

Assessment of Vocal Capacity of Finnish University Students

Timo Leino Anne-Maria Laukkanen Irma Ilomäki Elinita Mäki

Department of Speech Communication and Voice Research, University of Tampere, Tampere, Finland

Key Words

Vocal capacity • Voice training • Voice quality • Vocal fatigue • Airflow in phonation • Fundamental frequency • Equivalent sound level • Phonetogram • Long-term average spectrum • Vocally loading tests

Abstract

This article summarizes the results for a total of 189 female and 63 male university students (mean age 23 and 26 years, respectively) from the following vocal capacity tests: measurement of (1) vital capacity, (2) average airflow in phonation (on [o:]), (3) maximum duration of sustained blowing, (4) maximum phonation time (on [o:, z:]), (5) maximum duration of [s:], (6) the ratio of the duration of [s:] to [z:] (s/z ratio), (7) phonation quotient (the ratio between vital capacity and maximum phonation time), (8) the lowest possible tone, (9) mean fundamental frequency (F0) in normal and loud text reading and in shouting, (10) difference between the lowest tone and mean F0 in text reading and in shouting, (11) F0 range in text reading, (12) equivalent sound level of normal and loud reading and of shouting in maximum loudness, (13) phonetogram (voice range profile) and (14) long-term average spectrum of text reading. Voice quality is also evaluated perceptually, and additionally, data on subjects' self-reported vocal fatigue symptoms in general or after a vocally loading task were collected. The value of these measurements from the point of view of assessment of vocal capacity will be discussed.

Copyright © 2008 S. Karger AG, Basel

Introduction

Various parameters have been used in voice training tradition to facilitate the evaluation of vocal capacity and setting the goals of training. The present study aimed to collect average values for an extensive set of parameters and for a large number of subjects. To test the usefulness of these parameters in assessing vocal capacity, the parameter values were compared with subjective reports of vocal endurance and with perceptual evaluation of voice quality.

The traditional parameters include measurement of maximum phonation time and average airflow in phonation, the ratio between vital capacity (lung volume) and maximum phonation time (i.e. phonation quotient, PQ) as well as the ratio between the maximum duration of a voiceless fricative (typically [s:]) and the voiced counterpart (s/z ratio) [1–3]. These parameters have been used to describe phonation type along an axis from hypofunctional (breathy) to hyperfunctional (pressed phonation). Typically, hypofunctional phonation is characterized by a large average airflow and consequently a relatively short maximum phonation time and high PQ and s/z ratio, while the opposite is expected in hyperfunctional phonation [2, 3].

One traditional but much debated concept in voice training is 'optimal (speaking) pitch'. It has been claimed that in each speaker's pitch range there is a pitch (or a pitch range) at which the best sounding acoustic output can be obtained with the least effort [4–10]. This pitch

(or pitch range) should be located at approximately one fourth of the total pitch range from the bottom [4]. The search for the acoustic correlates of such an optimum pitch has failed [11, 12]. On the other hand, the status quo (i.e. how people on average speak) can be taken as a hypothetical 'norm'. Therefore, it is of interest to study in a sufficiently large number of subjects the relation of the average fundamental frequency (F0) in speech to the lowest possible tone each individual is capable of producing. F0 is naturally linked to loudness, the louder the voice, the higher the F0. According to voice training tradition, an excessive rise in F0 in loud phonation is regarded as harmful. A practical rule has been that the pitch should not rise over an octave when loudness is raised from conversational to loud speech. If it does it also tends to change the expression (the speaker may start sounding as if in trouble rather than being in control of the situation). The measurement of sound pressure level (SPL or its time average, the equivalent sound level, *Leq*) in speech quantifies the vocal dynamics. The phonetogram (voice range profile) has been used to describe the vocal capabilities in both dimensions of F0 and intensity in relation to each other [13]. The phonetogram has been reported to distinguish between normal and pathological voices and trained and untrained voices [14–20].

Voice quality has been defined as 'those characteristics which are present more or less all the time that a person is talking: it is a quasi-permanent quality running through all the sound that issues from his mouth' [21]. According to a broad definition as applied e.g. by Laver [22], voice quality is 'the characteristic auditory colouring of an individual speaker's voice'. 'Both laryngeal and supralaryngeal features will be seen as contributing to voice quality' [22].

Voice quality has been studied with the aid of long-term average spectrum (LTAS) analysis. LTAS provides information on the spectral distribution of the speech signal over a period of time [23]. When the sample duration is sufficiently long, e.g. from 30 to 60 s or longer [24–26], it yields a stabilized LTAS, where the individual speech sounds or prosodic aspects no longer affect the spectral shape but the LTAS thus reflects the average voice quality. LTAS has been found capable of distinguishing between normal and pathological voices [27], between different hoarseness levels [28], and between good and less good speaking voices [29–31]. It has also been used to describe the characteristics of different types of singing (opera, country, Chinese opera) and to indicate the effects of song training [32–35].

Perceptual evaluation has been regarded as the voice coaches' main tool. This is also true today, although nowadays the limitations of voice evaluation are also widely acknowledged. It is known that perceptual evaluation of voice quality cannot be used as a diagnostic tool to distinguish between different pathological states of the vocal organ; sometimes it cannot even reliably distinguish between voices produced by a normal and a diseased larynx. However, perceptual evaluation is naturally important from the perspectives of communication and aesthetics. Voice quality affects the listeners' impressions of the speaker [36, 37], and each culture has its vocal ideals which, in turn, vary according to the profession and vocal task. On the other hand, certain main characteristics of a well-functioning voice can be seen as more or less universal. These characteristics include sufficient loudness and absence of excessive effort. 'Maximum output with minimum effort' – i.e. combination of efficiency and economy – can be seen as the main goal in voice training, more or less independent of time and culture.

One of the most important aspects in defining vocal capacity is endurance, that is how well a voice preserves its characteristics and how well a vocalizer feels after a vocally loading task, e.g. a working day of a voice professional. Various procedures have been used to test vocal endurance [38, 39] or in general the effects of vocal loading [40, 41]. Nowadays each client's subjective evaluation of his/her well-being is increasingly focused on, also in medicine, and effort has been invested in testing and validating various questionnaires to be used, for instance, in the evaluation of vocal well-being [42, 43].

