From in-vivo data to mechanical replica: the case of sibilant noise

Y. Fujiso et A. Van Hirtum
GIPSA-lab, 11 rue des Mathématiques, 38402 Saint Martin D'Hères Cedex, France
annemie.van-hirtum@gipsa-lab.grenoble-inp.fr

Sibilant fricative speech production emerges from the interaction of a (jet) flow with the surrounding vocal tract (walls, tooth, articulators and lips). Consequently, sibilant fricative noise depends on geometrical as well as flow properties. Nevertheless, few studies deal with the impact of flow properties other than the volume flow rate (loudness level) on the noise produced. In this work, an experimental study is presented using a reconstructed constant oral cavity during human sibilant production. It is aimed to assess the potential impact of initial flow conditions upstream from jet formation and jet impact on sibilant fricative noise. The velocity distribution of the flow upstream from the sibilant groove is characterized by its statistical moments for different air supplies resulting in different inlet conditions. Spectral features of the noise produced are quantified. It is concluded that initial conditions can not be neglected.

1 Introduction

The underlying mechanism of sibilant fricative sound production is generally described as noise produced due to the interaction of a turbulent jet, issued from a constriction somewhere in the vocal tract, with a downstream wall or obstacle [13, 14]. Screening of geometric and flow parameters during human sibilant fricative phoneme production enabled to establish the major impact on acoustic sibilant features such as spectral properties or loudness level [6]. Flow and geometrical parameters during human sibilant production are inherently correlated and show a large inter- and intra-speaker variability [8] . To overcome these difficulties proper to human sibilant phoneme data, several studies rely on ex-vivo data – following experimental, numerical or modeling studies – in order to single out an individual sibilant production parameter (either geometric or flow) and to study its impact for sibilant sound production [13, 7, 5, 16, 11, 2]. The cited studies either do not mention flow conditions upstream from jet formation or suppose laminar upstream flow, whereas from literature focusing on jet development it is known that initial conditions, *i.e.* upstream flow conditions, will affect jet formation and downstream flow dynamics. Since jet formation is an essential part of sibilant fricative production, it is motivated to reflect on the potential impact of upstream flow conditions on jet formation and on the subsequent noise produced within the framework of sibilant fricative production. In order to ascertain decorrelated geometrical and flow properties and in order to ascertain a pertinent geometry, we rely not on a human speaker, but on a reconstructed – and hence constant – vocal tract geometry obtained from a cone-beam computed tomography (CT) scan during production of the phoneme /s/. We measured the acoustic outcome for airflow through this geometry for different upstream flow profiles.

2 Method

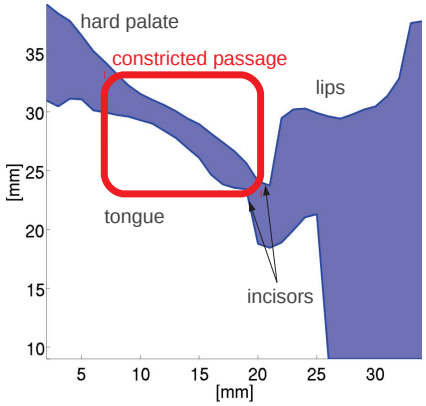
2.1 Reconstructed geometry

A cone-beam CT scan (CB MercuRay, 512 slices of 512× 512 pixels with accuracy ±0.1mm) was made for a single adult male subject 'KN' (Japanese native speaker, normal sitting position) while uttering phoneme /s/ at 'medium' loudness level (≈21l/min), so that the articulators positions (teeth, tongue, lips) corresponded to a sibilant /s/ sound articulation. The imaging process lasted about 10s. The oral cavity volume and its shape were reconstructed using a marching cube method and an optical modeling machine (SOUP 2 600GS, material TSR-829) with spatial accuracy ±0.1mm [10, 9]. Consequently, the combined maximal error resulting from the imaging process and the optical reconstruction yielded ±0.2mm. The reconstructed geometry (sectional CT scan data slice) is depicted in Fig. 1(a) and a sectional view and front of a three-dimensional plaster model is illustrated in Fig. 1(b). The length of the reconstructed portion is 37mm. The minimum area within the constricted passage of length $\simeq$ 22mm yielded 3.5mm^2 corresponding to a minimum hydraulic diameter of 2.1mm. The inlet portion of the reconstructed portion is of elliptical shape (major axis 17mm and minor axis 11mm) with hydraulic diameter 13mm. A mechanical vocal tract replica for sibilant /s/ was obtained by smoothly connecting an upstream circular duct, shown in Fig. 1(c) with length 120mm, inlet diameter 20mm and outlet diameter 8 mm, to the rigid reconstructed geometry (plaster or transparent) with elliptical inlet. The length of the connecting portion is 30mm so that based on the converging angle (9^o) the flow in the connecting portion is unstalled [1]. The upstream circular duct represents the portion of the vocal tract from the larynx up to the reconstructed portion characterising /s/. In the following, upstream flow conditions are determined at the outlet of the circular duct (exit diameter 8mm or aera $A = 50$mm^2). The total length of the mechanical replica was 18cm which corresponds to the averaged length of the vocal tract of an adult male subject [3, 14]. The complete replica is illustrated in Fig. 1(d).

