

An Exploration of Skin Acceleration Level as a Measure of Phonatory Function in Singing

Anick Lamarche and Sten Ternström

Stockholm, Sweden

Summary: Two kinds of fluctuations are observed in phonetogram recordings of singing. Sound pressure level (SPL) can vary due to vibrato and also due to the effect of open and closed vowels. Since vowel variation is mostly a consequence of vocal tract modification and is not directly related to phonatory function, it could be helpful to suppress such variation when studying phonation. Skin acceleration level (SAL), measured at the jugular notch and on the sternum, might be less influenced by effects of the vocal tract. It is explored in this study as an alternative measure to SPL. Five female singers sang vowel series on selected pitches and in different tasks. Recorded data were used to investigate two null hypotheses: (1) SPL and SAL are equally influenced by vowel variation and (2) SPL and SAL are equally correlated to subglottal pressure (P_S). Interestingly, the vowel variation effect was small in both SPL and SAL. Furthermore, in comparison to SPL, SAL correlated weakly to P_S . SAL exhibited practically no dependence on fundamental frequency, rather, its major determinant was the musical dynamic. This results in a non-sloping, square-like phonetogram contour. These outcomes show that SAL potentially can facilitate phonetographic analysis of the singing voice.

Key Words: Singing voice—Skin acceleration level—Phonetogram—Vocal function—Vowel variation—Across tone fluctuations—Differences between singing and speech.

INTRODUCTION

Rationales

The vocal folds are a vibrating system and phonatory problems are likely to be most pronounced

at certain frequencies and/or amplitudes of vibration. The phonetogram offers a convenient mapping of vocal effort and fundamental frequency (F_0), and might therefore be useful in delimiting problem areas. In speech, the phonetogram or the voice range profile is used extensively in research and clinical settings.^{1–13} This technique has also been applied to the classical singing voice.^{14–18} However, classical singers train to maximize vocal output by means of vocal tract modifications. This implies that, in a phonetogram of a singing voice, the relationship of sound pressure level (SPL) to F_0 and vocal effort differs from that in a speech phonetogram. This difference is important and needs to be considered in the interpretation of

Accepted for publication August 10, 2006.

This research is funded by Baxter & Alma Ricard Foundation.

From the Kungliga Tekniska Högskolan, Stockholm, Sweden.

Address correspondence and reprint requests to Anick Lamarche, Department of Speech Music and Hearing (TMH), School of Computer Science and Communication, Kungliga Tekniska Högskolan (KTH), Lindstedtsväg 24 SE-100 44 Stockholm, Sweden. E-mail: anick@nada.kth.se

Journal of Voice, Vol. 22, No. 1, pp. 10–22

0892-1997/\$34.00

© 2008 The Voice Foundation

doi:10.1016/j.jvoice.2006.08.005

phonetograms. Indeed, in recording phonetograms of singing voices, SPL can vary within tone due to vibrato and across tones in regard to singer specific vowel modifications. Since these variations, to a large extent, are consequences of the vocal tract acoustics and are not directly related to phonatory function, it would be useful to minimize them when phonation is the primary object of study. As a measure of vocal function, the electroglottogram or EGG has advantages. It is minimally influenced, if at all, by vowel production. On the other hand, Askénfelt et al as well as Baken demonstrated that EGG has limitations in depicting vocal fold oscillations. The EGG does not have any microphonic capacities and furthermore, it cannot provide any information for the open phase of vocal fold oscillation.^{19,20}

The skin acceleration level (SAL), if measured in the vicinity of the vocal folds, is another potential measure of phonatory activity. If the objective is to evaluate phonatory function, it is more relevant to estimate the intensity of the glottal source rather than the intensity of the radiated sound. In a 1983 study of chest wall vibrations, Sundberg noted that vibrations measured at the thyroid and the sternum are primarily determined by the voice source and to some extent modified by subglottal resonances. Hence it might be expected that vibrations measured at these locations would be less influenced by changes in the vocal tract. Moreover, it becomes a possible alternative for the vertical axis in the phonetogram and a replacement for SPL. SAL is mainly a measure of tissue vibrations rather than a measure of acoustic pressure and it is easily recorded near the vocal folds. One might also expect the vocal fold collisions to generate shock waves in the surrounding tissues. However, Sundberg investigated possible influences of colliding forces of the vocal folds and concluded that their contribution to vibrations recorded at the thyroid and sternum lamina is negligible.²¹

The subglottal pressure (P_S) drives the voice source. P_S is a main determinant of vocal loudness in speech and in singing and the literature demonstrates how P_S relates to SPL for both speech and singing. Therefore, it could be interesting to observe how SAL and SPL differ in their relationship to P_S . Generally, SAL seems to have the potential

to: (1) facilitate phonetographic analysis of the singing voice, (2) allow inclusion of all vowels in clinical evaluation, (3) address directly and unobtrusively the voice source, (4) allow singers more vocal and physical freedom during recordings, and (5) reduce influence of environmental noise on the recorded signal.

Earlier work

Accelerometers have occasionally been applied to speech and voice research, for example, in research on nasalization,^{22–24} F_0 extraction,^{19,25} frequency perturbation,²⁵ and alternative recording devices.²⁶ Recent studies have looked at SAL as an estimator for speech glottal characteristics²⁷ and also as an estimator for SPL for speech.²⁸ Švec et al showed that a near-to-linear relationship between SPL and SAL can be used to estimate long-term average SPL values in speech. Their data clearly show an established SAL correlation to long-term SPL: higher SPL corresponds to higher SAL in speech. The primary motivation for the current investigation was to find alternatives to the study of phonatory function in singers that would facilitate the interpretation of phonetograms. The questions to be answered were (1) does SAL vary less across vowels than does SPL? (2) is SAL more correlated to subglottal pressure than SPL? and finally (3) how does SAL measured in singers compare to findings for speech by Švec et al?

Hypotheses

The first question that was stated above leads to null hypothesis A: SPL and SAL are equally influenced by vowel variation. Our second question leads to null hypothesis B: SAL and SPL are equally correlated to P_S .

METHOD

To test these hypotheses, a number of singing tasks were designed to exercise variations in vowel, musical dynamic, and F_0 over a typical female singing range. Musical dynamic was included to obtain systematic variation in P_S . Through statistical analysis, the variance thereby incurred in SAL was compared to the variances incurred in SPL and P_S .

