. A second edition of the FDA (FDA-2) was provided by Enderby and Palmer in 2008, with improved psychometric properties (good feasibility, reliability, inter-rater agreement for the total score and construct validity)[10]. The FDA-2 has been translated and adapted into European Portuguese[11] and into French[12] and its use is widely spread in several countries. Moreover, some authors considered the FDA-2 as the only diagnostic test for dysarthria, being able to differentiate among different types of dysarthria[13]. However, an Italian version of the FDA-2 is currently unavailable.

The aim of this study is to perform a cross-cultural adaptation of the FDA-2 in Italian and to validate the Italian version of the FDA-2. The hypothesis was that the Italian version of the FDA-2 would yield similar psychometric properties to the original version.

Materials and methods

This validation study was carried out according to the Declaration of Helsinki and was approved by the Institutional Review Board (Comitato Etico dell'Insubria, Chair-person Dr Angelo Carenzi, n. 173/2020, approval date 02.03.2021). The original authors of the FDA provided their authorization to the back-translation and

validation process. Patient were recruited in a inpatient rehabilitation unit. All subjects included in the study gave their written informed consent. All data were collected prospectively. The study consisted of 4 different phases: item translation (phase 1), reliability analysis (phase 2), normative data generation (phase 3) and validity analysis (phase 4).

FDA-2

The FDA-2 is a clinical protocol for the measurement, differential description, and diagnosis of dysarthria. The tool is divided into seven sections, each containing a set of performance tasks related to speech function: reflexes (ratings for cough, swallow, and drool); respiration (rating at rest and in speech); lips (ratings for at rest, spread, seal, alternate, and in speech); palate (ratings for fluids, maintenance, and in speech); laryngeal function (ratings for time, pitch, volume, and in speech); tongue (ratings for at rest, protrusion, elevation, lateral, alternate, and in speech); intelligibility (ratings for words, sentences, and conversation). An additional session allows to describe potential influencing factors (including hearing, sight, teeth, language, mood, posture, speech rate, and sensation). The FDA-2 can be approximately administered in 20-30 minutes by qualified clinicians.

The FDA-2 rating scale has five "best-fit" descriptors ranging from "a" (normal) to "e" (severe abnormal):

a = normal age; b = mild abnormality noticeable to skilled observed; c = abnormality obvious but can perform task/movements with reasonable approximation; d = some production of task but poor in quality, unable to sustain, inaccurate or extremely labored; e = unable to undertake task/movement/sound. Moreover, the rating scale includes nine points, with "a" corresponding to 8 and "e" corresponding to 0.

These descriptors are unlikely to fit a patient's performance exactly. Rather, they give a general impression of the type and severity of dysarthria. The test results can be recorded on the FDA-2 rating form.

Phase 1: FDA-2 items translation and cross-cultural adaptation

Items of the original FDA-2 English version[10] were translated into Italian and then back-translated into English by qualified professional translators, according to Chiorri and colleagues[14]. Four Speech and Language Pathologists (SLPs) ensured the unanimity of the interpretation of the items and of the scoring system. Doubts concerning the test administration or scoring were discussed with the original author.

A literal translation was retained for most items. Concerning the words and sentences for the "in speech" tasks, the selection of Italian words and sentences was based on the occurrence of phonemes and the phonetic structure of words in Italian[15]. Words with different length containing Italian phonemes were selected following the original English word list for the "intelligibility" section[10]. The words frequency was higher than 10 per million[15]. The Italian version of the FDA-2 and the complete manual can be requested by contacting the corresponding author.

Phase 2: Reliability analysis

The original reliability testing process was followed to establish both the FDA-2 inter-rater and intra-rater reliability[10]. Nine people with different type and severity of dysarthria were recorded while performing the audible tests (respiration in speech, lips seal, palate in speech, maximum phonation time, voice pitch, voice volume, voice in speech, tongue alternate, tongue in speech). The inclusion criteria were age ≥18 and diagnosis of dysarthria. Patients with altered morphology of speech organs were excluded. Auditory tests have been used for reliability testing, since they are considered relatively more subjective and therefore difficult to evaluate. Seven samples of the maximum phonation time and 5 samples for each of the other above-mentioned audible tests were randomly selected. Thus, overall, 42 examples of FDA-2 tasks were independently judged by seven SLPs (offline reliability analysis). All the judges underwent a specific training of the scoring of the FDA-2 by a senior SLP. A cross section of examples from mild to severe was selected for each task by a senior SLP. Five examples of each of the tasks listed above and seven examples of maximum phonation time were presented to the judges. Digital recordings were presented via external speakers in a quiet room. Scoring instructions were given to the judges and each recording was played twice. In order to evaluate the FDA-2 intra-rater reliability, the same forty-two tests were presented in the same way to the same listeners after a six-week interval.

