Comparison of LPC analysis and impedance vocal tract measurements

M. Kob, J. Stoffers, Ch. Neuschaefer-Rube

Chair of Phoniatrics and Pedaudiology, RWTH Aachen University – University Hospital Aachen

Pauwelsstr. 30, 52074 Aachen, Germany

1 Introduction

The acoustic measurement of the resonances in the space between vocal folds and lips, the vocal tract, allows a non-invasive, objective analysis of the spectral energy distribution for different articulatory cases. Whereas a conventional LPC analysis is successful only when applied to more or less stationary voice signals, an external excitation with subsequent measurement of the vocal tract impedance at the mouth can give reliable results when the voice signal is not stable or even missing.

This contribution adresses the problem how to compare results from impedance measurements with LPC measurements. For a set of normal speakers, both measurements have been performed simultaneously. Based upon the evaluation of the LPC curves, similar values for resonance frequency, amplitude and bandwidth were derived. In a study of 81 normal speakers the results from the evaluations are compared.

2 Method

The measurement set-up and the software used to evaluate the measurements allow a sequence of measurements consisting of three parts: the LPC analysis of the voice signal, the impedance measurement during phonation, and the impedance measurement without phonation. The concept for measurement of the vocal tract impedance at the mouth, VTMI, is described in detail in [2], and the procedure of clinical measurements is described in [4].

Measurements were performed in a group of 35 female and 46 male healthy speakers, using a simplified set-up without velocity sensor. One reason is the problem of clipping which can occur in the velocity sensor at high sound velocities. Comparisons between this set-up and the original 2-sensor set-up showed that results from both methods yield comparable results for the performed task.

A sample rate of 22050 Hz was chosen, and subsequent evaluations were limited to the frequency range 100..5000 Hz.

2.1 LPC measurement

The “linear predictive coding” method (LPC) is a well-established method to identify the formant structure of a voice signal. The LPC curves are derived from a windowed part of the voice signal.

For the calculation of the LPC curves 28 coefficients were used to achieve a rather high pole density but not too many wrong identifications of formants. A Hamming window of 9525 samples was applied to the voice signal. The onset of the voice signal was automatically discarded.

2.2 Impedance measurement

All impedance measurements were performed using a linear swept sine from 250 Hz to 6000 Hz with a duration of 0.74 seconds. The signal-to-noise ratio was improved by application of a symmetric Hanning window of 15 ms length.

3 Evaluation and normalisation

3.1 Parameters

From the LPC curves the formant frequency, the formant bandwidth (3 dB decay), and the relative amplitude of the formants were calculated. The amplitude difference was calculated with respect to the highest formant in the frequency range 150..5000 Hz.

The resonances frequencies of the impedance curve were calculated from the minima of the impedance function $Z(f)$. Since no measures for the amplitude and bandwidth of the resonances could be directly derived from the impedance function, new measures had to be calculated: the slope between a local minimum and maximum near a resonance was used, as well as the amplitude difference between these points. These values were normalised as well by division of the local amplitude difference by the difference between the absolute maximum and minimum in the frequency range of the evaluation.

3.2 Normal ranges

From each three LPC measurements of the six phonemes /ɑ:/, /æ:/, /i:/, /o:/, /u:/, /l:/ the first four formants were evaluated with respect to mean values standard deviations of the formant frequencies, amplitudes and the bandwidths.

From the resonance and LPC curves the above parameter were automatically calculated and stored in an XML

Models and analysis of vocal emissions for biomedical applications. 4th international workshop. October 29-31, 2005 – Firenze, Italy.
Edited by C. Manfredi. ISBN 88-8453-320-1 (online)

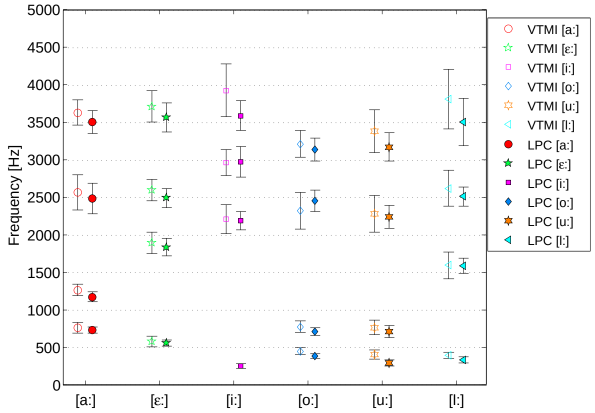

Figure 1: Male reference group: Comparison of formants and resonances from LPC and impedance analysis (automatic analysis, error bar corresponds to ± 1 standard deviation

