VOQUAL'03, Geneva, August 27-29, 2003

Aging Female Voices: an Acoustic and Perceptive Analysis

Markus Brückl, Walter Sendlmeier

Institut für Sprache und Kommunikation
Technische Universität Berlin, Germany
brueckl@kgw.tu-berlin.de

Abstract

This study examines the changes in adult female voices due to the aging process. Acoustic cues in voices that enable listeners to recognize a speaker's vocal age are specified as well as acoustic cues that straightly indicate the speaker's chronological age.

The analysed data are recordings of the voices of 56 female speakers differing in age. The recorded speech samples include sustained vowels, read speech and spontaneous speech. Our methods are acoustic analyses and perception tests. The perception tests are designed to analyse the influence of the vowel onset on the amount of information about aging transferred by sustained vowels. The acoustic evaluation comprises phonic parameters like measurements of stability of voice or of vocal tremor which are supposed to reflect voice qualities that are expected to vary with age. Changes in tempo of articulation are also investigated.

We found that increasing amplitude perturbation is an indicator of increasing age even on the basis of spontaneous speech. Reading rate decreases with increasing age, whereas there is no significant change in articulation rate of spontaneous speech in women's voices. Based on sustained vowels of female voices, the frequency tremor intensity index indicates age more accurately than F0 and amplitude perturbations. We also found evidence for the relevance of the vowel onset to recognize age more accurately.

1. Introduction

Our experiments investigate characteristics inherent in the speech signal that can be related to a concept of age. There has been a considerable amount of research concerned with acoustic phenomena of the aging voice as well as with the perception of aging voices. A comprehensive summary of the state-of-the-art can be found in Linville [3].

1.1. The accuracy of age estimations

The first studies investigating the aging voice were concerned with questions on the accuracy of age perception. According to Ptacek & Sander [6] listeners are able to assign single presented vowels correctly to two age groups (young vs. old) with a probability of 78%. Rating read speech, they even achieve 99%. Shipp & Hollien [9] found direct age estimations on read speech and chronological age correlating at r = 0.88. Factors influencing the accuracy of age estimations are the amount of information provided by the speech sample, the demanded precision of the age estimations, the dialect of the speakers and the age of the listeners. Linville [3] reports young and middle-aged adults to be the most accurate raters.

1.2. Acoustic measures of age

Acoustic examinations of the aging voice up to now included measurements of F0, intensity, perturbations (jitter, shimmer), spectral noise, speech tempo and formant frequencies.

F0. Relations of F0 and age are coherently reported: In female voices F0 is constantly decreasing with increasing age with a sharp decline around menopause. Men show a slightly decreasing curve followed by a strong increase starting around the 50th year.

Stability of vocal fold vibration. With increasing age fluent and controlled agility decreases. In speech production this affects vocal fold vibration, and thus is reflected in perturbation measurements. However, Ramig & Ringel [7] emphasize that frequency perturbations are likely to relate rather to physical fitness than to age. Orlikoff [4] found old men as a group to show increased perturbation values as well as increased variances regarding these measures. But in following investigations Orlikoff and colleagues demonstrated that perturbation measures depend on mean F0 and mean sound pressure levels. In this connection Linville asserts that "firm conclusions as to the effect of aging on jitter and shimmer levels are not now possible" ([3], p. 175) and further that "amplitude SD in female speakers with aging has yet to be investigated" ([2], p. 364).

Spectral noise. Ramig [8] also investigated vowel spectral noise as a function of age and physiological condition and found spectral noise to be related to fitness but not to age, at least in men's voices. Since spectral noise can be a result of perturbed vocal fold oscillations and since listeners attribute breathy voice to old speakers, "research is necessary to examine spectral noise as a correlate of perceived age estimates from women's voices" (Linville [3], p.197).

Speech tempo. Numerous studies demonstrated that older persons speak slower. Most of them concentrated on male speakers. If solely women were investigated, results are somewhat contradictory: Oyer & Deal [5] found a decline in articulation rate with increasing age in women. Hoit et al. [1] did not. Moreover, "studies have not been conducted correlating age estimates to speech rate in female speakers" (Linville [2], p. 372).

