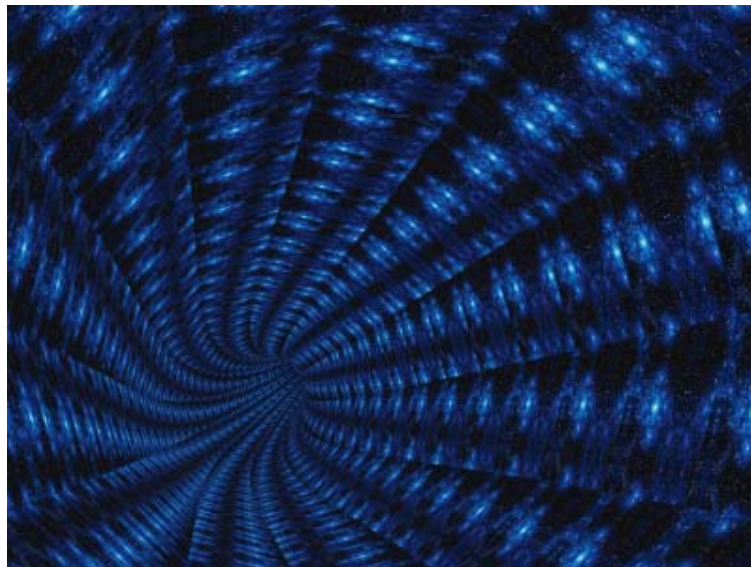


Multidimensional Representation of Personal Quality of Vowels and its Acoustical Correlates

(Matsumoto, Hiki, Sone, Nimura; 1973)



Pedro Davalos

CPSC 689-604

Feb 27, 2007

Outline

- Introduction
- Test 1: /a/
- Test 2: Hybrid
- Test 3: Vowels
- Conclusions

Intro: Goal

- Determine how acoustical properties influence recognizing speakers.

