Estimation of minimum oral tract constriction area in sibilant fricatives from aerodynamic data

Y. Fujiso
GIPSA-Lab, UMR CNRS 5216, Grenoble University, 11 rue des Mathématiques (BP46), 38402 Grenoble, France
yo.fujiso@grenoble-inp.fr

K. Nozaki
Osaka University Dental Hospital, 1-8 Yamadaoka, Suita, Osaka 565-0871, Japan
knozaki@dent.osaka-u.ac.jp

A. Van Hirtum[a)]
GIPSA-Lab, UMR CNRS 5216, Grenoble University, 11 rue des Mathématiques (BP46), 38402 Grenoble, France
annemie.vanhirtum@grenoble-inp.fr

Abstract: Speech screening of sibilant fricative phonemes is an important tool for oral health care. Nevertheless, screening as a function of quantitative geometrical markers is mostly limited to teeth features whereas the minimum area of the narrowed air passage upstream from the tooth is known to be a key production feature. The minimum area is estimated from non-invasive aerodynamic measurements using a laminar flow model. The influence of viscid flow losses on the area estimation is shown to be negligible. Current data suggest that speech screening is most effective for phoneme /s/, which supports common practice in oral health care.

© 2015 Acoustical Society of America

[AL]
Date Received: March 1, 2015 **Date Accepted:** May 27, 2015

1. Introduction

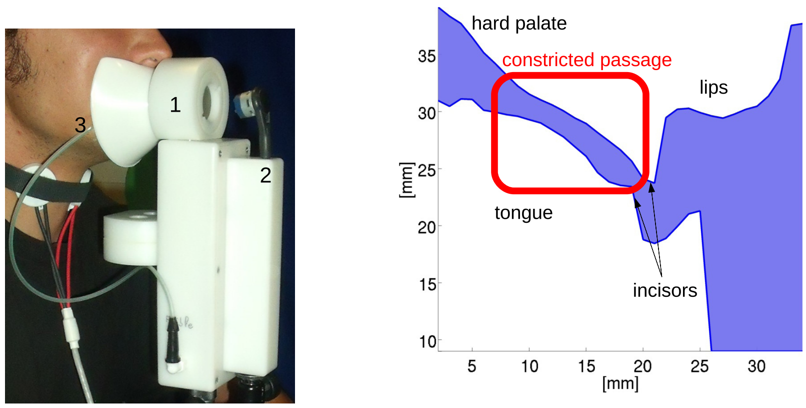
Health care practitioners dealing with health problems related to speech articulators (mouth, face, jaws, or teeth) often rely on speech screening at some stage during diagnostics, treatment or follow-up.[1–3] In dental practice, speech screening is mostly limited to the impact of geometrical teeth features on sibilant fricative production,[4,5] whereas a speech disorder following oral, maxillofacial, or orthognathic surgery is considered a secondary byproduct despite its severe impact on everyday oral communication skills.[1,6] In general, articulatory features related to the oral tract shape are not accounted for in clinical speech screening. Nevertheless, from speech production studies, it is known that the minimum oral tract constriction area A_c upstream from the teeth is directly related to the physical mechanism underlying sibilant fricative production. Indeed, a severely narrowed airway passage between the tongue and hard palate, i.e., sibilant groove illustrated in Fig. 1(b), is needed in order to create a turbulent jet, which will interact with downstream incisors resulting in sibilant fricative noise.[7] Moreover, the presence of a severe constriction is essential for sibilant fricative production, whereas the precise location of the constricted portion has far less influence on the perceived sibilant sound.[8]

Therefore, accurate and straightforward estimation of the minimum constriction area A_c provides a potential geometrical quantitative marker for integration in clinical speech screening procedures using sibilant fricatives. In the following, estimation of the minimum constriction area A_c from aerodynamic quantities is assessed. Aerodynamic data are measured non-invasively on seated human speakers without the need for expensive equipment compared to more commonly applied imaging techniques and with minimal discomfort for the speaker. As a reference, additional data for one of the speakers are obtained using cone-beam computed tomography (CT). Validation of the estimated minimum area is assessed by comparison with data reported in literature.

[a)]Author to whom correspondence should be addressed.

