

# Nueva herramienta para la evaluación perceptiva de la calidad de voz

Eugenia San Segundo Fernández

Department of Criminal Science & Technology  
Shanxi Police College, Taiyuan – China



Dpto. de Lengua Española y Lingüística General  
Universidad Nacional de Educación a Distancia



# 1. Introducción

# Cualidad de voz

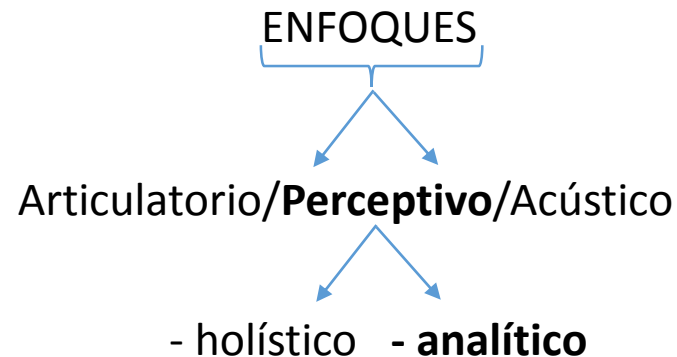
## 1. Definición estrecha:

- actividad de las cuerdas vocales
- análisis de vocales sostenidas



## 2. Definición amplia:

- “quasi-permanent quality resulting from a combination of long-term laryngeal and supralaryngeal features” (Laver, 1980)
- análisis de habla espontánea



# Enfoque perceptivo (I)

- Existencia de numerosos protocolos para evaluar la calidad de voz
  - GRBAS (*Grade + Roughness + Breathiness + Asthenia + Strain*) [Hirano 1981](#)
  - CAPE-V (*Consensus Auditory-Perceptual Evaluation of Voice*: 6 parámetros) [Kempster et al. 2009](#)
  - SVEA (*Stockholm Voice Evaluation Approach*: 13 parámetros) [Hammarberg 2000](#)
  - VPA (Vocal Profile Analysis: ~40 parámetros) [Laver 1980; Beck 2005, 2007](#)
- Uso del VPA:
  - Bastante extendido en el contexto de la **fonética judicial**
  - Menos conocido y aplicado en el **ámbito clínico**

# Enfoque perceptivo (II)

- Críticas al protocolo VPA:
  - demasiadas categorías (i.e. ajustes o 'settings' en inglés), lo cual implica un exceso de decisiones perceptivas que el experto debe tomar a la hora de evaluar una voz. [Webb et al. 2004](#); [McGlashan & Fourcin 2008](#)
- Carácter multidimensional de la voz **versus** desconocimiento de la naturaleza psicoacústica de las dimensiones del VPA:
  - ¿es recomendable usar el mismo tipo de **escala (ordinal)** para evaluar perceptivamente todas las dimensiones del VPA (e.g. 'nasalidad', 'protrusión labial' o 'tensión laríngea')?
  - o, por el contrario, ¿se prestarían mejor algunos de los ajustes articulatorios que recoge el VPA a su evaluación mediante una **escala visual analógica (EVA)**?

	FIRST PASS		SECOND PASS							
	Neutral	Non-neutral	SETTING	Moderate			Extreme			
				1	2	3	4	5	6	
<b>A. VOCAL TRACT FEATURES</b>										
1. Labial			Lip rounding/protrusion							
			Lip spreading							
			Labiodentalization							
			Extensive range							
			Minimised range							
2. Mandibular			Close jaw							
			Open jaw							
			Protruded jaw							
			Extensive range							
			Minimised range							
3. Lingual tip/blade			Advanced tip/blade							
			Retracted tip/blade							
4. Lingual body			Fronted tongue body							
			Backed tongue body							
			Raised tongue body							
			Lowered tongue body							
			Extensive range							
5. Pharyngeal			Pharyngeal constriction							
			Pharyngeal expansion							
6. Velopharyngeal			Audible nasal escape							
			Nasal							
			Denasal							
7. Larynx height			Raised larynx							
			Lowered larynx							
<b>B. OVERALL MUSCULAR TENSION</b>										
8. Vocal tract tension			Tense vocal tract							
			Lax vocal tract							
9. Laryngeal tension			Tense larynx							
			Lax larynx							
<b>C. PHONATION FEATURES</b>										
		SETTING	Present		Scalar Degree					
			Neutral	Non-neutral	Moderate Extreme					
					1	2	3	4	5	6
10. Voicing type	Voice									
	Falsetto									
	Creak									
	Creaky									
	Whisper									

