ISCA Archive
http://www.isca-speech.org/archive

LARYNX 1997
Marseille, France
June 16-18, 1997

VOICE QUALITY CORRELATES OF WORD STRESS AND TENSE VERSUS LAX VOWELS IN GERMAN

Michael Jessen and Krzysztof Marasek
Institute of Natural Language Processing, University of Stuttgart, Germany

RESUME

L'accentuation et la tension des voyelles allemandes ont été variées indépendamment l'une de l'autre et produites par 10 sujets. Le signal acoustique fut synchronisé et enregistré simultanément avec le signal électroglottographique. À l'aide des paramètres spectraux mésurés et à l'aide des mesures effectuées sur le signal électroglottographique il a pu être prouvé que l'accentuation et, à un degré moindre, la tension se manifestent dans la qualité du cycle vibratoire de la glotte.

INTRODUCTION

Results by [1] indicate that stressed vowels in American English are produced with a steeper and more rapid closing portion of the glottal cycle than unstressed vowels, which is likely to result from an increase in subglottal pressure with stress. Acoustically, this difference is reflected as a higher tilt in unstressed than in stressed vowels. Tense vowels in American English and several other languages have also found to be associated with higher tilt than lax vowels, though in the case of the tense/lax distinction spectral tilt is assumed to result from a breathy/modal voice distinction, respectively [2,3]. The goal of the present study is to investigate stress and tenseness under identical conditions in the attempt of replicating for German what has been reported for certain other languages. Furthermore, by using electroglottography we hope to obtain a more complete picture of the quality of the glottal cycle than what can be inferred from the acoustics alone.

METHOD

The following near-minimal pairs involving tense and lax vowels were selected, in which the segmental context was [t^h_l] throughout: *Ventil* [i:] 'valve' vs. *Tormentill* [ɪ] 'tormentilla', *Klientel* [e:] 'clients' vs. *Kartell* [ɛ] 'alliance', *Spital* [ɑ:] 'hospital' vs. *Metall* [a] 'metal', *Anatolien* [o:] 'Anatolia' vs. *Ayatollah* [ɔ] 'ayatollah', *Thulium* [u:] 'thulium' vs. *Schatulle* [ʊ] 'casket'. On the basis of each of the ten words both a variant with a stressed and one with an unstressed target vowel was triggered by appending the derivational suffixes *-isch* and *-ist*, respectively. The resulting 20 target words were read twice each by ten speakers of German, five female and five male [4 for details]. The acoustic signal and the electroglottographic signal (EGG signal) were recorded simultaneously on two different channels.

Every period of he EGG signal was segmented according to the temporal intervals (*a-f*) and temporal instances (t_x) illustrated in Fig. 1 [5 for details].

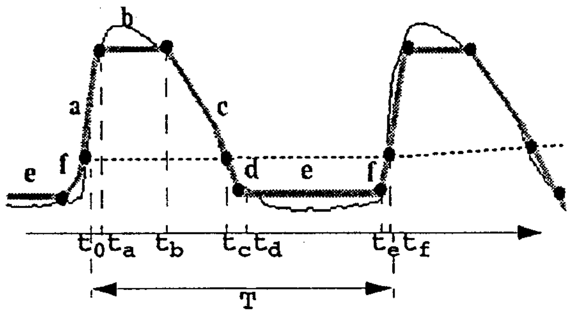

Figure 1. Description of the EGG signal

The maximum-contact phase *b* was defined to lie above 90% of peak-to-peak amplitude, while the no-contact phase *e* was below 10% peak-to-peak amplitude. The beginning of the closing phase t_0 corresponds to the steepest grow of the EGG signal. The most reliable way of determining the opening instant t_c was to define a straight line connecting t_0 of successive periods and to define t_c as the intersection between this line and the EGG signal. Tab. 1 presents a list of the obtained dependent variables (left) and their definition (right).

Table 1. Dependent variables in EGG signal.

Closing slope	slope between t_0 and t_a
Opening slope	slope between t_b and t_c
Open quotient I	$100 ((t_f - t_c) / T)$
Open quotient II	$100 ((t_e - t_d) / T)$
Duration of closing	$100 ((t_a - t_0) / T)$

The values of the dependent variables in Tab. 1 that are reported below were obtained by normalizing across speakers and by extracting the mean values across the entire duration of each target vowel, as defined from the beginning and end of F2.