In order to set objective goals for vocal evaluation and training, a set of vocal measurements has been carried out over the last 30 years for students attending courses in speech technique at the Department of Speech Communication and Voice Research at the University of Tampere. Some of the measurements have been included from the beginning, some have been added later, and some measurements have been tested for a shorter time and replaced with new ones after a while.

This article summarizes average data obtained in this long-term project for those subjects whose recording and measurement procedure have been comparable. (For instance, the recordings have been made digitally since 1983; analogous recordings have not been included in the present study.)

Subjects and Methods

Subjects

This study reports the results from a sample of 189 females and 63 males (mean age 23.4 years, SD 2.2 for females and mean 25.7 years, SD 3.6 years for males). All subjects were Finnish university students without any known pathologies of voice or hearing. They were tested and recorded while not suffering from cold or laryngitis. These subjects were chosen from a larger sample (approximately 1,000 students), since they represent the most complete set of all measurement data, recorded and analyzed in the same way. Most subjects were vocally untrained, 8.5% of the females ($n = 15$) and 4.8% of the males ($n = 3$) had received some training in singing or were engaged in artistic voice use on an amateur basis (amateur singing, elocution, oral interpretation, theater).

Recordings

All the recordings included in the present study were made in the same well-damped studio using a digital recorder and a Brüel & Kjær microphone (4165) at a distance of 40 cm from the subject's lips. The distance was controlled so that the subjects were standing on a special mark on the floor and they were watched on a TV monitor throughout the recording session.

The subjects (1) read two texts of approximately 1 min duration each (the same text for every subject) at habitual loudness and loudly, (2) shouted as loudly as possible on [a:] and 'apua!' ('help!' in Finnish), and (3) produced a phonetogram by phonating on [a:] at certain given pitches (c, e, g, a) throughout their phonation range both as softly and as loudly as possible. All the recordings were calibrated for further Leq measurements using a sound wave calibrator and level meter (Brüel & Kjær Frequency Analyzer 2120).

Analyses

Samples of text reading and shouting as well as the phonetogram samples were analyzed for mean F0 and Leq using a computerized signal analysis system, the Intelligent Speech Analyser, ISA (developed by Raimo Toivonen, MSc Eng). No A weighting was used in the Leq analyses. In shouting samples where F0 sometimes exceeded the limits of the automatic F0 analysis of the Intelligent Speech Analyser, RealSpeech 4 software (Tiger Electronics) and manual measurement were used.

LTAS were made with a Hewlett-Packard 3561A Signal Analyzer. A 400-point narrow-band FFT analysis was used. The frequency span was 10 kHz. Voiceless segments were excluded from the analysis by using a level trigger and an analog s-gate. The time record length was 40 ms. The display resolution was 25 Hz. Hanning window was used (the frequency band of 37.5 Hz). The spectra were normalized for the strongest peak and compared to each other according to the 'slope' (here: level differences between the strongest peak of the spectrum and the strongest peaks in the ranges of 1–2 kHz, 2–3 kHz, 3–4 kHz, 4–5 kHz and 5–10 kHz).

Tests

Vital capacity (maximum air volume during forced expiration, ml) and average airflow (ml/s) during the first 6 s of phonation on [o:] were measured with VitaloGraph (2022). Maximum duration of sustained blowing (s) after deep inhalation was measured during 'bubbling' through a straw into a cup filled with water. Maxi-

imum duration of phonation on [o:] and [z:] at habitual pitch and loudness was measured (in seconds). Additionally, maximum duration of the voiceless fricative [s:] (produced in a stable and non-forced manner) was measured (in seconds). The ratio of the duration of [s:] to [z:] (s/z ratio) was calculated. PQ was measured by dividing the vital capacity (ml) by the maximum duration of [o:] (in seconds). The lowest possible tone (Hz) the subjects could produce in a soft voice (vocal fry excluded) was measured with a real-time FFT spectrum display (HP Signal Analyzer).

Perceptual Evaluation

All samples of text reading at habitual loudness and loudly and samples of shouting ($n = 504$ for text reading and $n = 504$ for shouting, 1,008 samples in total) were evaluated by 2 experienced voice trainers. A digital recorder and Genelec Biamp (1019 A) loudspeaker were used. The samples were evaluated in a sound-treated studio in different sessions according to the listener's own judgment of their alertness and capability for performing a careful perceptual analysis. General voice quality was evaluated using a scale from -3 = very poor quality to $+3$ = excellent quality. Tightness of phonation was given points -1 = 'breathy' (hypofunctional), 0 = good, $+1$ = 'pressed' (hyperfunctional). Half of the reading samples (at habitual level and loud) were assessed twice in order to calculate intrarater reliability.

Vocally Loading Tests

Two types of vocally loading tests were carried out in the Department of Speech Communication and Voice Research: loading was obtained either by having the subjects read aloud from a book for 30 min, at 70 dB measured at a distance of 1 m (in a well-damped studio) or by having them shout numbers monotonically for 5 min at 90 dB, measured at a distance of 1 m. Sound level in the loading test was measured using either a Brüel & Kjær portable Sound Level Meter or a Brüel & Kjær Frequency Analyzer 2120. The type of loading test included in the speech technique course was varied over the years. Thus, which type of loading test the students did depended on the year when they joined the class. Of the female subjects of the present study, 13 did the 5-min loading test and 50 did the 30-min test. Of the males 4 did the short loading test and 8 the long test.