Vocal tremor. Measurements of vocal tremor do not seem to have been analysed yet with respect to age. Xue & Deliyski [10] report insufficient validity for the tremor measures applied in this study and they did not use them in a comparable study.

2. Aims

The primary objective of this experiment is to investigate several aspects of the aging voice that according to Linville [2, 3] have not been considered in detail, like amplitude standard deviation, spectral noise and articulation rate as a function of age or of age estimations in female voices.

Additional voice parameters supposed to vary with age are examined as well. The measured parameters are assumed to depict the voice qualities pitch, roughness, hoarseness, breathiness, speech tempo and, for the first time, vocal tremor. All parameters are correlated with the chronological age of the speaker as well as with the vocal age of the stimuli. A perception test is accomplished to estimate the vocal age of the acoustically analysed voice samples. The perception test yields direct age estimations and does not distinguish age groups.

In order to examine the relevance of the vowel onset on age-related information, two parts of equal length of each sustained vowel are analysed and perceptually rated: the first containing the onset, the second without onset.

3. Data and method

3.1. Speakers

The voices of 56 women, mostly middle-aged (from 20 to 87 years; AM = 49.77; SD = 16.01) were recorded. By means of this age distribution this investigation concentrates on changes occurring around menopause. All speakers judged themselves as healthy, especially as not suffering from hearing impairment or voice disease.

3.2. Voice samples

Each speaker provided 8 voice samples, assumed to differ in the amount and type of age-related information. They can be subdivided in (1-6) onset and quasi-stationary parts of the sustained vowels /a/, /i/ and /u/ (/a/-o, /a/-s, /i/-o, /i/-s, /u/-o, /u/-s), (7) read speech (r-sp) and (8) spontaneous speech (s-sp). For producing sustained vowels, speakers were instructed to keep pitch and loudness as constant as possible at their preferred levels. The onset samples are the first 2.2 seconds of a sustained vowel, the quasi-stationary samples are 2.2 seconds of the center of the same sustained vowels. That means that vowel onset is not actually defined in this study, but it is assumed that the starting section of a vowel contains 'onset' while the middle section does not. For read speech all speakers were reading the same visually presented path description applying their normal reading style. Spontaneous speech was created by describing a picture.

3.3. Listeners

15 young adult listeners (6 female, 9 male, between 22 and 35 years old (AM = 28.67; SD = 3.46), not hearing impaired) rated perceived age of the speech samples. The reliability of these judgements was tested separately for each voice sample group via Kendall's coefficient of concordance (W) (see Table 3). The lowest coefficient of concordance (W = 0.327; $p < 0.0005$) still denotes sufficient reliability and none of the listeners universally impaired all coefficients of concordance. Vocal age of a speech sample therefore is best approximated by the arithmetical mean of all listener's judgements on this speech sample.

3.4. Instrumentation

All voice samples were recorded on digital audio tape (DAT) with 16 bit at a sampling rate of 48 kHz using a Tascam DA-P1 recorder and an AKG C410 headset condenser microphone. For further computation the voices were digitally transferred to the hard disc of a PC. An audio CD ROM containing all voice samples was produced to be used in the perception test.

All stimuli are analysed with respect to their phonatory parameters via the Multi Dimensional Voice Program (MDVP) of KAY Elemetrics Corp. This program was originally developed to obtain measures indicating degree and kind of voice diseases on the basis of sustained /a/ realisations. Articulatory parameters are measured computer-aided.

3.5. Acoustic parameters

MDVP extracts 33 voice parameters. Only those parameters that hypothetically vary with age are examined in this experiment (see Table 1).