Additionally, eleven dysarthric patients based on the same inclusion criteria were tested using the FDA-2 by three trained judges (SLPs) in order to verify the inter-rater reliability. The three judges simultaneously attended to the assessment session, carried out by a fourth SLP, but independently completed the rating form (online reliability analysis).

Moreover, the eleven dysarthric patients were re-assessed using the FDA-2 in order to verify the test-retest reliability. Re-test was performed by one of the three judges within 24 hours from the first assessment to avoid

significant clinical modification.

Characteristics of the patients' sample for reliability analysis are reported in Table I.

Phase 3: Normative data

112 asymptomatic control subjects, 55 males and 57 females, aged between 18 to 88 years were included to establish normative data. Exclusion criteria were speech structures alterations and known neurological diseases. Moreover, for patients aged ≥70, the additional exclusion criteria of a Mini Mental State Examination ≤23 was added[16]. Subjects were divided into six age groups: 18 subjects in group A (18-29 years), 20 subjects in group B (30-39 years), 18 subjects in group C (40-49 years), 20 subjects in group D (50-59 years), 18 subjects in group E (60-69 years), 18 subjects in group F (≥70 years). At least 9 males and 9 females for each age group were recruited. All of them were assessed with the FDA-2 by the same SLP.

Phase 4: Known-group and criterion validity

For the validity analysis, inclusion criteria were: patients with dysarthria of any type[13] and sufficient cognitive status and language comprehension to follow clinicians instructions during the administration of the FDA-2. Forty-nine adult patients were recruited, 31 males and 18 females, aged on average 53 (range 20-81), exhibiting different type of dysarthria[13]: 11 patients had spastic dysarthria, 13 flaccid dysarthria, 13 ataxic dysarthria and 12 hypokinetic dysarthria.

For known-group validity, patients from the clinical group were assessed using the FDA-2 by a trained SLP. FDA-2 scores from the clinical group of patients with dysarthria were compared to the FDA-2 scores of the control group recruited for normative data generation.

To assess criterion validity, patients with dysarthria were also assessed with the Robertson Dysarthria Profile[7,17] and the Therapy Outcome Measure[18], within the same session of the FDA-2 administration and by the same SLP. The Robertson Dysarthria Profile is the only available test for dysarthria assessment in Italian. It is divided into eight domains (respiration, phonation, facial musculature, diadochokinesis, oral reflexes, articulation, intelligibility, prosody) for a total of 71 items, each rated on a 4-point scale from 1 (most deviant performance) to 4 (normal performance). For the purpose of the present study, only the items of the Robertson Profile assessing the same construct of the items included in the FDA-2 were considered. In particular, a total

of 45 items from the Robertson Profile were analyzed: 7 items for reflexes, 5 items for respiration, 4 items for lips, 2 items for palate, 12 items for laryngeal function, 9 items for tongue, and 6 items for intelligibility. The TOM is a scale exploring 4 different domains (impairment, disability, participation and well-being), each rated on a 6-point ordinal rating scale, with 0 representing severe dysfunction and 5 normal functioning for humans given age, gender and culture. Only the TOM impairment domain was considered for the present study.

Finally, analogously to the original study[10], the scores of each FDA-2 item were separately analyzed according to the dysarthria type (spastic, flaccid, ataxic, hypokinetic) to verify the ability of the FDA-2 to distinguish among different types of dysarthria.

Statistical Analysis

Statistical analyses were performed with the IBM SPSS Statistics 25.0® package for Windows (SPSS Inc, Chicago, IL).

Inter-, intra-, and test re-test reliability was calculated using the two-way random Intraclass Correlation Coefficient (ICC) for single measures with 95% confidence interval (CI95%). ICC values less than 0.5, between 0.5 and 0.75, between 0.75 and 0.9, and greater than 0.90 will be considered indicative of poor, moderate, good, and excellent reliability, respectively[19].

Since normality assumption was violated within the age groups based on Kolmogorov-Smirnov normality test, the non-parametric Kruskal-Wallis test with post hoc Dunn's test and Bonferroni correction for multiple comparisons was used to compare different age groups among normative data. Data are reported as median, interquartile range (IQ range), and range. Significance was set at $p<0.05$.