# Protocolo VPA

- VPA = *Vocal Profile Analysis*
- Laver (1980) – Universidad de Edimburgo
- Cada uno de los órganos del aparato vocal contribuye a la cualidad de voz del hablante
- Concepto fundamental = ‘setting’ (ajuste)
  - Honikman (1964) - L2 + base de articulación

## 33 ajustes:

- 22 **supralaríneos**
- 7 **laríneos**
- 4 **tensión muscular**

## Evaluación:

- 1 – neutro /no neutro
- 2 – escala del 1 al 6

# Simplified Vocal Profile Analysis (SVPA)

## Simplified Vocal Profile Analysis (SVPA)

A. Featural (tick the appropriate box)				
Major Setting Groups	Settings	Numerical Labels for One Neutral (N) and Two Non-Neutral Configurations		
		-1	0	+1
Vocal tract settings	Labial	Spreading	N	Rounding
	Mandibular	Close	N	Open
	Apical	Retracted	N	Advanced
	Dorsal	Backed and lowered	N	Fronted and raised
	Velopharyngeal	Denasal	N	Nasal
	Pharyngeal	Constricted	N	Expanded
	Laryngeal height	Lowered	N	Raised
	Overall muscular tension	Vocal tract tension	Lax	N
	Laryngeal tension	Lax	N	Tense
Phonation	Voice type	Whisper/Breathy	N	Creaky/Harsh
B. Holistic				
(fill with qualitative input; comments, etc)				

San Segundo, E. & Mompeán, J.A. (2017). A Simplified Vocal Analysis Protocol for the Assessment of Voice Quality and Speaker Similarity. *Journal of Voice*, 31 (5): 644.e11 - 644.e27

## 2. Objetivos



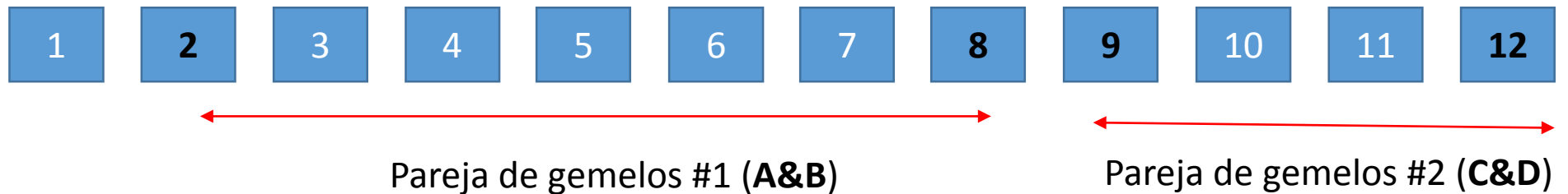
# Objetivos

1. Presentar una nueva herramienta para la evaluación perceptiva de la calidad de voz.
  - Adaptación del protocolo SVPA ([San Segundo & Mompeán 2017](#))
  - Consistente en:
    1. implementación de una escala visual analógica (EVA)
    2. creación de una versión online del protocolo, accesible para investigadores o profesionales de cualquier ámbito (e.g. clínico, forense)
2. Presentar los resultados obtenidos en una primera evaluación de voces usando esta herramienta:
  - **grado de acuerdo** entre oyentes (jueces perceptivos) – fiabilidad de la herramienta
  - métodos de visualización de **parecido entre hablantes**

### 3. Materiales y metodología

# Hablantes y Estímulos

- 12 hablantes varones (Corpus de Gemelos, [San Segundo 2014](#))
  - Hablantes nativos de español centro-norte peninsular
  - Edad comprendida entre 18 y 36 años (media: 26.67)



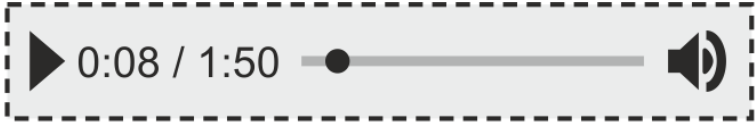
- Estímulos
  - Muestras de voz (90-120 segundos)
  - Conversaciones espontáneas semidirigidas