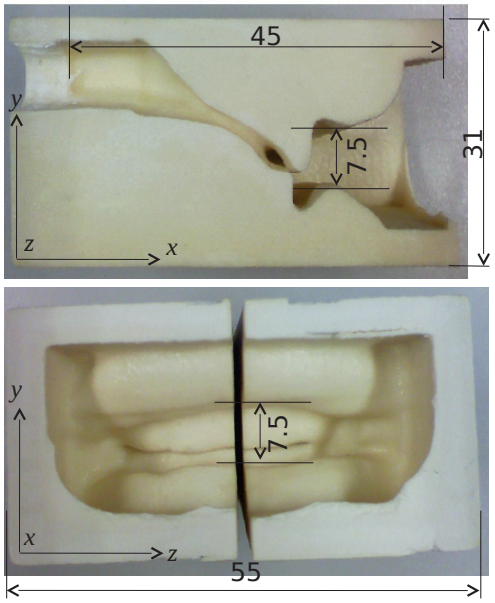
2.2 Flow supply and flow measurements

Airflow is provided to the replica illustrated in Fig. 1(d) in different ways. Firstly, a human subject blows into the replica. Secondly, a flow facility is used to provide airflow at a constant pressure. The flow facility consists of a an air compressor (Atlas Copco GA7), followed by a pressure regulator (Norgren type 11-818-987) providing an airflow at constant pressure. The volume flow rate Q is controlled by a secondary manual valve placed downstream the regulator and measured by a thermal mass flow meter (model 4043 TSI) with an accuracy of 2% of its reading. To homogenize the flow, a settling chamber is used with volume 0.4×0.4×0.5m^3 to which a series of 3 perforated plates with holes of diameter 1.5mm (spaced 0.8mm) are added. The walls of the settling chamber are tapered with acoustic foam (SE50-AL-ML Elastomeres Solutions) in order to avoid acoustic resonances. In addition, a grid can be insterted between the settling chamber and the replica.

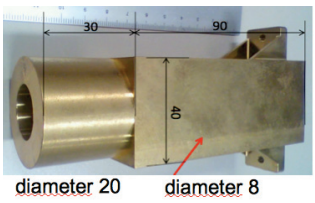
In order to characterise flow conditions, the flow velocity is measured at the exit of the circular upstream duct using hot-film (TSI 1201-20) in combination with a IFA 300 anemometer (TSI). The hot film is calibrated according to the technique described in [4] which results in a calibration curve fitted on a fourth order polynomial law. The hot film is mounted on a two dimensional stage positioning system (Chuo preci-



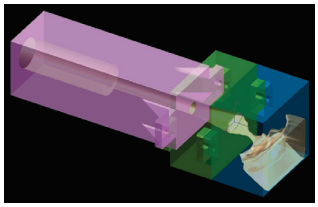
(a) centerline plane



(b) reconstructed geometry



(c) upstream circular duct



(d) complete replica

Figure 1: a) Illustration of CT scan data slice during an /s/ utterance showing the characterizing constricted airflow passage. b) Three-dimensional plaster reconstruction of a human oral cavity while pronouncing /s/ (dimensions in mm). c) Upstream circular duct and d) Complete replica consisting of upstream circular duct (pink, length 120mm) to which the three-dimensional reconstructed geometry (blue, length 37mm) is smoothly (green, length 30mm) mounted.

sion industrial co. CAT-C, ALS-250-C2P and ALS-115-E1P with accuracy of $\leq 4\mu$m). The velocity profile at the exit of the upstream circular duct (D = 8mm) shown in Fig. 1(c) is characterised by placing the hot-film downstream of the duct at a distance $\leq$ 0.5mm. The centerline velocity U_c at $y/D = 0$ is measured as well as the transverse velocity profile $-0.5 \leq y/D \leq 0.5$ with a spatial step $\Delta y/D = 0.11$ for human blowing and $\Delta y/D = 0.038$ for flow facility use. At each measurement position, velocity data are sampled

Table 1: Summary of flow conditions for human blowing used for acoustic measurements: loudness instruction and corresponding volume flow rate Q and Reynolds number Re_c at the exit of the circular tube shown in Fig. 1(c).