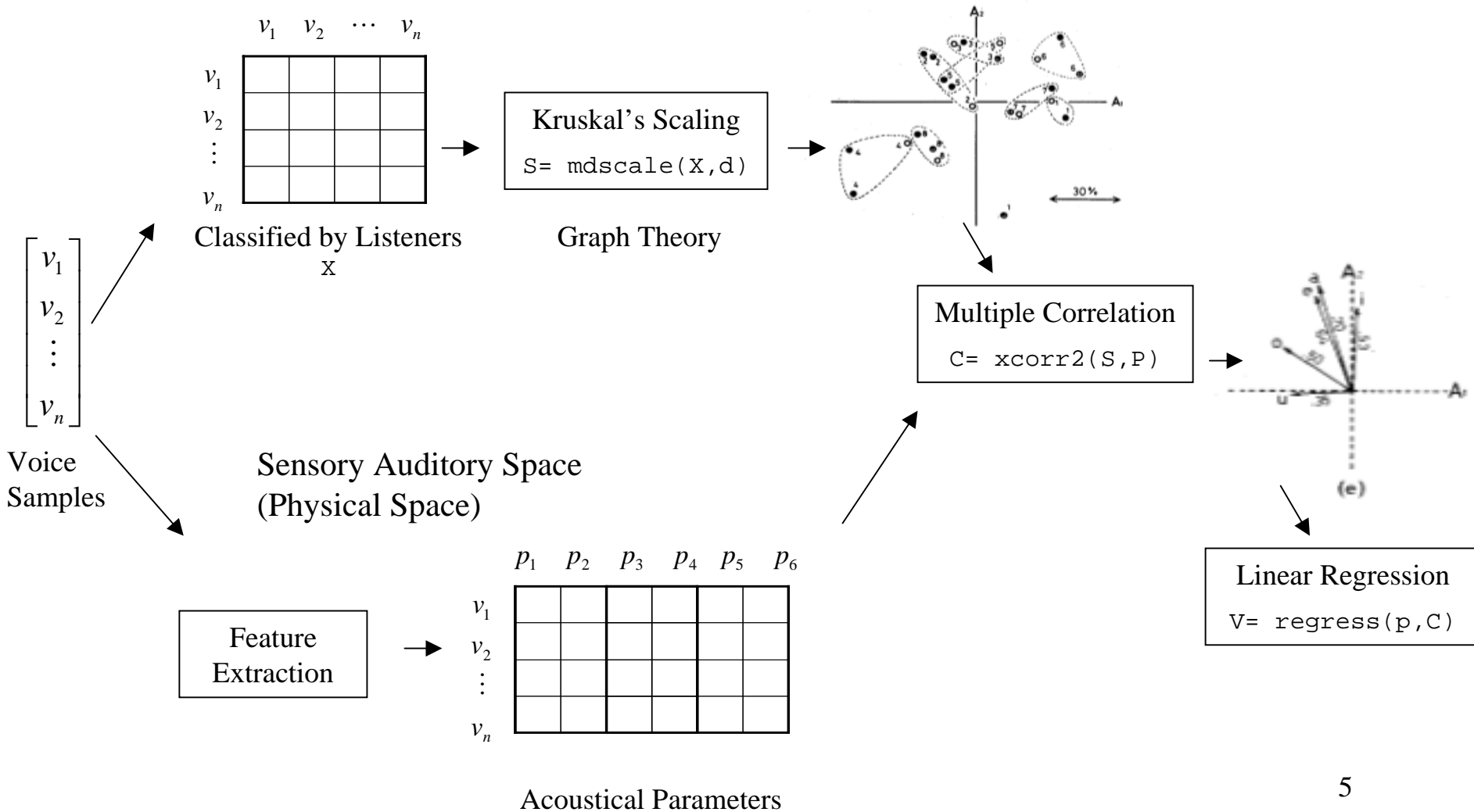
Intro: Background

- “Personal Quality”
 - Is NOT high quality as required to perform at an opera
 - It refers to the speaker’s characteristics and the voice attributes that allow speaker recognition



Intro: Approach

Psychological Auditory Space (PAS)

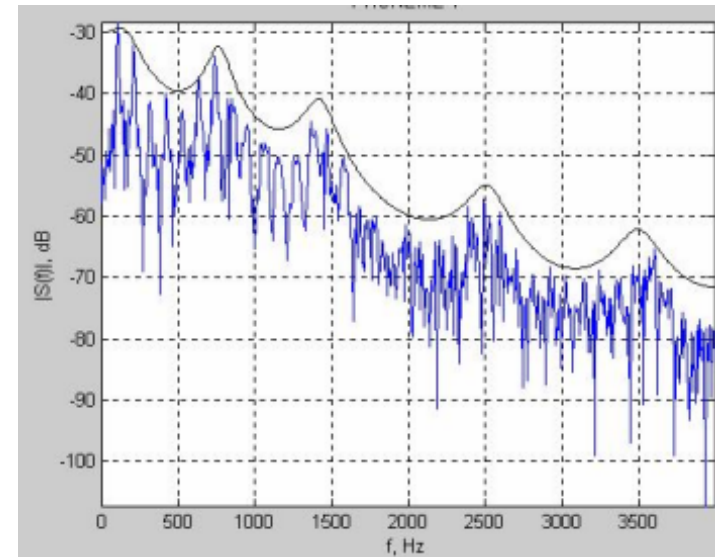
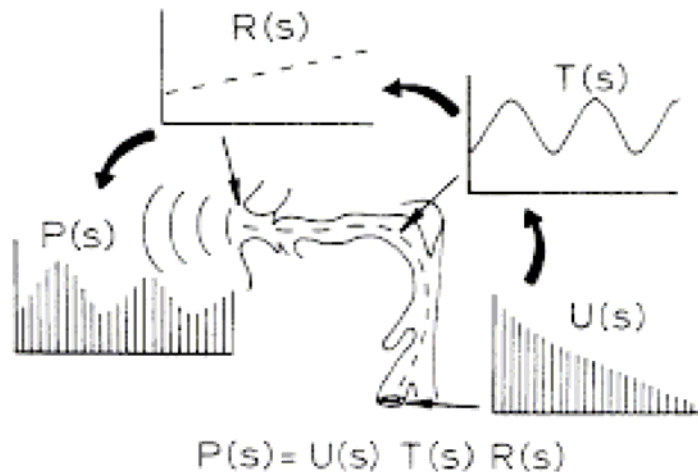