Each subject was instructed to warm up, before her arrival at the recording session, for a minimum of 5 minutes, and according to their personal warm-up routines. On arrival, the experimental procedure and tasks were explained. Subjects familiarized with the equipment and made a few trials. All recordings were performed at the NCVS Laboratories in Denver, Colorado. Recordings took place in a sound-isolated booth. Singers were asked to use a stage stance throughout the recording process. The experimenter was present to coach through different tasks as well as to monitor P_s signals on the oscilloscope.

Acoustic, aerodynamic, and accelerometric signals were recorded with the following equipment. Two accelerometers (Thin Case BU-7135; Knowles Acoustics, St. Louis, MO): one attached vertically at midline on the jugular notch and the other at midline on the sternum bone. Attachments and use of Mastisol surgical glue (Mastisol; Ferndale Laboratories, Ferndale, MI) and Suture-Strips (TS-3101; Derma Sciences, Elgin, IL) followed the protocol established in Popolo et al.²⁹ The airborne signal was recorded at 30 cm from the subject's mouth with the microphone of the sound level meter (Brüel & Kjaer 2238 Mediator, A weighted-slow; Brüel & Kjaer, Naerum, Denmark). Intraoral pressure during stop-plosives /p/ was measured with a pressure transducer (PT-series; Glottal Enterprises, Syracuse, NY).³⁰ Subjects were given the transducer to hold at the labial commissure during the performance of phonation tasks. The subjects familiarized themselves with the equipment and received brief oscilloscope feedback to facilitate the positioning of the

pressure transducer in their mouth and achieve a stable intraoral pressure during /p/ occlusions.

The microphone/sound level meter was connected through an amplifier to channel 0 of a model 4500 Kay CSL sound card (KayPentax, Lincoln Park, NJ). The pressure transducer was connected to channel 1 (DC) of the same card and the accelerometers to channels 2 and 3. A 20-dB attenuator pad (DGS pro-audio; Mouser Electronics, Mansfield, TX) was used when necessary to prevent clipping of the microphone signals (Figure 1 depicts the setup schematics). The sampling rate was 44 100 Hz. The four channels were recorded in synchrony and the resulting files were read and edited with *Cubase S.L.* (Version 1.07 build 97\2004 SE; Steinberg Media Technologies GmbH, Hamburg, Germany).

Calibration

Microphone and pressure transducer calibrations were performed at the beginning and the end of each subject's session. Accelerometer calibration followed NCVS-established calibration procedures for speech dosimetry (A. Starr, personal communication, August 2005). For the sound level calibrations, each subject phonated at three loudness levels and gains were adjusted to avoid clipping. The *Cubase S.L.* program was set to record position and a calibrator (Brüel & Kjaer 4231) was used to produce a 94-dB SPL re 20- μ Pa tone. Finally, pressure transducer calibrations were performed with a pneumotach calibration unit (Glottal Enterprises, Model MCU-4). Readings at 20, 10, and 5 cm water column were taken and recorded in the *Cubase S.L.* program.

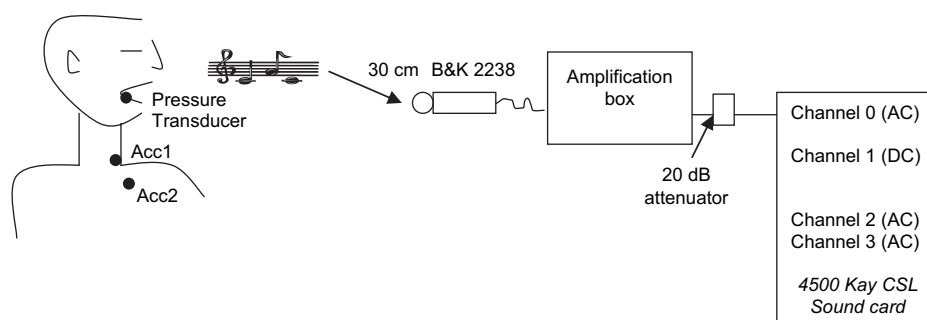


FIGURE 1. Schematic representation of the experimental setup. Accelerometers were attached at the jugular notch (Acc 1) and on the sternum bone (Acc 2) according to a protocol developed by Popolo et al.²⁹

Subjects and vocal tasks

Five female singers, three sopranos and two mezzo-sopranos, aged 20 to 30 years, participated in the recordings. Each singer had obtained a university certification in voice performance or formal classical training. Levels ranged from bachelor to DMA. It must be specified that only one singer met the criteria established for a professional singer.³¹ All singers reported good vocal health.

The subjects performed three tasks (Figure 2):

- (1) Sustain a tone at D5 (587 Hz) while singing a /pi pe pa po pu/ series in a slow tempo. This task was performed at three intensity levels (*piano*, *mezzo forte*, and *forte*). The exact task was then repeated at G5 (784 Hz).

- (2) Sing an ascending scale of an octave starting at a preferred F_0 and repeat each F_0 three times using the vowels /a/ and /i/. Again, this task was performed at all three intensity levels mentioned above, with /p/ occlusions preceding the vowel. Two subjects chose a C3 to C4 (131–262 Hz) scale, one a G4 to G5 (392–784 Hz), and two others D4 to D5 (294 Hz–587 Hz).
- (3) Arpeggiate an octave from F4 (349 Hz), repeating each F_0 three times. The task was performed at all three intensity levels and included /p/ occlusions and all /i e a o u/ vowels.

Each performance was carefully monitored and the tasks were repeated if, for example, the

Task 1-changing vowel

Task 2-changing frequency

Task 3-changing frequency and musical dynamic

FIGURE 2. Illustrations of the three tasks that subjects completed during recording.

oscilloscope displayed unstable P_S signals or if singers believed they could perform higher dynamic contrasts. At the end of each recording session, subjects filled out questionnaires concerning their voice and vocal experiences.

