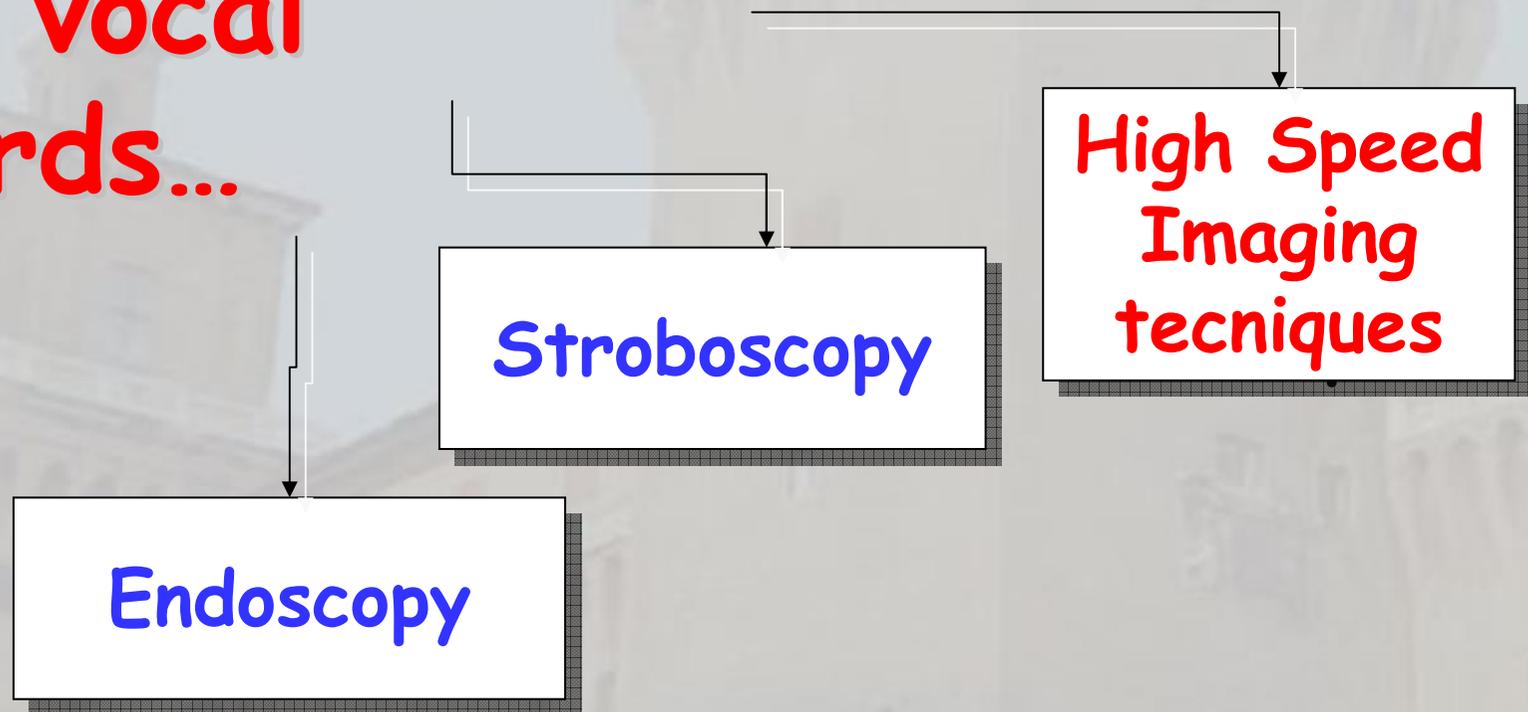


# DIGITAL VIDEOKYMOGRAPHY AND HIGH SPEED CAMERA

F. STOMEIO

*ORL FERRARA*

# Diagnostic of vocal cords...

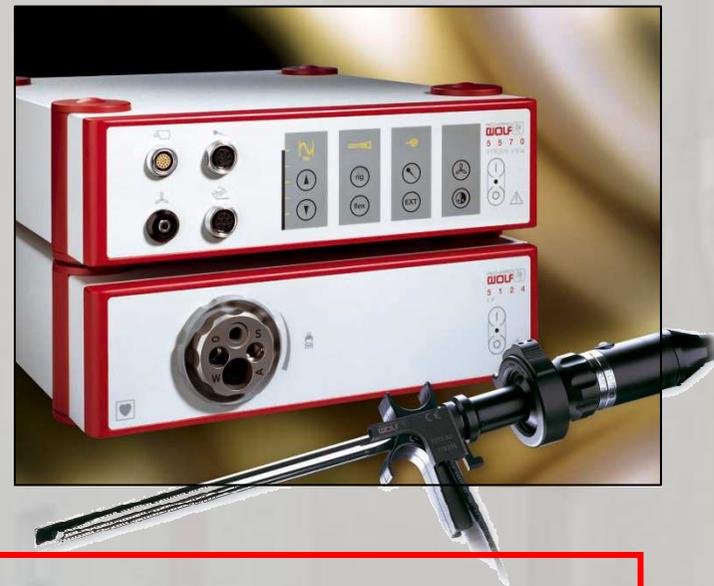
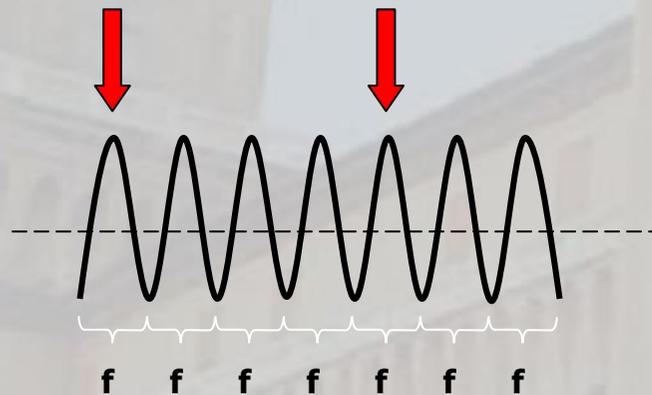


	<b>Telescope</b>	<b>Video-Stroboscope</b>	<b>High Speed Imaging</b>
<b>Morphological mutation</b>	X	X	X
<b>Vibration of the vocal cords</b>	---	X	XX
<b>Irregular vibration of the vocal cords</b>	---	---	X
<b>Phonation onset</b>	---	---	X

**ORL FERRARA**

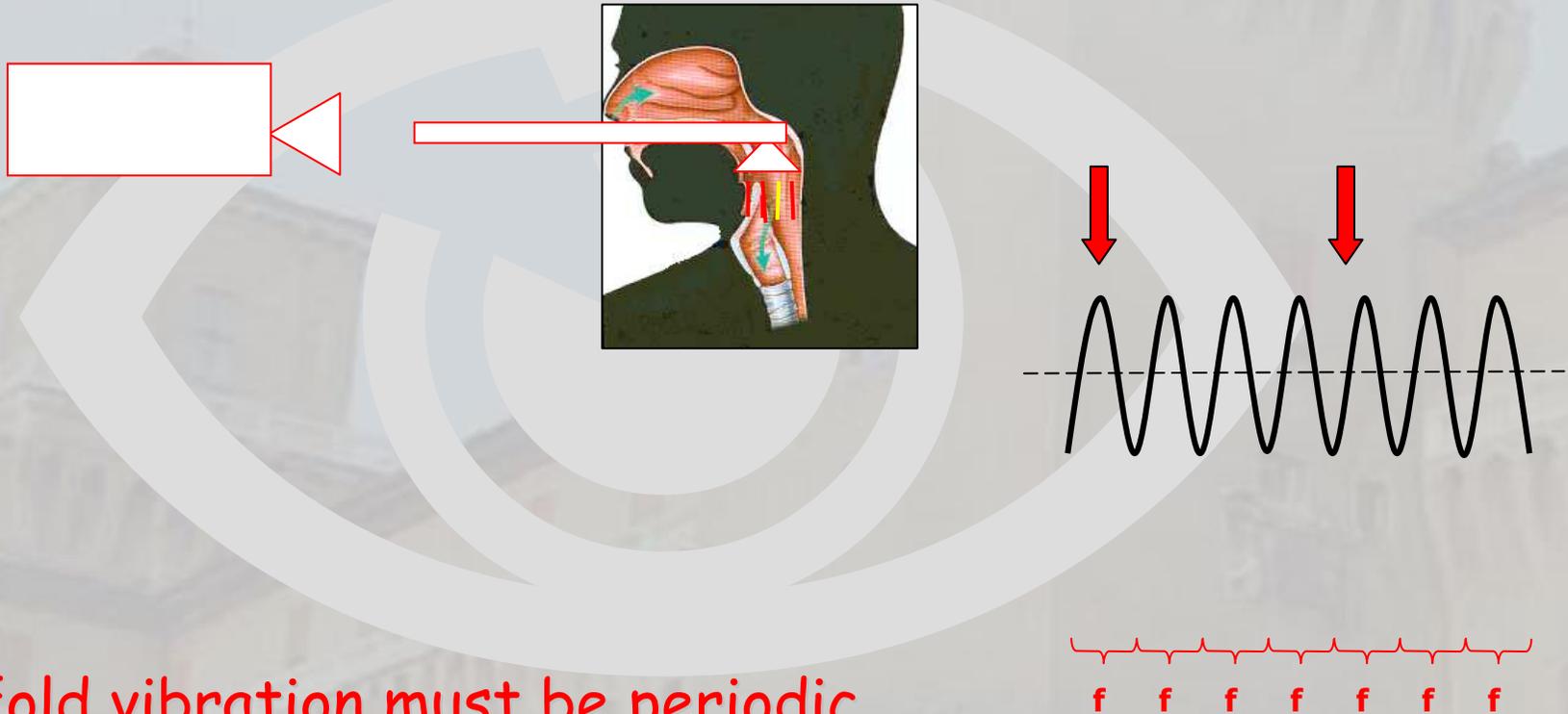
# VIDEOSTROBOSCOPY

- Needs a microphone
- Needs "Trigger" signal



Every frequency disturbance of the vibration disturbs as well the resulting stroboscopic image

# Videostroboscopy: needs a periodic signal



- vocal fold vibration must be periodic
- you can't appreciate vocal fold swinging if signal is aperiodic
- hardly appraised vocal fold with different frequencies (left/right asymmetries)

*ORL FERRARA*

*Why do we need technical aids to make the swinging of the vocal cords visual?*



*It's important to have a laryngoscopic method not limited to periodic vibration*

*ORL FERRARA*

*For normal acquisition video system  
phonation is the problem.....*

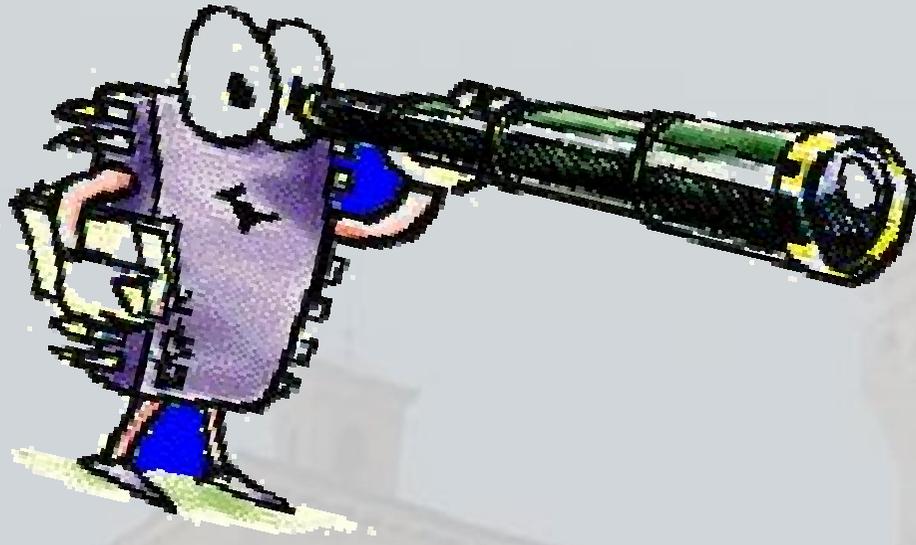


**20 pictures/sec**

**Frequency of human phonation  
100 - 800 Hz**



**ORL FERRARA**



According to Shannon theorem an acquisition (video) system to analyze correctly irregularity of swinging of the vocal cords needs acquisition speeds of 1000 - 4000 frames per second

*ORL FERRARA*

# High Speed Imaging

- Videochimography (VKG)
- High-Speed camera

*ORL FERRARA*

# Videokymography (VKG)

- This idea was brought to Kay by Harm Schutte, Francis Sram and Jan Svec.
- Given its unique capabilities, VKG is the ideal complement to stroboscopy, because stroboscopy needs quasi-periodic vibration to lock to voice, and cannot see voicing initiation.
- Although the full screen display of the VKG "image", constituted of single lines is not as intuitive as stroboscopy, VKG does allow direct viewing of vocal fold behaviours which may not be observable with a stroboscopic image.

*ORL FERRARA*

# WHICH BEHAVIOURS ?

Videokymographic images reflect important properties of vocal fold vibration including:

- open and closed phases of the glottal cycle,
- displacements of the upper and lower vocal fold margins,
- propagation of mucosal waves.

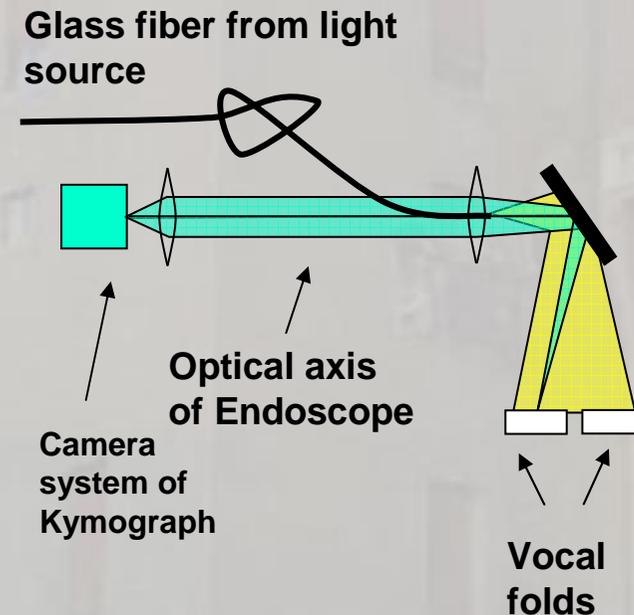
The high scan rate of VKG allows the direct observation of vocal fold motion, even if the motion is aperiodic.

- Thus, voicing initiation, diplophonia, biphonia, vocal fry, creaky voice, and aperiodicity can all be viewed directly.
- Even in normal quasi-periodic phonation, vocal asymmetry and mucosal waves are clearly visible with this powerful technique.

*ORL FERRARA*

# WHAT VKG IS ?

- A modified video camera, coupled to a standard rigid endoscope and a constant light source constitute VKG, that is capable of capturing high-speed motion such as vocal fold vibration.
- The VKG camera scans a single line at 8000 frames/second, according this way to Shannon theory; the recording allows review and analysis



*ORL FERRARA*

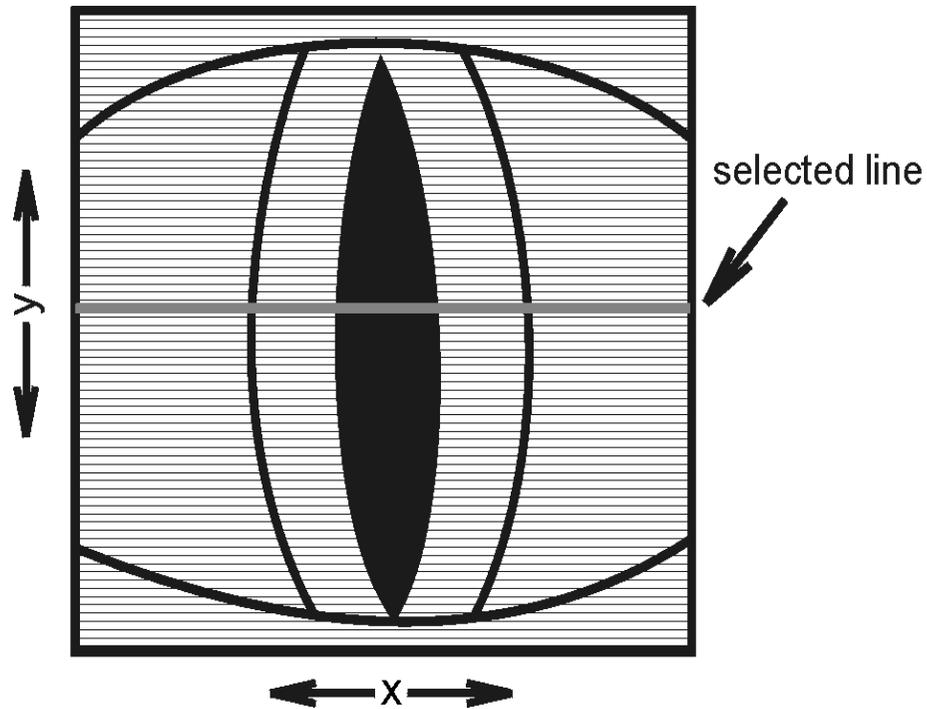
HOW DOES  
CLASSIC VKG  
WORK ?