Self-Evaluation

Two types of material have been included in the data for self-evaluation of vocal fatigue: (1) one group of subjects (30 females, 12 males) completed a questionnaire on symptoms of vocal fatigue experienced in general, and (2) another group of subjects (63 females, 12 males) reported on symptoms after a vocally loading test (consisting either of loud text reading for 30 min or of shouting numbers for 5 min). A three-point scale was used to report the severity of vocal fatigue symptoms (hoarseness, tiredness of throat) experienced after vocal loading in general or after the special vocally loading test (1 = mild, 2 = medium, 3 = severe). A four-point scale was used for the frequency of the vocal fatigue symptoms experienced in general (0 = never, 1 = seldom, 2 = sometimes, 4 = often, i.e. once a week or more often). Based on these answers, the subjects were classified into four groups: '0' = not suffering from vocal fatigue ($n = 3$ in females, $n = 5$ in males), '1' = mild vocal fatigue (i.e. mild symptoms seldom or sometimes or medium symptoms but seldom, $n = 33$ in females and $n = 7$ in males), '2' = medium vocal fatigue (i.e. mild symptoms often or

Table 1. Vital capacity (VC), body height, maximum duration of sustained blowing (bubbling into a glass of water, MDB), maximum duration of [o:], [z:] and [s:], s/z ratio, mean airflow in phonation during the first 6 s of [o:], and PQ

	VC ml	Height cm	Maximum duration, s				s/z	Airflow ml in 6 s	PQ ml/s
			MDB	[o:]	[z:]	[s:]			
Females									
Mean ± SD	3,755 ± 511	167.7 ± 5.6	48.5 ± 11.8	22.6 ± 6.0	22.7 ± 8.3	25.3 ± 9.7	1.2 ± 0.4	611 ± 253	177 ± 53
Range	2,600 – 5,430	152 – 180	24 – 104	10 – 51	8 – 50	6 – 62	0.5 – 3.0	100 – 1,500	67 – 400
Males									
Mean ± SD	5,335 ± 726	180.2 ± 5.6	64.9 ± 20.7	30.0 ± 10.1	29.0 ± 12.2	33.1 ± 14.5	1.1 ± 0.6	699 ± 249	197 ± 72
Range	3,750 – 7,200	171 – 192	37 – 112	12 – 55	14 – 60	15 – 77	0.4 – 3.1	300 – 1,400	88 – 389
Difference between males and females (Student's unpaired t test)									
	P < 0.001	P < 0.001	P < 0.001	P < 0.001	P < 0.001	P < 0.001	NS	P = 0.035	P = 0.039

NS = Nonsignificant ($P > 0.05$); SD = standard deviation; $n = 189$ female and $n = 63$ male university students.

severe symptoms seldom or medium sometimes, $n = 46$ in females and $n = 10$ in males) and '3' = severe vocal fatigue (i.e. medium or severe symptoms experienced often, $n = 11$ in females, $n = 2$ in males). In total, vocal fatigue classification was made for 93 females and 24 males.

Statistical Analyses

Spearman's correlation (SPSS-11) was used to calculate the intra- and interrater reliability of the perceptual evaluation and to study the relations between perceptual evaluation or self-evaluation and the acoustic variables. Pearson's correlation was used to study relations between acoustic variables for parameters with normal distribution; in other cases Spearman's correlation was used. Differences in the acoustic variables between males and females or between breathy and pressed voice types were studied with Student's t test (unpaired), when the variables showed a normal distribution and with Mann-Whitney U test, when the distribution was not normal.

Results and Discussion

Tables 1–7 summarize the results. The results for vital capacity, airflow during phonation, maximum duration of [o:], [s:], s/z ratio and PQ (table 1) were within the limits of those reported in the literature [2, 44, 45]. As could be expected, subject's height correlated strongly with vital capacity ($r = 0.73$, $P < 0.001$). Due to structural differences (larger lungs and a wider trachea) males naturally have a greater vital capacity and airflow than females. Maximum duration of fricative and vowel production was also longer for the males. The s/z ratio, however, did not differ significantly between the genders. PQ was somewhat higher for the males; they thus did not seem to use their vital capacity as fully as the females for speech sound production.

The male subjects' mean F0 in text reading at habitual loudness corresponded to the values reported, e.g. by Baken and Orlikoff [3], but the mean F0 of female subjects was about 2 semitones (st) lower, probably mainly due to cultural differences (table 2). The mean F0 in habitual reading was about 5 st above the lowest possible tone. This is in line with the results of Leino [46]. Åkerlund et al. [18] reported values of 18–20% or 7 st above the lowest F0 in the phonetogram, both for vocally trained and untrained females. The average F0 in untrained singers' text reading was about the same as in the present material (191.4 Hz) [i.e. fis in ref. 18]. This suggests that the subjects of Åkerlund et al. [18] reached relatively lower tones in the phonetogram; possibly vocal fry type of phonation was not excluded in the experiment. It is also noteworthy that the subjects were older than those of the present study [mean age 41 years in ref. 18]. Regarding the traditional concept of 'optimal pitch', according to the results of the present study, the mean F0 in habitual reading was about one fifth from the bottom of the total F0 range (in the phonetogram): 4.9 for the females and 5.1 for the males. The recommendation of Fairbanks [4], for example, was one fourth from the bottom of the total range. This discrepancy may be interpreted by suggesting that the students did not reach their real F0 maximum in the phonetographic registration. The 90% F0 range (in semitones) in reading was in line with the results summarized by Baken and Orlikoff [3]. Contrary to expectations, the males used a wider F0 variation in text reading compared to the females. For instance, the results for American conversational speech [47], the general description of voice therapy in transgender deformation [48] and the results obtained by Airas and Alku [49] con-

Table 2. F0 of the lowest possible tone, mean F0 and Leq in habitual and loud text reading and in shouting, range of F0 variation in text reading, relation between lowest possible tone and mean F0 in text reading and shouting (relative F0) and relation between mean Leq in reading and maximum Leq in shouting (in %)