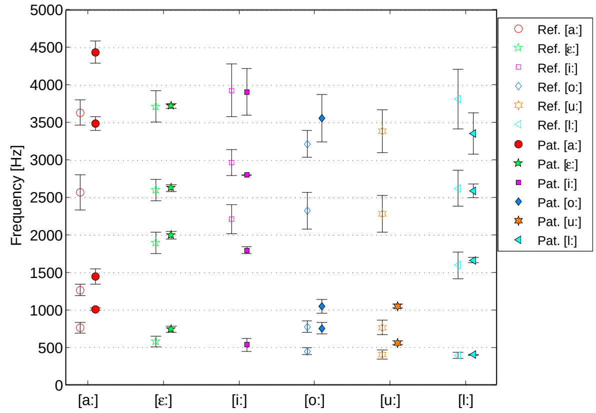

Figure 2: Comparison of vtmi measurement results of the reference group and a male patient with supraglottal tumour

tree. Whereas the formant values of the LPC curves were easily calculated and simply copied to the tree, the resonance evaluation was more complicated. The number of detected resonance frequencies was much higher than the desired number of about four resonances between 100 and 5000 Hz. The reason is the presence of small variations in the impedance curve which can easily be misinterpreted as resonances. A selection of the four most probable resonances was achieved by weighting the resonance features according to scores for relative amplitude between maximum and minimum, slope between neighboured extrema, and the relative distance to the LPC formant frequency. For the four resonances with the highest scores the related bandwidth and frequencies were calculated.

In Figure 1 the frequencies of the formants (LPC) and resonances (VTMI) are plotted for the male normal group. The results indicate a high comparability of the results from the LPC and from the impedance analysis method, both with respect to the mean frequencies as well for the standard deviations. The missing first resonance for the vowel /i/ is caused by the weak excitation of the vocal tract with the swept sine below 250 Hz.

4 Case study

Pathological alterations of the vocal tract configuration can lead to a change in the resonance structure of the vocal tract [5]. Examplarily for the results from an ongoing medical study we describe the application of the impedance method to the acoustic vocal tract characterisation of a 82 years old male patient (1.65 m, 55 kg), status after tonsillectomie.

Diagnosis: Expanded malignant tumour in the lower pharynx reaching down to the larynx on the left side (lower pharyn-larynx carcinoma, T4).

CT and endoscopy results: We observe a tumorous process in the left lower part of the pharynx, beginning on the lingual bone level, extending into the larynx over a length of ca. 4.5 cm, crossing the median line ventrally. A partial corrosion of the thyroid cartilage by the tumour is seen. The endoscopy of the airways shows a big exulcering tumour, extending from the vocal cord level up over the ventricular folds side with infiltration of the laryngeal epiglottis area. Tongue basis, tongue directed epiglottis area and sinus piriformis visually clear of tumour.

Particular aspects for LPC and impedance measurements: The upper and lower jaw is toothless. A distinctive disphonia (strong „hoarseness") is heard, the patient phonates with great effort and is not able to hold phonation for longer. The voice has a very high pitch. Figure 2 shows the results of the resonance frequency analysis of this patient compared to the standard values of the male reference group. It is evident that in several phonemes the resonances are shift to higher frequencies. For the phonemes /o/ and /u/, the third and fourth resonance are afflicted with high energy loss and have such low amplitude that they cannot be surely detected.

5 Discussion

Investigation of a reference group shows differences both in formant- and in resonance characteristics between male and female as well as between the two methods. The frequencies of LPC analysis and impedance measurements at the same subject are stongly correlated. Concerning the amplidue values, the impedance measurements shows a systematic lowering of the resonance amplitudes at frequencies below ca. 800 Hz. The deviation found in the patient's increase of the resonance values in the phomenes /ɑ/ and /æ/ indicates an acoustical decrease of the length of the vocal tract by a tumour caused narrowed supraglottal space. The absence of higher formants in /o/ and /u/ could be caused by a higher sound-absorption of the alterated tissue. A systematic examination of further patients with similar diseases is planned and should give additional clues in view of a correlation of physiological and acoustical properties of the vocal tract.

References

[1] M. Kob, Ch. Neuschaefer-Rube (2001): A method for measurement of the vocal tract mouth impedance. Conference CD-ROM, 2nd Int. Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications – MAVEBA, file "papers/26.pdf" 1-6.

[2] M. Kob und Ch. Neuschaefer-Rube (2002): A method for measurement of the vocal tract impedance at the mouth. Medical Engineering & Physics, 24, 467-471.

[3] M. Kob, Ch. Neuschaefer-Rube (2003): Acoustic analysis of overtone singing. Proceedings 3rd Int. Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications – MAVEBA, 187-190.

[4] M. Kob, J. Stoffers, M. Lievens, R. Katzer, Ch. Neuschaefer-Rube (2004): Application of impedance measurements for the diagnosis of articulatory dysfunction. Proceedings CFA/DAGA 2004, 1145.

[5] S. Koppetsch und K. Dahlmeier (2003): Funktionelle Störungen der Artikulation bei intra-oralen Tumoren – eine prä- und postoperative Langzeitstudie. Sprache–Stimme–Gehör, 155-160.