Intro: Acoustical Parameters

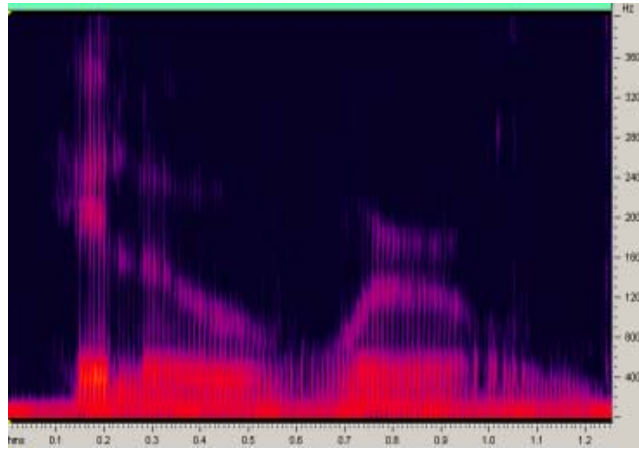
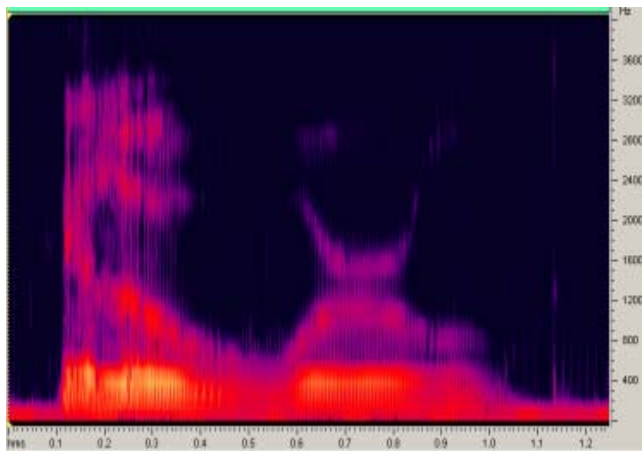
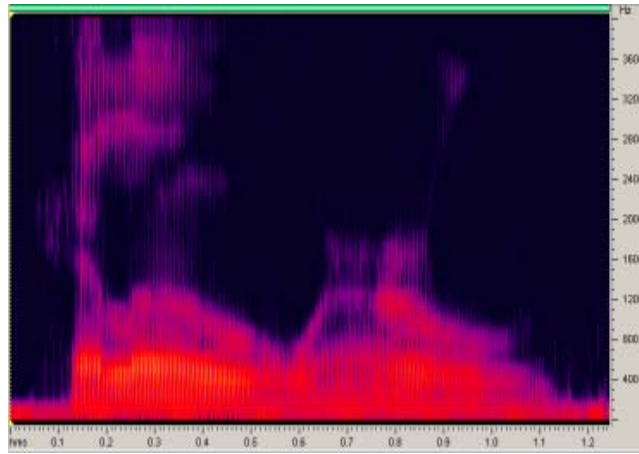
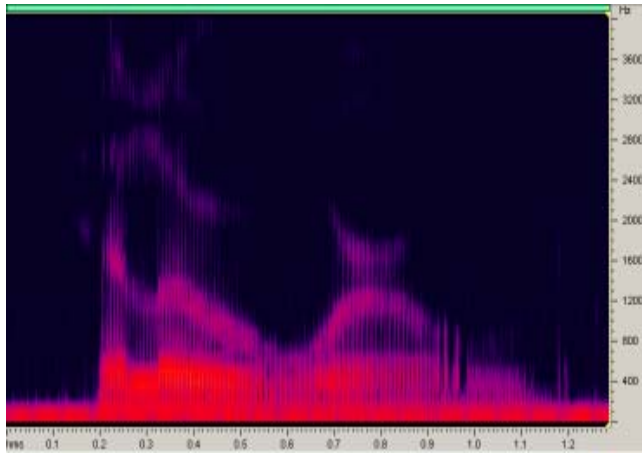
- Mean Fundamental Pitch Frequency
 - $\log \bar{F}_0$
- Fluctuation of Fundamental Pitch Period
 - $\sigma(\Delta T/\bar{T})$
- Slope of Glottal Source Spectrum
 - α
- Formant Frequencies
 - F_1, F_2, F_3

Glottal Source Characteristics - U(s)

Vocal Tract Characteristics - T(s)



Intro: ___ Recognition



Speech Recognition:
They are all the *same!*
→ “Hello World”