 © 2015 Acoustical Society of America

Redistribution subject to ASA license or copyright; see http://acousticalsociety.org/content/terms. Download to IP: 147.171.69.213 On: Wed, 08 Jul 2015 12:36:24



(a) (b)

Fig. 1. (Color online) Data acquisition: (a) EVA system capturing (1) oral airflow, (2) acoustic signal, and (3) intra-oral pressure. (b) Geometry of the air passage extracted from the CT data in the mid-sagittal plane for phoneme /s/.

2. Method

2.1 Aerodynamic data

Aerodynamic and acoustic measurements were conducted on two seated male adult subjects, designated as "YF" (native French speaker) and "KN" (native Japanese speaker). Subject "KN" pronounced several segments of sibilant /s/; subject "YF" pronounced several segments of sibilant /s/ and fricative /f/ for comparison. Three different loudness levels were assessed, namely "soft," "medium," and "loud." For each combination of speaker, phoneme, and loudness level more than 5 repetitions were acquired using the EVA (Evaluation Vocale Assistée, assisted vocal evaluation) system.[9] The EVA system, illustrated in Fig. 1(a), was used for simultaneous measurement of normalized radiated acoustic pressure (sampling frequency 25 kHz), oral airflow (U, sampling frequency 6250 Hz) and intra-oral pressure (ΔP, sampling frequency 6250 Hz). The EVA system was calibrated for the range of aerodynamic values against the following reference sensors: Brüel & Kjaer microphone (B&K 4192), volume flow meter (TSI 4000) and pressure sensor (Endevco 8507C). The positioning of the oral airflow meter and microphone of the EVA system is shown in Fig. 1(a). The intra-oral pressure was measured using a tube of outer diameter 5 mm and length 30 cm which was connected to the pressure sensor. The tube was inserted at the mouth corner and placed upstream from the minimum constriction between the hard palate and the tongue.

An example of aerodynamic and acoustic data measured with the EVA system during sibilant /s/ utterances is presented in Fig. 2. Each individual sibilant utterance was automatically detected from the measured oral airflow following a two-step threshold procedure. In a first step, single sibilant events were detected from positive values since negative values correspond to inspiratory breathing efforts in between sibilant utterances. In a second step, the spectrally quasi-steady noise portion of the spectrogram of the acoustic signal was detected using an objective threshold criterion based on the maximum oral airflow U for each individual sibilant event [$>0.6\max(U)$]. The mean duration of the selected portion varied from 0.3 s up to 2 s.

2.2 CT data

A cone-beam CT scan (CB MercuRay, 512 slices of 512×512 pixels with accuracy ± 0.1 mm) was made for the single seated adult male subject "KN" while uttering phoneme /s/ at "medium" loudness level (≈ 21 L/min), so that the articulators positions (teeth, tongue, lips) corresponded to a sibilant /s/ sound articulation. The imaging process lasted about 10 s. The oral cavity volume and its shape were reconstructed using a marching cube method.[10] The maximal error yielded ± 0.2 mm. A center slice of the oral tract volume is illustrated in Fig. 1(b). The minimum area within the constricted passage of length $L \simeq 22$ mm yielded 3.5 mm^2 corresponding to a minimum hydraulic diameter $D_c = 2.1$ mm. Note that a length of 22 mm corresponds to the upper end of the physiological range (10 mm up to 30 mm).

Redistribution subject to ASA license or copyright; see http://acousticalsociety.org/content/terms. Download to IP: 147.171.69.213 On: Wed, 08 Jul 2015 12:36:24

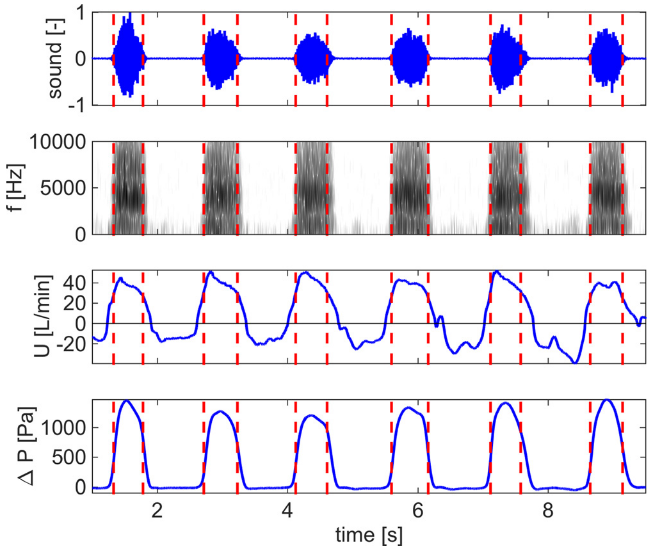


Fig. 2. (Color online) Illustration of quantities measured by EVA-system [normalized acoustic signal (sound), spectrogram of acoustic signal, oral flow rate U and intra-oral pressure ΔP] during consecutive /s/ utterances (subject "KN," "medium" loudness level). Vertical dashed lines indicate the analyzed signal portion.