IN TWO SEPARATE MODES

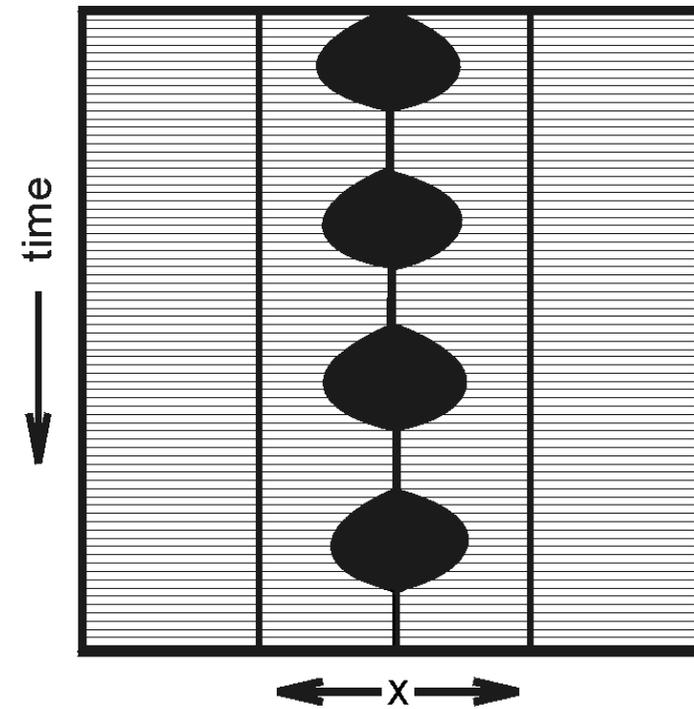
*ORL FERRARA*

# VIDEOKYMOGRAPHY (VKG)

STANDARD MODE

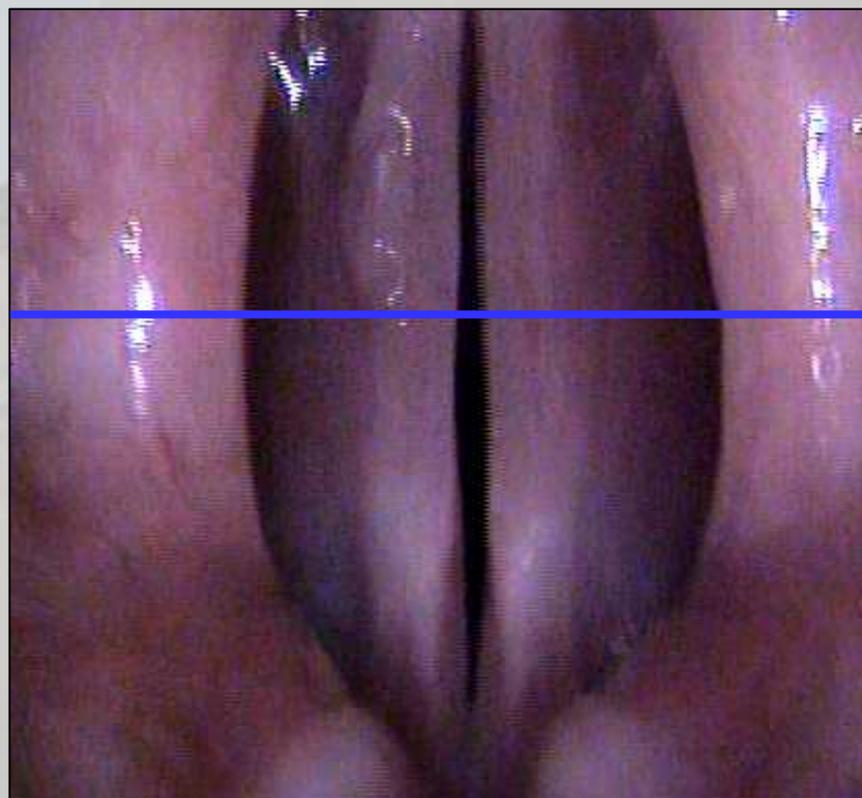
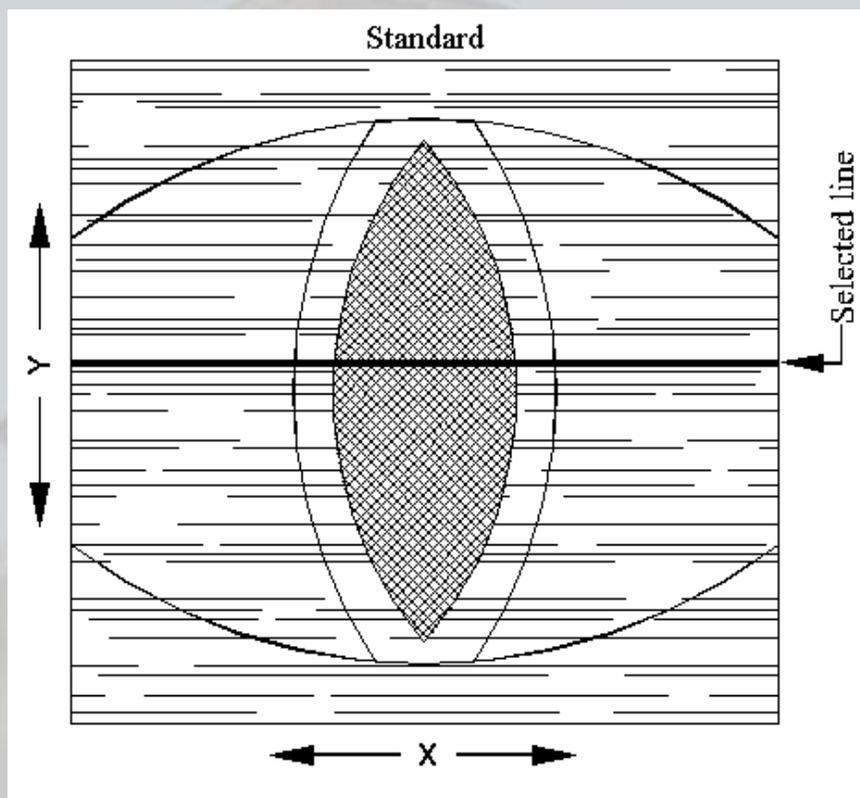


HIGH-SPEED MODE



Svec & Schutte, J.Voice 10(2): 201-205 (1996)

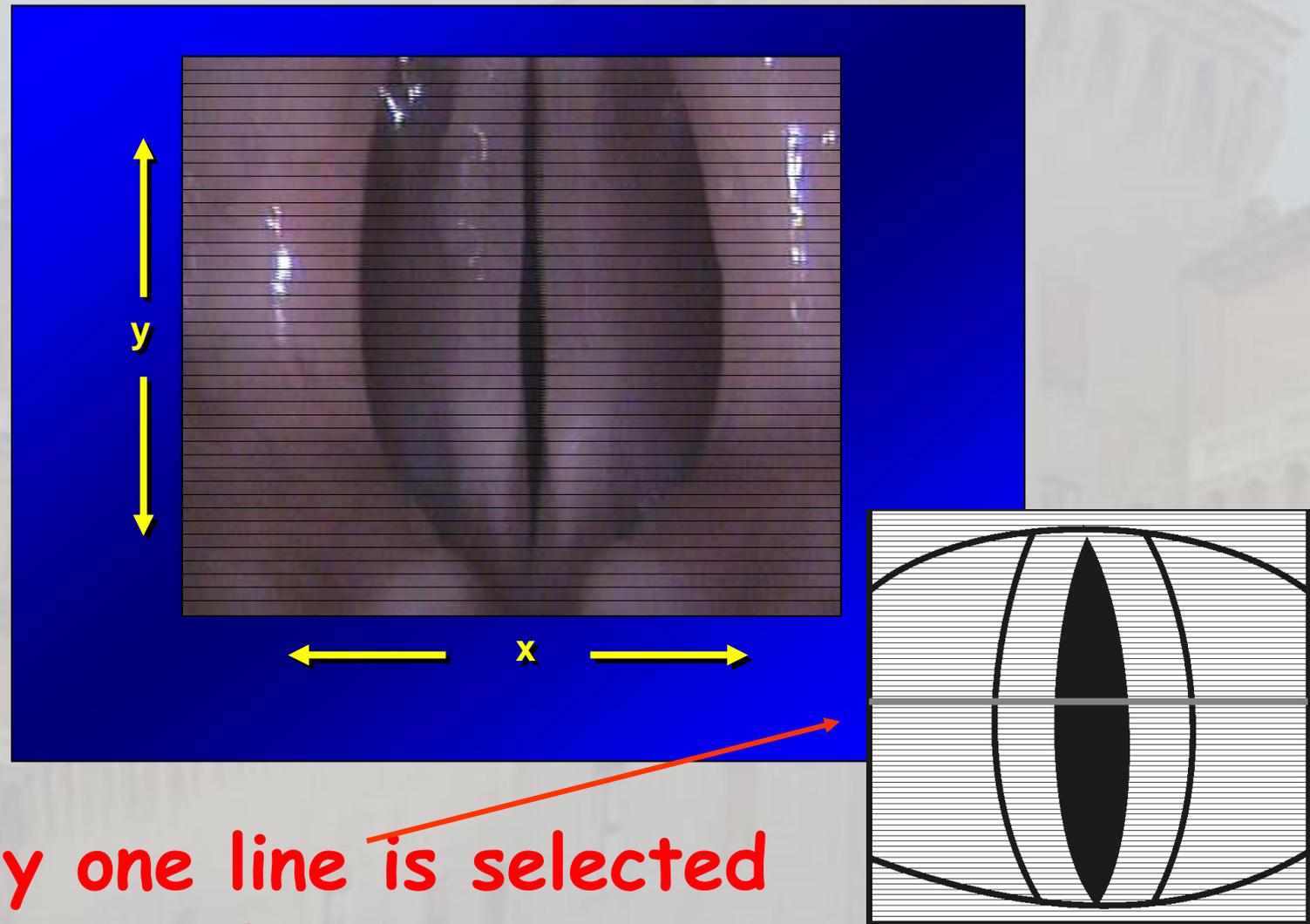
- **STANDARD MODE VISUALIZES THE WHOLE GLOTTIS.**
- **IMAGE IS MADE BY 262 LINES**



**ORL FERRARA**

**ACTIVE LINE**

One single horizontal line (active line) is positioned anywhere along the glottic length (transverse to it). This scan line serves at the point where the glottic waveform is sampled. This technique allows for a single point display of successive glottic cycles.



.....only one line is selected for the examination

ORL FERRARA

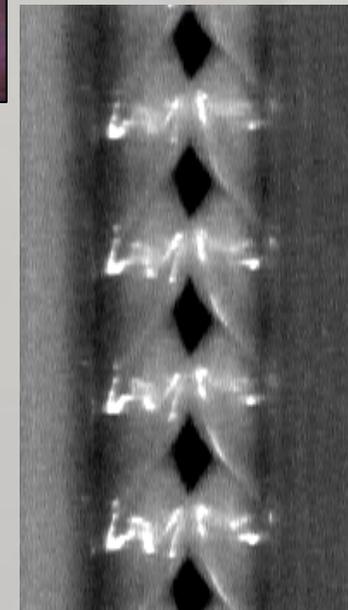
- HIGH-SPEED  
MODE ALLOWS  
THAT THE  
SELECTED LINE  
CAN BE  
VISUALIZED  
8000 TIMES  
FOR SECOND  
- IN THIS WAY  
ALL THE  
FREQUENCIES  
OF THE HUMAN  
VOICE CAN BE  
EVALUATED



← *Active Line*

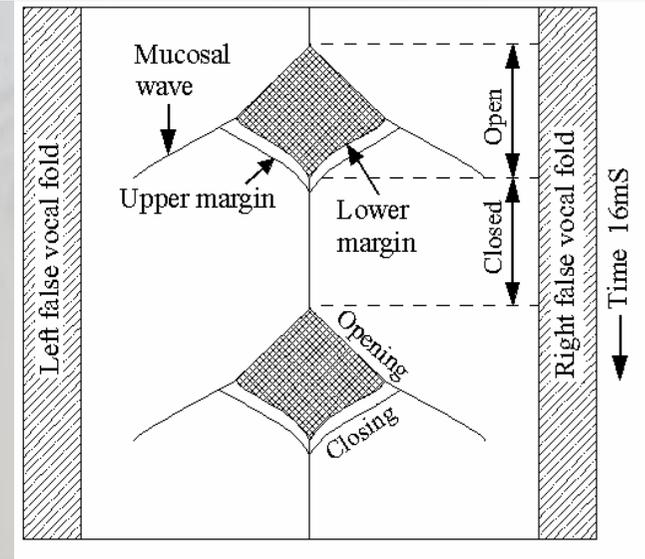
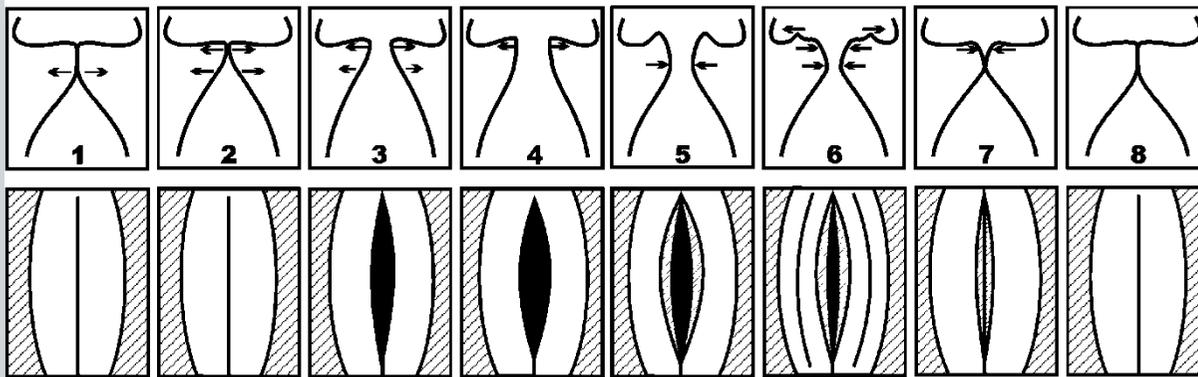
- IMAGES OF THE SELECTED LINE ARE PRESENTED IN REAL TIME
- Y AXIS IS TIME DIMENSION (successive line images of vocal fold vibration produced from this single glottic position are presented along the vertical "y" axis).

TEMPO

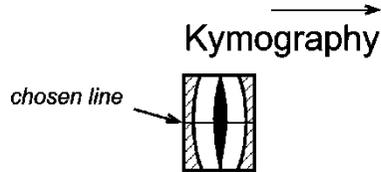


*ORL FERRARA*

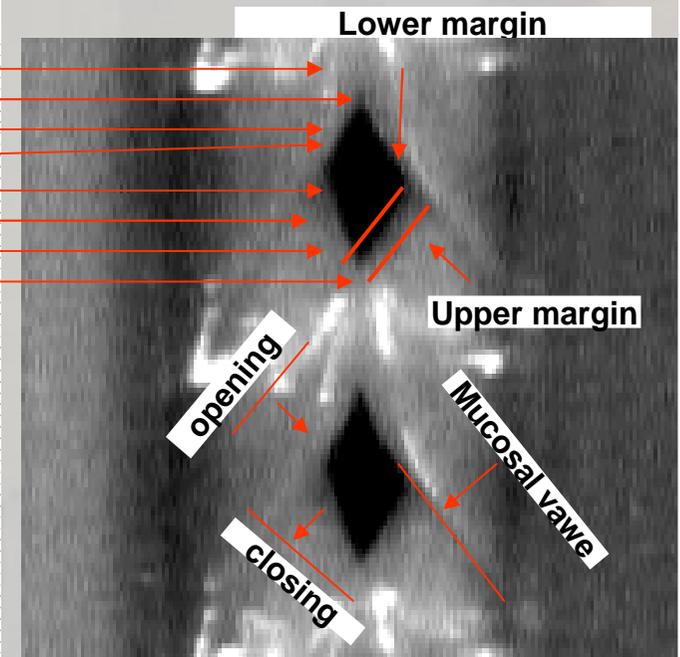
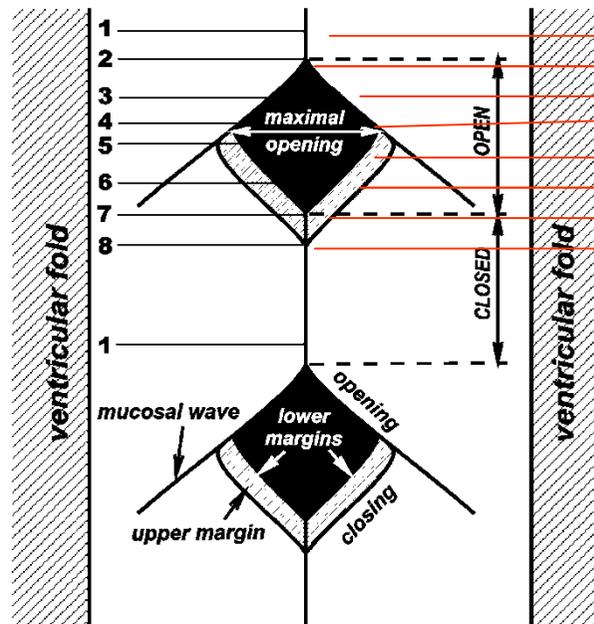
### Frontal section of the vocal folds during a phonatory cycle



### Laryngostroboscopy



- 1 - lower part of glottis starts to open
- 2 - upper part of glottis starts to open
- 3 - lower and upper parts of glottis open
- 4 - lower part of glottis is maximally open, upper part of glottis still opens
- 5 - lower part of glottis closes and is visible, upper part of glottis is maximally open
- 6 - lower and upper parts of glottis close, mucosal wave propagates on the surface
- 7 - lower part of glottis is closed
- 8 - upper part of glottis is closed

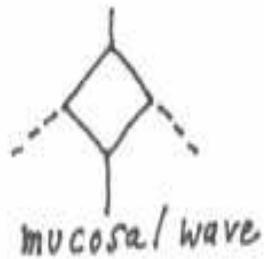


# What can be seen and evaluated in the kymograms?

- Mucosal Wave
- Amplitude
- Phase Closure
- Phase Symmetry
- Vibratory Behavior
- Periodicity
- Vertical Approximation
- Glottic Configurations