Data processing

Recorded files were truncated from 24- to 16-bit samples and they were losslessly compressed in *Flac* (Frontend 1.7.1, FLAC, <http://flac.sourceforge.net>). Each channel was saved separately and reopened as a .wav file and converted to .smp format with a file conversion utility (Audiofil; Hitech Development AB, Täby, Sweden). Files were then reorganized back into synchronized four-channel files. The pressure value corresponding to the onset of phonation was taken as the pressure immediately before the release of the plosive /p/ (Figure 3). In measuring P_S , pressure tokens were discarded if the /p/ occlusion and phonation were not perfectly aligned. This was seen in the case where singers did not always succeed in keeping a sustained *legato* from one plosive occlusion to the next. Tokens

were also discarded if they displayed instability, too much sharpness, or when a breath was taken. L_{eq} values were computed over the initial 200 milliseconds of each vowel sound, following the /p/. All signal manipulations and measurements were done using the Soundswell Signal Workstation 4.0 (Hitech Development AB, Täby, Sweden).

To make phonetograms of the microphone and accelerometer signals, the signal files were re-sampled to 16 kHz per channel. This was a requirement of the computerized phonetograph (Phog 2.0, Hitech Development AB, Täby, Sweden). Conventional phonetograms as well as SAL phonetograms were made of the complete recordings of each subject.

Statistical analysis

A univariate general linear model-based analysis of covariance (ANCOVA) was designed. Dependent variables were defined as SPL, SAL_N (SAL for notch), and SAL_S (SAL for sternum) and independent variables as F_0 , Dynamic, Vowel, and Subject. A univariate format was preferred to

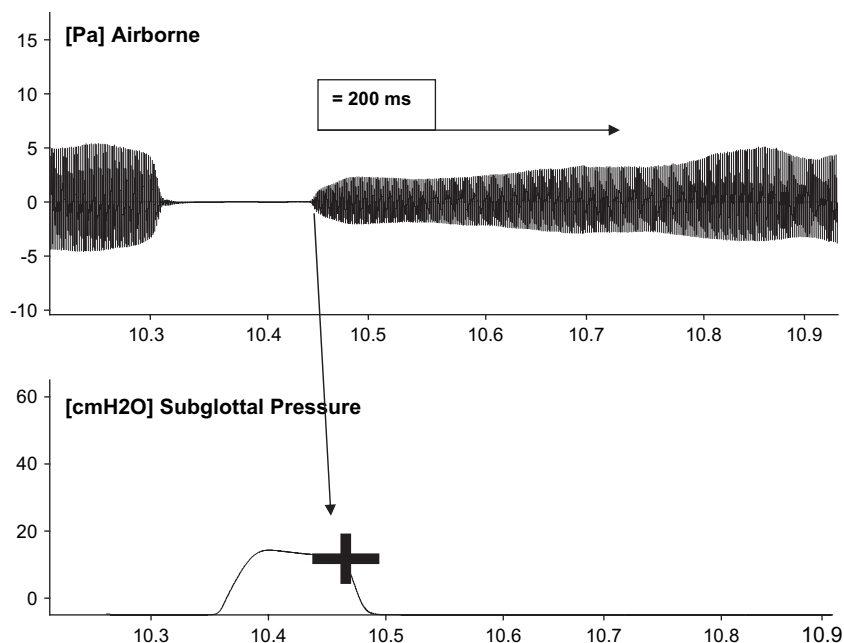


FIGURE 3. An example of the analysis points selected in audio and pressure signals. The intraoral pressure at *p*-release was used as an approximation of the subglottal pressure driving the first 200 milliseconds of subsequent phonation.

a multivariate to assess dependent variable behavior in isolation. Subject, Dynamic, and Vowel were treated as fixed factors, while F_0 was defined as a covariate.

The data were organized into two factorial designs, each with a balanced data set representing different levels of factors. For design 1, data recorded from the first task was combined with data from the third task. In the tasks for design 1, the subjects changed only the vowel or dynamic from token to token, while holding the F_0 constant (Table 1). For design 2, data from the second task were used (Table 2), in which subjects changed only the F_0 from token to token. Dividing the data into two groups by tasks should offer some insight as to the importance of tasks in the overall outcome. The division of the data also offers some indication of the reliability of the behavior observed across designs.

SAL was not calibrated against a reference level since SAL can be expected to vary from subject to subject, due to physiology and possible variations in transducer attachment. The intersubject variation in SAL is not relevant to this study. Rather, the SAL data were normalized by subject means, thereby excluding the expected variations in the gain of the SAL signals. This was done for each subject by computing the intrasubject average SAL within one design and one attachment (notch/sternum), and then subtracting the personal average from the raw SAL values. The SAL data were not normalized for standard deviation, since the variance in SAL is one of the outcomes of the experiment. Testing was performed using statistical software (SPSS Version 13.0; SPSS, Chicago, IL). The significance threshold was set to $P \leq 0.05$.

TABLE 1. Statistical Design 1 (60 Tokens per Subject) Used to Obtain a Balance Set of Data for a General Linear Model ANCOVA Analysis

Independent Variables	Statistical Label	Number	Definition
F_0	Covariate	4	349, 440, 587, 784 Hz
Subject	Fixed factor	5	2 mezzos, 3 sopranos
Vowel	Fixed factor	5	/pa/ /pe/ /pi/ /po/ /pu/
Dynamic	Fixed factor	3	<i>p—mf—f</i>

F_0 was defined as a covariate in the final model used.

RESULTS

Figure 4 depicts the collected data before overall statistical treatment and normalization of SAL. For each subject, the means and standard deviations obtained for three dependent factors are illustrated. SPL tended to be similar across all subjects, but both SAL signals showed some intersubject variation.

Design 1

The ANCOVA results for design 1 are given in Table 3. Interestingly, the chosen statistical model explained most of the variance in the data (see percentages in Table 3), provided that F_0 was defined as a covariate. It can be seen from the P values that the factors F_0 and Dynamic had a significant effect (at $P \leq 0.05$) on all the dependent variables: SPL, SAL_N , SAL_S , and P_S . This is, of course, as expected for SPL and P_S , since the voice output level rises with both F_0 and Dynamic. The expectations for SAL are not obvious. The factor Subject also had a significant effect on SPL and P_S (individuals differ in their choice of vocal power), but not on the two SAL measures, since they had been previously normalized. The Vowel factor was a significant source of variation in SPL and P_S ; however, the percentages of explained variance for the Vowel factor are much smaller than for other significant factor percentages. In Table 3, further comparisons of F values and percentage explained variance show that for SPL, F_0 was clearly the dominant source of variation (stronger even than Dynamic); whereas for SAL, Dynamic was the dominant source of variation. For all four dependent variables, Vowel was the weakest source of variation.