	Lowest F0 Hz	Mean F0 Hz	Leq dB	90% F0 range st	Relative F0 st	Leq vs. max %
<i>Females</i>						
Mean \pm SD	138.5 \pm 15.1					
Range	93 – 185					
Reading						
Mean \pm SD		194.0 \pm 17.2	67.1 \pm 3.9	9.3 \pm 2.2	5.9 \pm 1.7	61.9 \pm 4.7
Range		151 – 246	56 – 77	5 – 15	1 – 13	50 – 76
Loud reading						
Mean \pm SD		303.7 \pm 55.5	84.2 \pm 5.4		13.6 \pm 3.6	77.6 \pm 5.4
Range		196 – 493	73 – 99		5 – 23	64.1 – 93.4
Shouting						
Mean \pm SD		504.0 \pm 87.0	108.8 \pm 7.0		22.0 \pm 4.0	
Range		253 – 743	86 – 125		10 – 30	
<i>Males</i>						
Mean \pm SD	80.3 \pm 12.4					
Range	55 – 118					
Reading						
Mean \pm SD		109.8 \pm 14.7	68.9 \pm 3.8	10.1 \pm 3.3	5.3 \pm 2.2	62.5 \pm 4.7
Range		85 – 153	60 – 77	4 – 20	1 – 11	53.8 – 73.5
Loud reading						
Mean \pm SD		230.2 \pm 60.4	90.0 \pm 8.1		17.7 \pm 5.2	81.1 \pm 7.5
Range		131 – 370	69 – 102		7 – 29	62.7 – 99.0
Shouting						
Mean \pm SD		338.5 \pm 55.9	110.4 \pm 6.4		24.5 \pm 4.3	
Range		175 – 440	92 – 122		14 – 33	
Difference between males and females (Student's unpaired t test)						
Reading			P = 0.002	P = 0.037	NS	NS
Loud reading			P < 0.001		P < 0.001	P < 0.001
Shouting			NS		P < 0.001	
NS = Nonsignificant (P > 0.05); SD = standard deviation; n = 189 female and n = 63 male university students.						

cerning differences in voice source parameters in Finnish speakers would suggest that females tend to use a wider range in prosodic variables. Both for the males and females of the present study, habitual text reading loudness was about 62% of maximum Leq in shouting.

Maximum Leq in shouting (table 2) was somewhat higher than the sound level values reported, for example, by Ptacek et al. [50] (males of the corresponding age group 105.8 dB, females 106.2 dB; when the microphone distance of 30.5 cm used by Ptacek et al. [50] is compensated for the 40 cm that was used in the present study, we end up with values of 103.4 and 103.8 dB, respectively).

The difference in mean F0 between habitual and loud reading was on average 7.7 st for the females and 12.4 st

for the males. The corresponding mean Leq differences between habitual and loud reading were 17.1 dB for the females and 21.1 dB for the males. Thus, the subjects of the present study seemed to fit the pedagogical 'rule of thumb' concerning a rise of maximally one octave in F0 when raising loudness from habitual to loud.

When reading at habitual loudness and shouting were compared, the increase in F0 in relation to increase in Leq was on average 0.39 st/1 dB for the females and 0.46 st/1 dB for the males. This is somewhat lower than the values reported for increase in SPL when F0 is raised (7.5–12 dB/octave, i.e. 1.0–1.6 st/1 dB) [51]. On the other hand, the results for females are in line with those reported by Åkerlund et al. [18]. The males in the present study used

Table 3. LTAS characteristics of text reading (habitual loudness) and their relation to perceptual evaluation of phonation type ('tightness of phonation'; Spearman correlation) in 189 female and 63 male university students: level differences (dB) between the strongest spectral component (F1 region) and the regions of F0, 1–2 kHz, 2–3 kHz, 3–4 kHz, 4–5 kHz and 5–10 kHz

	F1–F0 dB	1–2 kHz dB	2–3 kHz dB	3–4 kHz dB	4–5 kHz dB	5–10 kHz dB
Females						
Mean ± SD	-1.2 ± 2.8	-12.8 ± 4.4	-25.1 ± 4.5	-30.6 ± 4.5	-34.5 ± 4.0	-40.7 ± 3.3
Range	-11 – +7	-26 – +7	-35 – -14	-40 – -20	-44 – -26	-52 – -32
Phonation type	r = 0.17 P = 0.019	r = 0.16 P = 0.035	r = 0.28 P = 0.000	NS	r = 0.19 P = 0.013	NS
Males						
Mean ± SD	1.4 ± 1.7	-13.7 ± 4.7	-23.9 ± 4.4	-29.0 ± 5.9	-38.9 ± 5.6	-41.3 ± 4.2
Range	0 – +8	-25 – -5	-33 – -10	-42 – -17	-54 – -25	-50 – -32
Phonation type	r = 0.44 P < 0.001	r = 0.43 P < 0.001	r = 0.51 P < 0.001	r = 0.58 P < 0.001	r = 0.50 P < 0.001	NS
Difference between males and females (Student's unpaired t test)	NS	NS	P = 0.053	P = 0.012	P < 0.001	NS
NS = Nonsignificant (P > 0.05); SD = standard deviation.						

Table 4. Differences between 'breathy' and 'pressed' voices in spectral parameters, maximum phonation time, PQ, air consumption in phonation and s/z ratio

	F1–F0 dB	1–2 kHz dB	2–3 kHz dB	3–4 kHz dB	4–5 kHz dB	5–10 kHz dB	[o:] sec	Airflow ml/s	PQ	s/z
Females										
Breathy (n = 128)										
Mean ± SD	-1.5 ± 3.0	-13.2 ± 4.5	-26.0 ± 4.4	-31.0 ± 4.3	-34.9 ± 3.9	-40.7 ± 3.2	22.1 ± 5.7	609.0 ± 247.7	179.8 ± 57.1	1.18 ± 0.44
Pressed (n = 47)										
Mean ± SD	-0.3 ± 2.3	-11.6 ± 3.9	-23.1 ± 4.3	-29.9 ± 4.9	-33.2 ± 3.8	-40.5 ± 3.4	23.9 ± 6.9	625.1 ± 283.9	168.9 ± 45.9	1.16 ± 0.32
Significance	P = 0.0018	P = 0.035	P < 0.001	NS	P = 0.01	NS	NS	NS	NS	NS
Males										
Breathy (n = 40)										
Mean ± SD	-0.81 ± 1.13	-14.9 ± 4.3	-25.4 ± 3.7	-31.3 ± 5.2	-41.0 ± 4.6	-41.6 ± 4.3	29.8 ± 10.1	716.7 ± 262.4	198.5 ± 78.4	1.37 ± 0.65
Pressed (n = 19)										
Mean ± SD	2.3 ± 1.7	-10.8 ± 4.2	-20.4 ± 4.1	-23.6 ± 4.4	-34.9 ± 5.7	-40.7 ± 3.4	31.7 ± 10.6	648.7 ± 228.7	191.6 ± 62.0	0.98 ± 0.27
Significance	P = 0.001	P < 0.001	P < 0.001	P < 0.001	P < 0.001	NS	NS	NS	NS	NS
Student's unpaired t test. SD = Standard deviation; NS = nonsignificant (P > 0.05).										

higher Leq in habitual and loud reading than the females; in loud reading the relative Leq (in % of maximum in shouting) was 3.5% higher in males. Since in shouting the Leq as such did not differ significantly between the genders and since the males raised F0 more with Leq, it seems that the males used F0 increase more than females as a

means of raising Leq. In general, a pure F0 rise may account for an approximately 4 dB rise in sound level [52].

