Acoustic Analysis of Adult Speaker Age

Schötz, Susanne

Published in:
Speaker Classification I, Lecture Notes in Computer Science

DOI:
10.1007/978-3-540-74200-5_5

2007

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Acoustic Analysis of Adult Speaker Age

Susanne Schötz

Lund University, Dept. of Linguistics and Phonetics,
Centre for Languages and Literature
susanne.schotz@ling.lu.se,
http://www.sol.lu.se

Abstract. Information about the age of the speaker is always present in speech. It is used as perceptual cues to age by human listeners, and can be measured acoustically and used by automatic age estimators. This chapter offers an introduction to the phonetic study of speaker age, with focus on what is known about the acoustic features which vary with age. The age-related acoustic variation in temporal as well as in laryngeally and supralaryngeally conditioned aspects of speech has been well documented. For example, features related to speech rate, sound pressure level (SPL) and fundamental frequency (F_0) have been studied extensively, and appear to be important correlates of speaker age. However, the relationships among the correlates appear to be rather complex, and are influenced by several factors. For instance, differences have been reported between correlates of female and male age, between speakers of good and poor physiological condition, between chronological age and perceived age, and also between different speech sample types (e.g. sustained vowels, read or spontaneous speech). More research is thus needed in order to build reliable automatic classifiers of speaker age.

Key words: Speaker age, Phonetics, Acoustic analysis, Acoustic correlates

1 Introduction

Every human being goes through the process of ageing. This is a very complex process, which affects us in numerous ways, including the way we speak. Our voices and speech patterns change from early childhood to old age. Although most changes occur in childhood and puberty, age-related variation can be observed throughout our adult lives into old age. Consequently, our age is reflected in our speech, and speaker age can be – and has been – studied using several methodological approaches, mainly acoustic analysis and perception experiments.

This chapter offers an introduction to the phonetic study of speaker age, with focus on acoustic variation. First, a summary is given of the age-related changes in the speech production mechanism, followed by short reviews of the study of speaker age from a perceptual and machine recognition perspective. The main part of the chapter comprises an overview of several known acoustic correlates of adult speaker age, including an overview of factors influencing these correlates.

2 Ageing of the speech production mechanism

From young adulthood to old age, the speech production mechanism undergoes numerous anatomical and physiological changes, which have not all been fully explored. For instance, there are substantial gender differences in the extent and timing of the ageing process [1, 2]. Moreover, the physiological differences between individuals seem to grow with advancing age [3]. It is also important, but sometimes difficult, to distinguish among age-related, disease-related and environment-related changes in speech. Linville [4, 2, 5] has provided excellent reviews of the numerous changes occurring in speech as we grow older. This section is mainly based on her work.

2.1 Respiratory system

Changes in the respiratory system affect speech breathing as well as the voice. The respiratory system reaches its full size after puberty but continues to change throughout adulthood to old age. Changes include decreased lung capacity (mainly due to loss of elasticity in lung tissue), stiffening of the thorax and weakening of respiratory muscles.

2.2 Larynx

The age-related changes of the larynx after it has reached its full size in puberty are numerous, and they affect mainly fundamental frequency and voice quality. Ossification of cartilages occurs later and is less extensive in females (fourth decade) than in males (third decade), while calcification probably occurs later than ossification in both females and males (cf. [6–9]).

Muscle atrophy occurs in all intrinsic laryngeal muscles. As research has focused on the vocal folds, we do not know to which extent other intrinsic muscles are affected. Whether there are any gender differences is also still unclear. The changes in the complex structure of the vocal folds with increased speaker age are substantial. Besides general degeneration and atrophy, the folds shorten in males (particularly after age 70). Also, the epithelium (the thin outer protective layer of tissue) thickens progressively in females, especially after age 70, while it thickens in males up to age 70 but then grows thinner again. The mucous glands reduce their secretions, leading to less hydrated vocal folds, particularly in males. There also seems to be some evidence of laryngeal nerve degeneration, as well as some changes in the blood supply to the laryngeal muscles.

2.3 Supralaryngeal system

Changes in the supralaryngeal system may also affect speech. The craniofacial skeleton grows continuously by about 3–5% from young adulthood to old age. Muscle atrophy occurs in the facial, mastication and pharyngeal muscles. A slight lowering of the larynx in the neck increases the length of the vocal tract.

Extensive degenerative changes occur in the temporomandibular joint, including a gradual reduction in size and reductions in blood supply. In the oral cavity, the mucosa grow thinner and lose elasticity, which is most apparent after age 70, and the mucosal surface roughens. Changes in the pharynx and soft palate include thinning of the epithelium, muscle atrophy and decreased sensation. The tongue surface becomes thinner and fissured, while the tongue muscles suffer from atrophy and fatty infiltration, beginning in the second or third decade.

2.4 Neuromuscular control

The effects of ageing on motor function can be observed in both the peripheral and the central nervous system. They may affect speech rate, co-ordination of articulators and breath support as well as the regulation of fundamental frequency (F_0). Peripheral changes include a type of "dying back" neuropathy, where the distal ends of the nerve fibres are affected earlier. Also, the number of motor units declines and conduction velocity slows down slightly.

Central changes include a decline in brain weight from age 20 to 90 by about 10% as well as a decrease in brain size. There are reports of decreases in the number of nerve cells in the cortex as well as age-related changes in these cells, which may slow down motor movements. In addition, dopamine levels in the brain may decline by up to 50%, leading to slower sensorimotor processes.

2.5 Female and male ageing

In addition to what has already been mentioned, a few more words deserve to be said about the differences between female and male ageing. These are often related to the timing and extent of age-related changes throughout life. One obvious difference is the dramatic changes occurring in males at puberty; another is that females experience greater changes around menopause. Nevertheless, the age-related changes in adults are generally greater in men than in women as regards (1) the extent of laryngeal structure change, (2) fine-motor control of laryngeal abductory and adductory movements, (3) tongue movements and (4) speech rate. It has also been noted that the mucous membranes in the larynx are more sensitive in females than in males and that females may thus be more vulnerable to age-related changes in this respect (P. Kitzing, personal communication, 31 January 2006). On the other hand, men and women display similar age-related changes in speech breathing.

3 Perception and automatic recognition of speaker age

Human listeners are able to judge speaker age at levels considerably better than chance. A large number of perception tests have been carried out with a various types of subjects, speech material and testing conditions. In recent years, a few studies on machine perception (or automatic recognition) of speaker age have emerged as well. This section briefly summarises human perception of age, and also describes a number of experiments on automatic recognition of speaker age.

3.1 Human perception of speaker age

Most people are able to estimate an individual's age from speech samples alone at accuracy levels significantly better than chance [10–12, 2], perhaps because of constant confrontation with this task throughout our lives, e.g. when listening to someone on the telephone or radio [13]. However, we are still unable to tell exactly how well listeners are able to judge speaker age. The numerous perception studies of speaker age have varied considerably in method and speech material, as well as in speaker and listener characteristics, and the results are often difficult to compare. Listeners' choice of cues and the accuracy obtained seem to depend on the type and length of the speech samples [14]. Moreover, the relationship of the perceptual cues used by listeners in age estimation with the acoustic correlates of chronological as well as perceived age has still not been fully established. In fact, the cues used by listeners to estimate speaker age do not always correspond to age-related changes which can be measured acoustically [4].