For the dependent variable SPL, there were no significant interactions between the fixed factors. For the two SAL measures and for P_S , there were significant but small interaction effects between Subject and Dynamic (1–3%). This means that different subjects produced slightly different increments in SAL and P_S between *piano*, *mezzo forte*, and *forte*. For P_S , there was also a significant interaction between Subject and Vowel; in other words, different subjects would exhibit different changes in P_S when changing vowel. Since we are not

TABLE 2. *Statistical Design 2 (144 Tokens per Subject) Used to Obtain a Balance Set of Data for a General Linear Model ANCOVA Analysis*

Independent Variables	Statistical Label	Number	Definition
F_0	Covariate	8	C major scale <262–523 Hz> D major scale <294–587 Hz> G major scale <392–784 Hz>
Subject	Fixed factor	5	2 mezzos 3 sopranos
Vowel	Fixed factor	2 × 3 repetitions	/papapa/ /pipipi/
Dynamic	Fixed factor	3	$p-mf-f$

F_0 was defined as a covariate in the final model used.

here concerned with individual behaviors, these interactions will not be discussed further.

Design 2

The ANCOVA results for design 2 are summarized in Table 4. Again, the statistical model seems to explain most of the variance (see percentages in Table 4). The overall pattern in the outcome was the same as was observed in design 1, with F_0 being the dominant source of variation for SPL, while Dynamic was the dominant source of variation for both the SAL measures. Generally, some significance levels were higher than those found for design 1 and Vowel variation presented a different pattern of significance; both SPL and SAL_N were significant for Vowel. On the other hand, SAL_S and P_S did not significantly change with Vowel. In this design, the F values were larger than those found for design 1, with the exception of Vowel for P_S but percentages of the explained variance are slightly lower. Nonetheless, design 2 confirms the most striking result of this study: in both statistical designs, F_0 was the dominant variation factor for SPL, while Dynamic was the dominant variation factor for SAL.

In the second design, practically all interactions between the fixed factors were significant, for all dependent variables, but their influence was small (1–3% of the variance explained). For the purpose of this study, these interactions do not seem to warrant a more detailed discussion.

Subglottal pressure

Finally, the P_S -SPL and P_S -SAL correlations were computed (Figures 5 and 6). SPL showed clearly *higher* correlation to P_S than did SAL.

This was true for both design 1 and design 2. The null hypothesis B, that SPL and SAL are equally correlated to P_S , is therefore rejected.

Examples of phonetograms are shown for one subject in Figures 7–9.

DISCUSSION

The first hypothesis of this study called for the investigation of the presence of vowel variation in SAL. Similar to the work done in Švec et al.,²⁸ vowel-induced variation in SAL and SPL was compared (Figures 10 and 11). Generally, speech phonetograms are recorded with the /a/ vowel to avoid variation in SPL between open and closed vowels. In the soprano singing tasks of the present

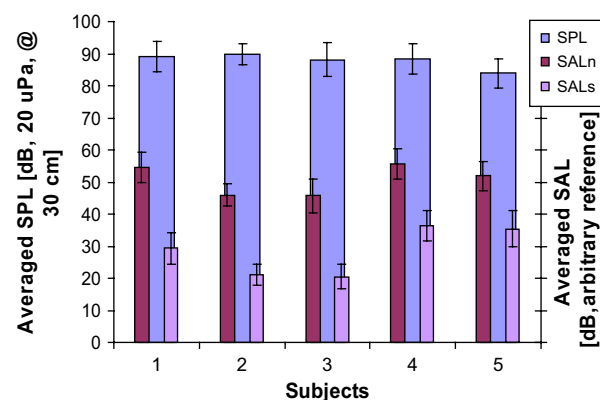


FIGURE 4. Subject means and standard deviations for three dependent variables: SPL, SAL_N , and SAL_S . For all subjects, SPL was quite similar. In this figure, SAL_N and SAL_S are not normalized and some intersubject variation can be observed. Intersubject differences in mean SAL were not important for the aims of this study and were canceled by normalizing to the subject mean.

TABLE 3. Design 1, Test Between-Subject Effects

Variables	SPL	SAL _N	SAL _S	P _S
<i>P</i>				
<i>F</i> ₀	0,000	0,006	0,000	0,000
Subject	0,000	(1,000)	(0,999)	0,000
Vowel	0,050	0,175	0,110	0,000
Dynamic	0,000	0,000	0,000	0,000
<i>F</i>				
<i>F</i> ₀	486,676	7,821	66,811	291,398
Subject	22,108	(0,010)	(0,025)	103,510
Vowel	2,406	1,603	1,911	6,074
Dynamic	62,287	155,722	171,808	90,457
% of Explained variance				
<i>F</i> ₀	49	1	9	24
Subject	9	0	0	34
Vowel	1	1	1	2
Dynamic	13	51	48	15
<i>R</i> ² (% of explained variance by model)	78	63	69	87
Observed power for $\alpha = 0.05$				
<i>F</i> ₀	1,000	,795	1,000	1,000
Subject	1,000	(,052)	(0,055)	1,000
Vowel	0,686	,490	0,572	0,985
Dynamic	1,000	1,000	1,000	1,000

Highest *F* values, all *F* values for vowel, and corresponding percentages of explained variance are given in bold to clearly depict the magnitude poles in the data. Values for the factor Subject and dependent variables SAL_N and SAL_S are in parentheses, since SAL was normalized for each subject. Frequency has a dominating influence on SPL variation but it is Dynamic which dominates SAL_N and SAL_S. Interestingly, the Vowel factor does not explain much of the variance for either SPL or SAL.

experiment, vowel variation was found to be an almost negligible source of SPL variation, when compared to the other experimental factors. *F*₀ was the dominant factor in terms of variation in SPL, and this result is supported by the literature. It is known that SPL in speech increases by approximately 9 dB per octave.³² The corresponding slope values observed in this study ranged from 20 to 30 dB per octave. Hence, in soprano singing, *F*₀ has a considerably stronger influence on SPL than it does in speech. This could be due to the *F*₁-*F*₀ matching that is conventional in high-pitched female singing. This matching would presumably become more precise with rising *F*₀.

