

# Voice quality analysis in forensic voice comparison: developing the vocal profile analysis scheme

Eugenia San Segundo, Paul Foulkes, Peter French  
Philip Harrison & Vincent Hughes

University of York & J P French Associates



IAFPA 2016  
University of York  
24-27 July



# 1. Introduction

- survey of practitioners (Gold & French 2011)
  - **voice quality (VQ):** one of most valuable features
    - 94% examine VQ
    - 68% of those do so ‘routinely’
    - 61% use recognised framework (e.g. VPA)
    - 21% perform “auditory analysis and provide some form of a verbal description”

	FIRST PASS		SECOND PASS						
	Neutral	Non-neutral	SETTING	moderate			extreme		
				1	2	3	4	5	6
A. VOCAL TRACT FEATURES									
			Lip rounding/protrusion						
1. Labial			Lip spreading						
			Labiodentalization						
			Extensive range						
			Minimised range						
2. Mandibular			Close jaw						
			Open jaw						
			Protruded jaw						
			Extensive range						
3. Lingual tip/blade			Minimised range						
			Advanced tip/blade						
4. Lingual body			Retracted tip/blade						
			Fronted tongue body						
			Backed tongue body						
			Raised tongue body						
			Lowered tongue body						
			Extensive range						
5. Pharyngeal			Minimised range						
			Pharyngeal constriction						
6. Velopharyngeal			Pharyngeal expansion						
			Audible nasal escape						
			Nasal						
7. Larynx height			Denasal						
			Raised larynx						
			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								
11. Laryngeal friction	Whisper								
	Whispery								
12. Laryngeal irregularity	Harsh								
	Tremor								

## Vocal Profile Analysis

- framework for systematic description of VQ
  - developed by Laver et al. (1981)
- modified by Beck (2007)
  - 25 supralaryngeal
  - 7 laryngeal
- comparison against ‘neutral setting’
  - clearly defined baseline with concrete acoustic and physiological correlates

# 1. Introduction

- **issues with VPA for FVC** (Nolan 2005, 2007)

- 1) lack of training

- 2) practical considerations of time

- 3) quality of samples (telephone trans., short)

- + courts need to know reliability of the method
- + analyses should rely on non-correlated features

# 1. Introduction

- **general issues with perceptual methods (VQ)**
  - bias and errors (Kent, 1997)
  - interrater disagreements (Kreiman et al. 2011)
    - VPA reliability with forensic data not reported yet
  - multidimensionality of VQ
    - dimension reduction (Bele, 2007)
    - dimensions difficult to isolate
      - interrelated dimensions
      - risk of overestimation

## 2. Research questions

*what changes can we make to improve VPA usability for FVC? → **simplified VPA***

1. how often do VPA settings occur? **frequency**
2. how reliable are VPA ratings across different analysts? **interrater agreement**
3. to what extent are VPA settings independent? **correlation tests**

# 3. Data

- **DyViS Corpus** (Nolan et al. 2009)
  - 100 male speakers
  - Standard Southern British English (SSBE)
  - 18-25 years old



## Task 2

information exchange  
over telephone

HQ, near-end recording  
(c. 10-15 mins)

## Manual editing:

Removed...

- ✓ Overlapping speech
- ✓ Background noise
- ✓ Extended pauses

# 4. Methods

## VPA simplified version

- reduced scalar degrees
  - ‘present’ features (1-3)
- reduced N settings
  - combined:
    - *fronted + raised*
    - *backed + lowered*
    - *creak + creaky*
    - *whisper + whispery*

	FIRST PASS		SECOND PASS				Notes
	Neutral	Non-Neutral	SETTING	Slight 1	Mark. 2	Extr. 3	
A. VOCAL TRACT FEATURES							
Labial			Lip rounding/protrusion				
			Lip spreading				
			Labiodentalisation				
			Extensive labial range				
			Minimised labial range				
Mandibular			Close jaw				
			Open jaw				
			Extensive mandibular range				
			Minimised mandibular range				
Lingual tip/blade			Advanced tongue tip/blade				
			Retracted tongue tip/blade				
Lingual body			Fronted tongue body				
			Backed tongue body				
			Extensive lingual range				
			Minimised lingual range				
Pharynx			Pharyngeal constriction				
			Pharyngeal expansion				
Velopharyngeal			Nasal				
			Denasal				
Larynx height			Raised larynx				
			Lowered larynx				
B. OVERALL MUSCULAR TENSION							
Vocal tract tension			Tense vocal tract				
			Lax vocal tract				
Laryngeal tension			Tense larynx				
			Lax larynx				
C. PHONATION FEATURES							
	SETTING	Present		Scalar Degree			
		Neutral	Non-neutral	Slight 1	Mark. 2	Extr. 3	
Voicing type	Voice						
	Falsetto						
	Creaky						
	Whispery						
	Breathy						
	Murmur						
	Harsh						
	Tremor						

