Glottal Source Parameterization: A comparative study

Ixone Arroabarren, Alfonso Carlosena

Universidad Publica de Navarra
Dept. Electrical and Electronic Engineering
Campus de Arrosadia, E-31006 Pamplona, Spain

ixone.arroabarren@unavarra.es, carlosen@unavarra.es

Abstract

In this paper two parameterization methods are jointly studied and compared, namely the Normalized Amplitude Quotient (NAQ) and the Parabolic Spectral Parameter (PSP), which are different ways of glottal source characterization, in the time and frequency domains, respectively. First, a theoretical study for the derivative of the glottal source is carried out based on the LF model, and it is proved that the two above mentioned parameters represent globally the glottal source, in the sense that they are dependent on the three main characteristics of the glottal source: the open quotient, the asymmetry coefficient and the spectral tilt. Considering this implicit dependence it is shown that both parameters are correlated. Last, natural speech signals, from the workshop database, are analyzed and the performance of both parameters is analyzed in terms of their robustness.

1. Introduction

In the study of speech production the inverse filtering of voiced speech signals is a widely used technique for the analysis of the glottal source excitation. This technique provides a noninvasive method to study voice quality. Inverse filtering is usually accomplished in two steps: In the first one the glottal source waveform is estimated, and in the second this signal is parameterized in a few numerical values which characterize a phonation type. This whole analysis can be practically implemented in several ways, which could be grouped in two main classes:

- In the first group, the two steps are combined in a single algorithm, such as for instance in [1, 2, 3]. There, a mathematical model for glottal source and an autoregressive (AR) model for the vocal tract response are considered, and then authors estimate simultaneously the vocal tract response and the glottal source model's parameters. In this way, the glottal source model's parameters parameterize a given phonation type. Several different algorithms follow this structure, but all of them are invariably time domain implementations that require Glottal Closure Instant Detection (GCI) [4], and, as such suffer from a high computational load, what makes them very cumbersome.
- The second group of procedures breaks down the glottal source estimation problem into two stages, as explained before, and regarding the first step, inverse filtering techniques are proposed, [5, 6, 7]. These algorithms remove from the speech signal the glottal source effect and the vocal tract response is obtained by either Linear Prediction, or alternatively Discrete All Pole modeling [8], which avoids the fundamental frequency dependence of the former.

Afterwards, the estimated glottal source excitation is described by a few parameters. This has been done in different ways: some researchers have considered again a mathematical model for the glottal source, and an optimization algorithm is proposed to fit this model to the inverse filtered glottal source waveform [9, 10]. These approaches require synchronization between the mathematical model and the glottal source waveform, and their computational load is also high.

Alternatively, and in a simpler way, the parameterization of the glottal source has been developed measuring several sensible time domain parameters such as: Open Quotient, Asymmetry Coefficient, or Closing Quotient. This is the most widely used parameterization method because these parameters can be defined without knowing the absolute flow values of the glottal volume velocity waveform. However, the direct measurement of these parameters is not usually an easy task because of the noise and the formant ripple typical of the inverse filtered signals [11]. In order to avoid this problem several efforts have been driven towards the search of spectral correlations of time domain parameters [12, 13].

In addition to the above mentioned approaches, new different parameters have been proposed to characterize the glottal source: the Normalized Amplitude Quotient (NAQ) [14], and the Parabolic Spectral Parameter (PSP) [15]. NAQ and PSP are by definition independent of signal power and of pitch. Both, allow to separate different phonation types. Also, it has been proved that the NAQ is correlated with the Closing Quotient [16], in such a way that the NAQ is proposed as a more robust way for measuring this timing feature.

The goal of this paper is to analyze the relation between these two normalized parameters and the three main, and most widely used, characteristics of the glottal source: Open quotient (OQ), Asymmetry coefficient (AC) and Spectral Tilt (TL). This could help to understand which of these parameters controls the two global parameters. Also, as both have been proposed as amplitude and fundamental frequency normalized parameters respectively, it would be interesting to know if both have some correlation, or on the contrary, they measure different features of the glottal source.

