

# Assessment of Voice using Acoustic Analysis

Peter Murphy



Department of Electronic and  
Computer Engineering, University  
of Limerick

# Overview

- ◆ Introduction/Acoustic analysis techniques
- ◆ Issues in the acoustic analysis of voice
- ◆ Conclusion/Work for COST 2103

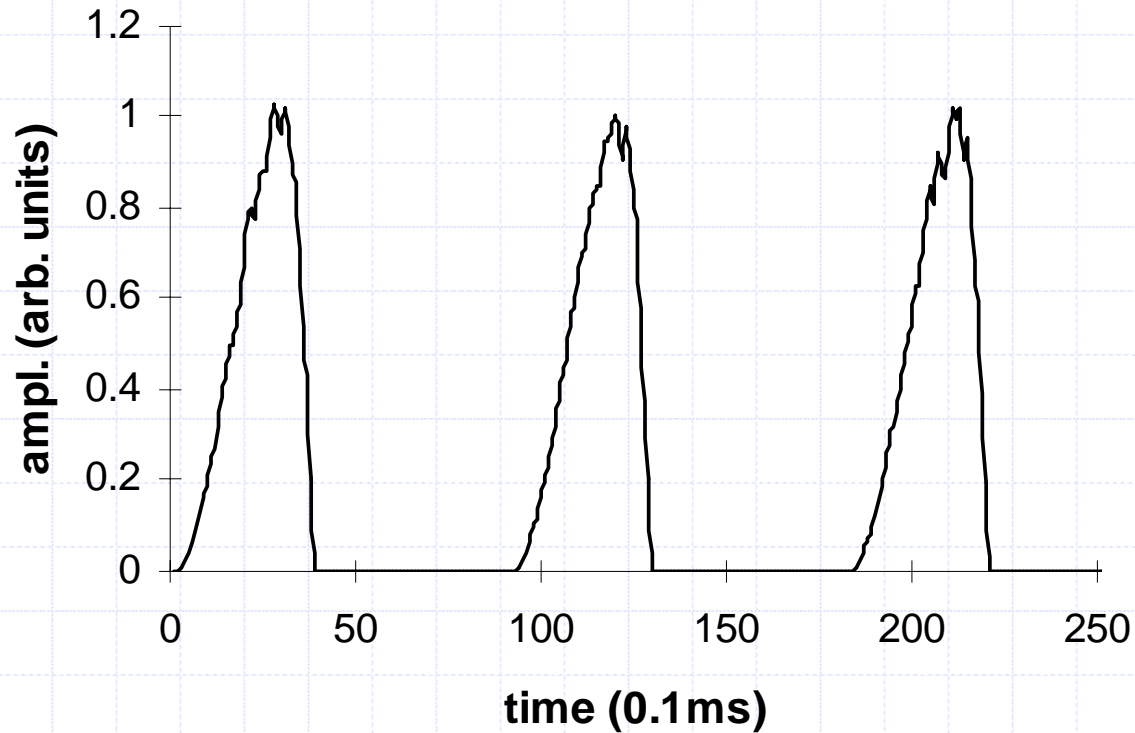
# Introduction – Acoustic analysis

- ◆ Early Methods- quantify noise/periodicity levels based on the visual inspection of spectrograms (Yanagihara, 1967)
- ◆ Indices automatically extracted from the acoustic speech signal
  - jitter
  - shimmer
  - noise (HNR)
  - waveshape
  - indicators of nonlinear behaviour

# Issues in the Acoustic analysis of Voice

- ◆ Non-independence of measures
- ◆ Perceptual correlation
  - auditory processing
- ◆ Correlation with production mechanism at glottis
  - inverse filtering