From a large number of studies concerning perception of speaker age, we have learned that human listeners are fairly good at estimating the age of an unknown (and unseen) speaker. Perceptual cues to speaker age include variation in pitch, speech rate, voice quality, articulation and phrasing. Moreover, it is likely that listeners use different acoustic cues and listening strategies when estimating the age of female and male speakers. For instance, F_0 seems somewhat more important for the age perception of female speakers than of male ones [15]. In addition, stimulus duration (i.e. longer speech samples, regardless of speech type) seems to be important when judging female speakers, while stimulus type (i.e. spontaneous speech, regardless of duration) seems to be more important in the case of male speakers [15].

Human perception of age is influenced by numerous phonetic as well as non-phonetic factors, e.g. the physiological state of the speaker, the age of the listener and the speech sample type. These factors have to be regarded in machine perception of age as well.

3.2 Automatic recognition of speaker age

Automatic recognition of age can be used to improve human–machine communication. If user age could be identified automatically, spoken dialogue systems could adapt their communication behaviour. For instance, the system could use more youthful language when talking to a teenager. It could also suggest age-adapted information, such as tourist attractions or directions.

As the number of children and elderly people who use computers in their daily lives increases, age-adapted speech recognition is becoming more important. Still, research on automatic age recognition is relatively scarce [16]. One explanation is that it certainly is not an easy task. Age cues are present in every phonetic dimension, and they are hard to separate from other speaker variation characteristics, such as physiological condition and dialect. This section summarises the relatively few attempts to build automatic age estimators.

Minematsu et al. [17, 18] built automatic classifiers of perceived age (PA, judged by 12 students) using linear discriminant analysis (LDA) and artificial neural networks (ANN) with mel frequency cepstral coefficients (MFCC), ΔMFCC and amplitude derivatives (ΔPower) as features. Eighty-six speakers (43 judged as elderly and 43 as non-elderly) were modelled using Gaussian mixture models (GMM) and normal distribution (ND). Elderly speakers were correctly identified in 90.9% of cases using LDA. The classifier was then improved by adding the features speech rate and local perturbation of power. This increased the identification rate to 95.3%.

Shafran et al. [16] used hidden Markov model (HMM) based classifiers with cepstral and F_0 features to recognise gender, age, dialect and emotion from a corpus consisting of 1,854 phone calls (65% female, 35% male callers) to a customer care system. The corpus contained a total of 5,147 utterances with an average length of 15 words divided into five age groups: (< 25, ≈ 25, 26–50, ≈ 50 and > 50). A trivial classifier assigning the most probable class label to all test points (33.3%) served as baseline. Results for age were 68.4% correct classifications using only cepstral features, and 70.2% correct using cepstral as well as F_0 features.

Minematsu et al. [19] conducted a study with male spekers (123 aged 6–12, 141 aged 20–60 and 143 aged 60–90). Thirty students in their early twenties estimated direct speaker age from single sentences. Each speaker was then modelled with GMM using MFCC, ΔMFCC and ΔPower as features. The two methods used for the machine estimations showed almost the same correlation between human judgements and machine estimation: the first method modelled PA as discrete labels (0.89), while the second one was based on the normal distributions of PA (0.88).

Müller et al. [20] compared six of the most common machine learning approaches for classification tasks – decision trees[1] (DT), ANN, k-nearest neighbour (kNN), naïve Bayes (NB) and support vector machines (SVM) – in a study of automatic classification of age group using jitter and shimmer as features. 393 speakers (about 10,000 utterances from 347 speakers over 60 years, about 5,000 utterances from 46 speakers under 60 years; gender distribution: 162 females, 231 males), were used in the study. All six methods performed significantly better than the baselines, which were simple classifiers always predicting the more frequently occurring class (elderly: 88%, male: 59%). ANN performed best with 96.57% correct age group estimations.

Müller et al. also used Bayesian networks (BN) to integrate a gender classifier with two age classifiers by first separately calculating the probability of a given speaker being female or male as well as being elderly or non-elderly, and then combining the results to obtain the most probable age and gender classification. This approach reduced errors likely to occur in a sequential classifier (gender first, then age), where failure to determine the correct gender strongly affects the performance of a gender-specific age classifier.

[1] C4.5 decision tree induction [21]

Müller [22, 23] further developed his approach for age and gender classification under the name AGENDER, with target applications such as mobile shopping and pedestrian navigation systems. Classification models were trained using the same five machine learning techniques as in [20], i.e. DT, ANN, kNN, NB and SVM, as well as an additional method: GMM. Features were extended to include jitter, shimmer, F_0, HNR (harmonics-to-noise ratio), speech rate (syllables per second), and pause duration and frequency. The number of speakers was increased to a total of 507 female and 657 male speakers, divided into four age classes for each gender. The majority of the speakers were children and seniors. The best accuracy for the four age classes was obtained with ANN (63.5%)

The author [15] carried out two studies with classification and regression trees (CART) to learn more about which acoustic-phonetic features are important in automatic age recognition. The first study used 50 features (e.g. measures of F_0, duration and formant frequencies) from the phoneme segments of 2,048 versions of one Swedish word (*rasa* [ˈʁɑ̀ːsa], 'collapse'), produced by 214 females and 214 males. The best CART for age group was 72% correct judgements, and the best correlation between direct chronological and estimated age was 0.45. Estimation accuracy was compared with that of human listeners. Although humans and CARTs used similar cues, the human listeners (mean error ± 8.89 years) were better judges of age than the CART estimators (± 14.45 years).

The second study used 748 speakers and 78 features to construct separate estimators of direct age for female, male and all speakers. CARTs were built for 390 single features, 13 feature groups (consisting of all features for one phonetic quality, e.g. F_1, B_1 and L_1) and five larger feature groups of all prosodic, all resonance, all inverse filtered, all spectral and all features. Results showed that F_0 and duration were the most important single features. Of 13 feature groups, F_0 and duration performed best for female speakers, while the formant groups of F2 and F3 were best for the male speakers. For the larger groups, the CART using all features was the best for female speakers, while the group with all prosodic features performed better for the male speakers. The best estimator of the second experiment (mean error ± 14.07 years) performed only marginally better than the one from the first study.

To sum up this section, automatic age estimation attempts have used MFCC as well as acoustic-phonetic features. The number and age range of speakers have varied among studies, as has the type of speech samples, the method used and the accuracy desired. In order to build reliable automatic age estimators, more knowledge is needed about how different acoustic features vary with speaker age for both genders as well as for different speech sample types and lengths.