The organization of the paper is as follows: First, the calculation of the NAQ and the PSP is reviewed. Then, the LF model is consider to extract empirical relations between these two parameters and OQ, AC and TL. This analysis will serve to establish the relationship between NAQ and PSP. In section three, the most critical aspects of the practical measurement of these parameters will be commented on. Finally, natural speech signals, from the workshop database, will be analyzed and a comparison of both parameters will be developed.

2. Glottal Source Parameterization

In this section we will shortly review the definition of NAQ and PSP

2.1. Normalized Amplitude Quotient

As it has been said before, the NAQ parameter was proposed in [14] as a timing measurement of the glottal source based on an amplitude quotient, in such a way that the direct measurement of time references was avoided. To see the rationale of this relationship, a simplified signal model is considered for the glottal source, as shown in Fig. 1.a:

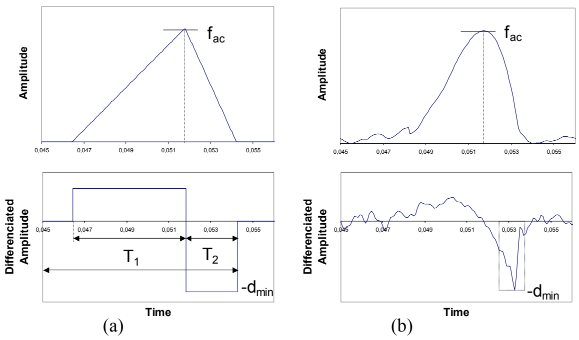

Fig.1. Definition of the NAQ. (a) Triangular-shaped glottal flow pulse. (b) Natural speech inverse filtered glottal source excitation

For the simplified glottal source waveform of Fig.1.a, the closing quotient can be expressed as:

$$CQ = \frac{T_2}{T} = \frac{f_{ac}}{T \cdot d_{\min}} \qquad (1)$$

It is evident that for this kind of signals the closing quotient can be calculated as an amplitude quotient instead of a time relationship, because f_{ac} represents the area of the rectangle in the derivative of the glottal source. In the case of the real glottal source waveform, expression (1) is not exactly true, but this intuitive idea allows the definition of the Normalized Amplitude Quotient as:

$$NAQ = \frac{f_{ac}}{T \cdot d_{\min}} \qquad (2)$$

It is possible to see that the definition of the NAQ is equal to the right side of (1). Regarding the relationship between the CQ and the NAQ in real signals, in [16] it is shown that both parameters are strongly correlated, and the NAQ provides a more accurate, consistent, and robust measurement than the CQ. It is also important to note that this new parameter is normalized in amplitude and thus signal power independent. By virtue of its correlation with CQ it is also independent on signal period (and thus frequency), This characteristic allows comparisons among different fundamental frequencies.

2.2. Parabolic Spectral Parameter

Unlike the time domain parameters, the parabolic spectral parameter purports to represent a spectral characteristic of the glottal source. In Fig. 2.b the spectrum of the glottal source waveform of Fig. 2.a is presented where a parabolic fitting is overimposed on the low frequency . Thus, in [15] this spectral feature is used as a quantification of the glottal source waveform:

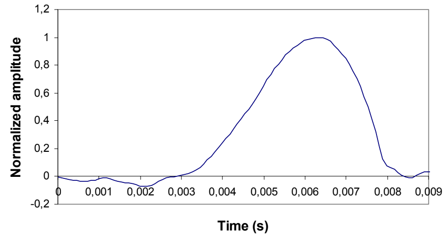

Fig. 2.a. Natural speech glottal source waveform

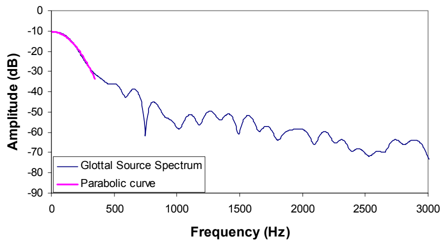

Fig. 2.b. Spectrum of the glottal source waveform of (a)