FDA-2 known-group validity was verified comparing the dysarthric patients' scores with normative data using Mann-Whitney U test. For criterion validity testing, Spearman correlations between the FDA-2, the Robertson dysarthria profile and the TOM impairment score were calculated. Finally, to analyze the ability of the Italian version of the FDA-2 to distinguish among different dysarthria types, a profile of the speech characteristics of each type of dysarthria (spastic, flaccid, ataxic, and hypokinetic) was derived. For each type of dysarthria, the mean and the standard deviation obtained by the patients in the items of the FDA-2 was computed and graphically represented. Analogously to the original study[10], no statistical analysis was performed, but the profiles of the different dysarthria types were qualitatively compared.

Results

Reliability

To analyze interrater reliability, ICCs were calculated based on the scores assigned by the 7 judges at the first presentation of audio-recordings. To analyze intra-rater reliability, the same analysis was made between scores assigned by each judge at the first presentation of audio-recordings and those assigned during the second presentation. Results are reported in Table II. Interrater and intrarater reliability were poor (ICC <0.5) for the item "lips in speech", moderate (ICC 0.5-0.75) for the item "palate in speech" and "laryngeal volume", and ranged from good to excellent (ICC >0.75) for the remaining items.

Furthermore, inter-rater reliability of the whole protocol was assessed using ICC based on the scores of the 3 judges during the face-to-face assessment and a re-test was performed by a judge within 24-hours from the 1st assessment. Both inter-rater and test-retest reliability were excellent (ICC >0.9) for all the sections and the total score (Table III).

Normative data

FDA-2's total and sections scores for each age group are in Table IV. P-Values of the Kruskal-Wallis test for the comparison among age groups are shown in Table IV, p-values of the post hoc analysis are reported in Table V.

The group of subjects aged >70 (F) scored significantly lower than all the other age groups in the total score and in the Tongue section. Moreover, the group >70 (F) had significantly lower score than: (i) the group 18-30 (A) in the sections Respiration, Lips, Palate, and Laryngeal; (ii) the group 30-40 (B) in the sections Lips and Laryngeal; (iii) the group 40-50 (C) in the sections Respiration, Palate, and Laryngeal; and (iv) the group 50-60 (D) in the Respiration section. The group 60-70 (E) significantly differed from the group 40-50 (C) in the respiration section. No significant difference among the age groups for the Reflexes section. Finally, no comparison was performed for the Intelligibility section because all subjects reached the highest score of 24/24.

Known-group and criterion validity

For known-group validity analysis, FDA-2 scores of the patients were compared with the scores of the normative sample. Results of the Mann-Whitney U test are shown in Table VI. A statistically significant difference between dysarthric patients and healthy subjects was found for the total score and the score of all the FDA-2 sections.

For criterion validity, the FDA-2 total score significantly correlated with both Impairment section of TOM impairment score ($r=0.75$, $p<0.05$; TOM impairment median 3, IQ range 2-4) and the Robertson Profile total score ($r=0.81$, $p<0.05$) and all selected sections ($r>0.51$, $p<0.05$). Correlation coefficients between FDA-2 and each section of the Robertson Profile are shown in Table VII.

The profile of the speech characteristics of each type of dysarthria, based on the FDA-2, is depicted in Figure 1. Patients with spastic dysarthria showed, on average, the worst scores on most of the items. Patients with flaccid dysarthria scored generally higher in the intelligibility compared to the other dysarthria types. Patients with ataxic dysarthria showed lower scores in the items requiring movements' coordination compared to other tasks. Patients with hypokinetic dysarthria exhibited lower scores in the laryngeal domain compared to the other domains.

Discussion

The social and psychological impact of dysarthria on patients' life is well known[5]. Comprehensive, valid, and reliable assessment tools play an essential role in supporting diagnosis, planning treatment and monitoring the patient's progress and enable both clinicians and researchers to work efficiently. To date, the Frenchay Dysarthria Assessment is one of the few dysarthria assessment tools that have been validated and standardized using appropriate psychometric criteria. This test was first published in English in 1983 and then translated into several languages. As reported by Enderby and Palmer, validity and reliability of the FDA-2 were highlighted by several studies conducted in different countries with different patient groups[10]. The aim of the present study was to provide the cross-cultural adaptation and validation of the FDA-2 into Italian and supply a comprehensive protocol of motor speech assessment for Italian dysarthric patients. The psychometric properties of the FDA-2 were studied in a group of 49 patients with dysarthria following acquired brain injury while normative

data were collected administrating the whole protocol to 112 healthy subjects. Results showed optimal inter-rater and test-retest reliability and known-group and criterion validity, in agreement with the original version of the test[10].