4 Acoustic correlates of adult speaker age

A large number of acoustic features vary with speaker age. This variation is most clearly observable in children, but information about adult speaker age can also be – and has been studied from an acoustic-phonetic perspective. Acoustic variation has been found in temporal as well as in laryngeally and supralaryngeally conditioned aspects of speech. Moreover, the relationships among the numerous acoustic correlates of speaker age appear to be rather complex, and are influenced by several factors, of which several are further described in Section 5.

There are several comprehensive overviews of acoustic correlates of age. For instance, [24] has summarised research up till 1987, and [4, 2] has provided excellent reviews of known acoustic aspects of the ageing voice. Based on these sources as well as on several other studies, this section gives an overview of the acoustic features usually related to speaker age. Furthermore, in an attempt to clarify which features have been found to be important age correlates, some of the reported acoustic variation with increased age is summarised in Table 1. Variation with chronological age (CA) as well as with perceived age (PA) in women and men is described.

4.1 General variation

Old women and men alike demonstrate a general higher intra-subject as well as inter-subject variation of acoustic features when compared with young speakers. For example, increased variation has been found in some F_0 measures, as well as in speech rate (e.g. phoneme duration and VOT), vocal sound pressure level (SPL), jitter, shimmer and HNR [25, 26, 2]. More age-related differences have been found for male than female speakers [27], and higher correlations of acoustic features with PA than with CA have generally been observed [28]. Moreover, correlations seem to vary with speech sample type [28].

4.2 Speech rate

Temporal – static as well as dynamic – aspects of speech are strongly affected by the age of the speaker. The speech rate is linked to segment (syllable, phoneme, sub-phoneme, etc.) duration, to the number of speech segments per time unit and also to pause duration and frequency. A large number of studies have found a 20–25% decrease with older CA in speaking and reading rates. Increases have been found in consonant, vowel and sub-phonemic (prevoicing, plosive closure and release, vowel transition) durations as well as in pause duration and frequency [29, 30, 24, 31–35, 2, 28, 15, 36]. Women often demonstrate a smaller decrease in speech rate with older CA than men, or none at all [15, 37]. This feature also appears to show a larger inter-speaker variation for female speakers [15]. Slower speech rates, a larger number of breaths and longer pause durations have been related to old male and female PA [28].

The results for the sub-phonemic segment voice onset time (VOT) are rather confusing. Some studies have found elderly (CA) women and men to exhibit shorter overall VOTs than younger people [29, 31, 38]. However, increased VOT with older male CA has also been observed [26]. Other researchers have reported only subtle differences and increased variation with advancing age [39, 2, 15]. It has also been suggested that age-related differences in VOT is related to phonetic context and perhaps even languages [40, 15].

Table 1. Some reported acoustic variation with increased chronological age (CA) and perceptual age (PA) in female and male adult speakers (*decr.*: decrease, *dur.*: duration, *flat.*: flatter, *freq.*: frequency, *incr.*: increase, *no*: no change, *sp.*: spectral, *steep.*: steeper). Please refer to the text for details (adapted from [15])

Group	*Feature*	*Variation with increasing adult age*			
		Female		*Male*	
		CA	*PA*	*CA*	*PA*
general	variation	incr.		incr.	
	overall changes	few	more	many	more
speech rate	syllables/second	decr. or no		decr.	decr.
	utterance dur.	incr.	incr.	incr.	incr.
	phoneme dur.	incr.		incr.	
	VOT	incr., decr. or no		incr., decr. or no	
	pause freq.&dur.	incr.	incr.	incr.	incr.
sound pressure level (SPL)	mean SPL	no		incr. or decr.	
	max. SPL range	decr.		decr.	
	amplitude SD	incr. or no	incr. or no	incr. or no	incr.
F_0	mean F_0	first no or decr., then decr., incr. or no	decr.	first decr., then incr.	first decr., then incr.
	F_0 range	first incr., then decr.	incr. or no	first incr., then decr. or no	
	F_0 SD	incr. or no	incr. or no	incr., decr. or no	incr.
tremor	vocal tremor	incr. or no	incr.	no	
jitter & shimmer	jitter	incr. or no	incr. or no	incr. or no	
	shimmer	incr. or no	incr. or no	incr. or decr.	
sp. noise	HNR	decr. or no		varying or no	
	NHR	incr. or no	incr. or no	incr. or varying	
sp. energy distribution	sp. tilt	flat. or no		steep., flat. or no	
	sp. tilt (LTAS)	steep. or varying		flat. or varying	
	sp. emphasis	no or varying		no or varying	
	sp. balance	no		no	
resonance	F_1	decr. or no	decr.	decr. or no	decr.
	F_2	incr., decr. or no	decr.	incr., decr. or no	decr.
	F_3–F_4	decr. or no.		decr. or no	
	F_1–F_3(LTAS)	decr.	no	decr.	decr.

4.3 Sound pressure level (SPL)

Conversational speech SPL (also called intensity level) appears to remain stable or decrease slightly with increased CA, but has also been reported to increase for men after age 70, even for speakers without hearing loss [41, 24, 42, 2, 15]. The habitual SPL range in vowels is likely to increase with advancing female and male CA, and may be an important correlate of speaker age [43, 15, 36]. However, the maximum vowel SPL range seems to decrease in both women and men, while minimum SPL levels increase for women (to the author's knowledge, no studies have been made concerning men) with advancing CA [31, 2].

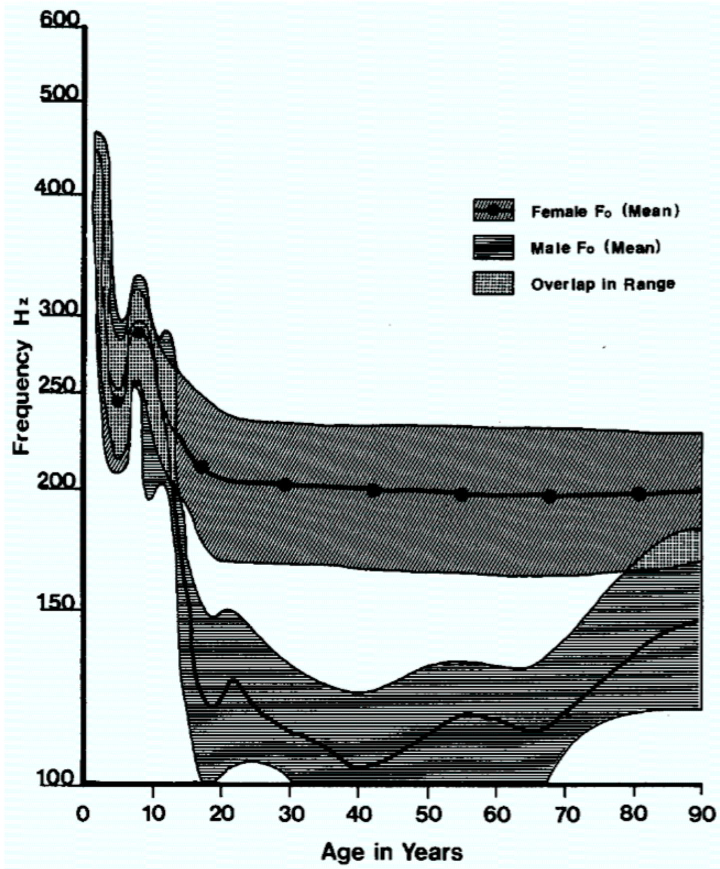

Fig. 1. Speaking F_0 and its standard deviation as a function of speaker age (1–90 years) for female and male speakers (source: [44])

