

FCT PROJECT “ASSISTIVE REAL-TIME TECHNOLOGY IN SINGING”

**STEPS TOWARDS A
FREQUENCY-DOMAIN APPROACH TO
GLOTTAL INVERSE FILTERING**

Sandra Dias and Aníbal Ferreira

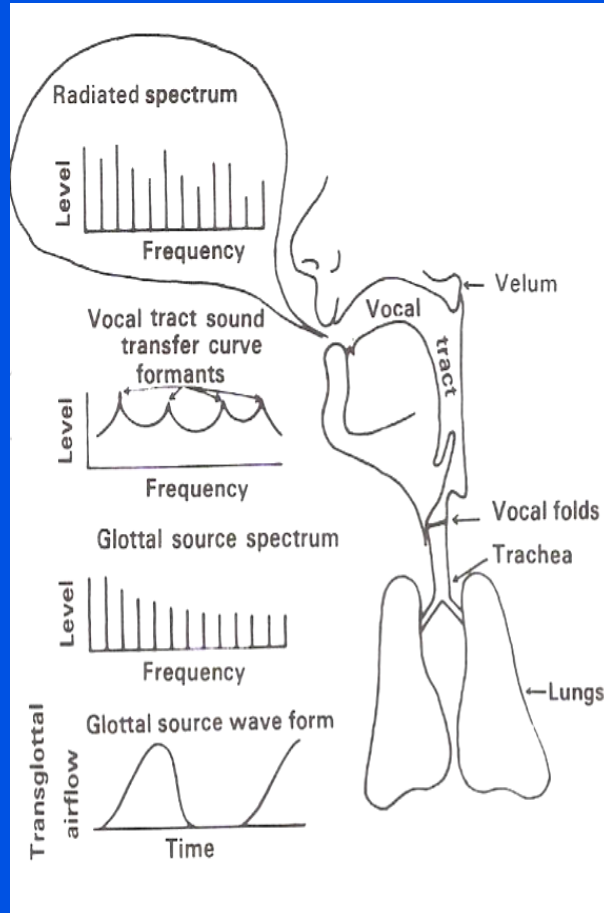
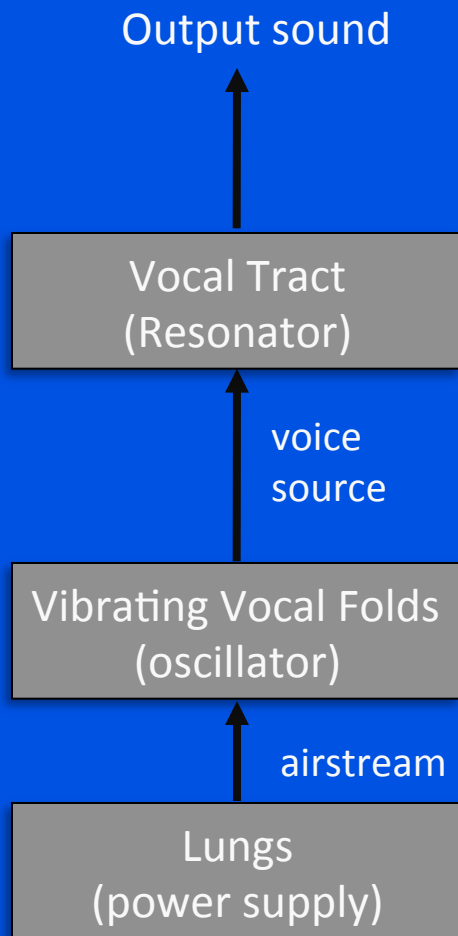
FEUP

TOPICS

- Source-Filter Model
- Inverse Filtering Techniques
- Frequency-domain approach to glottal inverse filtering
 - Integration in the frequency-domain
 - Speech signals acquisition and analysis
 - Future work