Loudness condition	≈Q [l/min]	≈Re_c [-]
Soft	11	2850
Medium	21	5250
Loud	45	8850

at 5kHz during $\geq$ 4s consecutively for human blowing and 30s consecutively for flow facility use. Statistical quantities (first up to the fourth moment) of the velocity distribution at each spacial position are obtained from instantaneous velocity measurements to which a 10Hz high pass filter and a 10kHz low pass filter is applied. The local mean velocity U, local turbulence intensity T_u, skewness S_u and flatness F_u are considered as a function of centerline velocity Reynolds number $Re_c = U_c D/\nu$, with kinematic viscosity of air $\nu = 1.5\times10^{-5}$m^2/s, in order to characterise the initial flow conditions upstream from the reconstructed portion [12].

2.3 Acoustic measurements

Acoustic measurements with the reconstructed geometry for different flow supplies described in the previous section 2.2, *i.e.* human blowing, flow facility with grid and flow facility without grid, were conducted using a quasi-anechoic chamber [15] by connecting the outlet of the replica to the opening into the quasi-anechoic chamber as illustrated in Fig. 2. The quasi-anechoic chamber has dimensions $2.07m\times2.10m\times2.14m$ (volume 9.3m^3) and is equipped with a pressure-field Brüel & Kjaer microphone (B&K 4192, frequency response 20Hz-20kHz, preamplifier B&K 2669 and +30dB B&K amplifier 5935). The microphone (sampling frequency 44.1-kHz) was located at the same height as the replica but with an angle of 37 ° with respect to its centerline, in the horizontal plane. The volume flow rates were chosen following loudness instructions suitable to characterize fricatives [6]: 'soft' (≈11l/min, bulk velocity U_b ≈4m/s), 'medium' (≈21l/min, bulk velocity U_b ≈7m/s) and 'loud' (≈45l/min, bulk velocity U_b ≈15m/s). A qualitative estimation of the volume flow rate associated with each loudness level was obtained using the volume flow meter (TSI 4000 series) in combination with the loudness instruction and confirmed by integration of the transverse mean velocity profile. An overview of the loudness instruction and corresponding volume flow rate Q and centerline velocity Reynolds number Re_c at the exit of the circular tube shown in Fig. 1(c), is given in Table 1.
Spectral features of the radiated acoustic pressure signals are derived from the acoustic spectra L_p as a function of physical frequency f. The acoustic spectra are obtained as the energy normalized sound pressure level (SPL) of the Welch averaged Power Spectral Density $P(f)$ of the measured acoustic pressure signal [17, 6]:

$$L_p(f) \approx 10 \cdot log_{10}\left(\frac{|P(f)|}{p_{ref}^2}\right), \quad (1)$$

with $p_{ref} = 2 \cdot 10^{-5}$Pa. Concretely, the time-averaging of the periodiograms is performed using Hamming windowed energy normalized time segments of fixed length (6ms) with

50% overlap. The sound pressure level spectra are parameterised by considering the total power of the acoustic pressure signal P, dynamic amplitude A_d, peak frequency f_m and spectral linear regression slopes S_1 ($f_{min} \leq f \leq f_m$, with f_{min} the frequency associated with minimum spectral amplitude, positive slope) and S_2 ($f_m \leq f \leq$ 20kHz, negative slope) in accordance with features exploded for human fricative phoneme characterisation [6, 14].

