Received: 16 July 2021 | Accepted: 5 May 2022

DOI: 10.1111/1460-6984.12737

International Journal of Language & Communication Disorders

RESEARCH REPORT

Validation and cross-linguistic adaptation of the Frenchay Dysarthria Assessment (FDA-2) speech intelligibility tests: Hebrew version

Michal Icht[1] | **Orly Bergerzon-Bitton**[1,2] | **Boaz M. Ben-David**[3,4,5]

[1]Department of Communication Disorders, Ariel University, Ariel, Israel

[2]The National Administration of Communication Disorders, Ministry of Health, Tel Hashomer, Israel

[3]Baruch Ivcher School of Psychology, Reichman University (IDC), Herzliya, Israel

[4]Department of Speech–Language Pathology, University of Toronto, Toronto, ON, Canada

[5]Toronto Rehabilitation Institute, University Health Networks (UHN), ON, Canada

Correspondence

Boaz M. Ben-David, Communication, Aging and Neuropsychology lab (CAN lab), School of Psychology, Reichman University (IDC), Herzliya, PO Box 167, Herzliya 4610101, Israel.

Email: boaz.ben.david@idc.ac.il

Abstract

'Dysarthria' is a group of motor speech disorders resulting from a disturbance in neuromuscular control. Most individuals with dysarthria cope with communicative restrictions due to speech impairments and reduced intelligibility. Thus, language-sensitive measurements of intelligibility are important in dysarthria neurological assessment. The Frenchay Dysarthria Assessment, 2nd edition (FDA-2), is a validated tool for the identification of the nature and patterns of oro-motor movements associated with different types of dysarthria. The current study conducted a careful culture- and linguistic-sensitive adaption of the two intelligibility subtests of the FDA-2 to Hebrew (words and sentences) and performed a preliminary validation with relevant clinical populations. First, sets of Hebrew words and sentences were constructed, based on the criteria defined in FDA-2, as well as on several other factors that may affect performance: emotional valence, arousal and familiarity. Second, the new subtests were validated in healthy older adults ($n = 20$), and in two clinical groups (acquired dysarthria, $n = 15$; and developmental dysarthria, $n = 19$). Analysis indicated that the new subtests were found to be specific and sensitive, valid and reliable, as scores significantly differ between healthy older adults and adults with dysarthria, correlated with other subjective measures of intelligibility, and showed high test–retest reliability. The words and sentences intelligibility subtests can be used to evaluate speech disorders in various populations of Hebrew speakers, thus may be an important addition to the speech–language pathologist's toolbox, for clinical work as well as for research purposes.

KEYWORDS

dysarthria, cross-language adaptation, Frenchay Dysarthria Assessment, Hebrew, intelligibility

What this paper adds

What is already known on the subject

- 'Dysarthria' is a group of disorders reflecting impairments in the strength, speed and precision of movements required for adequate control of the various speech subsystems. Reduced speech intelligibility is one of the main consequences of all dysarthria subtypes, irrespective of their underlying cause. Indeed, most individuals with dysarthria cope with communicative restrictions due to speech impairments. Thus, language-sensitive measurements of intelligibility are important in dysarthria assessment. The FDA-2's words and sentences subtests present standardized and validated tools for the identification of the nature and patterns of oro-motor movements associated with different types of dysarthria.

What this paper adds to existing knowledge

- The lack of assessment tools in Hebrew poses challenges to clinical evaluation as well as research purposes. The current study conducted a careful culture- and linguistic-sensitive adaption of the FDA-2 intelligibility subtests to Hebrew and performed a preliminary validation with relevant clinical populations. First, sets of Hebrew words and sentences were constructed, based on the criteria defined in FDA-2, as well as on several other factors that may affect performance: emotional valence, arousal and familiarity. Second, the new subtests were validated in healthy older adults ($n = 20$), and in two clinical groups (adults with acquired dysarthria, $n = 15$; and young adults with developmental dysarthria, $n = 19$).

What are the potential or actual clinical implications of this work?

- Analyses indicated that the new word and sentence subtests are specific, sensitive, valid and reliable. Namely, (1) they successfully differentiate between healthy individuals and individuals with dysarthria; (2) they correlate with other subjective measures of intelligibility; and (3) they show high test–retest reliability. The words and sentences intelligibility subtests can be used to evaluate speech disorders in various populations of Hebrew speakers. Thus, they may be an important addition to the speech–language pathologist's toolbox, for clinical and research purposes. The methods described here can be emulated for the adaptation of speech assessment tools to other languages.

INTRODUCTION

Dysarthria refers to a heterogenous group of motor speech disorders that result from damage to the central or peripheral nervous system. The disturbances in neuromuscular control may negatively affect one or more of the different speech subsystems, namely respiration, phonation, resonance, articulation and prosody (Enderby, 2013). The prevalence of dysarthria is relatively high. It is present in about 50% of stroke survivors (Mitchell et al., 2020), in about 90% of people with Parkinson's disease (Müller et al., 2001) and in up to 90% of individuals with cerebral palsy (Mei et al., 2014). This is true in Israel as well, in which about 15,000 strokes occur every year (ICDC, 2021), there are about 30,000 Parkinson's patients (The Israel Parkinson Association, 2021) and the incidence of cerebral palsy is relatively high (more than two cases per 1000 births), because it is more common in families with many children (e.g., from the Ultra-Orthodox, *Haredi*, community; Bartl et al., 2020).

Several methods and procedures have been offered for clinical assessment and diagnosis of dysarthria, evaluating the nature (or type) and severity of the speech impairments. Their linguistic and cultural adaptation is of clinical importance, as performance on speech tasks is crucial for correct assessment. Accurate and comprehensive clinical evaluation of dysarthria leads to the localization and differential diagnosis of neurological damage, guiding tailored rehabilitation for the specific abnormalities in different parts of the speech production system (Duffy, 2019).

Dysarthria can be classified into several types (e.g., Darley et al., 1975); each type is characterized by different perceptual features, and all by (some degree of) abnormal speech and reduced speech intelligibility (Duffy, 2019). Indeed, speech intelligibility is considered a primary indicator of dysarthria severity (Kent et al., 1989; Yorkston et al., 1999). The literature indicates that dysarthria is often characterized by imprecise articulation of segments (Carl et al., 2022; Icht, 2021), hypernasality, slowed rate and harsh vocal quality (Klasner & Yorkston, 2005). Imprecision of consonant production (e.g., reduced burst production, prolonged phonemes) may be related to impairments at the respiratory and/or articulatory levels (Allison & Hustad, 2018). Vowel centralization and reduced vowel space (Carl & Icht, 2021) are also common characteristics in dysarthric speech.

Given the key role of speech intelligibility in daily communication (Yorkston & Beukelman, 1981a), as well as in the evaluation and treatment of dysarthria, its assessment is clinically important (Ansel & Kent, 1992). Assessing the degree of clarity of the speaker is commonly performed by speech–language pathologists (SLPs), using various tools (Balzan et al., 2019). These often involve a variety of subtests using different speech materials, from phoneme (sounds in isolation) to words, phrases, sentences and spontaneous speech, aiming to gauge the effect of dysarthria on various levels of speech production (Weismer & Laures, 2002). For example, the Assessment of Intelligibility of Dysarthric Speech (ASSIDS; Yorkston & Beukelman, 1981b) is designed to evaluate intelligibility in single words and sentences.

One of the most prevalent commercially available tools for dysarthria diagnosis is the Frenchay Dysarthria Assessment, Second Edition (FDA-2; Enderby & Palmer, 2008). The original FDA tool (in English; Enderby, 1980; Enderby & Palmer, 1983) was developed using a focus group of SLPs. The tool was revised in 2008 (FDA-2), resulting in improved psychometric properties and a refinement of the domains of assessment. The FDA-2 was successfully translated into several languages (e.g., French, Ghio et al., 2019; European Portuguese, Cardoso et al., 2017; Arabic, Qutishat, 2015), and validated with different patient groups (e.g., patients with Duchenne muscular dystrophy, Hijikata et al., 2020; patients with Friedreich's ataxia, Eigentler et al., 2012; individuals with Parkinson's disease, Cardoso et al., 2017).

Currently, available tools to evaluate dysarthria in Hebrew, as well as a Hebrew version of the FDA-2, are lacking. Filling this existing gap, the goal of the current study is to perform and validate a Hebrew adaptation of the speech intelligibility subtests of the FDA-2, for both clinical and research purposes. The option for an adaptation (rather than constructing a new test) was decided based on the following factors: (1) the original test is well-known and acceptable; (2) for a cross-linguistic or cross-cultural comparisons, an adapted test is more suitable, producing an equivalent test in another language (Hambleton & Patsula, 1998); and (3) often adapting a test is considerably cheaper and faster than constructing a new test in a second language. The FDA-2 intelligibility section was chosen for adaptation because it is a familiar tool with a solid theoretical basis, translated into several languages and with good psychometric properties. Clinically, it is easy to use by SLPs.

The FDA-2 (Enderby & Palmer, 2008) includes a series of speech and articulatory evaluation tasks aiming to identify the specific dysarthria subtype which the patient appears to manifest (Darley et al., 1969; Darley et al., 1975). It uses different five-point scales by which the SLP rates the patient's performance on multiple tasks grouped under several sections: (a) reflexes, (b) respiration, (c) lips, (d) palate, (e) laryngeal, (f) tongue and (g) intelligibility.

A standard clinical assessment of dysarthria calls for culturally and linguistically sensitive diagnostic tools. Thus, a careful adaptation of the FDA-2 to languages different than the original one used (English) is called for. This adaptation is relatively simple for the abovementioned (a–f) sections, each consists of actions or tasks which are universal, oro-motor in nature and explicitly instructed. For example, the instructions for the lips' spread evaluation are: 'Ask the patient to give an exaggerated smile. Demonstrate and encourage the patient to exaggerate the attempt further.' Therefore, for these sections, the adaptation is based mainly on a standard translation methodology (translation, back-translation and experts' analysis; for an example in European Portuguese, see Cardoso et al., 2017). However, a plain word-for-word translation is clearly not suitable for the intelligibility (g) sections in which the patient is asked to read aloud lists of words and of phrases (short sentences; Ghio et al., 2019). Rather, the adaptation must consider the unique syllabic, phonetic, grammatic and pragmatic properties of the target language. Specifically, the word subset adaptation must control for factors such as word frequency, phonetic structure and grammatical type. The phrase subset adaptation should also consider morphosyntactic structure, tenses and modes, sentence predictability and pragmatics. In addition, words' arousal and valence were found to affect reading fluency (i.e., negative high-arousal words have been found to slow down reading, as compared with neutral low-arousal words; Ben-David et al., 2012), semantic processing and motor-related responses (Ben-David et al., 2016), as well as phoneme monitoring (Ben-David et al., 2016, 2011). Hence, a word's arousal and emotional valence may be also considered in a cross-cultural adaptation process.

The goal of the current study was twofold. First, adapting the FDA-2 word and sentence intelligibility subtests to Hebrew, carefully controlling for the core properties of the original version, and ensuring that the resulting Hebrew-adapted subtests are linguistically and culturally appropriate (Study 1). Second, verifying that the new subtests accurately identify individuals with dysarthria and correctly reject individuals without dysarthria (sensitivity and specificity of the subtests). Finally, assessing the validity and reliability of the adapted subtests. To that end, we administered the Hebrew adapted subtests to three groups of participants: healthy older adults, adults with acquired dysarthria and young adults with developmental dysarthria (Study 2).

STUDY 1: HEBREW ADAPTATION

The goal of the first experiment was to adapt the FDA-2 intelligibility subtests from the original English version to Hebrew. In these subtests, the patient is asked to read aloud a list of words and a list of phrases, and the examiner writes down what s/he has understood. The examiner's transcription is compared against the original lists, and intelligibility scores are calculated. The original test is comprised of relatively large pools of words and sentences to reduce the chance that clinicians and/or patients could memorize the stimuli through repeated exposure (e.g., on initial evaluation, and during therapy).