Reliability

Overall, results showed excellent inter-rater and test-retest reliability of the protocol, supporting the consistency and stability of the original version of the test[10] and similarly to the European Portoguese adaptation of the test[11]. On the one hand, different examiners who independently administer the FDA-2 will produce similar ratings in judging the same patient's performance; on the other hand, the same clinician will be consistent in his judgment over time.

However, some differences in terms of consistency were found considering each individual item: some of them were not scored consistently either between or within examiners. In particular, inter-rater and intra-rater reliability were poor for the item "lips in speech" and moderate for the item "palate in speech" and "laryngeal volume", while ranged from good to excellent for the remaining items. Low reliability coefficients were found for audible tasks were found also in the original version[10]. However, it should be noted that for the 2 items with the lowest ICC, "lips in speech" and "palate in speech", the assessment only based on audio-recordings may have negatively impact on the reliability. Indeed, the operational definitions for the item "lips in speech" include both information on the perceptual production of bilabial phonemes and on the amplitude of the lips movements during speech, that could not be assessed with audio-recordings. Analogously, for the "palate in speech" the manual of the FDA-2 suggests to place a finger on the bridge of the nose or to use a mirror under the patient's nose during the assessment to improve reliability. Thus, the reliability of these items may have been underestimated by the assessment condition.

Normative data

Overall, 112 healthy subjects, 55 males and 57 females, aged 18-88 years were included to provide normative scores. The effect of age was investigated. Concerning the total score, results highlighted the presence of significantly lower scores in the older group of subjects (age>70) than in the younger ones. Indeed, the progressive reduction of average scores combined to the progressive increase of interquartile ranges suggest that younger

subjects usually have better and more homogeneous abilities while seniors show a progressive drop of the performance and higher inter-subject variability[20].

To identify those parameters that are more influenced by age and with the greater impact on the total score, each single section of the protocol was considered. The Tongue, Respiration, and Laryngeal sections were more impaired by aging, analogously to previous findings[20]. Tongue function is known to decline with aging because of sarcopenia and changes in muscle fibers' composition[21-24]. In older subjects, literature shows that tongue strength is reduced[25-26] and tongue movements are slower and less regular in the rhythm[27]. In typical aging, inspiratory and expiratory muscle strength, lung volume, and pulmonary reserve are reduced[28]. The coordination between respiration and speech production is altered as well. Studies have demonstrated that older adults initiate speech at a higher lung volume, use a greater percent of their vital capacity per syllable, and produce fewer syllables per breath than younger adults[29-31]. Concerning laryngeal function, voice changes with normal aging have been extensively studied. Typical anatomical modifications are smaller diameters of vocal muscle fibers due to atrophy and thinner superficial lamina propria, leading to incomplete glottal closure and altered vibratory patterns[32-34]. The result is a poor voice quality, characterized by breathiness and strain, a reduced vocal intensity, a decreased phonation time, and altered acoustic parameters[35-39].

Conversely, all 112 subjects reached the highest score in the Intelligibility section, showing intact abilities to read word and sentences carefully and hold a conversation without reductions of comprehensibility. This finding supports the hypotesis that age-related changes to speech-related organs are well compensated in normal aging to preserve speech intelligibility[40-42].

Known-group and criterion validity

For the clinical validity analysis, 49 patients with different type of dysarthria were included in the study.

To analyze known-group validity, FDA-2 scores obtained by the group of patients were compared to those obtained by the normative sample. Results highlighted the presence of a statistically significant difference between dysarthric patients and healthy subjects for both the total score and the score of all the FDA-2 sections. In particular, the FDA-2 scores were significantly lower in the patient group than those found in the control group. Those findings indicate that the FDA-2 may be a sensitive tool ably to identify dysarthria and discriminate between dysarthric and non-dysarthric subjects. Moreover, a significant correlation was found between

the FDA-2 total score and both the Impairment section of TOM and the Robertson Profile, confirming the validity of the tool in detecting the severity of the impairment.