Speaker Recognition:
They are all the *different!*
→ s1, s2, s3, & s4

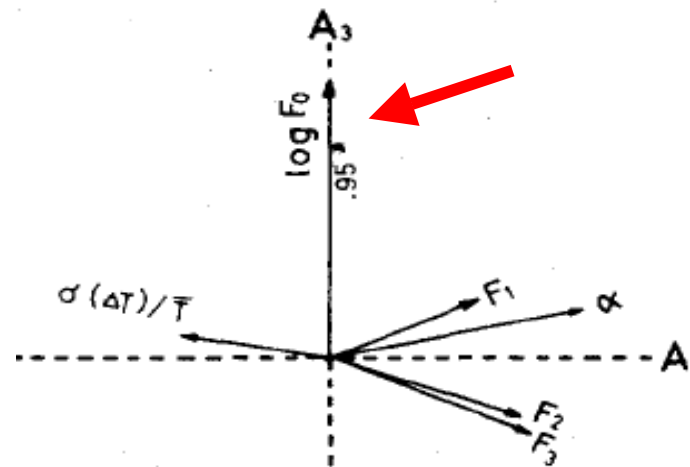
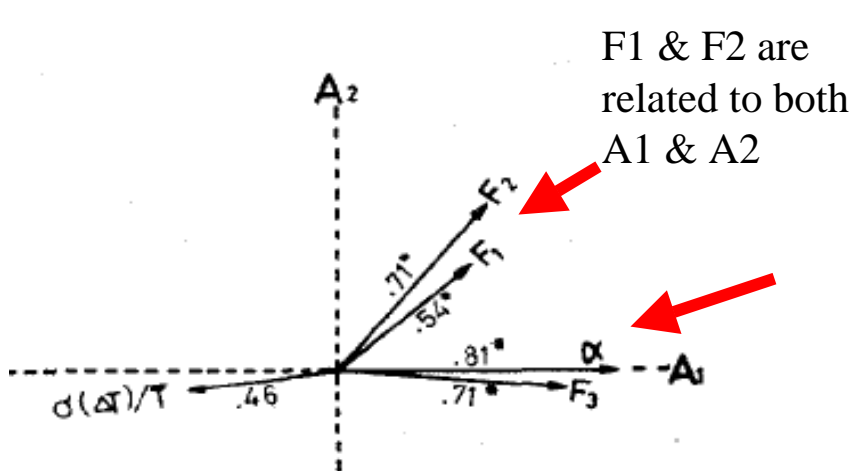
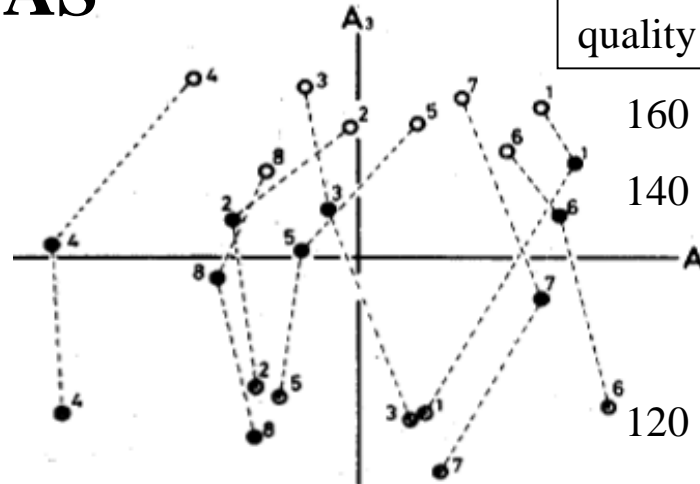
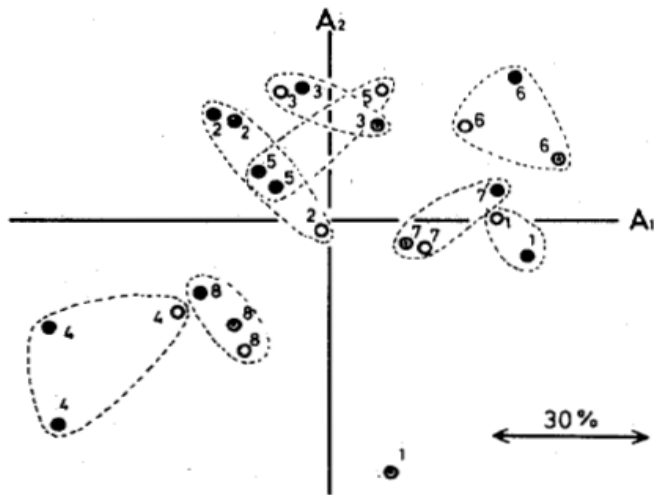
Test 1 - /a/: Specs

- Data Samples:
 - 8 speakers, vowel /a/ at 3 freq: 120, 140, & 160 Hz → 24 samples
- Listener Testing:
 - 6 listeners, listen 9 times to each pair twice (order) → 108 values/pair
 - Listeners classify voice pairs as “same talker” or “different talker”