The following steps were taken to preserve the set of criteria used to develop the original English version (Enderby & Palmer, 2008): (1) the number of items in the word list (116), and in the sentence list (50) matched the original version; (2) words familiarity was maintained; (3) phonetic distribution and the syllable types and structures are typical to Hebrew; and (4) sentences approximated the linguistic structures of Hebrew and satisfy morphosyntactic rules.

METHOD

Ethics and source authorization

A written authorization to adapt the FDA-2 intelligibility subsection for scientific purposes was given by the author (P. Enderby) and the publisher and owner of the rights (PRO-ED, Inc.). The study was approved by the university ethics committee, and electronically signed informed consent was obtained from each participant.

Participants

A total of 150 young adults (46 males; age range = 18–40 years, $M = 25$, SD = 3 years) were recruited for the adaptation process. They were Israeli university students or their peers, recruited by publishing a call for participation on campus and via social media (Facebook). Participants received either partial course credit or volunteered for the study. All participants were native Hebrew speakers (their language level was assessed during a preliminary interview with a research assistant, a trained SLP student). None of the participants reported on: (1) language or learning problems, (2) vision or hearing problems and/or (3) neurological disease, as confirmed by an on-line demographic questionnaire (taken from Hadar et al., 2016).

Participants were randomly divided into four groups: (1) 40 were asked to rate words on familiarity; (2) 40 were asked to rate the same words on emotional valence; (3) 40 participants were asked to rate the words on arousal; and (4) 30 participants were asked to complete short phrases (fill in the blank task) to assess final-word predictability. The three first groups did not significantly differ in terms of average age, $F < 1$, and gender distribution, $p > 0.05$ (see Table 1 for participants' data).

Materials

Word list

The first two authors (M.I., O.B.B., experienced SLPs) constructed a list of 200 Hebrew words (of which the most suitable 116 were chosen), in accord with the criteria used to construct the English FDA-2 intelligibility subtests:

- Frequency: the selected words all have a minimum frequency of 10 per million, selected from a widely used Hebrew corpus (based on norms of Hebrew newspaper

TABLE 1 Participants' data

Task	*N*	Mean age (years)	Gender (male, female)	Years of education (mean)
Familiarity rating	40	24.5	7, 33	13.8
Emotional valence rating	40	25.8	15, 25	14.2
Arousal rating	40	25.6	7, 33	15.1
Sentence predictability rating	30	29.0	17, 13	14.8

texts; Frost & Plaut, 2005; see also Mama & Icht, 2016). Note: the familiarity rating task in the current study aimed to further verify that the final chosen set includes words representative of everyday discourse as well.

- Phonetic structure: phonological structure that is adequately representative of Hebrew, with each consonant occurring in various positions (initial, middle and final) across the words. Vowels were also fully represented (Carl & Icht, 2021).
- Syllabic structure: words containing different number of syllables (one to four syllables) and different syllabic structures (mainly CV and CVC, the most frequent syllables in Hebrew nouns; Ben-David & Bat-El, 2016) were chosen, representative of Hebrew word structure.
- Grammatical types: nouns, adjectives and verbs were included.

Sentence list

For the sentence intelligibility section of the FDA-2, a list of 80 different Hebrew phrases (of which the most suitable 50 were chosen) was constructed by the two first authors, in accordance with the criteria used to construct the English FDA-2 intelligibility subtests:

- Short different carrier phrases (two to three words), so the listener must listen to a sentence, not just interpret the key word in a standard carrier phrase, each followed by a keyword.
- Keywords were phonetically balanced, to account for place, manner and position, approximating typical Hebrew, taken from the Hebrew monosyllabic phonetically balanced speech discrimination test (Auditory Word Discrimination Test; Putter-Katz et al., 2002). Note: in the English version, the vast majority of keywords are monosyllabic (48 of 50), a single word is bisyllabic ('noisy') and a single word is trisyllabic ('theatre'), while in Hebrew all 50 keywords were monosyllabic to avoid possible biases.
- Sentences satisfy morphosyntactic Hebrew rules, including questions and exclamation phrases as well, to reflect naturalistic language production (Friedmann & Lavi, 2006).

Word rating tasks

A total of 200 words were rated on familiarity, valence and arousal. Word order was randomized for each participant, such that no word order was repeated for two participants. Participants were encouraged to base their ratings on their immediate impression of the word. Word familiarity was rated using a four-point ordinal scale, with 1 indicating a highly non-familiar word and 4 indicating a highly familiar word (for a similar rating scale, see Parks & Toth, 2006). Emotional valence and arousal were rated using the nine-point Self-Assessment Manikin (SAM) scales (Bradley & Lang, 1994), a non-verbal pictorial assessment technique that directly measures the pleasure, and arousal associated with a person's affective reaction to a wide variety of stimuli. The Valence SAM scale ranges between 1, extremely positive, to 9, extremely negative; the Arousal SAM scale ranges between 1, extremely unaffected, to 9, extremely affected (for similar rating scales, see France et al., 1994).

Sentence completion task

The 80 sentences were rated on their predictability. The final word of each sentence was deleted and replaced by a fill-in blank space. Participants were asked to write the word that, in their opinion, best fits the blank space (for a similar procedure, see Geetha et al., 2014). Sentence's order was randomized for each participant. Predictability ratings were used to exclude highly predictable sentences because these may elevate intelligibility scores compared with sentences with low predictability (Geetha et al., 2014).

Procedure

As a first step, each participant logged into an on-line platform (Psychopy v.3; a free cross-platform package allowing a wide range of experiments in the behavioural sciences to be conducted), filled in the demographic questionnaires anonymously and was randomly assigned to one of the four groups. Following this, each participant completed his/her assigned rating task.

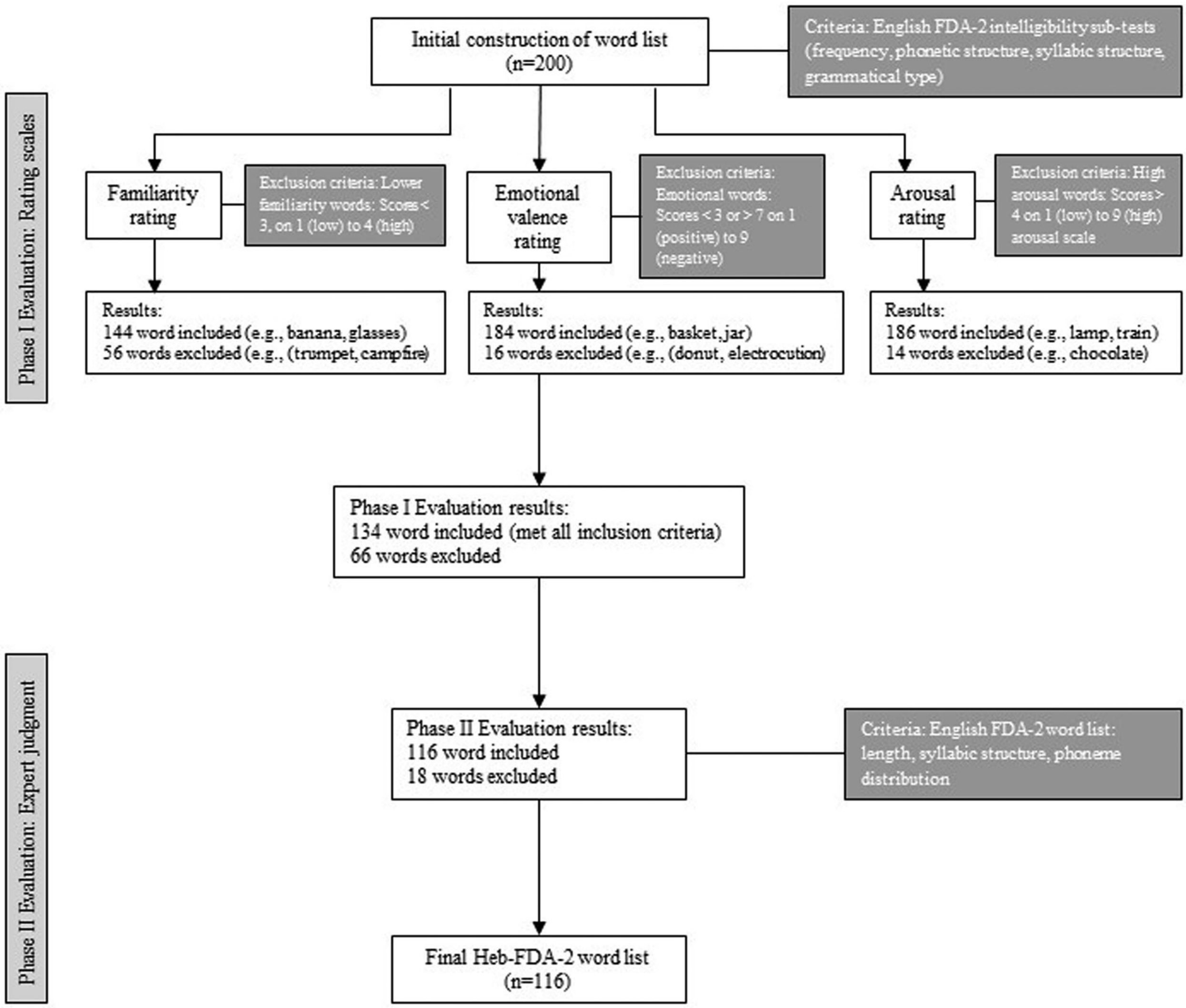

FIGURE 1 Dual-phase construction of the word list (Phase I: Rating scales; and Phase II: Expert judgment) and results

Based on the distribution of responses in each rating task, cut-off points were selected. We excluded words rated by at least 60% of the participants as less familiar (rated < 3 on the familiarity scale), non-neutral (emotional words, rated < 3.0 or > 7.0 on the valence scale), and highly arousing (rated > 4.0 on the arousal scale; Phase I Evaluation). The list of words that met all inclusion criteria was further analysed by the authors (considering their length, syllabic structure and phoneme distribution) to select the final sample of 116 words (Phase II Evaluation). Sentences that were completed by a relatively small number of different words (1 SD below the mean, or less) were considered to be highly predictable, and excluded.

Statistical analyses

A dual-phase process was used to select the final set of words from the initial set of 200 words. Participants' responses were averaged to create three scores for each of the 200 tested words, on the following scales: familiarity, emotional valence and arousal. The above-mentioned criteria were applied. The remaining words were assessed (by the first two authors, both experienced SLPs) according to three variables: length, syllabic structure and phoneme distribution to finalize the list. A Wilcoxon signed-rank test, a two-dependent samples non-parametric paired test, was used to compare the English and Hebrew words in terms of length (number of syllables). Spearman rank order correlation and Kandel's Tau were used to determine the phonetic balance of the Hebrew word list (Lammert et al., 2020).

A similar dual-phase procedure was applied to select the final set of sentences. As a first step, the sentence completion predictability was calculated based on the number of different words used to complete each sentence. As a second step, the phonetic make-up of the keywords in the remaining sentences was assessed by the first two authors.