*ORL FERRARA*

# Overview of kymogram waveform: kimographic patterns

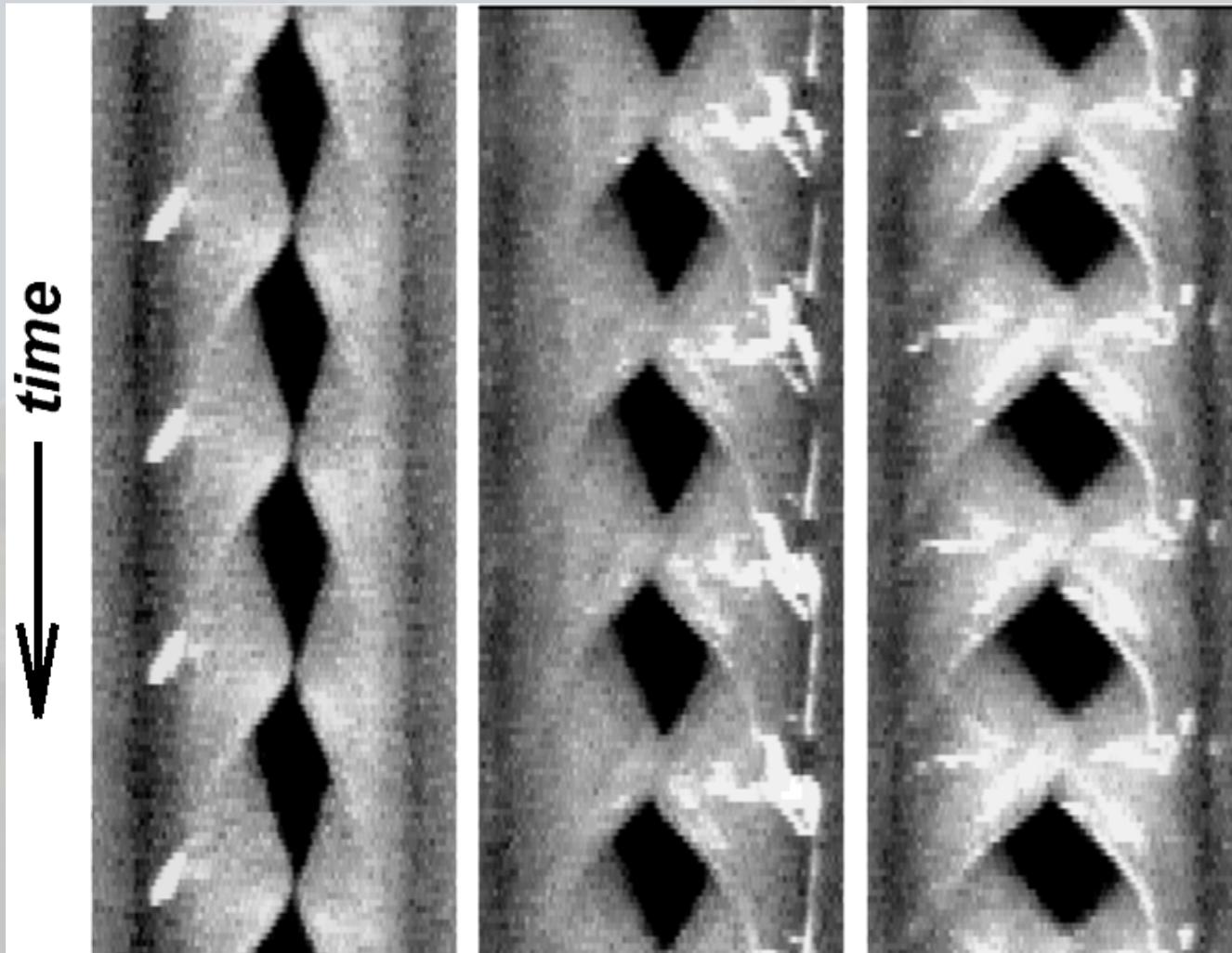


# Factors to be considered for the evaluation of the kymograms : vocal loudness

SOFT

MIDDLE

LOUD

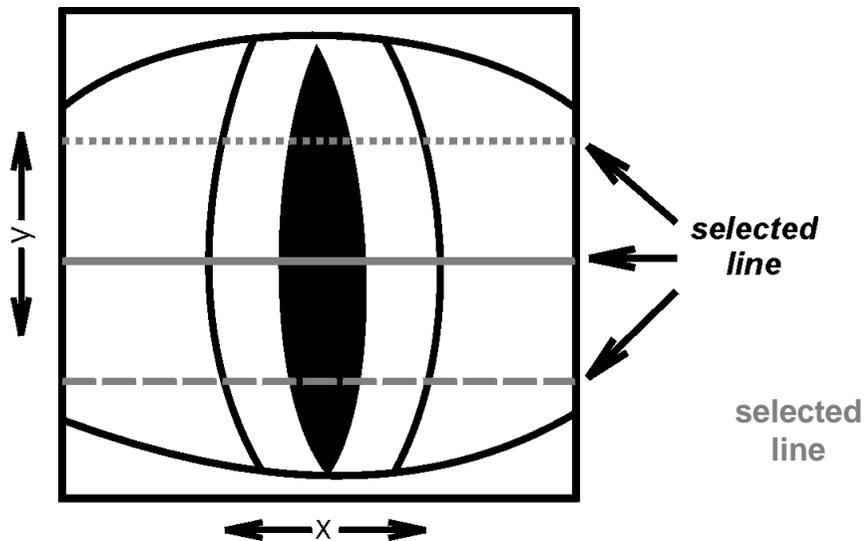


Measurement  
place:  
middle of the  
membraneous  
vocal fold

**ORL FERRARA**

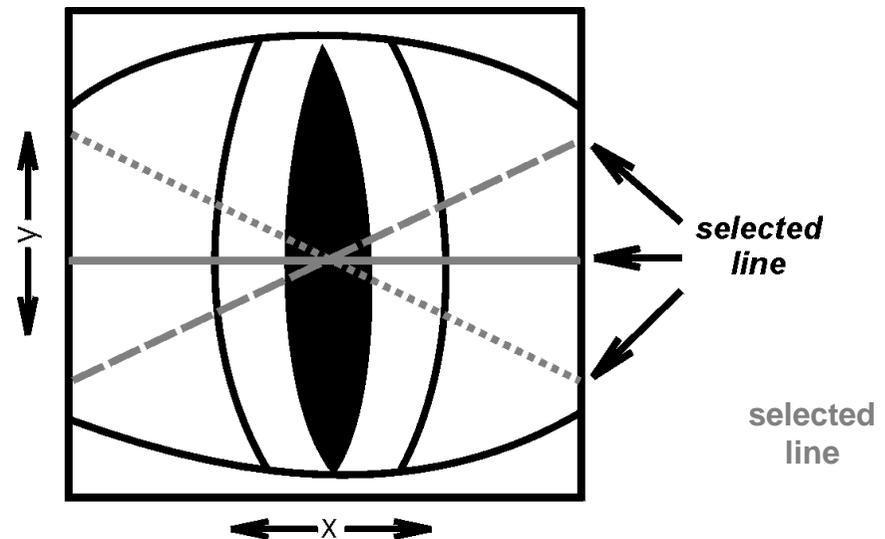
# Factors to be considered for the evaluation of the kymograms : positioning the measurement line

a) Along the glottis (anterior - middle - posterior):

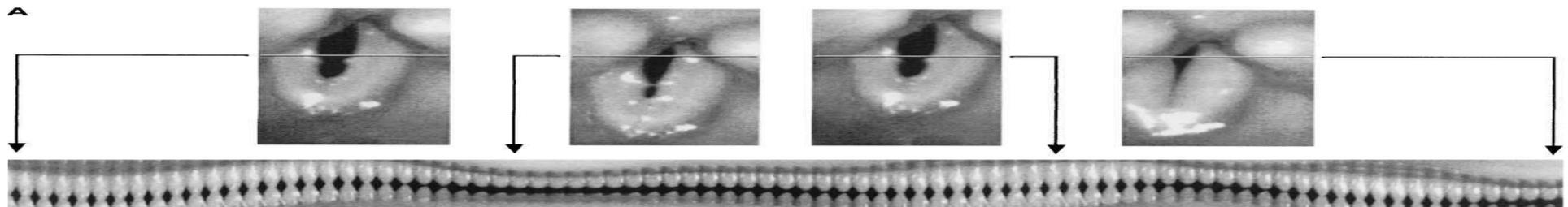


Can be adjusted by positioning and tilting the scope

b) Angle with respect to glottis (perpendicular – oblique):



Can be adjusted by rotating the camera head on the endoscope



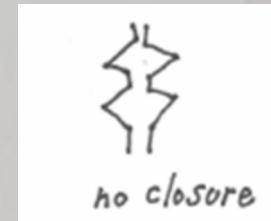
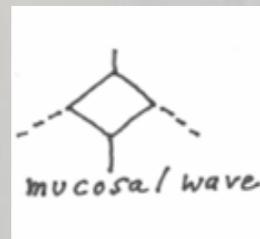
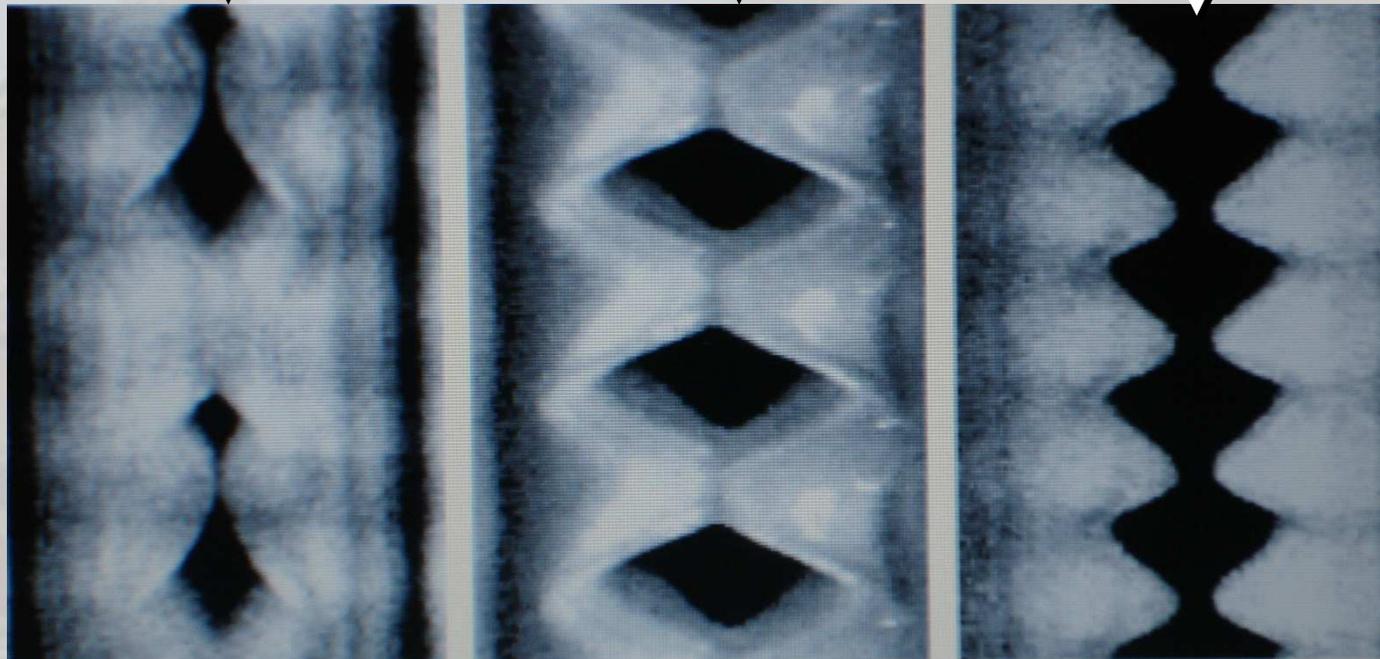
# Factors to be considered for the evaluation of the kymograms: voice registers

VOCAL FRY

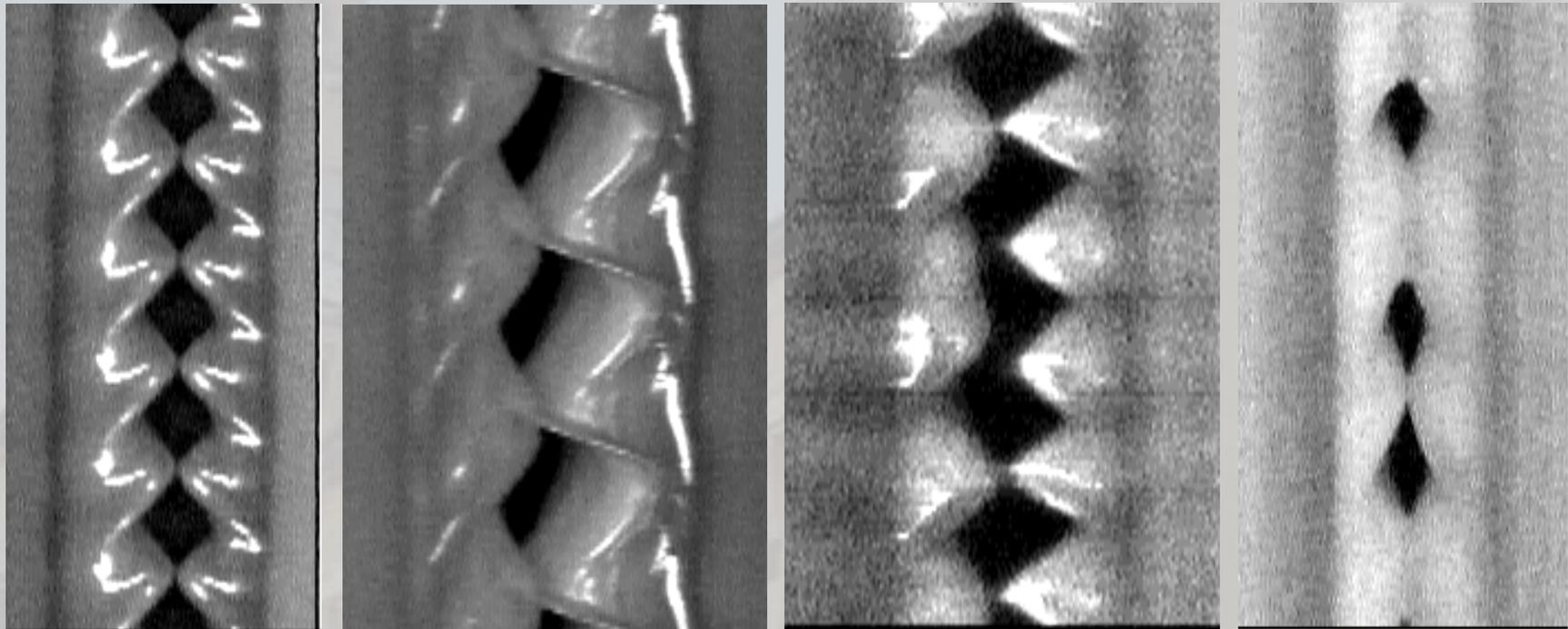
NORMAL VOICE

FALSETTO

T  
E  
M  
P  
O



# KIMOGRAMS EVALUATION : DESYNCHRONIZATION AND PHASE DIFFERENCES

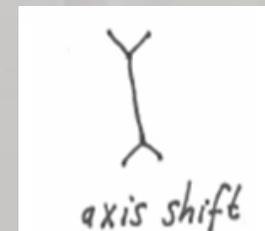
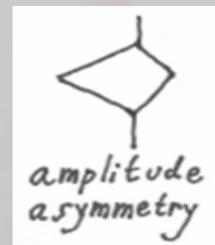
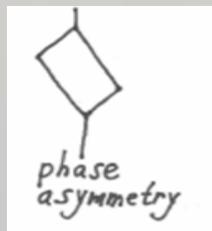
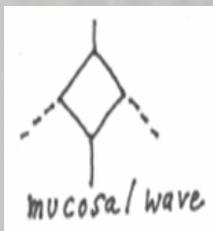
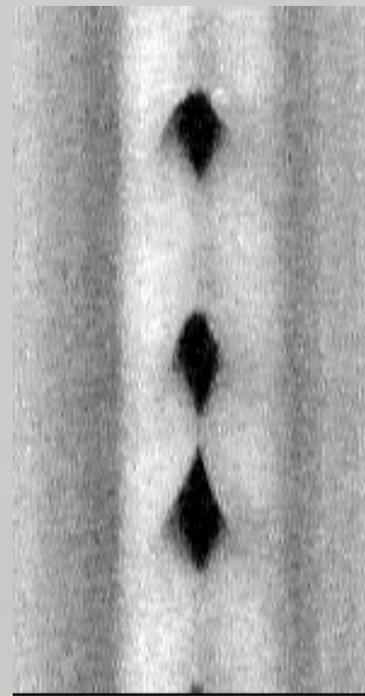
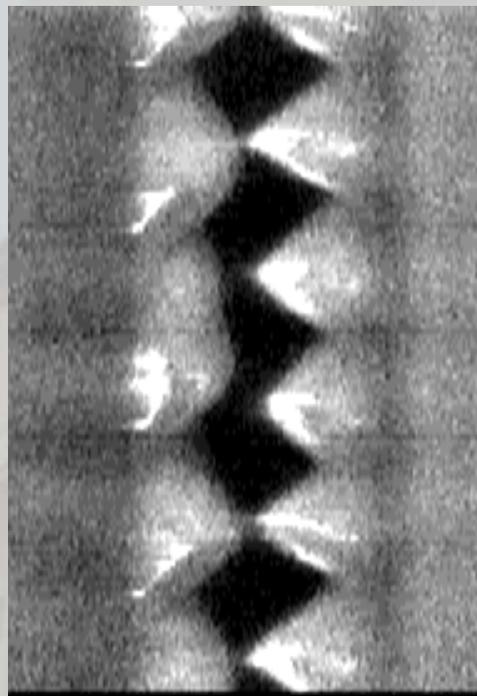
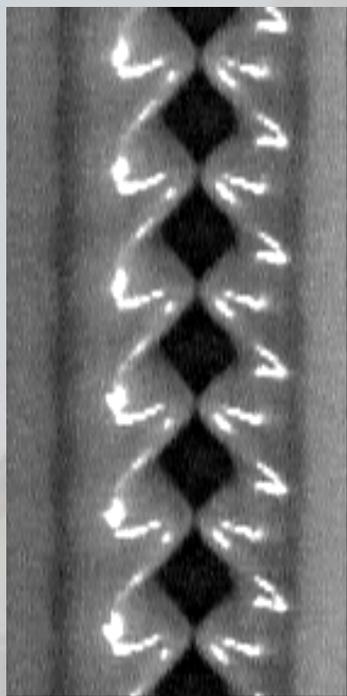