# Oyentes

- Oyentes (evaluadores de la calidad de voz):
  - 5 hablantes nativos de checo
  - Fonetistas (2 senior + 3 estudiantes de doctorado)
  - Conocimiento de español: 0 – intermedio bajo
  - Entrenamiento en VPA + SVPA durante 2 días
- Tarea: escuchar + analizar la calidad de voz de cada hablante
  - Mismo orden de escucha
  - Procedimiento: interfaz web, acceso abierto → →



# Metodología: el protocolo VPA-VAS\*

**Listener1**





**NEXT**



**Phonation Type**

Whisper/Breathy **70**  Creaky/Harsh 

**Larynx Tension**

Lax **60**  Tense 

**Velopharynx (nasality)**

Denasal **75**  Nasal 

<https://eugeniasansegundo.github.io/vas/tool.html>

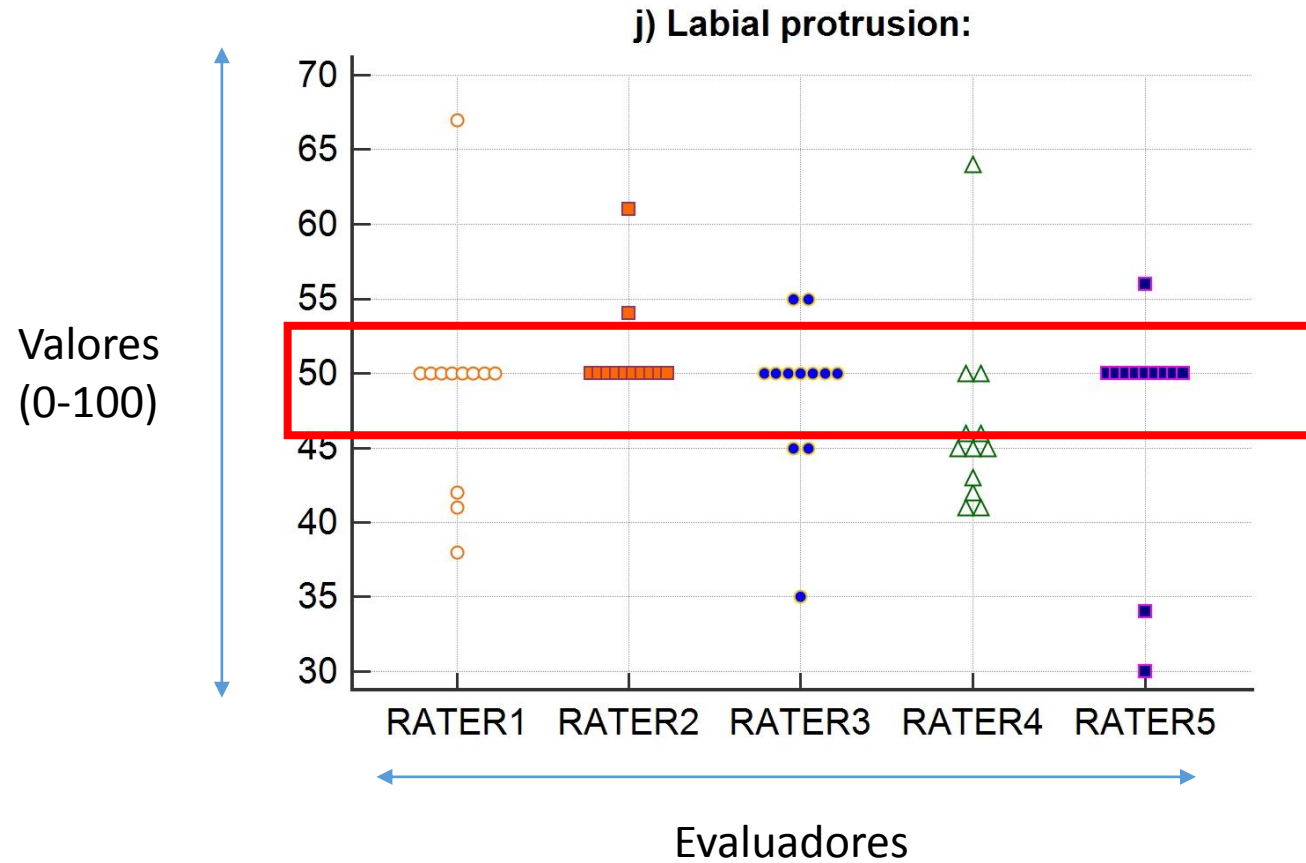
\*VAS = *Visual Analogue Scale* (español = Escala Visual Analógica)