The near absence of vowel variation in SPL and the strong *F*₀ dependency observed here both confirm the need for differentiation between speech and singing behaviors. Singers operate their vocal instrument characteristically on many different

levels, and a number of compensations can be at play in the production of an equal loudness and timbre across vowels. This raises the question whether vowel variation in *P*_S could be indicative of compensatory adjustments at the voice source. If so, we would expect the Vowel factor to be a stronger source of variation for *P*_S than for SPL. However, the ANCOVA results show that this was not the case. Rather, the present data suggest that if the singers systematically modify *P*_S with vowel, then such modifications are very small.

Overall, the results suggest that SAL cannot be proposed as a useful replacement for SPL merely on the grounds that it is a signal with reduced vowel variation. The first null hypothesis driving this work is, therefore, *not* rejected by our findings: in the singing voice tasks used, vowel changes caused little or no SPL variation, and in practice were negligible also in SAL.

TABLE 4. Design 2, Test Between-Subject Effects

Variables	SPL	SAL _N	SAL _S	P _S
<i>P</i>				
<i>F</i> ₀	0,000	0,000	0,000	0,000
Subject	0,000	(0,000)	(0,000)	0,000
Vowel	0,000	0,000	0,055	0,206
Dynamic	0,000	0,000	0,000	0,000
<i>F</i>				
<i>F</i> ₀	1863,362	145,144	503,107	1040,738
Subject	105,972	(11,081)	(38,642)	104,369
Vowel	36,538	57,639	3,702	1,604
Dynamic	539,606	763,143	838,468	534,905
% of Explained variance				
<i>F</i> ₀	44	5	16	28
Subject	10	2	5	11
Vowel	1	2	0	0
Dynamic	25	57	53	29
<i>R</i> ² (% of explained variance by model)	84	75	78	85
Observed power				
for $\alpha = 0.05$				
<i>F</i> ₀	1,000	1,000	1,000	1,000
Subject	1,000	(1,000)	(1,000)	1,000
Vowel	1,000	1,000	0,485	0,244
Dynamic	1,000	1,000	1,000	1,000

Highest *F* values, all *F* values for vowel, and corresponding percentages of explained variance are given in bold to clearly depict the magnitude poles in the data. Values for the factor Subject and dependent variables SAL_N and SAL_S are in parentheses, since SAL was normalized for each subject. Design 2 statistics show the trend observed in design 1 where Frequency is a dominant factor for SPL and Dynamic is the important factor for SAL. Here as well, the Vowel factor has a very small effect.

Although vowel variation was the primary topic of this study, other outcomes revealed some potentially useful aspects of SAL. The pronounced dependency of SPL on *F*₀ (20–30 dB per octave) is practically eliminated in SAL. For this study, the clear reduction of the influence of *F*₀ in SAL when compared to SPL is very interesting. *F*₀ remains a statistically significant source of variation across all dependent variables. However, although *F*₀ is significant for SAL, *F* values and percentages of explained variance are much lower than those for Dynamic, and thus indicate a weaker source of variation. The literature gives explanation for the reduced *F*₀ variation in SAL. In his study of chest wall vibrations, Sundberg demonstrated how sternum displacement amplitude lines up along a 12-dB slope when plotted according to *F*₀ and a constant vocal effort.²¹ Because acceleration is the second derivative of displacement, it is expected that the frequency related slope in SAL will have

12 dB less in inclination than for the displacement slope. This essentially agrees with the outcome of the present study.

For phonetography, these results point to the necessity for clear differentiation between the analysis of speech and high *F*₀ soprano singing. In speech studies, SPL estimation by SAL alone is successful whereas in the case of singing, this type of estimation would need to account for *F*₀.

In what concerns Dynamic, changes are somewhat smaller in SAL than in SPL (Figures 12 and 13). Nevertheless, the Dynamic variation in SAL is more explanatory than the one observed in SPL. SPL embeds a combination of different factors that work together in amplifying the voice. As seen in results obtained above, *F*₀ is the most important of these factors. Since there is a reduced *F*₀ effect, the source of variation in SAL is mostly attributed to the Dynamic factor. Indeed, results demonstrate clearly the dominant influence of

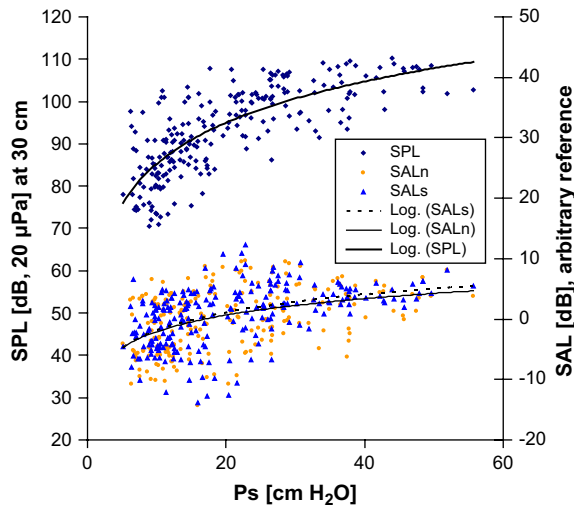


FIGURE 5. Correlations found for SPL, SAL_N, and SAL_S to P_S according to statistical design 1. The regression outcomes were $Y_{SPL} = 14 \ln(x) + 53$, $r^2 = 0,5968$, $Y_{SAL_N} = 4 \ln(x) - 11$, $r^2 = 0,1833$, and $Y_{SAL_S} = 4 \ln(x) - 12$, $r^2 = 0,2455$. SAL for both attachments was only very weakly correlated to subglottic pressure, whereas SPL followed trends documented in the literature. The null hypothesis that SPL and SAL are equally correlated to P_S is therefore rejected.

Dynamic as a source of variation in SAL. This points to the potential of SAL variation to display more immediate information and could result in interesting implications for phonetograms.

