GLOTTAL CORELLATES OF THE WORD STRESS AND THE TENSE/LAX OPPOSITION IN GERMAN

Krzysztof Marasek
Insitute of Natural Language Processing, Chair of Experimental Phonetics,
University of Stuttgart, Azenbergstraße 12, 70-174 Stuttgart, Germany
email: marasek@ims.uni-stuttgart.de

ABSTRACT

The influence of word stress and the tense/lax vowels opposition on the phonation process in German was investigated. A set of stimulus words in which tense and lax vowels occurs in both stressed and unstressed position was read by ten speakers of German. The phonation behaviour was investigated using the electroglottographic signal (EGG) recorded by the Laryngograph. The EGG signal was segmented and described using a set of timing and shape parameters. To compare the results with the literature the analysis concentrates on the Open and Speed Quotients. The ANOVA analysis shows that the main correlates of stress are the increase of fundamental frequency and the increase of closing and opening slopes steepness of the EGG signal. The Open and Speed Quotients depend primairly on other factors, like speakers sex or vowel type. Tenseness causes small changes in the Open Quotient, but the effect often interacts with stress.

1. GENERAL INTRODUCTION

In the auditory dimension stress is characterized by the higher prominence of the designated element. Prominence is, according to standard definition, a product of length, loudness, pitch (and its movement) and quality of the designated element. Those auditory quantities have correlates in acoustic signal properties such as: duration, intensity, fundamental frequency and spectral structure. Stress is also expressed by voice quality parameters that are caused by modifications of phonation type and vocal effort. Recent findings [10,11] show that the effects of stress are manifested as a change of spectral balance (tilt) emphasing higher frequencies of stressed vowels. The difference is caused by the change of glottal pulse shape.

The distinction between tense and lax vowels in German and related languages is primarily realized by duration and formant structure, and interacts with the realization of stress in the use of these parameters [6]. For some Asian languages that are also classified with a tense/ lax vowel distinction the distinction was found to be expressed with a clear difference in voice quality, including breathy phonation [7]. On these grounds it is justified to investigate the voice quality correlates of stress crossclassified with the influence of the tense/ lax distinction.

2. GLOTTAL PULSE DESCRIPTION

The production of speech is almost always modelled by a convolution of the excitation source (glottal pulses or noise) with time varying filter representing the vocal tract (sometimes coupled with nasal cavity) and lip radiation. The spectrum of the observed vowel could be regarded as the muliplication of the glottal pulse spectrum and vocal tract filter characteristics. A single vibratory cycle of the vocal folds could be (after [1]) regarded as consisting of three parts:1. an opening phase during which the folds pull apart increasing the glottal opening and enabling gradually increasing airflow through glottis; 2. a closing phase during which the folds come together reducing glottis opening thus reducing air flow; 3. a closed phase during which the glottis is closed or the glottal area is zero and there is no airflow through glottis. For certain phonation types (breathy voice, whisper), degree of vocal effort, or speakers there is no distinct closed phase. The waveform of the glottal volume velocity is responsible for the input drive of the vocal tract, so it affects mainly the voice quality. The glottal volume-velocity waveform could be described using certain parameters, for example using L-F model [3]. Of special interest are the Open (OQ) and Speed Quotients (SQ) having direct impact on the vowel spectra. In this contribution the effects of stress and tenseness are examined using the parameters derived from electroglottographic signals (EGG).

2.1. Description Of The EGG Signal

The EGG signal represents the relative contact area of the vocal folds [1,4]. The glottal waveform and the EGG signal have a comparable shape. The major difference is that the glottal flow will be more skewed and may have some formant ripple superimposed on it [2]. The phases of the EGG signal are presented on the Fig.1. As the vocal folds are separated the current flow between electrodes situated both sides of the speakers neck is minimal. When the vocal folds are closing, the electric current flow grows, a maximum is achieved by the complete contact of vocal folds.

The parametrisation of the EGG signal uses six straight line segments. Such a modelling relates the temporal features of the EGG waveform as well as motions of vocal folds including rolling and zippering actions [1,4,8]. The maximum-contact phase is defined to lie above 90% of peak-to-peak amplitude during the signal period, while the no-contact phase is defined to lie below the 10% level. The begining of the closing phase (segment (a) on Fig.1) is determined through the position of the peak change of the current flow (the steepest grow of the EGG). The closed phase (b) lies above 90% of the peak-to-peak amplitude. The opening phase ((c) and (d)) is tailored into two parts by the opening instant. The open segment (e) lies below 10% of the peak-to-peak amplitude and the last segment (f) connects the open phase with the point of steepest increase of the EGG.

