Subglottal resonances of adult male and female native speakers of American English

Steven M. Lulich[a] and John R. Morton
Department of Psychology, Washington University, One Brookings Drive, Saint Louis, Missouri 63130

Harish Arsikere
Department of Electrical Engineering, University of California Los Angeles, 405 Hilgard Avenue, Los Angeles, California 90095

Mitchell S. Sommers
Department of Psychology, Washington University, One Brookings Drive, Saint Louis, Missouri 63130

Gary K. F. Leung and Abeer Alwan
Department of Electrical Engineering, University of California Los Angeles, 405 Hilgard Avenue, Los Angeles, California 90095

This paper presents a large-scale study of subglottal resonances (SGRs) (the resonant frequencies of the tracheo-bronchial tree) and their relations to various acoustical and physiological characteristics of speakers. The paper presents data from a corpus of simultaneous microphone and accelerometer recordings of consonant-vowel-consonant (CVC) words embedded in a carrier phrase spoken by 25 male and 25 female native speakers of American English ranging in age from 18 to 24 yr. The corpus contains 17 500 utterances of 14 American English monophthongs, diphthongs, and the rhotic approximant [ɹ] in various CVC contexts. Only monophthongs are analyzed in this paper. Speaker height and age were also recorded. Findings include (1) normative data on the frequency distribution of SGRs for young adults, (2) the dependence of SGRs on height, (3) the lack of a correlation between SGRs and formants or the fundamental frequency, (4) a poor correlation of the first SGR with the second and third SGRs but a strong correlation between the second and third SGRs, and (5) a significant effect of vowel category on SGR frequencies, although this effect is smaller than the measurement standard deviations and therefore negligible for practical purposes. © 2012 Acoustical Society of America. [http://dx.doi.org/10.1121/1.4748582]

I. INTRODUCTION

In studies of speech production and perception, attention is generally focused on the acoustic properties of the larynx and vocal tract (i.e., supraglottal airway), such as the fundamental frequency (F0) and the frequencies of the first three formants (Stevens, 1998; Ladefoged and Maddieson, 1996; Liberman, 1996). Relatively little attention has been given to the acoustic properties of the subglottal airways until recently. The subglottal system, composed of the airways of the tracheobronchial tree and the surrounding tissues, forms an integral and important part of the human speech production apparatus. It is the battery that powers airflow through the larynx and vocal tract, allowing for the generation of most of the sound sources used in languages around the world (Stevens, 1998, pp. 55–126; Ladefoged and Maddieson, 1996, pp. 77–90). But it also contributes to the filter and to the nonlinear interaction between the filter and the phonation source (Stevens, 1998; Fant, 1960; Titze, 2006).

The subglottal resonances (SGRs) are the natural frequencies of the subglottal system and correspond to the complex conjugate pairs of poles in the subglottal input impedance (measured from the top of the trachea looking down). During speech, the subglottal system is acoustically coupled to the vocal tract via the larynx, and the amount of coupling depends on the configuration of the larynx (e.g., coupling is greater when the glottis is open) (Stevens, 1998; Chi and Sonderegger, 2007; Lulich, 2010) as well as the configuration of the vocal tract (e.g., coupling is greater when the formants are close to the SGRs) (Hanson and Stevens, 1995; Stevens, 1998). As a result of coupling, the SGRs are perturbed somewhat toward higher frequencies. The resulting SGRs can be measured from recordings of the vibration of the skin of the neck during phonation, much like speech formants are measured from microphone recordings. The lowest three resonances (Sg1, Sg2, and Sg3) have been claimed to play various roles in speech production and perception (Stevens, 1998; Lulich et al., 2007) and have been incorporated into automatic speaker normalization algorithms (Wang et al., 2009b; Wang et al., 2009a).

Studies have shown that the SGRs appear in vowel spectra as pole-zero pairs that can interrupt formant trajectories (Stevens, 1998; Chi and Sonderegger, 2007; Jung, 2009a,b; Lulich, 2010); they are perceptually salient (Lulich et al., 2007) and useful in speech recognition technologies (Wang et al., 2008; Wang et al., 2009b; Lulich and Chen, 2009;

---

[a]Author to whom correspondence should be addressed. Also at: Department of Speech and Hearing Sciences, Indiana University, Bloomington, IN 47405. Electronic mail: slulich@wustl.edu

Csapó and Németh, 2009); they play a role in defining vowel and consonant contrasts in several languages, including American English, Mexican Spanish, High German and Swabian German, Standard Korean, and Standard Hungarian (Stevens, 1998; Madsack *et al.*, 2008; Wokurek and Madsack, 2008, 2009, 2011; Jung, 2009b; Csapó *et al.*, 2009; Gráczi *et al.*, 2011; Lulich, 2010; Wang *et al.*, 2009b); and they can contribute to either stabilization or destabilization of vocal fold motion and various kinds of irregularities in the phonation sound source (Zhang *et al.*, 2006; Zañartu *et al.*, 2007; Titze, 2008; Titze *et al.*, 2008; Zañartu *et al.*, 2011). There has also been interest in the use of the subglottal input impedance by clinicians for monitoring aspects of lung health (Fredberg, 1978; Fredberg and Hoenig, 1978; Fredberg and Moore, 1978; Suki *et al.*, 1993; Habib *et al.*, 1994; Mansfield, 1996).

Given the diverse roles and effects of the subglottal input impedance and its resonances in speech production, it is critical that their actual properties be known for a large number of speakers. This has been difficult to achieve because typical methods for directly measuring the input impedance are quite invasive and/or technically challenging, requiring access to patients with laryngectomies (Ishizaka *et al.*, 1976), placing miniature pressure transducers below the glottis (Cranen and Boves, 1987), or using an endotracheal tube (Habib *et al.*, 1994). An alternative, noninvasive method is the use of an accelerometer placed against the skin of the neck (Cheyne, 2002; Chi and Sonderegger, 2007; Madsack *et al.*, 2008; Wokurek and Madsack, 2008, 2009; Csapó *et al.*, 2009; Lulich, 2010; Wokurek and Madsack, 2011; Gráczi *et al.*, 2011; Lulich *et al.*, 2011a; Lulich *et al.*, 2011b). In this case, the source of sound comes from the phonation volume velocity (as is the case in microphone recordings of vowels), and the motion of the neck tissues (and hence the accelerometer) is related to the pressure at the top of the trachea. This results in a frequency spectrum closely related to the input impedance but sampled by the source harmonics and partially shaped by the source spectral envelope and effects of acoustic coupling with the vocal tract. One goal of the present study is to establish how accelerometer measurements of SGRs compare with measurements using more invasive procedures.

Few studies have collected information about speaker height in addition to SGRs (Habib *et al.*, 1994; Jung, 2009a; Lulich *et al.*, 2011a), and few have measured SGRs separately in different vowels (Cranen and Boves, 1987; Wokurek and Madsack, 2008, 2011). Of these, Habib *et al.* (1994) and Jung (2009a) measured SGRs in fewer than 10 speakers, and Cranen and Boves (1987) measured SGRs in only two speakers (i.e., the authors themselves). Wokurek and Madsack (2008) and Wokurek and Madsack (2011) measured SGRs in 19 and 16 speakers, respectively, but did not report on the heights of any of their speakers. Most studies have not reported any standard deviations either within speakers or across speakers, and there is currently little empirical data regarding within- or across-speaker variability in SGR frequencies. No studies have reported the correlations of SGRs among each other or with F0, the formants, or speaker age. The present study therefore aims to fill in several gaps in our current knowledge of SGRs and their measurement using accelerometers.

In this paper, we present data from 50 adult native speakers of American English (25 males and 25 females) whose subglottal acoustics were recorded by an accelerometer placed against the skin of the neck at the same time that microphone recordings were being made of their speech. The data were analyzed to determine how the subglottal resonances vary by gender, height, age, and vowel and whether they are correlated with each other and with the fundamental frequency and formant frequencies.

In Sec. II, we present our methods for recording and analyzing microphone and accelerometer signals and give a basic description of the corpus. In Sec. III, we present the results of our analyses. Section IV provides a discussion of the results, and Sec. V concludes this study.

II. METHODS

A. Participants

A total of 50 adult native speakers of American English (25 male and 25 female) were recruited through the Washington University Psychology Department's online research participation website. The majority of participants were Washington University students, who represent a wide range of American English dialects. In general, most of these students are speakers of the mid-American English dialect, although we did not specifically inquire into or investigate the dialects of the speakers. The mean ages for male and female speakers were 20.47 and 20.58 yr, respectively. The youngest speakers were 18 yr old, and the oldest speakers were 24 yr old. None of the participants reported any history of speech or hearing disorders.

