Received: 23 March 2018 | Revised: 31 August 2018 | Accepted: 5 September 2018

DOI: 10.1002/cnm.3153

RESEARCH ARTICLE - APPLICATION

WILEY

Computational aeroacoustics to identify sound sources in the generation of sibilant /s/

Arnau Pont[1,2] | Oriol Guasch[2] | Joan Baiges[3] | Ramon Codina[3] | Annemie van Hirtum[4]

[1]Centre Internacional de Mètodes Numèrics en Enginyeria, Universitat Politècnica de Catalunya, Barcelona, Spain

[2]GTM—Grup de recerca en Tecnologies Mèdia, La Salle—Universitat Ramon Llull, Barcelona, Spain

[3]Departament d'Enginyeria Civil i Ambiental, Universitat Politècnica de Catalunya, Barcelona, Spain

[4]EGI, UMR CNRS 5519 Grenoble Alpes University, Grenoble, France

Correspondence
Oriol Guasch Fortuny, GTM—Grup de recerca en Tecnologies Mèdia, La Salle—Universitat Ramon Llull, Barcelona, Spain.
Email: oriol.guasch@salle.url.edu

Present Address
Oriol Guasch, C/ Quatre Camins 30, La Salle, Universitat Ramon Llull, Barcelona 08022, Catalonia, Spain

Funding information
EU-FET, Grant/Award Number: EUNISON 308874; Agència de Gestió d'Ajuts Universitaris i de Recerca, Grant/Award Number: 2015 FI-B 00227; GENIOVOX, Grant/Award Number: TEC2016-81107-P; Secretaria d'Universitats i Recerca del Departament d'Economia i Coneixement (Generalitat de Catalunya), Grant/Award Number: 2014-SGR-0590 and 2016-URL-IR-013; Spanish Government through the Ramón y Cajal, Grant/Award Number: RYC-2015-17367; ArtSpeech, Grant/Award Number: ANR-15-CE23-0024

Abstract

A sibilant fricative /s/ is generated when the turbulent jet in the narrow channel between the tongue blade and the hard palate is deflected downwards through the space between the upper and lower incisors and then impinges the space between the lower incisors and the lower lip. The flow eddies in that region become a source of direct aerodynamic sound, which is also diffracted by the speech articulators and radiated outwards. The numerical simulation of these phenomena is complex. The spectrum of an /s/ typically peaks between 4 and 10 kHz, which implies that very fine computational meshes are needed to capture the eddies producing such high frequencies. In this work, a large-scale computation of the aeroacoustics of /s/ has been performed for a realistic vocal tract geometry, resorting to two different acoustic analogies. A stabilized finite element method that acts as a large eddy simulation model has been adopted to solve the flow dynamics. Also, a numerical strategy has been implemented that allows the determination, in a single computational run, of the separate contribution of the sound diffracted by the upper incisors from the overall radiated sound. Results are presented for points located close to the lip opening showing the relative influence of the upper teeth depending on frequency.

KEYWORDS

computational aeroacoustics, finite elements, fricatives, large eddy simulation, Lighthill

1 | INTRODUCTION

In this work, we aim at better understanding the generation and radiation mechanisms of sibilant /s/, by means of computational aeroacoustics (CAA) performed on a realistic vocal tract geometry. In particular, we will explore the possibilities of CAA to determine the separate acoustic contributions from the flow noise generated near the lip opening and from its

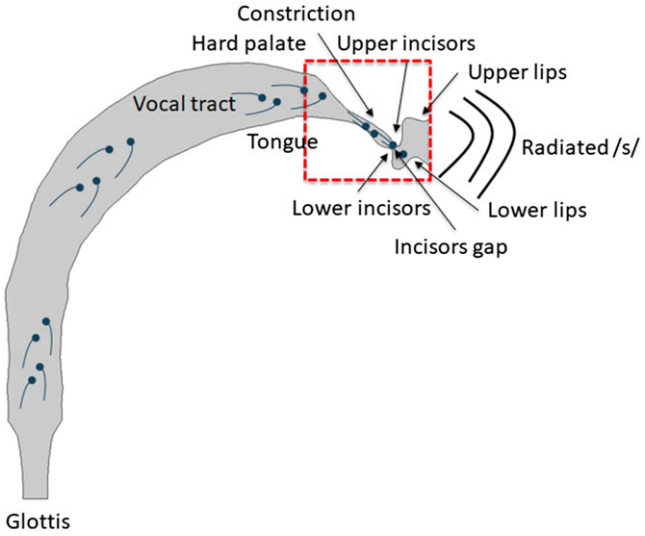

FIGURE 1 Schematic of the midsagittal section of the human vocal tract configured to generate an /s/

diffraction by the upper incisors. The influence of the latter on the characteristic high-frequency peaks of an /s/ realization will be analyzed. Besides, this work will constitute a first applied example of the numerical strategy reported in Guasch et al,[1] which avoids using hybrid finite element/boundary element CAA approaches. Also, the work constitutes another step towards demonstrating the potential of large-scale physics-based numerical simulations in speech production.

A schematic of the midsagittal section of the human vocal tract with the various speech articulators configured to generate an /s/ is presented in Figure 1. The airflow emanating from the lungs passes through the glottis and propagates within the vocal tract until it reaches the anterior part of the mouth. There, the tongue blade is pressed upwards to the hard palate leaving a small constriction where the airflow is accelerated and transformed into a turbulent jet. This passes through the space between the upper and lower incisors and impinges the cavity between the lower incisors and the lower lip. As a result, aerodynamic noise is generated, which in turn gets diffracted by all articulators (specially by the upper incisors) and becomes radiated outwards. The red dashed square in Figure 1 indicates the region where most of the physics of /s/ generation occurs. It comprises the domain of the vocal tract that will be considered for the numerical simulations in this work.

Sibilant /s/ sound has a characteristic wide-band spectral content that can be observed in measurements of uttered sibilants in previous studies.[2-4] Strong differences can be appreciated between those works at low frequencies, yet the general trends of the spectra are similar. They all exhibit a decay from low frequencies to a dip between approximately 1 to 3 kHz, followed by a strong level increase with frequency up to approximately 8 to 9 kHz, and two peaks whose locations change depending on each realization. It should be remarked that there are significant variations between the spectral shapes of speakers reported in the literature because of morphological differences. Therefore, one of the main goals in this work will consist in reproducing the basic tendencies obtained in most sibilant measurements,[3-7] with special attention to those in Nozaki et al.[2] The latter are performed on a mechanical replica upon which our computational domain relies.

A detailed analytical model to describe the physics behind the production of /s/ was proposed in Howe and McGowan.[8] In that model, the diffraction of the blocked wall pressure beneath a turbulent boundary layer (TBL) attached to the incisors was considered as the main noise contributor. A compact Green's function that accounted for a simplified geometry of the incisors and the vocal tract was deduced and convolved with a semiempirical model for the wall pressure wavenumber-frequency spectrum,[9] to predict the acoustic pressure at the far field. More recently, in Yoshinaga et al,[10] large eddy simulations (LES) were presented on a 3D realization of the geometry in Howe and McGowan[8] and lately compared with a realistic one in Yoshinaga et al.[11] Those works reported that it is actually in the cavity between the lower incisors and the lower lip where most of the aeroacoustic source terms concentrate. The simulations in the present work, also with a realistic vocal tract, support that conclusion. In addition, they essentially show that for points in the vicinity of the mouth, it is the diffraction of aerodynamic sound by the upper incisors that governs the high-frequency range of the spectrum, while the contribution to lower frequencies corresponds to the direct sound generated by the eddies within the lower incisors-lips cavity and its modulation by articulators other than the upper incisors. In fact, in the very recent work,[12] multimodal acoustic simulations were performed on a simplified vocal tract geometry, placing a monopole source

term close to the upper teeth. It was shown therein how the shape of the upper teeth region was responsible for the /s/ high-frequency peak. Our more complex CAA simulations that contemplate both, the flow dynamics and the aerodynamic acoustic waves generated by the flow-induced quadrupole source distribution on a realistic vocal tract, also seem to support that point of view (yet further investigations on the topic would be worthwhile in the future). Besides, and concerning the numerical simulations, we shall note that despite the articulators being in constant movement during speech,[13] stationary vocal tract walls are always assumed in computational models for simplicity.

To validate the above assertions, the numerical strategy in Guasch et al[1] was implemented. We note that in most hybrid approaches to CAA, a two-step procedure is followed (see, eg, Bailly et al[14]). First, an LES computation is carried out by means of a finite element (FEM) (or a finite volume) approach, to obtain the aerodynamic noise source terms. Secondly, these terms are input into an acoustic analogy that is solved using an integral formulation (see, eg, Curle[15] and Ffowcs-Williams and Hawkings[16]), which becomes discretized by a boundary element method (BEM). This procedure is advantageous whenever one has to compute the acoustic pressure at far-field points. Yet, if one is interested in the acoustic field at points closer to the turbulent source region that strategy may be not advantageous, because BEM matrices are fully populated while those of FEM are sparse. The proposal in Guasch et al[1] goes in this direction and only makes use of a FEM code. This allows one to solve, in a single computational run and with a sole computational mesh, an LES for the incompressible Navier-Stokes equations, together with a first-wave equation for the incident flow noise contribution and a second-wave equation for the sound diffracted by the incisors. The procedure circumvents an inconsistency related to the numerical solution of Curle's dipolar integral term for low Mach numbers,[15] given that the total pressure to be input in that integral cannot be obtained from an incompressible LES simulation.[1,17] It is also worthwhile mentioning that an option often found in the CAA literature is that of using a single FEM code, but with two different meshes, one for the flow dynamics and the other one for the acoustics. This is so because the former needs finer meshes than the second (an eddy at low Mach numbers typically generates sound with a wavelength much larger than its characteristic size). Resorting to such a procedure can save computational cost but requires, however, to interpolate between meshes and to smooth the acoustic source term as well, to avoid numerical instabilities when the considered computational fluid dynamics (CFD) domain is smaller than the acoustic one.[18]