SOURCE-FILTER MODEL

- Generation of voice sounds

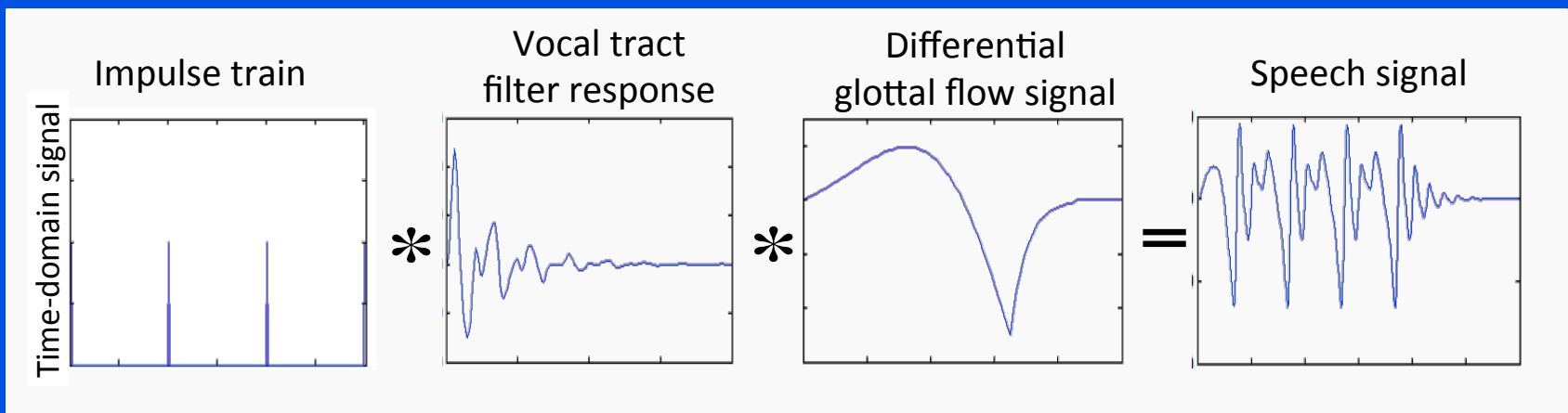
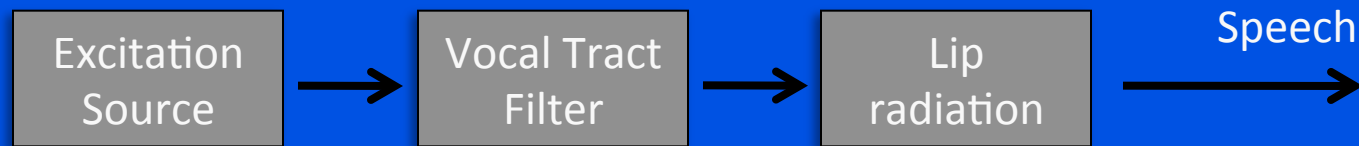


(adapted from: Johan Sundberg,
"The Science of the Singing Voice",
NIU Press, 1987)

SOURCE-FILTER MODEL

Based on Fant's source-filter theory:

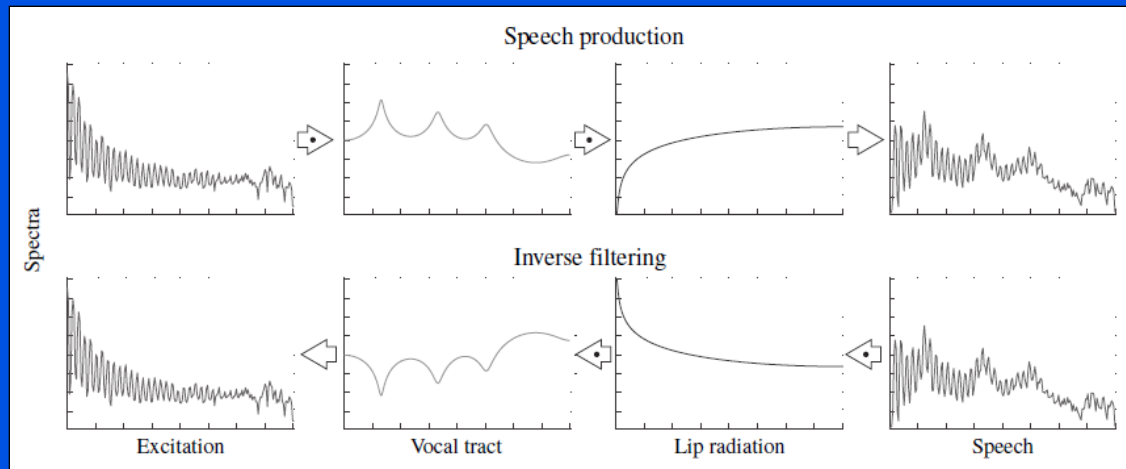
**The voice is the result of the convolution between
the excitation source and the filter system.**



INVERSE FILTERING TECHNIQUES

GOAL:

Estimate the glottal pulse by cancelling the spectral effects of the vocal tract and lip radiation on a speech signal.



