

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

Faustino Nunez-Batalla MD^{1,2}; Marta Morato,-Galan MD¹; Jose Luis Llorente-Pendas, MD^{1,2}

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 inexistent. 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, inter-rater reliability and CAPE-V versus GRABS judgments.

Results: Inter-rater reliability coefficients for the CAPE-V ranged from 0.93 for overall severity to 0.54 for intensity; intra-rater reliability ranged from 0.98 for overall severity to 0.85 for intensity. The comparison of judgments between GRABS and CAPE-V 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 voice 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 voice 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 intrarater 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 relationships between the two scales.

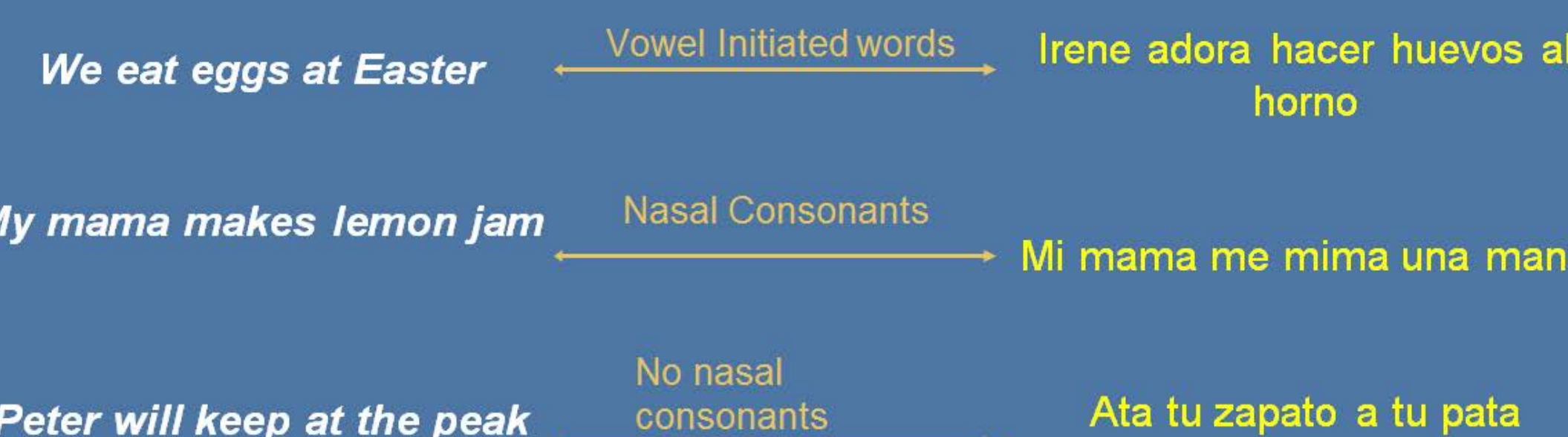
Methods and Materials

1. Development of the six specific sentences of task 2. Sentence 1: The sentence “Nuria ojeó una pajarita y una blusa amarillas” corresponds to the original sentence “The blue spot is in 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 llave 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 voiced stoppages/spasms 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 mano” corresponds to the original sentence “My mamma makes lemon jam” that includes numerous nasal consonants, thus providing the opportunity to assess hyponasality and possibly stimulability for resonant therapy. The last sentence: “Ata tu zapato 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 intraoral pressure and possibly hypernasality or nasal air emission.

2. Voice Stimuli. Dysphonic voice samples. Seventy two dysphonic patients and sixteen healthy volunteers were included in the study. All recordings included speech productions required 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 dysphonic voice samples represent a range of disorders, 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 dysphonic patients were selected for a pretreatment and post-treatment analysis of the CAPE-V items (longitudinal analysis).

Spanish adaptation



Results

Intrarater Reliability

	Task 1	Task 2	Task 3
Intraclass correlation (98% - 85%)	0.986	0.972	0.978
Global severity	0.981	0.969	0.972
Roughness	0.979	0.968	0.952
Strain	0.971	0.921	0.935
Tone	0.963	0.918	0.884
Intensity	0.961	0.866	0.861

Interrater Reliability

	Task 1	Task 2	Task 3
Intraclass correlation (93% - 54%)	0.935	0.879	0.833
Global severity	0.907	0.775	0.750
Roughness	0.894	0.773	0.769
Breathiness	0.848	0.698	0.738
Strain	0.795	0.710	0.808
Tone	0.783	0.646	0.617
Intensity			

CAPE-V versus GRABS judgements

	T1	T2	T3
Intraclass correlation (79.4%)	0.986	0.788	0.866
Grade - Global severity	0.950	0.750	0.842
Roughness - ASP	0.961	0.133	0.742
Breathiness - CA	0.912	0.819	0.799