It must be pointed out that the spectrum shown in Fig.2.b is a pitch synchronous spectrum, in other words, only one period of the glottal flow has been considered. Such period is taken form the minimum amplitude value of one period of the glottal source, to the same time event in the next period.
Regarding Fig.2.b the parabolic portion of the curve has been calculated according the mathematical expression:

$$y = ax^2 + b \qquad (3)$$

where, b controls the maximum of the parabola, and a accounts for its width. Thus, the PSP is based on the a parameter of the parabolic curve since it does not depend on the amplitude of the signal. However this value depends on the fundamental frequency, since if the fundamental period decreases the parabola stretches, and thus a increases. To avoid this dependence, the spectrum of the rectangular window is calculated, and as it is known, it has also a low frequency parabolic shape. Thus, the a value obtained for the glottal source is normalized to that obtained for the spectrum of the rectangular window whose length is the fundamental period of the glottal source, a_{max}:

$$PSP = \frac{a}{a_{\max}} \qquad (4)$$

For a practical implementation, the parabolic parameters, a and b, are calculated in such a way that the least squares error between the low frequency spectrum and the parabolic curve is minimized. Besides, there is another unknown term, that is the considered frequency range of this estimation. As this also depends on the glottal source waveform, a and b are calculated by an iterative procedure, for a given pitch synchronous glottal source spectrum: the considered frequency range is increased until the error exceeds a given limit. This level is an important parameter of the algorithm. The details of the algorithm are completely described in [15].

It is important to remark here that PSP by definition, does not depend on either signal power or fundamental frequency, This property, shared by NAQ, gives a clue of a likely relationship between these two parameters.

3. NAQ and PSP in the LF model

Once the glottal source parameters have been presented, it is interesting to study their relationship with other important glottal source parameters: the open quotient, the asymmetry coefficient and the spectral tilt. To this end, we have used a well known mathematical model for the derivative of the glottal source: the LF model. This model was proposed in [17], and according to it, the expression for the derivative of the glottal flow is:

$$u'(t)=\begin{cases} E_o e^{\alpha t}\sin\omega_g t & 0\le t\le t_e \\ \dfrac{-E_e}{t_a\varepsilon}\left(e^{-\varepsilon(t-t_e)}-e^{-\varepsilon(t_c-t_e)}\right) & t_e\le t\le t_c \end{cases} \qquad (5)$$

The eight direct synthesis parameters, (E_o, α, ω_g, E_e, t_a, ε, t_e, t_c), are determined by three independent timing parameters, R_g, R_k, R_a, besides fundamental frequency and the amplitude parameter, E_e. These timing parameters are related with the meaningful time instants, i.e.: T_e, which represents the time instant when the derivative of the glottal source has a negative peak, T_p, (which represents the instant when the glottal flow is maximum), and T_a, which is related with the time constant of the closing phase [17]:

$$R_g=\frac{T_o}{2\cdot T_p} \qquad R_k=\frac{T_e-T_p}{T_p} \qquad R_a=\frac{T_a}{T_o} \qquad (6)$$

These three parameters can be further related to the open quotient, asymmetry coefficient, and the normalized cut-off frequency of the spectral tilt as:

$$O_q=\frac{1+R_k}{2\cdot R_g} \qquad \alpha=\frac{1}{R_k+1} \qquad f_t=\frac{1}{2\pi R_a} \qquad (7)$$

Given a set of values of the timing parameters $\{R_g, R_k, R_a\}$, or a combination of $\{O_q, \alpha, f_t\}$, a system of non linear equations must be solved to generate the direct synthesis parameters of the model according to expression (5), what makes this model cumbersome to be used.