They can be grouped according to their supposed perceptive equivalent. Pitch as well as it's corresponding acoustic parameter fundamental frequency is supposed to decline in female voi-

	parameter group		short explanation	abbreviation [unit]
phonation	pitch		average fundamental frequency	**F0 [Hz]**
	stability of vocal fold vibration	F0 stability	standard deviation of the fundamental frequency	**STD [Hz]**
			average cycle-to-cycle variation in the duration of fundamental period	**Jita [μs]**
			average period-to-period variation of the pitch, normalized by F0	**Jitt [%]**
			relative jitter (Jitt) with smoothing factor of three periods	**RAP [%]**
			relative jitter (Jitt) with smoothing factor of five periods	**PPQ [%]**
			relative jitter (Jitt) with smoothing factor of 55 periods	**sPPQ [%]**
			relative standard deviation of F0	**vF0 [%]**
		amp. stability	average cycle-to-cycle variation in waveform amplitude	**ShdB [dB]**
			average period-to-period variation of the amplitude, normalized by the average amplitude	**Shim [%]**
			relative shimmer (Shim) with smoothing factor of five periods	**APQ [%]**
			relative shimmer (Shim) with smoothing factor of 55 periods	**sAPQ [%]**
			relative standard deviation of the peak-to-peak amplitude	**vAm [%]**
	spectral energy distribution		average ratio of the aperiodic energy between 1500 and 4500 Hz to the harmonic energy between 70 and 4500 Hz	**NHR**
			average ratio of the aperiodic energy between 2800 and 5800 Hz to the harmonic energy between 70 and 4500 Hz	**VTI**
			average ratio of the harmonic energy between 70 and 1600 Hz to the harmonic energy between 1600 and 4500 Hz	**SPI**
	tremor		frequency tremor intensity index	**FTRI [%]**
			amplitude tremor intensity index	**ATRI [%]**
articulation	speech tempo		duration of the analysed voice sample	**t [s]**
			accumulated duration of the breaks	**t(br) [s]**
			articulation rate: syllables per seconds of speech without breaks	**AR [syllables/s]**
			number of breaks	**N(br)**

Table 1: *Systematic overview and short explanation of the measured parameters.*

ces with increasing age.

Roughness or hoarseness, registered by perturbation measurements, is assumed to increase with age. The collected frequency and amplitude perturbation parameters measure either absolute variations or variations normalised by F0 or the mean amplitude, respectively. They also differ according to their time resolution, describing e.g. variations from one cycle to the next or variations between values that result from averaging 55 cycles or variations from one cycle to all other cycles in the analysed sample. Breathiness, measured by the parameters of spectral energy distribution, that are extracted from the long-term average spectrum as well as vocal tremor (frequency tremor intensity index and amplitude tremor intensity index) are also expected to increase with age.

The manually measured parameters of speech tempo (articulation rate, duration of read speech and accumulated duration of breaks) are expected to indicate slower speech in older voices. Articulation rate is calculated by dividing the spoken syllables by the sample duration less the accumulated duration of breaks.

4. Results and discussion

4.1. Age perception: relevance of vowel onset

The 'amount of information on age', which is controlled by the different stimuli, as well as 'smoking habits' and 'accent' are investigated according to their impact on age estimations. A multivariate analysis of variance testing the influence of these factors on the mean of age estimations as well as on the inter-listener variance of age estimations showed only the intra-speaker factor 'amount of information' to be significant.

The influence of the onset can also be shown in the correlations and regressions of vocal age, separately for each sample group, and chronological age (see Figure 1 and Table 2).

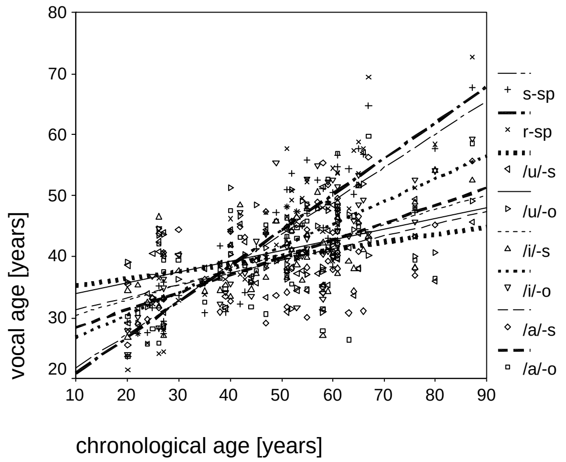

Figure 1: *Regression of the vocal age of each speaker's voice samples as a function of the speaker's chronological age. The regression lines per voice sample group show the listeners' tendency to rate difficult items conservatively: The less information a voice sample group provides, the smaller is the range of the estimations and the less steep are the regression lines.*