The phonetograms in Figures 7–9 exemplify clearly the type of results obtained when SAL is substituted for SPL on the phonetogram y-axis. The distribution shape of phonetogram changes from steeply inclined to horizontal and almost rectangular. SAL might therefore simplify the interpretation of the phonetogram, by showing results without the usual bias due to F_0 , which is even stronger in singing than in speech.

The second hypothesis concerned the subglottal pressure. Since P_S drives the vocal chords, the expectation was for P_S to have equal or more correlation to SAL at the notch and at the sternum than what is observed for SPL. This expectation was not borne out by the results. Nevertheless, the P_S -SPL relationship was similar to that which has been reported in previous literature. In speech, Fant originally established a 9.5-dB theoretical increase in SPL for every doubling in P_S .³³ In singers, Sjölander and Sundberg, in agreement with Schutte's studies,³⁴ observed that the decibel

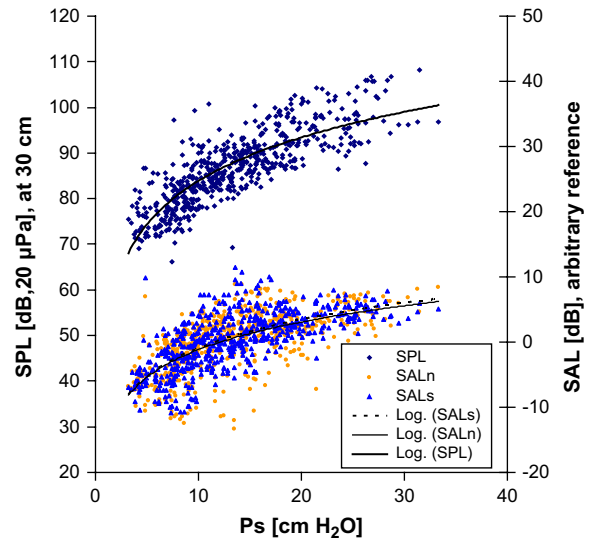


FIGURE 6. Depicted here are the correlations found for SPL, SAL_N, and SAL_S to P_S according to statistical design 2. Slopes and regressions are defined by the following equations: $Y_{SPL} = 13 \ln(x) + 52$, $r^2 = 0,6732$, $Y_{SAL_N} = 6 \ln(x) - 15$, $r^2 = 0,4171$, and $Y_{SAL_S} = 6 \ln(x) - 15$, $r^2 = 0,4771$. Clearly, SAL for both attachments is only very weakly correlated to subglottic pressure, whereas SPL reflects trends documented in the literature. The null hypothesis that SPL and SAL are equally correlated to P_S is therefore rejected.

increase was higher. They established an average of 12 dB.³⁵ It is interesting to note that those reports addressed only the male singing voice. According to our results, the relationship between

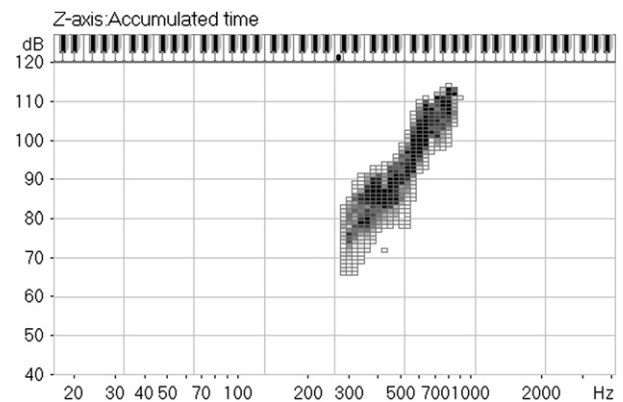


FIGURE 7. Aggregate phonetogram of all tasks performed by subject 2. The format below is the standard display used in clinics and in experiments, with SPL on the y-axis and log frequency on the x-axis. The phonetogram exhibits a pronounced slope with frequency, and shows a 20- to 30-dB increase per octave.

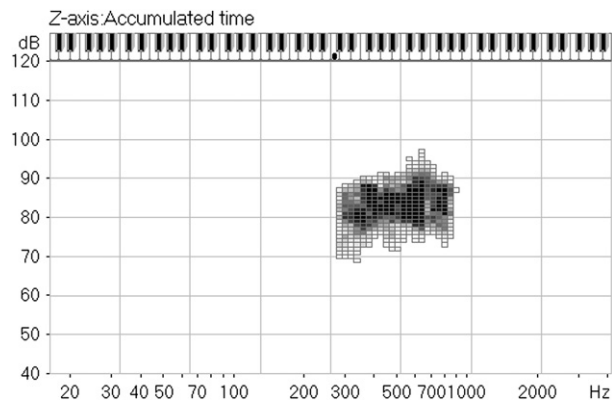


FIGURE 8. Alternative phonetogram of all tasks performed by subject 2, with SAL_N (at the jugular notch) on the y-axis rather than SPL. Although a slight slope remains, the observed dominance of the Frequency factor in the traditional phonetogram is almost gone. In this format, the Dynamic factor is the major source of level variation. This has the potential to simplify phonetogram interpretation.

SPL and P_S reported for male singers also holds true for female voices. From the regression equations shown in Figures 4 and 5, we find that SPL on average increased +12 dB per doubling of P_S in design 1 and +11 dB in design 2.

Unexpectedly, SAL showed weaker correlation to P_S than what was found for SPL. These findings oppose the null hypothesis above mentioned. A possible explanation could be the spectral

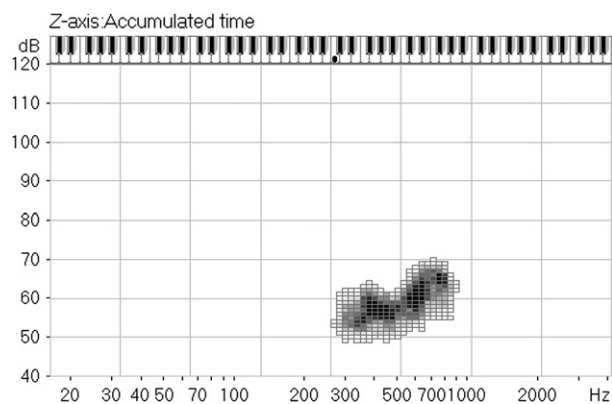


FIGURE 9. Alternative phonetogram of all tasks performed by subject 2, with SAL_N (at the jugular notch) on the y-axis rather than SPL. Although a slight slope remains, the observed dominance of the Frequency factor in the traditional phonetogram is almost gone. In this format, the Dynamic factor is the major source of level variation. This has the potential to simplify phonetogram interpretation.