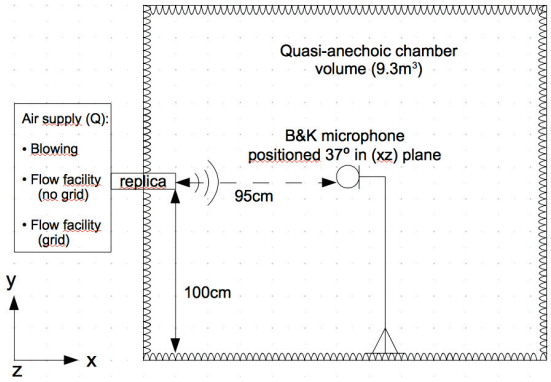


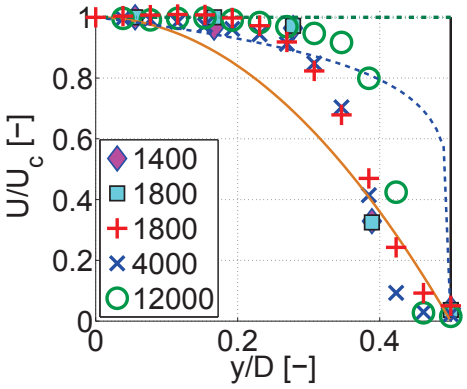
Figure 2: Overview of acoustic measurement setup. The microphone is positioned at an angle of 37 ° with respect to the main airflow direction in the horizontal plane (xz).

3 Results

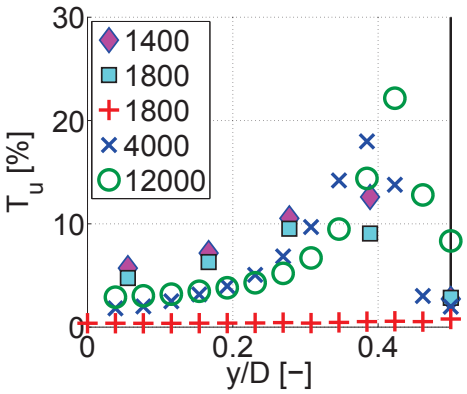
3.1 Upstream flow characterisation

In this section, we evaluate the impact of the used air supply on initial flow conditions from velocity measurements at the exit of the circular duct (shown in Fig. 1(c), following the methodology outlined in section 2.2). Transverse velocity profiles from the duct center ($y/D = 0$) up to the duct wall ($y/D = 0.5$) are illustrated in Fig. 3 for different air supplies (human blowing or flow facility use for $Re_c \approx 1800$ or $Q \approx$ 7l/min) and Reynolds numbers Re_c (flow facility for $Re_c \approx$ 1800, 4000, 12000 or $Q \approx$ 7, 18, 55l/min). Normalized mean transverse velocity profiles U/U_c (Fig. 3(a)) are characterized by an almost uniform ($U/U_c > 0.9$) core region ($|y/D| < 0.3$) regardless of flow supply or Reynolds number Re_c. The normalized mean velocity profile within the boundary layer enveloping the core region is observed to depend on the Reynolds number and less on the flow facility. The measured mean velocity profiles approximate therefore a top-hat velocity profile for which the boundary layer displacement thickness depends on Reynolds number [12]. As a result, mean velocity profiles characterizing laminar or turbulent pipe flow are not relevant as illustrated in Fig. 3(a). The corresponding transverse profiles of the turbulence intensity (Fig. 3(b)) are affected significantly by Reynolds number Re_c ($\Delta T_c \geq 2\%$) as well as air supply ($\Delta T_c \geq 4\%$) regardless of transverse position y/D, *i.e.* the turbulence intensity varies within the core region as well as within the boundary layer depending on flow supply or Reynolds number Re_c. Moreover, it is seen that turbulence intensities associated with human blowing are high compared to those resulting from flow facility use and this within the core region as well as within the boundary layer.

In order to characterize initial flow conditions further, velocity moments at the center of the duct ($y/D = 0$) are considered (Fig. 4) as a function of Reynolds number $500 < Re_c < 15000$ ($1 < U_c < 25$) and flow supply (see section 2.2: human blowing, use of flow facility without and with additional inlet grid). The centerline turbulence intensity $T_{u,c}$ (Fig. 4(a)) is seen to be largely affected by the flow supply since turbulence intensities observed for human blowing ($T_{u,c} > 3\%$) exceeds the turbulence level for turbulent flow when using flow supply without ($T_{u,c} \leq 2\%$) and with an inlet grid ($T_{u,c} \leq 3.2\%$). Increasing the Reynolds number Re_c is seen to result in an increase of the turbulence intensity $T_{u,c}$ regardless of the flow supply. It is further noted that, whereas for human blowing turbulent flow is observed regardless Re_c, increasing the Reynolds number is seen to trigger a transition of laminar ($T_{u,c} < 0.5\%$) to turbulent flow. The use of an inlet grid is seen firstly to decrease Reynolds numbers at which transition from laminar to turbulent regime occur (*e.g.* reduces the transition onset from $Re_{u,c} \approx 7000$ to $Re_{u,c} \approx 2000$) and secondly to favor the turbulence intensity level the turbulent regime as is expected for grid turbulence. The increase of the turbulence intensity with Reynolds number is in contrast with the decrease of turbulence intensity expected for pipe flow, but in case of human blowing the tendency can be expressed by a quadratic polynomial law for which the coefficient of determination yields