TABLE 2 Syllabic make-up of the final Hebrew word list: Length (number of syllables) and syllable structure

No. of syllables	Syllabic structure	Number of words (total = 116)	Suggested number of words to be selected by the clinician (total = 10)
Monosyllabic	Simple onset—CVC Complex onset—CCV; CCVC	509	21
Bisyllabic	CV.CV; CV.CVC; CVC. CV; CVC.CVC	24	3
Trisyllabic	CV.CV.CV; CV.CV.CVC; CVC.CV. CV; CVC.CV.CVC	21	2
Four-syllabic	CV.CV.CV.VC; CVC.CV.CV.CV; CV.CV.CV.CVC; CV.CVC.CV.VC; CVC.CV.CV.CVC	12	2

RESULTS

Word list

Figure 1 presents the dual-phase evaluation process of the word list, and its results. Listed are the number of words removed based on the different exclusion criteria: familiarity ($n = 56$), valence ($n = 16$) and arousal ($n = 14$); note that some words overlap (i.e., did not meet more than one inclusion criteria). Following the first evaluation phase, of the 200 words in the primary set, 66 words (33%) that did not meet at least one inclusion criteria were excluded.

In the next step (Phase II Evaluation), the set of 134 words was further evaluated by the authors to select the final sample of words. In line with the FDA-2 criteria, the words were analysed according to the following criteria:

- Words' length: one to four syllables, to verify that the word length is comparable in the English and Hebrew versions. The final sample matched the original FDA-2 characteristic, p = 0.89 (Wilcoxon signed-rank test, z = 0.135), with a high number of monosyllabic words to better assess speech intelligibility in dysarthria (Kent et al., 1989; Hustad, 2007).
- Syllabic structure: the final sample closely followed spoken Hebrew. The most frequent syllables in Hebrew nouns are CV and CVC. Syllables with complex onsets CCV(C) are relatively rare (about 4%), appearing mostly in word-initial position (Ben-David & Bat-El, 2016) (see Table 2 for the syllabic make-up of the final Hebrew word list—length and syllable structure).
- Phoneme distribution of target words: closely matched the Hebrew phonetic distribution (based on norms by Ben-David & Bat-El, 2016).

TABLE 3 Phonetic distribution of the final Hebrew word list as compared with the existing Hebrew norms

Hebrew consonants	Frequency norms (%) (Ben-David & Bat-El 2016)	Frequencies (%) of the final Hebrew word list
t	14.6	8.6
ʁ	9.7	10.9
m	9.1	12.4
n	8.8	6.5
k	6.8	5.9
l	6.6	7.7
x	5.4	5.6
s	5.1	3.8
j	5.0	1.8
d	4.2	4.4
ʃ	3.9	3.8
g	3.6	3.8
p	3.1	3.8
ts	3.1	3.6
v	3.0	5.3
f	2.8	3.6
b	2.6	3.6
z	2.3	1.8
ʔ	Not reported	2.7

Note: Hebrew consonants are reported as the International Phonetic Alphabet (IPA) transcription.

To ensure phonetic balance of the word list, we rank ordered phonemes in both sources, following Lammert et al (2020) (removing the phoneme /ʔ/ that was not counted by Ben-David & Bat-El, 2016) (Table 3; and see Appendix A in the additional supporting information). The word list

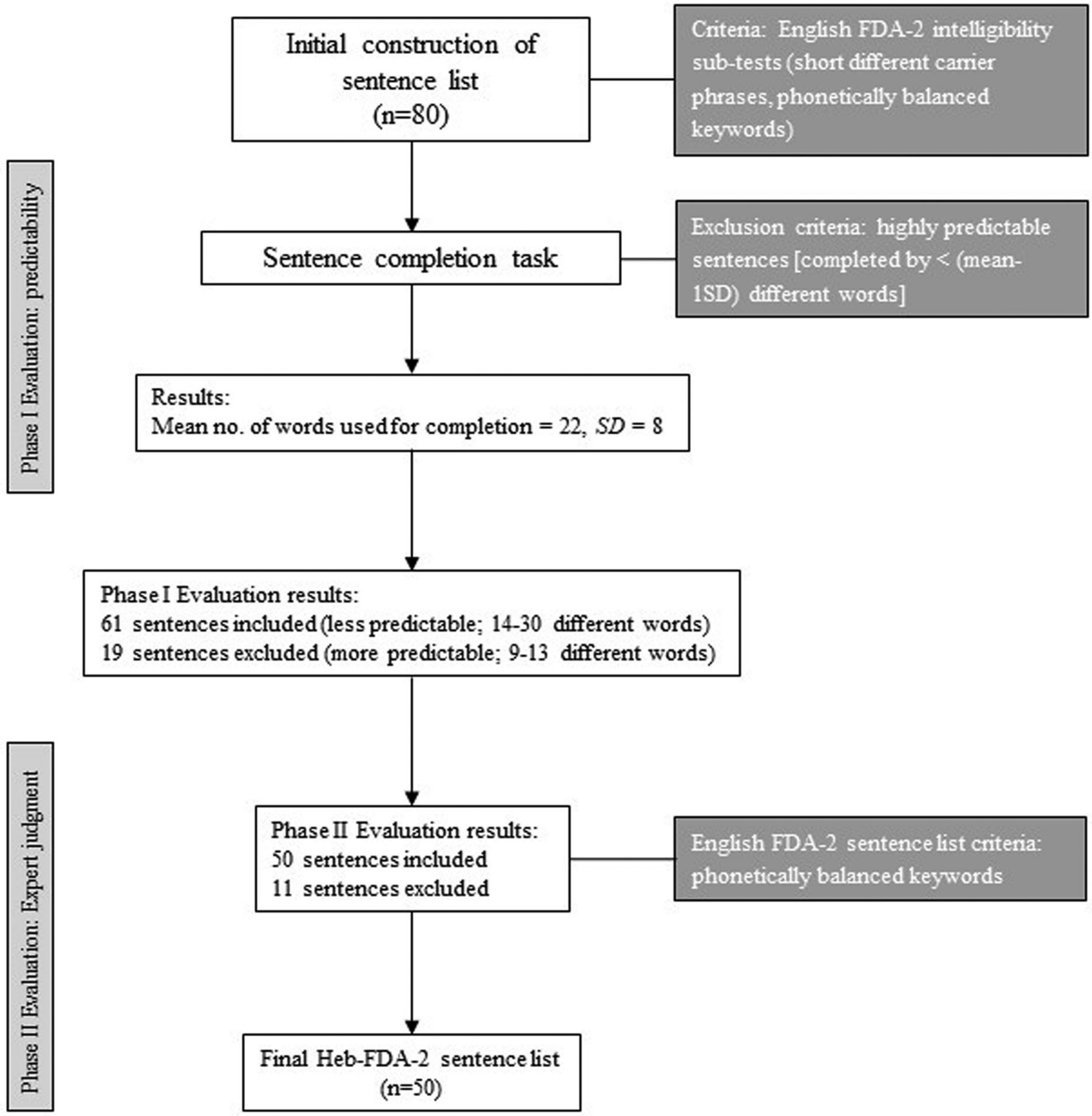

FIGURE 2 Dual-phase construction of the sentence list (Phase I: Sentence completion task; and Phase II: Expert judgment) and results

was found to match the norm for phonetic distribution in Hebrew, with a very high and significant Spearman rank order correlation, $\rho = 0.862$, $p < 0.001$ and a very high and significant Kandel's Tau, $\tau = 0.739$, $p < 0.001$.

Following this Phase II Evaluation, an additional 18 words were excluded, resulting in a final set of 116 words (mimicking the English version).

Sentence list

Figure 2 shows the dual-phase evaluation process of the sentence list, and its results. The participants' responses in the sentence task (fill-in blank task) were analysed (Phase I Evaluation). The number of different words used by the participants to complete the phrases was counted. The average number of different words used for phrases' completion was 22 (SD = 8, range = 9–30 different words). Based on these results, the relatively more predictable 19 sentences (that were completed by fewer than 14 different words; 1 SD below the mean) were excluded, resulting in a sample of 61 less predictable sentences. Finally, additional 11 sentences were filtered out in order to ensure that keywords (the final word in each phrase) are phonetically balanced in terms of place, manner and position (Phase II Evaluation), resulting in a final set of 50 sentences, mimicking the original FDA-2 (see Appendix B in the additional supporting information).

DISCUSSION

The present study aimed to describe the Hebrew adaptation of the FDA-2 intelligibility subtests, explain the criteria and choices made, and present the final word and sentences lists. The careful adaptation process followed a rigorous methodology and included an initial construction of two lists, words and sentences, based on the original FDA-2 English criteria (e.g., frequency, phonetic structure, syllabic structure, grammatical type). These lists were evaluated first using three different rating scales (to assess familiarity, emotional valence and arousal of the words, and predictability of the sentences), and then by expert opinions.

The final Heb-FDA-2 word list, as provided in Table 4, meets the criteria of the original English version as adjusted to the Hebrew language. Note that the selected words do not include a complete representation of the Hebrew linguistic diversity, since the list aims to provide a measure for dysarthria severity, rather than a complete analysis of the Hebrew phonological system (as can be found in an articulation test). Similarly, the final Heb-FDA-2 sentence list, as provided in Table 5, corresponds to the original tool in English. To validate the two new Heb-FDA-2 subtests, they were tested with clinical and non-clinical populations in Study 2.

STUDY 2: PRELIMINARY VALIDATION

The goal of Study 2 was to assess the validity of the Hebrew versions of the word and sentence subtests, Heb-FDA-2. Specifically, we attempted to verify that the new intelligibility subtests detect all individuals with dysarthria (who should therefore present low scores) and reject all non-clinical individuals. This would support the sensitivity and specificity of the tasks, respectively. Thus, the novel tests were administered to three groups of participants: healthy older adults, adults with acquired dysarthria and young adults with developmental dysarthria.

We also aimed to examine the validity and reliability of the new subtests. Participants also performed two commonly used word and sentence tasks in Hebrew (that have not been validated for intelligibility or for dysarthria assessment), for comparison. Construct validity was evaluated by measuring the correlations between the scores on the word and sentence subtests. Convergent validity was determined by comparing performance on the Heb-FDA-2 sentence task with subjective intelligibility scores assigned by an SLP for the clinical group of young adults with developmental dysarthria. Those participants were tested again after 3 months to establish test–retest reliability.

METHOD

Participants

Adults with acquired dysarthria

A total of 15 adults with acquired dysarthria were recruited for participation by contacting SLPs at several rehabilitation centres. Participants were nine males and six females, between the ages of 29 and 92 years (average age = 65.2 years, SD = 18.8 years), all had dysarthria (flaccid, spastic, ataxic or mixed, as assessed by their SLP) resulting from CVA (cerebrovascular accident, stroke, $n = 12$; post-acute or chronic phase), Parkinson's disease ($n = 2$) or following removal of a cerebellar tumour ($n = 1$). Cognitive status was confirmed by various tools: (1) The Hebrew version of the Mini-Mental State Examination (MMSE; Werner et al., 1999), $n = 3$; (2) the Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005), $n = 9$; (3) the Loewenstein Occupational Therapy Cognitive Assessment (LOTCA; Katz et al., 1989), $n = 1$; and (4) confirmation by the rehabilitation centres staff members, $n = 2$. To ensure a reliable performance of the cognitive tests, environmental factors were kept optimal (e.g., good lighting and face-to-face communication with the examiner). Additionally, participants were encouraged to use their personal hearing and/or visual aids (e.g., glasses; for comparable recommendations, see Ben-David & Icht, 2018). The participants (or their relatives or guardians) received written and oral information regarding the study and its goals, and gave their written or oral consent.