Examples of Inverse Filtering Techniques:

- Iterative Adaptive Inverse Filtering (IAIF/PSIAIF) – by Alku, 1992
- Javkin *et al.* Method - 1987
- Zeros of the Z-Transform (ZZT) and Complex Cepstrum (CC) – by Bozkurt *et al.* , 2005

FREQUENCY-DOMAIN APPROACH TO GLOTTAL INVERSE FILTERING

- Included in **Task 5** of the Project:

Robust real-time glottal pulse estimation from running singing

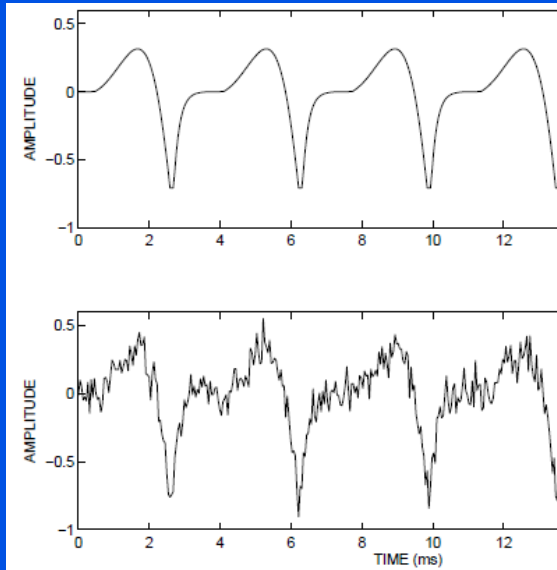
- **Principles:**

- Based on the Normalized Relative Delays (NRDs)
- The effects of lips/nostrils radiation can be accurately reversed in the frequency-domain
- Analysis – synthesis framework is able to implant any wave shape on a voiced sound, preserving the original fundamental frequency F_0 and time shift

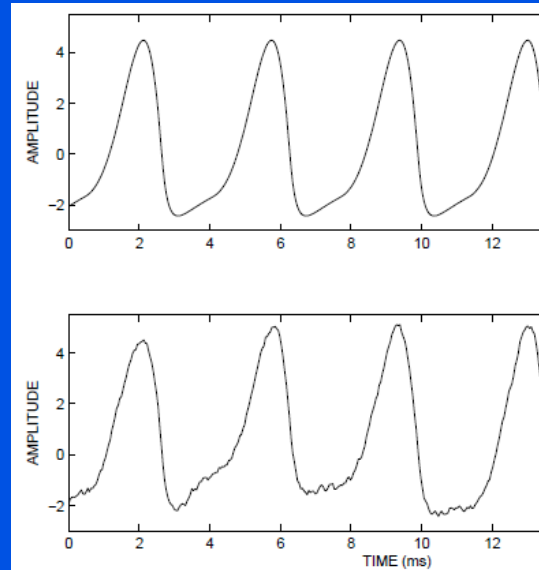
(Sandra Dias, Ricardo Sousa and Aníbal Ferreira. Glottal inverse filtering: a new road-map and first results. AFEKA: Speech Processing Conference, Israel, June 2011)

FREQUENCY-DOMAIN APPROACH TO GLOTTAL INVERSE FILTERING

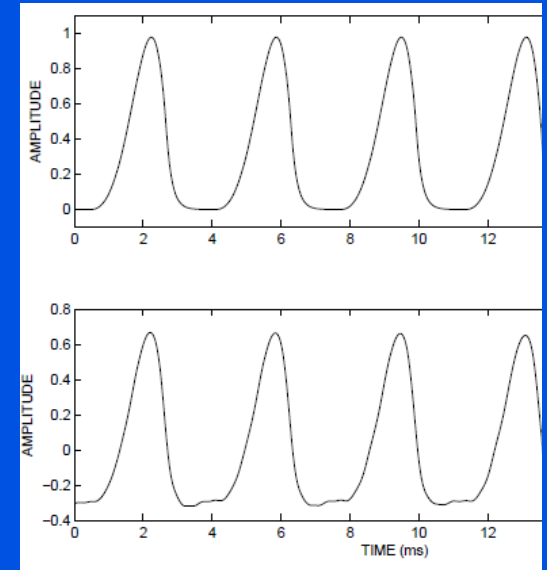
- Example:



Time representation of the derivative of the LF glottal pulse without noise (upper figure) and whit white noise at 9 dB SNR (lower figure).