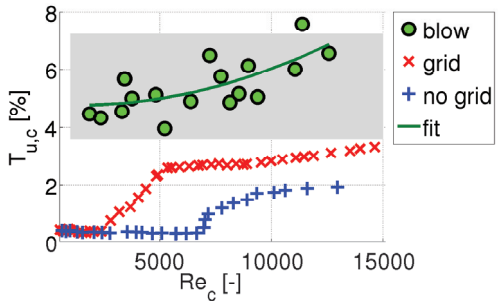


(a) normalized mean

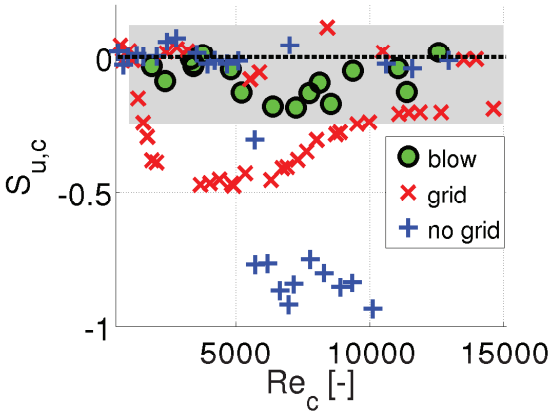


(b) turbulence intensity

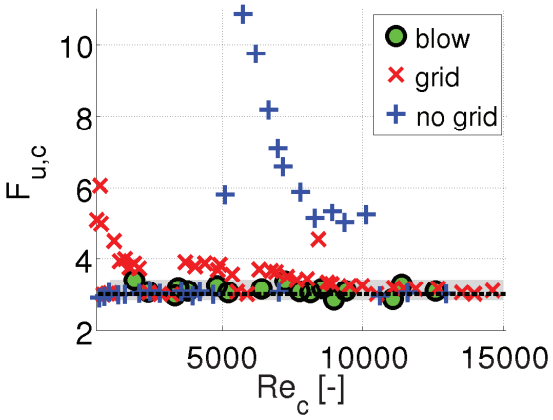
Figure 3: Transverse ($0 \leq y/D \leq 0.5$) velocity profiles as a function of Reynolds numbers Re_c and air supply – human blowing (◇, □) and flow facility use (+, ×, ○): a) normalized mean velocity U/U_c (symbols) and developed laminar (full line – parabolic) and turbulent (dashed line – power-law) pipe flows, b) turbulence intensity T_u.



(a) turbulence intensity $T_{u,c}$



(b) skewness $S_{u,c}$



(c) flatness $F_{u,c}$

Figure 4: Centerline initial velocity statistical features as a function of Reynolds number Re_c and flow supply (human blowing – filled ∘, flow facility without (+) and with (×) inlet grid). The gray shaded area indicates the feature range covering 95% of the values observed for human blowing: a) turbulence intensity $T_{u,c}$ and a quadratic fit (full line) of $T_{u,c}(Re_c)$ ($R^2 = 0.5$), b) skewness $S_{u,c}$ and $S_{u,c} = 0$ (dashed line) and c) flatness $F_{u,c}$ and $F_{u,c} = 3$ (dashed line) .

$R^2 = 0.5$. Moreover, it is observed that the turbulence intensity increases more rapidly when the flow is supplied by human blowing than when a flow facility is used. The skewness $S_{u,c}$ (Fig. 4(b)) and flatness $F_{u,c}$ (Fig. 4(c)) of the center velocity distribution as a function of Reynolds number further illustrates the impact of the flow supply on initial flow conditions. Values characterizing a Gaussian velocity distribution ($S_{u,c} = 0$ and $F_{u,c} = 3$) provide an approximation for human blowing regardless the Reynolds number Re_c and provide a poor description for the values observed when a flow facility is used, *e.g.* for Reynolds numbers Re_c associated with the transition regime.