Young adults with developmental dysarthria

A total of 19 young adults with developmental dysarthria participated, 11 males and eight females, between the ages of 21 and 39 years (average age = 26.4 years, SD = 5 years). All were patients of a single day centre that provides services for over 20 adults with physical disabilities (e.g., cerebral palsy, muscular dystrophy), in partnership with the Rehabilitation Department of the Israeli Ministry of Welfare. Participants were diagnosed with various types of dysarthria (flaccid, spastic, ataxic or mixed, as assessed by the day centre SLP) resulting from cerebral palsy ($n = 14$), Pelizaeus–Merzbacher disease ($n = 2$), amyotrophic

TABLE 4 Final Hebrew word list

No. of syllables	Hebrew Words (phonetic transcription)							
Mono-syllabic with simple onset (50 words)	mas	מס	ʁak	רק	ʃuk	שוק	xag	חג
	bat	בת	tsel	צל	bad	בד	xof	חוף
	ʁax	רך	kav	קו	dat	דת	ʃen	שן
	lul	לול	maʁ	מר	ʔiʃ	איש	sid	סיד
	din	דין	guʁ	גור	kam	קם	lax	לח
	tsax	צח	kal	קל	dan	דן	zol	זול
	ʔeʃ	אש	sug	סוג	mot	מוט	mits	מיץ
	ben	בן	vav	וו	xut	חוט	pil	פיל
	jom	יום	ʃum	שום	ʔof	עוף	ʃai	שי
	giʁ	גיר	ʔets	עץ	tal	טל	haʁ	הר
	siʁ	סיר	ken	כן	kaf	כף	min	מין
	neʁ	נר	tei	תה	guf	גוף	paʁ	פר
					ta	תא	zan	זן
Mono-syllabic with complex onset (9 words)	kvaʁ	כבר	pʁi	פרי	zvuv	זבוב	tlai	טלאי
	bli	בלי	dli	דלי	tsxok	צחוק	sʁox	שרוך
							tʁis	תריס
Bi-syllabic (24 words)	ʔaʁnak	ארנק	maʁʔa	מראה	tapuz	תפוז	safsal	ספסל
	mapa	מפה	pilpel	פלפל	maxʃev	מחשב	dubi	דובי
	tsamid	צמיד	ʁadjo	רדיו	ʃulxan	שולחן	kova	כובע
	madaf	מדף	tsipoʁ	ציפור	bakbuk	בקבוק	ʁegel	רגל
	baʁvaz	ברווז	kapit	כפית	bietsa	ביצה	luʔax	לוח
	maʁak	מרק	xatul	חתול	sulam	סולם	pita	פיתה
Tri-syllabic (21 words)	xoveʁet	חוברת	mitlabeʃ	מתלבש	banana	בננה		
	mistaʁek	מסתרק	mistakel	מסתכל	mivʁeʃet	מברשת		
	megeʁa	מגירה	telefon	טלפון	magevet	מגבת		
	ʁakevet	רכבת	mevaʃel	מבשל	mizvada	מזוודה		
	mexonit	מכונית	metsajeʁ	מצייר	xagoʁa	חגורה		
	madʁega	מדרגה	matate	מטאטא	menoʁa	מנורה		
	ʁeʃima	רשימה	miʁpeset	מרפסת	nadneda	נדנדה		
Four-syllabic (12 words)	ʔagvanija	עגבניה	mevugeʁet	מבוגרת	ʔofanaim	אופניים		
	kaduʁegel	כדורגל	miʃkafaim	משקפיים	laxmanija	לחמניה		
	tsipoʁnaim	ציפורניים	mexubeʁet	מחוברת	melafefon	מלפפון		
	mitstaʁefet	מצטרפת	mispaʁaim	מספריים	metsaxtseʔax	מצחצח		

lateral sclerosis ($n = 1$), hypoxic ischemic encephalopathy ($n = 1$) or Duchenne muscular dystrophy ($n = 1$). Intact or borderline cognitive abilities were confirmed by the day centre staff members (due to physical or sensory disabilities, standard cognitive tests were not administered). The parents or legal representatives of a potential 22 participants received written information regarding the study and gave their written consent. Three patients did not consent and were not approached. All 19 participants gave their oral consent. Clinical participants' demographic data are presented in Table 6.

Healthy older adults

A total of 20 older adults from the general population were recruited from two different independent-living retirement homes, and two community centres. Participants

 International Journal of Language & Communication Disorders

TABLE 5 Final Hebrew sentence list (and phonetic transcription of final keywords)

Phrase	Phonetic transcription of final word	Phrase	Phonetic transcription of final word
ראיתי שם **דוב**	dov	הוא אמר **כן**	ken
אתמול היה **צום**	tsom	השיער שלה **רך**	ʁax
היה לו **חום**	xom	השולחן היה **קל**	kal
זה ממש **נס**	nes	הילדה אכלה **פרי**	pʁi
בחוץ יש **חול**	xol	אתה מחזיק **כלי**	kli
הטבח הכין **עוף**	ʔof	אתמול ראיתי **נר**	neʁ
הוא מצא **גור**	guʁ	הגשר הזה **זז**	zaz
אתמול קנית **בול**	bul	שתיתי תה **מר**	maʁ
הילד רוצה **עוד**	ʔod	בגן יש **זבוב**	zvuv
דן צייר **קו**	kav	הוא קנה **רום**	ʁum
הוא אכל **דג**	dag	השעה הייתה **שש**	ʃeʃ
הוא נתן **לה**	la	הבגד היה **לח**	lax
יש לו **טלאי**	tlai	ראיתי בשדה **אש**	ʔeʃ
יש כאן **צל**	tsel	לילדה קוראים **שיר**	ʃiʁ
מותר לך **גם**	gam	החוט הזה **טוב**	tov
זה מאוד **קר**	kaʁ	הילד צייר **לב**	lev
אין פה **קיר**	kiʁ	האם החדר **ריק**?	ʁek
אני אוהב **כרוב**	kʁuv	אתה מרגיש **זר**?	zaʁ
בחדר יש **אור**	ʔoʁ	אפשר להזמין **תה**?	tei
אין לו **גבול**	gvul	מתי אתה **קם**?	kam
קניתי לך **זר**	zeʁ	איפה יש **פח**?	pax
הכלב שלי **נח**	nax	אתה מוכן **כבר**?	kvaʁ
תראה איזה **נוף**!	nof	האם זה **בוץ**?	bots
תני לי **גיר**	giʁ	יש פה **סיד**?	sid
אבא הוסיף **שום**	ʃum	מתי הוא **טס**?	tas

were 11 males and nine females, between the ages of 61 and 71 years (average age = 66.4 years, SD = 3). All participants reported obtaining a high school and/or college diploma and showed normal cognitive abilities on the MoCA test (average scores = 27.5, SD = 0.9). Participants were excluded if they reported (in an interview conducted by a research assistant, SLP student; Icht et al., 2019) one or more of the following diagnoses: (1) neurological disorders that may affect the speech production mechanisms; (2) structural or functional abnormalities of the oral mechanism; and (3) respiratory diseases (e.g., bronchial asthma, respiratory infection). All participants received oral and written information regarding the study and gave their written consent.

Inclusion criteria

To select participants for the clinical groups (adults with acquired dysarthria and young adults with developmental dysarthria), the SLPs of the rehabilitation centres and the day centre identified all participants as verbal, yet with speech impairments (e.g., reduced speech intelligibility, imprecise consonants or vowels, reduced voice quality). As tools for dysarthria assessment in Hebrew are lacking, dysarthria type and severity were assessed perceptually by the SLPs, using informal evaluations. An oral peripheral examination was also conducted identifying abnormalities in structure or function of oral and facial structures (for a similar selecting procedure, see Icht, 2021).

All participants in the three groups were adult native Hebrew speakers (or high-proficiency non-native speakers), with intact (or borderline for the clinical groups) cognitive abilities, and no history of severe hearing or visual problems (as indicated by self-report and by an interview with a research assistant, an SLP student). Exclusion criteria (as assessed by the medical staff) were the presence of severe cognitive impairment, visual and hearing impairment that could not be compensated for, associated aphasia or apraxia of speech, complications such as acute disease in major organs of the heart or brain, and severe psychological disorders.

TABLE 6 Participants' demographic and clinical data

Panel A: Adults with acquired dysarthria					
No.	**Age (years)**	**Gender**	**Aetiology and months/years post-onset; type of dysarthria; and severity**	**Cognitive status**	**Perceptual speech characteristics**
1	64	Male	CVA, 3 months; spastic; mild–moderate	MoCA[a] 24	Strained voice quality, slow speech rate, imprecise consonants
2	65	Female	CVA, 3 months; spastic; moderate	Intact per SLP report	Slow speech rate, imprecise consonants
3	70	Male	CVA, 12 months; spastic; moderate	MoCA 22	Strained voice quality, slow speech rate, imprecise consonants and vowels
4	92	Female	CVA, 3 months; spastic; moderate–severe	Borderline per SLP report	Strained voice quality, slow speech rate, low pitch, imprecise consonants and vowels
5	29	Female	CVA, 4 years; flaccid; mild	MoCA[a] 26	Reduced voice quality (breathy voice), imprecise consonants
6	84	Male	CVA, 3 months; flaccid; mild–moderate	MMSE[b] 26	Reduced voice quality (breathy voice), imprecise consonants
7	86	Female	3rd CVA, 7 months; flaccid; moderate–severe	MoCA[a] 24	Reduced voice quality (breathy voice), nasal speech, imprecise consonants and vowels
8	33	Male	CVA, 3 months; ataxic; moderate–severe	MoCA[a] 26	Irregular intonation (pitch and loudness), deviant rate (irregular)
9	62	Male	CVA, 16 years; mixed (spastic–ataxic); moderate	MoCA[a] 23	Slow or irregular speech rate, imprecise consonants and vowels
10	74	Male	CVA, 10 months; mixed (spastic–ataxic); severe	MMSE[b] 26	Irregular speech rate and pitch, imprecise consonants and vowels

(Continues)

International Journal of Language & Communication Disorders

TABLE 6 (Continued)

Panel A: Adults with acquired dysarthria					
No.	**Age (years)**	**Gender**	**Aetiology and months/years post-onset; type of dysarthria; and severity**	**Cognitive status**	**Perceptual speech characteristics**
11	75	Male	CVA, 8 months; mixed (spastic–ataxic); moderate–severe	MoCA[a] 22	Strained voice quality, low pitch, imprecise consonants and vowels
12	80	Male	CVA, 18 months; mixed (flaccid–spastic); mild	MoCA[a] 22	Reduced voice quality (breathy voice), imprecise consonants
13	46	Male	Brain tumour, 12 months; ataxic; moderate–severe	LOTCA[c] 90	Irregular intonation (pitch and loudness), slow speech rate
14	49	Female	Parkinson's disease, 12 years; hypo-kinetic; moderate	MoCA[a] 29	Reduced loudness, low pitch, irregular speech rate with occasional pauses
15	69	Female	Parkinson's disease, 6 years; hypo-kinetic; moderate	MMSE[b] 26	Reduced loudness, mono-pitch, irregular and slow speech rate

Panel B: Young adults with developmental dysarthria				
No.	**Age (years)**	**Gender**	**Aetiology; Type of dysarthria; severity**	**Perceptual speech characteristics**
1	39	Male	CP; spastic; moderate	Imprecise consonants and vowels, deviant rate (irregular)
2	28	Female	CP; spastic; moderate–severe	Short phrases, imprecise consonants and vowels
3	29	Female	CP; spastic; severe	Deviant voice quality (roughness), imprecise consonants and vowels, prolonged intervals, involuntary movements
4	23	Female	CP; spastic; moderate	Imprecise consonants and vowels, deviant rate (irregular)
5	23	Male	CP; spastic; moderate	Imprecise consonants and vowels, involuntary movements, strained voice quality
6	25	Female	CP; spastic; moderate–severe	Involuntary movements, imprecise consonants and vowels, strained voice quality and slow speech rate

(Continues)

TABLE 6 (Continued)