In the analysis of the acoustic signal a DFT-spectrum superimposed by a LPC-spectrum was created for each target vowel. The LPC-spectra were of the type Burg and of the order 20. A Hamming window was used with 40 ms duration for male and 25.6 ms duration for female speech. The center of the vowel was chosen as the point in time around which the spectra were created. Formant frequencies F1, F2, and F3 (as well as F0) were measured using the LPC-spectra, while formant intensities A1, A2, A3 as well as the intensities of the first two harmonics (H1, H2) were measured on the DFT-spectra. Fig. 2 provides an illustration of the measurements.

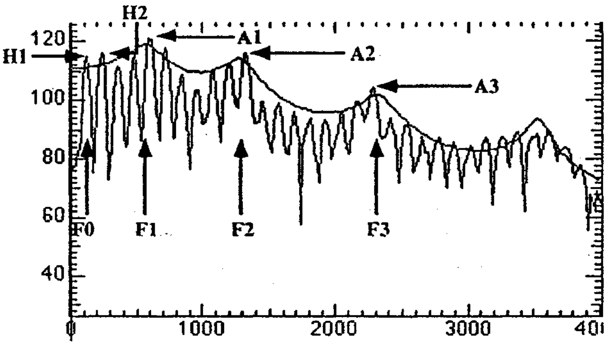

Figure 2. Measured acoustic parameters.

The DFT-spectrum and LPC-spectrum in Fig. 2 are superimposed upon each other. Amplitude parameters H1, H2, A1, A2, A3 (in Decibels) are represented with horizontal arrows, frequency parameters F0, F1, F2, F3 (in Hertz) with vertical arrows. The measured parameters in Fig. 2 were made subject to the following calculations:

$$H1^* = H1 - (20 \log_{10} (F1^2 / ((F1+F0)(F1-F0))))$$

$$H2^* = H2 - (20 \log_{10} (F1^2 / ((F1+2F0)(F1-2F0))))$$

$$A2^* = A2 - (40 \log_{10} (F1_n / F1_a) - 40 \log (\boldsymbol{\sqrt{(F2_a^2 - F1_n^2)} / \sqrt{(F2_a^2 - F1_a^2)}}))$$

$$A3^* = A3 - (20 \log_{10} (((F1_n^2\, F2_n^2) / (F1_a^2\, F2_a^2))\, ((F1_a^2\, F3_a^2) + (F1_a^2 - F3_a^2)) / ((F1_n^2 - F3_a^2)(F2_n^2 - F3_a^2))))$$

The formulae for H1*, H2*, and A2* are adopted from [1], the one for A3* builds upon [6]. In the boldface portion of the formula for A2* we deviated slightly from [1]. In the calculation of A2* and A3*, $F1_a$ and $F2_a$ refer to the *actual* F1- and F2-values measured in the same spectra in which the other parameters are measured. $F1_n$ and $F2_n$ refer to *neutral* values, i.e. formant frequency values derived by calculating the means across speakers, tenseness, and stress. Subsequently to these calculations, H2*, A1, A2*, and A3* were subtracted from H1*. Simplifying from [1], Table 2 presents a list of the obtained dependent variables in the acoustic study (boldface) with their physiological interpretation (left) and acoustic definition (right).

Table 2. Dependent variables (abbreviations in boldface) in acoustic signal.

Open quotient (**OQ**)	H1*-H2*
Degree of glottal opening (**GO**)	H1*-A1
Skewness of glottal pulse (**SK**), Durat. of closing portion (**RC**)	H1*-A2*, H1*-A3*

The first formant bandwidths (B1) were also measured, but since they were found never to depend significantly on stress or tenseness, no results for B1 are reported.

RESULTS

Four groups of t-test were performed separately for each set of vowels (A-vowels etc.), with the data pooled across individual speakers and sex. For example, in the first group of t-test the effect of stress on tense vowels was determined. The results are presented in Tab. 3, subdivided into the four groups of t-test. For each group the EEG results are reported first (first five rows) followed by the acoustic results (last four rows). Each cell contains the means and the standard deviations (in parentheses) of the unstressed (u) or lax (lx) tokens in the first line and the corresponding stressed (st) or tense (tn) tokens in the second line. Significant differences with $p < .05$ are presented in boldface, and those among them that approach $p < .01$ are also set in italics.