The estimations approximate chronological age more accurately, if vowel onset is provided. The strongest relation can be found between chronological age and the spontaneous speech stimuli, which are the items produced most naturally and which therefore should carry most information on age. The accuracy of age estimations on read speech stimuli (r = 0.862) corresponds with accuracies reported e.g. by Shipp & Hollien [9]. The accuracies of estimations on isolated vowels are as expected lower than those on s-sp and r-sp samples. If chronological age is to be estimated on the basis of these data, the best results will be obtained via linear regression of spontaneous speech stimuli by means of the following equation (1) (see Figure 2):

	/a/-o	/a/-s	/i/-o	/i/-s	/u/-o	/u/-s	r-sp	s-sp
r	.559	.443	.738	.603	.460	.344	.862	.864
p	.000	.000	.000	.000	.000	.005	.000	.000

Table 2: *Correlation coefficients (Pearson's r, one-tailed significance p) of the correlations of the chronological age and the vocal age of the different stimulus groups.*

$$chron.age = 1.37 * vocalage(s - sp) - 9.63 \qquad (1)$$

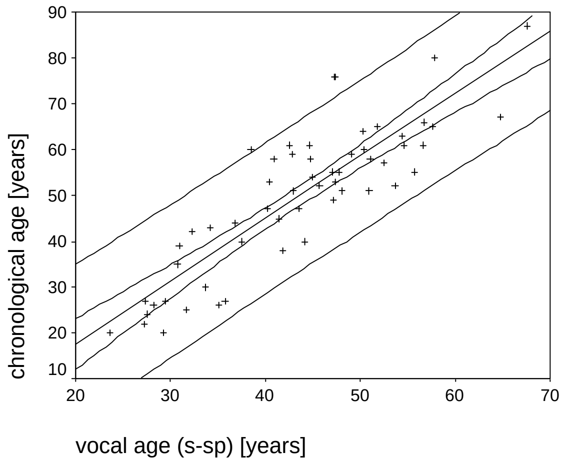

Figure 2: *Regression to predict the chronological age by means of the vocal age of the spontaneously spoken voice samples. The middle line is the regression line representing equation (1). The curved lines denote the 95% confidence interval for the regression line. The outer straight lines denote the 95% confidence interval for single cases.*

Vowel onset also contributes to appointing vocal age more accurately. If the concordance of age estimations is computed separately for each stimulus group, values corresponding to the correlations of Table 2 can be observed as shown in Table 3:

	/a/-o	/a/-s	/i/-o	/i/-s	/u/-o	/u/-s	r-sp	s-sp
W	.524	.438	.506	.371	.327	.336	.730	.654
p	.000	.000	.000	.000	.000	.000	.000	.000

Table 3: *Concordance coefficients (Kendall's W, one-tailed significance p) of listeners' judgements per voice sample group. The more information is provided by the speech samples the better the listener's estimations correspond.*

Listeners' estimations matched best for the speech samples. At least in the /a/ and /i/ groups concordance is better for the onset samples. /u/ realisations are generally hard to rate, which points to a relative lack of information on age in /u/ vowels.

Another interesting result of the perception test is that all listeners estimate the speakers as a group significantly younger than they are. The tendency to rate the perceived age lower than the actual chronological age of the speaker, can be shown as well for each voice sample group (see Figure 3).

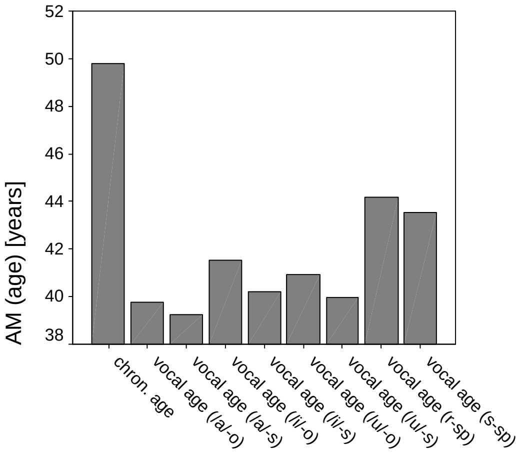

Figure 3: *Arithmetical mean (AM) of chronological age compared with the AM per voice sample group of vocal age values. Estimations are generally lower than chronological age and seem to rise towards chronological age the more information a speech sample carries.*

This effect could be explained by taking into account the negative bias on judging someone as too old and the listeners' endeavour to avoid it. In addition, old speakers as a group are rated more divergently than young speakers, which corresponds to former findings of increased variability in elderly speakers regarding most acoustic and physiological measures (e.g. [3] and [7]).