Panel B: Young adults with developmental dysarthria				
No.	**Age (years)**	**Gender**	**Aetiology; Type of dysarthria; severity**	**Perceptual speech characteristics**
7	21	Female	CP; spastic; mild	Deviant voice quality (roughness), imprecise consonants and vowels, monotone intonation
8	30	Male	CP; spastic; moderate–severe	Short phrases, voice stoppages, prolonged intervals, imprecise consonants and vowels
9	22	Female	CP; flaccid; moderate	Reduced voice quality (breathy voice), imprecise consonants and vowels
10	24	Male	CP; ataxic; mild	Imprecise consonants and vowels (mild)
11	39	Male	CP; mixed (spastic–ataxic); moderate	Deviant voice quality (roughness), imprecise consonants and vowels (mild)
12	28	Female	CP; mixed (spastic–ataxic); mild–moderate	Imprecise consonants and vowels, deviant rate (irregular)
13	24	Male	CP; mixed (primarily ataxic); mild	Imprecise consonants and vowels, inappropriate silences
14	25	Male	CP; mixed (primarily spastic); moderate	Reduced loudness, imprecise consonants and vowels, strained voice quality
15	24	Male	Pelizaeus–Merzbacher disease; mixed (spastic–ataxic); moderate	Imprecise consonants, voice tremor
16	24	Male	Pelizaeus–Merzbacher disease; mixed (primarily spastic); moderate–severe	Slow speech rate, mono-pitch, imprecise consonants and vowels
17	23	Male	Duchenne muscular dystrophy; mixed (spastic–ataxic); mild	Mono-pitch, slow speech rate, imprecise consonants and vowels
18	23	Female	Hypoxic ischaemic encephalopathy; mixed (primarily spastic); moderate	Reduced loudness, imprecise consonants and vowels
19	27	Male	Amyotrophic lateral sclerosis; spastic; mild–moderate	Mono-pitch, imprecise consonants

Notes: [a]For MoCA suggested cut-offs for impairment, see Rossetti et al. (2011).
[b]For MMSE suggested cut-offs for impairment, see Tombaugh et al. (1996).
[c]For LOTCA suggested cut-offs for impairment, see Annes et al. (1996).
CVA, cerebrovascular accident; and CP, cerebral palsy.

Materials

Heb-FDA-2 intelligibility subsections: Word and Sentence tasks

The newly developed word and sentence intelligibility subsections of the Heb-FDA-2 were administered. The administration procedure of the original FDA-2 was closely followed.

Comparison tasks in Hebrew: Word task

Assessing word-level intelligibility, the Hebrew monosyllabic phonetically balanced speech discrimination test was used (Auditory Word Discrimination Test; Putter-Katz et al., 2002). The test comprises of eight lists of 50 different monosyllabic words. This test was designed for hearing evaluation in clinical practice. Since there is no dedicated test in Hebrew for word-level speech intelligibility, this test is commonly used in clinical practice for speech evaluation. The words were printed on separate Intelligibility Cards, and the participants were asked to read aloud 10 random words.

Comparison tasks in Hebrew: Sentence task

Sentences of the 'Thousand Islands' reading passage were used (a different set for each participant). This is a Hebrew phonetically balanced and standardized reading passage (Amir & Levine-Yundof, 2013), commonly used in clinical practice as well as in research in the field of stuttering (Kronfeld-Duenias et al., 2016) and voice (Amir et al., 2014). Like the word task, this tool was not designed to assess sentence-level speech intelligibility (but rather, reading fluency and voice characteristics of complete passages); in fact, some of the sentences are rather complex and long. But, due to the absence of other validated and standardized sentences, it is widely used in clinical settings. The sentences were printed on separate Intelligibility Cards, and the participants were asked to read aloud 10 random sentences. The intelligibility tasks and materials are summarized in Table 7.

SLP subjective speech intelligibility evaluation

For the group of young adults with developmental dysarthria, the day centre SLP rated overall speech intelligibility a few days (up to 1 week) prior to the experimental session. Rating was done on a five-point scale, with 1 indicating completely unintelligible speech and 5 indicating completely intelligible speech.

Procedure

Participants were tested individually in a quiet room in the rehabilitation centres (clinical groups) or in the community centre they attended (control group). Two research assistants (SLP students) were present in the room throughout the experimental session, seated next to the participant. Upon arrival, each participant read and signed an informed consent form (or consented orally) and answered a short personal data questionnaire (taken from Ben-David & Icht, 2017; age, gender, place of birth, native language, years of education, vision and hearing status, etc.) These data were used to verify participants' inclusion criteria; relevant data are detailed in Table 6. Some of the participants performed the cognitive screening test (the Hebrew version of the MMSE, the LOTCA or the MoCA test). These data were available in the medical files for some of the participants. Others were unable to perform the tests, and intact or borderline cognitive functioning was verified by the staff members of the rehabilitation centre (Table 6). Following this, the participants were given a short explanation regarding the study and its goals.

The four tasks were administrated in a random order. In each task, the stimuli (words, sentences) were printed on separate Intelligibility Cards. The experimenter shuffled the cards, placed them facing down and selected 12 cards at random, avoiding looking at the cards. The clinician exposed the face of each card to the patient, who was asked to read the content of the card aloud (the first two cards were used as practise). The experimenter wrote down what the word or sentence was understood to be, and accuracy was determined by comparing with the Intelligibility Cards. The second research assistant provided logistical support and operational coordination.

Each task was terminated once all cards were read aloud. The total experimental session (four tasks) lasted approximately 10–15 min. At the end of the experimental session, the experimenter coded accuracy by comparing the cards against the spoken words or sentences. Participants were not informed of their performance on the different tasks.

To determine test–retest reliability, 13 young adults with developmental dysarthria performed the experimental tasks again following 3 months. The administration procedure was highly similar. We were not able to perform this with all the young adults with developmental dysarthria or with the group of older adults with acquired dysarthria due to drop-out and recurring health problems

TABLE 7 Intelligibility tasks and materials

	Heb-FDA-2 intelligibility subtests	Comparison tasks[a]
Words	Heb-FDA-2 word list: 116 familiar words, mono- to four-syllabic, controlled for emotional valence, arousal, syllabic structure and phonetic distribution	Auditory Word Discrimination Test: eight lists of 50 different monosyllabic words, phonetically balanced
Sentences	Heb-FDA-2 sentence list: 50 short different sentences, non-predictable, approximating the linguistic structures of Hebrew and satisfy morphosyntactic rules. Keywords are phonetically balanced	The 'Thousand Islands' reading passage: phonetically balanced and standardized reading passage comprises of relatively long and complex sentences

Note: [a]The comparison tasks are commonly used in clinical practice, although they have not been validated either for intelligibility or for dysarthria assessment, and they do not meet the original FDA-2 criteria.

common in these groups. Mainly, the sample of adults with acquired dysarthria (which included many stroke patients) was characterized by unstable medical–clinical status due to recurrent events, associated neurological illness and mortality. As the functional condition of some patients was unstable, test–retest was less suitable for this group.

Statistical analyses

The number of words interpreted correctly by the examiners in the word tasks (the Heb-FDA-2 word task, and the Auditory Word Discrimination Test) were counted. Similarly, the number of sentences interpreted correctly (the Heb-FDA-2 sentence task, and the sentences of the 'Thousand Islands' reading passage) were counted. Based on these raw scores, group means (SDs) were calculated for the three participant groups in each of the four experimental tasks. For the Heb-FDA-2 subtests, the intelligibility scores were used to assign grades (from a to e), as defined in the original FDA-2 assessment criteria.

To evaluate the accuracy of the Heb-FDA-2 tests, their sensitivity (i.e., ability to correctly identify those with dysarthria) and specificity (i.e., ability to correctly reject people who do not have dysarthria) were assessed using the following equations:

$$\text{Sensitivity} = \frac{\text{true positives}}{\text{true positives} + false\ \text{negatives}}$$

$$Specificity = \frac{\text{true negatives}}{\text{true negatives} + false\ \text{positives}}$$

Construct validity was assessed using the correlation between scores on the Heb-FDA-2 word and sentence subtests for the clinical groups. Convergent validity was calculated using the correlations between SLPs' intelligibility scores and the scores in both sentence tasks (Heb-FDA-2 sentence subtest, the comparison sentence task and the 'Thousand Islands' reading passage) for the group of young adults with developmental dysarthria. Test–retest reliability for the various tasks was also assessed using the correlations between first and second administrations of the tests for the group of young adults with developmental dysarthria.

RESULTS

The number of words and sentences correctly interpreted by the examiners (group means and SDs) in the different tasks for the three participant groups are listed in Table 8 (scoring was in accord with the original FDA-2 criteria). For the Heb-FDA-2 subtests, grades (from a to e) are also listed. Table 9 provides a summary of the validation measurements' results.

Tests' sensitivity and specificity

Analysing the Heb-FDA-2 subtests' scores revealed excellent accuracy levels, as the sensitivity was 0.91 (the tests correctly identified 91% of the people who have dysarthria; grades c–e), and the specificity was 1.0 (i.e., the tests correctly rejected 100% of the participants who do not have dysarthria; grade a). In other words, the Heb-FDA-2 intelligibility subtests can easily distinguish between clinical and non-clinical populations.

The expected ceiling effect was observable for all the non-clinical participants (healthy older adults), as they performed both subtests flawlessly, correctly producing all 10 words and 10 sentences in a fully intelligible manner. Mean scores for the clinical groups were lower, ranging from 5 to 8 on the word subtest, and from 5 to 6 on the sentence subtest. Importantly, only a few of the clinical participants (one of the 15 adults with acquired dysarthria; two of the 19 young adults with developmental dysarthria) scored flawlessly on both tests.

It is also notable that although all clinical participants completed the comparison sentence task (read aloud all the 'Thousand Islands' sentences), many were not able to

TABLE 8 Number of words and sentences correctly interpreted (group means and SDs) in the different tasks for the three participant groups; and FDA-2 grades[a] in the different tasks

Subtest/group		Healthy older adults ($n = 20$)	Adults with acquired dysarthria ($n = 15$)	Young adults with developmental dysarthria: First administration ($n = 19$)	Young adults with developmental dysarthria: Second administration ($n = 13$)
Heb-FDA-2	Word subtest	10 (0); a	8 (1.89); c	6.42 (3.24); d	5.57 (3.65); d
	Sentence subtest	10 (0); a	6.66 (2.52); d	5.63 (3.22); d	6 (3.04); d
Comparison	Word task[b]	10 (0); a	7.46 (2.21); c	6.68 (2.41); d	6.28 (3.24); d
	Sentence task[c]	10 (0); a	3.13 (3.40); e	4.10 (4.38); e	2.43 (3.67); e

Notes: aThe FDA-2 assessment criteria for the words and sentences tests are as follows: Grade a: 10 words are correctly interpreted by the therapist, with speech easily intelligible; Grade b: 10 words are correctly interpreted by the therapist, but they had to use particular care in listening and interpreting what was heard; Grade c: 7–9 words are interpreted correctly; Grade d: 5–6 words are interpreted correctly; and Grade e: fewer than five (i.e., 0–4) words are interpreted correctly.
[b]Auditory Word Discrimination Test.
[c]The 'Thousand Islands' reading passage.