Table 3 summarizes the results for LTAS analysis. Figure 1 illustrates the average differences between the male and female voices. The greatest LTAS differences between the genders were seen between 4 and 5 kHz. No special

spectral characteristics of good voice quality [29–31] were seen in the LTAS, and none of the voices were evaluated as really good by the listening panel. On the other hand, good voices cannot always be distinguished by LTAS, as also suggested by Bele [31], studying characteristics of good voices of Norwegian male actors and teachers. Interrater reliability of the perceptual evaluation was satisfactory ($r = 0.56$, $P < 0.001$). Intrarater reliability was high ($r = 0.75$, $P < 0.001$). The average slope of the spectrum correlated with the evaluation of ‘tightness’ of the voices (table 3), as can be expected [e.g. ref. 49, 50]; i.e. the voices were evaluated to be tighter, more pressed as the spectral slope was more gentle (see also fig. 2). Table 4 compares various parameters between speakers whose voice samples were evaluated as representing breathy or pressed voice type. It can be seen that the slope of the LTAS in habitual reading was the only characteristic that differed significantly between these voice types. This seems to suggest that phonation type in various tests (like maximally extended vowel phonation or phonation of a vowel in Vitalograph) may not be comparable to the phonation type used in connected speech. However, in addition to spectral slope, mean F0 in text reading also correlated negatively with tightness in females (table 5), suggesting that higher pitch was related to more breathy voice type (or perception of it, and thus yielding a negative value in the evaluation of tightness). In females, s/z ratio also correlated negatively with tightness, being higher for subjects with breathy voices (in text reading, table 5). In males, there was a positive correlation between tightness and Leq and between voice quality and Leq in habitual reading (table 5). For males, higher Leq in shouting was also related to self-reported symptoms of vocal fatigue. This seems to be in line with earlier results for female teachers [53] and for students in a vocally loading test [54]. On the other hand, for both genders there was a *negative* correlation between vocal fatigue and the relative Leq in habitual reading. For females, there was also a negative correlation between the relative F0 in habitual reading and vocal fatigue (i.e. those who used a lower pitch also reported more symptoms of vocal fatigue). In females, voice quality correlated negatively with relative F0 or relative Leq in shouting, indicating that shouting at high pitch and loudness relative to one’s capacity was evaluated as sounding worse. General voice quality did not correlate with self-reported vocal fatigue. On the other hand, in males, there was an indicative *negative* correlation between perceived tightness and self-reported vocal fatigue, suggesting that those with a more breathy type of voice production reported more vocal fatigue.

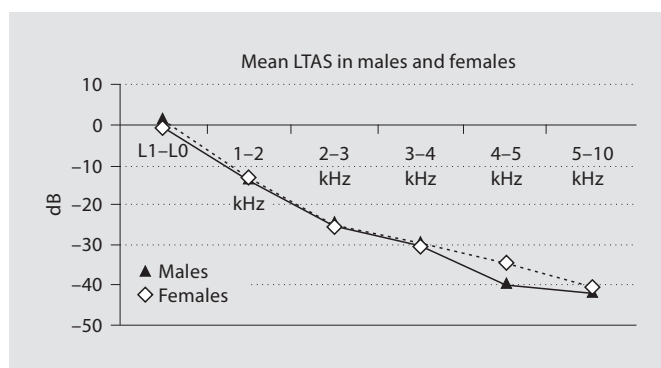


Fig. 1. Schematic mean LTAS for 189 females and 63 males.

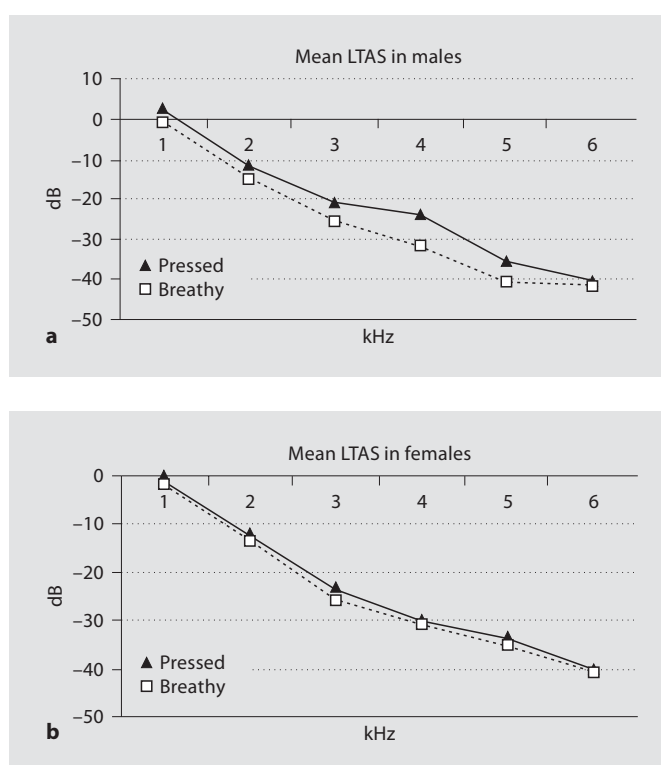


Fig. 2. Schematic mean LTAS differences between text reading samples evaluated as representing ‘breathy’ and ‘pressed’ types of phonation for males (a, n = 40 breathy, n = 19 pressed) and females (b, n = 128 breathy, n = 47 pressed).