B. The WashU-UCLA corpus

1. Recordings and stimuli

To investigate the acoustic properties of the participants' speech and subglottal resonances, recordings were made with a SHURE PG27 microphone and a K&K Sound HotSpot accelerometer while participants sat in a double-walled, sound attenuating booth. For all of the microphone and accelerometer signals, the sampling rate was 48 kHz, and the bit resolution was 16 bit/sample. Both the microphone and the accelerometer signals were recorded in MATLAB using a two-channel M-Audio MobilePre USB pre-amplifier connected to a computer running Windows VISTA. The microphone was placed roughly 20 cm in front of the speaker and slightly to the right to avoid distortion due to airflow during high airflow sounds (e.g., the fricative [s] or the release burst of the stop [p]). The speaker was instructed on how to hold the accelerometer him/herself against the skin of the neck at the cricoid cartilage below the level of the glottis. The placement of the accelerometer below the larynx ensures that the formants are not significantly present in the signal because of canceling with zeros introduced from the vocal tract on the other side of the laryngeal phonation source.

Directly in front of the speaker was a computer monitor that displayed sentences for the speaker to read. Various

consonant-vowel-consonant (CVC) words were embedded in the carrier phrase "I said a _____ again" and displayed on the monitor to be read aloud by the speaker. The CVC words were divided into two lists, and each list was recorded separately. The first list was composed of bVb, dVb, and gVb words, using seven of the American English vowels; this will allow for investigations of the effects of coarticulation between the vowel and its consonantal context with respect to subglottal resonances in other studies (cf. Gráczi *et al.*, 2011). The second list was an expanded version of the Peterson and Barney hVd word list (Peterson and Barney, 1952) and included 14 of the vowels of American English, including the rhotic approximant, [ɹ]. This allows for a direct comparison of our corpus with other corpora using the same word list. The complete set of CVC words in the corpus is given in Table I.

For each list, all of the relevant CVC words were randomly presented to participants at least 10 times each (the CVb word list was presented first, followed by a short break and then the hVd word list). Tokens that were judged by the experimenter to be of poor quality (e.g., because of mis-reading by the speaker, noise due to movements of the speaker, or the speaker becoming tongue-tied) were tagged and appended to the end of the list and re-recorded in the order in which they originally occurred. After the recordings were finished, each utterance was screened by another investigator to ensure that every CVC word was represented by 10 tokens with minimal background noise in the microphone signal. If this screening procedure resulted in fewer than 10 tokens of each CVC word, the speaker was discarded from the database and another speaker was recorded. This screening process resulted in 10 speakers being discarded. The speakers are identified numerically as speaker 9, speaker 10, etc. Speakers 1–8 were not included in the final database due to initial errors in the recording protocol. In all, 68 speakers were recorded. The final corpus therefore contains microphone and accelerometer signals from a total of 17 500 utterances (10 repetitions of 35 CVC words by 50 speakers). In addition, the corpus includes accelerometer signals of two tokens of the sustained vowel [a:] produced by each speaker, in which there was special emphasis on obtaining high quality accelerometer recordings with clear resonance structure up to the third subglottal resonance (Sg3). These high quality accelerometer recordings were obtained using WAVESURFER (Sjölander and Beskow, 2000) by allowing the speaker and the experimenter to interactively adjust the placement of the accelerometer as well as the vocal pitch (F0) and loudness of the speaker until the best quality signal was achieved. This was an important step because the quality of the accelerometer signals during the list recordings was expected to be somewhat variable.

Also recorded were the speaker's self-reported height (converted to centimeters), date of birth (converted to age in months), and gender.

2. Microphone spectrogram labeling

Using PRAAT (Boersma and Weenink, 2010), the start and end of the target vowel in each of the 17 500 microphone recordings were manually labeled by a single investigator (to ensure consistency). For monophthongs and the approximant [ɹ], the middle of the steady-state portion of the vowel was also labeled. This was generally near the vowel mid-point. For diphthongs, the nucleus was labeled either in the middle of the steady state, if a steady state existed, or just before the onset of rapid formant movements. Each microphone recording therefore has a corresponding label file saved in the PRAAT TextGrid format.

The start of each target vowel from the first (CVb) list was labeled where the formants become visible immediately after the initial plosive consonant. Vowels from the second (hVd) list were labeled when the formants were visible and the waveform demonstrated a significant deviation from previous aspirated pulses. In both lists, the end of the vowel was labeled at the point of closure of the final stop consonant. In general, the placement of each label was guided by inspection of the spectrogram and waveform and by listening. Labels were placed at the first zero-crossing of the waveform following the negative-going pulse marking the instant of vocal fold closure.

3. Microphone analysis methods

Microphone signals were down-sampled to 10 kHz and analyzed using WAVESURFER (Sjölander and Beskow, 2000). For automatic formant extraction, a 49 ms long Hamming window was used with a shift size of 5 ms. The LPC order was 12, and the pre-emphasis factor was 0.96. For automatic pitch (F0) extraction, a 7.5 ms long Hamming window was used with a shift size of 5 ms. The minimum and maximum pitch parameters were set to 60 and 400 Hz, respectively, and the entropic signal processing system (ESPS) method was used (Talkin, 1995). The values of F0, F1, F2, and F3 near the labeled steady-state part of the vowel were averaged over five frames centered around the label location. The extracted F0 and formant frequencies were not manually corrected.

C. Accelerometer analysis methods

The nine vowels [i, ɪ, ɛ, æ, a, ʌ, o, ʊ, u] in hVd words (4500 tokens in all) were analyzed in this paper. Each token was inspected, and if the accelerometer signal was of sufficient quality, the first three subglottal resonances (Sg1, Sg2, Sg3) were measured as follows.

Each token was down-sampled to between 6 and 10 kHz, depending on how noisy the signal was at high frequencies, pre-emphasized with a factor of 0.96, and a segment from the steady-state portion of the vowel was extracted. This segment was not necessarily centered at the steady-state label because the quality of the accelerometer

TABLE I. Complete list of CVCs recorded for the corpus. Various vowels (including the approximant [ɹ]) were recorded in up to four different consonant contexts. Phonological feature specifications are also given for the features [low] and [back] (Stevens, 1998, p. 253)

	i	ɪ	e	ɛ	æ	a	ʌ	o	ʊ	u	aɪ	aʊ	ɔɪ	ɹ
hVd	i	ɪ	e	ɛ	æ	a	ʌ	o	ʊ	u	aɪ	aʊ	ɔɪ	ɹ
bVb	i	—	—	ɛ	—	a	—	—	—	u	aɪ	aʊ	ɔɪ	—
dVb	i	—	—	ɛ	—	a	—	—	—	u	aɪ	aʊ	ɔɪ	—
gVb	i	—	—	ɛ	—	a	—	—	—	u	aɪ	aʊ	ɔɪ	—
[low]	—	—	—	—	+	+	+	—	—	—				—
[back]	—	—	—	—	—	+	+	+	+	+				+

recording could vary even within the time-course of the vowel, sometimes leading to weak or noisy signals immediately around the steady-state label which were unsuitable for measuring the SGRs. The length of each segment was approximately four pitch periods.

Manual measurements of the SGRs were guided by visual inspection of the discrete Fourier transform (DFT) spectrum, the linear predictive coding (LPC) spectrum, and an estimate of the wideband power spectral density (WPSD) (Nelson, 1997). In particular, each SGR was measured by choosing either the LPC peak or the WPSD peak depending on which of the two provided a more accurate representation of the DFT envelope. If neither spectral representation was satisfactory for a particular SGR, the SGR frequency was not measured. Hence it is important to note that not all three SGRs were necessarily measured in every vowel token chosen for analysis. In general, it was more difficult to measure Sg1 and Sg3 than to measure Sg2. While the measurement of Sg1 was sometimes difficult (especially for high-pitched speakers) due to its proximity to strong low-frequency harmonics, the measurement of Sg3 was always difficult owing to the attenuation of high frequencies caused by the low-pass nature of the neck tissues and skin. For Sg1, the WPSD measurements were generally more reliable than the corresponding LPC measurements. For Sg2 and Sg3, the two representations were usually more or less identical. Figure 1 shows the DFT spectrum, the LPC spectrum, and the estimated WPSD of a sample accelerometer segment from speaker 12.