## 4. Resultados

# Consistencia entre evaluadores

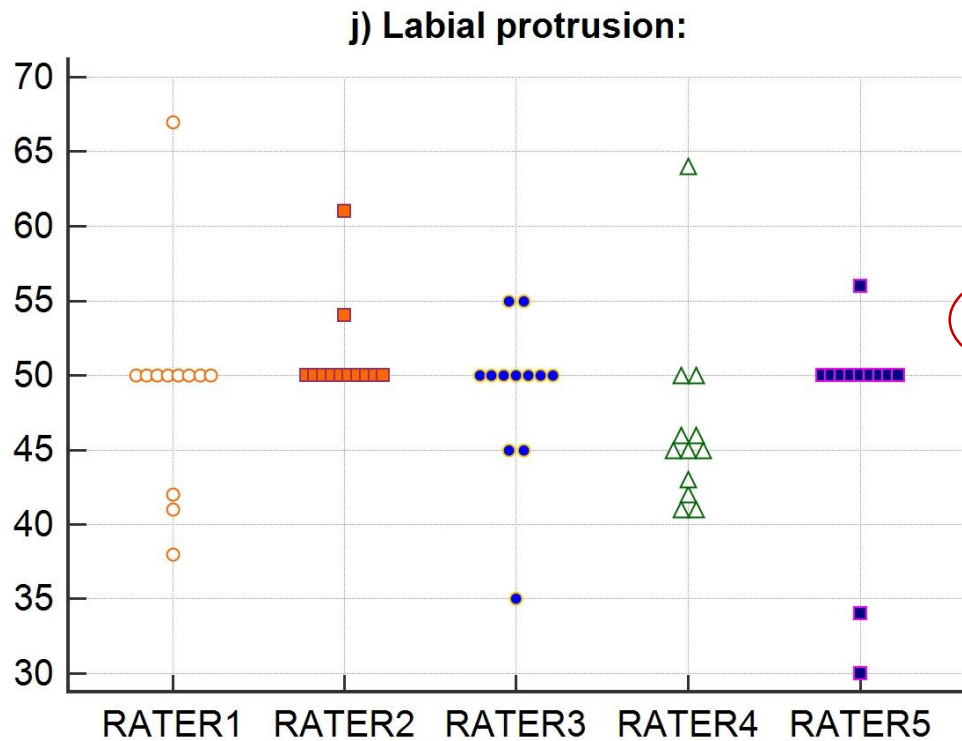
Parámetro	CCI	95% IC
<i>Phonation type</i>	0.786	0.507 – 0.930
<i>Larynx tension</i>	0.836	0.624 – 0.946
<i>Vocal tract tension</i>	0.588	0.052 – 0.865
<i>Larynx height</i>	0.676	0.255 – 0.894
<i>Pharynx expansion</i>	0.660	0.218 – 0.888
<i>Velopharynx (nasality)</i>	0.764	0.457 – 0.923
<i>Tongue body fronting</i>	0.448	-0.270 – 0.819
<i>Tongue tip fronting</i>	0.458	-0.247 – 0.822
<i>Mandible openness</i>	0.552	-0.031 – 0.853
<del><i>Labial protrusion</i></del>	-0.156	-1.661 – 0.620