4.4 Fundamental frequency (F_0)

F_0 patterns in speech related to CA are different for women and men, as shown in Figure 1. Female F_0 has been found to remain fairly constant until menopause, when a drop (of about 10–15 Hz) usually occurs. F_0 then remains stable into old age, but may also rise or lower slightly [45, 46, 27, 47, 4, 43]. Observations of decreasing F_0 from age 20 to 50 in females have also reported [15, 36]. A lower F_0 is also associated with older female PA [2, 28]. In males, F_0 lowers slightly (by about 10 Hz) from young adulthood to middle CA, but then rises considerably (about 35 Hz) with old CA [48, 46, 2, 15, 36]. Higher F_0 has been reported to be a cue to old male PA [33, 2]. However, there are also studies which have failed to find correlations between mean F_0 and CA in men [3]. Moreover, the way changes in F_0 relate to perceptual cues is not in line with the above findings. For instance, listeners have reported lower male F_0 to be a cue to older age [2].

Maximum phonational frequency range – i.e. the complete range of frequencies which a speaker can produce, from the lowest (without creak) to the highest tone (including falsetto) – expands in the lower end following menopause in females, but is restricted in both the upper and lower ends later in life [49, 2]. Contradictory findings suggest that men either undergo similar changes in F_0 range as women [50, 2], or that old and young males do not differ in F_0 range unless physiological condition and state of health are taken into account [3, 51]. A larger habitual F_0 range has been observed for the vowel /a/ in both women and men of old CA [43]. Relatively stable habitual F_0 range values for both genders until about the age of 60, followed by an increase (females) or decrease (males) have also been observed [15].

4.5 Variation in F_0 and amplitude

Fundamental frequency and amplitude instability and variation are related to various voice qualities. Jitter and shimmer (see p. 10) are often connected with harshness, hoarseness or vocal roughness, while increases in the more gross F_0 variation, as measured in standard deviation (F_0 SD), may cause vocal tremor or a “wobbling” voice quality [24, 52, 2].

Higher F_0 SD (with greater variation for females) has been found in both men and women with advancing CA and PA [24, 49, 33, 2, 43], but sometimes only a minor correlation has been reported, or none at all [3, 28, 15]. Substantial increases in fundamental amplitude standard deviation (Amp SD) have been demonstrated in older men and women, and have been associated with both CA and PA [53, 43]. However, relatively stable Amp SD values with advancing age have been reported as well [15], and Brückl and Sendlmeier [28] found a strong positive correlation with female CA and PA only in spontaneous speech but almost none in sustained vowels or read speech.

Jitter and shimmer are defined as period-to-period fluctuations in vocal fold frequency and amplitude, as shown in Figure 2, and they are considered to be correlates of rough or hoarse voice quality. These features have often been analysed in acoustic studies of age using a number of measures with varying results. Although sometimes no correlation with age has been found for jitter [51, 32, 54, 55, 15, 36], other researchers have reported increased jitter levels for older female and male CA (but not PA) [49, 53, 26, 43, 22]. However, higher and more variable jitter values seem to be more related to physiological health than to age [3, 53, 2, 28].

Higher shimmer levels have been found for older female CA and PA as well as for older male CA (independently of health) [3, 51, 53, 26, 43, 22]. However, stable (females) or decreasing shimmer levels after age 40 (males) have also been observed [15, 36]. Other studies have found shimmer to correlate strongly with CA and PA only in spontaneous speech samples (but not in read speech or in prolonged vowels) [28]. Other studies have observed correlations of shimmer and CA in sustained vowels, but only when 80-year-olds were compared with younger age groups [55].

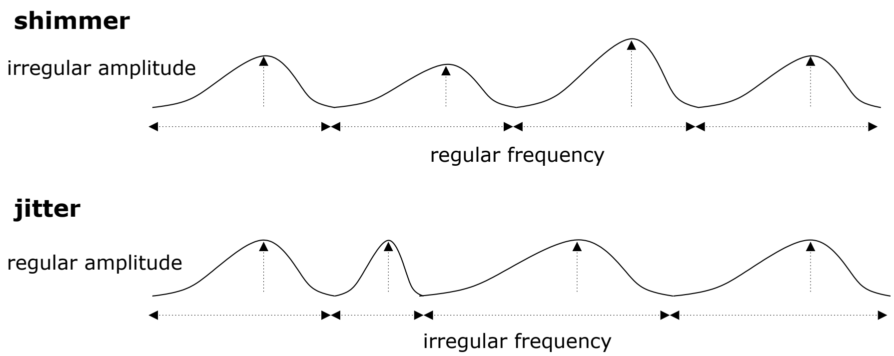

Fig. 2. Irregularities (microvariations) in vocal fold movements can be measured as shimmer (variation in amplitude) and jitter (variation in frequency) (after [22])

Linville [2] concludes that it is impossible to draw any firm conclusions as to the effect of ageing on jitter and shimmer since several factors, including sound pressure level, mean F_0, analysis system differences and individual health and fitness variables, appear to have a strong effect on these measures, especially in women. Moreover, the large number of measures used for these features and the differences in speech material used in various studies also appear to contribute to the problem with comparison of results.

4.6 Other voice measures

Spectral tilt (ST), spectral emphasis (SE) and spectral balance (SB) are all measures of the relative energy levels in different frequency bands of the spectrum [56, 57]. ST usually represents the slope – i.e. the difference between the energy levels of two different frequency bands – of the source (inverse filtered) spectrum in dB per octave. SE is a measure of the relative energy levels in the higher frequency bands, while SB is often measured in four contiguous frequency bands. The three measures have sometimes been defined differently [58, 59, 57, 60].

ST has been observed either to flatten (i.e. the energy in the frequency band 4–5 kHz increased with female and male CA) in some vowels, or to remain relatively stable until age 60 (females) or 80 (males), where an increase follows [15]. A longitudinal study found a steeper spectral tilt in old men compared with the same men when young [61]. SE and SB have been found not to correlate significantly with CA [59, 15].