# Glottal flow with noise



# HNR (Speech) is indirectly related to HNR (Glottal Flow)

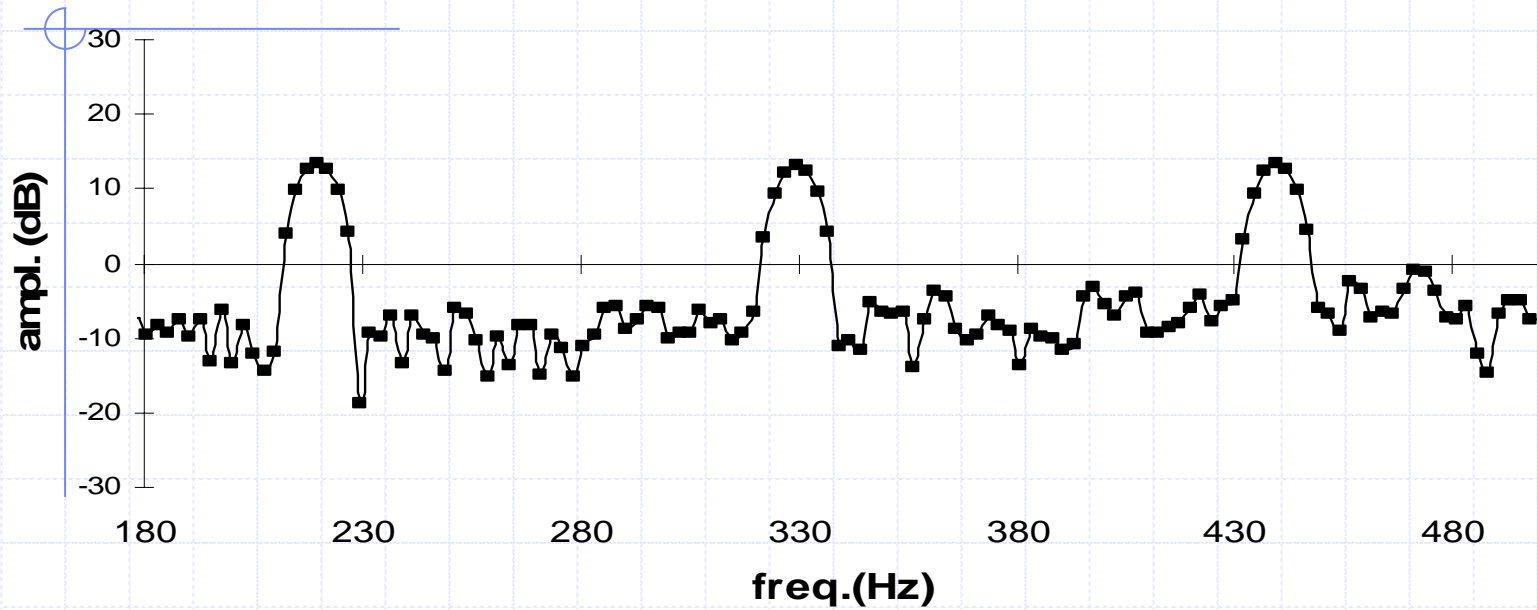
## Example methods:

- Relative harmonic intensity-Hiraoka *et al.*, 1984
- Normalised noise energy-Kasuya *et al.*, 1986
- Pitch synchronous (four period) analysis-Muta *et al.*, 1988