According to the analysis performed by Enderby & Palmer in the manual of the FDA-2, a functional profile for different types of dysarthria was defined. Analogously to the original data[10]: 1. patients with spastic dysarthria showed the lowest scores in most of the items, with a tendency for lower scores in speech tasks compared to non-speech tasks; 2. patients with flaccid dysarthria were more intelligible than the other dysarthria types and exhibited a tendency for lower scores in non-speech tasks, according to the site of the neurological damage to the lower motor neurons; 3. patients with ataxic dysarthria had the lowest score in the item of the tongue for alternating movements. However, some differences between the original data[10] and the results from the present study should be underlined. Specifically, patients with ataxic dysarthria seemed to exhibit a greater speech impairment compared to the original data, exhibiting lower laryngeal control and more compromised intelligibility. Conversely, patients with hypokinetic dysarthria from the present study had a less severe dysarthria, as well as swallowing impairment, than patients from the original study. As no data on the underlying diseases was available, it was hypothesized that the recruited sample for the ataxic and the hypokinetic dysarthria's groups differed among the two studies for disease severity.

Study limitations

There are several limitations in this study. A first limitation concerns the sample size. Indeed, 112 healthy subjects and 49 patients were included in the present study. Although it represents the largest sample of patients and healthy subjects used for the validation of a dysarthria assessment tool in Italian, the sample size is smaller compared to the one of the validation of the original version[10]. Additionally, no patients with hyperkinetic or mixed dysarthria were included in the validation process because of the lack of this patient in the clinical practice during the recruitment period. Future research studies should increase the strength of these results with larger samples and all types of dysarthria. Secondly, the responsiveness of the FDA-2 was not studied. Although this step was not performed for the original version either, it may provide useful information on the ability of the tool to record changes of dysarthria severity over time. Therefore, this psychometric property should be investigated in the future.

Conclusion

The study represents the first validation of the original English FDA-2 in Italian to clinically assess the presence and the severity of dysarthria. The Italian version of the FDA-2 yield satisfactory reliability and validity for its application in clinical practice and for trans-cultural research in dysarthria.

Conflicts of interest The authors certify that there is no conflict of interest with any financial organization regarding the material discussed in the manuscript

Funding None.

Authors' contribution Study conception and design: Valentina Riolo, Antonio Schindler, Franco Molteni. Data collection: Valentina Riolo, Bruna Agostinis, Megghi Confortola, Nicolò Schettino. Statistical analysis: Eleonora Guanziroli. Resources: Franco Molteni. Supervision: Valentina Riolo, Antonio Schindler. Paper writing: Nicole Pizzorni, Giulia Gilardone. Paper revision: Valentina Riolo, Nicole Pizzorni, Giulia Gilardone, Antonio Schindler. All authors read and approved the final version of the manuscript.

References

[1] Darley FL, Aronson AE, Brown JR. Motor speech disorders. Philadelphia: Saunders; 1975.

[2] Enderby, P. Disorders of communication: dysarthria. In: Barnes MP, Good DC. Handbook of clinical neurology (Vol. 110). Amsterdam: Elsevier; 2013. pp. 273-281.

[3] Palmer R, Enderby P. Methods of speech therapy treatment for stable dysarthria: A review. Int J Speech Lang Pathol 2009;9:140-153.

[4] Darley FL, Aronson AE, Brown JR. Differential diagnostic patterns of dysarthria. J Speech Hear Res 1969;12:246-269.

[5] MacKenzie C, Muir M, Allen C, Jensen A. Non-speech oro-motor exercises in post-stroke dysarthria intervention: a randomized feasibility trial. Int J Lang Commun Disord 2014;49:602-617.

[6] Kimberlin CL, Winterstein AG. Validity and reliability of measurement instruments used in research. Am J Health Syst Pharm 2008;65:2276-2284.

[7] Fussi F, Cantagallo A. Profilo di valutazione della disartria. Adattamento italiano del test Robertson, raccolta di dati normativi e linee di trattamento. Torino: Ed Omega; 1999.

[8] De Biagi F, Frico AC, Turrola A, Nordio S, Berta G, Meneghello F. Italian validation of a test to assess dysarthria in neurologic patients: a cross-sectional pilot study. Otolaryngology (Sunnyvale) 2018;8:1000343.

[9] Enderby P. Frenchay dysarthria assessment. Br J Disord Comm 1980;15:165-173.

[10] Enderby P, Palmer R. Frenchay dysarthria assessment–Second edition (FDA-2). Austin, TX: Pro-ed; 2008.

[11] Cardoso R, Guimarães I, Santos H, Loureiro R, Domingos J, de Abreu D, et al. Frenchay dysarthria assessment (FDA-2) in Parkinson's disease: cross-cultural adaptation and psychometric properties of the European Portuguese version. Neurology 2017:264:21-31.

[12] Ghio A, Giusti L, Blanc E, Pinto S. French adaptation of the "Frenchay Dysarthria Assessment 2" speech intelligibility test. Eur Ann Otorhinolaryngol Head Neck Disord 2020;137:111-116.