TABLE 9 Results of the initial validation process of the Heb-FDA-2 intelligibility subtests

Measurement	Description	Test statistics
Specificity	True negative rate	1.0
Sensitivity	True positive rate	0.91
Construct validity	Correlations between word and sentence tasks for both clinical groups	Heb-FDA-2 subtests first administration: $r = 0.779$, $p < 0.01$
		Heb-FDA-2 subtests second administration (young adults with developmental dysarthria): $r = 0.943$, $p < 0.001$
		Comparison tasks first administration: $r = 0.621$, $p < 0.01$
		Comparison tasks second administration (young adults with developmental dysarthria): $r = 0.527$, $p = 0.065$
Convergent validity	Correlations between SLPs' intelligibility scores and the scores in the sentence tasks, for the group of young adults with developmental dysarthria	Heb-FDA-2 sentence subtest: $r = 0.908$, $p < 0.001$
		The 'Thousand Islands' reading passage: $r = 0.665$, $p < 0.01$
Test–retest reliability	Correlations between first and second administrations of the tests for the group of young adults with developmental dysarthria	Heb-FDA-2 word subtest: $r = 0.844$, $p < 0.001$
		Auditory Word Discrimination Test: $r = 0.719$, $p = 0.006$
		Heb-FDA-2 sentence subtest: $r = 0.906$, $p < 0.001$
		The 'Thousand Islands' reading passage: $r = 0.608$, $p = 0.027$

correctly read even one of the sentences (six of the of the 15 adults with acquired dysarthria; eight of the 19 young adults with developmental dysarthria). As the number of sentences correctly interpreted by the experimenters was small, their grades were low (e). These low grades illustrate the incompatibility of this tool for the purpose of diagnosing speech intelligibility in dysarthria.

Tests' validity

Construct validity evaluates whether a tool really measures what it was designed to measure; in the current study, speech intelligibility in dysarthria. To achieve construct validity, it is important to ensure that the new tests are carefully developed, based on relevant existing knowledge.

To address this, the meticulous adaptation procedure for the Heb-FDA-2 subtests was used (as described in Study 1), carefully following the original FDA-2 criteria and knowledge on Hebrew linguistics. This procedure can be taken as forms of evidence for construct validity: (1) face validity: how suitable the content of the new subtests seems to be on the surface, and (2) content validity: assessing whether the new subtests are representative of all aspects of the construct.

Traditional criterion validity could not be assessed, as there are no available validated tools for evaluating speech intelligibility in dysarthria in Hebrew (word or sentence lists) to which the new subtests could be compared. However, as both Heb-FDA-2 subtests are designed to reflect the same underlying construct (speech intelligibility), we measured correlation between scores on the word and sentence subtests as a gauge for construct validity. High correlations can be taken to reflect high construct validity (internal consistency). Analyses were conducted for the clinical populations, as control participants scored at ceiling. The correlation between the two Heb-FDA-2 subtests was high, $r = 0.779$, $p < 0.01$, and the correlation for the comparison subtests was moderate, $r = 0.621$, $p < 0.01$, following classifications by Schober et al. (2018). We note that the two correlations were not statistically different, Olkin's $Z = 1.1$, $p > 0.05$ (Olkin, 1967).

To assess convergent validity, we tested how the Heb-FDA-2 sentence subtest scores correlate with an existing measure of the same construct, that is, SLPs subjective speech intelligibility scores. As aforementioned, for young adults with developmental dysarthria, daily speech intelligibility (i.e., ease of understanding) was subjectively assessed by the day centre SLP on a five-point scale. Analysis indicated significantly higher correlation between the SLP intelligibility scores and the Heb-FDA-2 sentence subtest, $r = 0.908$, $p < 0.001$, than between the SLP intelligibility scores and the comparison sentence task, $r = 0.665$, $p < 0.01$, $Z = 2.628$, $p = 0.004$ (following Ramseyer, 1979; https://www.psychometrica.de/correlation.html).

Tests' reliability

To measure test–retest reliability, 13 young adults with developmental dysarthria performed the tests again in the following 3 months. Reliability for the Heb-FDA-2 sentence subtest was very high, $r = 0.906$, $p < 0.001$, and significantly higher than reliability for the comparison sentence task, $r = 0.608$, $p = 0.027$, Olkin's $Z = 1.813$, $p = 0.035$ (Olkin, 1967). For the word subtests, test–retest correlations were moderate and significant for the comparison word task, $r = 0.719$, $p = 0.006$, and high and significant for the Heb-FDA-2 word subtest, $r = 0.844$, $p < 0.001$, the two correlations were not significantly different, Olkin's $Z = 1.08$, $p > 0.05$ (Olkin, 1967).

It is interesting to note that in the second administration, construct (or internal) validity (between the word and the sentence subtests) for the Heb-FDA-2 was extremely high, $r = 0.943$, $p < 0.001$, and was significantly higher than construct validity for the comparison tasks, $r = 0.527$, $p = 0.065$, $Z = 1.933$, $p = 0.027$.

DISCUSSION

In Study 2, the new Heb-FDA-2 word and sentence subtests were found to be accurate (specific and sensitive), valid and reliable:

- Sensitivity and specificity. The subtests accurately identify clinical (individuals with dysarthria) and reject non-clinical (healthy older adult) participants. This is of special importance as the current study tested healthy older adults as a control group, rather than young-adult college students.
- Construct validity. The two new subtests were highly correlated, gauging a similar ability—speech production.
- Convergent validity. The new sentence subtest was highly correlated with subjective intelligibility scores assigned to participants by their SLP.
- Reliability. The test–retest reliability of the new subtests was high, as repeated administration of the new tool within the group of young adults with developmental dysarthria yielded highly similar results.

These findings suggest that the new subtests provide a better clinical tool than the existing tools in Hebrew in terms of specificity, sensitivity, reliability, construct and convergent validity. This is of no surprise as the currently available tools in Hebrew were not designed to test speech intelligibility, whereas the two suggested Heb-FDA-2 subtests were designed to be administered together, providing a complete speech intelligibility gauge in Hebrew. The new subtests can be also used for research purposes, allowing for cross-language and cross-culture comparisons.

GENERAL DISCUSSION

The limited linguistic research on Hebrew, especially in clinical populations, and the lack of standard validated evaluation tools, poses challenges to clinical assessment as well as research in the field of speech intelligibility for individuals with dysarthria. To fill this gap, the present study adapted the FDA-2 intelligibility subtests (words and sentences) to Hebrew and validated it with two groups of individuals with dysarthria (acquired and developmental) and one control group (healthy older adults). We focused on the FDA-2 as this is a common and well-known tool, which has been previously translated and adapted to several other languages. The latter advantage offers cross-linguistic and cross-cultural comparisons, improving evaluation accuracy, as well as promoting research.

In a systematic adaptation process (Study 1), Hebrew versions of the words and sentence FDA-2 intelligibility subtests were constructed, Heb-FDA-2. The adaptation procedure followed a rigorous methodological cultural adaptation. It involved expert consulting, meeting the criteria of the English FDA-2 intelligibility task framework. The Hebrew word list (of 116 words) is phonetically balanced and controls for several other factors that may affect readability. These include emotional valence, arousal and familiarity of the words, as well as sentence predictability. The corpus is large enough to reduce probability of patients and clinicians learning the words with repeated exposures.

The Hebrew sentence list also corresponds to the original tool in English, comprises of 50 short and non-predictable phrases. First, the carrier phrases are all different, so the clinician has to listen to the whole sentence, rather than interpreting the keyword in a standard carrier phrase. Second, predictable sentences, those which had keywords that could be easily guessed, were excluded. This was important since highly predictable sentences may improve intelligibility scores relative to sentences with low predictability (Geetha et al., 2014). Indeed, for speakers with dysarthria, high semantic predictiveness improved sentence intelligibility (Garcia & Dagenais, 1998).

Interestingly, the literature suggests that for many individuals with dysarthria, sentence context tends to yield higher intelligibility scores than words in isolation (Hustad, 2007). Possibly, listening to sentences involves building of contextual knowledge which may facilitate listeners' ability to use top-down intrinsic linguistic information, better inferring words (that they may not otherwise have been able to identify in isolation). To reduce such possible effects, the final set of sentences were all low on semantic predictiveness, short and keywords (the final word in each phrase) were all monosyllabic (see also Kent et al., 1989). Third, keywords were phonetically balanced, in line with the original FDA-2 guidelines. Finally, the sentences follow the linguistic structures and morphosyntactic rules of spoken Hebrew.

In Study 2, the new Heb-FDA-2 subtests were found to be sensitive (correctly identifying 91% of the people who have dysarthria) and specific (correctly identifying 100% who do not have dysarthria), thus providing a better clinical tool than the existing tools in Hebrew. The new subtests were also found to be valid (high levels of construct and convergent validity) and reliable (test–retest reliability). Results are specifically encouraging given the high variability of the tested clinical participants in terms of aetiology, severity and duration (Table 6). These positive findings form the initial steps of a larger study that aims to adapt and normalize the entire FDA-2 for Hebrew. As a pilot test, a relatively small try-out of the adapted tests was conducted, using two medium-sized samples of individuals representative of the eventual target population (Hambleton & Patsula, 1999). Obviously, using the new Heb-FDA-2 subtests to evaluate larger samples of individuals with dysarthria is called for.

Limitations and future directions

It must be acknowledged that the word list does not represent the complete Hebrew linguistic diversity since it aims to provide an overall indication of dysarthria severity rather than a complete phonetic analysis of the Hebrew sounds and linguistic contexts. In addition, factors such as orthographic regularity and morpho-orthographic transparency were not controlled (note that Hebrew is characterized by a non-transparent mapping of phonology to the orthography; Schiff et al., 2020). However, as all words were highly familiar, no reading difficulties occurred. Similarly, a phonological dissimilarity profile for the word set was not generated (Mueller et al., 2003), given that such data are not readily available in Hebrew. Some of the monosyllabic words within the Heb-FDA-2 word list could possibly be grouped with minimal pairs (or near minimal pairs, e.g., giʁ-siʁ, kal-kav). Other words have phonological proximity with other common Hebrew words (not included in the word list). Therefore, minimal pair confusability might occur. However, this feature may form an advantage because it requires the patients to be as precise as possible in articulating the words (especially as the scoring procedure is based on number of words correctly interpreted by the examiner, rather than a multiple-choice format). SLPs may analyse the transcriptions to build individual perceptual error profiles that can be used in therapy.

As aforementioned, this study constituted part of a larger experimental project that aimed to perform a Hebrew adaptation (and normalization) of the full FDA-2, using a larger sample of participants. As this paper presented the initial steps of adaptation and validation of the intelligibility subtests, and as suggested by Hambleton and Patsula (1999), we conducted a pilot test using two different samples of participants with dysarthria. Clearly, future studies should increase the strength and generalization of the current results with larger samples (avoiding random error, as the full study population is not included, and selection bias), variety of aetiologies (e.g., multiple sclerosis, Parkinson's disease) with different severity levels of dysarthria, and different age groups. Further studies may wish to control for background factors as educational levels and occupational status. Importantly, future research is required to further establish the validity and reliability of the new subtests, assessing other psychometric properties (e.g., test–retest reliability in other clinical groups, interrater reliability, internal consistency), to minimize the chance for measurement errors. It may also be interesting to use the intelligibility subtests in longitudinal studies to understand the potential of speech production in dysarthria as a clinical marker of disease progression. Future studies may also wish to develop a computerized version of these subtests and telehealth solutions (e.g., Ben-David et al., 2021; Mama & Icht, 2020).

Clinical recommendations

We suggest using the new Heb-FDA-2 lists in clinical practice to reduce reliance on informal methods of intelligibility assessment and increase test validity and reliability. The Heb-FDA-2 words and sentence subtests may enable clinicians to obtain realistic and representative measures of intelligibility, providing essential information for developing appropriate intervention programmes. These subtests are important for research as well as clinical practice because valid and reliable Hebrew-language instruments are scarce, especially for people with dysarthria.

Finally, to ensure highest validity of the novel tool (given Hebrew linguistic characteristics), we suggest adopting a semi-random selection of the target words for the word list subtest, based on the number of syllables (presented as the rightmost column of Table 2). Specifically, the clinician randomly selects from five different card (word) types: two monosyllabic words with a simple onset, a single monosyllabic word with a complex onset, three bisyllabic, two trisyllabic and two four-syllabic words. Such procedure prevents possible unbalanced word selection (e.g., mostly polysyllabic words), which may hamper the evaluation process (different difficulty levels).