2.3 Estimation of minimum constriction area from aerodynamic data

The production of sibilant fricatives is governed by three physical quantities: minimum area of the narrowed airway passage A_c, quasi-static pressure drop across the constriction ΔP and oral airflow U.[7] Typical values of Mach number [$M^2 \sim O(10^{-4})$], Reynolds number [$Re \sim O(10^3)$], Strouhal number [$Sr \sim O(10^{-2})$], and mean aspect ratio [$O(10^{-1})$] allow one to model the flow as one-dimensional, inviscid, incompressible, laminar, and quasi-steady.[7,11] The relationship between the quantities characterizing sibilant production (A_c, ΔP, and U) is then derived from the inviscid Euler equation:

$$\Delta P = \frac{\rho U^2}{2c_s^2}\left(\frac{1}{A_c^2} - \frac{1}{A_0^2}\right), \tag{1}$$

with air density $\rho = 1.2\,\mathrm{kg/m^3}$, area upstream from the constricted portion A_0 and *ad hoc* constant $c_s > 1$ defining the position of flow separation along the diverging portion of the constriction as $c_s \times A_c$.[11] Further assuming $A_0 \gg A_c$ and denoting $k = c_s^2$ results in the following relationship:

$$A_c = \frac{U}{k\sqrt{\frac{2\Delta P}{\rho}}}. \tag{2}$$

Equation (2) has the same form as the so-called orifice equation which was empirically derived by Warren and DuBois[12] to provide a non-invasive estimation of the minimum constricted area A_c when the pressure drop ΔP is approximated by the pressure upstream from the constriction, i.e., the intra-oral pressure. In the empirical orifice equation the *ad hoc* constant $k \le 1$ combines inlet and viscous losses. In the current work, the constriction inlet is not sharp, so that considering first $k \le 1$ for losses in the orifice Eq. (2) and second $k \ge 1$ for flow separation Eq. (1), $k = 1$ seems a good compromise. With $k = 1$, Eq. (2) reduces to

$$A_c = \frac{U}{\sqrt{\frac{2\Delta P}{\rho}}} \tag{3}$$

from which the hydraulic diameter of the minimum constricted area yields $D_c = 2\sqrt{A_c/\pi}$. For the narrowed airflow passage characterizing sibilant sound production, the assumption of a circular cross-section within the narrowed passage is motivated from imaging studies[8,13,14] so that when accounting for viscous losses the pressure drop is proportional to volume flow rate U at a rate determined by $L(\nu\rho 8\pi/A_c^2)$ with L denoting the length of the constricted passage and kinematic viscosity of air $\nu = 1.5 \times 10^{-5}\ \mathrm{m^2/s}$.[15] The expression for the minimum area becomes then[15]

Redistribution subject to ASA license or copyright; see http://acousticalsociety.org/content/terms. Download to IP: 147.171.69.213 On: Wed, 08 Jul 2015 12:36:24

$$A_c = \sqrt{\frac{U^2 + 16\pi\nu UL}{\frac{2\Delta P}{\rho}}}, \tag{4}$$

where the right-hand term proportional to U^2 corresponds to Eq. (3) and the right-hand term proportional to U expresses the influence of viscosity. Since Eqs. (3) and (4) are derived theoretically for laminar flow, their use for sibilant sound production relies on the assumption that the flow, up to the point of separation, remains laminar, while a turbulent jet develops downstream from flow separation. Consequently, Eq. (3) (without viscous losses) and Eq. (4) (with viscous losses) provide the sought expression of minimum area A_c (or associated hydraulic diameter D_c) as a function of pressure drop ΔP and volume flow rate U.