4.2. Acoustic correlates of age

An overall comparison of the data reveals generally higher correlations of the acoustic parameters with vocal age than with chronological age (see Table 4 and Table 5).

F0. The average fundamental frequency (F0) of the /i/ samples is not correlated to both concepts of age. In the /a/ and /u/ samples the relation to perceived age is stronger than the actual one. The poor correlations of F0 and chronological age in /i/ and /u/ vowels can be explained by the intrinsic pitch phenomenon supposed to be responsible for the generally high pitch in these vowels (see Figure 4). Apart from isolated /i/ and /u/ vowels, decreasing F0 in women seems to be a reasonable predictor of increasing age.

In the speech samples correlations between F0 and chronological age are almost identical to those with vocal age. This argues for the speech samples carrying nearly the maximum of age-relevant information regarding F0.

The correlations between the age concepts and F0 in spontaneous speech are clearly stronger than in read speech, which argues for a general levelling F0 in reading voice compared to

	/a/-o	/a/-s	/i/-o	/i/-s	/u/-o	/u/-s	r-sp	s-sp
F0	**-.290**	**-.301**	-.039	-.054	-.176	-.212	**-.378**	**-.506**
Jita	.015	*-.014*	-.014	-.019	*-.045*	*-.060*	.160	**.344**
STD	-.009	*-.006*	**.396**	**.233**	*.054*	-.103	-.086	**-.244**
vF0	.093	*.056*	**.328**	.202	*.093*	.013	.050	-.072
ShdB	.124	.132	.109	*-.074*	*-.001*	*.036*	.218	**.519**
Shim	.086	.126	.039	-.007	*.007*	*.039*	.211	**.466**
APQ	.154	.186	.043	.002	*.009*	*.033*	.198	**.551**
sAPQ	**.374**	**.326**	**.296**	.138	-.015	-.046	.109	**.473**
vAm	.166	***.239***	.185	.192	.183	***.165***	.012	.220
FTRI	**.433**	**.291**	**.469**	**.434**	***.211***	.261	-.061	.038
AR							**-.362**	-.076
t							**.432**	
t(br)							**.400**	
N(br)							***.277***	

Table 4: *Correlation coefficients (Pearson's r) of the correlations of the speakers'* ***chronological age*** *and the acoustic parameters per speech sample group. One-tailed highly significant correlations (p < 1%) are printed in bold letters. The values in italics are non-parametric correlation coefficients (Kendall's τ). Parameters are excluded, if they offer no highly significant correlation with chronological age in any of the sample groups.*

	/a/-o	/a/-s	/i/-o	/i/-s	/u/-o	/u/-s	r-sp	s-sp
F0	**-.481**	**-.498**	-.016	-.105	**-.270**	**-.461**	**-.396**	**-.487**
Jita	.013	*.070*	-.195	-.019	*.064*	*.081*	.211	**.374**
vF0	.154	*.138*	.177	**.225**	***.195***	**.333**	.076	.019
ShdB	**.327**	**.377**	-.005	*.042*	*.036*	***.154***	**.231**	**.548**
Shim	**.256**	**.361**	-.077	.108	*.006*	***.156***	**.258**	**.505**
APQ	**.338**	**.419**	-.018	.131	*.036*	***.199***	**.232**	**.602**
sAPQ	**.418**	**.469**	**.278**	**.339**	.151	**.284**	.140	**.553**
vAm	.139	***.226***	.077	.194	.102	*.122*	.064	**.318**
SPI	**.418**	**.327**	-.065	-.052	-.153	.039	.135	.008
FTRI	**.364**	.208	**.366**	**.563**	***.216***	**.594**	-.040	-.009
AR							**-.415**	-.165
t							**.512**	
t(br)							**.495**	

Table 5: *Correlation coefficients (Pearson's r) of the correlations of the* ***vocal age*** *values of each speech sample group and the corresponding acoustic parameters. One-tailed highly significant correlations (p < 1%) are printed in bold letters. The values in italics are non-parametric correlation coefficients (Kendall's τ). Parameters are excluded, if they offer no highly significant correlation with the vocal age in any of the sample groups.*

'normal' voice, which leads to a loss of age-related information.