# Patrones de distribución de los valores a lo largo de la EVA

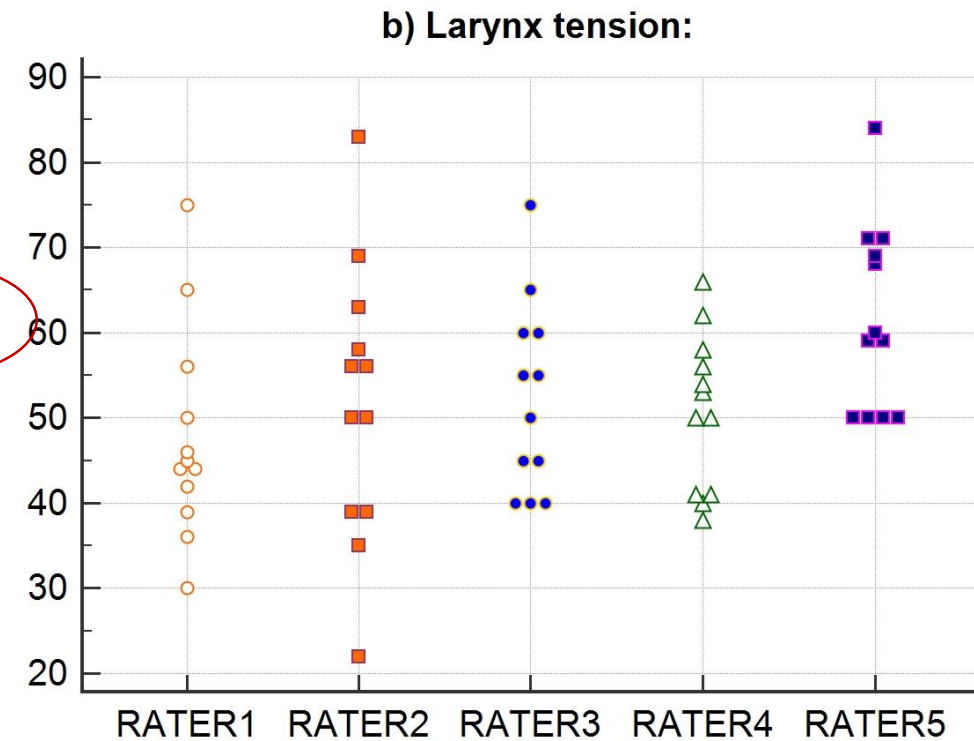




# Patrones de distribución de los valores a lo largo de la EVA

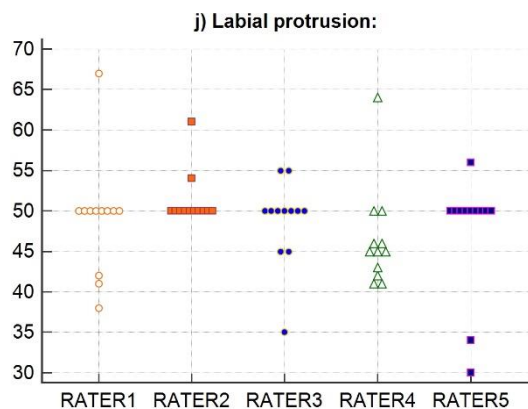
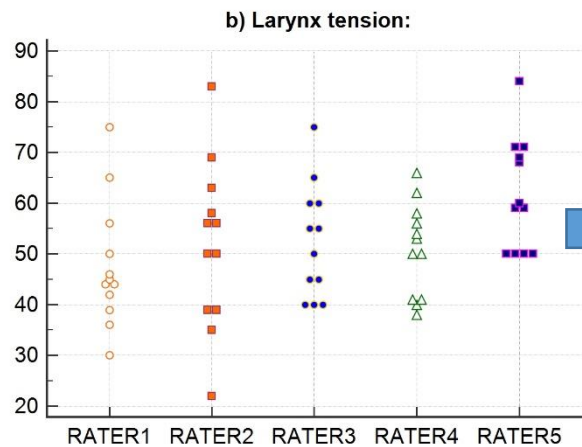


versus



- casi ningún hablante es percibido como neutro
- valores bien dispersos a lo largo de la escala
- distribución: casi todos los puntos de la escala se han usado (20-90)

# Consistencia entre evaluadores



Parámetro	CCI	95% IC
<i>Phonation type</i>	0.786	0.507 – 0.930
<i>Larynx tension</i>	0.836	0.624 – 0.946
<i>Vocal tract tension</i>	0.588	0.052 – 0.865
<i>Larynx height</i>	0.676	0.255 – 0.894
<i>Pharynx expansion</i>	0.660	0.218 – 0.888
<i>Velopharynx (nasality)</i>	0.764	0.457 – 0.923
<i>Tongue body fronting</i>	0.448	-0.270 – 0.819
<i>Tongue tip fronting</i>	0.458	-0.247 – 0.822
<i>Mandible openness</i>	0.552	-0.031 – 0.853
<del><i>Labial protrusion</i></del>	-0.156	-1.661 – 0.620

