

Percepción de la tensión del tracto vocal: correlatos acústicos

Eugenia San Segundo,
Sandra Schwab, Volker Dellwo, Lei He
& José Mompeán

Universidad de York, Inglaterra
Universidad de Zúrich, Suiza
& Universidad de Murcia, España

Esquema

1. Introducción/ Contexto

- cualidad de voz
- protocolo VPA (Vocal Profile Analysis)
 - tensión del tracto vocal
 - acuerdo entre evaluadores (ing: interrater agreement)

2. Objetivo de la investigación:

- correlatos acústicos-prosódicos de la tensión del tracto vocal

3. Materiales y metodología

4. Resultados

5. Conclusiones y líneas de trabajo futuro

1. Introducción

Cualidad de voz

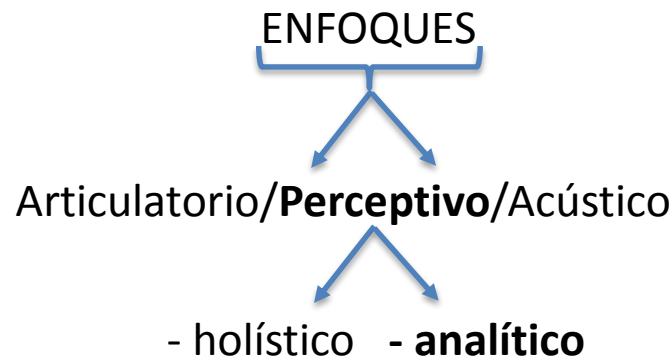
1) Definición estricta:

- 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)



Vocal Profile Analysis (VPA)

	FIRST PASS		SECOND PASS						
	Neutral	Non-neutral	SETTING	Moderate			Extreme		
				1	2	3	4	5	6
A. VOCAL TRACT FEATURES									
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						
			Minimised range						
			Pharyngeal constriction						
5. Pharyngeal			Pharyngeal expansion						
			Audible nasal escape						
6. Velopharyngeal			Nasal						
			Denasal						
			Raised larynx						
7. Larynx height			Lowered larynx						
B. OVERALL MUSCULAR TENSION									
8. Vocal tract tension			Tense vocal tract						
			Lax vocal tract						
9. Laryngeal tension			Tense larynx						
			Lax larynx						
C. PHONATION FEATURES									
	SETTING	Present		Scalar Degree					
		Neutral	Non-neutral	Moderate			Extreme		
				1	2	3	4	5	6
10. Voicing type	Voice								
	Falsetto								
	Creak								
	Creaky								
	Whisper								

- 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 **supralaríngenos**
- 7 **laríngenos**
- 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

Tensión del tracto vocal (acuerdo entre jueces)

- Coeficiente kappa de Cohen (κ)
- 2 jueces perceptivos → 24 hispanohablantes (♂)

- Consistencia interna (2 evaluaciones cada juez; intervalo de 1 semana).

Juez 1: 0.81 κ

Juez 2: 0.91 κ

TABLE 3.

Raw Agreement and Unweighted Kappa Results of Inter-Rater Agreement Between R1 and R2

Setting	Percentage Agreement	Cohen's Kappa	N Disagreements
Labial	75.00	0.55	6
Mandibular	50.00	0.06	12
Apical	54.17	0.11	11
Dorsal	91.67	0.78	2
Velopharyngeal	70.83	0.55	7
Pharyngeal	37.50	0.11	15
Larynx height	66.67	0.50	8
Vocal tract tension	41.67	0.13	14
Laryngeal tension	66.67	0.30	8
Voice type	66.67	0.42	8

BLE 4.