The estimated WPSD can be treated qualitatively as the envelope of the DFT spectrum. To obtain the estimated WPSD, the approach outlined in Umesh *et al.* (1999) was adopted. The WPSD was estimated by subdividing the vowel segment into several overlapping frames, calculating an autocorrelation function for each frame after applying a Hamming window, and obtaining the DFT of the averaged autocorrelation function. The overlap between successive frames was fixed at 80% of the frame size, and the frame size itself was

varied between 0.9 and 1.1 times the pitch period such that the resulting spectral envelope was of the best possible quality.

The total number of Sg1, Sg2, and Sg3 measurements was 657, 678, and 306, respectively.

D. Fundamental frequency and formant distributions

Table II gives the vowel-specific means and standard deviations of formant and fundamental frequency (F0) measurements made from the microphone recordings and separated by gender. The ranges of formant and F0 measurements are similar to those previously reported for adults by Peterson and Barney (1952) and Hillenbrand *et al.* (1995) and somewhat larger than those reported for adults (ages 25–50) by Lee *et al.* (1999). The extent of F0 variability is somewhat larger, and the mean F0 values for both males and females are slightly lower than those reported by the same three studies. The coefficients of variation (COVs = standard deviation normalized by the mean, $s/\bar{x}$) are generally between 10% and 30%.

E. Statistical analyses

MATLAB was used for all analyses and statistical calculations. In the statistical analyses, a significance level of $\alpha = 0.05$ was used. Correlation coefficients presented in tables are indicated by an asterisk when statistically significant.

III. RESULTS

A. Gender and vowel dependence of SGRs

Figure 2 shows the distributions of Sg1, Sg2, and Sg3 across all of the measurements made from the corpus (no averaging of any kind has been done) as well as the distributions of speaker heights. Table III gives the means and standard deviations for each SGR by speaker as well as the height, age, gender, and total number of measurements that were made of each SGR from each speaker. Note that the female SGR frequencies (660, 1513, and 2426 Hz, on average) are roughly between 11% and 20% higher than male SGR frequencies (554, 1327, and 2179 Hz, on average). The within-speaker COVs are generally less than 5%, with the corresponding 95% confidence intervals ranging from $\bar{x}\cdot(1 \pm 0.005)$ to $\bar{x}\cdot(1 \pm 0.03)$, indicating that repeated within-speaker measurements can be reliably made.

It is worth noting that the (within-gender) across-speaker COVs for SGRs are roughly 5% or 6%, about half the value of the smallest *within-vowel* COVs of formant and fundamental frequencies even though the SGRs were measured across all vowels. This indicates that the SGRs are much less variable than the formants and roughly independent of vowel category. Studies by Wokurek and Madsack (2008, 2011) and Cranen and Boves (1987), however, have reported small but statistically significant differences of some SGRs depending on the vowel. To test the hypothesis that SGRs depend on vowel category, we therefore conducted a 2 ([+low], [−low]) × 2 ([+back], [−back]) repeated measures ANOVA for each gender as well as an additional 2 ([+low], [−low]) × 2 ([+back], [−back]) × 2 (female, male) repeated measures ANOVA for the combined data. The individual SGR measurements were used as
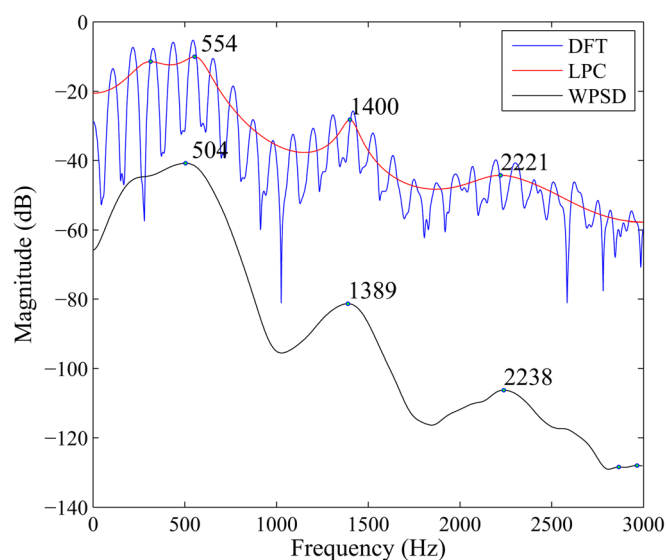

FIG. 1. (Color online) DFT, LPC, and WPSD spectra of a sample accelerometer signal during the vowel [a] produced by speaker 12. Estimated values of Sg1, Sg2, and Sg3 are indicated for both the LPC and WPSD spectra.

J. Acoust. Soc. Am., Vol. 132, No. 4, October 2012

Lulich *et al.*: Subglottal resonances   2595

	Females				Males			
	F0	F1	F2	F3	F0	F1	F2	F3
Vowel	$\bar{x}(s)$	$\bar{x}(s)$	$\bar{x}(s)$	$\bar{x}(s)$	$\bar{x}(s)$	$\bar{x}(s)$	$\bar{x}(s)$	$\bar{x}(s)$
i	215 (40)	379 (60)	2787 (381)	3393 (253)	121 (37)	282 (37)	2251 (154)	3020 (247)
ɪ	209 (41)	656 (137)	2109 (290)	3051 (217)	116 (33)	521 (86)	1771 (211)	2628 (206)
e	198 (41)	685 (229)	2019 (651)	2977 (254)	114 (34)	552 (161)	1671 (522)	2685 (257)
ε	210 (41)	580 (186)	1814 (336)	2913 (255)	119 (34)	465 (141)	1515 (317)	2602 (317)
æ	195 (43)	919 (149)	1667 (227)	2820 (236)	110 (32)	703 (65)	1443 (247)	2552 (243)
a	192 (41)	870 (130)	1584 (239)	2810 (213)	110 (32)	697 (58)	1309 (181)	2535 (268)
ʌ	207 (37)	629 (151)	1252 (402)	3012 (245)	114 (31)	525 (102)	1067 (320)	2609 (268)
o	211 (37)	458 (103)	2020 (809)	3115 (313)	118 (33)	369 (97)	1663 (613)	2724 (360)
ʊ	209 (39)	672 (122)	1775 (233)	2991 (211)	117 (34)	526 (69)	1475 (234)	2597 (282)
u	207 (43)	642 (264)	1444 (209)	2812 (210)	118 (37)	509 (205)	1223 (231)	2552 (343)
aɪ	205 (41)	698 (322)	1794 (288)	2804 (224)	116 (32)	536 (235)	1502 (268)	2540 (353)
aʊ	193 (41)	940 (152)	1678 (122)	2767 (227)	110 (30)	736 (76)	1440 (126)	2472 (234)
ɔɪ	205 (40)	604 (102)	1468 (577)	2989 (253)	113 (31)	520 (72)	1273 (464)	2604 (270)
ɹ	208 (39)	548 (66)	1362 (328)	2433 (507)	116 (30)	463 (39)	1210 (292)	2285 (573)

dependent variables. (In addition, the means, standard deviations, and number of measurements of each SGR across female speakers and across male speakers are given for each vowel separately in Table IV. The [low] and [back] feature values of each vowel are also given.)

When females were considered alone, Sg1 was found to be roughly 15 Hz higher in high vowels than in low vowels [main effect of vowel height: $F(1, 316) = 7.87$, $P = 0.0054$], and Sg2 and Sg3 were roughly 25 and 37 Hz higher in back vowels than in front vowels, respectively [main effect of vowel backness for Sg2: $F(1, 327) = 9.89$, $P = 0.0018$; for Sg3: $F(1, 149) = 4.67$, $P = 0.0324$]. When males were considered alone, the trends were similar for each of the SGRs, although smaller (3, 7, and 20 Hz, respectively) and not statistically significant ($P > 0.18$). When males and females were considered together, the vowel category dependence of each SGR was significant [$F(1, 649) = 5.93$ for Sg1, $F(1, 670) = 9.26$ for Sg2, $F(1, 298) = 4.54$ for Sg3, $P < 0.034$ for all SGRs], and there were no significant interactions between gender and either vowel height or vowel backness ($P > 0.148$).