**ORL FERRARA**

*From Svec, Sram & Schutte: Videokymography Workshop, The Voice Foundation Symposium Philadelphia 2002*

# KIMOGRAMS EVALUATION: DESYNCHRONIZATION AND PHASE DIFFERENCES

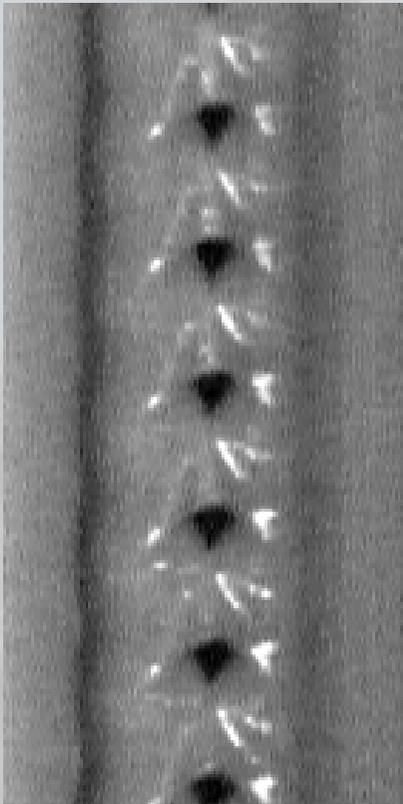
Da Svec, Sram & Schutte: Videokymography Workshop, The Voice Foundation Symposium Philadelphia 2002



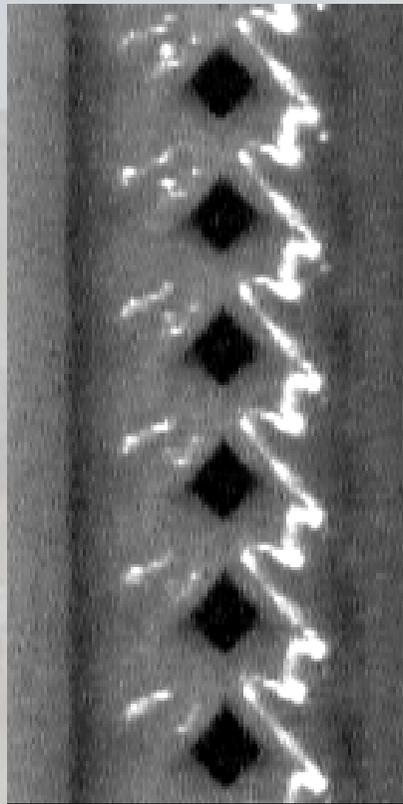
**ORL FERRARA**

# KYMOGRAMS EVALUATION : CLOSED QUOTIENT

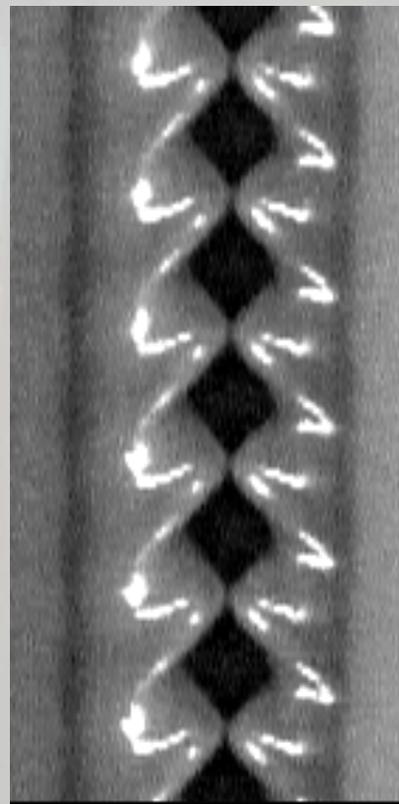
CQ > 50%



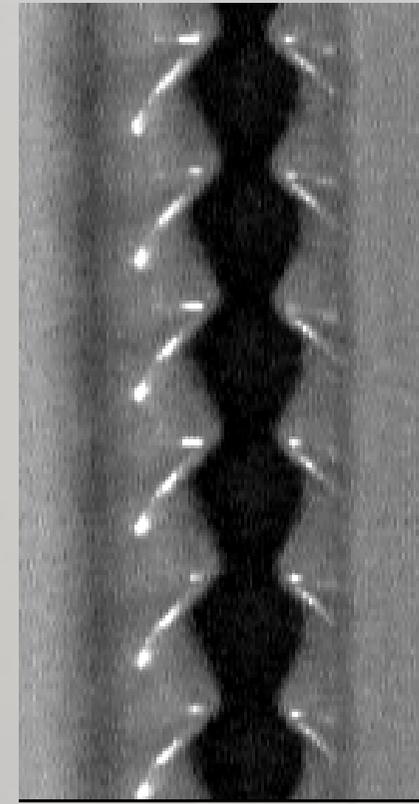
CQ ~ 40%



CQ ~ 10%



CQ = 0%

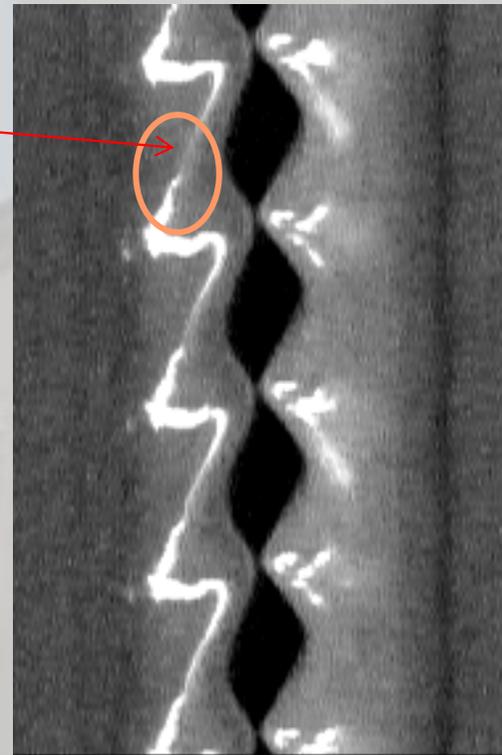
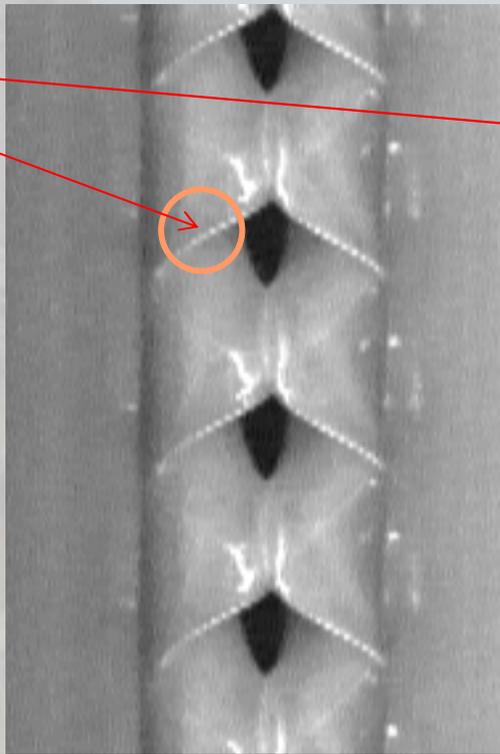
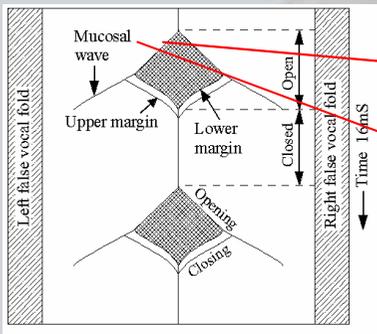


**ORL FERRARA**

# KYMOGRAMS EVALUATION: OPEN PHASE FEATURES - OPENING-TO-CLOSING TURN

Type 1: „abrupt“

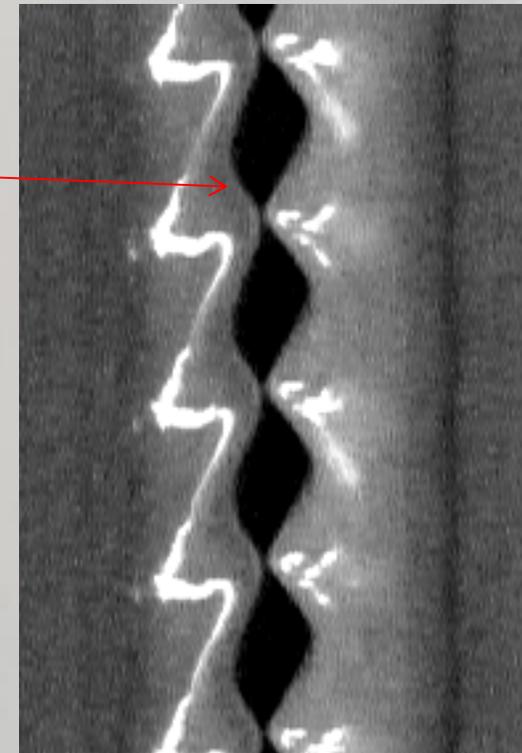
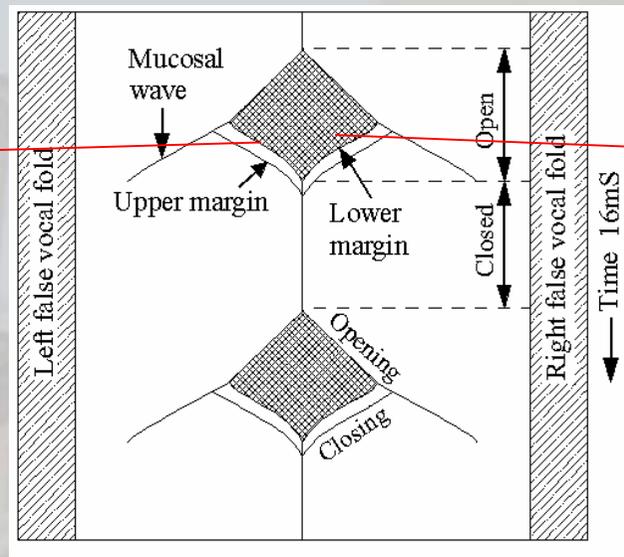
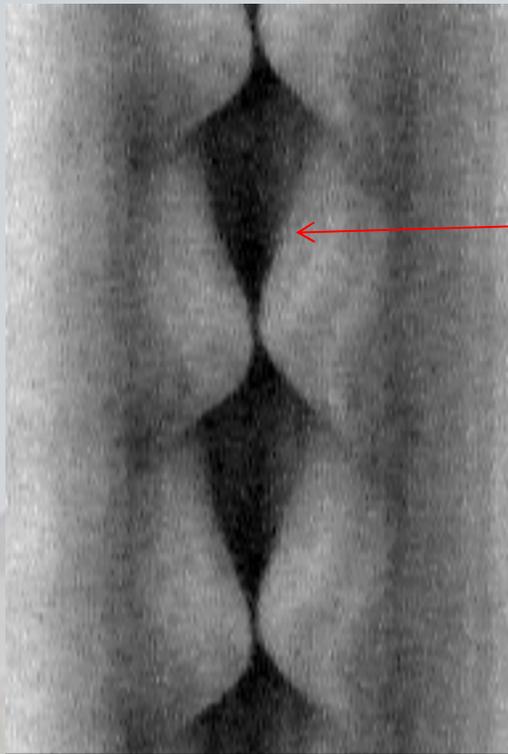
Type 2: „smooth“



Other types:  
disturbed,  
variable...

**ORL FERRARA**

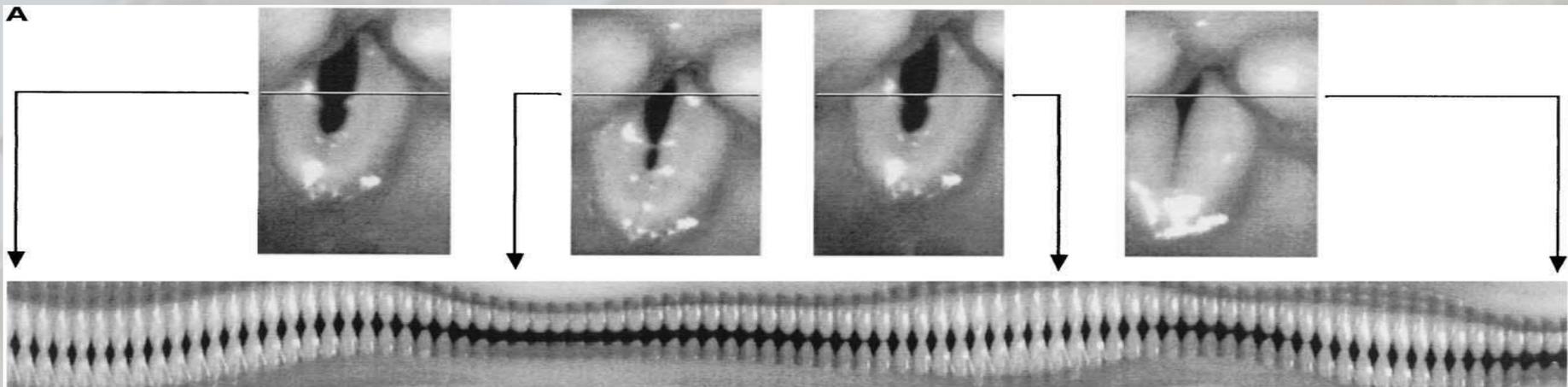
# KYMOGRAMS EVALUATION: LOWER MARGIN VISIBILITY



**ORL FERRARA**

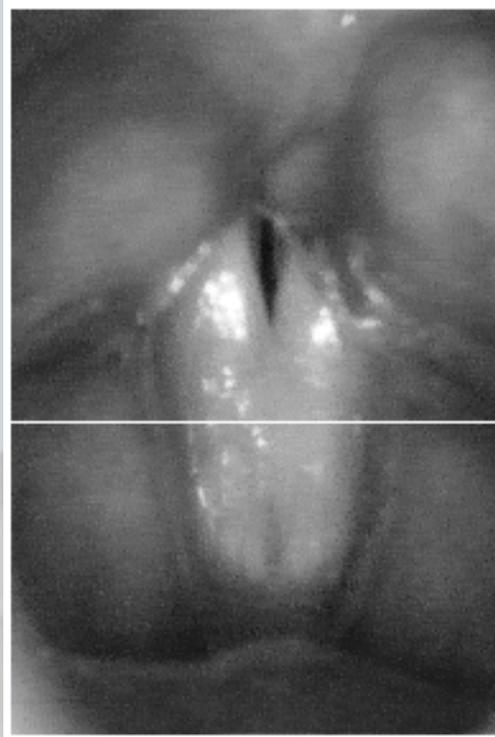
# CLASSIC VKG

- It's impossible to evaluate two different lines in the same recording
- It's difficult to keep constant the angle between the tip of the endoscope and the glottis.