3.2 Acoustic characterisation

In the previous section 3.1, a significant variation of initial flow conditions due to varying the flow supply is shown and quantified in terms of statistical moments. From the characterization of the initial velocity profile discussed in section 3.1, it is seen that assessed flow conditions when a flow facility is used correspond to a Reynolds number Re_c within the laminar, transition and turbulent regime. In order to search the influence of the variation of initial flow conditions on the variation of the radiated noise outcome we consider acoustic features derived from the gathered sound spectra as outlined in section 2.3 with the setup illustrated in Fig. 2. The power of the acoustic signal P, dynamic amplitude A_d, frequency with maximal acoustic energy f_m and spectral slopes $S_{1,2}$ are plotted in Fig. 5 as a function of Reynolds number Re_c for different air supply methods. For air supplied by human blowing we consider a limited number of flow conditions summarized in Table 1 resulting from the loudness instructions ('soft', 'medium' or 'loud') and we will vary the Reynolds number Re_c within this range for air supplied using the flow facility (with grid and without grid). The general tendencies of the acoustic features are similar regardless the air supply and are within the range of values observed for air supplied by human blowing. Nevertheless, important quantitative differences are seen to occur, which are observed to increase with Reynolds number Re_c when considering air supplied by the air facility with and without grid. This last observation points to a different contribution of the velocity fluctuations to the noise produced following different inlet condition, which following Fig. 4 becomes more important as the Reynolds number increases. Concretely, the measured acoustic power P (Fig. 5(a)) and dynamic amplitude A_d (Fig. 5(b)) increases strongly (P with ≈50% and A_d with ≈25%) for Reynolds numbers in the range $Re_c \approx 7000$ when a grid is used whereas the transition regime seems to limit P and A_d in absence of a grid. The same way, the magnitude of the negative spectral slope characterizing the high frequency range S_2 is more important (S_2 with ≈25%) in absence of a grid than when a grid is used. The variation in terms of positive spectral slope characterizing the lower frequency range S_1 is less pronounced. The same holds for the variation observed for the frequency f_m associated with maximal spectral energy. In both cases we observe an important dynamic with Reynolds number Re_c such as the shift in $f_m \approx 7.5$kHz for $Re_c \leq 3500$ to $f_m \approx 11$kHz for $Re_c > 3500$. Note, that despite the constant geometry small differences are still observed (≤ 10% for f_m). With respect to the features derived for human blowing it is interesting to note that for A_d and P observed values approximate values obtained for air supply with a flow facility without grid use, whereas for spectral slopes and in particular for S_2 acoustic features resulting from human blowing approximate features observed for air supplied with a flow facility with grid. Finally, it is seen that for high Reynolds numbers frequency associated with maximum spectral energy is increased compared to values obtained using a flow facility.

4 Conclusion

The current paper showed based on an analysis of acoustical features of noise produced from flow through a reconstructed and constant geometry (vocal tract during phoneme /s/) that the variation of flow properties upstream from the sibilant groove influence the spectral properties of the noise produced. This is an important finding since initial flow conditions are not considered in physical speech production studies. The current experimental results show that initial conditions need to be quantified and can not be neglected.

References

[1] R. Blevins. *Applied Fluid Dynamics Handbook*. Krieger publishing company, Malabar, 1992.

[2] J. Cisonni, K. Nozaki, A. Van Hirtum, X. Grandchamp, and S. Wada. Numerical simulation of the influence of the orifice aperture on the flow around a teeth-shaped obstacle. *Fluid Dynamics Research*, 45:1–19, 2013.

[3] G. Fant. *The acoustic theory of speech production*. Mouton, The Hague, 1960.

[4] X. Grandchamp, A. Van Hirtum, and X. Pelorson. Hot film/wire calibration for low to moderate flow velocities. *Meas. Sci. Technol.*, 21:115402, 2010.

[5] M.S. Howe and R.S. McGowan. Aeroacoustics of [s]. *Proceedings of the Royal Society A*, 461(2056):1005–1028, 2005.

[6] L.M.T. Jesus and C.H. Shadle. A parametric study of the spectral characteristics of European Portuguese fricatives. *J. Phonetics*, 30(3):437–464, 2002.