These results apparently differ from clinical tradition, according to which hyperfunctional (‘tight’) high-pitched and loud voice production has been regarded as a risk factor for voice problems. Experimental results for excised larynges also have shown that impact stress rises both

Table 5. Significant correlations (Spearman, $P < 0.05$) between acoustic parameters of vocal capacity, perceptual voice quality, 'tightness' of phonation and self-evaluated vocal fatigue for females ($n = 93$) and males ($n = 24$)

	F0, Hz	Leq, dB	Leq vs. max, %	s/z	Vocal fatigue
<i>Females</i>					
Voice quality					
Reading				$r = -0.18$ $P = 0.023$	
Loud reading					
Shouting					
Tightness in reading	$r = -0.15$ $P = 0.043$			$r = -0.21$ $P = 0.008$	
Vocal fatigue					
Reading			$r = -0.20$ $P = 0.059$ ($n = 93$)		
Loud reading					
Shouting					
<i>Males</i>					
Voice quality					
Reading		$r = 0.30$ $P = 0.014$			
Loud reading					
Shouting					
Tightness in reading	$r = 0.33$ $P = 0.006$		$r = 0.32$ $P = 0.01$		$r = -0.39$ $P = 0.057$
Vocal fatigue					
Reading			$r = -0.54$ $P = 0.007$ ($n = 24$)		
Loud reading					
Shouting		$r = 0.66$ $P < 0.001$ ($n = 24$)			

with F0, SPL and adduction [55]. However, the results of the present study may be interpreted to show that subjects who are used to soft hypofunctional type of voice production more easily contract symptoms of voice fatigue in a loading test requiring loud voice production.

Table 6 shows the average values measured for phonetograms. Both the decibel range and F0 range (st) measured in these vocally untrained university students were smaller compared to e.g. the values reported by Coleman et al. [56] and Sulter et al. [20]. This may reflect the problem of getting mainstream vocally untrained university students motivated in phonetographic measurements. Many students completely refused, claiming that they could not sing, repeat pitches after a model, etc. The values obtained by Gramming [17] for the mean dynamic range (the average decibel difference from softest to loudest at the same pitch) were comparable with those obtained in the present study. Both the lowest and the highest Leq values in males were significantly higher compared to those of the females, and the mean dynamic range was also greater in males. This may reflect a certain cultural difference due to which boys may be more used to loud phonation. On the other hand, as was seen, Leq in shouting did not differ significantly between the genders in this material (table 2).

Table 7 shows the relations of phonetographic and other parameters of vocal capacity. The lowest possible tone correlated with the lowest tone in the phonetogram both for the males and the females. The F0 range in text reading correlated with the F0 range in the phonetogram. Additionally, for the males, relative F0 (in relation to the lowest tone) correlated with F0 range in the phonetogram, indicating that those males who had a wider F0 range also used a relatively higher mean F0 in reading. The maximum dB in the phonetogram correlated with

Table 6. Characteristics of phonetograms for 189 female and 63 male university students

	F0 lowest Hz	F0 highest Hz	F0 range st	Leq lowest dB	Leq highest dB	Leq range, min-max, dB	Leq difference, dB
<i>Females</i>							
Mean \pm SD	147.1 \pm 19.8	757.3 \pm 227.3	28.7 \pm 6.7	53.0 \pm 5.1	92.6 \pm 8.9	39.6 \pm 10.7	19.6 \pm 6.7
Range	98 – 208	62 – 1318	11 – 40	40 – 68	72 – 115	15 – 68	4 – 38
<i>Males</i>							
Mean \pm SD	94.12 \pm 20.4	488.2 \pm 167.4	28.6 \pm 7.5	57.1 \pm 4.7	98.6 \pm 9.4	41.6 \pm 10.4	23.7 \pm 7.5
Range	65 – 196	185 – 880	7 – 43	45 – 70	74 – 119	20 – 62.5	8 – 38
Difference between males and females			NS	$P < 0.001$	$P < 0.001$	NS	$P < 0.001$

NS = Nonsignificant ($P > 0.05$); SD = standard deviation; Student's unpaired t test.

Leq in loud reading and shouting in both genders. In females, too, Leq in habitual reading correlated with maximum Leq and mean dynamic range in the phonetogram. However, in females there was a *negative* correlation between relative Leq in habitual reading and maximum Leq in the phonetogram, i.e. those females who had a higher dynamic capacity used relatively less of it in habitual reading. They also had higher Leq values as such, so they probably did not need to use as much effort as those with a less efficient vocal organ or vocal habits. On the other hand, it is likely that subjects who are used to softer voice production do not reach their absolute physiological limits and therefore the percentage of maximum values is not quite reliable.

In females vocal fatigue did not correlate with phonetographic parameters. In males, instead, vocal fatigue correlated negatively with F0 range in the phonetogram, i.e. less fatigue was reported by those who had a wider F0 range.

Similar results about the positive correlation of speech and phonetogram parameters have been reported [e.g. ref. 18]. The usefulness of the phonetogram in differentiating between normal and disturbed voices, trained and untrained voices and in quantifying the effects of vocal loading or voice training has been well documented [e.g. ref. 14]. However, the results of the present study suggest that for a simple and most practical screening test of vocal capacity of untrained students, it might be sufficient to use only registration and analyses of speech samples at habitual loudness and loudly.

Here the self-evaluation results have been used to assess vocal endurance. Further studies should concentrate for instance on the consistency of self-evaluation and on the validation of questionnaires to be used in quantifying the amount of subjective vocal fatigue.

Conclusions

This study presents average values for a large set of acoustic and phonatory parameters in vocally untrained Finnish university students ($n = 189$ females, $n = 63$ males). This data can hopefully be used in the assessment of vocal capacity and setting goals for vocal training.

(1) LTAS differentiated between breathy and pressed voice types. The greatest differences in LTAS between males and females were seen at 4–5 kHz.