The instant of the opening typically corresponds to a point of inflection in the EGG, where the EGG changes from a concave upwards to a concave downwards curve [1]. It is found primarily on the basis of first derivative of the EGG, but sometimes the maximum is quite weak (caused perhaps by the small deviation of the Laryngograph response at the very low frequencies). In such cases the opening instant is found at the crossing point of the EGG with the straight line connecting the closing phase markers of two following periods [8]. The additional problems in the EGG-signal segmentation are caused by vertical movements of the larynx. Those affect the baseline of the EGG and may vary quite rapidly, distoring the signal segmentation. To avoid this, a new algorithm was used [8].

The linearized model of the EGG signal was used to derive the Open and Speed Quotients. According to Fig.2 the Open Quotient was defined as $OQ1 = 100 \cdot \frac{(t_f - t_c)}{T}$, and the Speed Qutient was found as $SQ = 100 \cdot \frac{(t_a - t_0)}{(t_c - t_b)}$. An additional description of the EGG signal is given by the coefficients and durations of every segment. Additionally, the difference between true EGG-waveform and appropiate straight line segment is computed.

The evidence of the EGG-based description of the glottal signal is limited. Holmberg et al. [5] state a high correlation between Close Quotient derived from the EGG and inverse filtered glottal flow, Childers [1] reported also the use of the EGG for the OQ determination and as a basis for inverse filtering. The comparison between the OQ and the SQ measured by inverse filtered glottal flow waveform and the EGG is given in [2]. The OQ found in the EGG signal remain unchanged across varying voice intensities, although the OQ from volume flow shows significant decrease. The effect may be caused by an imprecise determination of the opening instant and the muscous bridging effect [1] in the EGG. To avoid this an additional measure of the Open Quotient was provided as $OQ2 = 100 \cdot \frac{(t_e - t_d)}{T}$. The SQ shows similiar trends for both methods, although the airflow waveform often exhibits a marked asymmetry not seen with the glottal area waveforms from high speed filming and photoglottography [2].

3. PHONETIC MATERIAL

The following near-minimal pairs involving tense and lax vowels were selected in which the segmental context was [t^h_l] throughout:

Ventil [i:] vs. Tormentill [I], Klientel [e:] vs. Kartell [ɛ], Spital [a:] vs. Metall [a], Anatolien [o:] vs. Ayatollah [ɔ], Thulium [u:] vs.Schatulle [ʊ].

On the basis of each of the ten words the variant with a stressed and unstressed target vowel was stimulated by appending the suffixes *-isch* and *-ist*, respectively. This triggers the following combinations of tenseness and stress: tense stressed, lax stressed, tense unstressed and lax unstressed. Those combinations are produced for a-, e-, i-, o- and u-vowels. The resulting 20 words were read twice each by ten native speakers of German, five female and five male, all with no history of speech, voice or hearing problems, with the age ranging from 24 to 34. The recordings consist of two-channels: one contains the speech signal, the second one the EGG signal. The recordings were made in a room in which no special acoustic precautions were taken. The signals were sampled with 16 KHz sampling frequency, 16-bits resolution and stored directly on a computer disk to avoid signal distortion. The EGG signal was supplied by a Laryngograph processor. No additional filtering (exlusive low pass filtering for sampling) nor signal modification was performed.

4. THE ANALYSIS OF THE EGG DATA

The recorded material was segmented and the vowels were isolated. The measurements of the described parameters were done automatically for the whole length of the vowel and for each parameter the mean, standard deviation, minimum and maximum were stored. The parameter values were stored also for the most representative pitch period of the vowel (found using the acoustical signal). Additional parameters were also collected: fundamental frequency, relative difference between successive pitch periods [8] and peak-to-peak amplitude. In the statistical analysis only the values at representative pitch period and the mean and the standard deviation for a whole vowel were used. No recordings were rejected due to poor quality, although the EGG recording of one female speaker was notably weaker.

5. RESULTS

The results are summarized in Table 1. A four-way ANOVA was computed for factors: vowel type (a-..u-), sex, stress, and tenseness with α=0.05.