Test 1 - /a/: Results

3D-PAS

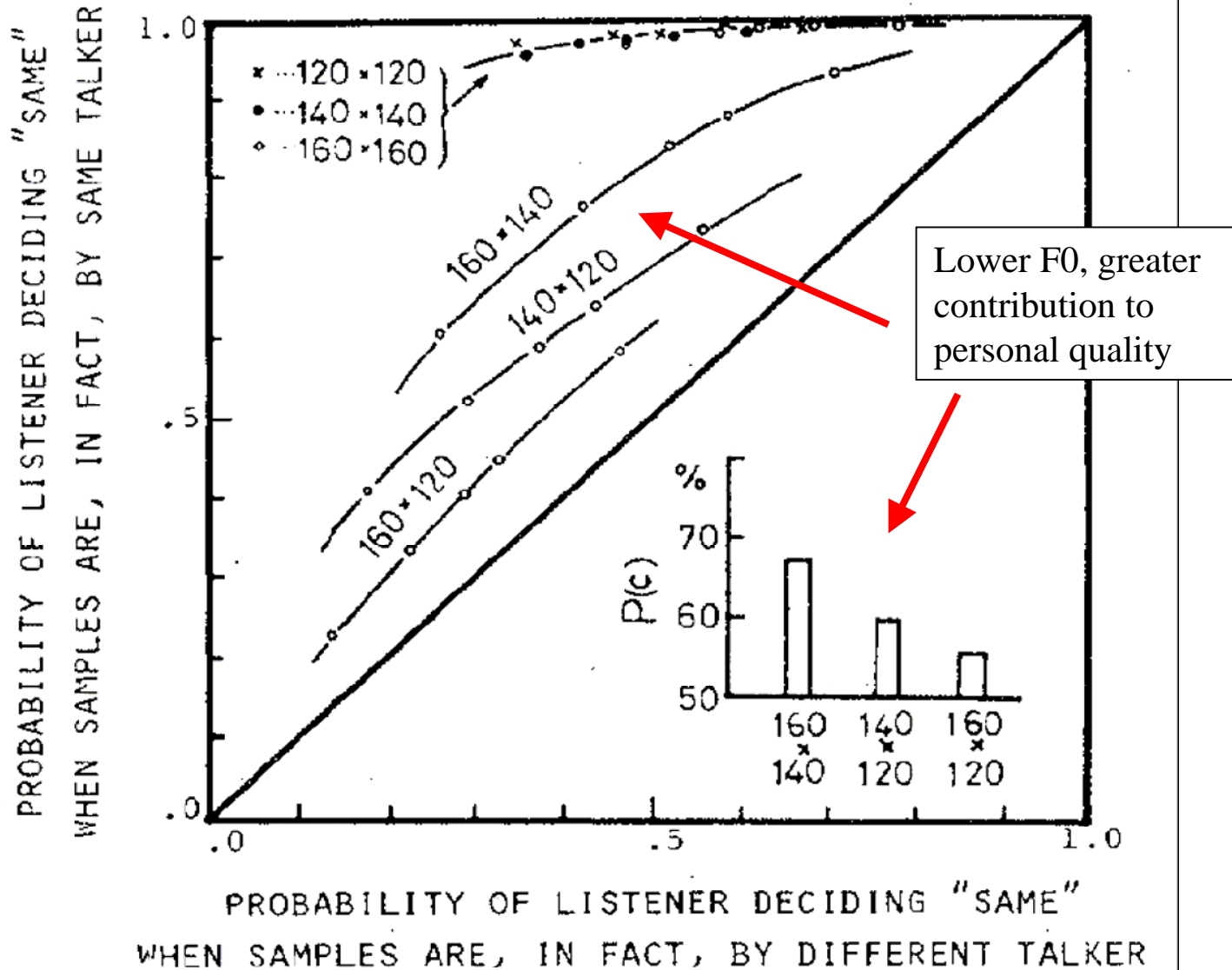
Lower F0, greater contribution to personal quality



Correlation between PAS and Acoustical Parameters

Test 1 - /a/: ROC

Receiver Operating Characteristics



Test 1 - /a/: Results (var)

TABLE I
Relation Between Various Sets of the Acoustical Parameters
and the Explained Variance of the Configuration of Voice Set
I on the PAS

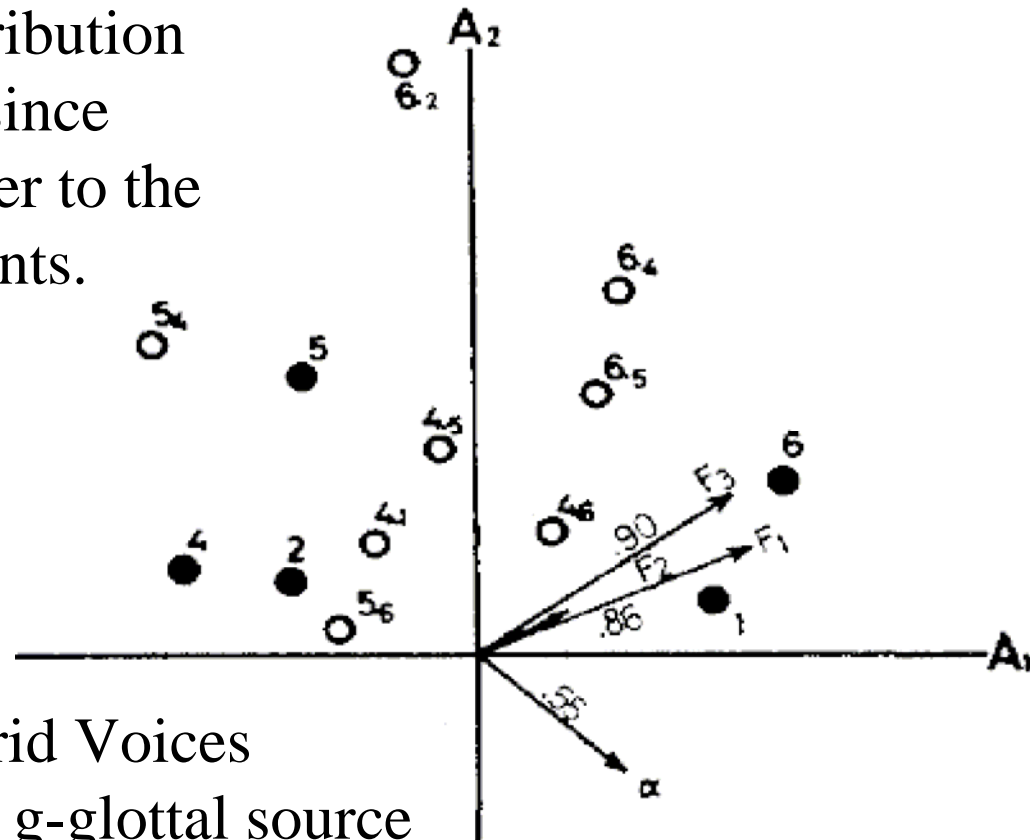
GLOTTAL SOURCE CHARACTERISTICS			VOCAL TRACT CHARACTERISTICS	EXPLAINED VARIANCE IN %
MEAN FUNDAMENTAL PITCH FREQUENCY	FLUCTUATION OF FUNDAMENTAL PITCH PERIOD	SLOPE OF GLOTTAL SOURCE SPECTRUM	FORMANT FREQUENCIES, F_1 , F_2 AND F_3	
X	X	X	X	86
X	X		X	84
X		X	X	84
X			X	81
X	X	X		71
X				55

