Validating the Spanish adaptation of the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V)

Faustino Nunez-Batalla MD1,2; Marta Morato-Galan MD1; Jose Luis Llorente-Pendas, MD1,2

1University of Oviedo, 2Hospital Universitario Central de Asturias (Spain)

Abstract

Purpose: The Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) was developed to promote a standardized approach to evaluating and documenting auditory-perceptual judgments of vocal quality. This tool was originally developed in English language and its Spanish version is still inexistente. The aim of this study was to develop a Spanish adaptation of CAPE-V and to examine the reliability and empirical validity of this Spanish version.

Method: To adapt the CAPE-V protocol to the Spanish language, we proposed 6 phrases phonetically designed according to the CAPE-V requirements. Prospective instrument validation was performed. The validity of the Spanish version of the CAPE-V was examined in 4 ways: intrarater reliability, interrater reliability and CAPE-V versus GRBAS

Results: Interrater reliability coefficients for the CAPE-V ranged from 0.93 for overall severity to 0.54 for intensity; intrarater reliability ranged from 0.98 for overall severity to 0.85 for intensity. The comparison of judgments between CAPE-V and GRBAS ranged from 0.86 for overall severity to 0.61 for breathiness.

Introduction

The Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) is a clinical and research tool developed to promote a standardized approach to evaluating and documenting auditory-perceptual judgments of vocal quality. The tool was created as a direct outcome of the Consensus Conference on Auditory-Perceptual Evaluation of Voice sponsored by the American Speech-Language-Hearing Association (ASHA) and the University of Pittsburgh.

The CAPE-V uses continuous visual analog scales for judgments of six parameters of voice: overall severity, roughness, breathiness, strain, pitch, and loudness. When using the CAPE-V, the clinician places a vertical tick mark on a 100-mm horizontal line to denote the severity of the disorder, with a higher value indicating greater severity. Thus, continuous interval data between 0 and 100 can be derived for each aspect of vocal quality and applied to statistical analysis where appropriate.

The CAPE-V also allows the clinician to note other voice features for a particular patient, as needed.

The CAPE-V stipulates that the individual whose voice is to be assessed perform three specific vocal tasks: Task 1 consists in sustain the vowels /a/ and /i/ three times each. Task 2 consists in reading six specific sentences with different phonetic contexts. Task 3 consists in converse naturally in response to the standard question (“Tell me about your voice problem”).

A Spanish version of the CAPE-V was developed adapting the specific sentences (task 2) with different phonetic context in order to create six sentences of varied speech contexts from which to assess different elements of vocal quality. Task 1 and 3 do not need language adaptation.

The purpose of the current study was twofold: first, to examine intra- and interrater reliability of experienced voice clinicians’ judgments of voice quality using the Spanish version of CAPE-V and GRBAS, and second, to establish the empirical validity of the Spanish version of CAPE-V by assessing the relationships between the two scales.

Methods and Materials

1. Development of the six specific sentences of task 2. Sentence 1: The sentence “Nuria ojé una pajarita y una blusa amarillas” corresponds to the original sentence “The blue spot is on the key again”, used to examine the coarticulatory influence of three vowels /a/, /i/, /u/. Sentence 2: The sentence “Marta multa mi moto más mágica” corresponds to the original sentence “How hard did he hit him?”, which provides a context to assess soft glottal attacks and voiceless to voiced transitions. Sentence 3: The sentence “La frase brilla en la mano” corresponds to the original sentence “We were away a year ago” that features all voiced phonemes and provides a context to judge possible vowel stoppages, pauses and one’s ability to link from one word to another. Sentence 4: The sentence “Irene adora hacer huevos al horno” corresponds to the original sentence “We eat eggs at Easter” includes several vowel-initiated words that may provoke hard glottal attacks and provides the opportunity to assess whether these occur. Sentence 5: The sentence “Mi mamá me mima una manzana” corresponds to the original sentence “My mamma makes lemon jam” that includes numerous nasal consonants, thus providing the opportunity to assess hypomania and possibly stimulability for resonant therapy. The last sentence: “Mi teta está a tu pata” corresponds to the original sentence “Peter will keep at the peak” contains no nasal consonants and provides a useful context for assessing intraral vocal pressure and possibly hypernasality or nasal air emission.

2. Voice Stimuli. Dyphonic voice samples. Seventy two dysphonic patients and sixteen healthy volunteers were included in the study. All recordings included speech productions recorded by the published CAPE-V protocol (Kempster et al., 2009), including sustained /a/ and /i/ vowels, six sentence repetitions, and a brief sample of conversational speech in response to a consistent question prompt. The same recording procedure was used to obtain all samples, using the KayPentax Computerized Speech Lab Model 4500 with a sampling rate of 22 kHz. A headset microphone was used, maintained at a distance of 5 cm from the speaker’s mouth. Ambient room noise was minimal. The dyphonic voice samples represent a range of disordens, ages, and severity levels. Two listeners screened all the voice samples and judged independently for severity using a 4-point Likert scale (1 = normal, 2 = mild dysphonia, 3 = moderate dysphonia, and 4 = severe dysphonia) and an analog-visual scale. Raters were asked to make judgments based on the conversational speech sample and the sustained vocal samples. They listened to the voice samples in a free-field environment they judged to be free of potential distraction and excessive ambient noise.

A group of dyphonic patients were selected for a pretreatment and post-treatment analysis of the CAPE-V items (longitudinal analysis).

Conclusions

The CAPE-V was developed to promote a standardized approach to evaluate and document auditory-perceptual judgment of voice quality (Kempster et al., 2009). This study provides evidence of the empirical validity of its Spanish version, which justifies the use of the CAPE-V in clinical practice.

References