In this work, the focus will be placed on the acoustic outputs of the strategy in Guasch et al,[1] rather than on the LES ones, which will be only described qualitatively together with the implemented FEM approach. It is to be noted that 3D LES simulations of flow passing around teeth-shaped obstacles to better understand the underlying physics of /s/ were already reported in Van Hirtum et al[19,20] and in Cisonni et al.[21] Simulations of flow passing through simplified geometries with constrictions of different sizes were also conducted in Cisonni et al.[22] The work in Nozaki[23] also resorted to LES to analyze the flow dynamics of /s/ in a realistic vocal tract geometry. As said before, more recently, Yoshinaga et al[10,11] presented simulations on a 3D realization of the geometry in Howe and McGowan[8] and compared it with realistic ones. On the other hand, the LES in the present work has been solved with the stabilized FEM method in Codina et al,[24] which behaves as an implicit LES model (see, eg, previous works[25,26]). In implicit LES methods, the additional terms included in the equations to avoid numerical instabilities simultaneously act as a turbulence model. This has been proven successful on well-known benchmark turbulence tests for the strategy in Codina et al[24] (see, eg, literature[27-29]), as well as through analytical reasoning.[30]

In what concerns the acoustic formulations to get the contributions from the direct flow noise in the lower incisors-lips cavity and from the upper incisors diffraction, one should ideally resort to approaches that could account for the unsteady flow acoustics in the vocal tract. The most relevant ones for that purpose are probably the linearized Euler equations (see, eg, Bailly and Bogey[31]), or some of its source-filtered counterparts to exclude the vorticity and entropy modes, like the acoustic perturbation equations (APEs), see Ewert and Schröder.[32] In Hueppe and Kaltenbacher,[33] a low Mach number version of the APE was introduced (see Guasch et al[34] for a full numerical solution retaining all terms). In the case of very low Mach number flows, like those encountered in speech production, the APE in Hueppe and Kaltenbacher[33] can be further simplified to the acoustic analogy in Roger.[35] The wave operator in this analogy is just the standard wave equation, like in Lighthill's analogy,[36] yet the double time derivative of the incompressible pressure field is used as the source term, instead of the double divergence of the Reynolds tensor of the incompressible velocity field. It can be shown that this allows the filtering of some *pseudosound* at the flow region (the term *pseudosound* refers to pressure fluctuations indistinguishable by a single microphone from proper sound; see, eg, Crighton et al[37]). In this piece of research, both Lighthill's and the analogy in Roger[35] will be employed.

To end this introductory section, we would like to remark that aside from static vowel sounds, for which an extensive literature is available (see, eg, literature[38-49]), few papers can be found addressing the numerical simulation of other speech

sounds. The reason for that is probably the complex physics beneath their generation and the associated high computational cost. An exception that has received some recent attention is that of vowel-vowel utterances in Guasch et al.[50] Also, attention has been paid to some unvoiced sounds (see, eg, Krane[51]) and in particular to fricative sounds. Related to the present work, we shall cite Anderson et al[52] where different CFD formulations were tested for fricative /ʃ/, using realistic two- (2D) and three-dimensional (3D) vocal tract geometries. That study compared the performance of a compressible LES, an incompressible LES, and a Reynolds-averaged Navier-Stokes (RANS) approach. Though a rather coarse mesh was used, the study reached some interesting conclusions. As expected, the RANS simulation provided no reliable results because it was unable to capture the rapid acoustic fluctuations, but the compressible LES and incompressible LES combined with an acoustic analogy yielded proper outputs. Another interesting and unexpected result was that although the flow field from 2D simulations did not match at all with the 3D one, that was not the case for the 2D acoustic pressure field, which was quite similar to the 3D one.

This paper is organized as follows. The methodology we followed to perform the simulations is detailed in Section 2, which includes a description of the realistic vocal tract geometry for sibilant /s/, the formulation of the acoustic analogies, and the splitting strategy between turbulent and diffracted sound. It also outlines the numerical approach used to solve the problem partial differential equations and includes specific details on how the numerical simulations were run. The results of the latter are presented in Section 3, with special emphasis on the characteristics of the generated aerodynamic sound. Conclusions close the paper in Section 4.

2 | METHODOLOGY

2.1 | Vocal tract model

The vocal tract geometry used for the simulations was obtained from the cone-beam computed tomography (CT) scan (CB MercuRay, 512 slices of 512 pixel × 512 pixel grid with accuracy ±0.1 mm) in Van Hirtum et al[53] and Fujiso et al,[54] from which a physical replica was constructed. The geometry in Figure 2 corresponds to the anterior portion of the vocal tract of an adult male Japanese native speaker with normal dentition (angle class I) without any speech disorder, in normal sitting position. The subject was instructed to sustain phoneme [s] at a medium loudness (with a flow rate approximately

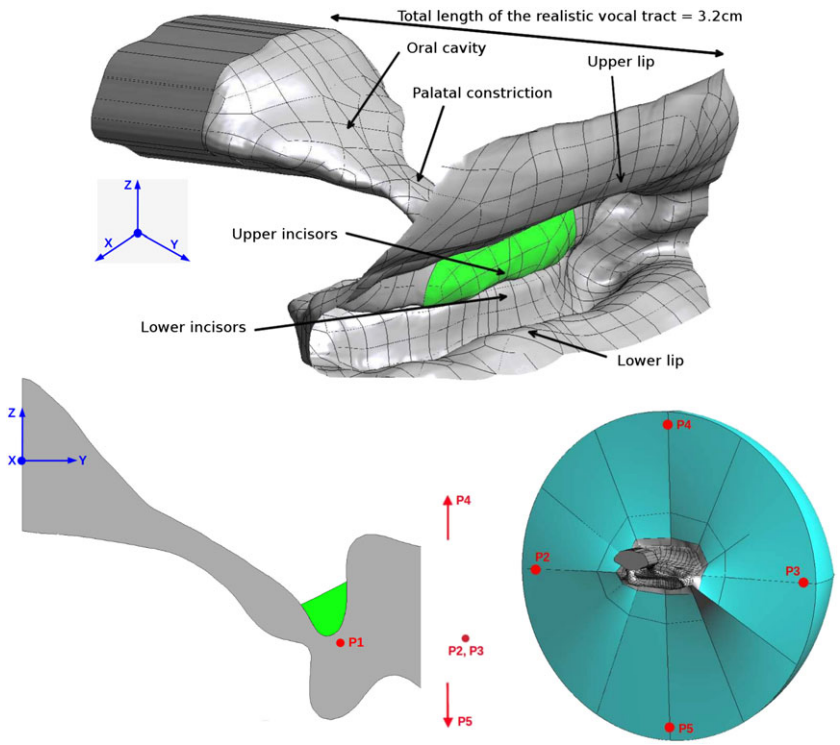

FIGURE 2 Three-dimensional realistic vocal tract geometry for /s/ with upper incisors highlighted in green. Location of the measuring points outside the flow region; *P*2 and *P*3 are placed along the x direction and *P*4 and *P*5 along z direction

21 L/min) during 10 seconds. The entire vocal tract was imaged, but only the portion containing the main tongue constriction and all the structures downstream of it, including the lip horn (approximately 32 mm), was reconstructed for the replica, see Figure 2 and dashed window in Figure 1. This simplification intends to focus on the region of the vocal tract where the generating mechanisms take place, as in Howe and McGowan.[8] This includes the constricted passage between the tongue blade and the hard palate (the section with minimum area has a hydraulic diameter of 2.1 mm), the lower and upper incisors (highlighted in green in the figure), and the lips. According to Van Hirtum et al,[53] the initial flow conditions upstream (geometry and Reynolds number) do have an influence on the modulation of the acoustic spectrum of fricatives but do not play an essential role in the physiological mechanisms that lead to the generation of this phoneme. The experimental mock-up has many similarities to that used in Nozaki et al,[2] so this reference case will be used in the qualitative validation of the numerical method.

2.2 | Problem formulation

2.2.1 | Acoustic analogies

As said in Section 1, in this work, the celebrated Lighthill acoustic analogy[36] and the analogy in Roger[35] will be used. For low Mach numbers, Lighthill's tensor can be well approximated by the double divergence of the Reynolds tensor of the incompressible velocity field. If one considers a computational domain Ω_v with outer boundary Γ_∞ and an embedded rigid body with boundary Γ_w and external normal $\boldsymbol{n}$, the Lighthill acoustic analogy to determine the aerodynamic noise generated by the flow past the body reads (see Figure 3)

$$\frac{1}{c_0^2}\frac{\partial^2 p}{\partial t^2} - \nabla^2 p = \rho_0 \left(\nabla \otimes \nabla\right) : \left(\boldsymbol{u}^0 \otimes \boldsymbol{u}^0\right) \text{ in } \Omega_v, \tag{1a}$$