The main effects of vowel category in our data suggest that the relative frequencies of neighboring SGRs and formants (e.g., Sg2 and F2) affect the SGR frequencies by small but statistically significant amounts, at least for females. Specifically, if F1 or F2 is lower than Sg1 or Sg2, respectively (i.e., if the vowel is [−low] or [+back]), then the SGR is pushed slightly higher in frequency, whereas if the formant is higher than the SGR, then the SGR is pushed slightly lower. This agrees with theoretical predictions of natural frequencies for coupled resonators (Stevens, 1998, pp. 142–148). That only the females showed significant main effects could be due to the fact that female speakers are more likely to have an opening at the posterior end of the glottis throughout the vocal fold vibration cycle (Södersten and Lindestad, 1960), resulting in a greater degree of acoustic coupling between the vocal tract and the subglottal airways and thus a greater amount of coupling between the two resonant systems (Hanson, 1996).

It should be noted that although SGRs appear to vary somewhat depending on the vowel, at least for female speakers, the absolute values of these differences are smaller than the SGR standard deviations, and it is therefore possible that random measurement error is responsible for the effects observed in this study. Further investigation of this phenomenon is warranted. In any event, the small magnitude of this effect suggests that for most practical purposes SGRs can be considered to be constant and independent of vowel category.

B. Correlations among SGRs, F0, and vowel formants

In some applications, it may be useful to know how the SGRs are correlated with each other or with the fundamental and formant frequencies (e.g., Wang *et al.*, 2008; Wang
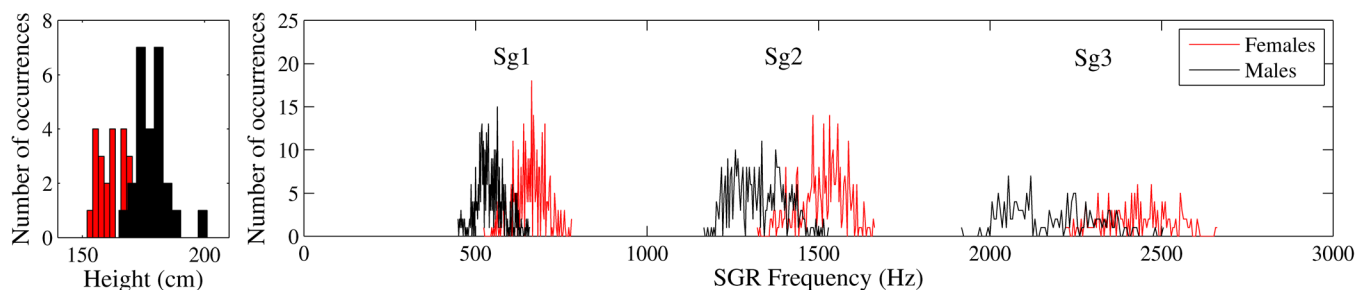

FIG. 2. (Color online) (Left) Distributions of speaker height for males and females. Twenty-five values in each distribution are plotted in 10 bins. (Right) Distributions of Sg1, Sg2, and Sg3 measurements for males and females. Up to 348 measurements of each SGR (for each gender) are plotted in 100 bins each.

TABLE III. Speaker ID, age in months, height in centimeters, number of measurements [Number] made of each SGR, and means [$\bar{x}$] and standard deviations [(s)] of SGR frequencies in Hertz for each speaker. The total number of measurements and the grand means and standard deviations for males and females separately are also given at the bottom. Data from female speakers are in the left half of the table, and data from males are in the right half.

Females			Sg1		Sg2		Sg3		Males			Sg1		Sg2		Sg3	
ID	Age	Height	Number	$\bar{x}$(s)	Number	$\bar{x}$(s)	Number	$\bar{x}$(s)	ID	Age	Height	Number	$\bar{x}$(s)	Number	$\bar{x}$(s)	Number	$\bar{x}$(s)
9	232	157	15	679 (37)	16	1562 (40)	7	2513 (63)	11	244	183	16	532 (43)	18	1269 (32)	7	2096 (38)
10	247	175	15	660 (26)	15	1562 (66)	3	2446 (131)	12	297	180	27	545 (22)	27	1387 (12)	17	2305 (50)
14	258	173	15	673 (19)	15	1530 (27)	10	2519 (49)	13	250	180	25	622 (21)	27	1402 (27)	6	2330 (26)
16	268	155	25	721 (25)	26	1535 (30)	15	2566 (53)	15	243	178	27	520 (36)	27	1254 (29)	17	2054 (55)
18	226	168	9	580 (24)	9	1478 (25)	0	—	17	231	178	9	588 (63)	9	1343 (36)	4	2217 (83)
19	255	163	9	637 (27)	9	1548 (31)	5	2548 (19)	21	239	201	27	531 (32)	27	1256 (23)	12	2106 (39)
20	269	170	9	719 (25)	9	1515 (29)	4	2494 (45)	22	263	173	9	513 (16)	9	1310 (29)	3	2208 (132)
24	279	168	9	641 (18)	9	1488 (15)	1	2336 (—)	23	266	178	9	534 (23)	9	1373 (25)	3	2284 (98)
25	244	168	9	623 (17)	9	1444 (27)	1	2293 (—)	29	267	170	12	565 (12)	12	1491 (26)	3	2430 (48)
26	239	170	8	642 (49)	9	1391 (19)	9	2315 (37)	31	238	165	9	567 (47)	9	1384 (52)	4	2304 (53)
27	261	160	14	590 (33)	15	1417 (50)	7	2277 (38)	38	248	175	9	545 (30)	9	1325 (30)	2	2246 (61)
28	254	170	12	625 (46)	11	1444 (33)	1	2273 (—)	41	239	178	9	543 (14)	9	1216 (10)	6	2050 (34)
32	248	163	11	710 (29)	11	1586 (20)	0	—	43	245	185	16	504 (39)	17	1306 (34)	7	2170 (75)
33	279	173	14	636 (35)	15	1382 (29)	11	2281 (40)	44	224	180	9	529 (33)	9	1226 (20)	8	2039 (44)
35	221	163	8	669 (22)	11	1506 (37)	2	2391 (66)	45	232	173	17	567 (26)	17	1425 (22)	5	2255 (66)
36	259	163	18	678 (22)	18	1540 (25)	16	2444 (39)	49	222	180	9	523 (25)	9	1306 (41)	6	2097 (67)
37	225	152	13	679 (68)	11	1584 (47)	2	2447 (8)	52	238	173	9	529 (22)	9	1283 (25)	4	2051 (7)
40	242	155	14	651 (28)	14	1513 (32)	4	2379 (78)	53	239	183	10	526 (25)	10	1234 (44)	4	2051 (92)
42	222	167	16	669 (25)	18	1531 (42)	4	2428 (38)	57	233	188	14	534 (26)	15	1282 (35)	7	2181 (70)
47	269	157	3	711 (57)	9	1610 (38)	6	2493 (70)	62	236	175	9	556 (41)	9	1430 (29)	4	2393 (22)
50	240	157	15	672 (22)	8	1553 (33)	10	2435 (57)	63	238	174	11	529 (28)	11	1370 (22)	5	2298 (62)
56	230	173	15	640 (42)	15	1436 (24)	6	2363 (25)	64	228	175	11	547 (23)	14	1333 (43)	9	2171 (76)
58	260	160	16	624 (29)	16	1496 (33)	13	2348 (43)	65	261	185	11	492 (30)	11	1244 (36)	3	2129 (105)
59	222	155	13	675 (42)	17	1593 (40)	3	2503 (6)	67	294	180	8	575 (20)	9	1299 (31)	2	2079 (33)
66	226	155	14	662 (19)	15	1534 (30)	12	2428 (46)	68	226	170	16	555 (29)	16	1390 (44)	6	2296 (108)
Overall:			319	**660** (47)	330	**1513** (69)	152	**2426** (101)	Overall:			338	**554** (42)	348	**1327** (77)	154	**2179** (126)

et al., 2009b). Figure 3 gives scatterplots showing the correlations among the SGRs for males and females, and Table V gives the *r* values for each correlation. When both males and females are considered together, the correlations among the SGRs are rather high ($r > 0.75$). Considered separately for males and females, Sg1 is not strongly correlated with Sg2 or Sg3 ($0.27 < r < 0.58$), but Sg2 is strongly correlated with Sg3 ($r > 0.73$). The poor correlations involving Sg1 are presumably due to the effects of the subglottal tissue resonance, which was shown to be in the vicinity of Sg1 (Lulich *et al.*,

2011a) and which may variably affect the Sg1 frequency (without altering the Sg2 and Sg3 frequencies) depending on the mechanical properties of the tissue and the proximity of Sg1 to the tissue resonance. The strong correlation between Sg2 and Sg3 was noted before (Lulich *et al.*, 2011a) and, together with the fact that the mean ratio between Sg3 and Sg2 (1.62) is roughly equal to the expected relation 5/3, is evidence that at frequencies far from the tissue resonance the subglottal system behaves like an equivalent uniform tube closed at the glottal end and open at the distal end.