The most important effects are caused by the sex factor. The effect of sex on the F0 does not need any explanation, although the female voices were more stable (smaller jitter). The mean of the females OQ1 is about 5% greater as the males, which confirms [5,10]. Also the OQ2 shows significant (ca. 15%) difference between male and female means. The average SQ for males is twice as low as for females, so the closing portion of the males EGG waveforms is shorter and the overall shape of glottal pulse is more skewed. This confirms [5], that the female voices have stronger negative tilt at high frequencies. There was no significant difference between measured peak-to-peak values of the EGG in relation to sex.

The F0 grows from open to closed and from front to back vowels, although there is a drop for o-vowels. The standard deviation of the Speed Quotient reduces significantly along the same line, though the mid vowels have higher Speed Quotient (of about 12%) as the other ones. The Speed Quotient tends to decrease along the front-back dimension. This stands in opposition to Sluijter [10], who did not find differences in SQ across vowel groups.

The vowel type factor interacts with the stress and tenseness factors.

The mean of the Open Quotient (OQ1) depends on stress and tenseness among all vowels, although stress interacts with sex. The OQ1 is about 7% larger for stressed vowels for female speakers, while for male speakers the difference does not exist. The effect of tenseness is significant at α=0.065. The OQ1 is of about 3% larger for stressed vowels and also about 3% shorter for tense vowels. The OQ for a choosen pitch period shows similiar pattern. The standard deviation of the OQ1 (OQ1STD) is smaller for tense back vowels as for the front vowels. The OQ2 shows similar patterns as the OQ1, although no effect of tenseness on mean OQ2M was found, only on its standard deviation. No relation between the OQ1 and the fundamental frequency was found, but the OQ2 increases with increasing F0. The relationship is stronger for the female speakers (r^2=0.22, for males r^2=0.13) with a linear regression slope of 0.16% per 1 Hz.

For the Speed Quotient the multiple ANOVA analysis shows significant interactions between vowel, sex, stress and tenseness. The significant difference was found only for /a/ vowels for which the SQ was notable smaller, confirming the expectation about lower spectral tilt for stressed vowels.

The duration of single phases of the EGG period shows a complex dependency pattern. The closing phase duration shortens only for stressed a-vowels, remaing unchanged (or even longer) for other vowel groups. The duration of the closing phase depends primarly on speaker's sex, being twice as long for females as for males. The differences caused by tenseness varied between vowel groups. The opening phase duration shows a regular pattern - it increases for stressed vowels.

The most important dependencies are found in the amplitude measurements of the EGG phases. The energy contained in the acoustical signal depends upon the maximal rate of glottal airflow decrease during the vocal cycle [3]. The negative slope of inverse filtered glottal flow is steepest during the EGG closing phase, which suggests [8] the use of the EGG contact closing slope as the indicator of the source signal energy gain. The measure is, of course, a relative one and depends on current system settings. Orlikoff [8] showed a strong correlation between this measure and the SPL. Regarding the stress, those findings are fully supported by the results of present and past [6] experiments. The closing slope is significantly greater for stressed tokens across all vowels and speakers. As the slope of the closing phase gets steeper the amplitudes at the mid and high frequencies increase relative to the amplitudes at low frequencies [10] and cause changes in the spectral tilt of the speech signal. Across all recordings the signal phase "end of opening"((d) on Fig.1) was shorter for stressed vowels than for unstressed ones. It seems to rely on the diffrencies on the EGG shape, for stressed vowels the open phase is more flat. For the tense/lax opposition no sharp tendency within timing and slopes parameters was found.

5. SUMMARY

The increase of the steepness of the closing slope of the EGG causes less negative spectral tilt of stressed vowels. The significant difference on skewness of the glottal pulse expressed by the Speed Qutient was observed only in case of /a/ vowels, although the strength of the excitation was notably higher for all stressed tokens. The timing properties, amplitudes and slopes of the EGG signals differ according to the speakers sex and vowel group. The effects of tenseness are weak. They may be interpreted as interactions with other factors.