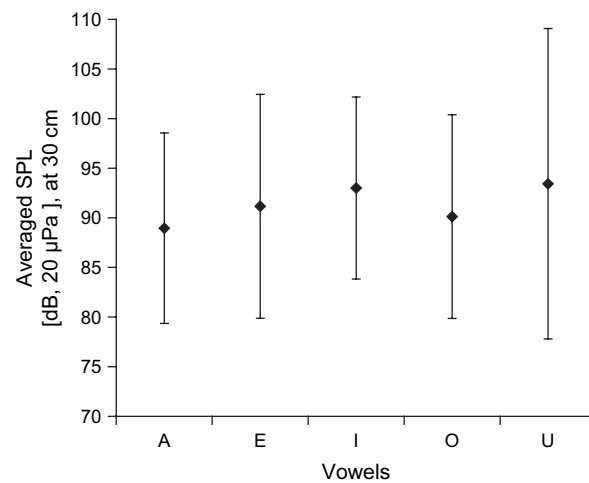


FIGURE 10. Overall means and standard deviations in SPL, by vowel. The means changed less than 4 dB between vowels, which would probably be negligible in practice.

characteristics of the skin acceleration signal. SAL is dominated by the level of the first partial. If the effect of increasing P_S is mostly to boost the rather weak higher partials, then there would be very little effect on the overall signal level. This issue would require further study. Until this is clarified, the SAL correlation to P_S does not in itself support the use of SAL as an alternative to SPL or as a method for vocal function quantification.

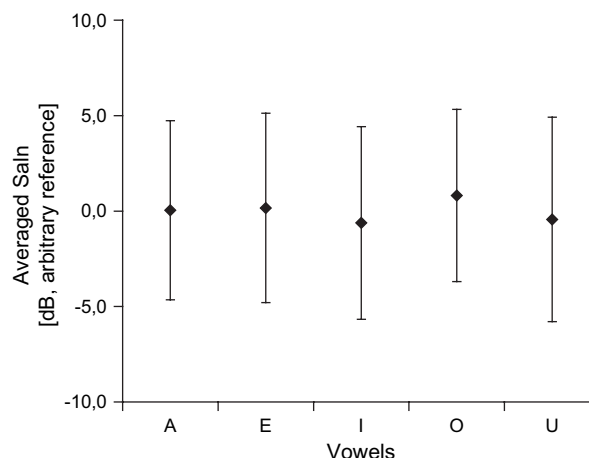


FIGURE 11. Overall means and standard deviations in SAL_N , by vowel. In comparison to the vowel variation observed in SPL, the variation in SAL_N was even smaller, and did not have any important effect on the signal. The outcome for SAL_S was practically identical.

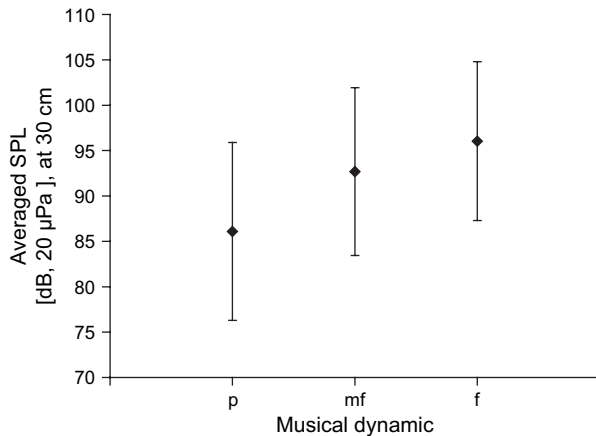


FIGURE 12. Overall means and standard deviations in SPL, by Dynamic. The Dynamic variation in SPL was considerable, with a 10-dB difference in the means; however, it was smaller than the Frequency variation, which dominated SPL.

CONCLUSION

From this study, it appears that SAL does have the potential to (1) facilitate phonetographic analysis of the singing voice, (2) allow singers more vocal and movement freedom during recordings, and (3) reduce influence of environmental noise on the recorded signal. Nonetheless, SAL remains an indirect assessment of vocal function. It would be necessary to investigate further the spectral properties of the skin acceleration signal to assess in detail

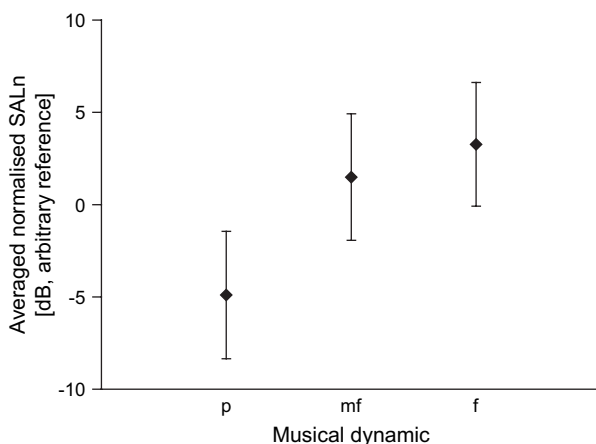


FIGURE 13. Overall means and standard deviations in SAL_N , by Dynamic. The Dynamic factor in SAL_N (7 dB) was important in that it dominated the variation in SAL_N : the standard deviations were smaller than the differences in the means. Again, the same was observed for SAL_S .

the behavior of the fundamental in SAL and how precisely dynamic is displayed.

SAL proves to be potentially useful to phonetography, but for different reasons than those initially expected. There are two main observations which have important consequences. First, SAL offers a signal which is minimally influenced by F_0 and therefore, is able to clearly illustrate effects of the musical dynamic. And second, vowel variation, when addressing the high singing voice, is practically negligible in both SPL and SAL. This finding is important in that it underpins differences between speech and singing. In singing, phonetograms might not be as influenced by the use of different vowels as they are in speech. This fact would allow for much more freedom in performing phonetograms of the singing voice given that singing tasks involving different vowels and song/aria excerpts could be used.