Therefore, for the analysis at hand, the three independent parameters have been simply numerically varied in order to generate different glottal pulse waveforms, and then the NAQ and the PSP have been measured. In Fig. 3 the relationship between NAQ and the three parameters is graphically depicted:

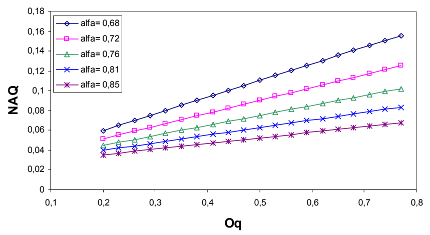

Fig 3.a. NAQ versus Open Quotient. f_t= 7

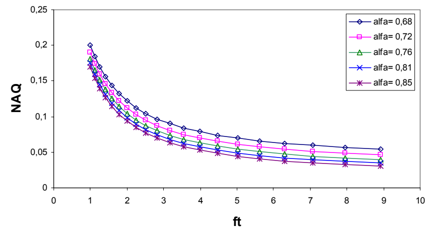

Fig 3.b. NAQ versus normalized cut-off frequency of the Spectral Tilt. O_q= 0,2

In Fig. 3.a it is possible to see that as the open quotient increases, NAQ increases. Also, the asymmetry coefficient has been changed and, as it is shown, that as it increases the NAQ decreases. On the other hand, in the case of the Spectral Tilt, as the cut-off frequency increases, or the closing phase is sharper, the NAQ decreases.

In the case of the PSP, it is important to note that the maximum error considered for the PSP calculation has been 1E-2:

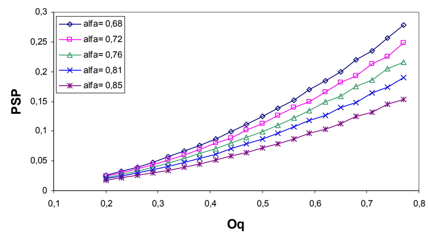

Fig 4.a. PSP versus Open Quotient. f_t= 7

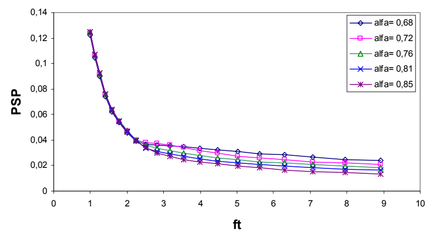

Fig 4.b. PSP versus normalized cut-off frequency of the Spectral Tilt. O_q= 0,2

Comparing Fig. 3 and Fig.4, it is evident that both parameters are directly related because their dependence on O_q α and f_t, is very similar, thought not exactly the same. In the case of the NAQ in the LF model it is possible to obtain the mathematical expression of the parameter in terms of the direct synthesis parameters given by (5):

$$NAQ=\frac{-\omega_g\left(e^{\frac{\alpha\pi}{\omega_g}}+1\right)}{\left(\alpha^2+\omega_g{}^2\right)\sin\omega_g t_e\cdot e^{\alpha t_e}\cdot t_o} \qquad (8)$$

Looking at (8), it is clear that the NAQ is a global parameter because it depends on all of the characteristics of the glottal pulse. Even though in (8), the NAQ is expressed in terms of the parameters of the opening phase, the value of α depends also on the parameters of the closing phase, because of the non linear set of equations that must be solved to generate the direct synthesis parameters from the timing parameters. In the case of the PSP parameter, it is difficult to extract a counterpart expression of (8), but in [18] a unified study of the glottal source spectrum has been developed, .There, it is shown that the low frequency spectral area depends on the three parameters, (O_q, α, f_l) in the way shown by Fig. 4.a and Fig.4.b. This analysis further supports the empirical evidence that NAQ and PSP are strongly correlated parameters

4. Practical considerations

In section 3 the NAQ and the PSP have been analyzed from a theoretical point of view. Since no practical aspects have been mentioned (such as sampling frequency, extraction of a single fundamental period of the glottal source, and the maximum error in the PSP calculation) in this section of the paper the effect of these three factors will be analyzed.

4.1. Sampling frequency

Even though NAQ is an amplitude quotient measurement, it will be affected by the sampling frequency. Looking at Fig. 1, the NAQ is based on the values of the maximum value of the glottal flow, f_{ac}, and the negative peak of the derivative of the glottal flow, d_{min}. The measurement of f_{ac} does not depend too much on the value of the sampling frequency because the glottal waveform is rather smooth, at least with medium or high O_q. However, in the case of the derivative of the glottal flow d_{min} is a peak value, therefore, as sampling frequency decreases the peak value will not be correctly estimated. Thus, for a correct calculation of NAQ, a high sampling frequency is mandatory, or a more elaborate interpolation or estimation algorithm.