The age-related variation of the energy distribution in long-term average spectra (LTAS) has also been studied to some extent. An LTAS is an averaged spectrum of all voiced sounds across a relatively long speech sample. Elderly women have been observed to have higher spectral levels at 320, 6080, 6240, 6400, 6560 and 6720 Hz but lower levels at 3040 and 3200 Hz than young women, and a tendency for older women to have higher levels at 160 Hz has been found as well [62]. Somewhat higher female LTAS levels with advancing age have been observed for 160 (but only from age group 40 to 70), 320 and 2240–2560 Hz, while slightly lower levels with increased age were found at about 5920–7200 Hz [15]. Differences in spectral amplitude have been found between old and young men, too. Old males have demonstrated higher LTAS levels at 160 Hz and lower levels at 1600 Hz than young males [62]. Moderately higher LTAS amplitudes with advancing male age have also been found at 160 (but only for age group 40 to 70), 320 and 1760–2080 Hz [15]. A strong spectral attenuation of high frequencies has also been observed in LTAS at older CA and PA in males, but not in females [63].

Spectral noise is defined as the unmodulated aperiodic energy in vowel spectra [2]. It is considered an acoustic correlate of breathy and harsh or hoarse voice quality [64, 65], and has been analysed using various methods. Visual analysis of spectral noise in spectrograms has shown that this feature is much more strongly correlated with physiological condition than with CA [66].

The harmonics-to-noise ratio (HNR) is a measure that quantifies the amount of additive noise in the voice signal, and it can be calculated in several ways [67, 68]. The ratio reflects the dominance of the periodic level over the aperiodic one, as quantified in dB. HNR has sometimes been reported to decrease with older female CA [54], or to increase with younger male CA [69], while other researchers have failed to find strong correlations with CA in females [69] or both genders [70, 15, 36]. No studies exist (to the author's knowledge) of HNR in relation to PA.

Other measures of spectral noise used in acoustic studies of speaker age include the parameters VTI, SPI and NHR of the commercial voice quality analysis software Multi-Dimensional Voice Program (MDVP, see e.g. [71]). Voice turbulence index (VTI) is a measure of the relative energy level in high-frequency noise. It is calculated as the average ratio of the inharmonic spectral energy in the 2.85.8 kHz range to the harmonic spectral energy in the 0.074.5 kHz range. Soft phonation index (SPI) measures the relative energy in low-frequency noise, calculated as the average ratio of the lower (0.071.6 kHz) to the higher (1.64.5 kHz) frequency harmonic energy. The noise-to-harmonics ratio (NHR) is the average ratio between noise in the frequency band 1.5–4.5 kHz and the harmonic energy in the frequency band 0.07–4.5 kHz; it is sometimes referred to as a low-frequency harmonics-to-noise ratio [72]. Increased values for all three features in women and men of older CA have been reported [43]. Other researchers have failed to find strong correlations of these features with female and male CA [73, 15], though weak (NHR but not VTI) and strong (SPI) positive correlations with female PA [28] or have also been observed.

Vocal tremor can be measured using the MDVP parameters FTRI (intensity of the strongest frequency modulation) and ATRI (intensity of the strongest amplitude modulation). FTRI (but not ATRI) has been found to increase with both female CA and PA in vowels, but not in read or spontaneous speech [28].

4.7 Resonance measures

Research has revealed that age-related changes in the supralaryngeal structures provide acoustic cues to adult speaker age [10, 13, 74]. However, there are relatively few studies of the age-related changes in the vocal tract resonance features.

Formant frequencies in vowels have been reported to lower with female and male CA and PA owing to increased vocal tract length [75, 76]. There also seems to be a trend towards vowel centralisation (or reduction) for old CA [77, 78]. It appears that some old speakers centralise more than others, suggesting an increase in formant frequency variation across speakers of old CA [2]. Moreover, different results have been observed for different vowels. F_1 has been found to decrease with older female age in [yː], and to drop substantially with age for both genders at about age 40 in [ɛː], while other vowels ([a], [ɑː] and [uː]) did not vary much with age in either gender [15, 36]. In the same study, F_2 was found to increase with advancing age in [ɑː] and [ɛː] for both genders. In [a] and [uː], F_2 tended to decrease slightly, interrupted by increases and peaks at age group 40 in both genders. A fairly stable F_2 was observed in [yː]. F_3, F_4 and F_5 have been found to show somewhat different patterns depending on gender and vowel quality. Often (but far from always) decreases were observed from age class 20 to 30, followed by little change or a very slight increase, with an occasional rise or fall after age 80 for one or both genders [15]. For PA, formant information seems to lose its significance when F_0 information is present [52].

Energy peaks in long-term average spectra (LTAS), corresponding to the average formant frequencies across all vowels in a speech sample, have been studied by Linville and Rens [79]. They found a significant lowering of peaks 1, 2 and 3 (corresponding to F_1–F_3) with old female CA, and a significantly lower peak 1 (F_1) in old male CA. Moreover, the age-related lowering of peaks was greater in females than in males.

To sum up this section, previous research has found numerous acoustic correlates of chronological and perceptual speaker age. Some features, such as measures of F_0 and speech rate, have been found to be more important than others and have thus been investigated to a larger extent. In addition, there are also a number of factors which may also influence acoustic analysis of speaker age. Some of these factors are described in the following section.

5 Factors which may influence acoustic analysis of speaker age

Several factors (besides age) may affect the analysis outcome in acoustic studies of speaker age. These are often related to the material and the methods used, and may contribute to the divergent and sometimes even contradictory results found in different studies. Differences have been reported between correlates of female and male age, between speakers of good and poor physiological condition, between chronological age (the age of a speaker as measured in time from birth) and perceived age (the mean age of a speaker as estimated by a group of listeners), and also between different speech sample types (e.g. sustained vowels and read or spontaneous speech). This section offers a brief overview of some of the factors which may influence analysis results.

5.1 Speaker-related factors

Speaker-related factors include physical (anatomical and physiological) attributes such as gender, race, weight, health and physiological condition. Women and men differ in several vocal characteristics. Some can be explained by anatomical differences while others, such as the paralinguistic use of breathy voice quality, appear to be learned behaviours [80]. Differences in body physiology, vocal training and medical condition may also affect the age-related variation in speech [3, 81–83], including effects of medication [61] and cigarette smoking [84]. For instance, smokers generally exhibit lower F_0 than non-smokers [85], while professional sopranos and tenors have a higher F_0 than age-matched non-singers [44]. Furthermore, age-related differences in habitual F_0 seem less prominent or even absent in singers and other voice professionals [82].

Cultural, social and psychological factors, including speaker language, dialect, sociolect, emotional state and attitude, may influence and even mask age-related acoustic variation. For instance, there are language-related, dialectal and attitudinal differences in habitual F_0, HNR and shimmer levels [47, 86]. Moreover, consideration must also be given to the fact that voice settings are more or less subject to swings in fashion [7], and that the pronunciation of a language is constantly changing. Young individuals often wish to speak differently from their parents [87]. One example is the increased use of the more open allophones [æː] and [œː] of the /ɛ/ and /ø/ phonemes in Swedish [88]. Another example is the growing use of the glottal stop in British English [87].