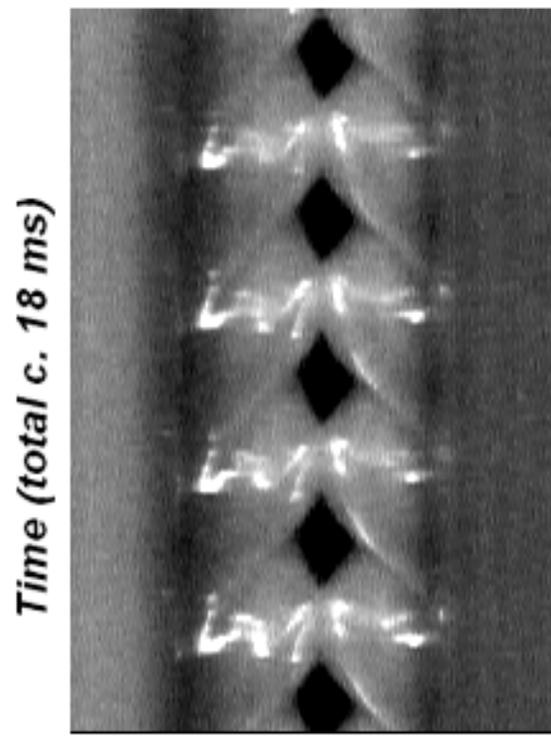


# CLASSIC VKG

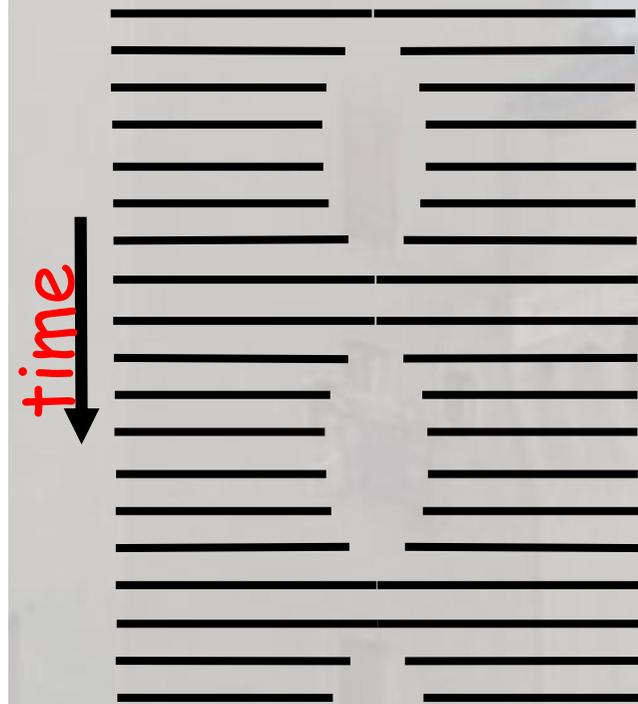
STANDARD mode



HIGH-SPEED (VKG)  
mode



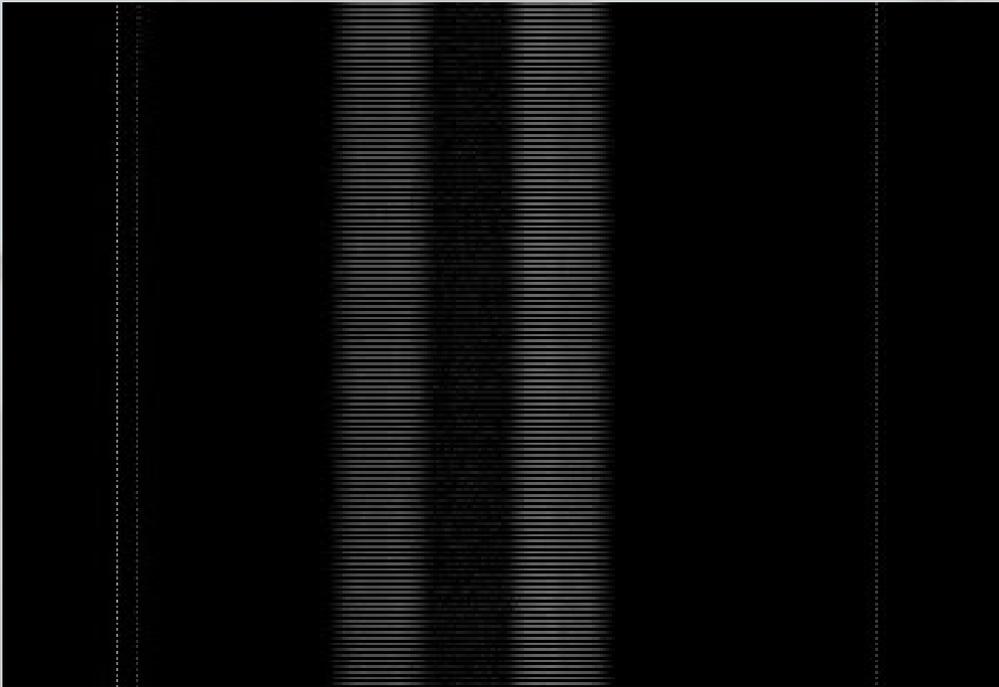
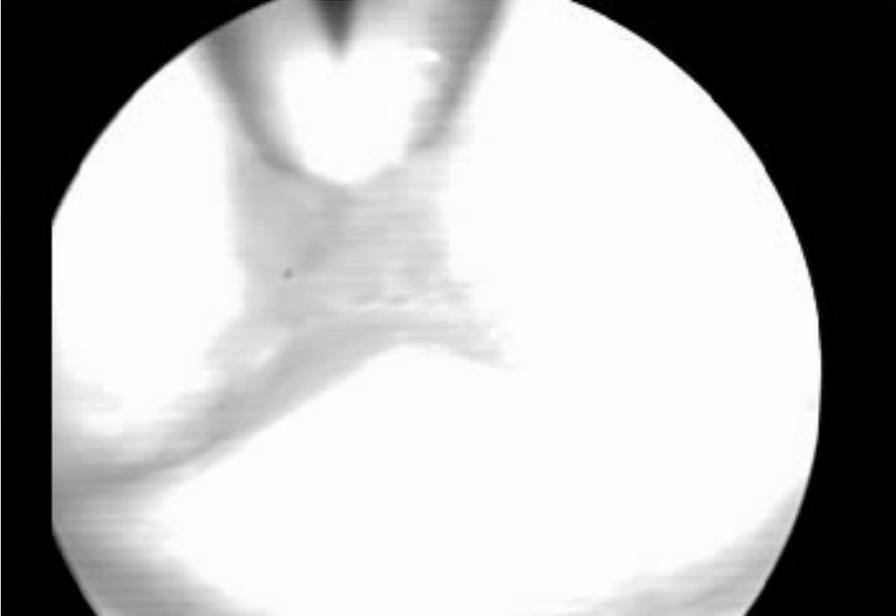
Time (total c. 18 ms)



- The development from analogic system makes difficult the successive analysis
- Temporal limit (18.4 msec)

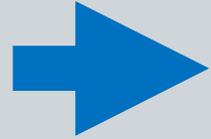
**ORL FERRARA**

*Da Svec, Sram & Schutte: Proc. LP '2000 Prague.*

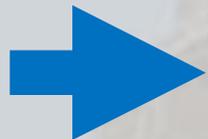


***ORL FERRARA***

# FOR THESE REASONS:



**NEW GENERATION OF VKG**



**DIGITAL KIMOGRAMS FROM  
HIGH SPEED CAMERA**

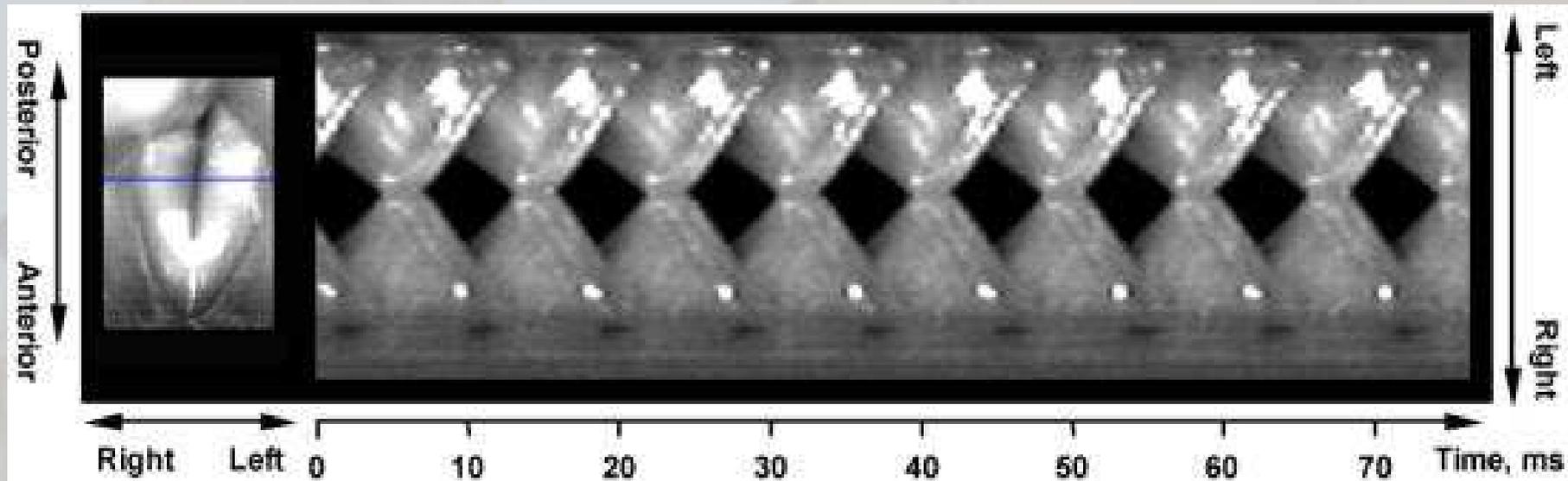
*ORL FERRARA*



## II<sup>nd</sup> GENERATION VKG

WITH A NEW CCD CAMERA:

1. Recording time more than 18.4 msec (40 msec.)

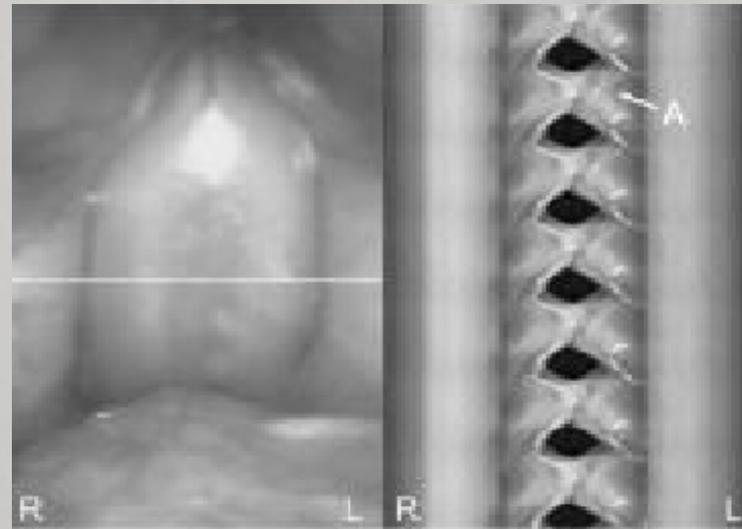
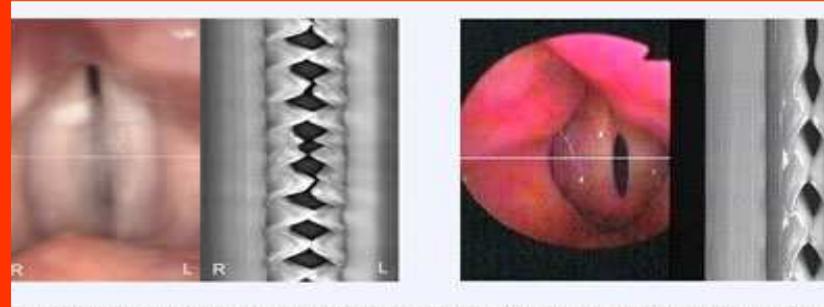


*ORL FERRARA*

# II<sup>nd</sup> GENERATION VKG

WITH A NEW  
CCD CAMERA:

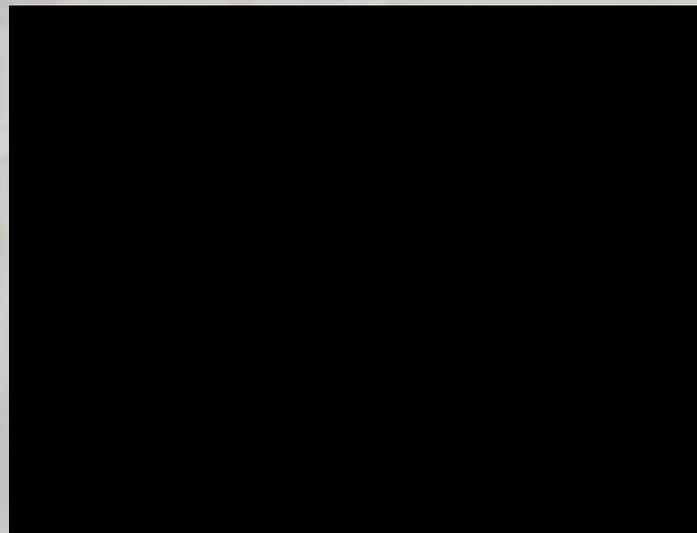
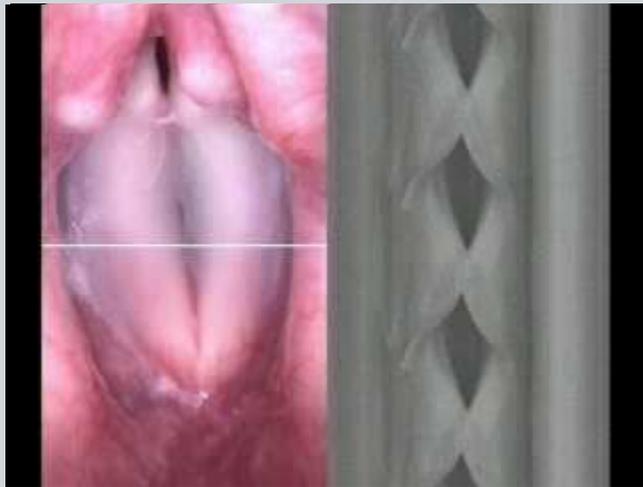
2. It's possible to have at the same time the laryngoscopic image as well the kymographic image
3. In every moment you can know where the selected line is



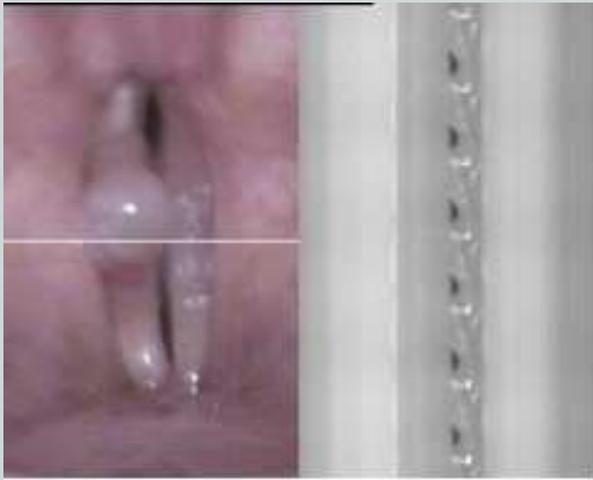
*ORL FERRARA*

# New Generation Video Kymography

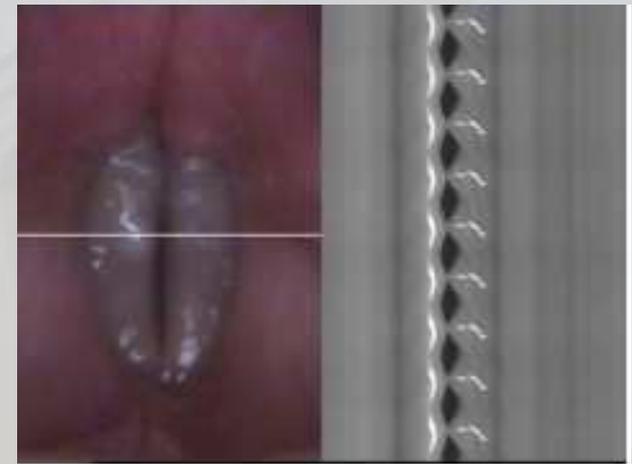
Now a split Screen with both full image and Line scan image displayed to provide better positioning



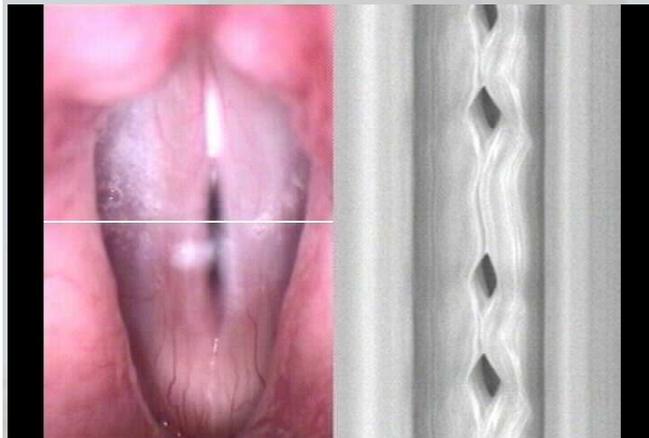
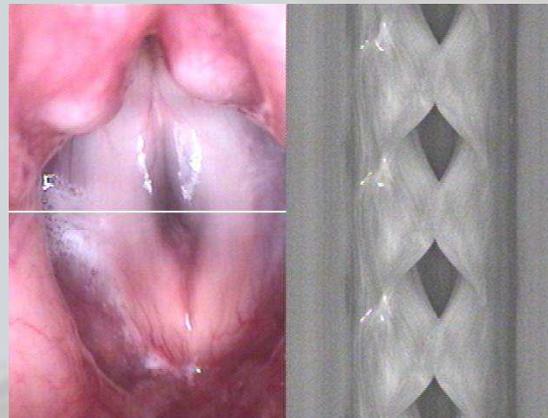
*ORL FERRARA*



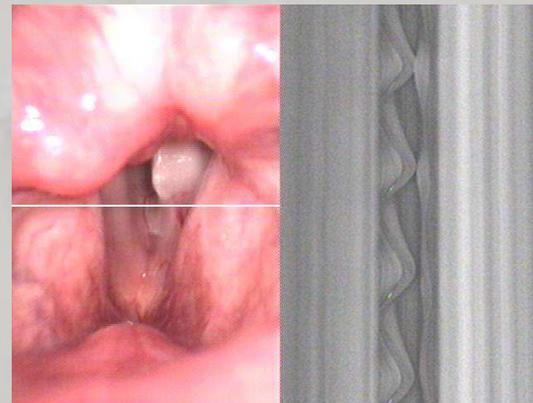
Polipo cordale destro in edema di Reinke



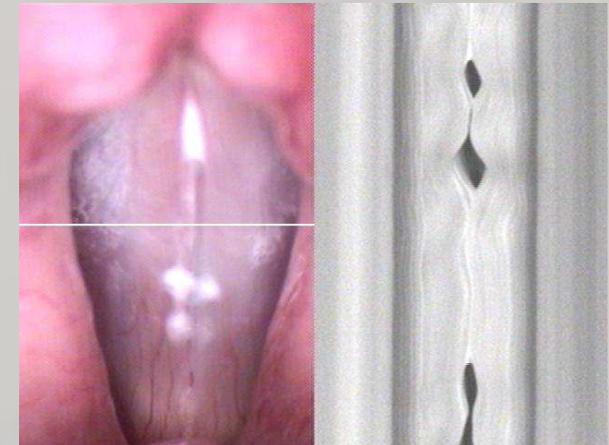
Cisti sottomucosa c.v. destra



Fonazione aperiodica



Paresi cordale sinistra

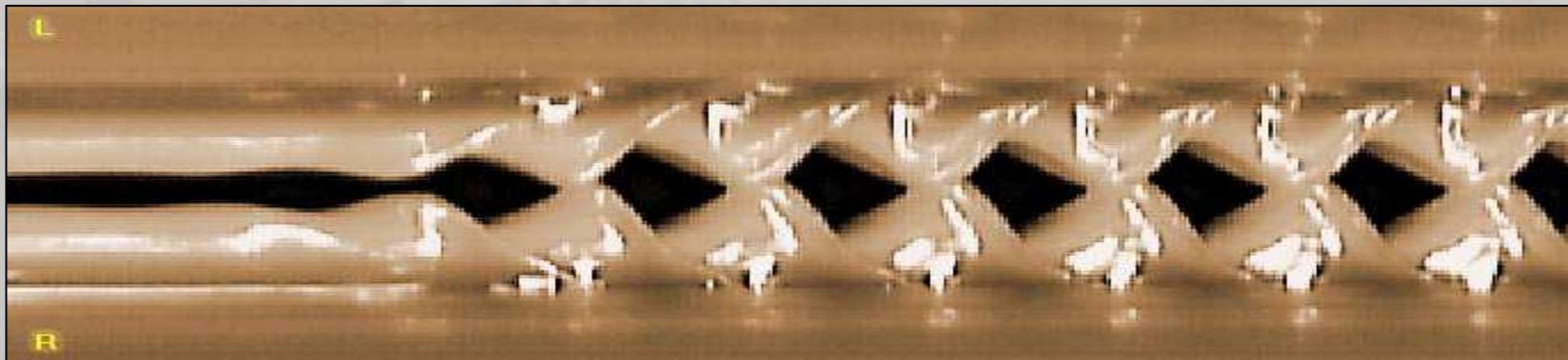


Fonazione molto aperiodica

**ORL FERRARA**

# DIGITAL KIMOGRAMS

- It's necessary to use the **HIGH SPEED CAMERA** (functional imaging)
- Digital kimograms can be reconstructed as a single image from digital highspeed exposures with the aid of image processing algorithms
- In this way a three-dimensional process (glottic vibration) became a two-dimensional process (the digital kimoogram); third dimension is time.