TABLE IV. Number of measurements [Number], means [$\bar{x}$, in Hertz] and standard deviations [(s), in Hertz] of SGR frequencies for each vowel, separated by gender. The ± value for the features [low] and [back] are also given for each vowel.

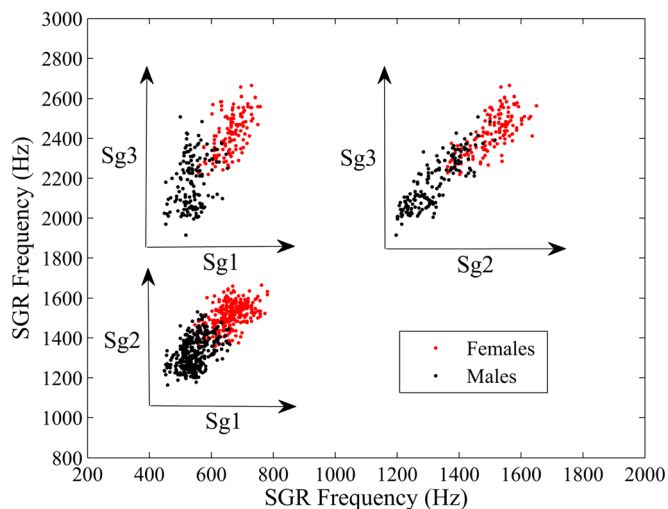|			Females						Males					
			Sg1		Sg2		Sg3		Sg1		Sg2		Sg3	
Vowel	[low]	[back]	Number	$\bar{x}$ (s)	Number	$\bar{x}$(s)	Number	$\bar{x}$(s)	Number	$\bar{x}$(s)	Number	$\bar{x}$(s)	Number	$\bar{x}$(s)
i	−	−	33	664 (53)	41	1516 (73)	19	2405 (104)	38	545 (46)	39	1329 (67)	27	2183 (109)
ɪ	−	−	37	667 (46)	37	1489 (72)	22	2394 (99)	37	549 (43)	37	1327 (87)	20	2160 (130)
ɛ	−	−	30	659 (46)	33	1499 (67)	19	2428 (104)	36	550 (41)	37	1325 (76)	12	2180 (163)
æ	+	−	45	655 (44)	42	1494 (73)	18	2406 (74)	41	541 (48)	42	1313 (72)	16	2153 (128)
a	+	+	44	652 (46)	49	1529 (54)	18	2466 (120)	45	538 (39)	46	1327 (77)	17	2153 (120)
ʌ	+	+	32	645 (44)	32	1533 (66)	13	2435 (112)	34	548 (40)	36	1321 (81)	16	2206 (139)
o	−	+	36	667 (51)	33	1513 (72)	17	2419 (111)	38	541 (44)	38	1332 (83)	14	2182 (102)
ʊ	−	+	32	672 (45)	32	1514 (73)	14	2444 (86)	35	547 (35)	37	1330 (75)	15	2228 (138)
u	−	+	30	667 (45)	31	1531 (62)	12	2462 (90)	34	542 (43)	36	1342 (77)	17	2176 (122)

FIG. 3. (Color online) Scatter plots of Sg1 vs Sg2, Sg1 vs Sg3, and Sg2 vs Sg3 for males and females. The number of data points in each subplot and the correlation coefficients are given in Table V.

TABLE V. r values (and number of data points) for the correlations among SGRs for males and females both separately and combined.

	Males	Females	Combined
Sg1 vs Sg2	0.4585* (337)	0.5195* (292)	0.8026* (629)
Sg1 vs Sg3	0.2781* (146)	0.5727* (131)	0.7572* (277)
Sg2 vs Sg3	0.8474* (154)	0.7389* (145)	0.9133* (299)

None of the SGRs is strongly correlated with F1, F2, or F3, either when males and females are considered together ($|r| < 0.44$) or separately ($|r| < 0.48$). However, when averages of each speaker's SGRs are considered and not separated by gender, somewhat stronger correlations are observed, as shown in Table VI. For females and males considered separately, the SGRs are not strongly correlated with F0 ($r < 0.48$ when all data are considered, $r < 0.7$ when averaged data are considered), but when male and female data are combined, the correlations are somewhat stronger ($0.65 < r < 0.85$), both when non-averaged and averaged data are considered (see Table VI).

C. Height and age dependence of subglottal resonances

Figure 4 shows the relation between the SGRs and speaker height. The solid lines show the relation for quarter-

wavelength resonances of a tube closed at one end and open at the other, whose length is given by the speaker's height, h (in centimeters), divided by a scaling factor, $k_a = 8.802$, using the equation:

$$\text{SgN} = \frac{(2N-1)c}{4h/k_a}. \tag{1}$$

For Sg2 and Sg3, the wave propagation velocity is $c = c_0 = 35\,900$ cm/s, which is the speed of sound in fully saturated air at body temperature. For Sg1, the wave propagation velocity is increased to $c = c_w = 46\,586$ cm/s as a result of the nearby wall tissue resonance (Lulich *et al.*, 2011a).

The value for k_a was obtained by allowing it to vary and minimizing the sum of squared errors (SSE) between the predicted and measured values of Sg2 and Sg3 for all speakers combined (except for Sg3 of speakers 18 and 32 because no Sg3 measurements were obtained). Subsequently, the value of c_w was obtained by minimizing the SSE between predicted and measured Sg1. We also calculated k_a and c_w for males and females separately. These values, and the associated root-mean-square (RMS) errors, ε_{rms}, are given in Table VII.

It is interesting to note that the difference in k_a between males and females indicates a somewhat longer trachea relative to height for males. This is consistent with data reported by Chong *et al.* (2006) for Chinese adults and Munguía-Canales *et al.* (2011) for Mexican adults. In those studies,

TABLE VI. Correlation coefficients between SGRs, F0, and the formants. The values not in parentheses were calculated by including all of the available data, whereas the values in parentheses were calculated based on each speaker's average SGR, F0, and formant frequencies. The averages were made irrespective of vowel category. In general, only when averages were used and the two genders were combined were the correlations somewhat strong (bold type). The number of data points in each correlation was based on the number of SGR measurements made as given in Table III. Thus for the averaged data, there were 25 data points for the Sg1 and Sg2 correlations as well as for male Sg3 correlations, there were 23 points in the female Sg3 correlations, and for the combined-gender correlations, there were 50 or 48 data points, accordingly. For the non-averaged data, the number of data points for each correlation can be found in the bottom line of Table III.

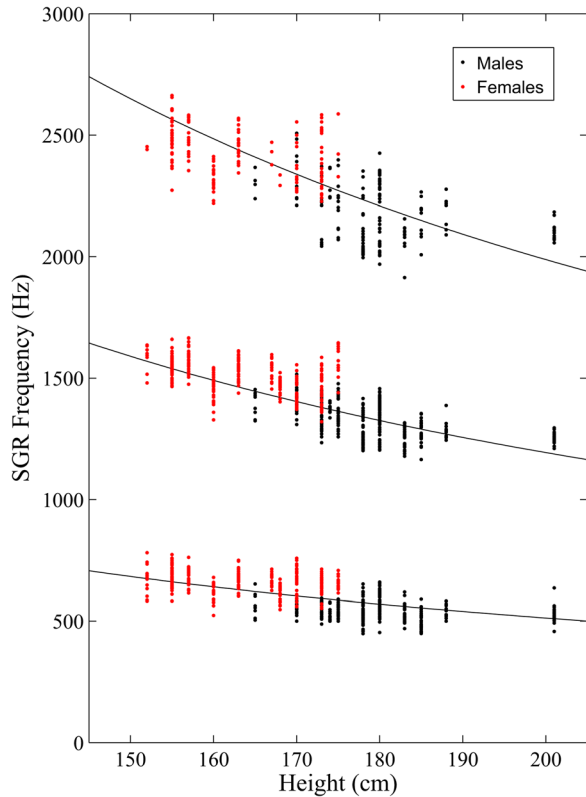|		Sg1	Sg2	Sg3
Females	F0	0.1767* (0.4235*)	0.2057* (0.2686)	0.0752 (0.2191)
	F1	−0.0577 (0.0072)	0.0473 (0.2284)	0.0437 (0.2120)
	F2	0.0054 (−0.0639)	−0.1313* (−0.3360)	−0.2041* (−0.3226)
	F3	−0.0304 (−0.0616)	−0.0238 (−0.0007)	−0.1639* (−0.1082)
Males	F0	0.4725* (0.6947*)	0.4098* (0.4458*)	0.3221* (0.4190*)
	F1	0.0076 (0.3503)	−0.0304 (0.3301)	−0.0389 (0.2330)
	F2	0.0549 (−0.0717)	−0.0171 (0.1585)	−0.0295 (0.2204)
	F3	0.1004 (0.0475)	0.0780 (0.2447)	0.0548* (0.3277)
Combined	F0	0.7260* (**0.8479***)	0.7310* (**0.7674***)	0.6558* (**0.7265***)
	F1	0.2586* (**0.6107***)	0.2762* (**0.6371***)	0.2932* (**0.5833***)
	F2	0.2544* (**0.7031***)	0.1939* (**0.6240***)	0.1124* (**0.5783***)
	F3	0.4205* (**0.6917***)	0.4301* (**0.6912***)	0.3841* (**0.6418***)

FIG. 4. (Color online) Relation between height and SGRs for males and females. The quarter-wavelength resonances of a uniform tube with length equal to h/k_a [see Eq. (1)] are shown by solid lines.