5.2 Speech-material-related factors

Speech-material-related factors include the number and age distribution of the speakers and the duration and speech type (and number of speech types) of the speech samples analysed. Fewer speakers will yield less reliable results, as will an unbalanced (for age) speech corpus. Valid measurements of some features are obviously obtained more easily from certain speech types. For instance, formant frequencies are more reliably measured in sustained vowels than in connected speech, and calculations of the average number of syllables per second are more reliable in longer speech samples. Moreover, studies which have used more than one speech type have sometimes found contradictory results for different speech types. One example is Brückl and Sendlmeier [28], who found that vocal tremor correlated with age in sustained vowels, but not in read or spontaneous speech.

5.3 Methodological factors

Methodological factors, such as differences in recording and analysis equipment and techniques, may strongly influence the outcome of acoustic analyses. One example concerns the vocal effort made by speakers to adapt to the distance to a listener or a microphone, which may affect speech rate, voice quality, measures of F_0 and even some formant frequencies [89]. Different measurement techniques could also be one reason why, for instance, it has not yet been possible to draw any firm conclusions as to the effect of ageing on jitter and shimmer [2].

Another major methodological factor in acoustic studies of speaker age is whether the findings are related to chronological or perceived age. In automatic age recognition applications, the goal is in many cases to identify speakers' actual CA, and not the mean PA as estimated by a group of listeners. However, if only CA is considered in an acoustic study, no knowledge about the relative importance of the correlates to listeners will be gained [11]. On the other hand, when the acoustic correlates of PA are examined, the age judgements of a group of listeners – often quite small – will have to be trusted. Since PA is a subjective measure, results may not be reliable, as listener characteristics (gender, age, etc.) affect the age estimates. Thus, the purpose of each study or application will have to determine whether CA or PA is chosen as the frame of reference.

A connected question is whether we should use archival recordings in combination with recent ones of the same speakers (longitudinal studies) or speech samples from different speakers recorded close in time (cross-sectional studies). Although it may be tempting to use longitudinal data because of the invariant speaker-specific parameters, several aspects which may affect the results should be regarded. Differences in recording equipment and technical sound quality may yield unreliable results. Moreover, voice communication habits may change over time, one example being that Australian women aged 18–25 years recorded in 1993 had significantly lower F_0 levels than women of the same age recorded in 1945 [90]. Another example concerns VOT and F_0 SD. Several cross-sectional studies have reported that VOT decreased and F_0 SD increased in males with advancing age. However, in a longitudinal study of male speakers recorded twice over a period of 30 years, [91] found the opposite results.

In spite of the numerous factors which may affect acoustic analysis, different studies have agreed on several acoustic correlates of speaker age. However, many experiments have varied in the number and choice of speakers and acoustic features, as well as in speech material and method. Some studies have reduced the effect of certain factors by controlling variables or by using a large material.

In summary, speaker age is a very complex characteristic of speech. It leaves traces in all acoustic-phonetic dimensions and it is influenced by numerous other factors, such as physiological condition. Studying it is by no means a trivial task. The studies carried out so far have varied greatly in the type of speech material used (read, spontaneous, prolonged vowels etc.) as well as in analysis method (number and kind of of features investigated). More research is needed with a larger and more systematically varied material and methods to fully explore the age-related acoustic variation in speech and to identify optimal combinations of features for automatic recognition of speaker age.

References

1. J. M. Beck, "Organic variation of the vocal apparatus," in *The Handbook of Phonetic Sciences*, W. J. Hardcastle and J. Laver, Eds., pp. 256–297. Blackwell Publ., Oxford, 1997.
2. S. E. Linville, *Vocal Aging*, Singular Thomson Learning, San Diego, CA, 2001.
3. L. A. Ramig and R. L Ringel, "Effects of physiological aging on selected acoustic characteristics of voice," *Journal of Speech and Hearing Research*, vol. 26, pp. 22–30, 1983.
4. S. E. Linville, "The aging voice," in *Voice Quality Measurement*, R. D. Kent and M. J. Ball, Eds., pp. 359–376. Singular Thomson Learning, San Diego, CA, 2000.
5. S. E. Linville, "The aging voice," *The American Speech-Language-Hearing Association (ASHA) Leader, Oct. 19*, pp. 12, 21, 2004.
6. A. Jurik, "Ossification and calcification of the laryngeal skeleton," *Acta Radiol Diagn.*, vol. 25, pp. 17–22, 1984.
7. P. Lindblad, *Rösten*, Studentlitteratur, Lund, 1992.
8. R. A. Dedivitis, M. Abrahão, M. J. Simães, O. A. Mora, and O. W. Cervantes, "Aging histological changes in the cartilages of the cricoarytenoid joint," *Acta Cir Bras [serial online] 19*, vol. Retrieved 16 August 2006, from http://www.scielo.br/acb, 2004.
9. M. Mupparapu and A. Vuppalapati, "Ossification of laryngeal cartilages on lateral cephalometric radiographs," *The Angle Orthodontist*, vol. 75, pp. 196–201, 2005.
10. P. H. Ptacek and E. K. Sander, "Age recognition from voice," *Journal of Speech and Hearing Research*, vol. 9, pp. 273–277, 1966.
11. W. J. Ryan and K. W. Burk, "Perceptual and acoustic correlates of aging in the speech of males," *Journal of Communication Disorders*, vol. 7, pp. 181–192, 1974.
12. R. Huntley, H. Hollien, and T. Shipp, "Influences of listener characteristics on perceived age estimations," *Journal of Voice*, vol. 1, pp. 49–52, 1987.
13. T. Shipp and H. Hollien, "Perception of the aging male voice," *Journal of Speech and Hearing Research*, vol. 12, pp. 703–710, 1969.
14. L. Ramig, "Aging speech: Physiological and sociological aspects," *Language and Communication*, vol. 6, pp. 25–34, 1986.
15. S. Schötz, *Perception, Analysis and Synthesis of Speaker Age.*, Ph.D. thesis, Travaux de l'Institut de linguistique de Lund 47. Lund: Dept. of Linguistics and Phonetics, Lund University, 2006.
16. I. Shafran, M. Riley, and M. Mohri, "Voice signatures," in *Proc. of The 8th IEEE Automatic Speech Recognition and Understanding Workshop*. 2003, St. Thomas, U.S. Virgin Islands.
17. N. Minematsu, M. Sekiguchi, and K. Hirose, "Automatic estimation of one's age with his/her speech based upon acoustic modeling techniques of speakers," in *Proc. of ICASSP 2002*. 2002, pp. 137–140, Orlando, FL.
18. N. Minematsu, M. Sekiguchi, and K. Hirose, "Performance improvement in estimating subjective agedness with prosodic features," in *Proc. of Speech Prosody 2002*. 2002, pp. 507–510, Aix-en-Provence.
19. N. Minematsu, K. Yamauchi, and K. Hirose, "Automatic estimation of perceptual age using speaker modeling techniques," in *Proc. of Eurospeech*. 2003, pp. 3005–3008, Geneva.
20. C. Müller, F. Wittig, and J. Baus, "Exploiting speech for recognizing elderly users to respond to their special needs," in *Proc. of Eurospeech 2003*. 2003, pp. 1305–1308, Geneva.