F0 stability. As it was expected according other findings, there are only sporadic respectively minor correlations of F0 perturbations and age. The most clearly found are those between the relative (vF0) and the absolute (STD) F0 standard deviation and chronological age in the /i/-o samples, as well as those between vF0 and perceptive age in the /u/-s samples. However, these significant correlations seem to be rather randomly distributed, which points to a general but very faint relation between F0 perturbation measures and age. This relation could very plausibly be reduced to F0 perturbation measures being a measure of physical fitness, which itself is related to age. But this assumption is speculative, since the physical fitness of the speakers was not documented in this study. The correlation between absolute Jitter (Jita) and chronological age in the spontaneous speech samples can be ascribed to the strong relation of F0 to age in these sample groups, as especially Jita depends

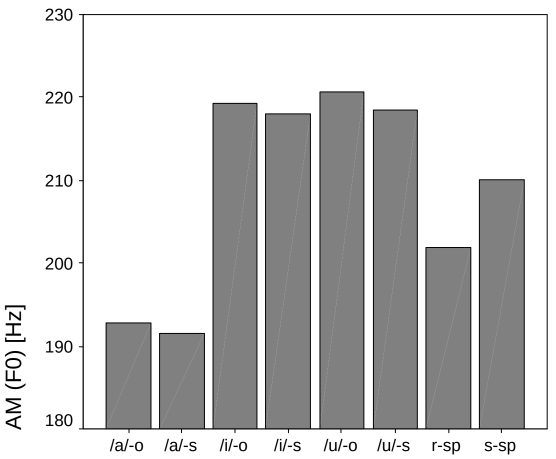

Figure 4: *The arithmetical mean values of the fundamental frequency per voice sample group.*

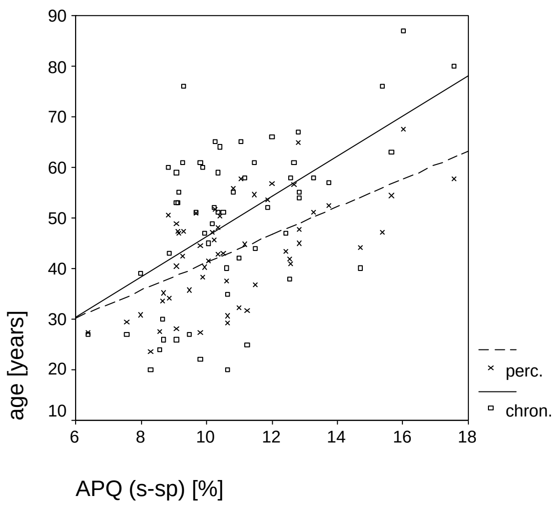

Figure 5: *Regressions of the age estimates on spontaneous speech items respectively of the speaker's chronological age as a function of the APQ values of the spontaneous speech items.*

strongly on F0.

Amplitude stability. It is striking that the amplitude perturbation measures show the strongest correlations with chronological and perceptive age in the spontaneous speech samples, although these measures have rather been designed to be applied in the analysis of sustained /a/ vowels. Most notably, in read speech samples the amplitude perturbation measures show only marginal correlations with vocal age and almost do not correlate with the speakers' chronological age. The question arising is why amplitude perturbation measures only increase with increasing age when a picture is described and not when a text is read, the more so since these differences can not be explained by greater intensity changes from one word to the next. The measure reflecting amplitude perturbations with a temporal magnitude of words is vAm. This parameter is the only amplitude perturbation measure that does not correlate with chronological age in spontaneous speech. The strongest correlation is assessed by APQ (see Figure 5), which compares amplitude differences at a temporal resolution of about 25 ms.