[13] Duffy JR. Motor speech disorders: Substrates, differential diagnosis, and management. St Louis, MO: Elsevier; 2013.

[14] Chiorri C. [Teoria e tecnica psicometrica. Costruire un test psicologico]. New York: McGraw-Hill; 2011. p.24

[15] Bambini V, Trevisan M. [EsploraCoLFIS: Un'interfaccia web per ricerche sul corpus e lessico di frequenza dell'italiano scritto]. Quaderni del Laboratorio di Linguistica della Scuola Normale Superiore 2012;11:1-16. http://linguistica.sns.it/esploracolfis/home.htm

[16] Cockrell JR, Folstein M. Mini Mental State Examination. Hoboken: John Wiley & Sons; 2002.

[17] Robertson SJ. Dysarthria Profile. Tucson: Communication Skills Builders; 1982.

[18] Enderby P. Therapy Outcome Measure. San Diego, CA: Singular; 1997.

[19] Terry LB, Koo K, Mae Y. Li. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med 2016;15:155-163.

[20] Wallace GL. Assessment of oral peripheral structure and function in normal aging individuals with the Frenchay. J Commun Disord 1991;24:101-109.

[21] Yamaguchi A, Nasu M, Esali Y, et al. Amyloid deposits in the aged tongue. A post mortem study of 107 individuals over 67 years of age. J oral pathol 1982;11:237–244.

[22] Bassler R. Histopathology of different types of atrophy of the human tongue. Pathol res pract 1987;182:87–97.

[23] Nakayama M. Histological study on aging changes in the human tongue. J otolaryngol Jap 1991;94:541–555.

[24] Roos MR, Rice CL, Vandervoort AA. Age-related changes in motor unit function. Muscle Nerve 1997;20:679–690.

[25] Hara K, Tohara H, Kobayashi K, et al. Age-related declines in the swallowing muscle strength of men and women aged 20-89 years: A cross-sectional study on tongue pressure and jaw-opening force in 980 subjects. Arch Gerontol Geriatr 2018;78:64-70.

[26] Machida N, Tohara H, Hara K, et al. Effects of aging and sarcopenia on tongue pressure and jaw-opening force. Geriatr Gerontol Int 2017;17:295-301.

[27] Hirai T, Tanaka O, Koshino H, Yajima T. Ultrasound observations of tongue motor behaviour. J Prosth Dent 1991;65:840–844.

[28] Lowery EM, Brubaker AL, Kuhlmann E, Kovacs EJ. The aging lung. Clin Interv Aging 2013;8:1489-96.

[29] Hoit JD, Hixon TJ. Age and speech breathing. Journal of Speech and Hearing Research. 1987;30:351–366.

[30] Huber JE, Spruill I, John Age-related changes to speech breathing with increased vocal loudness. J Speech Lang Hear Res 2008;51:651–668.

[31] Huber JE. Effects of utterance length and vocal loudness on speech breathing in older adults. Respir Physiol Neurobiol 2008;164:323-30.

[32] Honjo, I, Isshiki, N. Laryngoscopic and voice characteristics of aged persons. Arch Otolaryngol 1980;106:149-150.

[33] Pontes, P, Brasolotto, A, Behlau, M. Glottic characteristics and voice complaint in the elderly. J Voice 2005;19:84-94.

[34] Martins RH, Benito Pessin AB, Nassib DJ, Branco A, Rodrigues SA, Matheus SM. Aging voice and the laryngeal muscle atrophy. Laryngoscope 2015;125:2518-21.

[35] Hodge, FS, Colton, RH, Kelley, RT. Vocal intensity characteristics in normal and elderly speakers. J Voice 2001;15:503-511.

[36] Golub, JS, Chen, P-H, Otto, KJ, Hapner, E, Johns, MM. Prevalence of perceived dysphonia in a geriatric population. J Am Geriatr Soc 2006;54:1736-1739.

[37] Maslan, J, Leng, X, Rees, C, Blalock, D, Butler, SG. Maximum phonation time in healthy older adults. J Voice 2011;25:709-713.

[38] Dehqan, A, Scherer, RC, Dashti, G, Ansari-Moghaddam, A, Fanaie, S. The effects of aging on acoustic parameters of voice. Folia Phoniatr Logop 2012;64:265-270.

[39] Vaca M, Mora E2, Cobeta I2. The Aging Voice: Influence of Respiratory and Laryngeal Changes. Otolaryngol Head Neck Surg 2015;153:409-13.