In the case of the PSP, the dependence on the sampling frequency is not so clear and, intuitively, one could think that it does not depend on this factor because is a low frequency feature. However the problem comes from a step that must be taken earlier, being the extraction of a single fundamental period of the glottal source. Such factor will be analyzed in the next subsection.

4.2. Extraction of one fundamental period

For the calculation of the PSP, the glottal source spectrum is calculated. To that end, one period of the signal must be considered, and this must be taken from the minimum amplitude value of one period of the glottal source, to the same time instant in the next period. This condition must be fulfilled because the FFT algorithm [19] is used for the spectrum calculation. When the glottal source is considered as explained before, the FFT calculated spectrum matches the continuous time Fourier Transform spectrum of the glottal source waveform, at least in a low frequency region. However, when the FFT is used, the spectrum is also affected by the window spectrum, being the rectangular window in our case. Extracting the glottal period from the minimum value to the next minimum value, the window's effect is compensated and thus the low frequency spectrum matches its continuous time Fourier Transform. In any other case, the window's spectrum would affect the obtained spectrum as it is shown in Fig.5. It is shown that the low frequency parabolic shape is lost with a wrong single period extraction.

It is interesting to remark that although PSP is a spectral parameter it has a very strong time domain requirement, specially in natural speech signals where noise and formant ripple could make it a difficult task to extract the right period of the glottal flow.

Coming back to the sampling frequency effect, it could also affect the period selection in the case of a low sampling frequency. However, the sampling frequency requirement is not as critical as for the NAQ. On the other hand, the NAQ will not be affected by the period extraction, because no time reference is used.

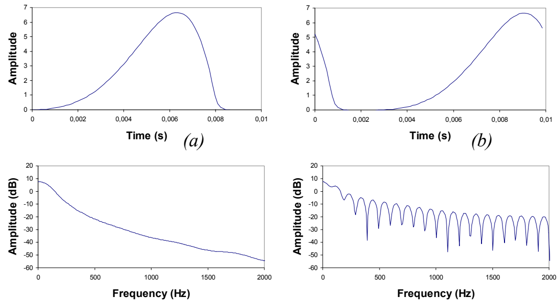

Fig. 5. *(a)* Glottal source FFT spectrum for a right period selection. *(b)* Glottal source FFT spectrum for a wrong period selection

4.3. Maximum error in the PSP calculation

When PSP was presented in [15], the parabolic curve fitting was accomplished as an iterative algorithm, because the parabolic shape low frequency range of the spectrum of the glottal source changes with the glottal source characteristics and the fundamental frequency. Therefore, the considered frequency range is increased until a maximum error is reached. The recommended value of the error is 1E-2, because for this value, the frequency range considered in the case of the parabolic fitting for the rectangular window's spectrum falls into the main lobe. However, it is interesting to know which is the effect of changing this value:

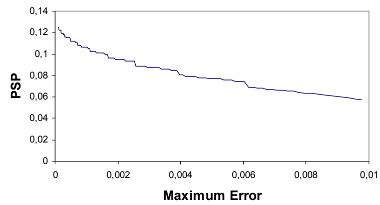

Fig. 6.a. Dependence of the PSP on the maximum error

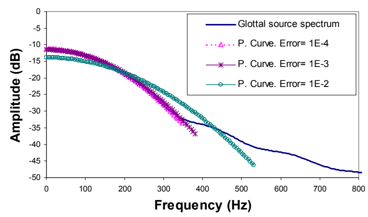

Fig. 6.b. Parabolic curve estimation for several maximum error values

Looking at Fig. 6.a, as the maximum increases the PSP decreases. This is evident looking at Fig. 6.b., where the parabolic fitted curve has been represented for three maximum error values. As the maximum error increases the considered frequency range increases, what makes the *a* parameter of the fitted parabola decrease. Moreover, as the tolerated error increases, the considered spectrum is not so close to the parabolic curve, what would reflect in a loss of resolution for different phonation discrimination.