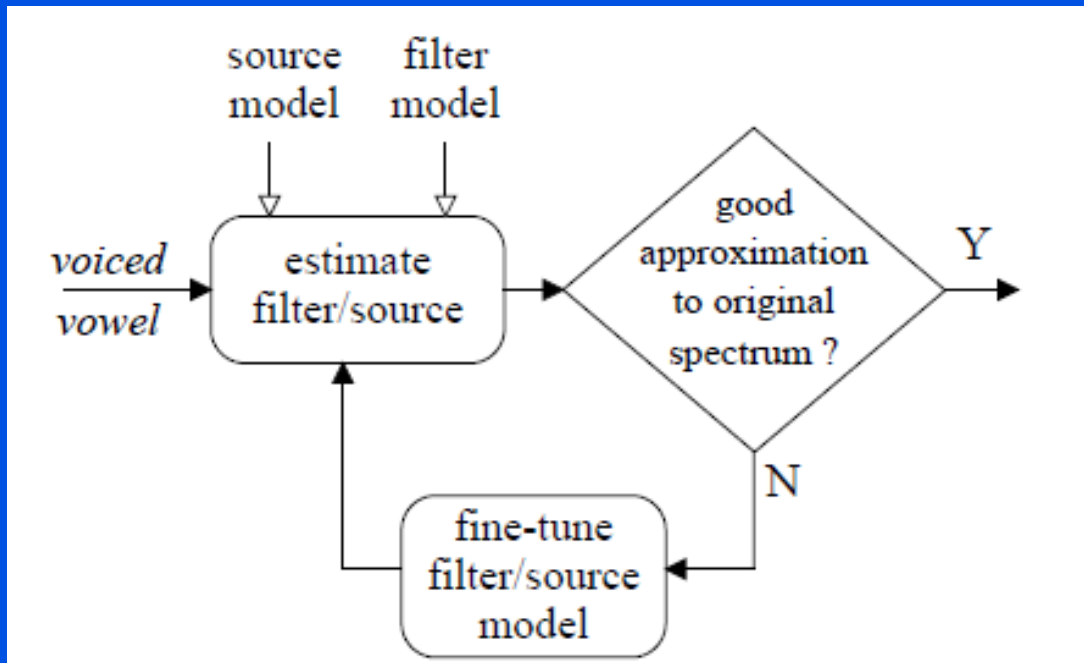
Corresponding results of a first-order time-domain integrator.



Corresponding results of integration in the frequency-domain.