Among all the patterns observable in Tab. 3, we can for reasons of space only emphasize a few points. For almost all combinations of vowel type and tenseness, unstressed vowels manifest higher values of SK than stressed ones. In most cases this difference is significant. This reflects a higher spectral tilt of unstressed compared to stressed vowels and is consistent with the findings of [1]. According to Tab. 2 the results for SK suggest a larger

Table 3. Results of the electroglottographic and acoustic measurements

		A-vowels	E-vowels	I-vowels	O-vowels	U-vowels
effect of stress on lax tokens						
Closing slope	u	87.0 (19.0)	100.7 (13.3)	**87.3** (12.2)	101.5 (11.7)	**103.4** (28.2)
	st	95.9 (18.3)	124.6 (46.6)	**103.8** (10.0)	101.1 (12.0)	**125.8** (22.1)
Opening slope	u	-74.6 (13.1)	**-89.7** (10.9)	**-80.8** (13.9)	-89.0 (12.1)	**-104.8** (20.7)
	st	-79.9 (15.1)	**-114** (23.1)	**-123.9** (20.4)	-97.9 (8.0)	**-139.2** (22.2)
Open quotient I	u	100.8 (3.0)	***99.1*** (5.3)	**93.6** (8.0)	96.7 (11.4)	99.5 (13.2)
	st	102.2 (2.6)	***104.9*** (7.5)	**102.8** (3.1)	102.6 (8.7)	102.5 (6.6)
Open quotient II	u	92.9 (11.6)	102.1 (15.1)	***90.9*** (13.1)	96.8 (7.4)	***94.8*** (17.6)
	st	94.9 (14.0)	108.9 (20.1)	***107.8*** (21.2)	102.7 (7.4)	***110.7*** (12.8)
Duration of closing	u	87.1 (15.3)	90.8 (10.3)	107.8 (18.2)	90.2 (8.1)	107.0 (24.7)
	st	84.8 (14.7)	92.9 (22.1)	105.8 (11.1)	95.5 (14.9)	102.9 (10.1)
OQ	u	2.5 (3.3)	1.7 (3.0)	3.7 (3.1)	2.5 (2.6)	8.2 (10.3)
	st	2.7 (3.0)	4.4 (5.7)	7.2 (6.5)	3.2 (3.5)	5.5 (4.0)
G0	u	-1.2 (4.5)	-1.6 (3.9)	-5.2 (7.6)	-1.4 (4.3)	1.5 (23.7)
	st	-3.3 (4.6)	-2.5 (7.3)	-5.4 (3.3)	-3.5 (4.6)	-4.0 (2.2)
SK	u	***12.5*** (6.6)	***16.0*** (6.7)	18.5 (6.5)	***11.4*** (10.1)	***17.6*** (5.0)
	st	***1.1*** (7.9)	***5.2*** (6.5)	14.8 (5.0)	***-0.9*** (5.7)	***9.0*** (4.2)
RC	u	***22.7*** (5.3)	***21.2*** (4.5)	***23.8*** (5.6)	23.7 (10.2)	**24.0** (10.5)
	st	***15.4*** (6.7)	***11.8*** (4.9)	***17.5*** (6.9)	21.1 (6.6)	**30.8** (8.0)
effect of stress on tense tokens						
Closing slope	u	91.0 (28.1)	**91.2** (9.3)	**84.6** (10.3)	***90.8*** (7.8)	96.6 (27.2)
	st	89.1 (15.9)	**101.8** (14.9)	**94.0** (8.8)	***105.2*** (14.7)	125.3 (51.1)
Opening slope	u	-85.2 (12.6)	**-77.5** (7.4)	**-92.4** (10.6)	**-91.2** (10.2)	**-104.5** (23.4)
	st	-72 (12.5)	**-95.4** (12.4)	**-119.7** (24.7)	**-115.9** (15.0)	**-152.6** (40.3)
Open quotient I	u	98.4 (6.9)	95.7 (6.2)	***96.5*** (5.8)	98.8 (5.6)	96.5 (7.6)
	st	100.9 (4.9)	99.7 (5.1)	***102.8*** (7.0)	98.6 (4.9)	99.3 (8.3)
Open quotient II	u	99.4 (11.0)	95.1 (14.0)	94.4 (14.5)	96.9 (10.8)	95.4 (12.9)
	st	94.2 (14.9)	103.3 (18.8)	106.8 (18.9)	104.6 (9.1)	107.3 (19.9)
Duration of closing	u	**102.9** (15.9)	97.6 (5.9)	111.8 (16.9)	***99.2*** (8.6)	110.5 (20.7)
	st	**84.5** (13.8)	94.0 (7.4)	108.8 (12.7)	***111.8*** (15.6)	114.8 (20.8)
OQ	u	2.7 (3.5)	2.9 (5.4)	4.0 (3.8)	3.4 (4.7)	6.3 (9.5)
	st	1.9 (3.9)	4.6 (9.7)	25.4 (14.3)	2.1 (2.5)	15.9 (17.7)
G0	u	**0.0** (6.6)	-3.9 (4.8)	-4.7 (2.8)	***-3.1*** (2.5)	**-3.8** (6.1)
	st	**-4.1** (5.7)	-5.3 (3.7)	-8.4 (9.8)	***-5.9*** (2.8)	**-11.4** (12.2)
SK	u	***13.2*** (7.0)	19.2 (7.5)	19.6 (8.7)	***15.1*** (5.4)	***16.5*** (6.2)
	st	***-1.0*** (9.5)	15.2 (6.1)	19.9 (8.7)	***10.1*** (5.0)	***7.3*** (11.9)
RC	u	22.7 (6.5)	***22.0*** (5.7)	21.4 (7.2)	***32.7*** (11.3)	***36.6*** (9.5)
	st	18.6 (7.5)	***15.9*** (5.1)	18.7 (9.0)	***49.0*** (7.7)	***46.7*** (11.4)
effect of tenseness on unstressed tokens						
Closing slope	lx	87.0 (19.1)	100.7 (13.3)	87.3 (15.2)	***101.5*** (11.7)	103.0 (28.2)
	tn	91.0 (28.1)	91.2 (9.3)	84.6 (10.3)	***90.8*** (7.9)	96.6 (27.3)
Opening slope	lx	-74.6 (13.1)	**-77.5** (7.4)	***-80.8*** (13.9)	-89.0 (13.1)	-104.8 (20.7)
	tn	-85.2 (13.6)	**-89.7** (10.9)	***-92.4*** (10.6)	-91.2 (10.1)	-104.5 (23.4)
Open quotient I	lx	100.8 (3.0)	99.2 (5.3)	93.6 (8.0)	96.8 (11.3)	99.5 (13.2)
	tn	98.4 (6.9)	95.2 (6.3)	96.5 (5.8)	98.8 (5.6)	96.5 (7.6)
Open quotient II	lx	92.9 (11.7)	102.1 (15.1)	91.0 (13.1)	96.8 (7.4)	94.8 (17.6)
	tn	99.4 (11.0)	95.1 (13.9)	94.5 (14.5)	96.9 (10.8)	95.6 (13.0)
Duration of closing	lx	***87.1*** (15.3)	90.8 (10.3)	107.8 (18.2)	**90.2** (8.1)	107.0 (24.7)
	tn	***102.9*** (15.9)	97.6 (5.9)	111.9 (16.9)	**99.2** (8.6)	110.5 (20.8)