Test 2 – Hybrid: Specs

- Data Samples:
 - 5 speakers, vowel /a/ at 140 Hz.
 - Data set altered by generating fixed glottal source (removing fluctuation of fundamental pitch period variable)
- 6 listeners repeat 10 trials each pair

Test 2 – Hybrid: Results

- F3 became similar to F1 & F2
- Vocal Tract has greater contribution than Glottal (other than F0) since hybrid voices tend to be closer to the original with the same formants.



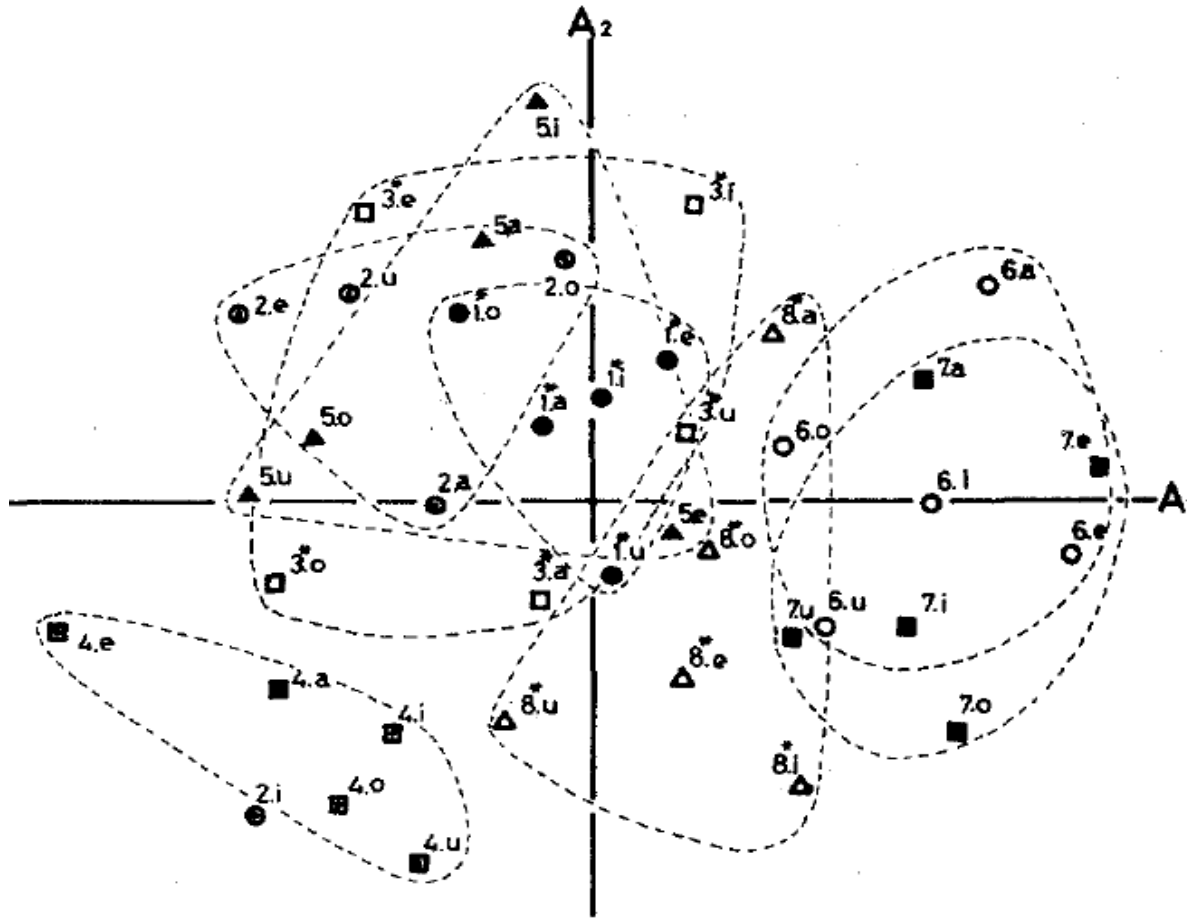
2D PAS of Hybrid Voices

V_g : V-Formant, g-glottal source

Test 3 – Vowels: Specs

- Data Samples
 - 8 Speakers, 5 vowels (40 Voices) all at 164 Hz
- Listeners
 - 13 people listened 3 times to all voice pairs
 - (78 Samples)

Test 3 – Vowels: Results (PAS)