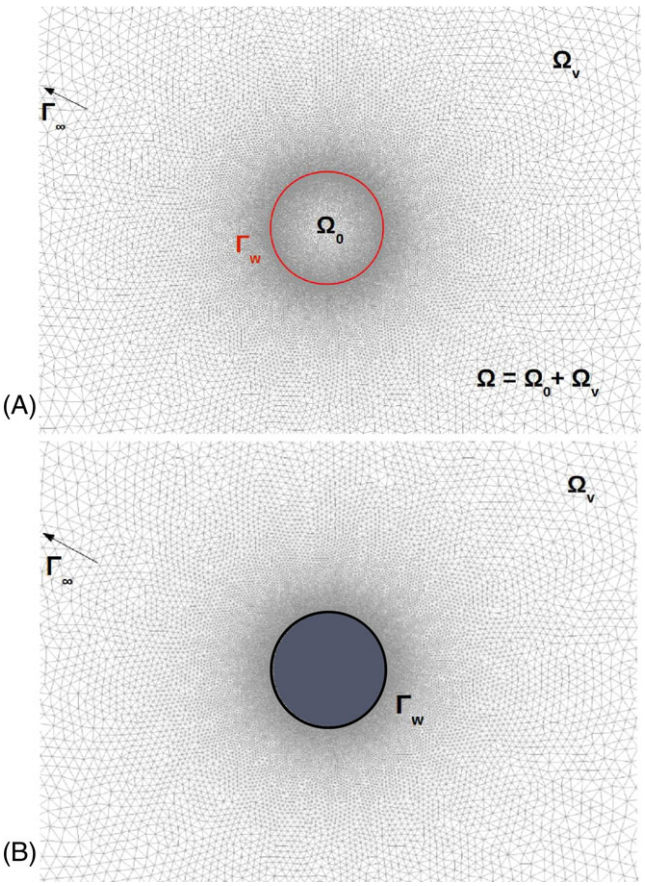

FIGURE 3 Domains for computing A, the incident acoustic pressure and B, the diffracted acoustic pressure in the proposed splitting strategy for aeroacoustics

$$\frac{\partial p}{\partial \boldsymbol{n}} = -\frac{1}{c_0}\frac{\partial p}{\partial t} \text{ on } \Gamma_\infty,\ t > 0, \tag{1b}$$

$$\frac{\partial p}{\partial \boldsymbol{n}} = 0 \text{ on } \Gamma_w,\ t > 0, \tag{1c}$$

$$p = 0,\ \frac{\partial p}{\partial t} = 0 \text{ in } \Omega_v,\ t = 0. \tag{1d}$$

In Equation 1, $p(\boldsymbol{x}, t)$ stands for the acoustic pressure fluctuations and $\boldsymbol{u}^0(\boldsymbol{x}, t)$ for the incompressible velocity field obtained, eg, from a CFD computation. ρ_0 stands for the flow density, c_0 for the speed of sound, and $\otimes$ for the tensor product. In the second line, (1b) introduces a Sommerfeld nonreflecting condition on Γ_∞, and in the third one, 1c expresses a rigid wall assumption for the immersed body. The initial conditions are set in (1d). The source term $\rho_0(\nabla \otimes \nabla) : \left(\boldsymbol{u}^0 \otimes \boldsymbol{u}^0\right)$ in Equation 1 is often rewritten as $\rho_0\left(\nabla \otimes \boldsymbol{u}^0\right) : \left(\nabla \otimes \boldsymbol{u}^0\right)^\top$ for computations ($\top$ denotes transpose), given that $\nabla \cdot \boldsymbol{u}^0 = 0$ (see, eg, Guasch and Codina[55]).

With regard to the acoustic analogy in Roger,[35] it can be obtained from the low Mach APEs in Hueppe and Kaltenbacher,[33] by simply neglecting the mean velocity field and combining the momentum and continuity equations to get the scalar wave equation for the acoustic pressure. Alternatively, the analogy was originally derived from the following straightforward reasoning. Taking the divergence of the Navier-Stokes momentum conservation equation results in the Poisson equation for the incompressible pressure, p^0, namely, $\nabla^2 p^0 = -\rho_0(\nabla \otimes \nabla) : \left(\boldsymbol{u}^0 \otimes \boldsymbol{u}^0\right)$, which allows one to replace Equation (1a) with

$$\frac{1}{c_0^2}\frac{\partial^2 p}{\partial t^2} - \nabla^2 p = -\nabla^2 p^0. \tag{2}$$

This equation is interesting for the following reason. It is well known that the acoustic pressure predicted by Lighthill's analogy is only valid far away from the source region, where no flow motion occurs. If one needs to determine the acoustic pressure field close to the generation area, it becomes necessary to filter out from p the *pseudosound* induced by the nonacoustic pressure fluctuations $p^0(\boldsymbol{x}, t)$. In the case of a flow with very small mean convection velocity, in Roger,[35] it was proposed to do so by splitting the pressure into its incompressible *pseudosound* and acoustic, $p^a(\boldsymbol{x}, t)$, components, ie, $p(\boldsymbol{x}, t) = p^0(\boldsymbol{x}, t) + p^a(\boldsymbol{x}, t)$. Inserting this factorization in Equation 2 results in the following wave equation for the acoustic pressure fluctuations

$$\frac{1}{c_0^2}\frac{\partial^2 p^a}{\partial t^2} - \nabla^2 p^a = -\frac{1}{c_0^2}\frac{\partial^2 p^0}{\partial t^2}, \tag{3}$$

which is to be supplemented with the boundary and initial conditions (1b) to (1d), now for $p^a(\boldsymbol{x}, t)$.

2.2.2 | Incident and diffracted contributions to the acoustic pressure

Let us consider the problem depicted in Figure 3, where a flow impinges a cylinder resulting in the generation of aerodynamic noise. We would like to separate the noise generated by the turbulent wake from its diffraction by the cylinder. To that purpose, one could typically resort, as said before, to Curle's formulation.[15] However, in the case of low Mach number flows, a severe difficulty appears when trying to account for the rigid body (ie, the cylinder or in our speech problem the incisors) contribution to the far field. The reason is that the surface integral in Curle's formulation involves the total pressure, which includes both the incompressible and acoustic fluctuations. Obviously, the latter cannot be obtained from an incompressible CFD simulation.

Though recently some proposals have been made to at least partially mitigate that problem in the framework of integral formulations (see, eg, Martínez-Lera et al[17]), in Guasch et al,[1] a very different approach was suggested. The latter considered that the acoustic dipole distribution of Curle's surface integral corresponds, in fact, to the diffraction of the turbulent noise generated by the jet flow (see, eg, Gloerfelt et al[56]). On the one hand, the proposed methodology circumvented the total pressure difficulty in Curle's surface term. On the other hand, it avoided the need to resort to integral formulations, so that one can obtain the flow field in the domain, the noise generated by the wake past the cylinder, and the noise contribution from the latter, in a single finite element computational run.

The cornerstone of the method in Guasch et al[1] consists in splitting the acoustic pressure, $p(\boldsymbol{x}, t)$, in Equation 1 (the same holds for $p^a(\boldsymbol{x}, t)$ in Equation 3), into incident and diffracted components $p(\boldsymbol{x}, t) = p_i(\boldsymbol{x}, t) + p_d(\boldsymbol{x}, t)$. This leaves one with two wave equations, one for $p_i(\boldsymbol{x}, t)$ and the other one for $p_d(\boldsymbol{x}, t)$, which are solved subsequently in slightly different domains (see Figure 3). The procedure goes as follows. Once an acoustic source term $s(\boldsymbol{x}, t)$ is obtained from an incompressible CFD computation, for instance,

$$s(\boldsymbol{x}, t) = \begin{cases} \rho_0 \left(\nabla \otimes \nabla\right) : \left(\boldsymbol{u}^0 \otimes \boldsymbol{u}^0\right) \\ -c_0^{-2} \quad \partial^2 p^0 / \partial t^2, \end{cases} \tag{4}$$

$s(\boldsymbol{x}, t)$ is used as the inhomogeneous term in the wave equation for the incident pressure component, namely,

$$\frac{1}{c_0^2} \frac{\partial^2 p_i}{\partial t^2} - \nabla^2 p_i = s \text{ in } \Omega, \tag{5a}$$

$$\frac{\partial p_i}{\partial \boldsymbol{n}} = -\frac{1}{c_0} \frac{\partial p_i}{\partial t} \text{ on } \Gamma_\infty, \ t > 0 \tag{5b}$$

$$p_i = 0, \ \frac{\partial p_i}{\partial t} = 0 \text{ in } \Omega, \ t = 0, \tag{5c}$$

with $\Omega := \Omega_v \cup \Omega_0$, Ω_0 being the volume occupied by the rigid body (see Figure 3A). In other words, the problem is solved as if the rigid body (the cylinder in the figure) was absent. After having computed $p_i(\boldsymbol{x}, t)$, the rigid body is inserted again into the computational domain, which leaves one with Ω_v (see Figure 3B). The diffracted pressure $p_d(\boldsymbol{x}, t)$ is then obtained knowing the value of the incident pressure on the boundary Γ_w of the rigid body. That is to say solving

$$\frac{1}{c_0^2} \frac{\partial^2 p_s}{\partial t^2} - \nabla^2 p_s = 0 \text{ in } \Omega_v, \tag{6a}$$

$$\frac{\partial p_s}{\partial \boldsymbol{n}} = -\frac{\partial p_i}{\partial n} \text{ on } \Gamma_w, \ t > 0, \tag{6b}$$

$$\frac{\partial p_s}{\partial \boldsymbol{n}} = -\frac{1}{c_0} \frac{\partial p_s}{\partial t} \text{ on } \Gamma_\infty, \ t > 0, \tag{6c}$$

$$p_s = 0, \ \frac{\partial p_s}{\partial t} = 0 \text{ in } \Omega_v, \ t = 0. \tag{6d}$$

Note that the summation of problems Equations (5) and (6) recovers the original Lighthill analogy in Equation (1) or that in Equation 3, depending on the selected source term in (4).