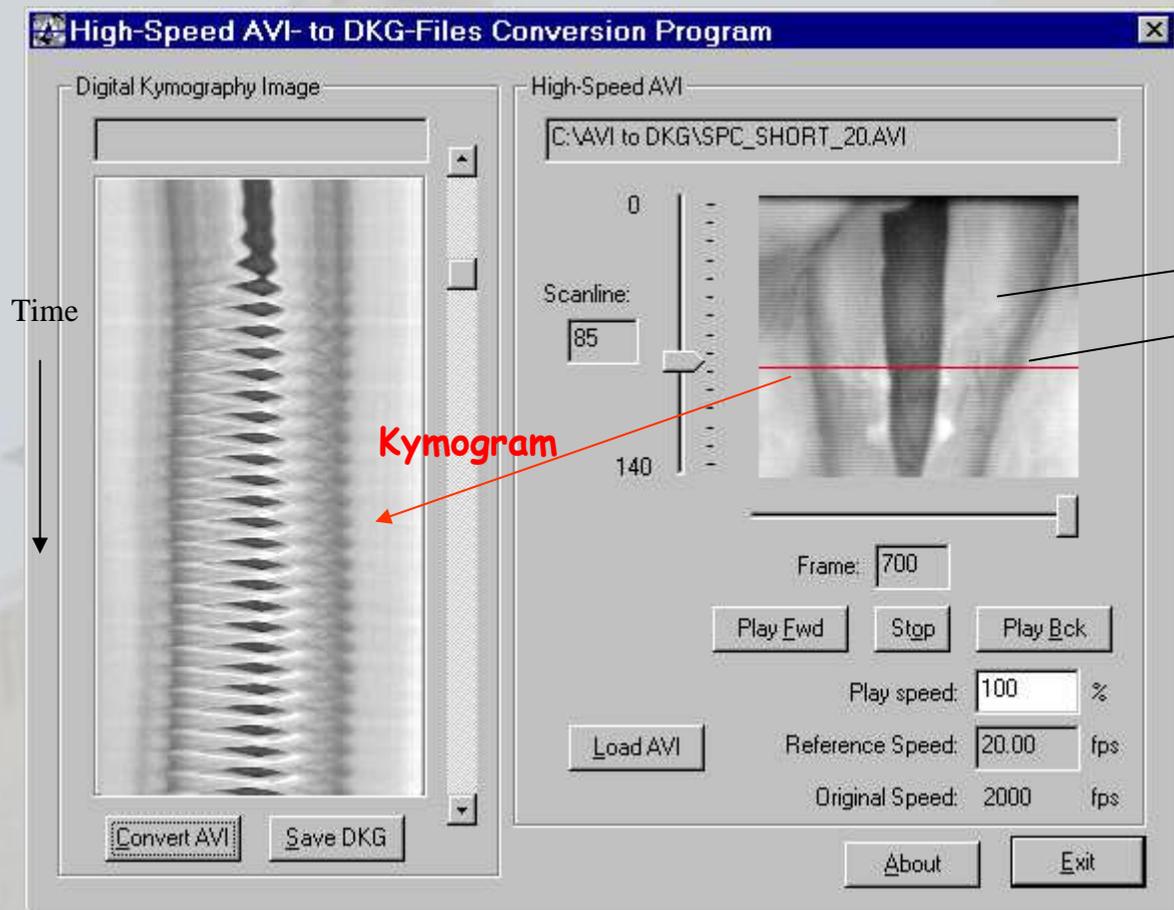
time



# High Speed Capture of the Glottic Image and Generation of the Line Scan Image

Figure 1 (right) is an example of high speed video imaging of vocal fold vibration at a rate of 2000 frames per second. A total of 2.048 seconds or 4096 frames may be captured and saved. The kymogram is generated from this high speed image. Kymogram generation requires the positioning of the scan line at a single glottic point. Figure 1 (left) is an example of a kymogram generated from the scan line depicted on the right. Kymograms may be generated at any glottic position.

Figure 1



Glottic high speed image  
Kymography scan line

Kymogram

ORL FERRARA

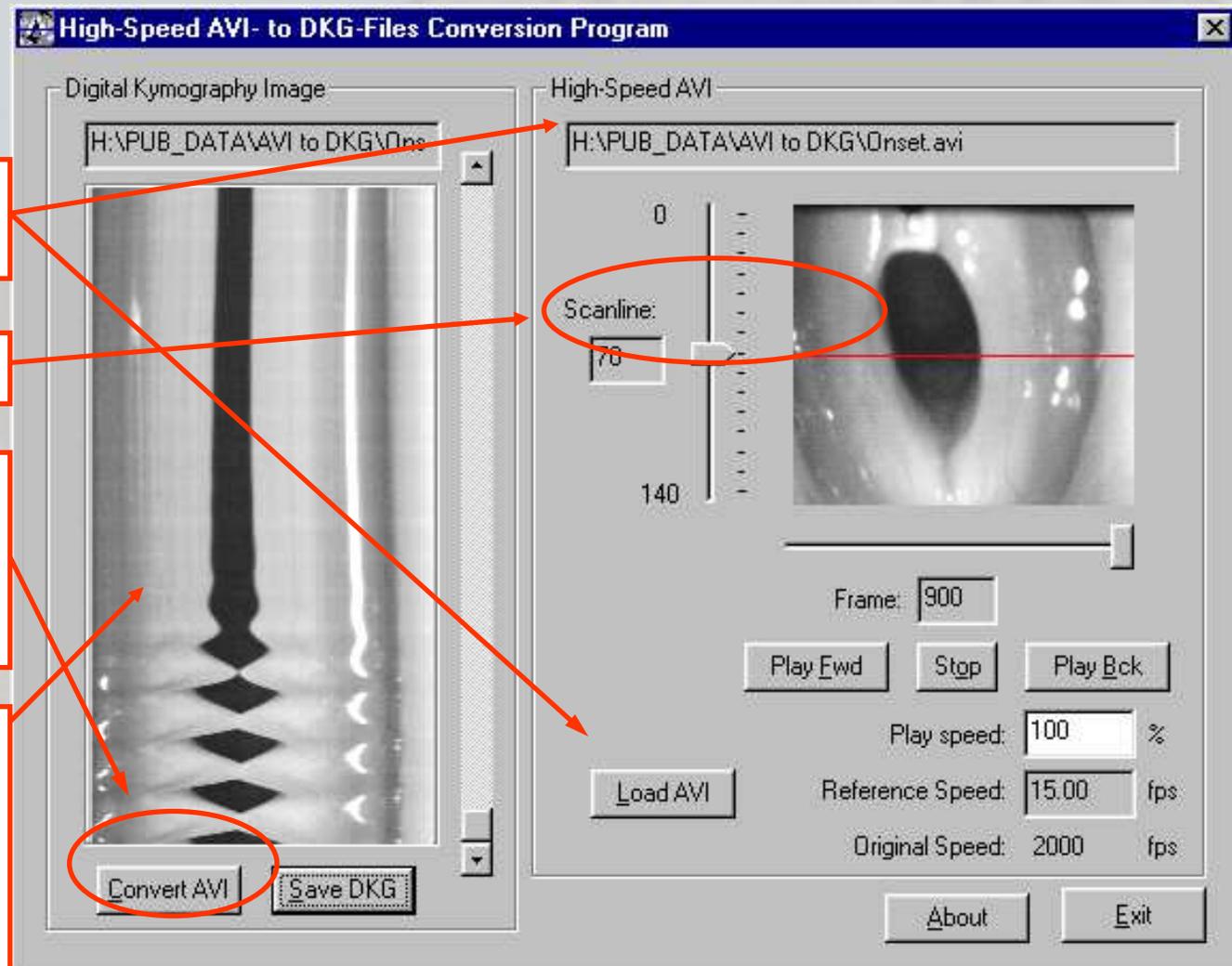
# Converting HSV to Digital Kymography

1. Load high-speed video AVI file.

2. Select line for analysis.

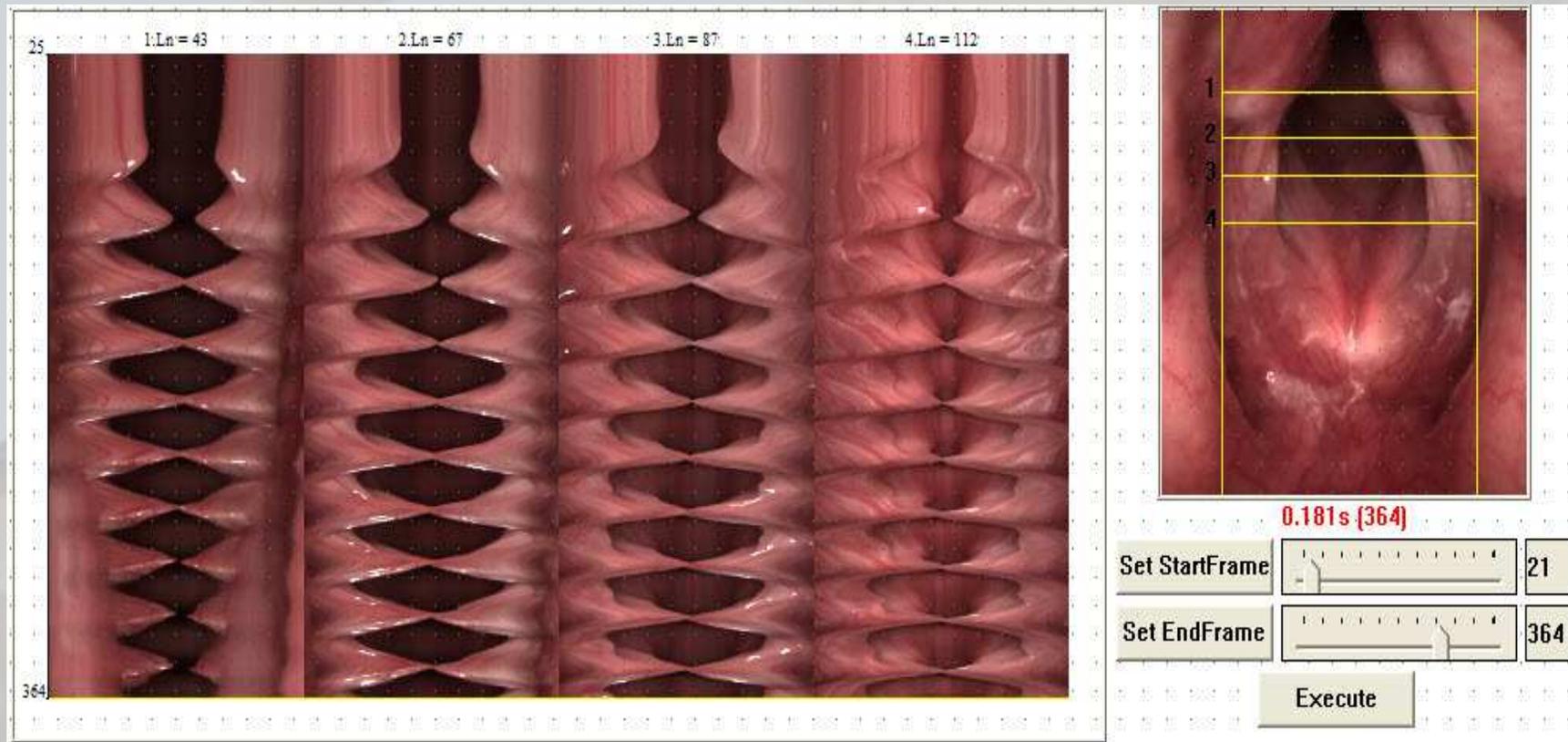
3. Press **Convert AVI** button to extract selected line from all stored 4096 frames