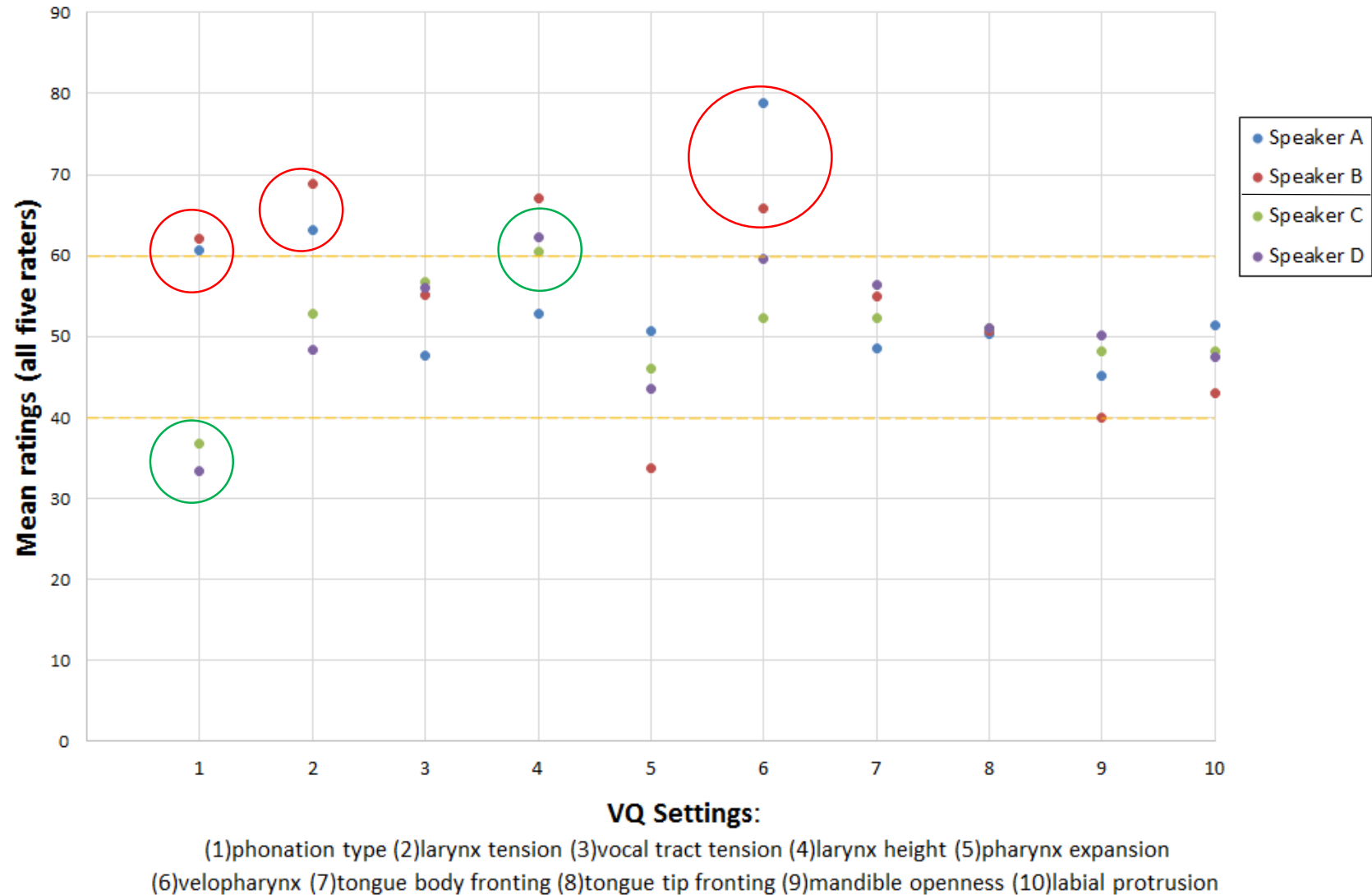
CCI: Coeficiente de Correlación Intra-Clase

# Parecido entre hablantes

- Explorar la utilidad de la nueva herramienta en contextos de comparación entre pares de voces:
  - Comparaciones entre dos muestras de voz (dubitada e indubitada) – F. Judicial
  - Comparaciones de un mismo hablante en dos sesiones diferentes – F. Clínica
- Gemelos:
  - Casos extremos de similitud entre hablantes (tracto vocal + fonación)
  - Útiles para mostrar si los evaluadores detectan sus similitudes
  - Los oyentes no sabían que había gemelos en el test!