FREQUENCY-DOMAIN APPROACH TO GLOTTAL INVERSE FILTERING

- Analysis-by-synthesis approach to filter/source fine-tuning



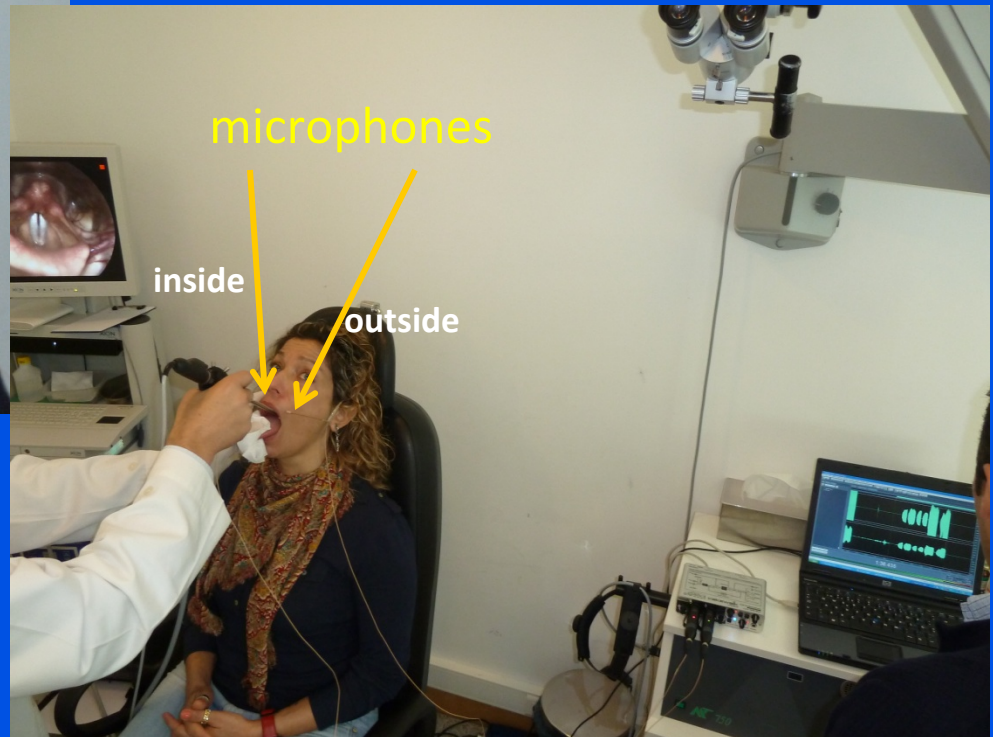
FREQUENCY-DOMAIN APPROACH TO GLOTTAL INVERSE FILTERING

- Speech signals acquisition



microphones

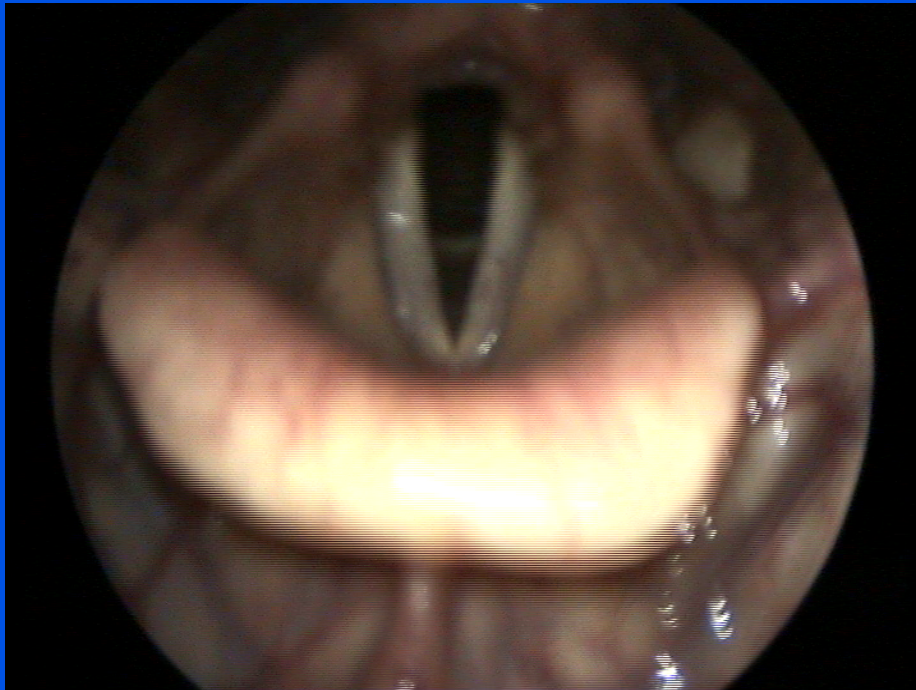
Capture, at the same time, the acoustic signal with a microphone as close as possible to the glottis and the speech signal outside.



➡ Estimate accurately the vocal tract transfer function.