When applied to the production of /s/, the incisors will play the role of the cylinder. Then, Equation (5) will provide the acoustic contribution $p_i(\boldsymbol{x}, t)$ from the jet flow exiting the mouth and being modulated by the vocal tract articulators (others than the upper incisors) to the total sibilant sound. Hereafter, we will refer to this acoustic contribution as the *incident* acoustic field. Besides, the contribution from the aerodynamic sound diffracted by the upper incisors will be given by the solution $p_s(\boldsymbol{x}, t)$ to Equation (6) and we will term it the *diffracted* field. Let us note that in this work, the factorization into incident and diffracted components has been only performed for Lighthill's analogy. No splitting into components has been carried out for the alternative acoustic analogy in Equation 3, albeit that was possible too.

2.3 | Numerical strategy

All the partial differential equations in this work were solved using the method of lines; ie, the spatial discretization was carried out with the FEM, while the finite difference method was adopted for the time discretization. A custom developed software was used in all computations.

As regards the incompressible Navier-Stokes equations, it is well known that the Galerkin FEM solution suffers from strong instabilities for convection dominated flows and for small time steps at the beginning of evolutionary processes. Moreover, the spatial discrete problem has to satisfy an inf-sup compatibility condition that implies the use of different

interpolation spaces for the incompressible pressure and velocity fields. An efficient way to circumvent all these difficulties is to resort to residual-based stabilized variational multiscale (VMS) methods, see previous studies.[57,58] The unknown variables in the problem weak form become split into large components, resolvable by the finite element mesh, and subgrid scales whose effects onto the large ones must be modeled. In this work, the subscales have been chosen orthogonal to the finite element space and the stabilization parameters have been obtained from a Fourier analysis of the subgrid scale equation, see the orthogonal subgrid scale (OSS) method in Codina et al[24] and Codina.[59] The OSS stabilization procedure allows one to use equal interpolation spaces for the pressure and velocity unknowns (linear P1/P1 elements were employed in our simulations). Besides, and as mentioned in Section 1, the OSS method also acts as an implicit LES model. As regards time discretization, an implicit third-order backward differentiation formula was implemented. In this way, one can sidestep the constraints of the CFL condition and dimension the time step according to the highest frequency to be resolved and the intrinsic time step dictated by the stabilization parameter. We finally note that the stabilized variational Navier-Stokes equations for the problem at hand have been solved using the second-order fractional step scheme presented in Codina and Badia.[60]

At each discrete time step of the computation, the acoustic source terms in Equation 4 were computed, in the same computational mesh, by postprocessing the incompressible pressure and velocity output from the fluid dynamics computation. In the case of Lighthill's analogy, the acoustic waves for the incident and diffracted acoustic pressure, Equations 5–6, were then solved. As mentioned before, the contribution analysis was not performed for the analogy in Roger[35] as it would had yielded very similar results to those of Lighthill (see Section 3.2 below). Therefore, only Equation 3 was solved in this case.

From a computational point of view, the spatial discretization of the wave equation poses no particular problem given that the Laplacian of the acoustic pressure gives rise to a coercive term in the variational form of the problem. Therefore, the main difficulty with the numerical solution of the wave equation arises from the time discretization, which should prevent numerical dissipation as a wave propagates. Given that the focus in this work is on the aerodynamic acoustic pressure at points only a few wavelengths from the mouth, a second-order BDF2 scheme proves accurate enough to approximate the time derivatives in the wave equation and its boundary conditions. Another issue to consider is that of imposing a proper nonreflecting condition at the outer boundary of the computational domain. This often requires the use of a perfectly matched layer (PML) formulation. However, the outward propagating waves in the present simulations are nearly spherical and impinge the radiation boundary in the normal direction, which makes a PML unnecessary. Spurious reflections were avoided by simply imposing the Sommerfeld condition in Espinoza et al,[61] which has been tested in a wide range of problems and geometries.

Interested readers are referred to Guasch et al[1] and references therein for a fully detailed description of the numerical strategy we have just outlined.

2.4 | Numerical simulations

To perform the numerical simulations, the vocal tract in Figure 2 was set in a circular, rigid, flat baffle. A hemispherical domain was attached to it to allow the flow to emanate from the mouth and the acoustic waves to propagate outwards (see Figure 4A).

The following boundary conditions were applied to solve the incompressible Navier-Stokes equations. A velocity of $(0, 2.4, 0)$ m/s was prescribed at the inlet section. This was scaled from the blower in Van Hirtum et al,[53] where the role of laminar/turbulent initial and inlet flow conditions is discussed in detail, and corresponds to a Reynolds number of $\mathrm{Re} = 8850$, according to the inlet section diameter. No turbulence was prescribed at the inlet. Nonslip conditions were applied to the whole vocal tract and baffle surfaces, while the hemispherical surface was considered as an open boundary.[21] As regards the acoustic computations, the vocal tract and baffle were assumed rigid, whereas as mentioned, a nonreflecting boundary condition was applied at the hemispherical boundary and at the flow inlet. The first-order Sommerfeld boundary condition in Espinoza et al[61] did not led to any spurious reflection in the present example. The following values were taken for the physical parameters appearing in the equations: an air density of $\rho_0 = 1.2\mathrm{kg/m}^3$, a kinematic viscosity of $\nu = 1.5 \times 10^{-5}\ \mathrm{m}^2/\mathrm{s}$, and a sound speed of $c_0 = 343$ m/s.

At low Mach numbers, M, an eddy of characteristic size l essentially radiates sound of wavelength $\lambda \sim \mathcal{O}(l/M)$. Taking into account that, according to the inlet flow speed, the computed Mach number is $M \sim 0.007$ (though it can locally reach values up to $M \sim 0.23$) and that frequencies up to 12 kHz are to be captured to reproduce the physics of /s/, very fine spatial meshes are required for the simulations. Therefore, a computational mesh of 45 million linear tetrahedral elements with equal interpolation for all variables (approximately 36 million degrees of freedom for the CFD problem and approximately

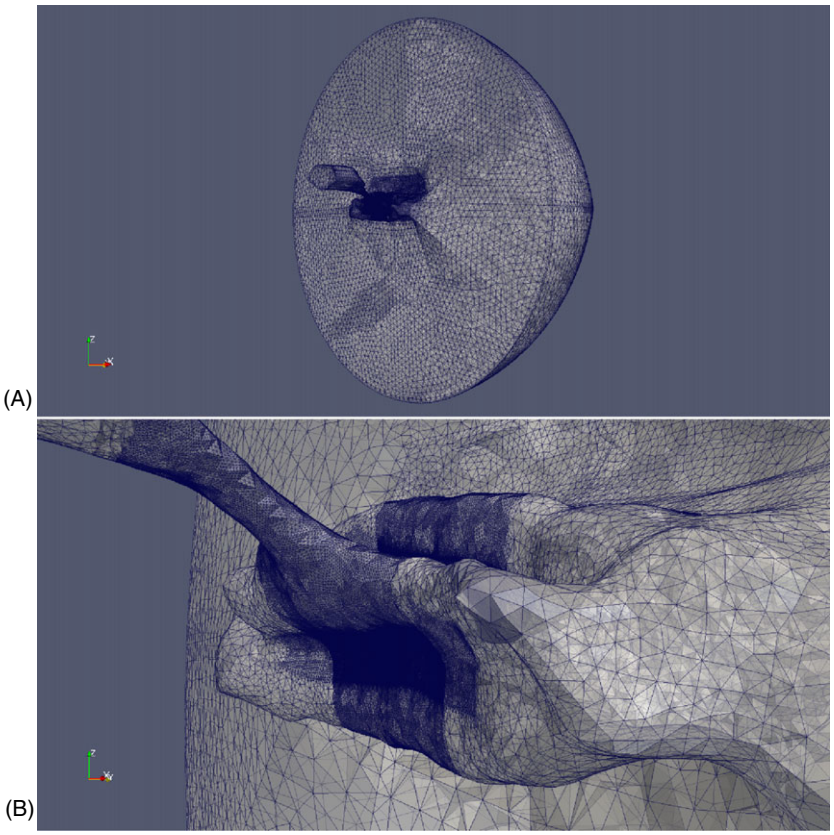

FIGURE 4 A, Meshed computational domain including the vocal tract geometry and the far field. B, Zoom of the refined mesh region

9 million for the acoustics) was used in this work (see Figure 4A,B for a general view and mesh refinement details). The mesh size ranges from $h = 0.025$ mm close to the incisors, where the smallest turbulent flow scales are expected, to $h = 2.5$ mm outside the mouth. The latter guarantees having about 12 nodes per wavelength at 12 kHz ($\lambda \sim 30$ mm). Obviously, more optimal meshes could have been obtained by resorting to adaptive strategies (see, eg, Kannan et al[62] for an example in the biological context), though this issue lies out of the scope of the current paper.