3. Results and discussion

Figure 3 illustrates the variability of the measured quantities (ΔP and oral flow rate U) for each of the assessed loudness levels ("soft," "medium," and "loud"), subject ("KN" or "YF") and phoneme (/s/ or /f/). An increased loudness level corresponds to an increased oral flow rate U. Figure 3 shows that the resulting variation (standard variation σ) of measured aerodynamic values ΔP and U is smaller than 10% of the mean value (denoted with an over-line bar) for all assessed loudness levels, subjects, and phonemes. The variation increases with U, and thus with loudness level, for both /s/ and /f/. The repeatability for /f/ is less than for /s/, in particular, the variation in U associated with /f/ is larger than that observed for /s/. In addition, the parameter space spanned by the mean values $\overline{U}$ and $\overline{\Delta P}$ for different loudness conditions is smaller for /f/ than for /s/ so that the use of different loudness conditions is more effective, in terms of aerodynamic variability, for /s/ than for /f/. Therefore, from the current data, sibilant fricative /s/ seems a better candidate to be included in a health care protocol than fricative /f/ based on repeatability and extent of the parameter space associated with aerodynamic observations of intra-oral pressure ΔP and oral flow rate U. This finding supports the use of phoneme /s/ in dental health care.[3,5,8]

The hydraulic diameter D_c is estimated from the overall mean values of intra-oral pressure $\overline{\Delta P}$ and oral flow rates $\overline{U}$ shown in Fig. 3 following Eq. (3). An upper limit for the standard deviation of D_c, σ_{Dc}, is obtained from ($\overline{U}$, σ_U) and ($\overline{\Delta P}$, $\sigma_{\Delta P}$) following the logarithmic differentiation method as[16]

$$\frac{\sigma_{D_c}}{D_c} \leq \frac{1}{2}\left|\frac{\sigma_U}{\overline{U}}\right| + \frac{1}{4}\left|\frac{\sigma_{\Delta P}}{\overline{\Delta P}}\right|. \tag{5}$$

The resulting estimates of the minimum hydraulic diameter D_c and the variation σ_{Dc} for each phoneme (/s/ and /f/), each loudness level (or volume flow rate) and subject ("KN" or "YF") are illustrated in Fig. 4. The estimated diameters range from 2.5 mm up to 5 mm for phoneme /s/ and from 5.5 mm up to 6.2 mm for phoneme /f/. Estimated values are well in the range of values reported in literature based on different imaging techniques (x-ray, magnetic resonance imaging or ultrasound)[7,13,14,17,18] for both phonemes. Because of the larger standard deviation of measured aerodynamic quantities for /f/ than for /s/ (Fig. 3), the standard deviation of the estimated minimum diameter is smaller for /s/ than for /f/. Consequently, the minimum constriction area can be straightforwardly and accurately estimated from aerodynamical measurements

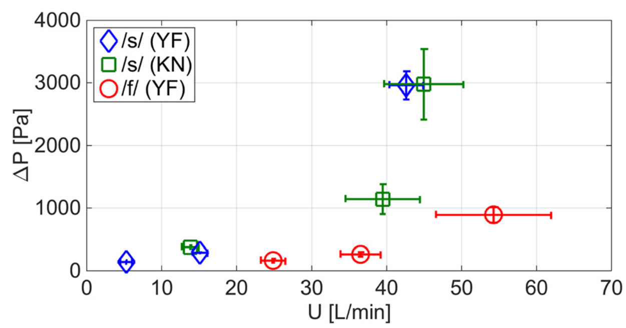


Fig. 3. (Color online) Overall mean $\overline{U}$ and $\overline{\Delta P}$ (symbols) and overall variation σ_U and $\sigma_{\Delta P}$ (horizontal and vertical whiskers) of intra-oral pressure ΔP and oral flow rates U for all assessed loudness levels ("soft," "medium," and "loud") for subjects "YF" (/s/ and /f/) and "KN" (/s/).