# 4. Methods

## Perceptual evaluation:

- Three analysts:



- Two stages:

### 1. Blind perceptual assessment of voices



### 2. Calibration procedure

- joint listening
- disagreement typology:
  - setting reassignment  
e.g. *lowered larynx* ~  
*expanded pharynx*
  - proper disagreement  
e.g. missed presence of a setting  
e.g. different scalar degree

## 5. Results: setting frequency (1)

- based on the mode per setting → **agreed version**

### **Absent settings**

Labiodentalization

Extensive labial range

Minimised labial range

Open jaw

Protruded jaw

Extensive mandibular range

Backed tongue body

Audible nasal escape

Falsetto

Tremor



**100%  
NEUTRAL**

## 5. Results: setting frequency (2)

### Rare settings (<10%)

Lip spreading (5)

Lip rounding (1)

Close jaw (1)

Min. mandibular range (4+1)

Retracted tongue tip (1+1)

Extensive lingual range (3)

Min. lingual range (0+1)

Pharyngeal constriction (3)

Pharyngeal expansion (3)

Denasal (1+3)

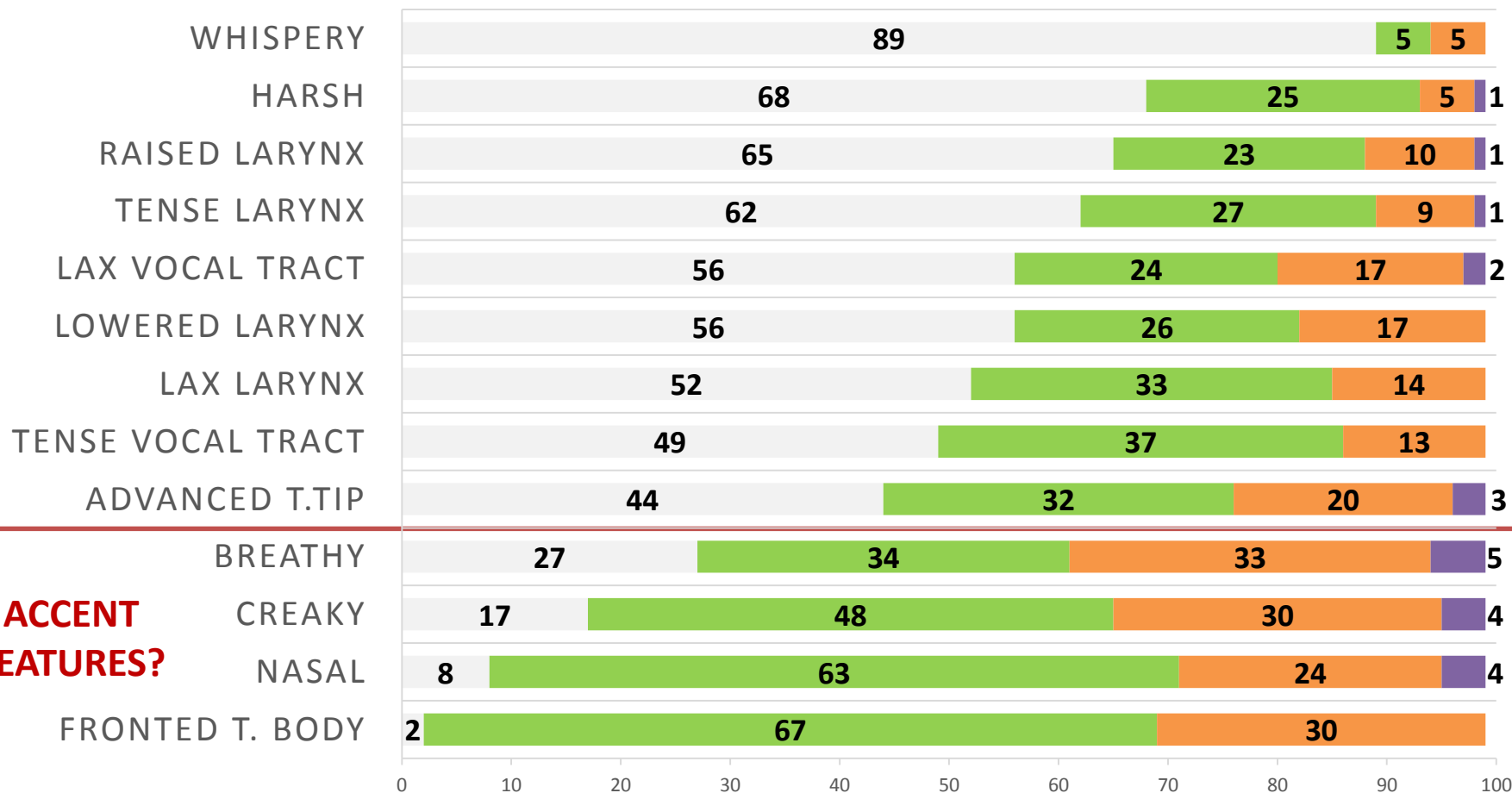
1-5%  
**NON NEUTRAL**



\*(N cases in brackets: slight + moderate)

# 5. Results: setting frequency (3)

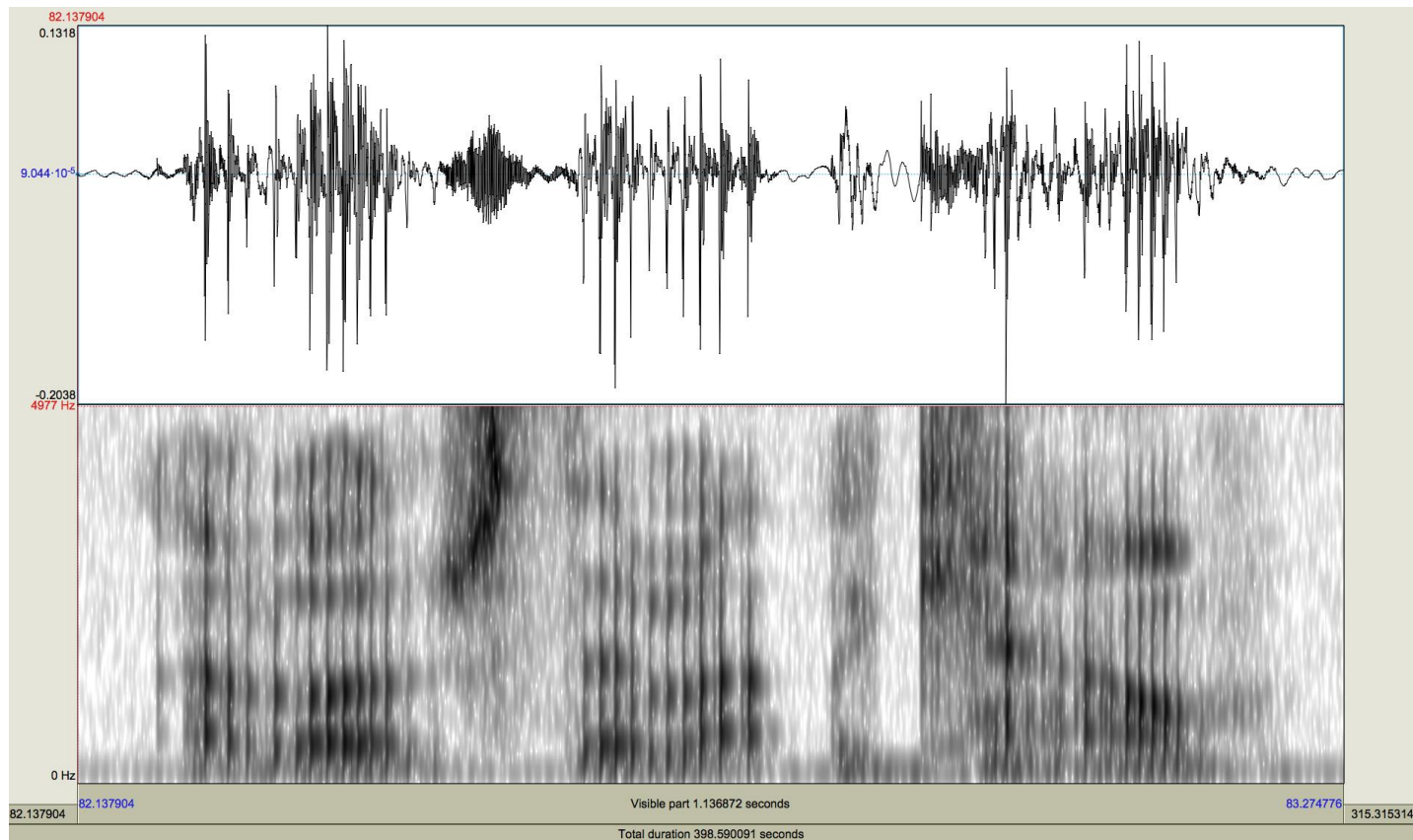
Neutral Slight Moderate Extreme



**ACCENT  
FEATURES?**

# 5. Results: setting frequency (3)

- example *creakiness* – degree “3” 