As explained in the previous section, a total of three equations are to be solved in the same finite element computational run (four if the acoustic analogy in Roger[35] is chosen). Performing such computations with very large models is complex. To that purpose, a domain decomposition with an message passing interface (MPI) distributed memory scheme was carried out so as to run the problem at the MareNostrum computer cluster, of the Barcelona Supercomputing Centre (BSC). A period of 10.8 milliseconds with a time step of $\Delta t = 5 \times 10^{-6}$ seconds (dimensioned according to the highest audible frequency) was simulated. This sufficed to reach a statistical stationary state. The computational run lasted approximately 30 hours using 256 processors. A Biconjugate Gradients solver with Pilut preconditioner of the Hypre library was used to solve the FEM algebraic matrix systems, all of them integrated in PETSc.[63]

3 | RESULTS

3.1 | Flow field and acoustic sources

The qualitative results from the CFD simulation confirm the general theoretical framework describing the mechanisms of /s/ generation and provide some further insight to it as well. The jet flow is strongly accelerated in the constriction between the tongue blade and the hard palate, where the flow diverts (see Figure 5A). This can also be appreciated from the flow speed profile in Figure 5B (the cut plane is indicated with a dashed line in Figure 5A), where it can be observed how the tongue closes the tract in the transverse direction, forcing the flow to concentrate in a small region before impacting the upper incisors. The flow is firmly accelerated again at the interteeth space reaching local Mach numbers up to $M \sim 0.2$. Eddies are shed past the upper incisors and impinge on the lower lips. As a consequence, a strong turbulent flow develops

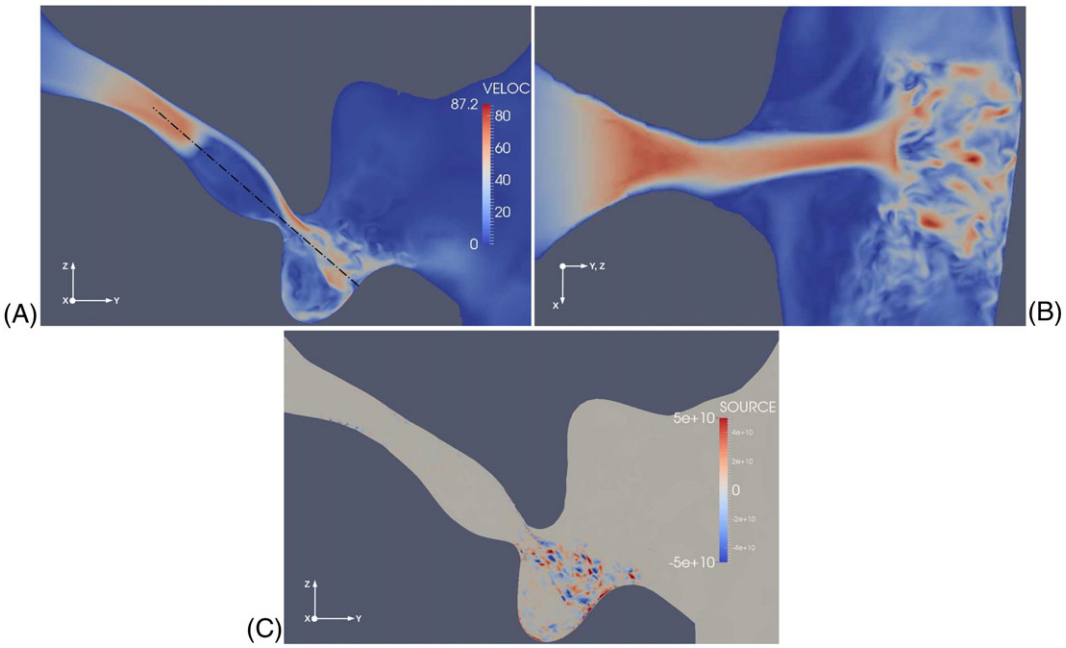

FIGURE 5 Snapshot of A, flow speed profile in (m/s) on the vertical midplane cut; B, flow speed profile on a plane tangent to the tract; and C, Lighthill's acoustic source term at $t = 0.0108$ second in $\mathrm{kgm^{-3}s^{-2}}$

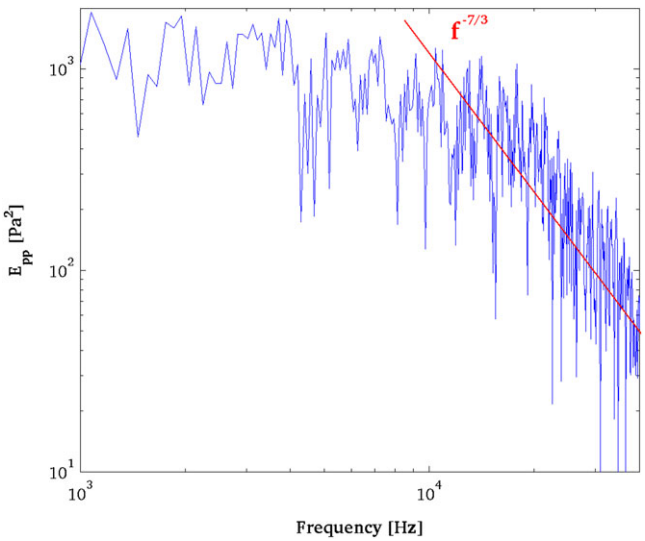

FIGURE 6 Spectrum of the incompressible pressure at a near-field point showing the right $f^{-7/3}$ dependence at the turbulent inertial subrange

in the cavity between the lower incisors and the lower lip, where most sound sources concentrate. This is shown in Figure 5C where a snapshot of Lighthill's source term, $\rho_0 (\nabla \otimes \nabla) : \left(\boldsymbol{u}^0 \otimes \boldsymbol{u}^0\right)$, is presented. The strongest quadrupole sources can be found in the direct path between the edge of the upper incisors, where flow separation takes place, and the top side of the lower lips.

To check that the LES is able to capture all turbulent scales down to the inertial subrange, the incompressible pressure spectrum, E_{pp}, has been plotted for an arbitrary point with coordinates $P1 = (0, 0.0205, -0.012)^{\mathsf{T}}$, located just in front of the upper incisors, see Figure 6. The origin of coordinates is placed at the center of the flat section leading to the realistic vocal tract geometry in Figure 2. According to Kolmogorov's theory for isotropic turbulence, the energy spectrum at the inertial subrange behaves as $E \sim k^{-5/3}$, while the pressure spectrum behaves as $E_{pp} \sim k^{-7/3}$ (see, eg, Pope[64]). Taylor's hypothesis of frozen turbulence allows one to show that the latter also exhibits the same power dependence with frequency, namely, $E_{pp} \sim f^{-7/3}$. Note that, at the near field, the assumptions of fully developed isotropic turbulence do

not apply due to the presence of walls. However, it is observed in Figure 6 that the present CFD simulation manages to reproduce the behavior of the pressure spectrum at the inertial subrange of a fully developed turbulent flow, therefore showing that the stabilized FEM formulation, combined with the fine computational mesh, corresponds to a large eddy simulation approach.

Quantitative details of the flow dynamics simulation are left out of the scope of this paper. The reader is referred to previous studies,[19-23] for some works dealing with them. In contrast and as stated in Section 1, few papers seem to exist reporting numerical results on the aeroacoustics of /s/. Therefore, this will be the focus of the next subsections and the main output from this work.

3.2 | Incident and diffracted sound contributions to the total acoustic field

The quadrupole sources depicted in Figure 5B directly radiate sound which propagates outside the mouth. The emitted sound is modulated by the speech articulators, and in particular, that diffracted by the upper incisors is believed to be the predominant contributor to the acoustic far field.[8] However, it will be shown that the incident sound component cannot be neglected for points in the vicinity of the mouth.

The numerical strategy in Section 2.2 has been applied to compute the contributions from the incident and diffracted sound. For illustrative purposes, however, before proceeding with a quantitative analysis of the results, we present a

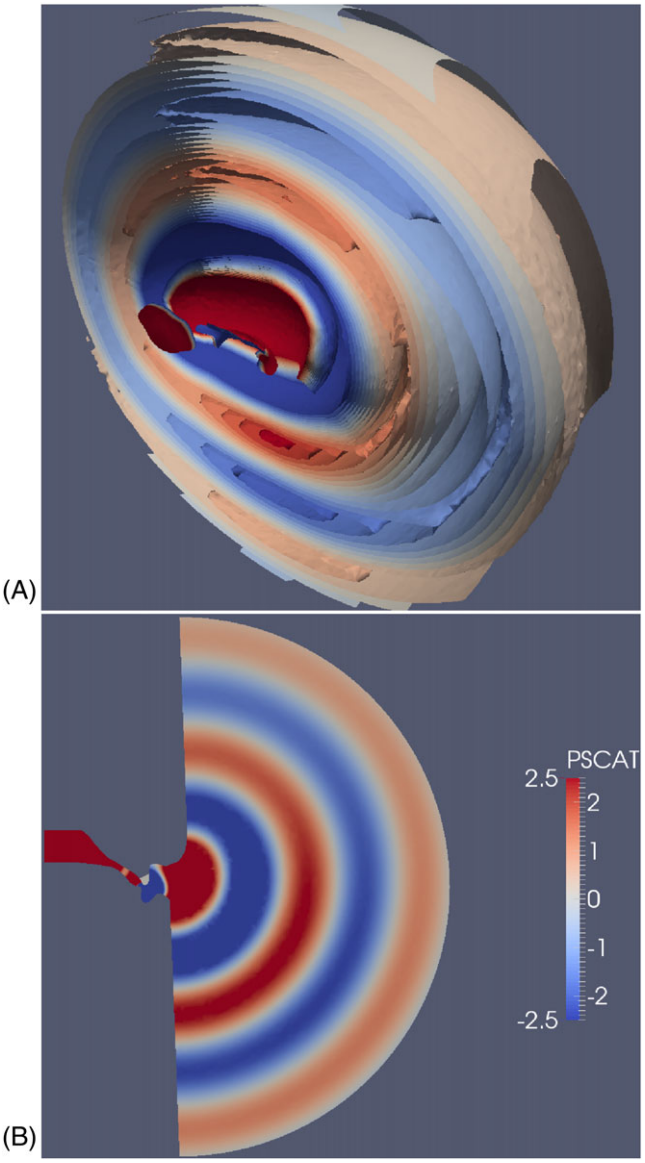

FIGURE 7 A, Total acoustic pressure isosurfaces and B, midcut showing diffracted pressure wavefronts. Units in Pa

snapshot of the spherical acoustic wavefronts emanating from the mouth in Figure 7A, whereas a midcut showing the propagation of the acoustic waves resulting from the incisors' diffraction is detailed in Figure 7B.