CONFLICT OF INTEREST

The authors declare that they have no competing interests.

DATA AVAILABILITY STATEMENT

Original data will be provided by the authors upon request.

FUNDING

This work was supported in part by Ariel University (grant number RA1700000298).

REFERENCES

Allison, K.M. & Hustad, K.C. (2018) Acoustic predictors of pediatric dysarthria in cerebral palsy. *Journal of Speech, Language, and Hearing Research*, 61(3), 462–478.

Amir, O., Lebi-Jacob, N. & Harari, O. (2014) The effect of in vitro fertilization treatment on women's voice. *Journal of Voice*, 28(4), 518–522

Amir, O. & Levine-Yundof, R. (2013) Listeners' attitude toward people with dysphonia. *Journal of Voice*, 27(4), 524–e1.

Annes, G., Katz, N. & Cermak, S. (1996) Comparison of younger and older healthy American adults on the Loewenstein Occupational Therapy Cognitive Assessment. *Occupational Therapy International*, 3, 157–173.

Ansel, B.M. & Kent, R.D. (1992) Acoustic–phonetic contrasts and intelligibility in the dysarthria associated with mixed cerebral palsy. *Journal of Speech, Language, and Hearing Research*, 35(2), 296–308.

Balzan, P., Vella, A. & Tattersall, C. (2019) Assessment of intelligibility in dysarthria: development of a Maltese word and phrase list. *Clinical Linguistics & Phonetics*, 33(10–11), 965–977.

Bartl, H., Keller, H., Zohar, N. & Wahle, N. (2020) Many children, many risks? Listening to the voices of families with many children from the Ultra-Orthodox (Haredi) community in Israel. *Context-informed perspectives of child risk and protection in Israel*. Cham: Springer, pp. 105–129.

Ben-David, A. & Bat-El, O. (2016) Paths and stages in acquisition of the phonological word in Hebrew. *Acquisition and Development of Hebrew: From Infancy to Adolescence*, 19, 39–68.

Ben-David, B.M., Chajut, E. & Algom, D. (2012) The pale shades of emotion: a signal detection theory analysis of the emotional Stroop task. *Psychology*, 3, 537–541.

Ben-David, B.M., Durham, N.A.-M. & Van Lieshout, P.H.H.M. (2016) The Linguistic Acoustic Threat Effect (LATTE): screening tool for the impact of semantic threat in speech processing after a brain injury. *Brain Injury*, 30(2), 237–239.

Ben-David, B.M. & Icht, M. (2018) The effect of practice and visual feedback on oral-diadochokinetic rates for younger and older adults. *Language and Speech*, 61(1), 113–134.

Ben-David, B.M. & Icht, M. (2017) Oral-diadochokinetic rates for Hebrew-speaking healthy aging population: non-word vs. real-word repetition. *International Journal of Language & Communication Disorders*, 52(3), 301–310

Ben-David, B.M., Mentzel1, M., Icht, M., Gilad, M., Dor, Y.I., Ben-David, S., et al. (2021) Challenges and opportunities for telehealth assessment during COVID-19: IT-RES, adapting a remote version of the Test for Rating Emotions in Speech. *International Journal of Audiology*, 60(5), 319–321. https://doi.org/10.1080/14992027.2020.1833255

Ben-David, B.M., Moral, M.I., Namasivayam, A.K., Erel, H. & van Lieshout, P.H. (2016) Linguistic and emotional-valence characteristics of reading passages for clinical use and research. *Journal of Fluency Disorders*, 49, 1–12.

Ben-David, B.M., Van Lieshout, P.H.H.M. & Leszcz, T. (2011) A resource of validated affective and neutral sentences to assess

identification of emotion in spoken language after a brain injury. *Brain Injury*, 25, 206–220.
Bradley, M.M. & Lang, P.J. (1994) Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1), 49–59.
Cardoso, R., Guimarães, I., Santos, H., Loureiro, R., Domingos, J., De Abreu, D., et al. (2017) Frenchay dysarthria assessment (FDA-2) in Parkinson's disease: cross-cultural adaptation and psychometric properties of the European Portuguese version. *Journal of Neurology*, 264(1), 21–31.
Carl, M. & Icht, M. (2021) Acoustic vowel analysis and speech intelligibility in young adult Hebrew speakers: developmental dysarthria vs. typical development. *International Journal of Language and Communication Disorders*, 56(2), 283–298.
Carl, M., Levy, E.S. & Icht, M. (2022) Speech treatment for Hebrew speaking adolescents and young adults with developmental dysarthria: a comparison of *m*SIT and Beatalk. *International Journal of Language and Communication Disorders*, 57(3), 660–679.
Darley, F.L., Aronson, A.E. & Brown, J.R. (1969) Differential diagnostic patterns of dysarthria. *Journal of Speech and Hearing Research*, 12(2), 246–269.
Darley, F.L., Aronson, A.E. & Brown, J.R. (1975) *Motor speech disorders*, Vol. 304, Philadelphia: Saunders.
Duffy, J.R. (2019) *Motor speech disorders-E-book: substrates, differential diagnosis, and management*. Missouri: St. Louis, Elsevier Health Sciences.
Eigentler, A., Rhomberg, J., Nachbauer, W., Ritzer, I., Poewe, W. & Boesch, S. (2012) The scale for the assessment and rating of ataxia correlates with dysarthria assessment in Friedreich's ataxia. *Journal of Neurology*, 259(3), 420–426.
Enderby, P. (1980) Frenchay dysarthria assessment. *British Journal of Disorders of Communication*, 15(3), 165–173.
Enderby, P. (2013) Disorders of communication: dysarthria. In: Barnes, M.P. & Good, D.C., *Handbook of clinical neurology: Neurologica. Rehabilitation*, 110(3), 273–282.
Enderby, P. & Palmer, R. (1983) *Frenchay dysarthria assessment, PRO-ED*. Austin: CIT0007.
Enderby, P. & Palmer, R. (2008) *Frenchay dysarthria assessment–Second edition (FDA-2)*. Austin, TX: Pro-Ed.
Folstein, M.F., Folstein, S.E. & McHugh, P.R. (1975) Minimental state: a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189–198.
France, K.R., Shah, R.H. & Park, C.W. (1994) The impact of emotional valence and intensity on ad evaluation and memory. *Advances in Consumer Research (ACR) North American Advances*, 21, 583–588.
Friedmann, N. & Lavi, H. (2006) On the order of acquisition of A-movement, Wh-movement and VC movement. In A. Belletti, E. Bennati, C. Chesi, E. Di Domenico & I. Ferrari (Eds), *Language Acquisition and Development*, pp. 211–217. Cambridge: Cambridge Scholars Press/CSP.
Frost, R. & Plaut, D.C. (2005) The word-frequency database for printed Hebrew. Retrieved March 10, 2017, from http://word-freq.mscc.huji.ac.il/index.htm
Garcia, J.M. & Dagenais, P.A. (1998) Dysarthric sentence intelligibility: contribution of iconic gestures and message predictiveness. *Journal of Speech, Language, and Hearing Research*, 41(6), 1282–1293.
Geetha, C., Kumar, K.S.S., Manjula, P. & Pavan, M. (2014) Development and standardisation of the sentence identification test in the Kannada language. *Journal of Hearing Science*, 4(01), 18–26.
Ghio, A., Giusti, L., Blanc, E. & Pinto, S. (2010) French adaptation of the "Frenchay Dysarthria Assessment 2" speech intelligibility test. *European Annals of Otorhinolaryngology, Head and Neck Diseases*, 137(2), 111–116.
Hadar, B., Skrzypek, J.E., Wingfield, A. & Ben-David, B.M. (2016) Working memory load affects processing time in spoken word recognition: evidence from eye-movements. *Frontiers in Neuroscience*, 10, 220.
Hambleton, R.K. & Patsula, L. (1998) Adapting tests for use in multiple languages and cultures. *Social Indicators Research*, 45(1–3), 153–171.
Hambleton, R.K. & Patsula, L. (1999) Increasing the validity of Adapted Tests: myths to be avoided and guidelines for improving test adaptation practice. *Journal of Applied Testing Technology*, 1(1), 1–16.
Hijikata, N., Kawakami, M., Wada, A., Ikezawa, M., Kaji, K., Chiba, Y., et al. (2022) Assessment of dysarthria with Frenchay dysarthria assessment (FDA-2) in patients with Duchenne muscular dystrophy. *Disability and Rehabilitation*, 44(8), 1443–1450.
Hustad, K.C. (2007) Effects of speech stimuli and dysarthria severity on intelligibility scores and listener confidence ratings for speakers with cerebral palsy. *Folia Phoniatrica et Logopaedica*, 59(6), 306–317.
Icht, M. (2021) Improving speech characteristics of young adults with congenital dysarthria: an exploratory study comparing articulation training and the Beatalk method. *Journal of Communication Disorders*, 93, 106147.
Icht, M., Bergerzon-Biton, O. & Mama, Y. (2019) The production effect in adults with dysarthria: improving long-term verbal memory by vocal production. *Neuropsychological Rehabilitation*, 29(1), 131–143.
Israeli Center for Disease Control (ICDC) (2021) Ministry of Health. Israel. https://www.health.gov.il/UnitsOffice/ICDC/Chronic_Diseases/stroke/Pages/default.aspx
Katz, N., Itzkovich, M., Averbuch, S. & Elazar, B. (1989) Loewenstein Occupational Therapy Cognitive Assessment (LOTCA) battery for brain-injured patients: reliability and validity. *American Journal of Occupational Therapy*, 43(3), 184–192.
Kent, R.D., Weismer, G., Kent, J.F. & Rosenbek, J.C. (1989) Toward phonetic intelligibility testing in dysarthria. *Journal of Speech and Hearing Disorders*, 54(4), 482–499.
Klasner, E.R. & Yorkston, K.M. (2005) Speech intelligibility in ALS and HD dysarthria: the everyday listener's perspective. *Journal of Medical Speech–Language Pathology*, 13(2), 127–140.
Kronfeld-Duenias, V., Amir, O., Ezrati-Vinacour, R., Civier, O. & Ben-Shachar, M. (2016) The frontal aslant tract underlies speech fluency in persistent developmental stuttering. *Brain Structure and Function*, 221(1), 365–381.
Lammert, A.C., Melot, J., Sturim, D.E., Hannon, D.J., DeLaura, R., Williamson, J.R., et al. (2020) Analysis of phonetic balance in standard English passages. *Journal of Speech, Language, and Hearing Research*, 63(4), 917–930
Mama, Y. & Icht, M. (2016) Auditioning the distinctiveness account: expanding the production effect to the auditory modality reveals the superiority of writing over vocalising. *Memory (Hove, England)*, 24(1), 98–113.

Mama, Y. & Icht, M. (2021) Overcoming COVID-19 challenges: a remote adaptation of the production effect task. *Journal of the International Neuropsychological Society*, 27(8), 855–856.

Mei, C., Reilly, S., Reddihough, D., Mensah, F. & Morgan, A. (2014) Motor speech impairment, activity, and participation in children with cerebral palsy. *International Journal of Speech–Language Pathology*, 16(4), 427–435.

Mitchell, C., Gittins, M., Tyson, S., Vail, A., Conroy, P., Paley, L. & Bowen, A. (2021) Prevalence of aphasia and dysarthria among inpatient stroke survivors: describing the population, therapy provision and outcomes on discharge. *Aphasiology*, 35(7), 950–960.

Mueller, S.T., Seymour, T.L., Kieras, D.E. & Meyer, D.E. (2003) Theoretical implications of articulatory duration, phonological similarity, and phonological complexity in verbal working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(6), 1353.