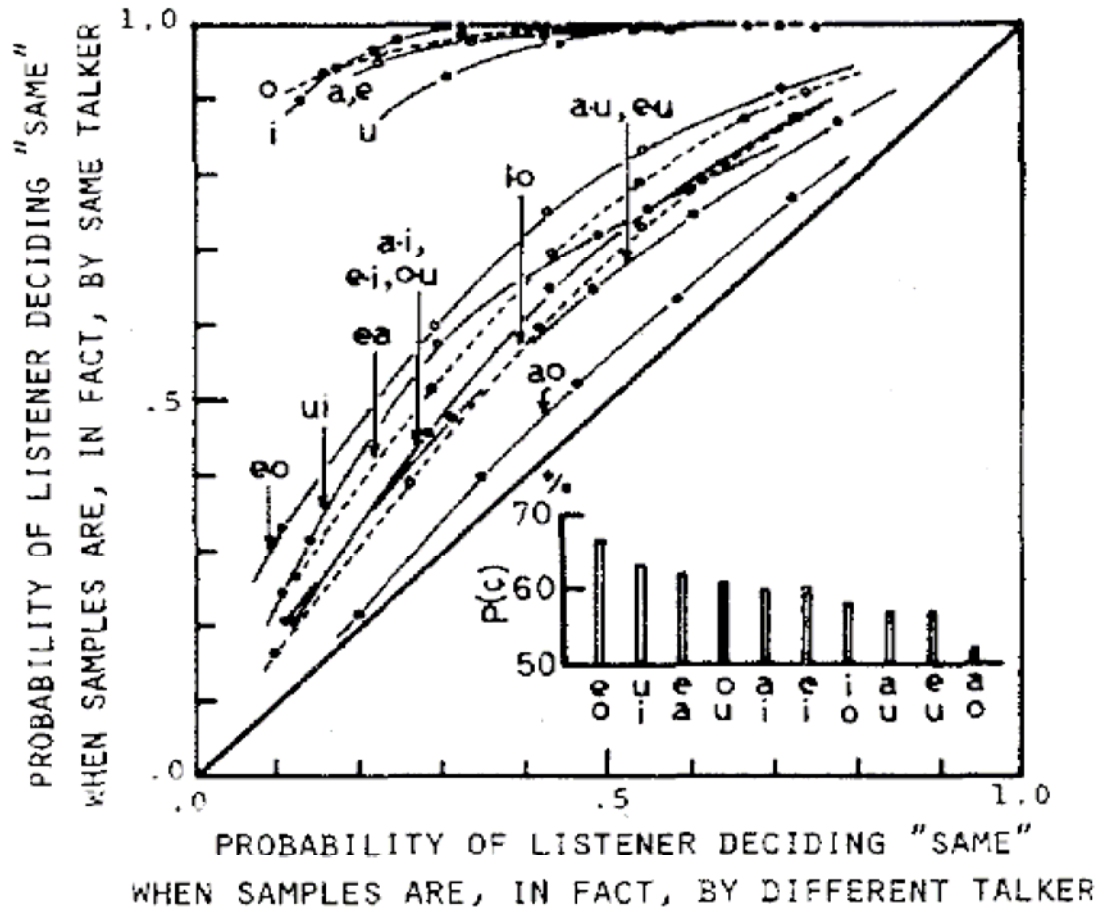


Since Talkers are clustered,

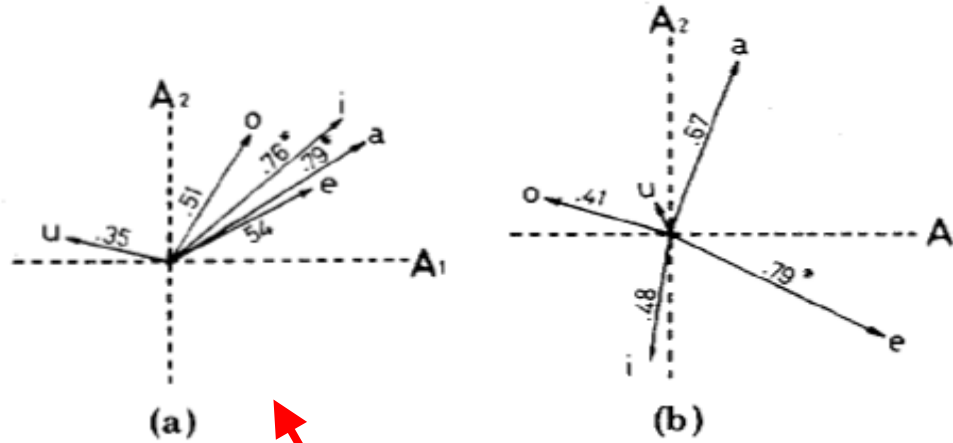
The perceptual cues of personal quality
common to different vowels is involved in
listener judgment

Test 3 – Vowels: Results (ROC)

Receiver Operating Characteristics



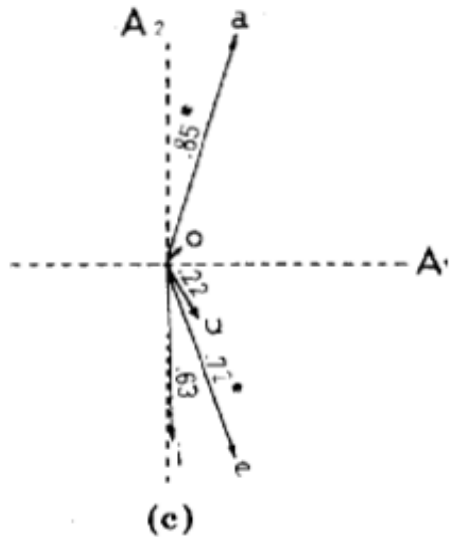
Test 3 – Vowels: Results (xcorr)



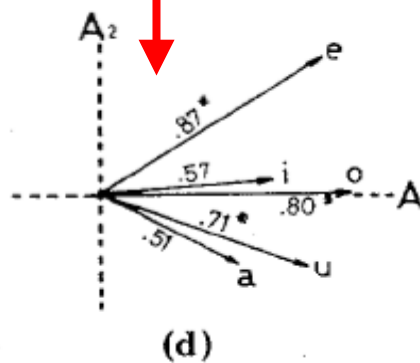
F1

F2

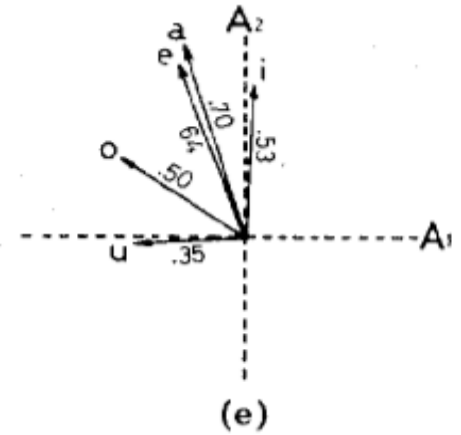
Large Correlations
and similar directions



F3



α : Slope of glottal
source spectrum



$\sigma(\Delta T/T)$: Rapid
fluctuation of pitch period

Test 3 – Vowels: Results (var)

TABLE II
Relation Between Various Sets of the Acoustical Parameters
and the Explained Variance of the Configuration of Voice Set
II on the PAS

GLOTTAL SOURCE CHARACTERISTICS		VOCAL TRACT CHARACTERISTICS	EXPLAINED VARIANCE IN %				
FLUCTUATION OF FUNDAMENTAL PITCH PERIOD	SLOPE OF GLOTTAL SOURCE SPECTRUM	FORMANT FREQUENCIES, F_1 , F_2 AND F_3	/i/	/e/	/a/	/o/	/u/
X	X	X	97	99	97	93	94
X	X		80	80	79	88	80
		X	71	80	85	62	45

Conclusions (1)

- F0 is the relative most significant contributor to perception of personal quality
- Vocal Tract and Glottal Characteristics contribute to different perceptual dimensions from each other with F0 constant
- Vocal Tract contributions to perception of personal quality varies with different vowels
- The perceptual dimensions of F0, F1, α -slope of glottal, and fluctuation of F0 period are independent of vowel

Conclusions (2)

- Authors claim success because...
 - Talkers cluster on the A1-A2 PAS
 - The P(c) from the listeners was about 60-70%
 - There is uniformity of the results despite different listeners
 - Acoustical parameters were found to influence perception of personal quality
- Future Work:
 - Evaluate other Parameters