## 5. Results: correlation tests (1)

- based on the mode per setting → **agreed version**

POSITIVE CORRELATIONS	Contingency Coefficient
RAISED LARYNX - TENSE LARYNX	0.58
NASAL - TENSE LARYNX	0.58
HARSH - TENSE LARYNX	0.57
LAX LARYNX - LOWERED LARYNX	0.52
CREAKY - LAX LARYNX	0.45
ADVANCED TONGUE TIP - FRONTED TONGUE BODY	0.41
CREAKY - LOWERED LARYNX	0.35

## 5. Results: correlation tests (2)

NEGATIVE CORRELATIONS	Contingency Coefficient
LAX VOCAL TRACT - TENSE VOCAL TRACT	0.61
LAX LARYNX - TENSE LARYNX	0.57
LOWERED LARYNX - RAISED LARYNX	0.51
LAX LARYNX - RAISED LARYNX	0.47
CREAKY - RAISED LARYNX	0.44
LOWERED LARYNX - TENSE LARYNX	0.46

# 5. Results: interrater measures

- based on absolute scores:

	HARSH	RAISED LARYNX	LOWERED LARYNX	TENSE LARYNX	LAX LARYNX	ADVANCED TONGUE TIP	LAX VT	TENSE VT	CREAKY	BREATHY	NASAL	FRONTED TONGUE BODY
<b>Average pairwise agreement</b>	75%	74%	67%	67%	62%	59%	59%	55%	52%	52%	43%	36%
Agreement raters 1 & 3	74%	73%	70%	66%	69%	56%	55%	55%	41%	42%	36%	43%
Agreement raters 1 & 2	75%	78%	62%	69%	66%	55%	66%	53%	49%	49%	43%	33%
Agreement raters 2 & 3	76%	71%	71%	68%	51%	66%	58%	59%	65%	64%	49%	31%
<b>Fleiss' kappa</b>	0.43	0.46	0.41	0.34	0.31	0.35	0.29	0.22	0.31	0.31	0.13	0.01

- more realistic definition of disagreement:

- disagreement about presence/ absence (0-1)
- disagreement beyond 1 scalar degree (1-3)

CREAKY	BREATHY	NASAL	FRONTED TONGUE BODY
71%	66%	58%	40%

## 6. Discussion: **setting frequency**

- useful for typicality and LR calculation
  - e.g. absent settings (in this population)
    - phonatory settings: *false* *setto*, *tremor*
    - supralaryngeal settings: *open jaw*, *protruded jaw*, *audible nasal escape*
      - mostly linked to pathological conditions (Beck, 2007)
  - e.g. rare settings
    - supralaryngeal settings: *lip spreading*, *lip rounding*, *denasal*
      - need to consider non-contemporaneous recordings:
        - within- speaker differences?

## 6. Discussion: correlation

- results according to phonetic theory
  - *harsh* ~ *tense larynx*
  - *creaky* ~ *lax larynx* ~ *lowered larynx*
- other deserve further exploration
  - *nasal* ~ *tense larynx*

...but correlations < .60 suggest that further VPA simplifications → not necessary!

## 6. Discussion: interrater

- overall % agreement = good
  - some settings easier to agree upon? more salient?
  - *harshness* also high % agreement in previous studies (Beck 2005: 84%)
- lower % agreement may have simple solutions:
  - increase training
  - search for acoustic correlates
    - e.g. different types of creaky? (Keating et al. 2015)
    - e.g. prosodic correlates of vocal tract tension?

## 5. Conclusion & Future work

- first attempt at simplifying VPA for FVC
- overall good interrater agreement
  - systematic patterns (individuals/listening strategies)
- promising speaker discriminatory value
  - to what extent is a speaker's profile variable across recordings? / how useful is VPA for speaker discrimination?
  - complement to ASR? (e.g. detection of differences between speakers in falsely accepted trials; González-Rodríguez et al. 2014. )

# Thanks! Questions?



THE UNIVERSITY *of York*

J P French Associates  
Forensic speech and acoustics laboratory