(2) In females, s/z ratio and mean F0 correlated negatively with perceived tightness of the voice. None of the

Table 7. Correlations between phonetogram and other parameters of vocal capacity and self-reported vocal fatigue in females and males

	Phonetogram			
	lowest Hz	range st	Leq max dB	mean Leq diff., dB
<i>Females</i>				
Habitual reading				
F0 range, st		$r = 0.22$ $P = 0.002$		
Leq, dB			$r = 0.18$ $P = 0.014$	$r = 0.15$ $P = 0.049$
Leq % max			$r = -0.19$ $P = 0.01$	$r = -0.19$ $P = 0.01$
Loud reading				
Leq, dB			$r = 0.51$ $P < 0.001$	$r = 0.37$ $P < 0.001$
Leq % max			$r = 0.12$ $P = 0.085$	NS
Shouting				
Leq, dB			$r = 0.38$ $P < 0.001$	$r = 0.35$ $P < 0.001$
Lowest tone		$r = 0.62$ $P < 0.001$		
Vocal fatigue			NS	NS
<i>Males</i>				
Habitual reading				
Relative F0, st		$r = 0.41$ $P = 0.008$		
F0 range, st		$r = 0.33$ $P = 0.01$		
Leq, dB				
Leq % max				
Loud reading				
Leq, dB			$r = 0.60$ $P = 0.000$	$r = 0.45$ $P = 0.001$
Leq % max				
Shouting				
Leq, dB			$r = 0.39$ $P = 0.002$	
Lowest tone		$r = 0.52$ $P < 0.001$		
Vocal fatigue			$r = 0.45$ $P = 0.028$	

tests, however, differentiated between breathy and pressed voice types in either gender, suggesting somewhat different voice use in text reading and in phonation tests.

(3) Higher Leq in shouting correlated with symptoms of vocal fatigue in males.

(4) The use of *lower Leq in relation to capacity* in habitual reading correlated with self-reported vocal fatigue. This may either reflect the subjects' cautiousness or lack of experience of loud phonation, which in turn may contribute to fatigue.

(5) Males with a more hypofunctional type of voice production reported more symptoms of vocal fatigue.

(6) Males seemed to use a relatively larger increase in F0 per increase in Leq, suggesting that compared to fe-

males they used F0 more as a means of increasing loudness.

(7) Phonetographic characteristics correlated with parameters measured from speech. Therefore, registration and analyses of speech samples alone may be sufficient for a routine voice screening test for vocally untrained students.

(8) The smallest informative set of vocal capacity tests could be: lowest possible tone, mean F0 and Leq in habitual reading and shouting, s/z ratio, perceptual evaluation of tightness and self-reported symptoms of vocal fatigue.

(9) The self-evaluation procedure in the assessment of vocal endurance warrants further studies.

References

- Michel JF, Wendahl R: Correlates of voice production; Travis LE (ed): Handbook of Speech Pathology. Englewood Cliffs, Prentice-Hall, 1971, pp 465–479.
- Hirano M: Clinical examination of voice; in Arnold GE, Winckel F, Wyke BD (eds): Disorders of Human Communication 5. Vienna, Springer, 1981.
- Baken RJ, Orlikoff R: Clinical Measurement of Speech and Voice, ed 2. San Diego, Singular, 2000.
- Fairbanks G: Voice and Articulation Drillbook. New York, Harper & Row, 1960.
- Anderson V: Training the Speaking Voice, ed 2. New York, Oxford University Press, 1961.
- Cooper M: Modern Techniques of Vocal Rehabilitation. Springfield, Thomas, 1973.
- Eisensohn J: Voice and Diction: A Program for Improvement, ed 3. New York, Macmillan, 1974.
- Moncur JB, Brackett IP: Modifying Vocal Behavior. New York, Harper & Row, 1974.
- Fisher HB: Improving Voice and Articulation, ed 2. Boston, Houghton Mifflin, 1975.
- Aronson AE: Clinical Voice Disorders: An Interdisciplinary Approach, ed 2. New York, Thieme, 1985.
- Minifie FD: Against the clinical use of optimal pitch; in Van Lawrence L (ed): Transcripts of the Twelfth Symposium: Care of the Professional Voice, Part II: Vocal Therapeutics, Respiration, Medical Voice Formation. New York, Voice Foundation, 1984, pp 148–154.
- Stone RE: Issues in clinical assessment of laryngeal function: contraindications for subscribing to maximum phonation time and optimum fundamental frequency; in Bless DM, Abbs JH (eds): Vocal Fold Physiology: Contemporary Research and Clinical Issues. San Diego, College-Hill Press, 1983, pp 410–424.
- Damste PH: The phonetogram. Pract Otorhinolaryngol (Basel) 1970;32:185–187.
- Schutte HK: Untersuchungen von Stimmqualitäten durch Phonetographie. HNO-Praxis 1980;5:132–139.
- Komiyama S, Watanabe H, Ryu S: Phonographic relationship between pitch and intensity of the human voice. Folia Phoniatr 1984;36:1–7.
- Gramming P: The Phonetogram: An Experimental and Clinical Study; doct diss University of Lund, Malmö, 1988.
- Gramming P: Vocal loudness and frequency capabilities of the voice. J Voice 1991;5:144–157.
- Åkerlund L, Gramming P, Sundberg J: Phonetogram and averages of sound pressure levels and fundamental frequencies of speech: comparison between female singers and nonsingers. J Voice 1992;6:55–63.
- Awan SN: Superimposition of speaking voice characteristics and phonetograms in untrained and trained vocal groups. J Voice 1993;7:30–37.
- Sulter AM, Schutte HK, Miller DG: Differences in phonetogram features between male and female subjects with and without vocal training. J Voice 1995;9:363–377.
- Abercrombie D: Elements of General Phonetics. Edinburgh, Edinburgh University Press, 1967.
- Laver J: The Phonetic Description of Voice Quality. Cambridge, Cambridge University Press, 1980.
- Löfqvist A, Mandersson B: Long-time average spectrum of speech and voice analysis. Folia Phoniatr 1987;39:221–229.
- Prytz S: Long-time-average-spectra (LTAS) analyses of normal and pathological voices; in Buch NH (ed): Proc Int Assoc Logop Phon Congr, Copenhagen, August 1977. Herning, Organizing Committee of the Congress, 1978, pp 459–475.
- Majewski W, Rothman H, Hollien H: Acoustic comparisons of American English and Polish. J Phonet 1977;5:247–251.
- Nolan F: The Phonetic Bases of Speaker Recognition. Cambridge, Cambridge University Press, 1983.
- Dejonckere PH, Villarosa D: Analyse spectrale moyennée de la voix: comparaison de voix normales et de voix altérées par différentes catégories de pathologies laryngées. Acta Otorhinolaryngol Belg 1986;40:426–435.
- Wendler J, Doherty ET, Hollien H: Voice classification by means of long-term speech spectra. Folia Phoniatr 1980;32:51–60.
- Leino T: Long-term average spectrum study on speaking voice quality in male actors; in Friberg A, Iwarsson J, Jansson E, Sundberg J (eds): SMAC93, Proc Stockholm Music Acoustics Conf, July–August, 1993. Stockholm, Royal Swedish Academy of Music, 1994, No 79, pp 206–210.
- Nawka T, Anders LC, Cebulla M, Zurakowski D: The speaker's formant in male voices. J Voice 1997;11:422–428.
- Bele IV: Professional Speaking Voice: A Perceptual and Acoustic Study of Male Actors' and Teachers' Voices; doct diss University of Oslo, 2002.