## Basic process for all methods:

- sum energy at harmonics
- sum noise
- calculate harmonic-to-noise index

# Relative harmonic intensity-Hiraoka *et al.*, 1984

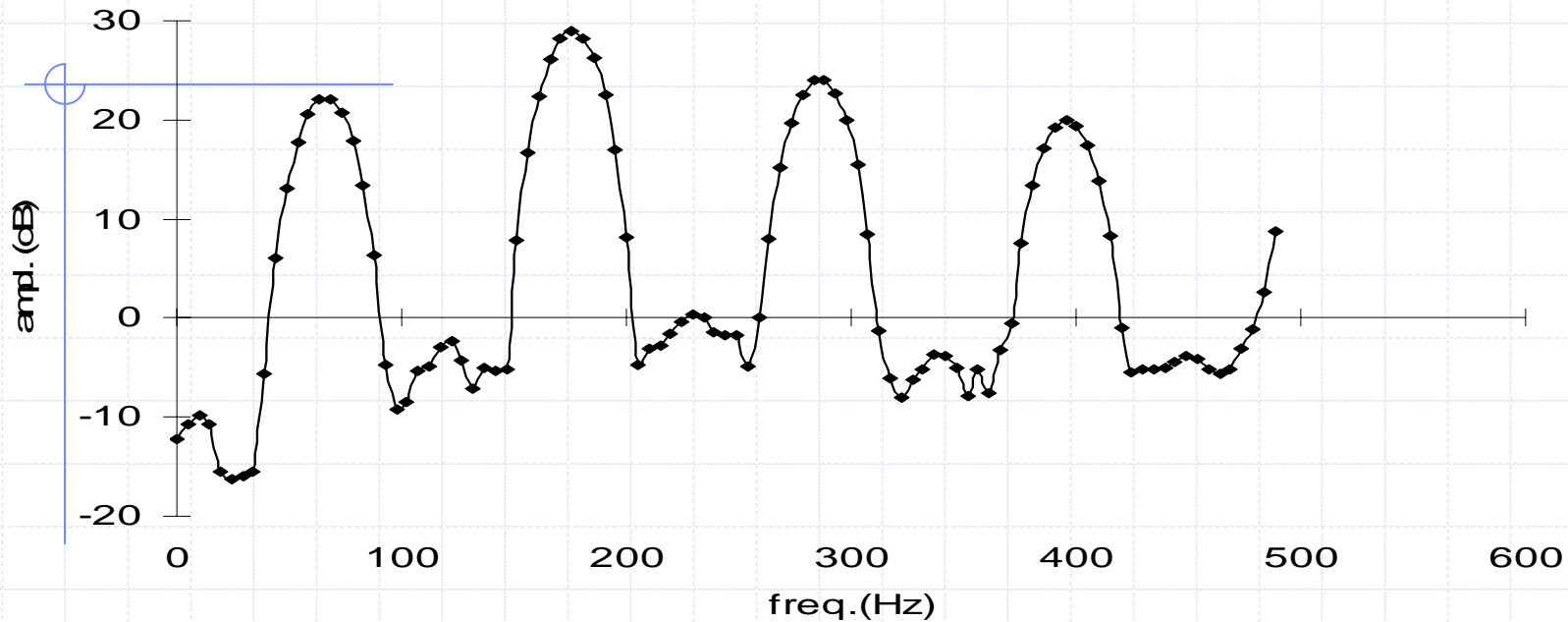


$$H_r = \left( \frac{\sum_{i>2} h_i}{S} \right) \times 100(\%)$$

Sum harmonic energy (except at  $f_0$ )