FIG. 5. (Color online) Relation between age and SGRs for males and females.

mean values for height and trachea (or vocal folds-to-carina) length were reported separately for males and females. From these mean values, k_a could be estimated with ratios between male and female k_a values falling between 0.956 and 0.991. In our study, the ratio is $8.720/8.874 = 0.983$.

The reason for a difference in c_w between males and females is less clear. One possibility is that the tissue resonance in females is closer to Sg1 because the smaller (narrower) trachea (Griscom and Wohl, 1986) yields a higher tissue resonance frequency, resulting in a higher wave propagation velocity in the vicinity of Sg1 (Lulich et al., 2011a).

It is also worth noting that there is substantial individual variation in the relationship between height and SGRs and that the correlation between height and SGRs does not appear to be particularly strong within-gender. This has recently been investigated by Arsikere et al. (2012).

In the age range represented in this database (18–24 yr), SGRs do not correlate with age, either for males and females

separately or combined ($|r| < 0.21$) as shown in Fig. 5 and in Table VIII.

IV. DISCUSSION

This paper presents the first large-scale study of SGRs and their dependencies on a number of factors. Fifty speakers (25 males, 25 females) participated in the experiment, which involved simultaneous microphone recordings of speech and accelerometer recordings of subglottal acoustics, resulting in 17 500 utterances in the entire corpus. This study focused on monophthongs in hVd context of which there were 2250 tokens for females and 2250 tokens for males. Of these, Sg1 was measured in 319 female tokens and 338 male tokens; Sg2 was measured in 330 female tokens and 348 male tokens; and Sg3 was measured in 152 female tokens and 154 male tokens. The results of the present study indicate that the SGRs are, on average, in the vicinity of 660, 1513, and 2426 Hz for females and 554, 1327, and 2179 Hz for males. Table IX summarizes the results of this study in comparison with previous studies.

TABLE VII. Values of k_a and c_w, and their associated RMS errors, ε_{rms}, for females and males considered separately and combined. The RMS errors are calculated for each SGR separately, as well as for Sg2 and Sg3 combined.

	k_a	c_w (cm/s)	ε_{rms} (Hz)			
			Sg1	Sg2	Sg3	Sg2+Sg3
Females	8.874	48,431	37.7	75.8	114.3	96.2
Males	8.720	44,324	28.4	59.6	104.4	85.0
Combined	8.802	46,586	46.0	72.6	108.9	92.2

TABLE VIII. Correlation coefficients between the mean SGRs and age. The number of data points for each correlation are 25 for female and male Sg1 and Sg2 correlations and for the male Sg3 correlation, 23 for the female Sg3 correlation, 50 for the combined male-female Sg1 and Sg2 correlations, and 48 for the combined male-female Sg3 correlation.

	Sg1	Sg2	Sg3
Females	0.0968	−0.2004	−0.0981
Males	0.0606*	0.1765*	0.1946
Combined	0.0702	0.0334*	0.0945*

J. Acoust. Soc. Am., Vol. 132, No. 4, October 2012

Lulich et al.: Subglottal resonances   2599

van den Berg (1960) (VDB in Table IX) famously obtained values of the first three SGRs that have never been replicated [see Lulich et al. (2011a) for a possible explanation of these values]. Ishizaka et al. (1976) (IMK), measured the input impedance of the subglottal airways through the tracheostomata of five male Japanese patients. Their Sg1 and Sg2 values are relatively high. However, Ishizaka et al. (1976) noted that "an average ratio of the tracheal lengths between Japanese and Western anatomies" was 0.941. Multiplication of their measured SGRs by this factor yields 602, 1317, and 1976 Hz, which are more in line with other studies, including the present one.

Habib et al. (1994) (HB) measured the subglottal input impedance of seven male and two female subjects through an endotracheal tube as the subjects were undergoing a bronchoscopic alveolar lavage procedure. They reported Sg1 for each speaker and Sg2 for each of the males. Oddly, their subject number 7 was the tallest of the males, and yet his reported SGRs (732 and 1670 Hz) are more in line with our female data than with our male data. For the two females in their study, the reported Sg1 frequencies are abnormally high: 920 and 926 Hz, respectively. One possible explanation for the abnormally high SGRs in their female subjects and male subject number 7 is that for these three subjects the end of the endotracheal tube was advanced further into the trachea than for the other subjects (Habib et al., 1994 commented on precisely this with respect to the two female subjects), resulting in a shortening of the "acoustic length" of the equivalent uniform tube representing the subglottal airways. One way to evaluate this possibility for subject number 7 is to examine the ratio of Sg2/Sg1. Habib et al. (1994) reported values of Sg2/Sg1 between 2.36 and 2.53 for subject numbers 1–6, and 2.28 for subject number 7. According to our data (III), the mean value of Sg2/Sg1 for males is 2.45 and that for females is 2.30, which is close to the value for

subject number 7. Thus it seems that the data of Habib et al. (1994) are in good agreement with our own data, and the discrepancies can be accounted for by the artificial shortening of the trachea based on placement of the endotracheal tube. For this reason, their subject number 7 and their two female subjects were excluded from the calculation of the Sg1 and Sg2 means and standard deviations given in Table IX.

Cranen and Boves (1987) (CB) used miniature pressure transducers passed through the posterior glottal opening to measure subglottal pressure waveforms directly (Cranen and Boves, 1985).

Additional studies of subglottal resonances have made use of non-invasive measurements using various kinds of accelerometers placed against the skin of the neck below the glottis: Hanson (1996) (HA); Harper et al. (2001) (HR); Cheyne (2002) (CH); Chi (2005), Sonderegger (2004), and Chi and Sonderegger (2007) (CS); Jung (2009a) (JE for English data, JK for Korean data); Wokurek and Madsack (2009) (WM); and Gráczi et al., 2011 (GR). Wokurek and Madsack (2011) reported only means and standard deviations averaged across both genders: 586 and 69 Hz for Sg1 and 1325 and 122 Hz for Sg2.

Three studies have estimated SGRs on the basis of their effects on vowel or aspirated sound spectra, usually by noting the frequency of "extra" spectral peaks, or the frequency at which a formant transition becomes discontinuous. Fant et al. (1972) (FA) reported peaks in the vicinity of 500, 1000, and 1500–1700 Hz. Except for the peak around 1000 (which could be the so-called "chest resonance") (cf. Wokurek and Madsack, 2011), these peaks could be due to Sg1 and Sg2. Klatt and Klatt (1990) (KK) identified peaks near 700, 1700, and 2500 Hz for females and near 1500 and 2200 Hz for males (though with considerable variability). Both Fant et al. (1972) and Klatt and Klatt (1990) made their measurements in aspirated sound segments, during which

TABLE IX. Means [$\bar{x}$] and standard deviations [(s)] (in Hertz) of the SGRs for males and females in the present study and in previous studies. The results of previous studies are in good agreement with our own.