Redistribution subject to ASA license or copyright; see http://acousticalsociety.org/content/terms. Download to IP: 147.171.69.213 On: Wed, 08 Jul 2015 12:36:24

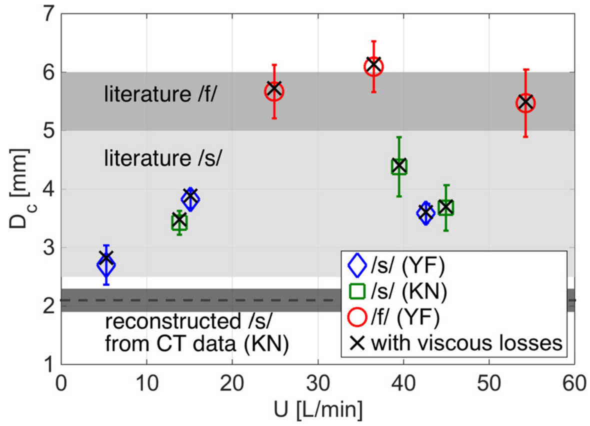


Fig. 4. (Color online) Minimum constriction hydraulic diameters D_c as a function of oral flow rate U for subjects "YF" (/s/ and /f/) and "KN" (/s/): using Eq. (3) (◇, □, ○) and using Eq. (4) (×). Vertical whiskers indicate the overall variation σ_{Dc}. The corresponding range reported in Refs. 13, 14, 17, and 18 is indicated for phoneme /s/ (light grey shade) and phoneme /f/ (dark grey shade). Oral cavity reconstruction from CT data (subject "KN," phoneme /s/, 21 L/min) yields $D_c = 2.1 \pm 0.2$ mm (horizontal dashed line and black shade).

during phoneme /s/ utterances compared to values obtained with imaging methods, which often rely on expensive equipment and speakers in supine instead of seated position. Note, that the presence of the tube required to measure the pressure drop does not influence the value of the minimum area.

Inviscid flow is supposed in Eq. (3), so that following Eq. (1) the oral flow rate U is proportional to the square root of the measured intra-oral pressure ΔP. However, for small volume flow rates through a severely constricted area, viscosity is known to affect the flow[15] and in this case a linear relationship between oral flow rate U and intra-oral pressure ΔP occurs for which the proportionality constant depends on the cross-section shape of the constricted area as well as on the streamwise extent of the constricted portion.[15] The pressure-flow measurements shown in Fig. 3, suggest such a change of slope for linear segments connecting respectively "soft-medium" and "medium-loud" levels, although more data are needed to establish the potential onset of the viscid flow regime. Therefore, the effect of viscosity on the estimated hydraulic diameter D_c is approximated as expressed by Eq. (4). The length of the constricted portion is set to 22 mm based on the CT scan illustrated in Fig. 1(b). It is seen (crosses in Fig. 4) that accounting for viscous losses does not improve the accuracy of the estimation of the hydraulic diameter of the minimum constriction area. Consequently, for the phoneme /s/ an inviscid laminar flow model, Eq. (3) can be used to estimate the order of magnitude of hydraulic diameter D_c for all loudness conditions.

Finally, it is noted (dashed line in Fig. 4) that, regardless of the loudness condition, D_c values reported in literature or estimated from aerodynamic measurements, are greater than hydraulic diameter $D_c = 2.1$ mm characterizing the minimum area ($3.5\,\mathrm{mm}^2$) of the reconstructed oral tract geometry from CT data described in Sec. 2.2. The discrepancy is suggested to be because of the different duration of the phoneme /s/ during aerodynamic measurements (natural, <2 s) and the imaging process (sustained, ≈ 10 s). The prolonged timespan required for CT volume imaging favors a lower oral flow rate, which in turn results in smaller D_c values following Eq. (3). Considering the wide range of $U(D_c)$ values represented in Fig. 4 (estimated or taken from literature), it is seen that a decrease in oral flow rate alone cannot explain the low value of D_c of the reconstructed geometry. Although no conclusion can be made from the current data, it seems that the effect of a lower flow rate is reinforced by hyper-articulation in order to produce a sustained sibilant /s/ during several seconds, which is not a natural speech condition.

4. Conclusion

Oral air flow and intra-oral pressure were measured during sibilant fricative utterances of two seated speakers using a simple non-invasive measurement technique. The minimum area of the narrowed passage upstream from the teeth, characterizing sibilant fricative phonemes, was estimated from aerodynamic quantities using a laminar inviscid flow assumption. It is shown that accounting for viscid flow losses is negligible compared to the uncertainty resulting from aerodynamic data statistics. Estimated values correspond well with values reported in literature. Therefore, the proposed estimation of aerodynamic quantities is judged to be accurate and thus provides a non-invasive