The only mentionable correlations of amplitude perturbation measures and chronological age apart from the spontaneous speech are those of sAPQ in /a/ realisations. In all sample groups vAm is only poorly correlated with chronological age.

With respect to the perceptive age there are more significant and stronger correlations of the amplitude perturbation measures than with chronological age. Especially in the /a/ samples listeners presume a relation between amplitude stability and age, although it cannot be found in reality to the supposed extent. The perceptive age, just as the chronological age, does not seem to be related to vAm but strongly to APQ respectively to sAPQ.

Spectral energy distribution. The noise-to-harmonic ratio (NHR) which is supposed to be a measure of spectral noise is not correlated with chronological age. A significant ($p < 5\%$) correlation of NHR and perceived age can be shown in the /a/ samples, but the corresponding coefficients are below 0.3, which indicates that NHR is hardly relevant as a measure of vocal age.

Breathiness, reflected by the voice turbulence index (VTI), is not correlated to either of the age concepts.

The soft phonation index (SPI) shows the same correlation pattern as NHR, but differs in the strength of its relations. SPI is highly significantly correlated to perceived age in the /a/ samples. That means that listeners relate a proportional increase of the modulated energy in the lower frequency range (70 - 1600 Hz) to increasing age, at least in sustained /a/ realisations.

Vocal tremor. The measures of vocal tremor, which do not seem to have been analysed with respect to age until this study, offer surprising results: FTRI, defined as intensity of the strongest frequency modulation, seems to be the most reliable measure of chronological and vocal age in all vowel samples as a group. Only in the /a/-s samples the FTRI correlates with chronological age insignificantly lower than F0, which is still supposed to be the measure related most strongly to age. As a correlate of vocal age the FTRI is not significant in the /a/-s samples but in both /i/ groups, where F0 is without meaning to predict chronological and vocal age.

In the speech samples FTRI is uncorrelated to both concepts of age. The amplitude tremor intensity index does not seem to be of any importance to indicate age.

Speech tempo. The articulation rate (AR) as a measure of speech tempo decreases in read speech with increasing chronological and vocal age. AR in spontaneous speech cannot be considered as systematically correlated with either of the age concepts. But even in read speech the tempo measures considering no phonological information, namely the duration of speech breaks (t(br)) and the total duration required to read the given text (t), are better indicators of age than the articulation rate.

5. Conclusion

5.1. The primarily observed acoustic measures of age

Amplitude SD as a function of chronological age. Linville suggests that amplitude perturbation measures that reflect more gross fluctuations over time may be better discriminators of age than the cycle-to-cycle variations (cf. [3], p. 176 and p. 196). This can only partially be confirmed: Our data show the best correlations of amplitude perturbations and age if the fluctuations are averaged over 5 respectively 55 periods (APQ and sAPQ). Correlations get weaker if the time resolution of the measure is more precise as well as if it is less precise. The relative standard deviation of the peak-to-peak amplitude (vAm), the measure that probably corresponds best to what Linville named "Amp SD", reflects the most gross fluctuations over time but is nearly uncorrelated to both concepts of age.

Spectral noise as a function of perceived age. The noise-to-harmonic ratio (NHR), which is assumed to indicate hoarseness, only shows little importance for the perception of advanced age in sustained /a/ realisations and no correlation with chronological age in female voices. Breathiness, measured by the voice turbulence index (VTI), is uncorrelated with chronological and perceived age in female voices. Our results therefore confirm for female voices the results discovered by Ramig [8] for male voices.

Articulation rate as a function of perceived age. Tempo measures only seem to reveal information on age in read speech, suggesting that the slowing of women's articulation with age is better described as decline in reading or cognitive performance but cannot be explained by an altered physiology of the speaking apparatus. Moreover, the incorporation of the number of spoken syllables into the calculation of speech tempo leads to an impairment of its ability to predict a speaker's chronological and perceived age, indicating either that it is not actually the speech tempo that is correlated to age or that 'syllables per second' is not the most accurate measure of speech tempo.