Longitudinal CAPE-V item analysis

	GRABS	CAPEV	Mean differences (Pre and Post treatment)
Global Severity	Yes	Yes	
Roughness	No	Yes	Statistically significant difference Following treatment
Breathiness	No	Yes	
Strain	No	No	

Discussion

Auditory-perceptual scales are used clinically to assess voice quality. To establish the empirical validity of the CAPE-V, we compared experienced raters’ judgments of voice quality to judgments they made using another commonly used auditory-perceptual scale, the GRBAS. The extent to which raters’ judgments can discern voice quality attributes reflects the instrument’s content validity, while the agreement between raters’ judgments using the CAPE-V versus GRBAS reflects empirical (concurrent) validity. These findings demonstrate that intra- and interrater reliability coefficients for the CAPE-V are higher than those for the GRBAS. The strong correlations between the two scales suggest that the spanish version of the CAPE-V is empirically valid.

Strengths of the CAPE-V include its use of a defined elicitation protocol, use of a consistent and ecologically valid conversational speech probe, and the inclusion of phonetically diverse speech contexts. The CAPE-V’s visual analog scales also yield interval-level data, allowing the use of parametric statistics, and thereby bringing greater power. Consequently, the CAPE-V may be more sensitive to small differences within and across auditory-perceptual judgments than the GRBAS would provide. However, the GRBAS may be faster to administer in clinical settings. Globally, the GRBAS appears to be the most widely used auditory-perceptual tool.

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.

Contact

Faustino Nunez-Batalla, MD
Hospital Universitario Central de Asturias
Email: fjnunezb@gmail.com
Website:

References

1. Kempster GB, Gerratt BR, Verdolini Abbott K, Barkmeier-Kramer J, Hillman RE. Consensus Auditory Perceptual Evaluation of Voice: Development of a standardized clinical protocol. Am J Speech Lang Pathol. 2009; 18: 124-132.
2. Zraick RJ, Kempster GB, Connor NP, Thibeault S, Klaben BK, Bursac Z, Glaze LE. Establishing validity of the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V). Am J Speech-Language Pathol 2011; 20: 14-22.
3. Alexandre NMC, Guirardello Ede B. Cultural adaptation of instruments utilized in occupational health. Rev Panam Salud Publica 2002;11:109-111.
4. Beaton DE, Bombardier C, Guillemin F, Bosi-Ferraz M. Guidelines for the process of cross-cultural adaptation of self-reports measures. Spine. 2000;25:3186-3191.
5. Karnell M, Melton S, Childes J, Coleman T, Dailey S, & Hoffman H. Reliability of clinician-based (GRBAS and CAPE-V) and patient-based (V-RQLQ and IPVI) documentation of voice disorders. Journal of voice. 2007; 21:576-590
6. Zraick R, Wendel D, Smith-Olinde L. The effect of speaking task on perceptual judgment of the severity of dysphonic voice. J Voice. 2005; 19:574-581.
7. Maryn Y, Corthals P, Cauwenberge PV, Roy N, de Bodt M. Toward improved ecological validity in the acoustic measurement of overall voice quality: combining continuous speech and sustained vowels. J Voice. 2010; 24(5); 540-555.
8. Dejonckere PH, Obbens C, de Moor GM, Wieneke GH. Perceptual evaluation of dysphonia: reliability and relevance. Folia Phoniatr (Basel). 1993;45:76-83.
9. de Krom G. Consistency and reliability of voice quality ratings for different types of speech fragments. J Speech Hear Res. 1994;37:985-1000.
10. Wolfe VI, Cornell R, Fitch J. Sentence/vowel correlation in the evaluation of dysphonia. J Voice. 1995; 9:297-303.

ANEXO A: PROTOCOLO CAPE-V (ESPAÑOL)

Paciente: _____
Fecha: _____

Tarea 1: Vocales Sostenidas /a/ y /e/
Tarea 2: Frases
Nuria ojeó una pajarita y una blusa amarillas
Marta multa mi moto más mágica
La llave brilla en la mano
Irene adora hacer huevos al horno
Mama me mima una mano
Ata tu zapato a tu pata

Tarea 3: Habla natural ¿Cómo percibe su voz? ¿Cuáles son sus aficiones?

SEVERIDAD GENERAL _____ /100

ASPEREZA _____ /100

CALIDAD AÉREA _____ /100

TENSIÓN _____ /100

TONO _____ /100

INTENSIDAD _____ /100

Epígrafes a utilizar: T1 (tarea 1); T2 (tarea 2); T3 (tarea 3)
C (consistente) | I (intermitente)

REALIZADO POR: _____