FREQUENCY-DOMAIN APPROACH TO GLOTTAL INVERSE FILTERING

- Speech signals acquisition



Data:

- 3 female voices
- 3 male voices

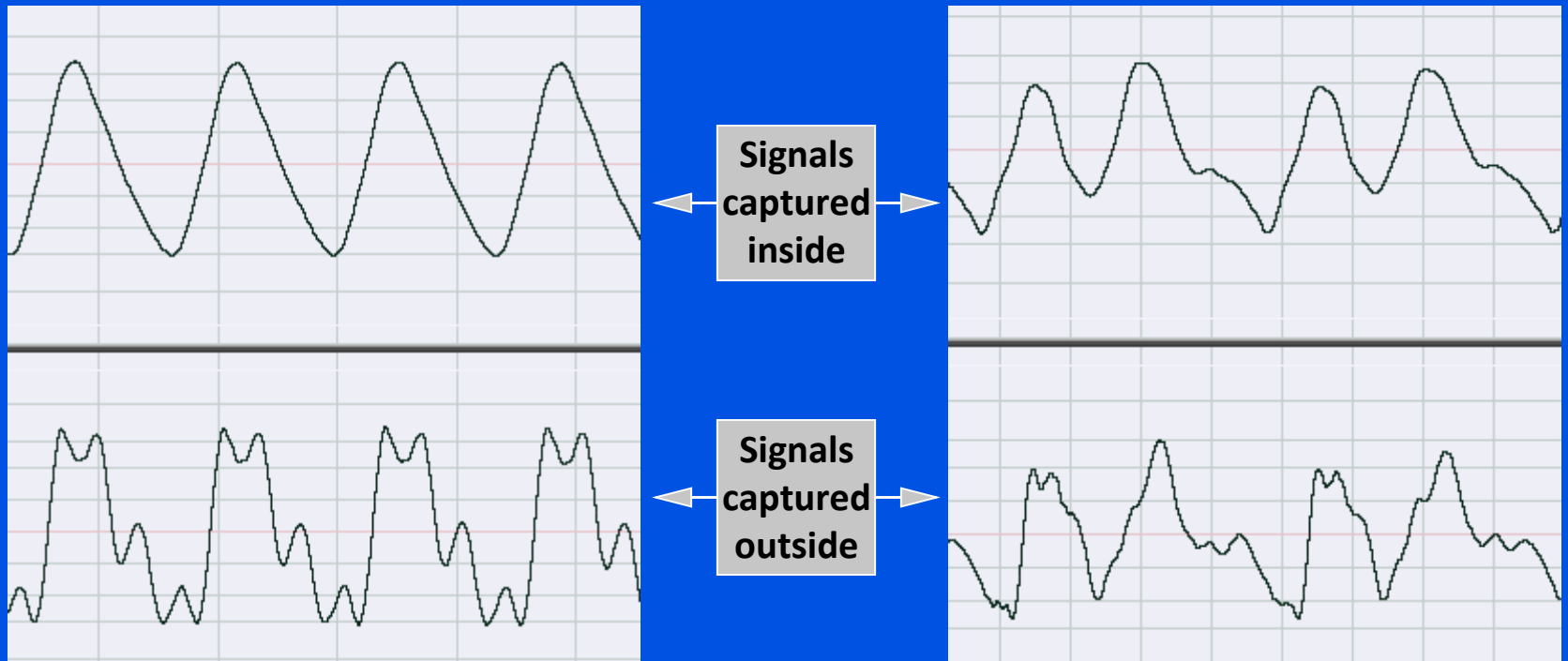
Vowels:

- /a/
- /i/

FREQUENCY-DOMAIN APPROACH TO GLOTTAL INVERSE FILTERING

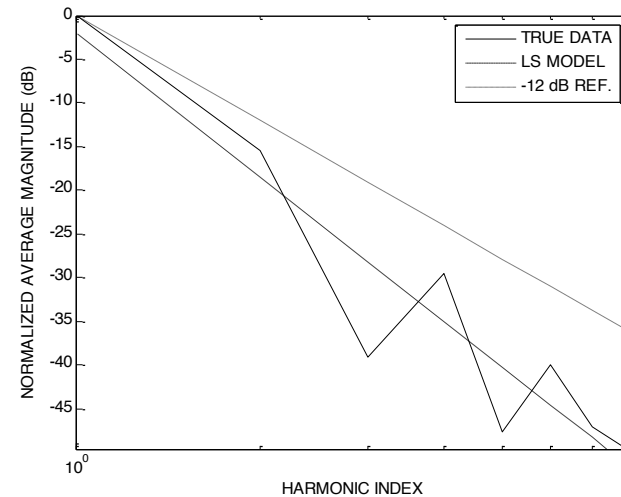
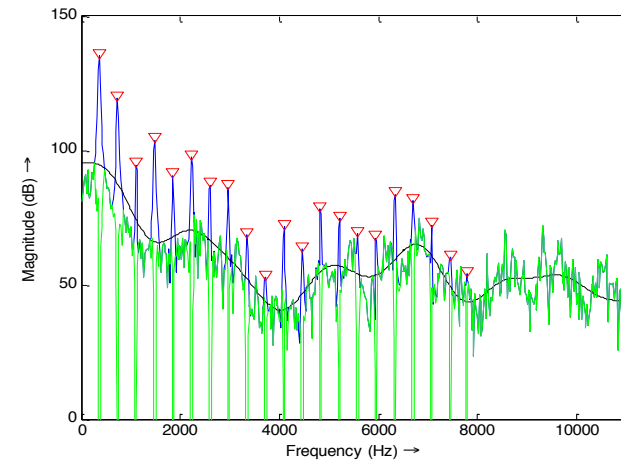
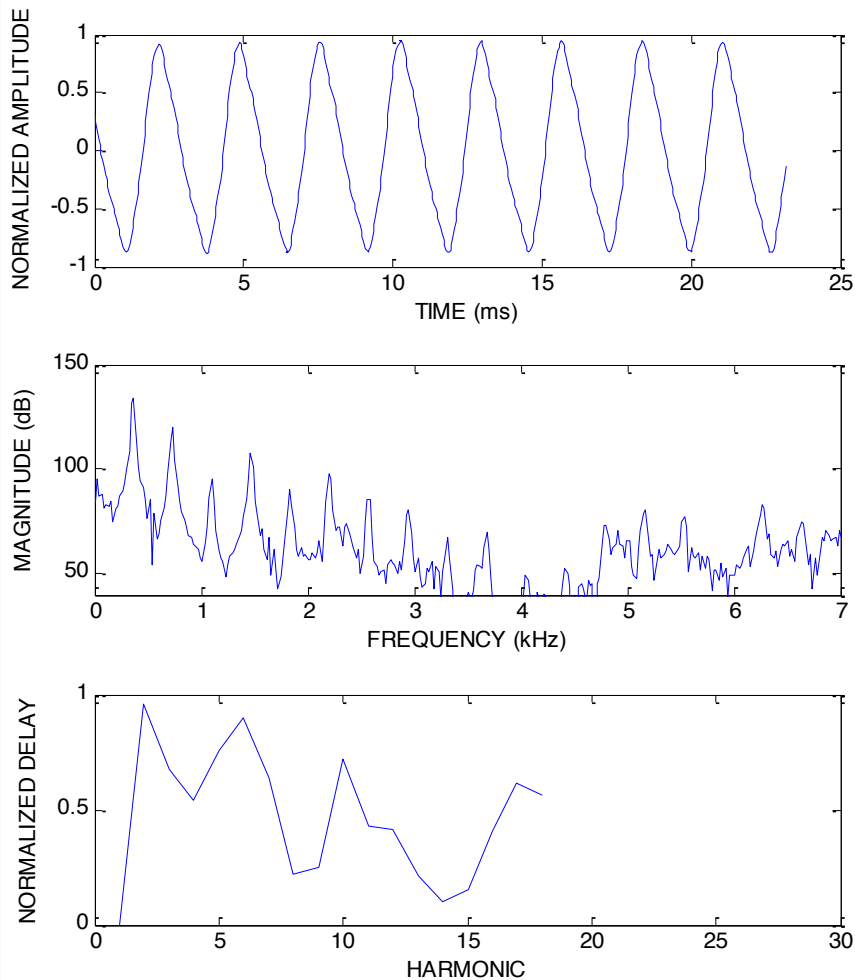
- Speech signals analysis

Examples (vocal /i/)



FREQUENCY-DOMAIN APPROACH TO GLOTTAL INVERSE FILTERING

- Speech signals analysis



Decay:
16 dB/oct

FREQUENCY-DOMAIN APPROACH TO GLOTTAL INVERSE FILTERING

- Future work

- Analyse and compare the speech signals captured inside and outside, in frequency-domain, in order to:
 - obtain a model of the glottal pulse per subject and vowel
 - estimate the vocal tract transfer function (or frequency response) per subject and vowel.
- Develop an algorithm to estimate the glottal pulse using only the voice signal captured outside the mouth
- Submit a paper to Interspeech (April, 2012) with the results.