Based on this effect, one could think that choosing a lower maximum error is the best option, however, sometimes the low frequency spectrum of the glottal source has not a perfect parabolic shape, because the glottal period has not been properly extracted, as in Fig. 5.b.. In this kind of spectra, a lower tolerated error would drive to a wrong PSP value. Therefore, these two factors have to be taken into account when the maximum error is selected.

5. Natural speech analysis

In the previous sections, the two global glottal source parameterizations have been reviewed, their critical practical factors have been identified, and the correlation between them has been demonstrated. To support the analysis, the LF model of the derivative of the glottal source was used. In this section we will prove this correlation making use of real natural-speech signals. NAQ and PSP will be calculated for various voice qualities. The analyzed voice material corresponds to speech and electroglotographical signals available in the database "*Speech database 1 (American English)*", whose author is Christophe d'Alessandro, and is provided for the meeting VOQUAL'03.

For this study, only speech recordings are considered. These are stored in *.wav* format, recorded with a bandwidth of 20 KHz, which was further reduced to 6 KHz. In order to remove low frequency ambient noise the signals are filtered by a high pass linear phase FIR filter, and their cut-off frequency has been fitted to be a 75% of the fundamental frequency.

In order to parameterize the glottal source pulse for these signals, an inverse filtering algorithm is first applied in order to obtain the glottal source signal. First, every period of the glottal waveform has been extracted, and finally the NAQ and the PSP have been obtained for each period of the signal.

5.1. Inverse filtering

We have applied the inverse filtering technique proposed in [7]. This technique makes use of a more accurate spectral model of the source. This spectral model is based on the KLGLOTT88 time domain model [20], which is decomposed into two blocks: a first one were a basic voicing waveform is generated, and that is controlled by the Open quotient; and a second one were the Spectral Tilt is included as a first order low pass filter.

The inverse filtering approach uses the glottal decomposition of the KLGLOTT88 model, in such a way that the basic voicing wave is removed from the speech signal using the following function:

$$G(f) = \frac{27AV}{2O_q(2\pi f)^3}\left[\frac{je^{-j2\pi f O_q T_o}}{2} + \frac{1+2e^{-j2\pi f O_q T_o}}{2\pi f O_q T_o} + 3j\frac{1-e^{-j2\pi f O_q T_o}}{\left(2\pi f O_q T_o\right)^2}\right] \quad (9)$$

and the spectral tilt and the vocal tract response are combined in an (N+1)th order all pole system. The algorithm follows the next steps:

1. The short time spectrum is calculated. The window length must be long enough to detect the spectral peaks of the signal. A 3,5 o 4 period time length is a good reference, and the hamming window will provide a good spectral resolution. As a result, a set of spectral points is obtained: $\{(f_i(Hz), A_i(dB))\}$
2. The basic voicing effect is removed from the spectral peaks of the signal, using (9). Once the spectral peaks have been detected, $20\log|G(f_i)|$ is calculated for every peak and a new set of points are obtained
$$\{(f_i(Hz), A_i(dB) - 20\log|G(f_i)|)\}$$
3. DAP modeling [8] is then applied to the resulting set of points, in order to calculate the global transfer function where the vocal tract response and the spectral tilt are included.
4. The global transfer function is decomposed, and a real pole representing the spectral tilt is eliminated from the calculated transfer function.
5. The speech signal is then filtered by the inverse of the vocal tract response, and then the derivative of the glottal source is obtained

These steps are followed for a set of values of the open quotient, and the best solution will give the minimum formant ripple in the deconvolved signal. The detailed algorithm is explained in [7].