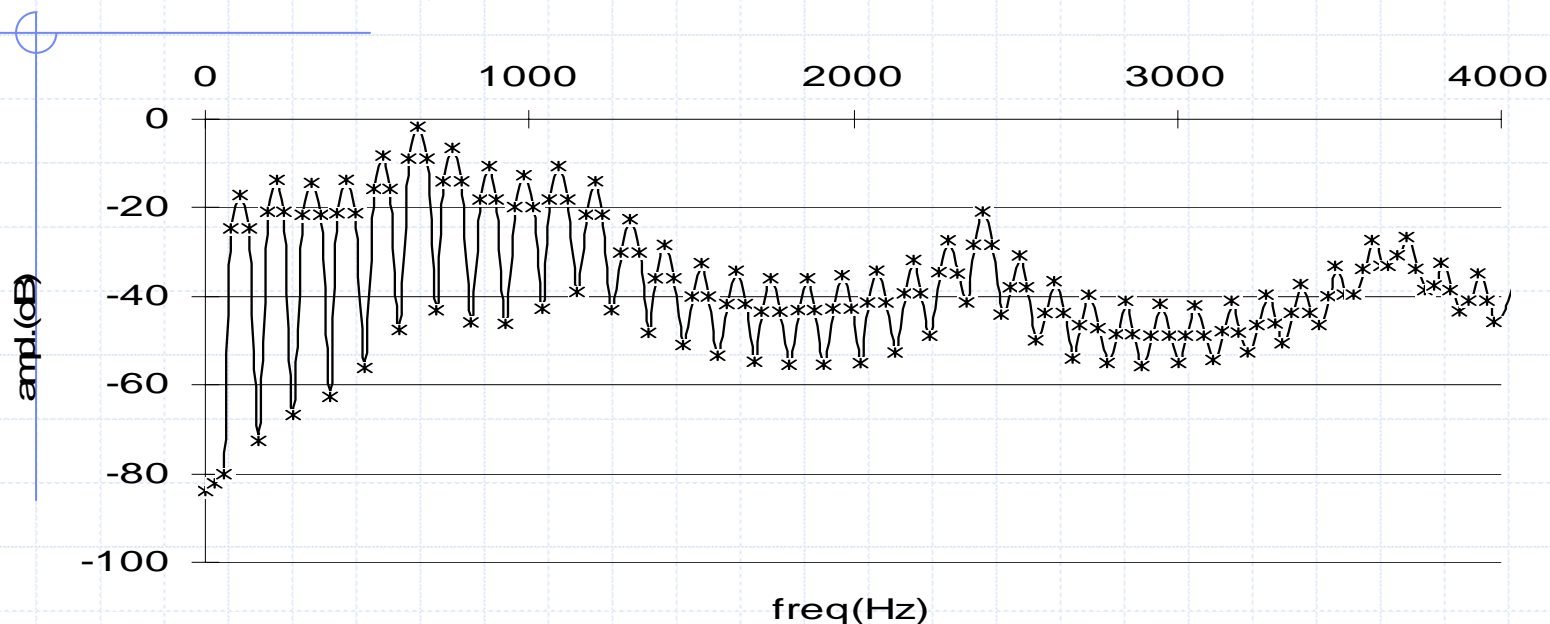
Divide by total signal energy

# Normalised noise energy-Kasuya *et al.*, 1986



Sum energy at harmonics. Sum energy at ‘between harmonics’ (i.e. noise). Estimate noise at harmonics by interpolating adjacent ‘between harmonic’ estimates. Divide total noise energy by total signal energy.

# Pitch synchronous (four period) analysis— Muta *et al.*, 1988

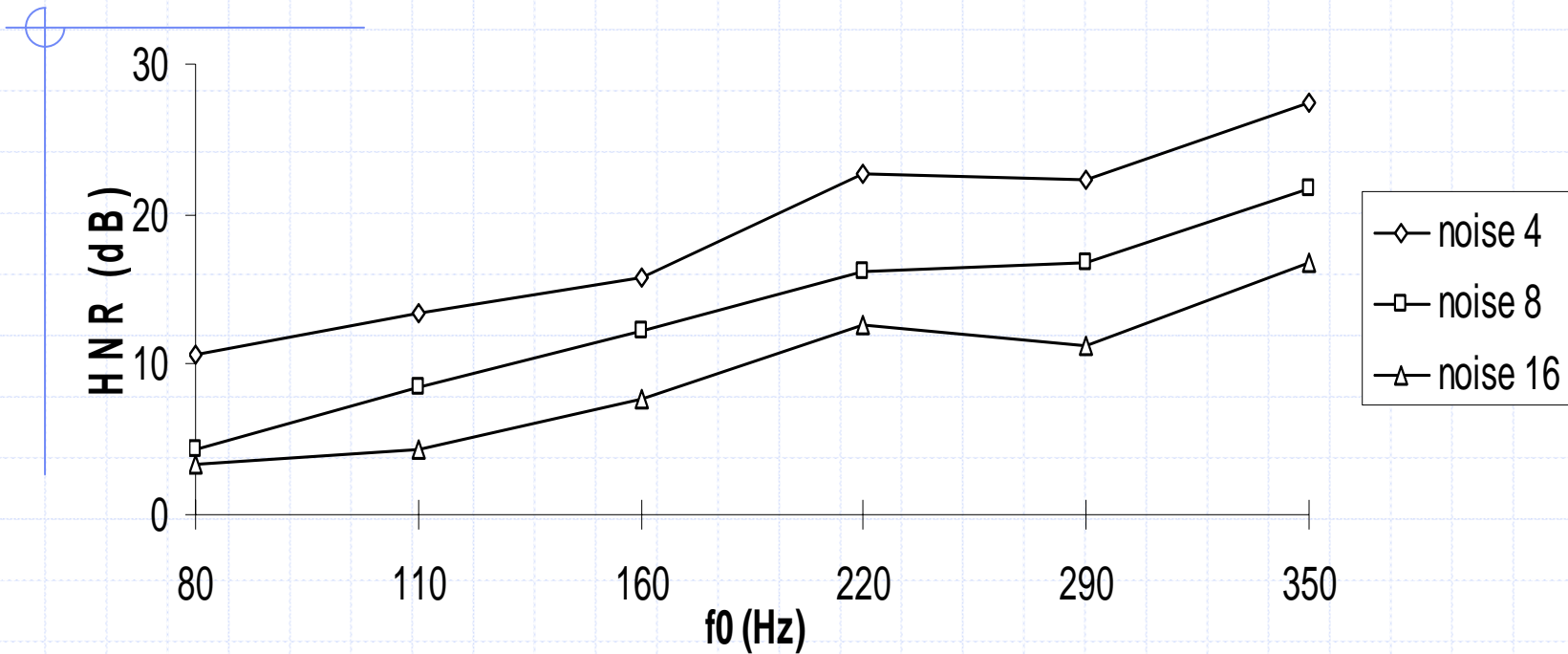


Scheme: three points—harmonic, one point-noise

Sum harmonic and noise energies for the first 16 harmonics

Divide summed noise energy by summed harmonic energy

# Results



Harmonics-to-noise ratio vs f0 for the synthesized vowel a/ for three levels of glottal additive noise std. dev. 4%, 8% and 16%

# Discussion

◆ As a consequence of source-filter theory the HNR of the glottal source is not equal to the HNR of the speech signal. For equal noise levels at the glottis the HNR varies in the speech signal as  $f_0$  varies (as shown in results for Hiraoka et al., 1984).