Linear Weighted Kappa Results of Inter-Rater Agreement Between R1 and R2

Setting	Observed Kappa (κ)	Standard Error	95% Confidence Interval		Maximum Possible κ †	Proportional κ to Maximum Possible*
			Lower Limit	Upper Limit		
Labial	0.53	0.17	0.20	0.86	0.80	0.66
Mandibular	0.11	0.15	0	0.41	0.56	0.20
Apical	0.14	0.12	0	0.39	0.28	0.50
Dorsal	0.79	0.13	0.52	1	0.79	1
Velopharyngeal	0.59	0.14	0.32	0.86	0.90	0.66
Pharyngeal	0.19	0.12	0	0.42	0.49	0.38
Larynx height	0.66	0.11	0.37	0.83	0.70	0.66
Vocal tract tension	0.21	0.16	0	0.51	0.82	0.25
Laryngeal tension	0.37	0.13	0.11	0.63	0.65	0.57
Voice type	0.43	0.17	0.09	0.76	0.94	0.45

* Maximum possible linear weighted kappa given the observed marginal frequencies.

† Observed kappa as proportion of maximum possible.

$\kappa = 0.40 - 0.80$
moderado – sustancial

2. Objetivo de la investigación

Búsqueda de correlatos acústico-prosódicos

- ¿Por qué?

Jueces diferentes → atienden a aspectos fonéticos diferentes

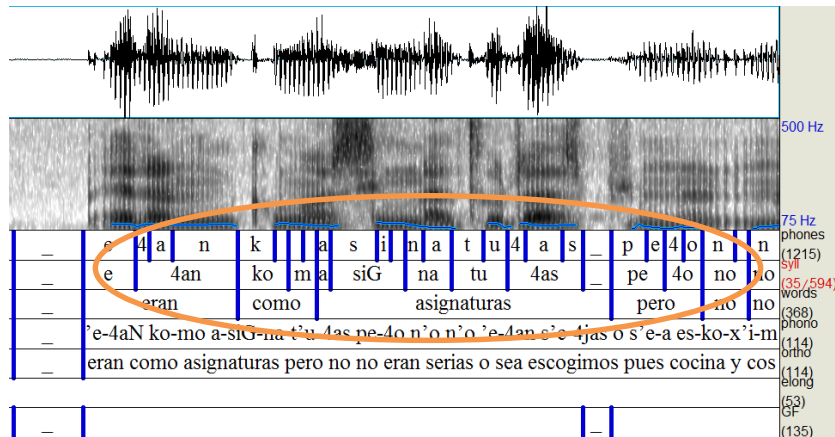
- ¿Qué correlatos?

	Lax vocal tract	Tense vocal tract
Perceptual corr.	<ul style="list-style-type: none">- Increased nasality- Min. Lip/Jaw/Tongue Range	<ul style="list-style-type: none">- Reduced nasality- Ext. Lip/Jaw/Tongue Range+ Pharyngeal constriction
Spectral corr.	<ul style="list-style-type: none">- Damping of high frequency energy in the spectrum- Broad formant peaks	<ul style="list-style-type: none">- Less absorption of high frequency energy by the vocal tract walls, resulting in sharper formant peaks.
Prosodic corr.	<ul style="list-style-type: none">- Slower tempo- Phonetic undershoot: consonantal and vocalic reduction- Higher incidence of unstressed syllables	<ul style="list-style-type: none">- Faster tempo- Precise articulation, fortition- As for consonants, the stop releases of tense speakers are perceived as sharper

3. Materiales y metodología

Materiales

1. Grabaciones de voz



- 24 hablantes ♂ (20-36 años)
- español centropeninsular
- conversación espontánea
participante – investigador
(10 min) → 2 min

2. Evaluación perceptiva

Simplified Vocal Profile Analysis (SVPA)

A. Featural (tick the appropriate box)		Numerical Labels for One Neutral (N) and Two Non-Neutral Configurations		
Major Setting Groups	Settings	-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	Tense
	Laryngeal tension	Lax	N	Tense
Phonation	Voice type	Whisper/Breathy	N	Creaky/Harsh
B. Holistic (fill with qualitative input; comments, etc)				