OQ	lx	2.5 (3.3)	1.7 (3.0)	3.7 (3.1)	2.5 (2.6)	8.2 (10.3)
	tn	2.7 (3.5)	2.9 (5.4)	4.0 (3.8)	3.4 (4.7)	6.3 (9.5)
G0	lx	-1.2 (4.5)	-1,6 (3.9)	-5.2 (7.6)	-1.4 (4.3)	1.5 (23.7)
	tn	0.0 (6.6)	-3.9 (4.8)	-4.7 (2.8)	-3.1 (2.5)	-3.8 (6.1)
SK	lx	12.5 (6.6)	16.0 (6.7)	18.5 (6.5)	11.4 (10.1)	17.6 (5.0)
	tn	13.2 (7.0)	19.2 (7.5)	19.6 (8.7)	15.1 (5.4)	16.5 (6.2)
RC	lx	22.7 (5.3)	21.2 (4.5)	23.8 (5.6)	***23.7*** (10.2)	***24.0*** (10.5)
	tn	22.7 (6.5)	22.0 (5.7)	21.4 (7.2)	***32.7*** (11.3)	***36.6*** (9.5)
effect of tenseness on stressed tokens						
Closing	lx	95.1 (15.9)	124.6 (46.6)	***103.8*** (10.0)	101.1 (13.0)	125.8 (32.1)
slope	tn	81.9 (18.3)	101.8 (14.5)	***94.0*** (8.8)	105.3 (24.7)	125.3 (51.1)
Opening	lx	-79.6 (12.5)	***-114.0*** (23.1)	-123.9 (20.3)	**-97.9** (8.0)	-139.2 (22.4)
slope	tn	-72.6 (15.1)	***-95.3*** (13.5)	-119.7 (24.8)	**-115.9** (15.0)	-152.6 (40.3)
Open	lx	102.2 (2.6)	99.7 (5.1)	102.8 (3.1)	102.6 (8.7)	102.5 (6.6)
quotient I	tn	100.9 (4.9)	105.0 (7.5)	102.8 (7.0)	98.6 (4.9)	99.3 (8.3)
Open	lx	94.9 (14.0)	108.9 (20.1)	107.7 (21.1)	102.7 (7.4)	100.6 (17.8)
quotient II	tn	94.2 (14.9)	103.3 (18.7)	106.8 (18.9)	104.6 (9.1)	107.3 (19.9)
Duration of	lx	84.8 (14.7)	92.9 (22.1)	105.8 (11.1)	**94.5** (14.8)	102.9 (10.1)
closing	tn	84.5 (13.8)	94.0 (7.4)	108.6 (12.7)	**111.8** (15.6)	114.8 (20.8)
OQ	lx	2.7 (3.0)	4.4 (5.7)	7.2 (6.5)	3.2 (3.5)	5.5 (4.0)
	tn	1.9 (3.9)	4.6 (9.7)	25.4 (14.3)	2.1 (2.5)	15.9 (17.7)
G0	lx	-3.3 (4.6)	-2.5 (7.3)	-5.4 (3.3)	-3.5 (4.6)	**-4.0** (2.2)
	tn	-4.1 (5.7)	-5.3 (3.7)	-8.4 (9.8)	-5.9 (2.8)	**-11.4** (12.2)
SK	lx	1.1 (7.9)	***5.2*** (6.5)	**14.8** (5.0)	***-0.9*** (5.7)	9.0 (4.2)
	tn	-1.0 (9.5)	***15.2*** (6.1)	**19.9** (8.7)	***10.1*** (5.0)	7.3 (11.9)
RC	lx	15.4 (6.7)	***11.8*** (4.9)	17.5 (6.9)	***21.1*** (6.6)	***30.8*** (8.0)
	tn	18.6 (7.5)	***15.9*** (5.1)	18.7 (9.0)	***49.0*** (7.7)	***46.7*** (11.4)