[7] M. Krane. Aeroacoustic production of low-frequency unvoiced speech sounds. *J. Acoust. Soc. Am.*, 118(1):410–427, 2005.

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

[9] K. Nozaki. Numerical simulation of sibilant [s] using the real geometry of a human vocal tract. In *High Performance Computing on Vector Systems 2010*, pages 137–148. Springer, 2010.

[10] K. Nozaki, T. Akiyama, H. Tamagawa, S. Kato, Y. Mizuno-Matsumoto, M. Nakagawa, Y. Maeda, and S. Shimojo. The first grid for oral and maxillofacial region and its application for speech analysis. *Methods of information in medicine*, 44:253–256, 2005.

[11] K. Nozaki, M. Nakamura, H. Takimoto, and S. Wada. Effect of expiratory flow rate on the acoustic characteristics of sibilant /s/. *Journal of Computational Science*, 3(5):298–305, 2012.

[12] S. Pope. *Turbulent flows*. Cambridge University Press, 2005.

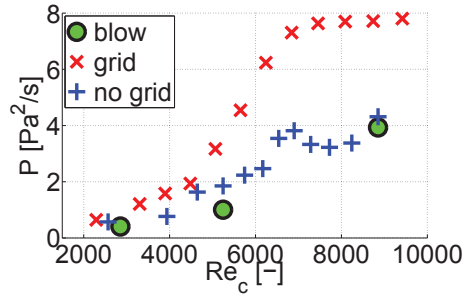
[13] C.H. Shadle. *The acoustics of fricative consonants*. PhD thesis, Massachusetts Institute of Technology, 1985.

[14] K Stevens. *Acoustic Phonetics*. MIT Press, London, 1998.

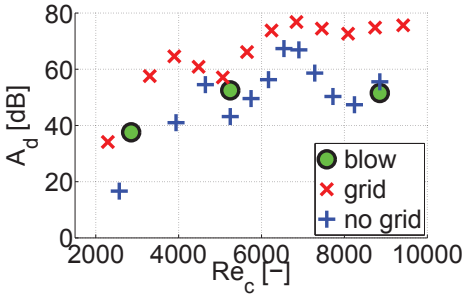
[15] A. Van Hirtum and Y. Fujiso. Insulation room for aeroacoustic experiments at moderate Reynolds and low Mach numbers. *Applied Acoustics*, 73(1):72–77, 2012.

[16] A. Van Hirtum, X. Pelorson, O. Estienne, and H. Bailliet. Experimental validation of flow models for a rigid vocal tract replica. *J. Acoust. Soc. Am.*, 130:2128, 2011.

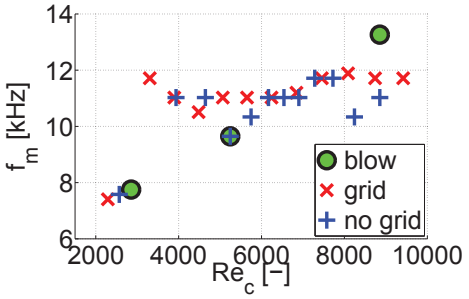
[17] P. Welch. The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodiograms. *IEEE Trans. Audio and Electroacoustics*, 15:70–73, 1967.



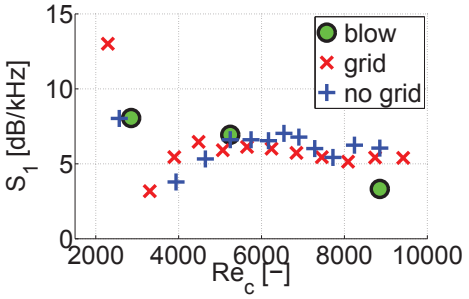
(a) acoustic power P



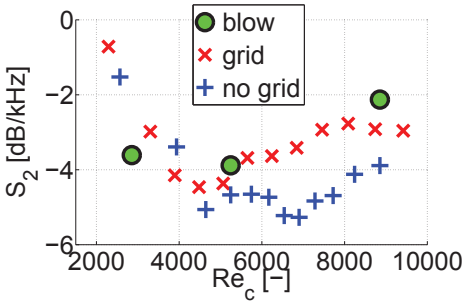
(b) dynamic amplitude A_d



(c) maximal spectral frequency f_m



(d) positive spectral slope S_1



(e) negative spectral slope S_2

Figure 5: Spectral features of acoustic signal as a function of air supply and resulting flow conditions (human blowing, compressed air with and without inlet grid).