Figures 8 and 9 contain the main outputs of the simulation. In Figure 8, we present time occurrences of the acoustic pressure at two points located at different distances from the mouth exit. The first one is point $P1$, which is affected by the flow emanating from the mouth and was already introduced in the previous Section 3.2. The second one, $P5$, is not influenced by the flow and is placed at $P5 = (0, 0.035, -0.085)^{\mathsf{T}}$ (see Figure 2). Figure 8A plots the incident and diffracted contributions to point $P1$ using Lighthill's acoustic analogy. As known, Lighthill's analogy is not valid for points in the source region so the strong fluctuations from the incident contribution in Figure 8A (dashed blue line) correspond to pseudosound rather than to acoustic fluctuations. These are one order of magnitude higher than the acoustic contribution from the incisors' diffraction (red continuous line). The contributions for point $P5$ are presented in Figure 8B. In this case, Lighthill's acoustic analogy is perfectly valid to compute the generated aerodynamic acoustic pressure. As observed, at $P5$, the acoustic pressure from the diffracted component is clearly higher than that provided by the direct turbulent flow contribution. In Figure 8C, a comparison is presented between the contribution of the incisors' diffraction to point $P1$ (dashed blue line), already plotted in Figure 8A, and the total acoustic pressure (continuous red line) computed at the same point with the acoustic analogy in Roger,[35] instead of Lighthill's one. Both contributions have amplitudes of the same order showing that the analogy in Equation 3 is capable of filtering the pseudosound and extracting the acoustic component from the flow pressure fluctuations. When evaluated at a point outside the acoustic source region like $P5$, Lighthill's analogy in Equation 1a and the analogy in Equation 3 should yield almost the same results. This is what is observed in Figure 8D where the time evolution for the total acoustic pressure at $P5$ using both analogies is plotted. The two curves in the figure show very similar trends in terms of amplitude and phase.

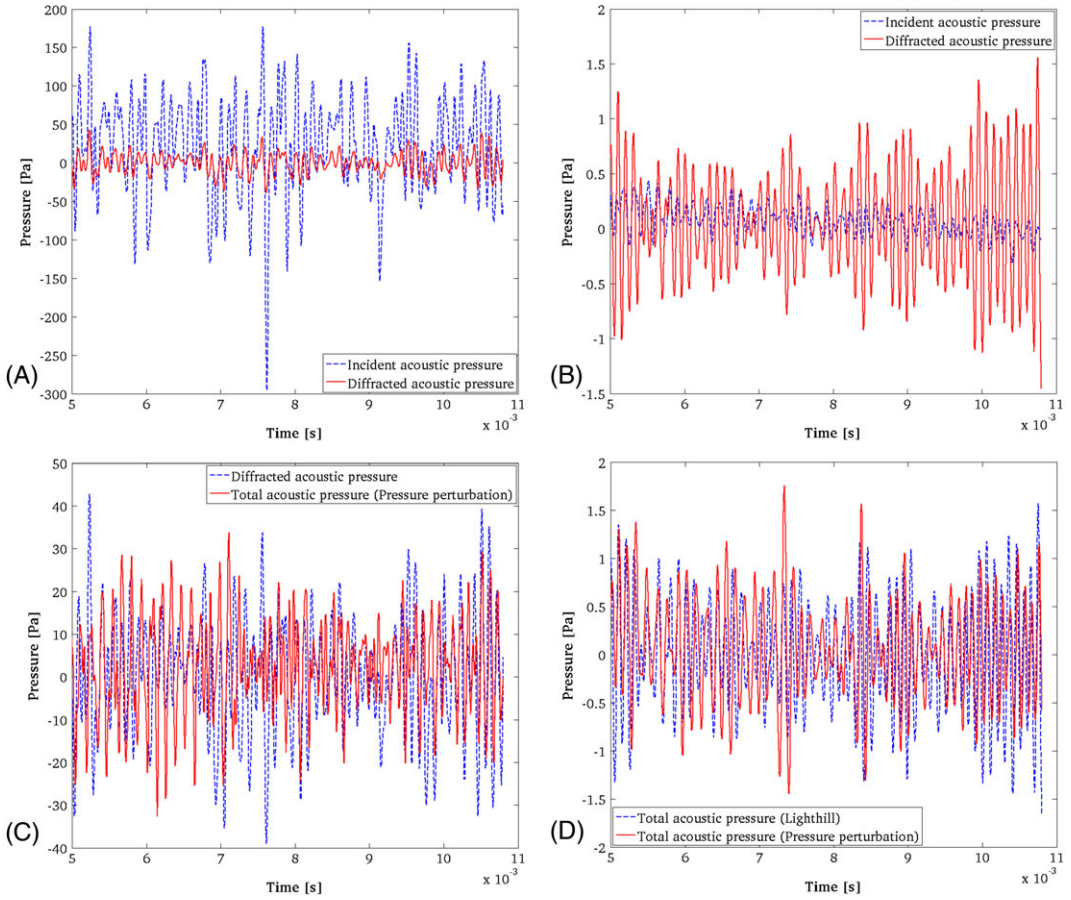

FIGURE 8 Time evolution of the acoustic pressure. A, Incident pseudosound (dashed blue) and diffracted (continuous red) components at the near-field point $P1$ from Lighthill's analogy; B, incident turbulent (dashed blue) and incisor diffracted (continuous red) contributions to the far-field point $P5$; C, incisors' diffracted contribution using Lighthill's analogy (dashed blue) versus total acoustic pressure (continuous red) using the analogy in Roger[35] at the near field; and D, total acoustic pressure at the far field using Lighthill's analogy (dashed blue) and the analogy in Roger[35] (continuous red)

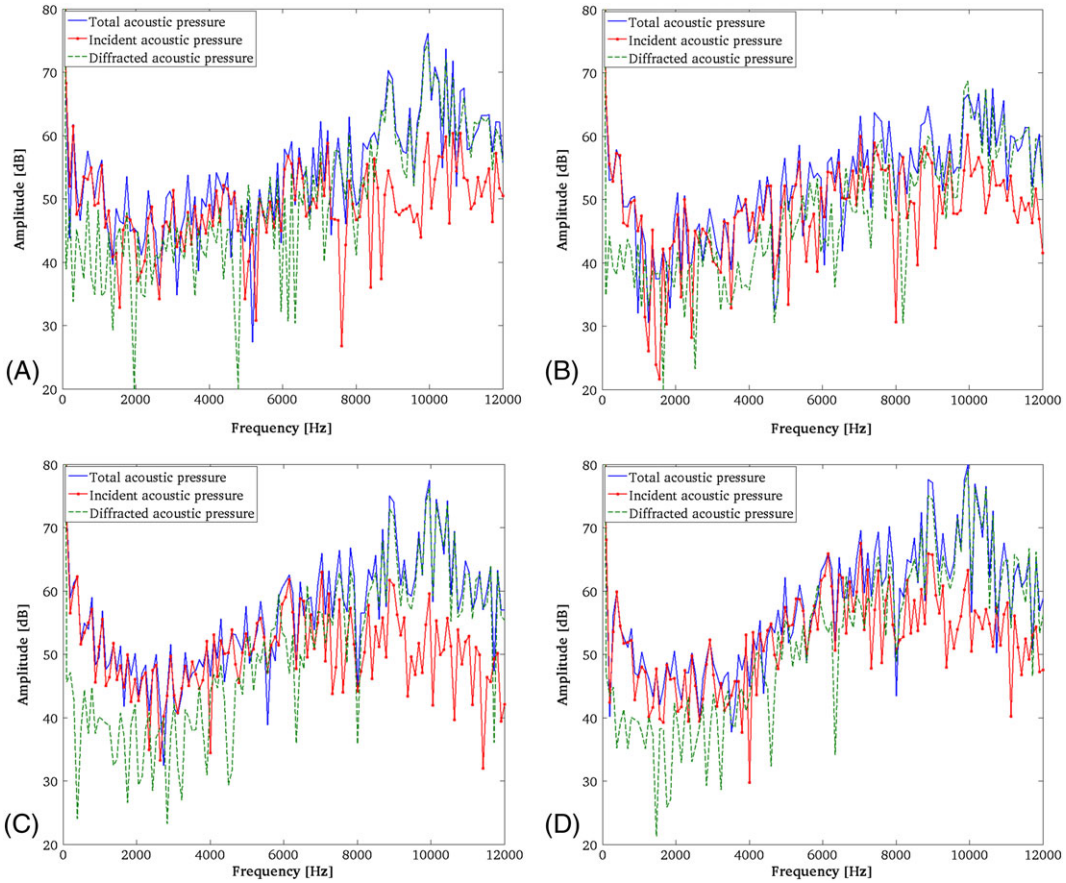

FIGURE 9 Spectra of the incident, diffracted, and total components of the acoustic pressure at the far field. A, Point *P*2; B, point *P*3; C, point *P*4; and D, point *P*5

On the other hand, the spectra in decibel (ref. 2×10^{-5} Pa) at four points well separated from the flow exiting the mouth are plotted in Figure 9. Figure 9A-D respectively corresponds to the following points: $P2 = (-0.075, 0.035, -0.0125)^{\top}$, $P3 = (0.075, 0.035, -0.0125)^{\top}$, $P4 = (0, 0.035, 0.06)^{\top}$, and *P*5, see Figure 2.