5.2 Glottal source period extraction

As it has been analyzed before, the glottal source period selection is a critical aspect for the PSP calculation, and it is interesting to show how this has been realized. For that case, the derivative of the glottal source has been considered, as shown in Fig. 7:

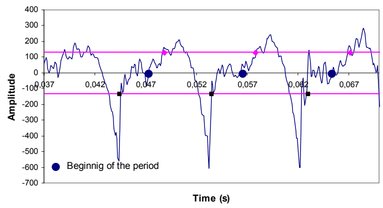

Fig. 7. Glottal source period extraction. Vowel "a", normal phonation

For a given signal two thresholds are established, one positive and the other one negative, in order to detect the time instants where the derivative of the glottal source is growing in both opening and closing phases. Once this has been done, the time intervals limited by this two kind of instants have been considered, and for each one of them the average value is obtained. Finally, for each interval the last crossing point of the signal by this average value will give the beginning of the period. In this way, the noise and formant ripple typical from the natural speech signals is somehow avoided

5.3 Glottal source parameterization

The analyzed database is the American English speech database. There, the sentence "*Sh<u>e</u> (has) l<u>e</u>ft for a great p<u>a</u>rty today*" has been recorded with various vocal qualities. For our analysis the underlined vowels have been extracted, in such a way that the different source-tract interaction [11] appearing in the inverse filtered signals is taken into account. It is known that this effect produces differences between the inverse filtered glottal source waveforms corresponding to different vowels for the same voice quality.

Under these conditions, and for each vowel, the NAQ and the PSP have been calculated for each recording provided in the database. After calculating these values, the average value and the standard deviation have been obtained for each vowel and each voice quality. The results are presented in Fig. 8.

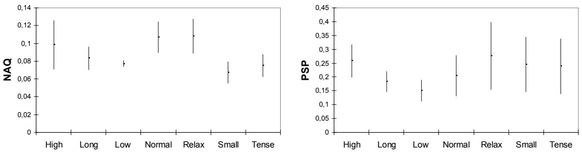

Fig. 8.a. Vowel [ɑ:]

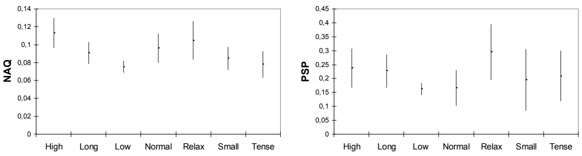

Fig. 8.b. Vowel [e]

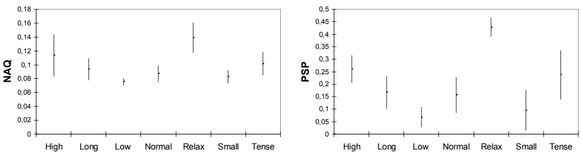

Fig. 8.c. Vowel [I]

Looking at Fig. 8 the graphs corresponding to different vowels are slightly different, because of the source-tract interaction. However the variation of the NAQ and PSP for the different voice qualities is quite similar. Now, comparing the results obtained for the NAQ and PSP for the same vowel it is possible to see that for the three cases both parameters are correlated, although, as in the simulation case, the values are different. Besides, the standard deviation for the PSP is larger than for the NAQ, what could be expected from the conclusions drawn in section 4.

6. Conclusions

The main conclusion from the analysis presented in this paper is that the Normalized Amplitude Quotient and the Parabolic Spectral Parameter are two global parameters, in the sense that both depend on the three main characteristics of the glottal source: Open quotient, asymmetry coefficient and spectral tilt. Also, it has been proved by simulations and natural speech signal's analysis that both parameters are correlated, or in other words, in spite of the different way they are defined they parameterize the glottal source in the same way.

By analyzing practical factors that could affect the measurement of the parameters, it is possible to conclude that the NAQ is a more robust parameter than the PSP. Moreover, NAQ is completely free from absolute time domain references, what makes it more robust to noise and formant ripple than the PSP. On the other side, even though the PSP is a spectral measurement, it depends on how the glottal period is extracted in time. In addition, the accuracy of this parameter depends on the value of the maximum tolerated error.

7. Acknowledgements

The Gobierno de Navarra and the Universidad Pública de Navarra are gratefully acknowledged for financial support.