21. J. R. Quinlan, *C4.5: Programs for Machine Learning*, Morgan Kauffman, San Mateo, CA, 1993.
22. C. Müller, *Zweistufige kontextsensitive Sprecherklassifikation am Beispiel von Alter und Geschlecht*, Ph.D. thesis, Computer Science Institute, Saarland University, 2005.
23. C. Müller, "Automatic recognition of speakers' age and gender on the basis of empirical studies," in *Proc. of Interspeech 2006*, Pittsburgh, PA, 2006.
24. H. Hollien, "Old voices: What do we really know about them?," *Journal of Voice*, vol. 1, pp. 2–13, 1987.
25. R. J. Morris and W. S. Brown, "Age-related differences in speech variability among women," *Journal of Communication Disorders*, vol. 27, pp. 49–64, 1994.
26. W. Decoster, *Akoestische kenmerken van de ouder wordene stem*, Ph.D. thesis, Leuwen: Leuwen University Press (Summary in English), 1998.
27. M. B. Higgins and J. H. Saxman, "A comparison of selected phonatory behaviours of healthy aged and young adults," *Journal of Speech and Hearing Research*, vol. 13, pp. 1000–1010, 1991.
28. M. Brückl and W. Sendlmeier, "Aging female voices: An acoustic and perceptive analysis," in *Proc. of VOQUAL'03*. 2003, pp. 163–168, Geneva.
29. B. Benjamin, "Phonological performance in gerontological speech," *Journal of Psycholinguistic Research*, vol. 11, pp. 159–167, 1982.
30. E. Oyer and L. Deal, "Temporal aspects of speech and the aging process," *Folia Phoniatrica (Basel)*, vol. 37, pp. 109–112, 1985.
31. R. J. Morris and W. S. Brown, "Age-related voice measures among adult women," *Journal of Voice*, vol. 1, pp. 38–43, 1987.
32. W. S. Brown, R. J. Morris, and J. F. Michel, "Vocal jitter in young adult and aged female voices," *Journal of Voice*, vol. 3, pp. 113–119, 1989.
33. T. Shipp, Y. Qi, R. Huntley, and H. Hollien, "Acoustic and temporal correlates of perceived age," *Journal of Voice*, vol. 6, pp. 211–216, 1992.
34. J. D. Amerman and M. M. Parnell, "Speech timing strategies in elderly adults," *Journal of Voice*, vol. 20, pp. 65–67, 1992.
35. E. B. Slawinski, "Acoustic correlates of [b] and [w] produced by normal young to elderly adults," *Journal of the Acoustical Society of America*, vol. 95(4), pp. 2221–2230, 1994.
36. S. Schötz and C. Müller, "A study of acoustic correlates of speaker age," in *Speaker Classification*, vol. 4343 of *Lecture Notes in Computer Science / Artificial Intelligence*. Springer, Heidelberg - Berlin - New York, 2007, this issue.
37. J. Hoit, K. Hixon, M. Altman, and W. Morgan, "Speech breathing in women," *Journal of Speech and Hearing Research*, vol. 32, pp. 353–365, 1989.
38. K. Stölten and O. Engstrand, "Effects of sex and age in the Arjeplog dialect: A listening test and measurements of preaspiration and vot," in *Proc. of Fonetik 2002*. 2002, vol. 44, pp. 29–32, TMH-QPSR.
39. L. Petrosino, R. D. Colcord, K. B. Kurcz, and R.J. Yonker, "Voice onset time of velar stop productions in aged speakers," *Journal of Perceptual and Motor Skills*, vol. 76, pp. 83–88, 1993.
40. G. Neiman, R. Kluch, and E. Shuey, "Voice onset time in young and 70-year-old women," *Journal of Speech and Hearing Research*, vol. 26, pp. 118–123, 1983.
41. W. J. Ryan, "Acoustic aspects of the aging voice," *Journal of Gerontology*, vol. 27, pp. 256–268, 1972.
42. R. J. Morris and W. S. Brown, "Age-related differences in speech intensity among adult females," *Folia Phoniatrica (Basel)*, vol. 46, pp. 64–69, 1994.

43. S. A. Xue and D. Deliyski, "Effects of aging on selected acoustic voice parameters: Preliminary normative data and educational implications," *Educational Gerontology*, vol. 21, pp. 159–168, 2001.
44. W. S. Brown, R. J. Morris, H. Hollien, and E. Howell, "Speaking fundamental frequency characteristics as a function of age and professional singing," *Journal of Voice*, vol. 5, pp. 310–315, 1991.
45. H. Hollien and T. Shipp, "Speaking fundamenal frequency and chronological age in males," *Journal of Speech and Hearing Research*, vol. 15, pp. 155–159, 1972.
46. P. Kitzing, *Glottografisk frekvensindikering: En undersökningsmetod för mätning av röstläge och röstomfång samt framställning av röstfrekvensdistributionen.*, Ph.D. thesis, Lund University, Malmö, 1979.
47. H. Traunmüller and A. Eriksson, "The frequency range of the voice fundamental in the speech of male and female adults. [manuscript]," Retrieved 2 January 2006, from http://www.ling.su.se/staff/hartmut/aktupub.htm, 1995.
48. E. Mysak, "Pitch and duration characteristics of older males," *Journal of Speech and Hearing Research*, vol. 2, pp. 46–54, 1959.
49. S. E. Linville, "Acoustic-perceptual studies of aging voice in women," *Journal of Voice*, vol. 1, pp. 44–48, 1987.
50. P. H. Ptacek, E. K. Sander, W. H. Maloney, and C. C. R. Jackson, "Phonatory and related changes with advanced age," *Journal of Speech and Hearing Research*, vol. 9, pp. 350–360, 1966.
51. R. L. Ringel and W. J. Chodzko-Zajko, "Vocal indices of biological age," *Journal of Voice*, vol. 1, pp. 31–37, 1987.
52. S. E. Linville, "The sound of senescence," *Journal of Voice*, vol. 10, pp. 190–200, 1996.
53. R. Orlikoff, "The relationship of age and cardiovascular health to certain acoustic characteristics of male voices," *Journal of Speech and Hearing Research*, vol. 33, pp. 450–457, 1990.
54. C. T: Ferrand, "Harmonics-to-noise ratio: An index of vocal aging," *Journal of Voice*, vol. 16, pp. 480–487, 2002.
55. E. Shuey, J. Herr-McCauley, C. Prohaska, and K. Martin, "Perturbation measures and chronologic age," Presented at the annual convention of the American Speech-Language-Hearing Association (ASHA), November 13–15, Chicago, IL, 2003.
56. N. Campbell, "Loudness, spectral tilt, and perceived prominence in dialogues," in *Proc. of ICPhS 95*. 1995, vol. 3, pp. 676–679, Stockholm.
57. M. Heldner, "Spectral emphasis as an additional source of information in accent detection," in *Prosody 2001: ISCA Tutorial and Research Workshop on Prosody in Speech Recognition and Understanding*, M. Bacchiani, J. Hirschberg, D. Litman, and M. Ostendorf, Eds. 2001, pp. 57–60, Red Bank, NJ: ISCA.
58. A. M. C. Sluijter and V. J. van Heuven, "Spectral balance as an acoustic correlate of linguistic stress," *Journal of the Acoustical Society of America*, vol. 100, pp. 2471–2485, 1996.
59. H. Traunmüller, "Perception of speaker sex, age, and vocal effort," in *PHONUM, Reports in Phonetics 4*, pp. 183–186. Umeå University, 1997.
60. E. O'Leidhin and P. Murphy, "Analysis of Spectral Measures for Voiced Speech with Varying Noise and Pertubation Levels," in *Proc. of ICASSP*, 2005, vol. 1, pp. 869–872.
61. W. Decoster and F. Debruyne, "Changes in spectral measures and voice onset time with age: A cross-sectional and a longitudinal study," *Folia Phoniatrica et Logopaedica*, vol. 49, pp. 269–280, 1997.