closing slope for stressed compared to unstressed vowels. These predictions are borne out by the EGG results, since the EGG-based closing slope parameter shows the pattern consistent with Tab. 2. The acoustic parameter RC, and to a more limited extent also GO, and the EGG-based opening slope parameter also show a striking influence of stress.

For most combinations of vowel type and stress, tense vowels show larger values for SK and RC than lax vowels. This result is consistent with [2,3], according to which tense vowels (or "ATR"-vowels) exhibit a higher spectral tilt than lax vowels. Telling from the number of significant cases this effect is somewhat less robust than the dependency of tilt on stress. The EGG measurements show a longer duration of closing in tense than lax vowels and a smaller closing slope. The effect is more consistent and more often significant for duration of closing, than closing slope, which is consistent with the fact that the acoustic results show a stronger effect for RC than SK (see Tab. 2).

Contrary to [3], tense and lax vowels do not exhibit a reliable difference in open quotient, neither acoustically nor in the EGG signal. In several languages tense vowels are more breathy than lax ones, and therefore expected to show a higher open quotient [3]. It is possible that this discrepancy with our findings is due to phonetic differences between languages.

REFERENCES

[1] Sluijter, A.M.C. 1995. *Phonetic correlates of stress and accent*. The Hague: Holland Academic Graphics [Doctoral Diss., Leiden].

[2] Halle, M. & Stevens, K.N. 1969. On the feature "Advanced Tongue Root". *MIT Research Laboratory of Electronics Quarterly Progress Report* 94: 209-215.

[3] Kingston, J. et al. 1997. Integrality in the perception of tongue root position and voice quality in vowels. *JASA* 101: 1696-1709.

[4] Jessen, M., Marasek, K., Schneider, K. & Claßen, K. 1995. Acoustic correlates of word stress and the tense/lax opposition in the vowel system of German. *Proc. Intern. Congr. Phonetic Sciences* 13, 4: 428-431.

[5] Marasek, K. 1995. An attempt to classify Lx signals. *Proc. of Eurospeech* 1995. 1729-1731.

[6] Hanson, H. M. 1995. *Glottal characteristics of female speakers*. Ph.D. Diss., Harvard.

We like to acknowledge the contribution of Kathrin Claßen to the acoustic measurements.