6. REFERENCES

1. Childers, D.G. "A critical review of electroglottography", *CRC Critical Reviews in Biomedical Engineering*: Vol. 12, Issue 2, 131-161, 1984.
2. Dromey, Ch., Sathopoulos E.T., Sapienza C.M., "Glottal Airflow and Electroglottographic Measures of Vocal Function at Multiple Intensities", *Journal of Voice,* Vol.6, No.1, 44-54, 1992.
3. Fant, G. "The LF-model revisited. Transformations and frequency domain analysis" *STL-QPSR* 2-3/1995: 119–156, 1995.
4. Fourcin, A.J., "Normal and pathological speech: phonetic, acoustic and laryngographic aspects" in*: Fuctional surgery of the larynx and pharynx, eds: Singh W., Soutar D.,* Butterworth Heinemann, 1993.
5. Holmberg E.B., Hillman R.E., Perkell J.S. "Measures of the glottal airflow waveform, EGG, and acoustic spectral slope for female voice", *Proc. of XIIIth ICPhS, Stockholm, Sweden, 14-19 August*: 3:178-3:181, 1995.
6. Jessen, M., Marasek K., Schneider K., Classen K., "Acoustic correlates of word stress and the tense/lax opposition in the vowel system of german", *Proc. of XIIIth ICPhS, Stockholm, Sweden, 14-19 August*: 4:428-4:432, 1995.
7. Maddieson, I. & Ladefoged, P., "Tense" and "lax" in four minority languages of China, *Journal of Phonetics*, Vol. 13: 433-454, 1985.
8. Marasek, K. "An attempt to classify Lx signals", *Proc of Eurospeech'95, Madrid, Spain, 18-21 September*: 1729-1732, 1995.
9. Orlikoff R.F., "Assesment of the Dynamics of Vocal Fold Contact From the Electroglottogram: Data from Normal Male Subjects", *Jornal of Speech and Hearing Research*, Vol. 34, 1066-1072, 1991.
10. Sluijter A., *Phonetic Correlates of Stress and Accent*, Holland Institute of Generative Linguistics, *Leiden*, 1995.
11. Stevens K.N. Hanson H.M., "Classification of Glottal Vibration from Acoustic Measurement", *in: Vocal Fold Physiology Voice Quality Control, eds: Fujimura O., Hirano M.*, Singular Publishing Group, San Diego , 1995.

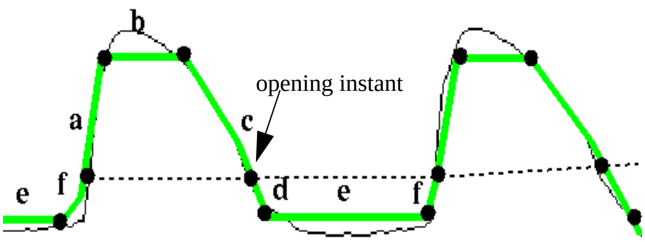

Figure 1: The description of the EGG signal shape using 6 straight segments. Dotted line connects the closing phase markers.

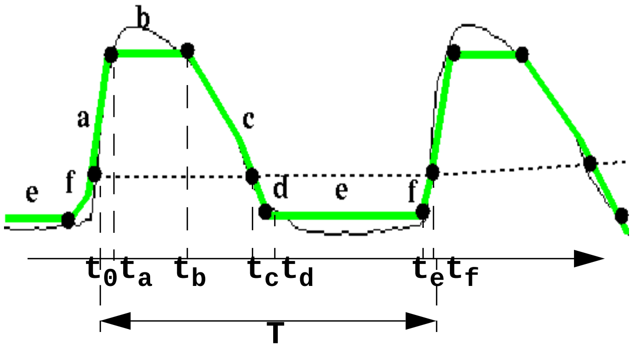

Figure 2: Timing of the EGG signal phases.

Table 1: The mean values of groups given by unstressed/stressed and lax/tense (given in parenthesis) oppositions. Only significant (α=0.05) and the most extend relations are given. Explanation of variable names: OQ1 - Open Quotient given by method I (see text), OQ2 - Open Qutient II, OQ2STD - standard deviation of the OQII along the whole wovel duration, F0 - fundamental frequency, SQ - Speed Quotient, SQSTD - standard deviation of the SQ, ECT - end of closing duration, SOT - start of opening duration, ECA - end of closing slope, SOA -start of opening slope.

Variable	female					male				
	a	e	i	o	u	a	e	i	o	u
OQ1	58.7/62.6 (61.3/59.9)					- (56.3/54.9)				
OQ2	37/41.9					-				
OQ2STD	(4.3/3.3)					(5.2/4.4)				
F0	205/228					118/134				
SQ	42					23				
	59/29		(36/42)	(35/39)		24/19				
ECT	9					5				
	9.2/8.2							4.5/6.0 (4.6/6.0)	4.6/5.5 (4.7/5.5)	(5/6.5)
SOT	24					21				
	22/28	20/23 (20/22.8)	-	-	-	19/22.5	-			
ECA	2267/2690									
SOA	-439/-540									