4. Digital Kymographic (DKG) analysis for 2 seconds is generated. Image is 140 pixels (wide) by 4096 pixels (length).



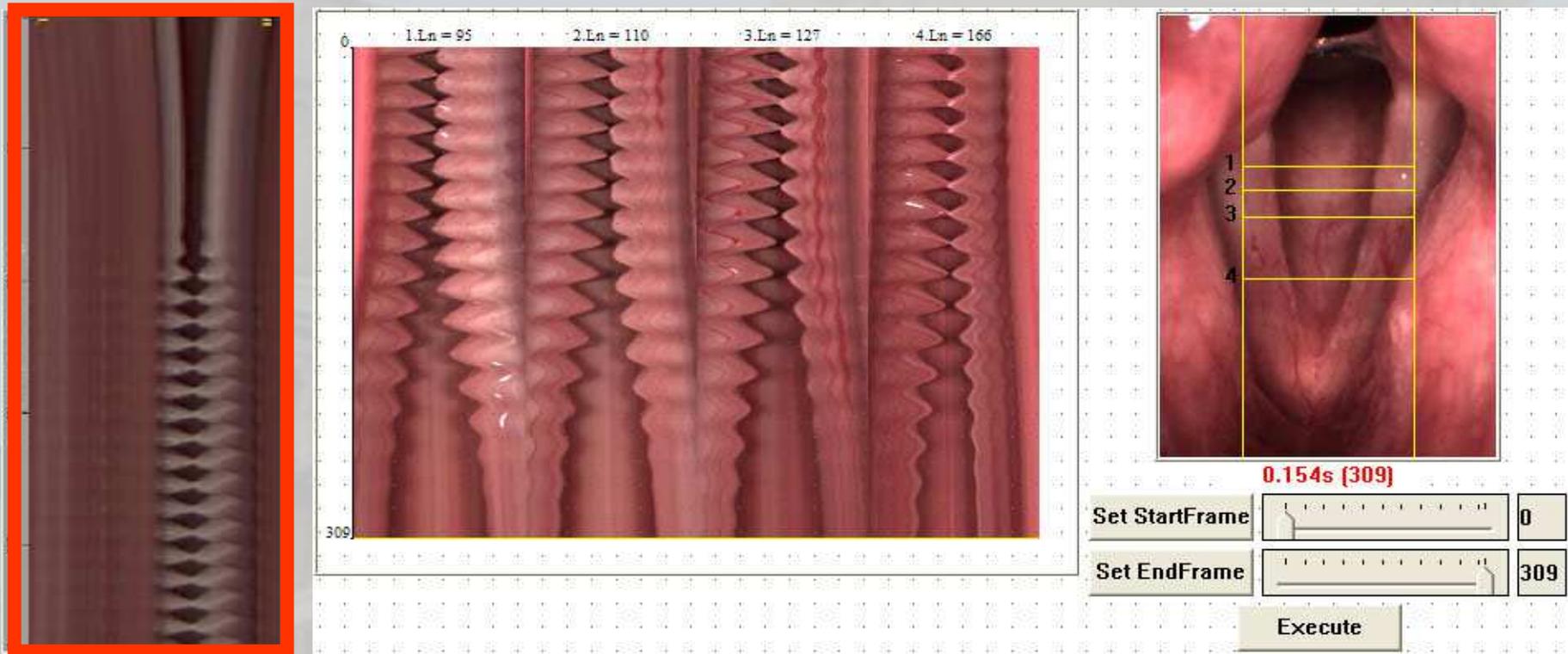
*ORL FERRARA*

# Loud Low Frequency Breathy Voice Kymography Display



**ORL FERRARA**

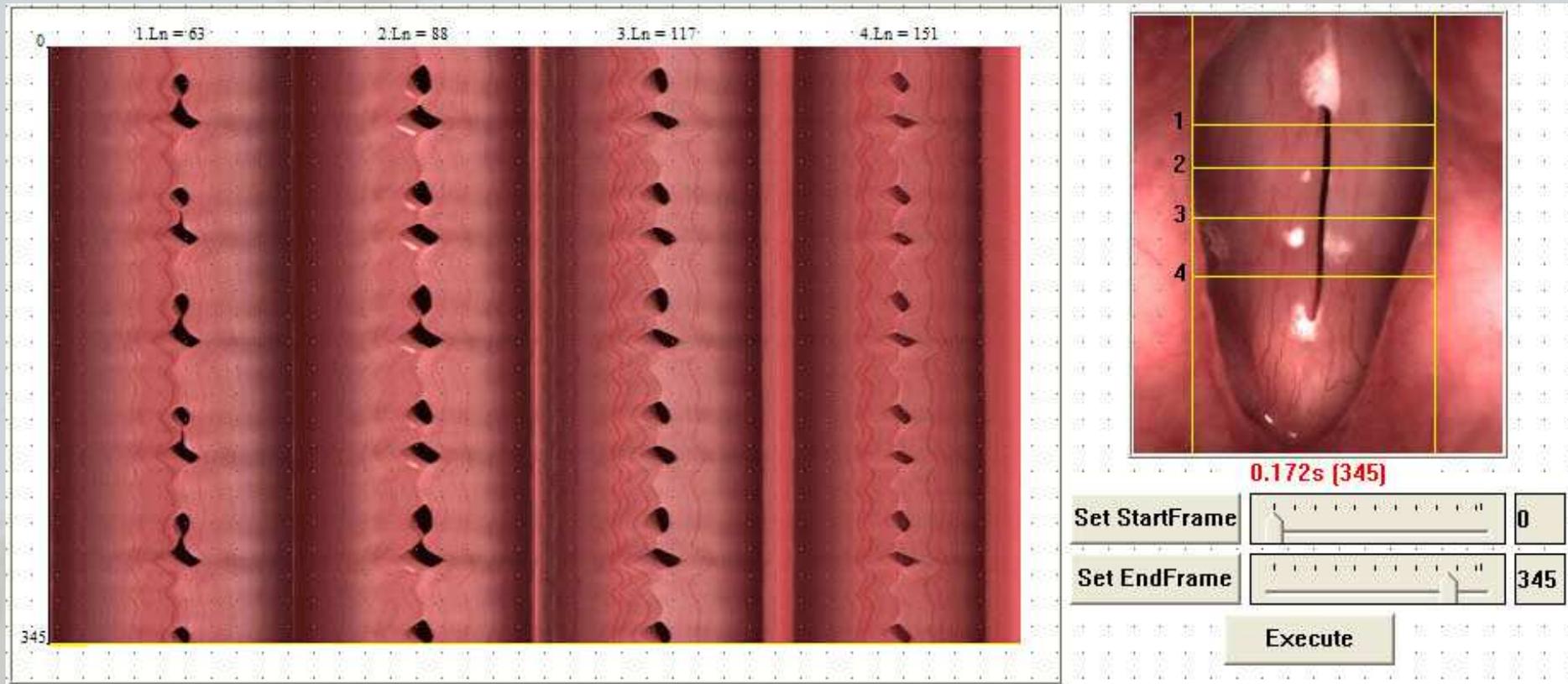
# Voicing Onset/Offset Kymography



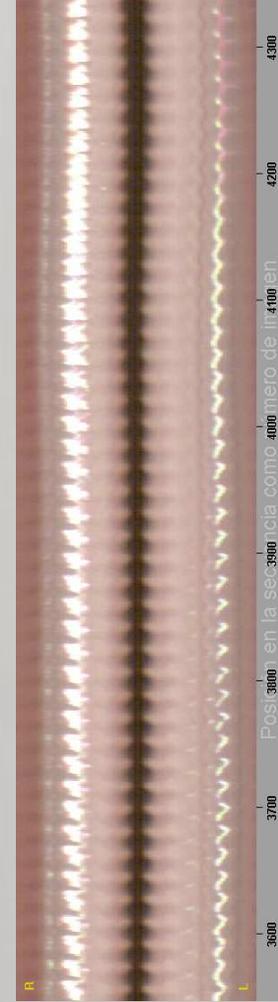
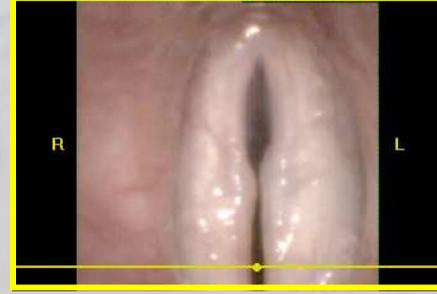
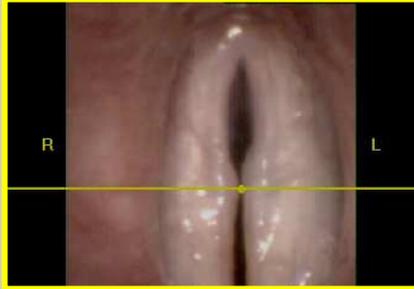
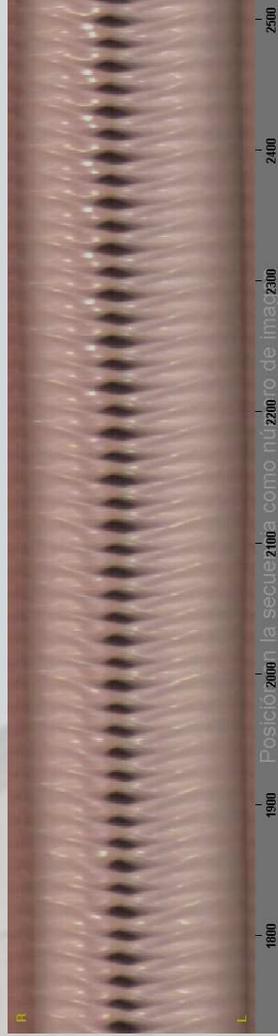
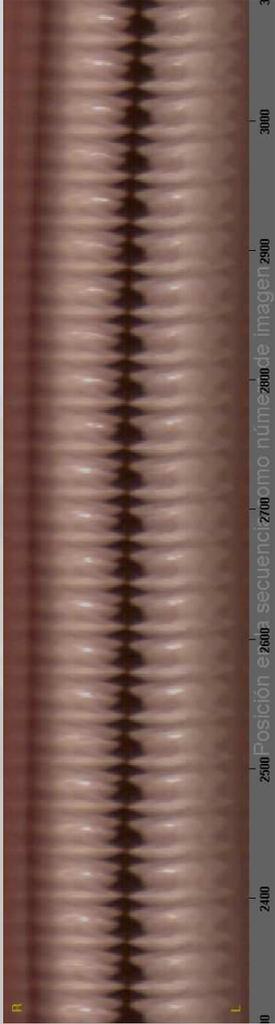
**ORL FERRARA**

# Glottal Fry kymography

Note the  
Aperiodic  
behavior



**ORL FERRARA**



**ORL FERRARA**

# II<sup>nd</sup> generation VKG promises to fill a key role in broadening the understanding of phonatory dynamics

- High-speed line scanning (8000 frames/second) of the vocal folds
- Information obtained similar to high-speed film at a fraction of the cost
- Accurate representation of physiology in all vocal behaviors
- Usable with most constant light sources and rigid endoscopes



*ORL FERRARA*

# ADVANTAGES OF DIGITAL KIMOGRAPHY

Many drawbacks of classic VKG are passed by digital kimograms that allow :



- To chose the line to select after recording
- To obtain many kimograms from the same recording modifying the position of the active line
- Compared to HSC, VKG is based on a single line of the glottic plane, but gives a better spatial resolution (pixels) as well a better time resolution (8000 fps)

*ORL FERRARA*

## CONCLUSIONS:

### KYMOGRAPHIC IMAGING

- new possibilities for the diagnostic of voice disorders
- large amount of informations about the dynamic behavior of the vocal folds
- a lot of the informations are new and not yet explored

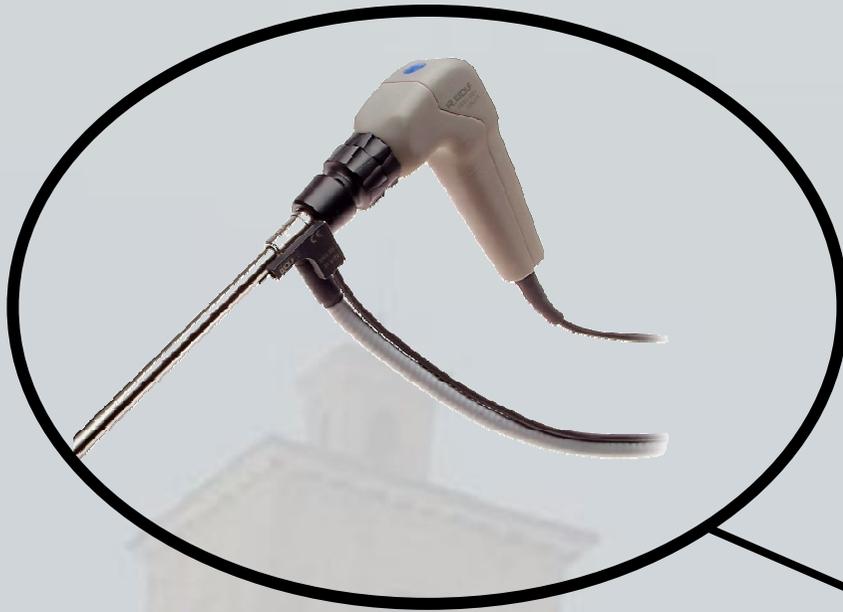
## CONCLUSIONS:

### KYMOGRAPHIC IMAGING

- informations useful for basic research as well as clinical practice
  - but VKG is a time consuming method
  - .....VKG is to be considered as a complementary method to stroboscopy

# High Speed Imaging

- Developed for the first time in the Bell laboratories in 1938
- This system was time-consuming and expensive, so it couldn't be
- In 1980 first digital high speed camera was built for sport medicine and crash test examinations



High speed cameras now available are different for:

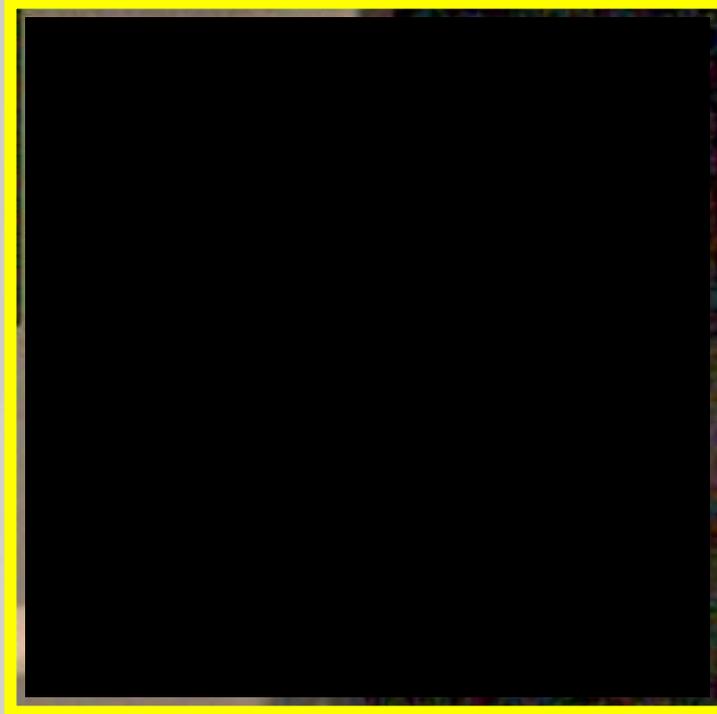
Spatial resolution

Time resolution

Exposure duration

# RESOLUTION

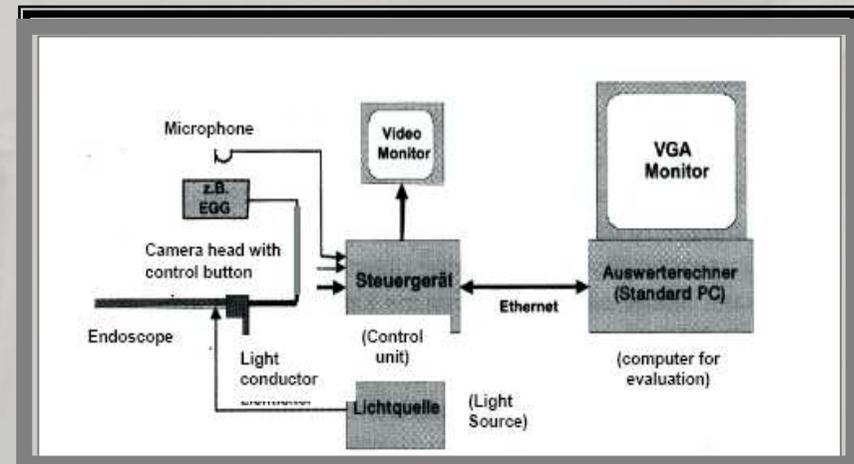
- Time resolutions can go from 1000 to 4000 frames per second with a highest spatial resolution of 1536 x 1024 pixels (with 1000 fps)
- To evaluate vocal fold vibration we need a time resolution of 4000 frames per second (with a worsening spatial resolution to 256 x 256 pixels) so it's possible to acquire this movements according to Shannon theorem



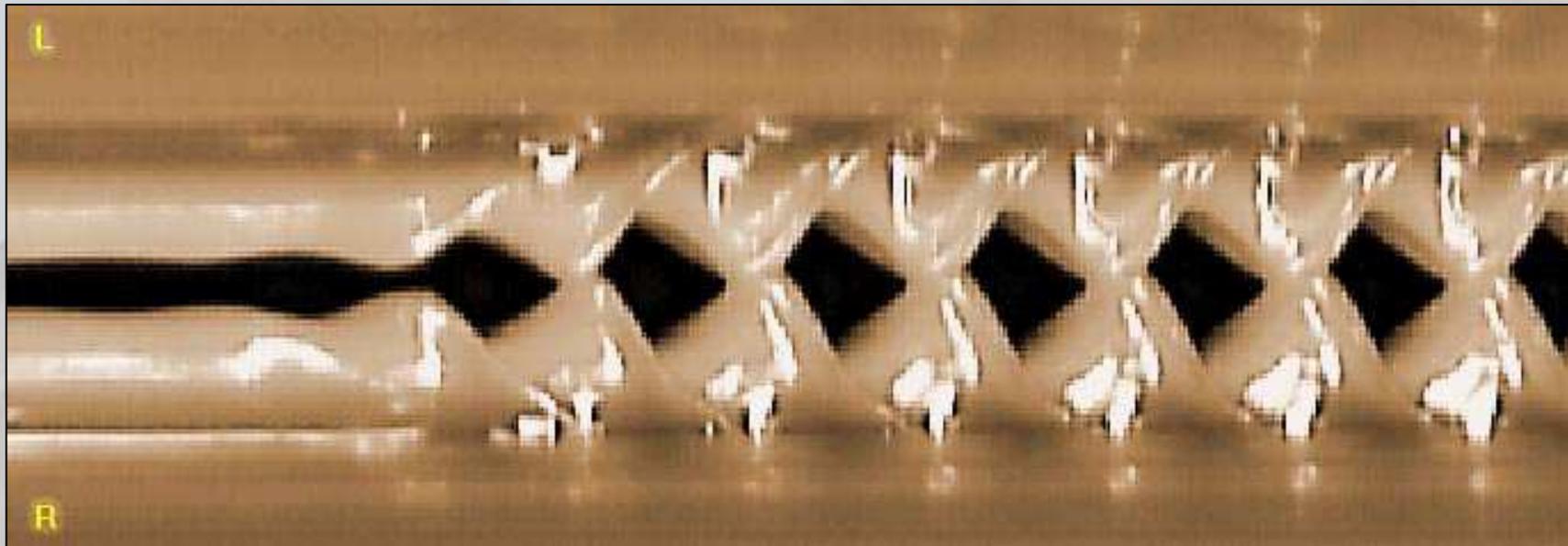
- Exposure duration currently needed in clinical routine are no longer than 2 seconds
- This exposure duration allows to appreciate short events like voice onset and offset, incomplete vocal closure (order of magnitude of 300-500 msec).
- Exposure time duration is limited by camera memory



- Data are archived by a control module on a commercially available PC.
- From digital high speed exposure with the aid of image processing algorithms space-time curves are reconstructed.
- Characteristic parameters can be derived :
  - 1 Initial response
  - 2  $F_0$
  - 3 Amplitude
  - 4 Open and close quotients