62. S. E. Linville, "Source characteristics of aged voice assessed from long-term average spectra," *Journal of Voice*, vol. 16, pp. 472–479, 2002.
63. R. Winkler, M. Brückl, and W. Sendlmeier, "The aging voice: an acoustic, electroglottographic and perceptive analysis of male and female voices," in *Proc. of ICPhS 03*. 2003, pp. 2869–2872, Barcelona.
64. A. McAllister, J. Sundberg, and S. Hibi, "Acoustic measurements and perceptual evaluation of hoarseness in children's voices," *Logopedics Phoniatrics Vocology*, vol. 23, pp. 27–38, 1998.
65. J. Kreiman and B. R. Gerratt, "Perception of aperiodicity in pathological voice," *Journal of the Acoustical Society of America*, vol. 117, pp. 2201–2211, 2005.
66. L. A. Ramig, "Effects of physiological aging on vowel spectral noise," *Journal of Gerontology*, vol. 38, pp. 223–225, 1983.
67. G. de Krom, "A cepstrum-based technique for determining a harmonics-to-noise ratio in speech signals," *Journal of Speech and Hearing Research*, vol. 36, pp. 224–266, 1993.
68. P. Boersma, "Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound," in *Proc. of the Institute of Phonetic Sciences*. 1993, vol. 17, pp. 97–110, University of Amsterdam.
69. C. C. Wang and H. T. Huang, "Voice acoustic analysis of normal Taiwanese adults," *J Chin Med Assoc*, vol. 67, pp. 179–184, 2004.
70. S. Schötz, "Prosodic cues in human and machine estimation of female and male speaker age," in *Nordic Prosody. Proc. of the IXth Conference, Lund, 2004*, G. Bruce and M. Horne, Eds., pp. 215–223. Frankfurt am Main: P. Lang, 2006.
71. D. Deliyski and C. Gress, "Intersystem reliability of MDVP for DOS and Windows 95/98," Paper presented at the 1998 Annual Convention of American Speech-Language-Hearing Association, San Antonio, Texas, 1998.
72. G. Pereira Jotz, O. Cervantes, M Abrahao, F. A. Parente Settanni, and E. Carrara de Angelis, "Noise-to-harmonics ratio as an acoustic measure of voice disorders in boys," *Journal of Voice*, vol. 16, pp. 28–31, 2002.
73. E. Shuey, J. Herr-McCauley, and M. Anders, "Indices of turbulence in aging voice," Presented at the annual convention of the American Speech-Language-Hearing Association (ASHA), November 17-20, Philadelphia, PA, 2004.
74. R. Jacques and M. Rastatter, "Recognition of speaker age from selected acoustic features as perceived by normal young and older listeners," *Folia Phoniatrica (Basel)*, vol. 42, pp. 118–124, 1990.
75. W. Endres, W. Bambach, and G Flösser, "Voice spectrograms as a function of age, voice disguise, and voice imitation," *Journal of the Acoustical Society of America*, vol. 49, pp. 1842–1848, 1971.
76. S. E. Linville and H. Fisher, "Acoustic characteristics of perceived versus actual vocal age in controlled phonation by adult females," *Journal of the Acoustical Society of America*, vol. 78, pp. 40–48, 1985.
77. M. Rastatter and R. Jacques, "Formant frequency structure of the aging male and female vocal tract," *Folia Phoniatrica (Basel)*, vol. 42, pp. 118–124, 1990.
78. M. Rastatter, R. McGuire, J. Kalinowski, and A. Stuart, "Formant frequency characteristics of elderly speakers in contextual speech," *Folia Phoniatrica et Logopaedica*, vol. 49, pp. 1–8, 1997.
79. S. E. Linville and J. Rens, "Vocal tract resonance analysis of aging voice, using long-term average spectra," *Journal of Voice*, vol. 15, pp. 323–330, 2001.
80. D. Klatt and L. Klatt, "Analysis, synthesis, and perception of voice quality variations among female and male talkers," *Journal of the Acoustical Society of America*, vol. 87, pp. 820–857, 1990.

81. R. Orlikoff, “Heartbeat-related fundamental frequency and amplitude variation in healthy young and elderly male voices,” *Journal of Voice*, vol. 4, pp. 322–328, 1990.
82. R. T. Sataloff, D. C. Rosen, M. Hawksha, and J. R. Spiegel, “The three ages of voice: the aging adult voice,” *Journal of Voice*, vol. 11, pp. 156–160, 1997.
83. J. González, “Formant frequencies and body size of speaker: a weak relationship in adult humans,” *Journal of Phonetics*, vol. 32, pp. 277–287, 2004.
84. A. Braun and T. Rietveld, “The influence of smoking habits on perceived age,” in *Proc. of ICPhS 95*. 1995, vol. 2, pp. 294–297, Stockholm.
85. J. González and A. Carpi, “Early effect of smoking on voice: A multidimensional study,” *Medical Science Monitor*, vol. 10, pp. 649–656, 2004.
86. A. Wagner and A. Braun, “Is voice quality language-dependent? Acoustic analyses based on speakers of three different languages,” in *Proc. of ICPhS 03*. 2003, pp. 651–654, Barcelona.
87. P. Roach, *Phonetics*, Oxford University Press, Oxford, 2001.
88. L-G. Andersson, “Språket, Vetenskapsradion [radio programme],” Article retrieved 24 August 2006, from http://www.sr.se, 2006.
89. H. Traunmüller and A. Eriksson, “Acoustic effects of variation in vocal effort by men, women, and children,” *Journal of the Acoustical Society of America*, vol. 107, pp. 3438–3451, 2000.
90. C. Pemberton, P. McCormack, and A. Russell, “Have women’s voices lowered across time? A cross sectional study of Australian women’s voices,” *Journal of Voice*, vol. 12, pp. 208–213, 1998.
91. W. Decoster and F. Debruyne, “Longitudinal voice changes: facts and interpretation,” *Journal of Voice*, vol. 14, pp. 184–193, 2000.