[40] Caruso AJ, Mueller PB, Shadden BB. Effects of aging on speech and voice. Phys Occup Ther Geriatr 1995;13:63-79.

[41] Goozée JV, Stephenson DK, Murdoch BE, Darnell RE, Lapointe LL. Lingual kinematic strategies used to increase speech rate: comparison between younger and older adults. Clin Linguist Phon 2005;19:319-34.

[42] Bennett JW, van Lieshout PH, Steele CM. Tongue control for speech and swallowing in healthy younger and older subjects. Int J Orofacial Myology 2007;33:5-18.

Table I. Demographic and clinical characteristics of the patients' sample for reliability analysis

Note. Offline analysis was performed by 7 judges on audio-recordings; online analysis was performed by 3 judges during face-to-face assessment session. Values are median (range), n/N (%), or as otherwise indicated.

		Offline analysis (n = 9)	**Online analysis (n = 11)**
Age (years)		46.3 (22-73)	48.2 (22-66)
Sex	M	4/9 (44.4%)	7/11 (63.6%)
	F	5/9 (55.6%)	4/11 (36.4%)
Education (years)		13.6 (5-18)	10.7 (5-13)
Medical diagnosis	Stroke	3/9 (33.3%)	3/11 (27.3%)
	Severe acquired injury	2/9 (22.2%)	3/11 (27.3%)
	Cerebellar ataxia	1/9 (11.1%)	2/11 (18.2%)
	Neurosurgical sequelae	1/9 (11.1%)	2/11 (18.2%)
	Guillain-Barré syndrome	1/9 (11.1%)	-
	Multiple sclerosis	1/9 (11.1%)	-
	Amyotrophic lateral sclerosis	-	1/11 (9%)

Table II. Interrater and intrarater reliability of the audible tasks based on the assessment of audio-recordings by 7 judges

Note. Reliability is reported as ICC (CI95%)

Item	**Inter-rater reliability**	**Intra-rater reliability**
Respiration in speech	0.932 (0.799-0.992)	0.855 (0.732-0.924)
Lips in speech	0.478 (0.146-0.899)	0.467 (0.163-0.690)
Palate in speech	0.622 (0.279-0.938)	0.840 (0.706-0.916)
Tongue in speech	0.743 (0.431-0.963)	0.779 (0.605-0.882)
Laryngeal time	0.942 (0.851-0.988)	0.919 (0.861-0.954)
Laryngeal pitch	0.843 (0.599-0.979)	0.814 (0.662-0.902)
Laryngeal volume	0.653 (0.313-0.945)	0.585 (0.317-0.767)
Laryngeal in speech	0.894 (0.706-0.986)	0.831 (0.690-0.911)

Table III. Inter-rater and test-retest reliability of the whole FDA-2 based on the face-to-face assessment by 3 judges

Note. Reliability is reported as ICC (CI95%)

FDA-2 section	Inter-rater reliability	Test-retest reliability
Total	0.992 (0.978-0.998)	0.996 (0.984-0.999)
Reflexes	0.941 (0.848-0.982)	0.998 (0.992-0.999)
Respiration	0.983 (0.953-0.995)	0.995 (0.983-0.999)
Lips	0.947 (0.863-0.984)	0.985 (0.944-0.996)
Palate	0.967 (0.911-0.990)	0.990 (0.964-0.997)
Laryngeal	0.975 (0.932-0.992)	0.995 (0.982-0.999)
Tongue	0.962 (0.900-0.989)	0.925 (0.748-0.979)
Intelligibility	0.971 (0.921-0.991)	0.988 (0.956-0.997)

Table IV: Comparison of the FDA-2 scores among age groups in the normative sample

Note. Significant p-values of the Kruskal-Wallis test are reported in bold

No comparison was performed for the Intelligibility section because all subjects scores 24/24