Digital videokymograms are acquired from digital exposure



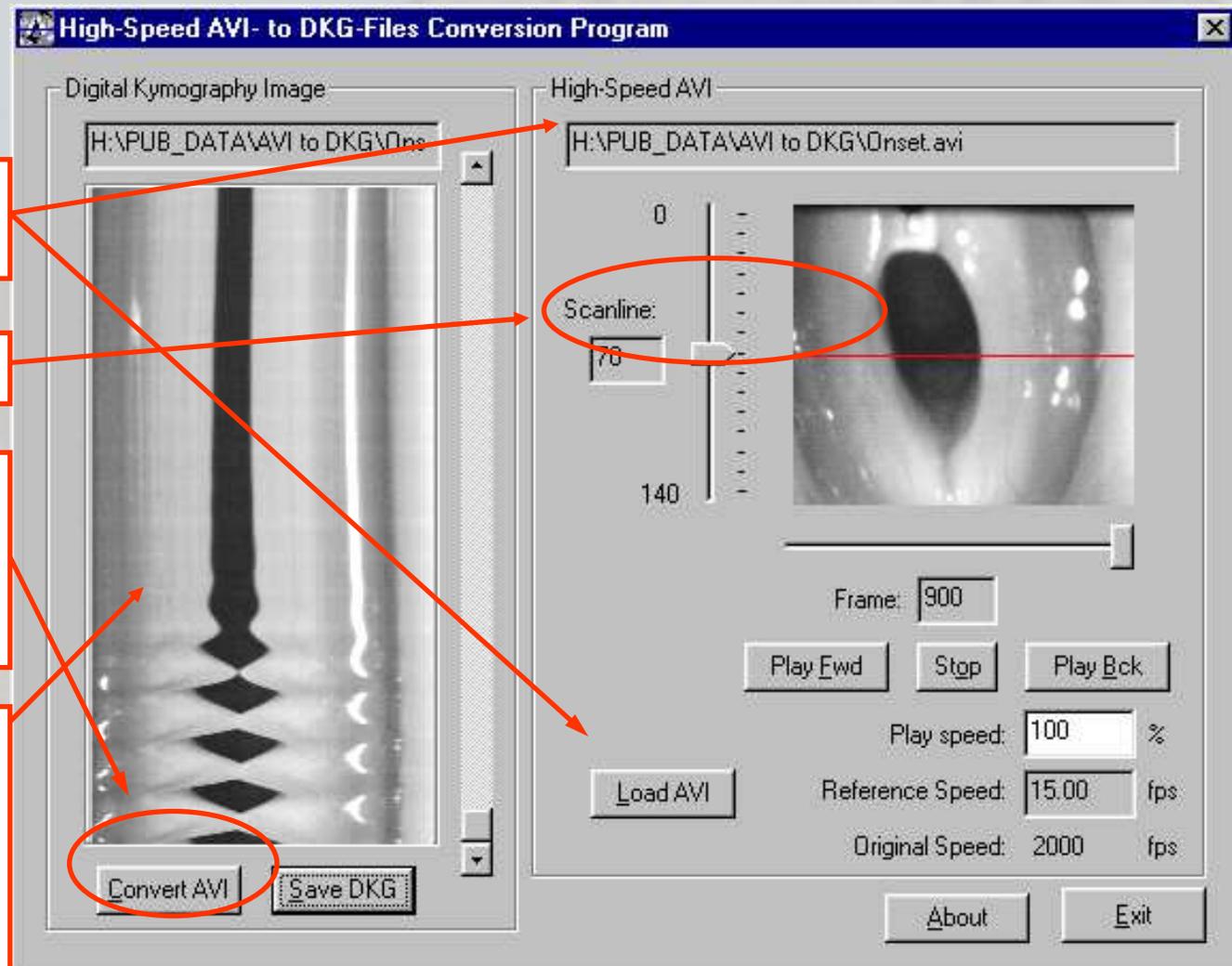
# Converting HSV to Digital Kymography

1. Load high-speed video AVI file.

2. Select line for analysis.

3. Press **Convert AVI** button to extract selected line from all stored 4096 frames

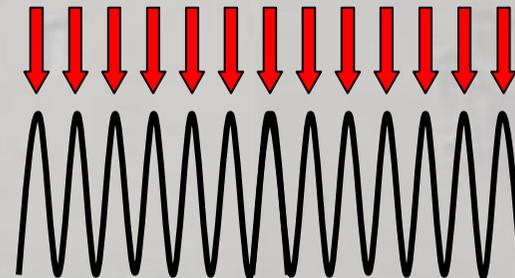
4. Digital Kymographic (DKG) analysis for 2 seconds is generated. Image is 140 pixels (wide) by 4096 pixels (length).



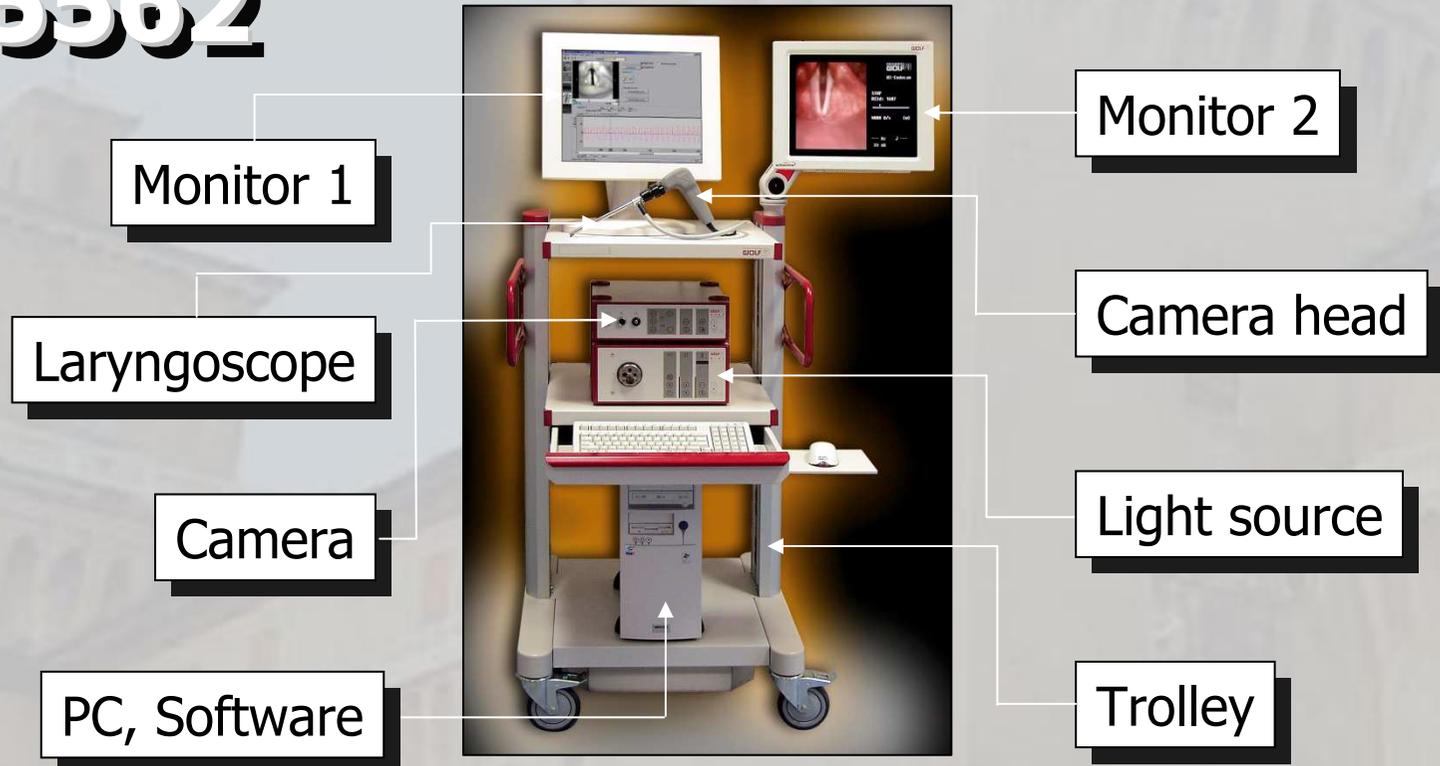


# HRES ENDOcam 5562

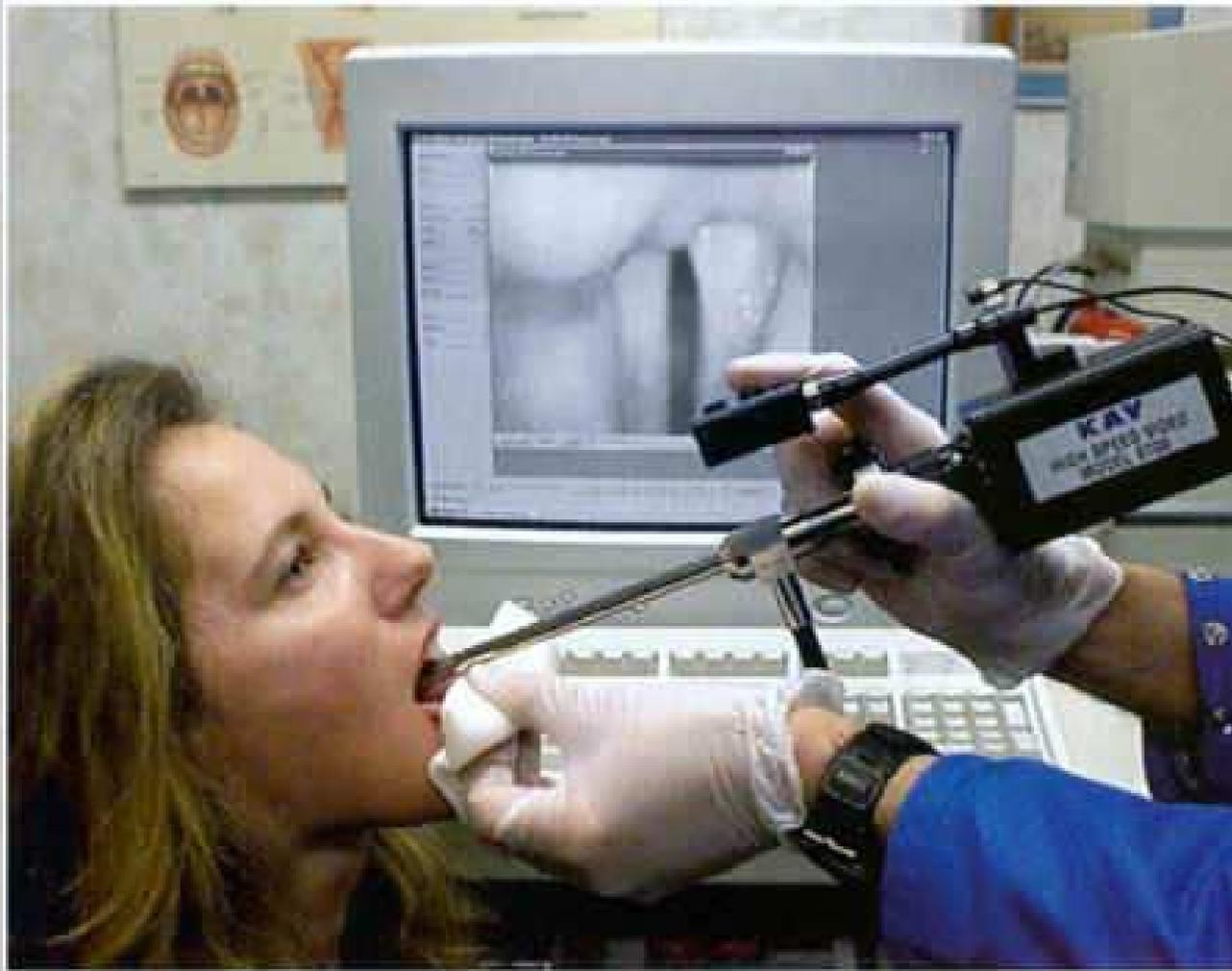
**4000** frames/sec



# HRES ENDOcam 5562

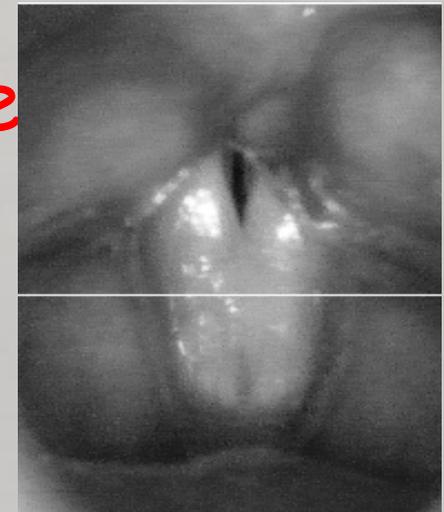


# High Speed Video System 9700



## ADVANTAGES and NEW POSSIBILITIES

- Time resolution up to 4000 frames per second i.e. evaluation of frequencies of vocal fold vibration up to 800 Hz is possible.
- In contrast with videokymography high speed camera acquire images of the vibration process on the whole glottis
- Several digital kymograms can be extracted from one single high speed recording



# ADVANTAGES and NEW POSSIBILITIES



→ Improvement of the conventional stroboscope system

→ Possibility to see the swing of the vocal cords in real time (slow motion)

→ Diagnostic without voice (hoarseness patients)

→ Accurate evaluation of the changing of registers



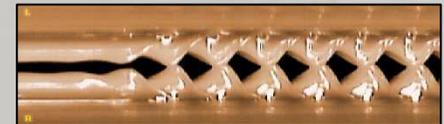
# New medical information

→ For the first time it is possible to record the transient response of the vocal cords.

→ Evaluation of very short voicing segments and glottic spasms

→ Evaluation of irregular/aperiodic vibration

→ Compile of a digital kymograms



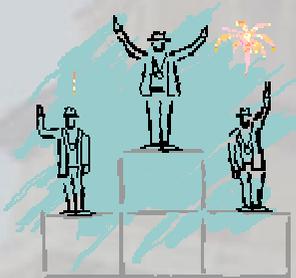
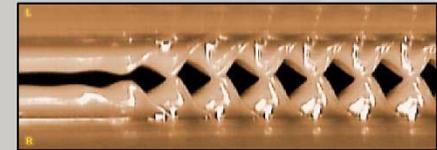
# Documentation System

→ digital high speed recordings of oscillations of vocal folds

→ administration of the corresponding patient data

→ generation of digital kymograms

→ computation of glottograms



# DISADVANTAGES

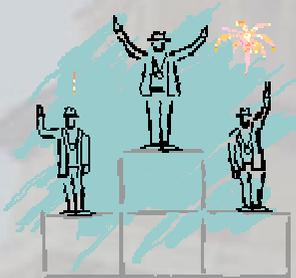
- Low spatial resolution (256 x 256 pixels) than stroboscopy (700 x 500 pixels)
- Short exposure duration (2 seconds)
- Lower image resolution for the study of morphological changes (polyps, nodules, etc) as well color deviations compared to endoscopy or stroboscopy

# Conclusion

→ With the new system we start a new period of vocal cords diagnostic.

→ This is the first system in the world, where you can see the real swinging of the vocal cords.

→ With this system you get a diagnostic of patients without sound or with hoarseness problems.

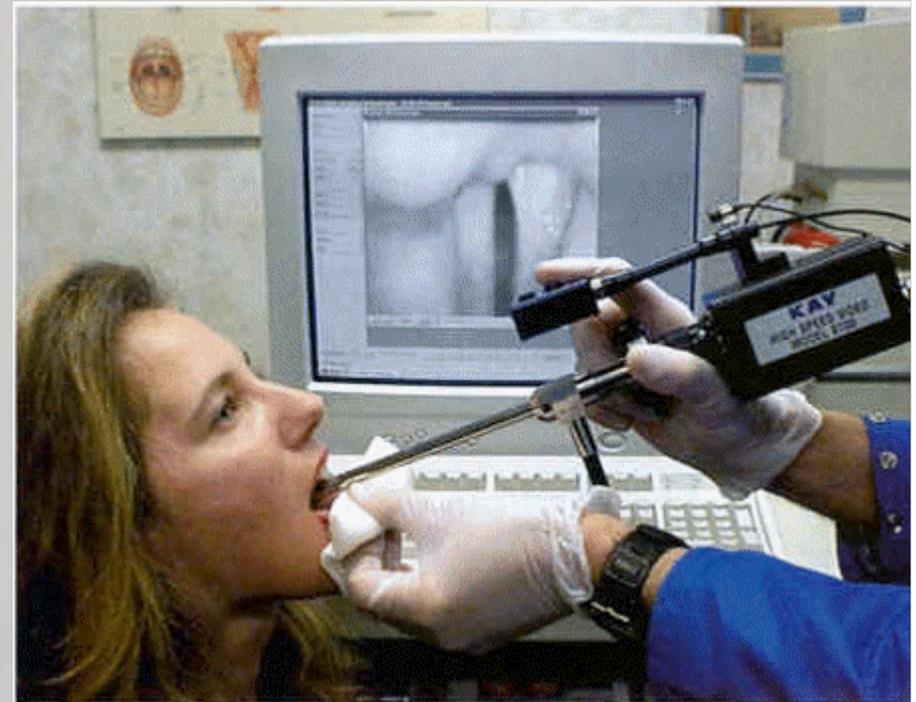


# Our View of High-Speed Video

- HSV is a clinical research tool.
- HSV helps us understand what stroboscopy can and cannot see.

Presently, HSV does not replace stroboscopy.

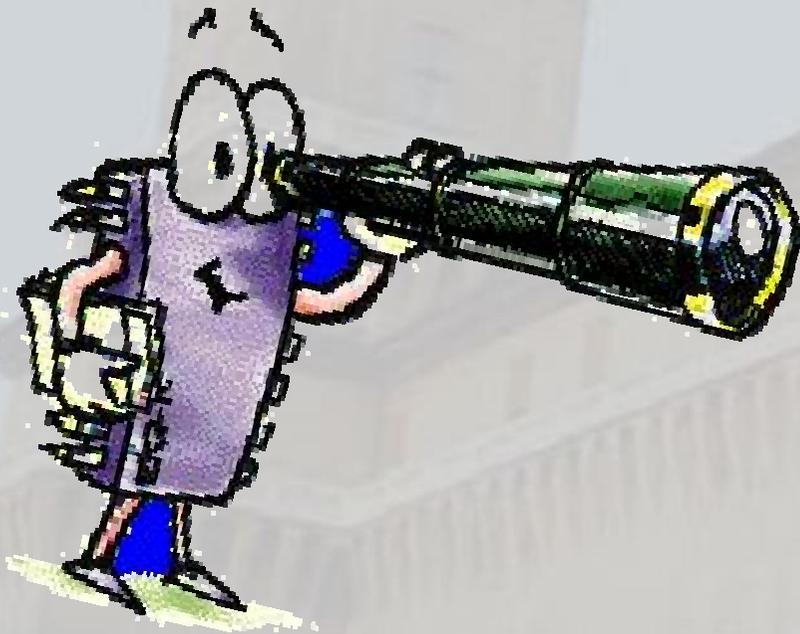
HSV does point clearly to the utility of VKG as a practical alternative in viewing some voicing behavior.





*HRES ENDOcam  
THE VISIBLE VOICE*

# GOAL HRES-ENDOCAM



**Physical visibility of  
the voice disturbed**

**Movement analysis  
of the vocal cords**

**Visual of irregular  
swinging**