The figures contain the total acoustic pressure at these points, together with the separate contributions from the incident aerodynamic noise and from the sound diffracted by the incisors. As regards the total acoustic pressure, very similar tendencies can be appreciated in all figures. The spectra first decrease with frequency and exhibit a dip close to 2 kHz; then they increase linearly showing two marked peaks close to 9 and 10 kHz and finally decrease again with frequency. In what concerns the contributions, it can be checked that in the range up to 2 kHz, the incident component clearly dominates. The size of the incisors is small compared with the acoustic wavelengths at those frequencies so diffraction is not so important. However, as the frequency increases, the situation changes and both the incident and diffracted components contribute almost the same from 2 to approximately 8 kHz, with a slight predominance of the former. Above approximately 8 kHz, the incident contribution begins to decay and the spectra become mostly driven by the diffracted component. The diffracted sound is of a dipolar nature and radiates more efficiently than the quadrupolar noise. At several wavelengths from the incisors, the dipolar component will be responsible of the overall sound.[8]

Another remarkable result from Figure 9 is that the two peaks close to 9 and 10 kHz are essentially captured by the diffracted component of the acoustic pressure, for which they must be attributed to the influence of the upper incisors. This conclusion is supported by Yoshinaga et al,[12] where a monopole source is placed in the incisors gap near the upper incisors of a simplified vocal tract geometry for reproducing the characteristic spectral shape of the fricative /s/, while accounting for the propagation of higher-order acoustic modes.

3.3 | Comparison with measurements in literature

The spectra for the total acoustic pressure in Figure 9A-D basically exhibit the same spectral trends as those described in the Introduction of Nozaki et al,[2] which correspond to experimental recordings obtained with a vocal tract mechanical

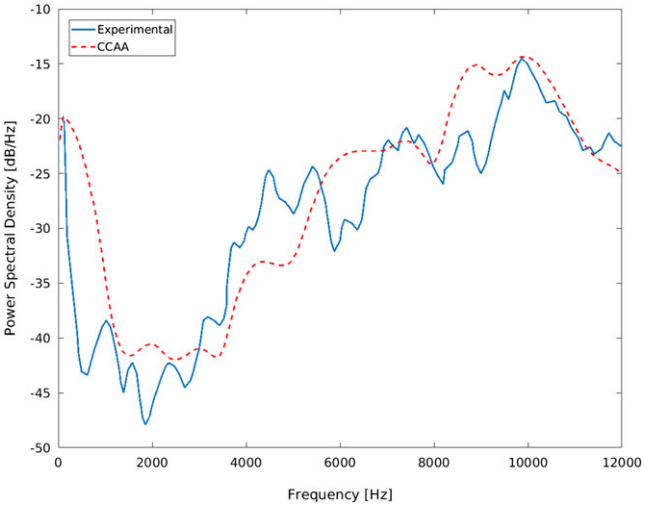

FIGURE 10 Welch averaged power spectral density level at point *P*3. Simulation versus experimental data in Nozaki et al[2] (different boundary and initial conditions were used in simulations and experiments)

replica very similar to that used in our simulations. The level of the Welch averaged power spectral density at *P*3 has been plotted, after proper scaling, in Figure 10 and compared with that in Nozaki et al.[2] Even though the initial and boundary conditions were different in the computations and experiments, the numerical simulation is able to reproduce all relevant characteristics of the acoustic spectrum in Nozaki et al[2]: The lowest amplitude is found at approximately 2 kHz after a steep fall, which is followed by a sudden growth up to a plateau around 6 kHz. Beyond this region, the main peak is found at 10 kHz, and for higher frequencies, the slope becomes negative. Moreover, the amplitude range between the peak and the valley frequencies is approximately 30 dB in both cases. The essentials of sibilant /s/ (see Jesus and Shadle[4]) are thus well recovered, namely, the dip at mid-frequencies followed by a positive spectral slope, the frequency position of the spectral peaks and the dynamic amplitude. More involved comparisons with experiments, like the directivity patterns in Yoshinaga et al.[65], are left for future developments.

4 | CONCLUSIONS

In this work, a large-scale numerical simulation of the aeroacoustics of a single example of sibilant fricative /s/ has been carried out. Lighthill's acoustic analogy and an analogy that allows filtering pseudosound to some extent at the source region have been implemented. A stabilized FEM that acts as an implicit large eddy simulation model has been used to solve the incompressible Navier-Stokes equations. The acoustic field has been resolved resorting to a splitting strategy that allows one to distinguish the acoustic contribution of the diffracted sound by the upper incisors from the incident one, in a single computational run.

The spectra from the contributions to the total acoustic pressure at points located close to the mouth exit (yet out of the flow wake) reveal a significant different behavior depending on the frequency range. At the lowest side of the spectrum, the acoustic pressure level decreases and the incident sound directly dominates. This is the case up to approximately 2 kHz. From approximately 2 to approximately 8 to 9 kHz, the level strongly increases and the contributions from the incident sound and its diffraction by the upper incisors are very similar. At higher frequencies, the diffracted sound becomes dominant and the shape of the upper teeth region seems to be physically responsible for the peaks in that frequency range. This finding is expected to impact speech production studies, since it provides a potential explanation for spectral features up to 12 kHz. Though one should bear in mind that the results reported in this work correspond to a single vocal tract geometry extracted from a subject while uttering a particular realization of /s/ and therefore lack of general validity, comparisons with subject recordings in the specialized literature seem to confirm the main trends of the obtained results.

ACKNOWLEDGMENTS

The authors gratefully acknowledge K. Nozaki (Osaka University) for providing the realistic geometry of sibilant /s/ used in the simulations. The authors also thankfully acknowledge the computer resources, technical expertise, and assis-

tance provided by the Red Española de Supercomputación (RES-BSC) and by the Swedish National Infrastructure for Computing (SNIC) at the PDC Centre for High Performance Computing (PDC-HPC). This work has been partially supported by the EU-FET grant EUNISON 308874. Moreover, the first author would also like to thank the Agència de Gestió d'Ajuts Universitaris i de Recerca for the predoctoral FI grant no. 2015 FI-B 00227. The second author would like to acknowledge the project GENIOVOX (TEC2016-81107-P) and the Secretaria d'Universitats i Recerca del Departament d'Economia i Coneixement (Generalitat de Catalunya) under grant refs. 2014-SGR-0590 and 2016-URL-IR-013. The third author acknowledges the support of the Spanish Government through the Ramón y Cajal grant RYC-2015-17367. The fourth author gratefully acknowledges the support received from the Catalan Government through the ICREA Acadèmia Research Program. The fifth author thanks the support from the ArtSpeech project (ANR-15-CE23-0024).

ORCID


Arnau Pont http://orcid.org/0000-0001-8358-2111

REFERENCES

1. Guasch O, Pont A, Baiges J, Codina R. Concurrent finite element simulation of quadrupolar and dipolar flow noise in low Mach number aeroacoustics. *Comput Fluids*. 2016;133:129-139.
2. Nozaki K, Yoshinaga T, Wada S. Sibilant/s/simulator based on computed tomography images and dental casts. *J Dent Res*. 2014;93, 2:207-211.
3. Nittrouer S, Studdert-Kennedy M, McGowan RS. The emergence of phonetic segments evidence from the spectral structure of fricative-vowel syllables spoken by children and adults. *JSLHR 32*. 1989;1:120-132.
4. Jesus LMT, Shadle CH. A parametric study of the spectral characteristics of European Portuguese fricatives. *J Phonetics 30*. 2002;3:437-464.
5. Shadle CH. The effect of geometry on source mechanisms of fricative consonants. *J Phonetics*. 1991;19:409-424.
6. Badin P. Fricative consonants: acoustic and x-ray measurements. *J Phonetics*. 1991;19(3–4):397-408.
7. Narayanan SS, Alwan AA, Haker K. An articulatory study of fricative consonants using magnetic resonance imaging. *J Acoust Soc Am 98*. 1995;3:1325-1347.
8. Howe MS, McGowan RS. Aeroacoustics of [s]. *Proc R Soc A*. 2005;461, 2056:1005-1028.
9. Howe MS. *Acoustics of Fluid-Structure Interactions*. Cambridge, UK: 204–231. Cambridge University Press; 1998.
10. Yoshinaga T, Koike N, Nozaki K, Wada S. Study on production mechanisms of sibilants using simplified vocal tract model. In: INTER-NOISE and NOISE-CON Congress and Conference Proceedings, **250, 1**, Institute of Noise Control Engineering; 2015; Reston (VA), USA:5662-5669.
11. Yoshinaga T, Nozaki K, Wada S. A relationship between simplified and realistic vocal tract geometries for Japanese sibilant fricatives. In: ISSP 2017 - The 11th International Seminar on Speech Production October, 16-19; 2017; Tianjin, China.
12. Yoshinaga T, Van Hirtum A, Wada S. Multimodal modeling and validation of simplified vocal tract acoustics for sibilant /s/. *J Sound Vib*. 2017;411:247-259.
13. Gracco VL, Lofqvist A. Speech motor coordination control: evidence from lip, jaw, and laryngeal movements. *J Neuroscience*. 1994;14, 11:6585-6597.
14. Bailly C, Bogey C, Gloerfelt X. Some useful hybrid approaches for predicting aerodynamic noise. *C R Mec*. 2005;333, 9:666-675.
15. Curle N. The influence of solid boundaries upon aerodynamic sound. *Proc R Soc Lond A*. 1955;231, 1187:505-514.
16. Ffowcs-Williams JE, Hawkings DL. Sound generation by turbulence and surfaces in arbitrary motion. *Phil Trans Roy Soc A*. 1969;264, 1151:321-342.
17. Martínez-Lera P, Schram C, Bériot H, Hallez R. An approach to aerodynamic sound prediction based on incompressible-flow pressure. *J Sound Vib*. 2014;333, 1:132-143.
18. Martínez-Lera P, Schram C. Correction techniques for the truncation of the source field in acoustic analogies. *J Acoust Soc Am*. 2008;124, 6:3421-3429.
19. Van Hirtum A, Grandchamp X, Cisonni J, Nozaki K, Bailliet H. Numerical and experimental exploration of flow through a teeth-shaped nozzle. *Adv Appl Fluid Mech*. 2012;11:87-117.
20. Van Hirtum A, Grandchamp X, Pelorson X, Nozaki K, Shimojo S. LES and in vitro experimental validation of flow around a teeth-shaped obstacle. *Int J Appl Mech*. 2010;2, 2:265-279.
21. Cisonni J, Nozaki K, Van Hirtum A, Grandchamp X, Wada S. Numerical simulation of the influence of the orifice aperture on the flow around a teeth-shaped obstacle. *Fluid Dyn Res*. 2013;45, 2:025505.
22. Cisonni J, Nozaki K, Van Hirtum A, Wada S. A parameterized geometric model of the oral tract for aero acoustic simulation of fricatives. *Int J Info Elec Eng*. 2011;1, 3:223.
23. Nozaki K. *Numerical Simulation of Sibilant [s] Using the Real Geometry of a Human Vocal Tract*, High Performance Computing on Vector Systems. Heidelberg, Germany: Springer-Verlag; 2010;137-148.
24. Codina R, Principe J, Guasch O, Badia S. Time dependent subscales in the stabilized finite element approximation of incompressible flow problems. *Comput Methods Appl Mech Eng*. 2007;196, 21:2413-2430.