		Males			Females		
	Method	Sg1 $\bar{x}(s)$	Sg2 $\bar{x}(s)$	Sg3 $\bar{x}(s)$	Sg1 $\bar{x}(s)$	Sg2 $\bar{x}(s)$	Sg3 $\bar{x}(s)$
Present study	Accelerometer	554 (42)	1327 (77)	2179 (126)	660 (47)	1513 (69)	2426 (101)
VDB	Input impedance	314	890	1390	-	-	-
IMK	Input impedance	640	1400	2100	-	-	-
HB	Input impedance	586 (53)	1426 (96)	-	-	-	-
CB	Miniature pressure transducer	510 (26)	1355 (61)	2290 (200)	-	-	-
HA	Accelerometer	-	-	-	625	1540	2110
HR	Accelerometer	644	1220	2059	-	-	-
CH	Accelerometer	566	1400	-	628	1450	-
CS	Accelerometer	606	1359	-	680	1499	-
JE	Accelerometer	541	-	-	651	-	-
JK	Accelerometer	-	1322	-	-	1545	-
WM	Accelerometer	519 (12)	1368 (26)	-	568 (21)	1554 (35)	-
GR	Accelerometer	581	1319	-	562	1484	-
FA	Spectrum	500	1500-1700	-	-	-	-
KK	Spectrum	-	1500	2200	700	1700	2500
LU	Formant discontinuity	-	1486	-	-	-	-

the glottal area is larger than during vowels. An increase in the glottal area (either via a posterior gap or an abducted glottis) leads to increased coupling between the subglottal and supraglottal airways, with a concomitant increase in SGR frequencies (Lulich, 2010), and this may explain the relatively high values reported by these two studies. Lulich (2010) (LU), measuring Sg2 as the frequency at which an F2 transition became discontinuous, reported a value of 1486 Hz for one male speaker.

Within-speaker standard deviations of SGRs have been reported by Chi and Sonderegger (2007), Chi (2005), Sonderegger (2004), Jung (2009a), and Gráczi et al. (2011). These are summarized in Table X.

Overall, the results of previous studies are in good agreement with our own, regardless of the methodology used to measure the SGRs. Among males, Sg1 measured by HR and CS and Sg2 measured by HB, are relatively high in frequency compared with the remaining studies and our own (the high values of IMK, FA, and KK can be accounted for by other factors, as described above). Among females, Sg1 measured by WM and GR, Sg2 measured by CH, and Sg3 measured by HA, are relatively low in frequency. WM report much smaller standard deviations than other studies, including the present one, suggesting that their subject population was relatively uniform.

Wokurek and Madsack (2008, 2011) and Cranen and Boves (1987) reported small but significant differences in Sg1 and Sg2 depending on the vowel. Wokurek and Madsack (2008) found that Sg2 was on average 4 Hz higher in [i] and [u] than in [a], and Wokurek and Madsack (2011) found that Sg1 and Sg2 could be as much as 20 Hz higher in long vowels than in homorganic short vowels. Cranen and Boves (1987) found that Sg1 was approximately 20 Hz higher in the high vowels [i] and [u] than in the low vowel [a]. In the present study, we found that for females, Sg1 was approximately 15 Hz higher in high vowels than in low vowels in good agreement with Cranen and Boves (1987), and Sg2 was approximately 25 Hz higher in back vowels than in front vowels in contrast to Wokurek and Madsack (2008). For males, we did not find any significant vowel dependence of the SGRs. As was noted in Sec. III A, a significant dependence of SGRs on vowel category may require a relatively large posterior glottal opening, which is common in females but less common in males. The fact that Cranen and Boves (1987) had passed pressure transducers through the posterior glottal opening of their subjects may have led to an increase in its cross-sectional area, therefore revealing the vowel dependence in their male speakers.

TABLE X. Range [min-max] of within-speaker standard deviations [s] (in Hertz) for Sg1, Sg2, and Sg3 measurements in the present study and in previous studies. The results of previous studies are in good agreement with our own.

	Sg1 s	Sg2 s	Sg3 s
Present study	12–68	10–66	6–132
CS	-	12–30	-
JE	40–55	-	-
JK	-	40–70	-
GR	30–73	50–80	-

V. SUMMARY AND CONCLUSION

In this paper, we have presented the first large-scale study of subglottal resonances. Overall, our data are in good agreement with previous studies that used relatively small numbers of subjects and restricted sets of vowels. On average, female SGRs are between 11% and 20% higher than male SGRs, and the COVs across males and females are on the order of 5%. Although the SGRs are roughly constant, there is a statistically significant tendency for female SGRs to vary with vowel categories. These variations are hypothesized to be due to increased coupling with the vocal tract via the posterior glottal opening, and the effects of the coupled subglottal and vocal tract resonators on the natural frequencies of the combined system. Although the variations are statistically significant, they are also small compared with the possible measurement error and with the standard deviations of the SGR measurements. For practical purposes, the SGRs can therefore be considered constant.

Although Sg2 and Sg3 are strongly correlated with each other, neither is strongly correlated with Sg1, presumably because of the wall tissue resonance present in the vicinity of Sg1. None of the SGRs is strongly correlated with any of the formants or the fundamental frequency.

All three SGRs have been found to be successfully described as quarter-wavelength resonances of an equivalent uniform tube the length of which is related to the height of the speaker, and the parameters of this model (k_a and c_w) are similar for both males and females. There is no correlation between the SGRs and speaker age within this age group (18–24 yr).

The results of this study provide normative data on SGRs that can be used in future studies of the effects of subglottal acoustics on various aspects of speech production, perception, and technology.

ACKNOWLEDGMENT

Arsikere, H., Leung, G. K. F., Lulich, S. M., and Alwan, A. (2012). "Automatic estimation of the first three subglottal resonances from adults' speech signals with application to speaker height estimation," Speech Commun. http://dx.doi.org/10.1016/j.specom.2012.06.004.

Boersma, P., and Weenink, D. (2010). "PRAAT: Doing phonetics by computer (version 5) [computer program]," p. 42.

Cheyne, H. A. (2002). "Estimating glottal voicing source characteristics by measuring and modeling the acceleration of the skin on the neck," Ph.D. thesis, MIT, Cambridge, MA.

Chi, X. (2005). "The quantal effect of sub-glottal resonance on vowel formant," RQE Report, MIT, pp. 1–31.

Chi, X., and Sonderegger, M. (2007). "Subglottal coupling and its influence on vowel formants," J. Acoust. Soc. Am. 122, 1735–1745.

Chong, D. Y. C., Greenland, K. B., Tan, S. T., Irwin, M. G., and Hung, C. T. (2006). "The clinical implication of the vocal cords-carina distance in anaesthetized Chinese adults during orotracheal intubation," Br. J. Anaesth. 97, 489–495.

Cranen, B., and Boves, L. (1985). "Pressure measurements during speech production using semiconductor miniature pressure transducers: Impact on models for speech production," J. Acoust. Soc. Am. 77, 1543–1551.

Cranen, B., and Boves, L. (1987). "On subglottal formant analysis," J. Acoust. Soc. Am. 81, 734–746.

Csapó, T. G., Bárkányi, Z., Gráczi, T. E., Bőhm, T., and Lulich, S. M. (2009). "Relation of formants and subglottal resonances in Hungarian vowels," in Proceedings of Interspeech, pp. 484–487.

J. Acoust. Soc. Am., Vol. 132, No. 4, October 2012

Lulich et al.: Subglottal resonances   2601

Csapó, T. G., and Németh, G. (**2009**). "Mássalhangzó-magánhangzó kapcso-latok automatikus osztályozása szubglottális rezonanciák alapján (Auto-matic consonant-vowel classification based on subglottal resonances)," in *Magyar Számtógépes Nyelvészeti Konferencia (Sixth Conference on Hungarian Computational Linguistics)*, pp. 226–237.

Fant, G. (**1960**). *Acoustic Theory of Speech Production: With Calculations Based on X-Ray Studies of Russian Articulations* (Mouton, The Hague), pp. 1–328.

Fant, G., Ishizaka, K., Lindqvist, J., and Sundberg, J. (**1972**). "Subglottal formants," STL- QPSR **1**, 1–12.

Fredberg, J. J. (**1978**). "A modal perspective on lung response," J. Acoust. Soc. Am. **63**, 962–966.

Fredberg, J. J., and Hoenig, A. (**1978**). "Mechanical response of the lungs at high frequencies," J. Biomech. Eng. **100**, 57–66.

Fredberg, J. J., and Moore, J. A. (**1978**). "The distributed response of complex branching duct networks," J. Acoust. Soc. Am. **63**, 954–961.

Gráczi, T. E., Lulich, S. M., Csapó, T. G., and Beke, A. (**2011**). "Context and speaker dependency in the relation of vowel formants and subglottal resonances—evidence from Hungarian," in *Proceedings of Interspeech*, pp. 1901–1904.