In the process of this investigation other pertinent questions were encountered. For example, are the SPL variations that are due to vibrato smaller in SAL phonetogram output? Since it is established that vowel variation for the singing voice has a minimal influence, it would be interesting to explore also the other type of fluctuations observed in real-time acquisition of phonetograms.

Acknowledgments: The authors wish to thank Dan Mattson and Mattias Heldner for statistical advice, Johan Sundberg, Jan Svec, and Svante Grandqvist, the NCVS team, and especially Andrew Starr for recording assistance.

REFERENCES

1. Damste PH. The phonetogram. *Pract Otorhinolaryngol (Basel)*. 1970;32:185–187.
2. Dejonckere PH. Effects of phoniatric therapy on the phonetogram. *Acta Otorhinolaryngol Belg*. 1980;34:309–317.
3. Gramming P, Åkerlund L. Non-organic dysphonia II. Phonetograms for normal and pathological voices. *Acta Otolaryngol*. 1988;106:468–476.
4. Gramming P, Sundberg J. Spectrum factors relevant to phonetogram measurement. *J Acoust Soc Am*. 1988;83:2352–2360.
5. Pabon JP, Plomp R. Automatic phonetogram recording supplemented with acoustical voice-quality parameters. *J Speech Hear Res*. 1988;31:710–722.

6. Titze IR. Acoustic interpretation of the voice range profile (phonetogram). *J Speech Hear Res.* 1992;35:21–34.
7. Coleman RF. Sources of variation in phonetograms. *J Voice.* 1993;7:1–14.
8. Sulter AM, Wit HP, Schutte HK, Miller DG. A structured approach to voice range profile (phonetogram) analysis. *J Speech Hear Res.* 1994;37:1076–1085.
9. Behrman A, Agresti CJ, Blumstein E, Sharma G. Meaningful features of voice range profiles from patients with organic vocal fold pathology: a preliminary study. *J Voice.* 1996;10:269–283.
10. Hilgenheger C, Sarfati J, Rey E, Sittel C, Eckel HE. Value of the relative phonetogram (RP) for the evaluation of organic dysphonias. *Rev Laryngol Otol Rhinol (Bord).* 1999;120:231–237.
11. Dejonckere PH. Efficacy of voice therapy assessed with the Voice Range Profile (Phonetogram). *Rev Laryngol Otol Rhinol (Bord).* 2003;124:285–289.
12. Södersten M, Ternström S, Bohman M. Loud speech in realistic environmental noise: phonetogram data, perceptual voice quality, subjective ratings, and gender differences in healthy speakers. *J Voice.* 2005;19:29–46.
13. Ma E, Robertson J, Radford C, Vagne S, El-Halabi R, Yiu E. Reliability of speaking and maximum voice range measures in screening dysphonia. *J Voice.* May 4, 2006. In press [Epub ahead of print].
14. Klingholtz F. The voice of the singer in the phonetogram. *Laryngorhinootologie.* 1989;68:62–66.
15. Pedersen MF. A longitudinal pilot study on phonetograms/voice profiles in pre-pubertal choir boys. *Clin Otolaryngol.* 1993;18:488–491.
16. Awan SN. Superimposition of speaking voice characteristics and phonetograms in untrained and trained vocal groups. *J Voice.* 1993;7:30–37.
17. 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.
18. Hunter EJ, Svec JG, Titze IR. Comparison of the produced and perceived voice range profiles in untrained and trained classical singers. *J Voice.* 2005. In press [Epub ahead of print].
19. Askenfelt A, Gauffin J, Sundberg J, Kitzing P. A comparison of contact microphone and electroglottograph for the measurement of vocal fundamental frequency. *J Speech Hear Res.* 1980;23:258–273.
20. Baken RJ. Electroglottography. *J Voice.* 1992;6:98–109.
21. Sundberg J. Chest wall vibrations in singers. *J Speech Hear Res.* 1983;25:329–340.
22. Stevens K, Kalikow D, Willemain T. A miniature accelerometer for detecting glottal waveforms and nasalization. *J Speech Hear Res.* 1975;18:594–599.
23. Lippmann RP. Detecting nasalization using a low-cost miniature accelerometer. *J Speech Hear Res.* 1981;24:314–317.
24. Horii Y. An accelerometric measure as a physical correlate of perceived hypernasality in speech. *J Speech Hear Res.* 1983;26:476–480.
25. Horii Y. Some voice fundamental frequency characteristics of oral reading and spontaneous speech by hard-of-hearing young women. *J Speech Hear Res.* 1982;25:608–610.
26. Coleman RF. Comparison of microphone and neck-mounted accelerometer monitoring of the performing voice. *J Voice.* 1988;2:200–205.
27. Cheyne II, HA. *Estimating glottal voicing source characteristics by measuring and modeling the acceleration of the skin on the neck* [doctoral thesis]. Massachusetts: Institute of Technology; 2002.
28. Švec JG, Titze IR, Popolo PS. Estimation of sound pressure levels of voice speech from skin vibration of the neck. *J Acoust Soc Am.* 2005;117:1386–1394.
29. Popolo PS, Svec JG, Hunter EJ, Starr AC, Rogge-Miller K, Nix J, Titze IR. Teachers' Guide to Voice Dosimetry. NCVS Memo No. 04 2004; NCVS Online Technical Memo.
30. Rothenberg M. Interpolating subglottal pressure from oral pressure. *J Speech Hear Disord.* 1982;47:218–224.
31. Bunch M, Chapman J. Taxonomy of singers used as subjects in scientific research. *J Voice.* 2000;14:363–369.
32. Titze IR, Sundberg J. Vocal intensity in speakers and singers. *J Acoust Soc Am.* 1992;91:2936–2946.
33. Fant G. Preliminaries to analysis of the human voice source. *KTH Speech Music Hear Quart Prog Stat.* 1982; Report 4:1–27.
34. Schutte H. *The efficiency of voice production* [doctoral dissertation]. The Netherlands: Groningen University; 1980.
35. Sjölander P, Sundberg J. Spectrum effects of subglottal pressure variation in professional baritone singers. *J Acoust Soc Am.* 2004;115:1270–1273.