25. Gravemeier V. The variational multiscale method for laminar and turbulent flow. *Arch Comput Methods Eng*. 2006;13,2:249-324.
26. Sagaut P. *Large Eddy Simulation For Incompressible Flows: An Introduction*. Heidelberg, Germany: 1–290, Springer-Verlag; 2006.
27. Colomés O, Badia S, Codina R, Principe J. Assessment of variational multiscale models for the large eddy simulation of turbulent incompressible flows. *Comput Methods Appl Mech Eng*. 2015;285:32-63.
28. Principe J, Codina R, Henke F. The dissipative structure of variational multiscale methods for incompressible flows. *Comput Methods Appl Mech Eng*. 2010;199, 13:791-801.
29. Massarotti N, Nithiarasu P, Codina R, Principe J, Ávila M. Finite element approximation of turbulent thermally coupled incompressible flows with numerical sub-grid scale modelling. *Int J Num Meth Heat Fluid Flow*. 2010;20, 5:492-516.
30. Guasch O, Codina R. Statistical behavior of the orthogonal subgrid scale stabilization terms in the finite element large eddy simulation of turbulent flows. *Comput Methods Appl Mech Eng*. 2013;261:154-166.
31. Bailly C, Bogey C. Contributions of computational aeroacoustics to jet noise research and prediction. *Int J Comput Fluid Dyn*. 2004;18, 6:481-491.
32. Ewert R, Schröder W. Acoustic perturbation equations based on flow decomposition via source filtering. *J Comput Phys*. 2003;188, 2:365-398.
33. Hueppe A, Kaltenbacher M. Spectral finite elements for computational aeroacoustics using acoustic perturbation equations. *J Comput Acoust*. 2012;20, 2:1240005.
34. Guasch O, Sánchez-Martín P, Pont A, Baiges J, Codina R. Residual-based stabilization of the finite element approximation to the acoustic perturbation equations for low Mach number aeroacoustics. *Int J Numer Meth Fluids*. 2016;82, 12:839-857.
35. Roger M. The acoustic analogy some theoretical background. *Lect Ser-Von Karman Inst Fluid Dyn*. 2000;2:A1—A32.
36. Lighthill MJ. On sound generated aerodynamically. *I Gen Theory, Proc R Soc Lond A*. 1952;211, 1107:564-587.
37. Crighton DG, Dowling AP, Ffowcs-Williams JE, Heckl M, Leppington FG. Modern methods in analytical acoustics. *J Fluid Mech*. 1996;318:410-412.
38. Arnela M, Guasch O. Finite element computation of elliptical vocal tract impedances using the two-microphone transfer function method. *J Acoust Soc Am*. 2013;133, 6:4197-4209.
39. Speed M, Murphy DT, Howard DM. Three-dimensional digital waveguide mesh simulation of cylindrical vocal tract analogs. *IEEE Trans Audio Speech Lang Process*. 2013;21, 2:449-455.
40. Vampola T, Horáček J, Švec JG. FE modeling of human vocal tract acoustics. Part I: production of Czech vowels. *Acta Acust United Ac*. 2008;94, 3:433-447.
41. Vampola T, Horáček J, Švec JG. Modeling the influence of piriform sinuses and valleculae on the vocal tract resonances and antiresonances. *Acta Acust united Ac*. 2015;101, 3:594-602.
42. Arnela M, Dabbaghchian S, Blandin R, Guasch O, Engwall O, Van Hirtum A, Pelorson X. Influence of vocal tract geometry simplifications on the numerical simulation of vowel sounds. *J Acoust Soc Am*. 2016;140, 3:1707-1718.
43. Takemoto H, Mokhtari P, Kitamura T. Acoustic analysis of the vocal tract during vowel production by finite-difference time-domain method. *J Acoust Soc Am*. 2010;128, 6:3724-3738.
44. Arnela M, Blandin R, Dabbaghchian S, Guasch O, Alías F, Pelorson X, Van Hirtum A, Engwall O. Influence of lips on the production of vowels based on finite element simulations and experiments. *J Acoust Soc Am*. 2016;139, 5:2852-2859.
45. Arnela M, Guasch O, Alías F. Effects of head geometry simplifications on acoustic radiation of vowel sounds based on time-domain finite-element simulations. *J Acoust Soc Am*. 2013;134, 4:2946-2954.
46. Blandin R, Arnela M, Laboissière R, Pelorson X, Guasch O, Van Hirtum A, Laval X. Effects of higher order propagation modes in vocal tract like geometries. *J Acoust Soc Am*. 2015;137, 2:832-843.
47. Blandin R, Van Hirtum A, Pelorson X, Laboissière R. Influence of higher order acoustical propagation modes on variable section waveguide directivity: application to vowel [α]. *Acta Acust united Ac*. 2016;102, 5:918-929.
48. Degirmenci NC, Jansson J, Hoffman J, Arnela M, Sánchez-Martín P, Guasch O, Ternström S. A unified numerical simulation of vowel production that comprises phonation and the emitted sound. *Interspeech*. 2017;2017:3492-3496.
49. Fleischer M, Pinkert S, Mattheus W, Mainka A, Alexander M, Mürbe D. Formant frequencies and bandwidths of the vocal tract transfer function are affected by the mechanical impedance of the vocal tract wall. *Biomech Model Mechanobiol*. 2015;14, 4:719-733.
50. Guasch O, Arnela M, Codina R, Espinoza H. A stabilized finite element method for the mixed wave equation in an ALE framework with application to diphthong production. *Acta Acust united Ac*. 2016;102, 1:94-106.
51. Krane MH. Aeroacoustic production of low-frequency unvoiced speech sounds. *J Acoust Soc Am*. 2005;118, 1:410-427.
52. Anderson P, Green S, Fels S. Modeling fluid flow in the airway using CFD with a focus on fricative acoustics. In: Proc 1st International Workshop on Dynamic Modeling of the Oral, Pharyngeal and Laryngeal Complex for Biomedical Applications; 2009:146-154.
53. Van Hirtum A, Fujiso Y, Nozaki K. The role of initial flow conditions for sibilant fricative production. *J Acoust Soc Am*. 2014;136, 6:2922-2925.
54. Fujiso Y, Nozaki K, Van Hirtum A. Towards sibilant physical speech screening using oral tract volume reconstruction: some preliminary observations. *Appl Acoust*. 2015;96:101-107.
55. Guasch O, Codina R. Computational aeroacoustics of viscous low speed flows using subgrid scale finite element methods. *J Comput Acoust*. 2009;17, 3:309-330.
56. Gloerfelt X, Pérot F, Bailly C, Juvé D. Flow-induced cylinder noise formulated as a diffraction problem for low Mach numbers. *J Sound Vib*. 2005;287, 1:129-151.

57. Hughes TJR, phenomena Multiscale. Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods. *Comput Methods Appl Mech Eng*. 1995;127, 1:387-401.
58. Hughes TJR, Feijóo GR, Mazzei L, Quincy JB. The variational multiscale method a paradigm for computational mechanics. *Comput Methods Appl Mech Eng*. 1998;166, 1:3-24.
59. Codina R. Stabilized finite element approximation of transient incompressible flows using orthogonal subscales. *Comput Methods Appl Mech Eng*. 2002;191, 39:4295-4321.
60. Codina R, Badia S. On some pressure segregation methods of fractional-step type for the finite element approximation of incompressible flow problems. *Comput Methods Appl Mech Eng*. 2006;195, 23:2900-2918.
61. Espinoza H, Codina R, Badia S. A Sommerfeld non-reflecting boundary condition for the wave equation in mixed form. *Comput Methods Appl Mech Eng*. 2014;276:122-148.
62. Kannan R, Guo P, Przekwas A. Particle transport in the human respiratory tract: formulation of a nodal inverse distance weighted Eulerian–Lagrangian transport and implementation of the Wind–Kessel algorithm for an oral delivery. *Int J Numer Meth Biomed Eng*. 2016;32(6).
63. Balay S, Gropp WD, McInnes LC, Smith BF. *Efficient Management of Parallelism in Object Oriented Numerical Software Libraries*, Modern Software Tools in Scientific Computing 163–202. Basel: Birkhäuser Press; 1997.
64. Pope SB. *Turbulent Flows*. Cambridge, UK: 238–261. Cambridge University Press; 2000.
65. Yoshinaga T, Van Hirtum A, Nozaki K, Wada S. Influence of the lip horn on acoustic pressure distribution pattern of sibilant /s/. *Acta Acust united Ac*. 2018;104, 1:145-152.

How to cite this article: Pont A, Guasch O, Baiges J, Codina R, van Hirtum A. Computational aeroacoustics to identify sound sources in the generation of sibilant /s/. *Int J Numer Meth Biomed Engng*. 2019;35:e3153. https://doi.org/10.1002/cnm.3153