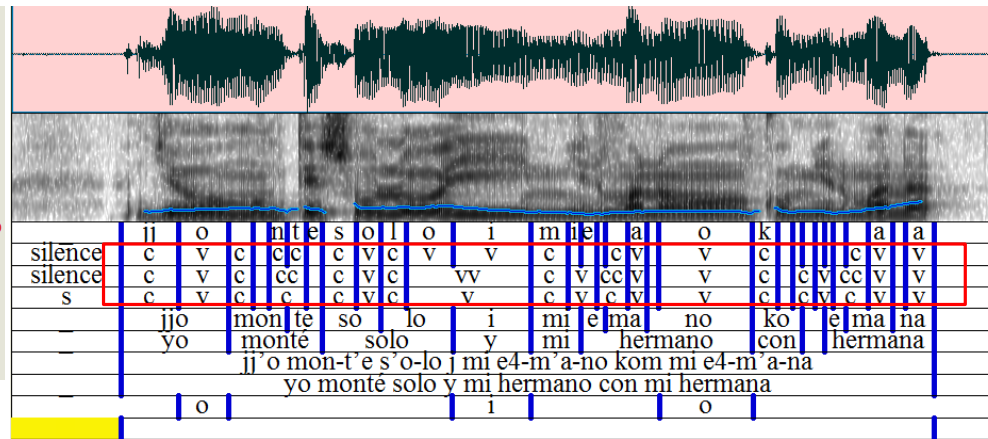
- 24 hablantes x 2 jueces (SVPA)
(solo los “ratings” correspondientes a
tensión del tracto vocal):

“laxo” - “neutro” - “tenso”

San Segundo, E.(2014). *Forensic speaker comparison of Spanish twins and non-twin siblings: A phonetic-acoustic analysis of formant trajectories in vocalic sequences, glottal source parameters and cepstral characteristics*, Tesis doctoral (UIMP-CSIC)

Metodología:

Alineación y etiquetado



- Praat plugin *CV tier creator* (V. Dellwo)
 - [p, t, k, m, n,] \rightarrow C
 - [a, e, i, o, u, j, w] \rightarrow V



Metodología: Análisis acústico

(1) Velocidad de articulación:
sílabas / seg.

(2) Variables rítmicas:

Nombre	Definición
mean {Con Vow C V Syl}Ln	Duración media normalizada de {Consonantes Vocales Intervalos Consonánticos Intervalos Vocálicos Sílabas}
delta {Con Vow C V Syl}	Desviación estándar de <i>mean</i> {Con/Vow/C/V/Syl}Ln
varco {Con Vow C V Syl}	Coeficiente de variación de {Consonantes Vocales Intervalos Consonánticos Intervalos Vocálicos Sílabas}
nPVI_ {Con Vow C V Syl}	Índice de variabilidad entre intervalos consecutivos {Consonantes Vocales Intervalos Consonánticos Intervalos Vocálicos Sílabas}

Dellwo, V. (2009) Choosing the right rate normalization methods for measurements of speech rhythm. *Proceedings of AISV*, 13-32.

Metodología:

Análisis acústico

(3) Variables relacionadas con la intensidad:

Nombre	Definición
mean {M P}	Media de {intensidad media pico de intensidad} entre sílabas consecutivas
stdev {M P}	Desviación estándar de <i>mean</i> {M/P}
varco {M P}	Coeficiente de variación de {intensidad media pico de intensidad} entre sílabas consecutivas

(4) Variables relacionadas con alargamientos (C & V)

$$\frac{\{n|dur\} \text{ alargamientos } \{C|V\} \text{ en cada GF}}{\text{duración del GF}}$$