Griscom, N. T., and Wohl, M.E. (**1986**). "Dimensions of the growing trachea related to age and gender," Am. J. Roentgenol. **146**, 233–237.

Habib, R. H., Chalker, R. B., Suki, B., and Jackson, A. C. (**1994**). "Airway geometry and wall mechanical properties estimated from subglottal input impedance in humans," J. Appl. Physiol. **77**, 441–451.

Hanson, H. M. (**1996**). "Measurements of subglottal resonances and their influence on vowel spectra," J. Acoust. Soc. Am. **100**, 2656–2656.

Hanson, H. M., and Stevens, K. N. (**1995**). "Sub-glottal resonances in female speakers and their effect on vowel spectra," in *Proceedings of the XIIIth International Congress of Phonetic Sciences*, Stockholm, Vol. 3, pp. 182–185.

Harper, V. P., Kraman, S. S., Pasterkamp, H., and Wodicka, G. R. (**2001**). "An acoustic model of the respiratory tract," IEEE Trans. Biomed. Eng. **48**, 543–550.

Hillenbrand, J., Getty, L. A., Clark, M. J., and Wheeler, K. (**1995**). "Acoustic characteristics of American English vowels," J. Acoust. Soc. Am. **97**, 3099–3111.

Ishizaka, K., Matsudaira, M., and Kaneko, T. (**1976**). "Input acoustic-impedance measurement of the subglottal system," J. Acoust. Soc. Am. **60**, 190–197.

Jung, Y. (**2009a**). "Acoustic articulatory evidence for quantal vowel categories: The features [low] and [back]," Ph.D. thesis, MIT, Cambridge, MA, pp. 1–142.

Jung, Y. (**2009b**). "Subglottal effects on the vowels across language: Preliminary study on Korean," J. Acoust. Soc. Am. **125**, 2638.

Klatt, D. H., and Klatt, L. C. (**1990**). "Analysis, synthesis, and perception of voice quality variations among female and male talkers," J. Acoust. Soc. Am. **87**, 820–857.

Ladefoged, P., and Maddieson, I. (**1996**). *The Sounds of the World's Languages* (Wiley-Blackwell, New York), pp. 1–425.

Lee, S., Potamianos, A., and Narayanan, S. (**1999**). "Acoustics of children's speech: Developmental changes of temporal and spectral parameters," J. Acoust. Soc. Am. **105**, 1455–1468.

Liberman, A. (**1996**). *Speech: A Special Code* (The MIT Press, Cambridge, MA), pp. 1–458.

Lulich, S. M. (**2010**). "Subglottal resonances and distinctive features," J. Phonetics **38**, 20–32.

Lulich, S. M., Alwan, A., Arsikere, H., Morton, J. R., and Sommers, M. S. (**2011a**). "Resonances and wave propagation velocity in the subglottal airways," J. Acoust. Soc. Am. **130**, 2108–2115.

Lulich, S. M., Arsikere, H., Morton, J. R., Leung, G. K. F., Alwan, A., and Sommers, M. S. (**2011b**). "Analysis and automatic estimation of children's subglottal resonances," in *Proceedings of Interspeech*, pp. 2817–2820.

Lulich, S. M., Bachrach, A., and Malyska, N. (**2007**). "A role for the second subglottal resonance in lexical access," J. Acoust. Soc. Am. **122**, 2320–2327.

Lulich, S. M., and Chen, N. F. (**2009**). "Automatic classification of consonant-vowel transitions based on subglottal resonances and second formant frequencies," Proc. Meet. Acoust. **6**, 060005.

Madsack, A., Lulich, S. M., Wokurek, W., and Dogil, G. (**2008**). "Subglottal resonances and vowel formant variability: A case study of High German monophthongs and Swabian diphthongs," in *Proceedings of LabPhon11*, pp. 91–92.

Mansfield, J. P. (**1996**). "Theory and application of acoustic reflectometry in the human body," Ph.D. thesis, Purdue University, Indiana, pp. 1–208.

Munguía-Canales, D. A., Ruiz-Flores, J., Vargas-Mendoza, G. K., Morales-Gómez, J., Méndez-Ramirez, I., and Murata, C. (**2011**). "Tracheal dimensions in the Mexican population," Cir. Cir. **79**, 465–510.

Nelson, D. (**1997**). "Correlation based speech formant recovery," in *Proceedings of ICASSP*, pp. 1643–1646.

Peterson, G. E., and Barney, H. L. (**1952**). "Control methods used in a study of the vowels," J. Acoust. Soc. Am. **24**, 369–381.

Sjölander, K., and Beskow, J. (**2000**). "WAVESURFER—an open source speech tool," in *Proceedings of ICSLP*, Vol. 4, pp. 464–467.

Södersten, M., and Lindestad, P.-A. (**1960**). "Glottal closure and perceived breathiness during phonation in normally speaking subjects," J. Speech Hear. Res. **33**, 601–611.

Sonderegger, M. (**2004**). "Sublottal coupling and vowel space: An investigation in quantal theory," undergraduate thesis, MIT, Cambridge, MA, pp. 1–73.

Stevens, K. N. (**1998**). *Acoustic Phonetics* (MIT Press, Cambridge, MA), pp. 1–607.

Suki, B., Habib, R. H., and Jackson, A. C. (**1993**). "Wave propagation, input impedance, and wall mechanics of the calf trachea from 16 to 1,600 Hz," J. Appl. Physiol. **75**, 2755–2766.

Talkin, D. (**1995**). "A robust algorithm for pitch tracking (RAPT)" in *Speech Coding and Synthesis*, edited by W. B. Kleijn and K. K. Paliwal (Elsevier Science, New York), pp. 495–518.

Titze, I. R. (**2006**). *The Myoelastic-Aerodynamic Theory of Phonation* (National Center for Voice and Speech, Denver, CO), pp. 1–424.

Titze, I. R. (**2008**). "Nonlinear source-filter coupling in phonation: Theory," J. Acoust. Soc. Am. **123**, 2733–2749.

Titze, I. R., Riede, T., and Popolo, P. (**2008**). "Nonlinear source-filter coupling in phonation: Vocal exercises," J. Acoust. Soc. Am. **123**, 1902–1915.

Umesh, S., Cohen, L., Marinovic, N., and Nelson, D. J. (**1999**). "Scale transform in speech analysis," IEEE Trans. Speech Audio Process. **7**, 40–45.

van den Berg, J. (**1960**). "An electrical analogue of the trachea, lungs and tissues," Acta Physiol. Pharmacol. Neerl. **9**, 361–385.

Wang, S., Alwan, A., and Lulich, S. M. (**2008**). "Speaker normalization based on subglottal resonances," in *Proceedings of ICASSP*, pp. 4277–4280.

Wang, S., Lee, Y.-H., and Alwan, A. (**2009a**). "Bark-shift based nonlinear speaker normalization using the second subglottal resonance," in *Proceedings of Interspeech*, 1619–1622.

Wang, S., Lulich, S. M., and Alwan, A. (**2009b**). "Automatic detection of the second subglottal resonance and its application to speaker normalization," J. Acoust. Soc. Am. **126**, 3268–3277.

Wokurek, W., and Madsack, A. (**2008**). "Messung subglottaler resonanzen mit beschleunigungssensoren (Measurement of subglottal resonances with accelerometers)," Fortschr. Akustik-Proc. DAGA 125–126.

Wokurek, W., and Madsack, A. (**2009**). "Comparison of manual and automated estimates of subglottal resonances," in *Proceedings of Interspeech*, 1671–1674.

Wokurek, W., and Madsack, A. (**2011**). "Accelerometer sensor based estimates of subglottal resonances: Short vs. long vowels," in *Proceedings of Interspeech*, pp. 2821–2824.

Zañartu, M., Mehta, D. D., Ho, J. C., Wodicka, G. R., and Hillman, R. E. (**2011**). "Observation and analysis of *in vivo* vocal fold tissue instabilities produced by nonlinear source-filter coupling: A case study," J. Acoust. Soc. Am. **129**, 326–339.

Zañartu, M., Mongeau, L., and Wodicka, G. R. (**2007**). "Influence of acoustic loading on an effective single mass model of the vocal folds," J. Acoust. Soc. Am. **121**, 1119–1129.

Zhang, Z., Neubauer, J., and Berry, D. A. (**2006**). "The influence of subglottal acoustics on laboratory models of phonation," J. Acoust. Soc. Am. **120**, 1558–1569.