5.2. Unexpected results

Frequency tremor intensity index. Since Xue & Deliyski [10] explicitly excluded the tremor measures from the compilation of age-related MDVP-parameters from their study ("[...]they lack sufficient validity report in the current literature"), further investigations seem to be necessary as a raised frequency tremor intensity index (FTRI) is the best predictor of increased chronological and vocal age on the basis of sustained vowels in this study. The relation of advanced age and a quivering voice seems obvious and the accumulation of highly significant correlations with perceived and chronological age across the different vowel types can not be considered as a random effect, even if it is not quite clear whether this measure really reflects what it was originally held to do.

Perturbation measurements applied to spontaneous speech. Since the amplitude perturbation measures of spontaneous speech yield highly significant middle-high to high correlations with both chronological and vocal age as well as the strongest correlation of this study, the computation of these measures should be checked more closely as to their relation to age. As all parts of the signal that MDVP identifies as quasi-periodic are considered for the calculation of these values, it is not easy to interpret them. Which are the amplitude differences that influence this measure, which contribute to its age-indicating ability and which impair it? These questions are to be answered in future investigations.

But again, as with the frequency tremor intensity index, the discovered correlations are far from being considered as random effects. Therefore, the amplitude perturbation measures applied to spontaneous speech are good indicators of age. APQ alone is able to explain about one third of the variance of the age values. Furthermore it may be assumed that the correlations can be improved by choosing a smoothing factor between 5 and 55 cycles.

5.3. Relevance of vowel onset

The age estimations are more precise respectively more concordant, if they result from speech samples that contain vowel onset. Hence, there has to be information on age carried by the vowel onset that cannot be found in the quasi-stationary part of the same vowel. This additional information is to be extracted from the manner of how vocal folds start to oscillate. Therefore, for the analysis of aging voices it seems to be advisable to examine sustained vowels with vowel onset and not only the quasi-stationary section of the signal.

5.4. Summary

Our experiments show that the parameters correlating with chronological and perceived age are the amplitude perturbation measures, at best normalised by the mean amplitude and with a time resolution between 20 and 250 ms, the frequency tremor intensity index and F0. Conditionally correlated are measures of frequency perturbation and speech tempo. Thus, this study confirms that adult female speakers speak more quiveringly, lower, rougher and read aloud more slowly as they grow older.

We also found evidence for the relevance of the vowel onset to recognize age more accurately.

6. Acknowledgements

This work was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, SE 462/5-1).

7. References

[1] **Hoit**, J., **Hixon**, K., **Altman**, M. & **Morgan**, W., "Speech breathing in women", *Journal of Speech and Hearing Research*, 32:353-365, 1989.

[2] **Linville**, S. E., "The aging voice". In Kent, R. D. & Ball, M. J. (Eds.) *Voice Quality Measurement*. Singular Thomson Learning, San Diego, 2000.

[3] **Linville**, S. E., *Vocal Aging*, Singular Thomson Learning, San Diego, 2001.

[4] **Orlikoff**, R., "The relationship of age and cardiovascular health to certain acoustic characteristics of male voices", *Journal of Speech and Hearing Research*, 33:450-457, 1990.

[5] **Oyer**, E. & **Deal**, L., "Temporal aspects of speech and the aging process", *Folia Phoniatrica (Basel)*, 37:109-112, 1985.

[6] **Ptacek**, P. & **Sander**, E., "Age recognition from voice", *Journal of Speech and Hearing Research*, 9:273-277, 1966.

[7] **Ramig**, L. & **Ringel**, R., "Effects of physiological aging on selected acoustic characteristics of voice", *Journal of Speech and Hearing Research*, 26:22-30, 1983.

[8] **Ramig**, L., "Effects of physiological aging on vowel spectral noise", *Journal of Gerontology*, 38:223-225, 1983.

[9] **Shipp**, T. & **Hollien**, H., "Perception of the aging male voice", *Journal of Speech and Hearing Research*, 12:703-710, 1969.

[10] **Xue**, S. A. & **Deliyski**, D., "Effects of aging on selected acoustic voice parameters: preliminary normative data and educational implications", *Educational Gerontology*, 21:159-168, 2001.