Müller, J., Wenning, G.K., Verny, M., McKee, A., Chaudhuri, K.R., Jellinger, K., Poewe, W. & Litvan, I. (2001) Progression of dysarthria and dysphagia in postmortem-confirmed parkinsonian disorders. *Archives of Neurology*, 58(2), 259–264

Nasreddine, Z.S., Phillips, N.A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J.L. & Chertkow, H. (2005) The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695–699.

Parks, C.M. & Toth, J.P. (2006) Fluency, familiarity, aging, and the illusion of truth. *Aging, Neuropsychology, and Cognition*, 13(2), 225–253.

Putter-Katz, H., Peled, M., Schaik, M., Sachartov, E., Feldman, I., Said, L.A.B., Miran, D. & Kushnir, D. (2002) A comparison between vocal reaction time and word recognition measures of children with APD and age-matched peers using auditory word discrimination test. *Journal of Basic and Clinical Physiology and Pharmacology*, 13(2), 97–104.

Qutishat, D. (2015) Development and psychometric evaluation of an assessment of dysarthria for Arabic speakers. PhD thesis, University of Sheffield.

Ramseyer, G.C. (1979) Testing the difference between dependent correlations using the Fisher Z. *The Journal of Experimental Education*, 47(4), 307–310.

Rossetti, H.C., Lacritz, L.H., Cullum, C.M. & Weiner, M.F. (2011) Normative data for the Montreal Cognitive Assessment (MoCA) in a population-based sample. *Neurology*, 77(13), 1272–1275.

Olkin, I. (1967) Correlations revisited. In: Stanley, J.C. (Ed.) *Improving experimental design and statistical analysis*. Chicago, IL: Rand McNally, pp. 102–128.

Schiff, R., Rosenstock, S. & Ravid, D. (2020) Morpho-orthographic complexity in affix spelling in Hebrew: a novel psycholinguistic outlook across the school years. *Frontiers in Psychology*, 11, 868.

Schober, P., Boer, C. & Schwarte, L.A. (2018) Correlation coefficients: appropriate use and interpretation. *Anesthesia & Analgesia*, 126(5), 1763–1768.

The Israel Parkinson Association (2021) https://www.parkinson.org.il/

Tombaugh, T.N., McDowell, I., Kristjansson, B. & Hubley, A.M. (1996) Mini-Mental State Examination (MMSE) and the Modified MMSE (3MS): a psychometric comparison and normative data. *Psychological Assessment*, 8(1), 48.

Weismer, G. & Laures, J.S. (2002) Direct magnitude estimates of speech intelligibility in dysarthria: effects of a chosen standard. *Journal of Speech, Language, and Hearing Research*, 45, 421–433.

Weismer, G., Jeng, J.Y., Laures, J.S., Kent, R.D. & Kent, J.F. (2001) Acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders. *Folia Phoniatrica et Logopaedica*, 53(1), 1–18.

Werner, P., Heinik, J., Mendel, A., Reicher, B. & Bleich, A. (1999) Examining the reliability and validity of the Hebrew version of the Mini Mental State Examination. *Aging Clinical and Experimental Research*, 11(5), 329–334.

Yorkston, K.M. & Beukelman, D.R. (1981a) Communication efficiency of dysarthric speakers as measured by sentence intelligibility and speaking rate. *Journal of Speech and Hearing Disorders*, 46(3), 296–301.

Yorkston, K.M. & Beukelman, D.R. (1981b) *Assessment of intelligibility of dysarthric speech*. Tagard, OR: C.C. Publications.

Yorkston, K.M., Beukelman, D.R., Strand, E.A. & Bell, K.R. (1999) *Management of motor speech disorders in children and adults*, 2nd edition, Austin, TX: Pro Ed.

How to cite this article: Icht M., Bergerzon-Bitton O. & Ben-David B. M. (2022) Validation and cross-linguistic adaptation of the Frenchay Dysarthria Assessment (FDA-2) speech intelligibility tests: Hebrew version. *International Journal of Language & Communication Disorders* , 1023–1049. https://doi.org/10.1111/1460-6984.12737

APPENDIX A: Characteristics of Hebrew FDA-2 word list

Target consonant	Articulation place and manner	Mono-syllabic	Bi syllabic				Tri-syllabic			Four-syllabic		
		Initial	Final	Initial	Middle	Final	Initial	Middle	Final	Initial	Middle	Final
p	Bilabial; Plosive	pil paʁ pʁi[1]	/	pilpel pita	mapa pilpel tsipoʁ kapit tapuz	/	/	miʁpeset	/	/	tsipoʁnaim mispaʁaim	/
b		bat ben bad bli[1]	/	baʁvaz bakbuk bietsa	bakbuk dubi	/	banana	mitlabeʃ	/	/	mexubeʁet	/
f	Labiodental;/ Fricative		ʔof kaf guf xof	/	safsal	madaf	/	telefon	/	/	mitstaʁefet miʃkafaim ʔofanaim melafefon	/
v		vav kvaʁ[1] zvuv[1]	kav vav zvuv	/	baʁvaz kova	maxʃev	/	xoveʁet ʁakevet mevaʃel mivʁeʃet magevet mizvada	/	/	ʔagvanija mevugeʁet	/
t	Alveolar; Plosive	tei tal ta tlai[1] tʁis[1]	bat dat mot xut	tapuz	xatul pita	kapit	telefon	mistaʁek mitlabeʃ mistakel matate	xoveʁet ʁakevet mexonit miʁpeset mivʁeʃet magevet	/	mitstaʁefet	mitstaʁefet mevugeʁet mexubeʁet
d		din dat dan dli[1]	bad sid	dubi	madaf ʁadjo	tsamid	/	madʁega mizvada nadneda	/	/	kaduʁegel	/

(Continues)

Target consonant	Articulation place and manner	Mono-syllabic	Bi syllabic				Tri-syllabic			Four-syllabic		
		Initial	Final	Initial	Middle	Final	Initial	Middle	Final	Initial	Middle	Final
s	Alveolar; Fricative	siʁ sug sid sʁox[1]	mas tʁis	sulam safsal	safsal	/	/	mistaʁek mistakel miʁpeset	/	/	mispaʁaim	/
z		zol zan zvuv[1]	/	/	/	baʁvaz tapuz	/	mizvada	/			
k	Velar; Plosive	kav kal ken kam kaf kvaʁ[1]	ʁak ʃuk tsxok	Kapit kova	bakbuk	ʔaʁnak maʁak bakbuk	/	ʁakevet mistakel	mistaʁek	kaduʁegel	miʃkafaim	/
g		giʁ guʁ guf	sug xag	/	ʁegel	/	/	megeʁa madʁega magevet xagoʁa	/	/	ʔagvanija kaduʁegel mevugeʁet	/
x	Velar; Fricative	xut xag xof tsxok[1]	ʁa Xtsa Xla Xsʁox	xatul	maxʃev ʃulxan	luʔax	xoveʁet xagoʁa	mexonit	/	laxmanija	mexubeʁet met-saxtseʔax	metsaxtseʔax
ʁ	Velar; Approxi-mant	ʁa Xʁak pʁi[1] sʁox[1] tʁis[1]	giʁ siʁ neʁ maʁ guʁ haʁ paʁ kvaʁ	ʁadjo ʁegel	ʔaʁnak baʁvaz maʁak maʁʔa	tsipoʁ	ʁakevet ʁeʃima	xoveʁet mistaʁek megeʁa madʁega miʁpeset mivʁeʃet xagoʁa menoʁa	metsajeʁ	/	kaduʁegel tsipoʁnaim mitstaʁefet mevugeʁet mexubeʁet mispaʁaim	/
l	Alveolar; Lateral approxi-mant	lul la Xbli[1] dli[1] tlai[1]	lul tsel kal tal zol pil	luʔax	pilpel ʃulxan sulam	pilpel xatul safsal ʁegel	/	mitlabeʃ telefon	mistakel mevaʃel	laxmanija	melafefon	kaduʁegel

(Continues)

Target consonant	Articulation place and manner	Mono-syllabic	Bi syllabic				Tri-syllabic			Four-syllabic		
		Initial	Final	Initial	Middle	Final	Initial	Middle	Final	Initial	Middle	Final
j	Palatal; Approxi-mant	jom	/	/	ʁadjo	/	/	metsajeʁ	/	/	ʔagvanija laxmanija	/
ts	Alveolar; Affricate	tsa Xtsel tsxok[1]	ʔets mits	tsamid tsipoʁ	bietsa	/	/	metsajeʁ	/	tsipoʁnaim	mitstaʁefet met-saxtseʔax	/
ʃ	Postalveolar; Fricative	ʃum ʃuk ʃen ʃai	ʔeʃ ʔiʃ	ʃulxan	maxʃev	/	/	ʁeʃima mevaʃel mivʁeʃet	mitlabeʃ	/	miʃkafaim	/
m	Bilabial; Nasal	mas maʁ mot mits min	jom ʃum kam	mapa madaf maʁak maʁʔa maxʃev	tsamid	sulam	mistaʁek megeʁa mexonit madʁega mitlabeʃ mistakel mevaʃel metsajeʁ matate miʁpeset mivʁeʃet magevet mizvada menoʁa	ʁeʃima	/	mitstaʁefet mevugeʁet miʃkafaim mexubeʁet mispaʁaim melafefon metsaxtseʔax	laxmanija	tsipoʁnaim miʃkafaim mispaʁaim ʔofanaim
n	Alveolar; Nasal	neʁ	din ben ken dan ʃen min zan	/	ʔaʁnak	ʃulxan	nadneda	mexonit banana menoʁa nadneda	telefon	/	ʔagvanija tsipoʁnaim ʔofanaim laxmanija	melafefon
ʔ	Glottal; Plosive	ʔeʃ ʔets ʔiʃ ʔof	/	ʔaʁnak	maʁʔa luʔax	/	/	/	/	ʔagvanija ʔofanaim	/	/
h[2]	Glottal; Fricative	haʁ	/	/	/	/	/	/	/	/	/	/

1 Complex onset.
[2]Often pronounced as [ʔ].

APPENDIX B: Characteristics of target (final) words in Hebrew FDA-2 sentence list

Target consonant	Articulation place and manner	Mono-syllabic			
		Initial	Final		
p	Bilabial; Plosive	pʁi[1]	pax	/	
b		Bul	bots	/	
f	Labiodental; Fricative	/	ʔof	nof	
v		gvul[1] zvuv[1]	kvaʁ[1]	dov kav kʁuv	zvuv tov lev
t	Alveolar; Plosive	tlai[1] tov	tei tas	/	
d		dov	dag	ʔod	sid
s	Alveolar; Fricative	sid	nes	tas	
z		zeʁ zaz	zvuv[1] zaʁ	zaz	
k	Velar; Plosive	kav kaʁ kiʁ kʁuv[1] ken	kal kli[1] kam kvaʁ[1]	ʁek	
g		guʁ gam	gvul[1] giʁ	dag	
x	Velar; Fricative	xom	xol	na Xʁax	la Xpax
ʁ	Velar; Approximant	kʁuv[1] ʁa Xpʁi[1]	ʁum ʁek	guʁ kaʁ kiʁ ʔoʁ zeʁ giʁ	neʁ maʁ ʃiʁ zaʁ kvaʁ

(Continues)

l	Alveolar; Lateral approximant	La tlai[1] kli[1]	la Xlev	xol bul tsel	gvul kal
j	Palatal; Approximant	/	/		
ts	Alveolar; Affricate	tsom	tsel	bots	
ʃ	Postalveolar; Fricative	ʃum ʃeʃ	ʃiʁ	ʃeʃ	ʔeʃ
m	Bilabial; Nasal	maʁ	tsom xom gam	ʃum ʁum kam	
n	Alveolar; Nasal	nes nax	nof neʁ	ken	
ʔ	Glottal; Plosive	ʔof ʔod	ʔoʁ ʔeʃ	/	

[1]Complex onset.