# Parecido entre hablantes

Voces similares:  
evaluación de CV  
coincidente



## 5. Discusión y conclusiones

# Discusión

- Consistencia (acuerdo) entre evaluadores
  - Acuerdo entre *moderado* y *muy bueno* en la evaluación perceptiva de la mayoría de los parámetros de calidad de voz usando la nueva herramienta propuesta:
    - CCI medio = 0.64
    - resultados especialmente buenos resultados teniendo en cuenta:
      - (1) escaso entrenamiento de los oyentes
      - (2) no nativos
  - Diferencias entre dimensiones:
    - **tensión laríngea, tipo de fonación, nasalidad vs. protrusión labial, adelantamiento lingual**
    - **distinta naturaleza psicoacústica de las dimensiones perceptivas** (Stevens 1975)
- EVA  - **continuos protéticos**: c. cuantitativo aditivo (dimensión varía en magnitud o cantidad)
- **continuos metatéticos**: c. cualitativo sustitutivo (implica cambio en la calidad)



Escalas de intervalos iguales o escalas ordinales

# Conclusiones

- Escalas Visuales Analógicas (EVA) → capturar “continuo” subyacente
  - Estudios previos para algunas dimensiones:
    - Nasalidad = protética ([Cheng 2007](#))
    - Voz aspirada (‘breathy’) = metatética ([Sewall et al. 1999](#))
  - Naturaleza psicoacústica de la mayoría de los atributos perceptivos no se ha abordado aún.
- Primer estudio que implementa una EVA para el protocolo VPA
  - Estudios previos sobre idoneidad de distintos tipos de escalas en otros protocolos perceptivos
    - Escala GRBAS: escala ordinal de 4 puntos vs. EVA [Wuyts et al 1999](#)
    - Escala GRBAS (ordinal) vs. CAPE-V (de tipo EVA) [Nemr et al. 2012](#)

# Gracias por la atención

The screenshot shows the homepage of the Journal of Voice, published by The Voice Foundation. The header includes the journal logo, navigation links (Articles and Issues, For Authors, Journal Info, Subscribe, Society Info, Best Papers, Videos), a search bar, and links for RSS Feeds, Mobile, Login, Register, Claim Subscription, and Subscribe. The main content area features a red-bordered box highlighting an 'Article in Press' titled 'A Computer-Based Tool for the Assessment of Voice Quality Through Visual Analogue Scales: VAS-Simplified Vocal Profile Analysis' by Eugenia San Segundo and Radek Skarnitzl. To the right of the article title is a 'ScienceDirect' link. Below the article title, there are links for 'Open Access', 'PlumX Metrics', and the DOI: <https://doi.org/10.1016/j.jvoice.2019.10.007>. On the right side, under 'Article Tools', there are links for PDF (1 MB), Download Images (.ppt), Email Article, Add to My Reading List, Export Citation, Create Citation Alert, and Cited by in Scopus (0).

**JOURNAL OF VOICE**  
THE VOICE FOUNDATION

RSS Feeds Mobile

Login | Register | Claim Subscription | Subscribe

Articles and Issues ▾ For Authors ▾ Journal Info ▾ Subscribe Society Info ▾ Best Papers Videos

All Content ▾ Search [Advanced Search](#)

< Previous Article [Articles in Press](#) Next Article >

Access this article on [ScienceDirect](#)

Article in Press

**A Computer-Based Tool for the Assessment of Voice Quality Through Visual Analogue Scales: VAS-Simplified Vocal Profile Analysis**

[Eugenia San Segundo](#)\*, [Radek Skarnitzl](#)

Open Access PlumX Metrics

DOI: <https://doi.org/10.1016/j.jvoice.2019.10.007>

Article Tools

- [PDF \(1 MB\)](#)
- [Download Images \(.ppt\)](#)  
[About Images & Usage](#)
- [Email Article](#)
- [Add to My Reading List](#)
- [Export Citation](#)
- [Create Citation Alert](#)
- [Cited by in Scopus \(0\)](#)

<https://eugeniasansegundo.github.io/vas/tool.html>