8. References

[1] H. Fujisaki, M. Ljungqvist, "Proposal and evaluation of models for the glottal source waveform", *in Proceedings of the ICASSP*, 1986

[2] A. Isaksson, M. Millnert, "Inverse glottal filtering using a parameterized input model", *Signal Processing*, n °18, pp. 435-445, 1989

[3] M. Frölich, D. Michaelis, W. H. , "SIM-simultaneous inverse filtering and matching of a glottal flow model of acoustic speech signals", *Journal of the Acoustical Society of America*, Vol. 110, n° 1, pp. 479-488, July 2001

[4] Ch. Ma, Y. Kamp, L. F. Willems, "A Forbenius Norm Approach to Glottal Closure Detection from the Speech Signal", *IEEE Transactions on Speech and Audio Processing*, Vol. 2, nº2, pp. 258-265, April 1994

[5] P. Alku, U.K. Laine, "A new glottal LPC method for voice coding and inverse filtering", *in Proceedings of the IEEE International Symposium on Circuits and Systems*, Vol. 3, pp. 1831 -1834, 8-11 May 1989

[6] P. Alku, E Vilkman, "Estimation of the glottal pulseform based on Discrete All Pole modeling", *in Proceedings of the ICSLP*, pp. 1619-1622, September 1994, Jokohama Japan

[7] I. Arroabarren, A. Carlosena, "Glottal Spectrum Based Inverse Filtering", *in Proceedings of EUROSPEECH*, September 1-4, 2003, Geneva, Switzerland

[8] A. El-Jaroudi, J. Makhoul, "Discrete All Pole Modeling", *IEEE Transactions on Signal Processing*, Vol. 39, nº 2, pp. 411-423, February 1991

[9] K. Krishnamurthy, "Glottal source estimation using a sum-of-exponentials model", *IEEE Transactions on Signal Processing*, Vol. 40, nº 3, pp. 682-686, March 1992

[10] E. L. Riegelsbergerand, A. K. Krishnamurthy, "Glottal Source Estimation: Methods of Applying the LF-Model to Inverse Filtering", *in Proceedings of the ICASSP*, Vol. 2, pp. 542–545, 1993

[11] D. G. Childers, Chun-Fan Wong, "Measuring and modeling vocal source-tract interaction", *IEEE Transactions on Biomedical Engineering*, Vol. 41, nº 7, pp. 663 -671, July 1994

[12] Doval, C. d'Alessandro, "Spectral correlates of glottal waveform models: An analytic study", *in Proceedings of the ICASSP*, pp. 446-452, April 1997

[13] N. Henrich, Ch. d'Alessandro, B. Doval, "Spectral correlates of voice open quotient and glottal flow asymmetry: theory,limits and experimental data", *in Proceedings of the EUROSPEECH*, September 2001

[14] P. Alku, T. Bäckström, "Normalized amplitude quotient for parametrization of the glottal flow", *Journal of the Acoustical Society of America*, Vol. 112, nº 2, pp. 701-710, August 2002

[15] P. Alku, H. Strik, E. Vilkman, "Parabolic spectral parameter - a new method for quantification of the glottal flow", *Speech Communication*, nº 22, pp. 67-79, 1997

[16] T. Backstrom, P. Alku, E. Vilkman, "Time-domain parameterization of the closing phase of glottal airflow waveform from voices over a large intensity range", *IEEE Transactions on Speech and Audio Processing*, Vol. 10, nº 3, pp. 186 -192, March 2002

[17] G. Fant, J. Liljencrants, Q. Lin, "A four-parameter model of glottal flow", *STL-QPSR*, Vol. 85, nº 2, pp. 1-13, 1985

[18] I. Arroabarren, A. Carlosena, "Unified Analysis of Glottal Source Spectrum", *in Proceedings of EUROSPEECH*, September 1-4, 2003, Geneva, Switzerland

[19] A. V. Oppenheim, R. W. Schafer, "Digital Signal Processing", *Prentice Hall, Englewood Cliffs, NJ,* 1975

[20] H. Klatt, L. C. Klatt, "Analysis, synthesis and perception of voice quality variations among female and male talkers", *Journal of the Acoustical Society of America*, Vol. 87, nº 2, pp. 820-857, February 1990