- 32 Flach M, Schwickardi H, Dickopf G, Pabst F: Zur Beurteilung sängerischer Stimmentwicklung mittels LTAS und Stimmfeldmessung. *Folia Phoniatri* 1989;41:4–5.
- 33 Rossing T, Sundberg J, Ternström S: Acoustic comparison of voice use in solo and choir singing. *J Acoust Soc Am* 1986;79:1975–1981.
- 34 Wang S: Singing voice: bright timbre, singer's formant, and larynx positions; in Friberg A, Iwarsson J, Jansson E, Sundberg J (eds): *SMAC93, Proc Stockholm Music Acoustics Conf, July–August 1993*. Stockholm, Royal Swedish Academy of Music, 1994, No 79, pp 313–322.
- 35 Cleveland T, Sundberg J, Stone RE: Long-term-average spectrum characteristics of country singers during speaking and singing. *J Voice* 2001;15:54–60.
- 36 Addington DW: The relationship of selected vocal characteristics to personality perception. *Speech Monogr* 1968;35:492–503.
- 37 Blood GW, Mahan W, Hyman M: Judging personality and appearance from voice disorders. *J Commun Disord* 1979;12:63–68.
- 38 Buekers R: Voice Performances in Relation to Demands and Capacity: Development of a Quantitative Phonometric Study of the Speaking Voice; PhD thesis University of Maastricht, 1998.
- 39 Wendler J, Seidner W: *Lehrbuch der Phoniatrie*. Leipzig, Thieme, 1987.
- 40 Lira EA, Bonilla R, Castro L, Figari A, Maraboli E, Munon O, Pinto E, Salinas S, Schwalm E, Vera S: Un estudio experimental sobre los efectos del abuso vocal en seis voluntarios. *Proc XV Congreso Internacional de logopedia y foniatria*. Buenos Aires, August 1971.
- 41 Vilkmán E, Lauri ER, Alku P, Sala E, Sihvo M: Effects of prolonged oral reading on F₀, SPL, subglottal pressure and amplitude characteristics of glottal flow waveforms. *J Voice* 1999;13:303–315.
- 42 Ma E, Yiu E: Voice activity and participation profile: assessing the impact of voice disorders on daily activities. *J Speech Lang Hear Res* 2001;44:511–524.
- 43 Simberg S: Prevalence of Vocal Symptoms and Voice Disorders among Teacher Students and Teachers and a Model of Early Intervention, doct diss University of Helsinki, Faculty of Behavioural Sciences, Department of Speech Sciences, 2004.
- 44 Isshiki N: Regulatory mechanism of voice intensity variation. *J Speech Hear Res* 1964;7:17–29.
- 45 Koike Y, Hirano M: Significance of vocal velocity index. *Folia Phoniatri* 1968;20:285–296.
- 46 Leino T: In search of 'optimal pitch': the lowest possible tone as the reference in the evaluation of speaking pitch in Finnish speakers. *Proc 24th World Congr Int Assoc Logop Phoniatri*, August 1998.
- 47 Syrdal AK: Acoustic variability in spontaneous conversational speech of American English talkers. <http://www.asel.udel.edu/icslp/cdrom/vol1/582/a582.pdf>.
- 48 Voice therapy (trans) Gender transitioning. From Wikipedia, the Free Encyclopedia [http://en.wikipedia.org/wiki/Voice_therapy_\(trans\)](http://en.wikipedia.org/wiki/Voice_therapy_(trans)).
- 49 Airas M, Alku P: Emotions in vowel segments of continuous speech: analysis of the glottal flow using the normalised amplitude quotient. *Phonetica* 2005;62:1–21.
- 50 Ptacek P, Sander E, Malone W, Jackson C: Phonatory and related changes with advanced age. *J Speech Hear Res* 1966;9:353–360.
- 51 Stone R, Krause-Ferch P: Intra-subject variability in F₀-SPL voice profiles. *J Speech Hear Disord* 1982;47:134–137.
- 52 Alku P, Vintturi J, Vilkmán E: Measuring the effect of fundamental frequency raising as a strategy for increasing vocal intensity in soft, normal and loud phonation. *Speech Commun* 2002;38:321–334.
- 53 Rantala L, Vilkmán E: Relationship between subjective voice complaints and acoustic parameters in female teachers' voices. *J Voice* 1999;13:484–495.
- 54 Mäki E, Niemi HM, Lundén S, Laukkanen AM: F₀, SPL and vocal fatigue in a vocally loading test. *Proc 25th World Congr Int Assoc Logop Phoniatri*, Montreal, August 2001.
- 55 Jiang JJ, Titze IR: Measurement of vocal fold intraglottal pressure and impact stress. *J Voice* 1994;2:132–144.
- 56 Coleman RF, Mabis JH, Hinson JK: Fundamental frequency: sound pressure level profiles of adult male and female voices. *J Speech Hear Res* 1977;20:197–204.