Redistribution subject to ASA license or copyright; see http://acousticalsociety.org/content/terms. Download to IP: 147.171.69.213 On: Wed, 08 Jul 2015 12:36:24

quantitative method to include an articulation parameter in speech screening protocols for sibilant fricatives. Moreover, on the basis of the uncertainty of the estimated minimum area, it is concluded that phoneme /s/ is more suited for speech screening than phoneme /ʃ/ which supports common practice in oral health care. Finally, it is noted that a more detailed validation remains to be done since the number of speakers in the current study is limited and speakers have good oral health. Nevertheless, validation is not straightforward since most studies reporting on the minimum area lack quantitative information about associated aerodynamic quantities and the diversity in screening protocols and applied techniques (speaker in seated or supine position, natural or sustained phonemes, different phoneme loudness conditions, different acoustic environment, etc.). Future studies combining imaging techniques and aerodynamic measurements applied during a same protocol in an identical acoustic environment are of interest.

References and links

[1]A. Lee, T. Whitehill, V. Ciocca, and N. Samman, "Acoustic and perceptual analysis of the sibilant sound /s/ before and after orthognathic surgery," J. Oral Maxillofac. Surg. **60**, 364–372 (2002).

[2]E. Bond and W. Lawson, "Speech and its relation to dentistry. I. Speech and speech defects," Dent. Pract. Dent. Rec. **19**(5), 75–82 (1968).

[3]S. Pound, "Let /s/ be your guide," J. Prosthes. Dent. **38**(5), 482–489 (1977).

[4]E. Mehringer, "The use of speech patterns as an aid in prosthodontic reconstruction," J. Prosthes. Dent. **13**(5), 825–836 (1963).

[5]C. Runte, M. Lawerino, D. Dirksen, F. Bollmann, A. Lamprecht-Dinnesen, and E. Seifert, "The influence of maxillary central incisor position in complete dentures on /s/ sound production," J. Prosthes. Dent. **85**(5), 485–495 (2001).

[6]J. Hufnagle, P. Pullon, and K. Hufnagle, "Speech considerations in oral surgery. Part II. Speech characteristics of patients following surgery for oral malignancies," Oral Surg. Oral Med. Oral Pathol. **46**, 354–361 (1978).

[7]G. Fant, *The Acoustic Theory of Speech Production* (Mouton, The Hague, the Netherlands, 1960), 1002 pp.

[8]S. McLeod, A. Roberts, and J. Sita, "Tongue/palate contact for the production of /s/ and /z/," Clin. Ling. Phonetics **20**(1), 51–66 (2006).

[9]EVA2 datasheet, Technical report, http://www.sqlab.fr/doc/PlaquetteEvaENU0501.pdf (Last viewed 2009).

[10]K. Nozaki, T. Yoshinaga, and S. Wada, "Sibilant /s/ simulator based on computed tomography images and dental casts," J. Dent. Res. **93**, 207–211 (2014).

[11]J. Cisonni, A. Van Hirtum, J. Willems, and X. Pelorson, "Theoretical simulation and experimental validation of inverse quasi-one-dimensional steady and unsteady glottal flow models," J. Acoust. Soc. Am. **124**(1), 535–545 (2008).

[12]D. Warren and A. DuBois, "A pressure-flow technique for measuring velopharyngeal orifice area during continuous speech," Cleft Palate J. **1**, 52–71 (1964).

[13]S. Narayanan, A. Alwan, and K. Haker, "An articulatory study of fricative consonants using magnetic resonance imaging," J. Acoust. Soc. Am. **98**, 1325–1347 (1995).

[14]C. Shadle and C. Scully, "An articulatory-acoustic-aerodynamics analysis of /s/ in VCV sequences," J. Phonetics **23**(3), 53–66 (1995).

[15]B. Wu, A. Van Hirtum, and X. Luo, "Pressure driven steady flow in constricted channels of different cross section shapes," Int. J. Appl. Mech. **5**, 1–19 (2013).

[16]N. Timm, *Handbook of Statistics* (Motilal Banarsisdass Publisher, Delhi, India, 1980), pp. 41–87.

[17]N. Lass, Ed., *Contemporary Issues in Experimental Phonetics* (Elsevier, New York, 1976), pp. 105–137.

[18]K. Stevens, "Airflow and turbulence noise for fricative and stop consonants: Static considerations," J. Acoust. Soc. Am. **50**, 1180–1192 (1971).

Redistribution subject to ASA license or copyright; see http://acousticalsociety.org/content/terms. Download to IP: 147.171.69.213 On: Wed, 08 Jul 2015 12:36:24