Fda-2 section		Age group						P
		A(18-30)	B(30-40)	C(40-50)	D(50-60)	E(60-70)	F (>70)	
Total	Median	206.5	206	206.5	205.5	205	197.5	**<0.001**
	Iq range	2.5	2.5	3	5	5.25	9.75	
	Min-max	203-208	202-208	203-208	195-208	191-208	178-206	
Reflexes	Median	24	24	24	24	24	24	**0.036**
	Iq range	0	0	0	1	0	1	
	Min-max	22-24	23-24	24-24	21-24	20-24	21-24	
Respiration	Median	16	16	16	16	16	15	**<0.001**
	Iq range	0	1	0	0	2.25	2	
	Min-max	14-16	15-16	16-16	15-16	11-16	10-16	
Lips	Median	40	40	40	39.5	40	39	**0.013**
	Iq range	0.75	0.75	1	1	1	2.5	
	Min-max	38-40	39-40	39-40	38-40	36-40	33-40	
Palate	Median	24	24	24	24	24	23.5	**0.010**
	Iq range	0	0	0	0	0	1	
	Min-max	23-24	22-24	22-24	21-24	23-24	21-24	
Laryngeal	Median	32	32	32	30.5	30	29	**<0.001**
	Iq range	1	1.75	2	3	3	3.25	
	Min-max	30-32	28-32	27-32	25-32	26-32	22-32	
Tongue	Median	48	48	48	48	48	44.5	**<0.001**
	Iq range	0	0	0.25	1	1	5.25	
	Min-max	46-48	45-48	45-48	43-48	43-48	35-48	

Intelligibility	Median	24	24	24	24	24	24	
	Iq range	0	0	0	0	0	0	
	Min-max	24-24	24-24	24-24	24-24	24-24	24-24	

Table V: P-values of the post hoc analysis for the comparison of the FDA-2 scores among different age groups in the normative sample

Note. Significant p-values are reported in bold

No comparison was performed for the Intelligibility section because all subjects scores 24/24

Age groups for post-hoc comparison		**FDA-2 section**						
		Total	Reflexes	Respiration	Lips	Palate	Laryngeal	Tongue
A	B	0.999	0.999	0.999	0.999	0.999	0.999	0.999
A	C	0.999	0.999	0.999	0.999	0.999	0.999	0.999
A	D	0.999	0.999	0.999	0.999	0.999	0.577	0.999
A	E	0.779	0.999	0.841	0.999	0.999	0.220	0.999
A	F	**<0.001**	0.999	**0.002**	**0.026**	**0.022**	**<0.001**	**<0.001**
B	C	0.999	0.999	0.328	0.999	0.999	0.999	0.999
B	D	0.999	0.419	0.725	0.999	0.999	0.999	0.999
B	E	0.999	0.999	0.999	0.999	0.999	0.999	0.999
B	F	**<0.001**	0.298	0.098	**0.018**	0.311	**0.001**	**<0.001**
C	D	0.999	0.183	0.999	0.999	0.999	0.999	0.999
C	E	0.999	0.999	**0.034**	0.999	0.999	0.999	0.999
C	F	**<0.001**	0.122	**<0.001**	0.231	**0.012**	**0.006**	**<0.001**
D	E	0.999	0.999	0.090	0.999	0.999	0.999	0.999
D	F	**<0.003**	0.999	**<0.001**	0.999	0.053	0.210	**<0.001**
E	F	**0.010**	0.999	0.999	0.123	0.131	0.542	**0.001**

Table VI: Comparison of FDA-2 scores between dysarthric patients and healthy subjects

Note. Significant p-values are reported in bold. Data are reported as median (IQ range)

FDA-2 section	**Dysarthric patients**	**Healthy subjects**	**p**
Total	151 (128.5-168.5)	206 (205.75-207.25)	**<0.001**
Reflexes	18 (12.5-20)	24 (24-24)	**<0.001**
Respiration	13 (8.5-14.5)	16 (15-16)	**0.015**
Lips	31 (27.5-33.5)	40 (39.75-40)	**<0.001**
Palate	21 (18-23)	24 (24-24)	**0.023**
Laryngeal	19 (13-24.5)	32 (31-32)	**0.001**
Tongue	34 (28-40)	48 (48-48)	**0.003**
Intelligibility	15 (12-23)	24 (24-24)	**0.030**

Table VII: Criterion validity: correlation between the FDA-2 and the Robertson Profile
Note. All correlations were significant $p<0.05$. Descriptive statistics of the Robertson Profile scores are reported as median (IQ range)

Section	Robertson Profile score	r
Total	132 (105-155)	0.81
Reflexes	23 (17-27)	0.75
Respiration	12 (8-16)	0.60
Lips	12 (10-14)	0.52
Palate	6 (6-8)	0.69
Laryngeal	28 (21.5-42)	0.82
Tongue	30 (24.5-35.5)	0.70
Intelligibility	18 (16-21.5)	0.74

FIGURE CAPTION
Figure 1: Functional profile of dysarthria based on the FDA-2 according to the type of dysarthria
a. Spastic; b. Flaccid; c. Ataxic; d. Hypokinetic
Note. The dot represents the mean score, the bar represents the standard deviation