Análisis estadístico

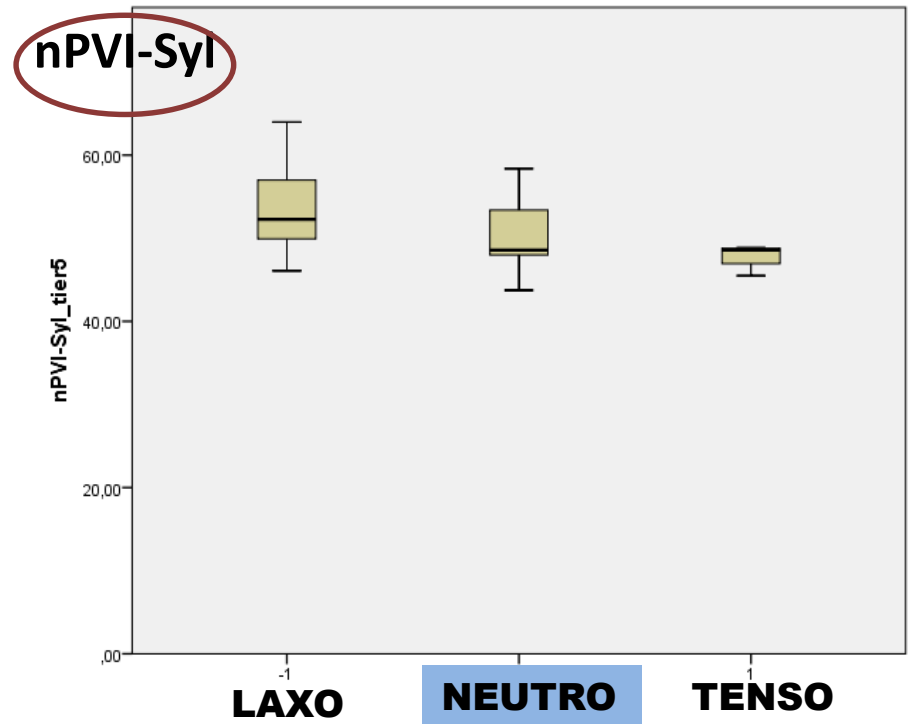
- SPSS (v.24)
 - ✓ Coeficiente de correlación de *Spearman*
 - Modelos de regresión

4. Resultados

Correlaciones

	Eval. percep. Juez 1	Eval. percep. Juez 2
varcoM	0.344*	n.s.
meanConLn	n.s.	0.362*
meanCLn	n.s.	0.395*
nCV	n.s.	-0.358*
varcoSyl	n.s.	-0.396*
deltaSylLn	n.s.	-0.464*
nPVI-Syl	n.s.	-0.490**

* $p < .05$, ** $p < .01$



varcoM = coeficiente de variación de **intensidad media** entre sílabas consecutivas

mayor variabilidad en la intensidad silábica → mayor tensión del tracto vocal

meanConLn = duración media de las duraciones consonánticas normalizadas

meanCLn = duración media de las duraciones de intervalos consonánticos normalizados

mayor duración de consonantes → mayor tensión del tracto vocal

deltaSylLn = desv. Est. de las duraciones silábicas normalizadas

nPVI-Syl = índice de variabilidad entre sílabas consecutivas

mayor variabilidad en la duración silábica → menor tensión del tracto vocal

5. Conclusiones

Conclusiones

- **Primer acercamiento** a la caracterización cuantitativa de un aspecto de CV
= tensión del tracto vocal
....desde un punto de vista acústico-prosódico (!)
 - **Discusión de los resultados**
 - Diferentes correlaciones entre variables acústicas y evaluaciones perceptivas explican (en parte?) los desacuerdos entre jueces
 - Oyentes diferentes prestan atención a pistas acústicas diferentes
 - Juez 1 = variabilidad de intensidad entre sílabas
 - Juez 2 = variabilidad duracional entre sílabas
- ✓ mayor duración de consonantes → mayor tensión del tracto vocal
- ✓ mayor variabilidad en la duración silábica → menor tensión del tracto vocal

	Lax vocal tract	Tense vocal tract
Prosodic corr.	<ul style="list-style-type: none">- Slower tempo- Phonetic undershoot: consonantal and vocalic reduction- Higher incidence of unstressed syllables	<ul style="list-style-type: none">- Faster tempo- Precise articulation, fortition- As for consonants, the stop releases of tense speakers are perceived as sharper

- **Implicaciones**
 - Mejor acuerdo entre expertos → importante para fonética clínica y judicial