◆ Filter-free HNR - gmHNR (cepstrum)

◆ Perturbation-free HNR

# Inverse filtering

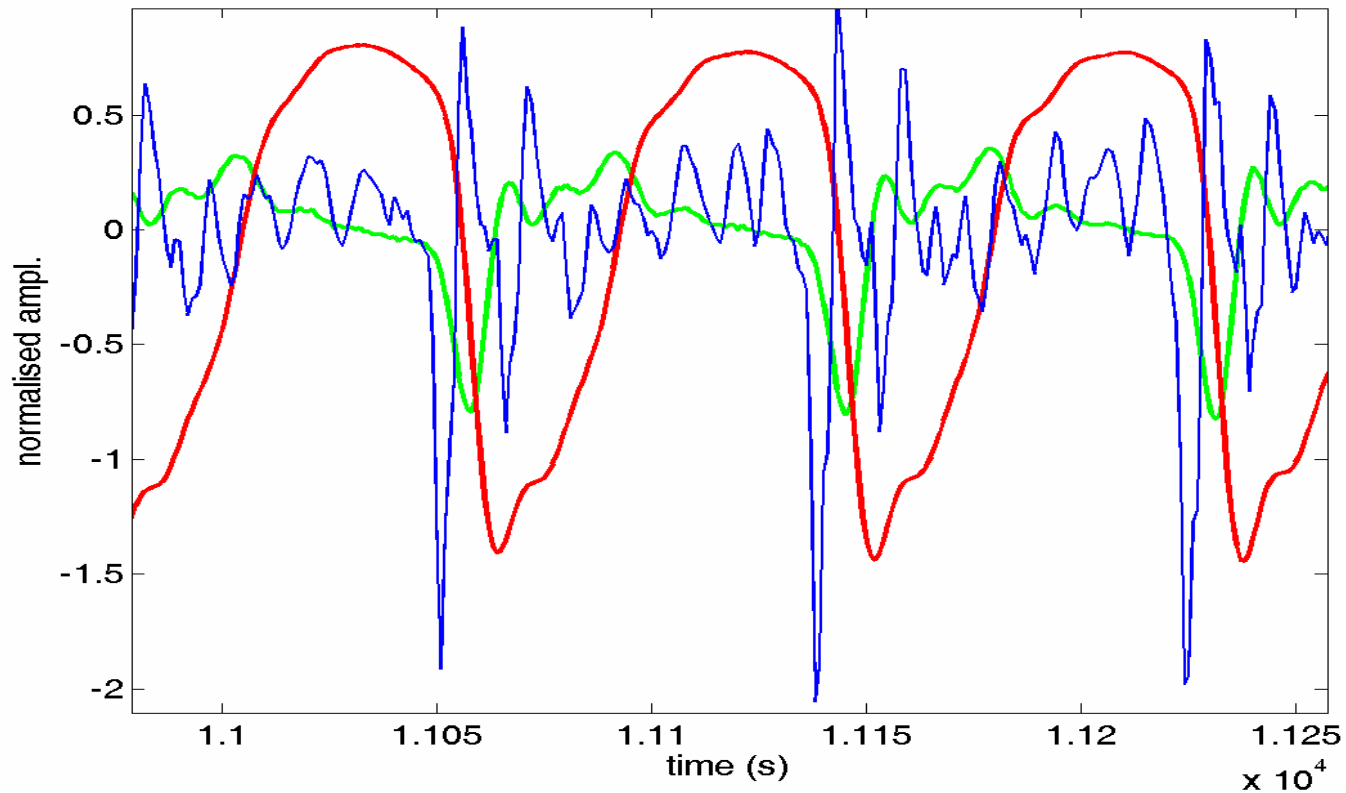
$$G(z)P(z) = \frac{S(z)}{AV(z)R(z)}$$

- ◆ Remove the effect of the vocal tract

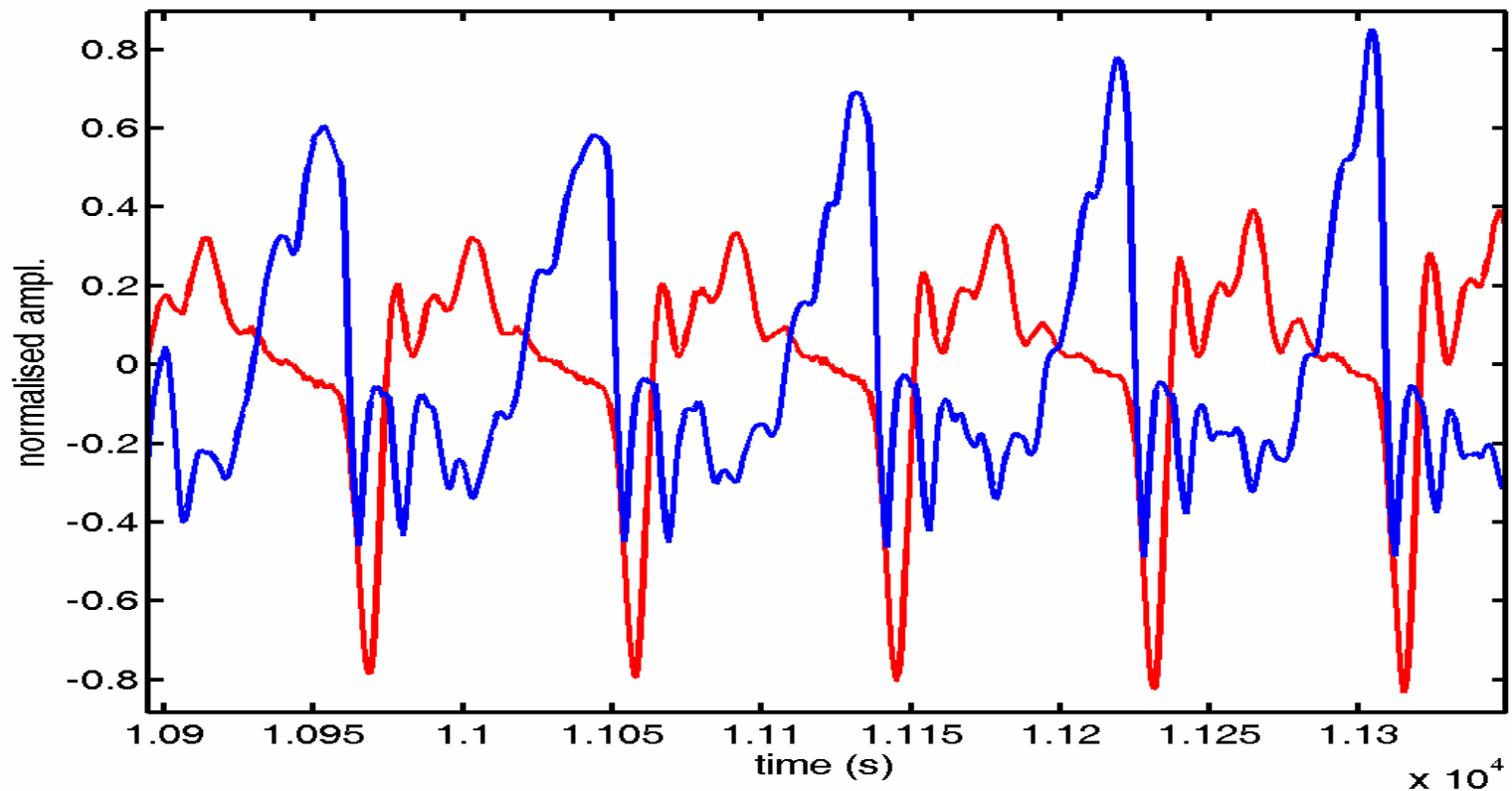
- ◆ Review

- (i) A more extensive test of evaluation can be introduced through using more detailed synthesis data derived from a time-varying filter function or through using model simulation
- (ii) A more thorough understanding of source-filter interaction is required

# Timing relationship between Speech(-), EGG (-)and DEGG (-)



# Timing relationship between volume velocity at the lips ( $v_{vl,-}$ ) and differentiated electroglottograph signal (DEGG, -)



Department of Electronic and  
Computer Engineering, University  
of Limerick

# Conclusion/Future Work

◆ New Methods should be developed to

- (i) maximally differentiate between normal and pathological voice
- (ii) specifically characterize the noise of turbulent origin

How?

- (i) Synthesis, real speech, modelling
- (ii) Multiple signals