# Dynamics of plosive consonants via imaging, computations, and soft electronics

Jin-Tae Kim[a,1] , Wei Ouyang[a,1] , Hanul Hwang[b,1] , Hyoyoung Jeong[a,c,1] , Soohyeon Kang[d], Sanjeeb Bose[b,e] , Sung Soo Kwak[a,f], Xiaoyue Ni[g] , Hyeonsu Kim[h], Jaehong Park[h] , Hope Chen[i], Alan Soetikno[i], Joohee Kim[a], Shuai Xu, Leonardo P. Chamorro[d,2] , and John A. Rogers[a,2]

A quantitative understanding of the coupled dynamics of flow and particles in aerosol and droplet transmission associated with speech remains elusive. Here, we summarize an effort that integrates insights into flow-particle dynamics induced by the production plosive sounds during speech with skin-integrated electronic systems for monitoring the production of these sounds. In particular, we uncover diffusive and ballistic regimes separated by a threshold particle size and characterize the Lagrangian acceleration and pair dispersion. Lagrangian dynamics of the particles in the diffusive regime exhibit features of isotropic turbulence. These fundamental findings highlight the value in skin-interfaced wireless sensors for continuously measuring critical speech patterns in clinical settings, work environments, and the home, based on unique neck biomechanics associated with the generation of plosive sounds. We introduce a wireless, soft device that captures these motions to enable detection of plosive sounds in multiple languages through a convolutional neural network approach. This work spans fundamental flow-particle physics to soft electronic technology, with implications in monitoring and studying critical speech patterns associated with aerosol and droplet transmissions relevant to the spread of infectious diseases.

flow-particle physics | experimental fluid mechanics | computational fluid dynamics | skin-integrated electronics | machine learning

Aerosol- and droplet-based transmission of infectious diseases, ranging from *Mycobacterium tuberculosis* to common cold, and associated flow and particle dynamics are poorly understood, as evidenced by relative lack of quantitative insights into the spread of COVID-19 (1, 2). Recent studies provide social distancing guidelines based on various conditions, including one-to-one exposure to infectious particles (3, 4), turbulent gas clouds (5), space and time dependence of virus transmission from speech-driven aerosol transport (6), indoor airborne transmission (7, 8), and relative humidity (9, 10). Particle dynamics of aerosols/droplets have been investigated primarily under intensive respiratory activities, including coughing and sneezing (11). The few studies of particle size distributions during speaking (12, 13) show significant variability, possibly due to the complex mechanisms of speech (14). Recent work shows that phonetic characteristics of bilabial closure (BC) plosive sounds such as "P" generate unique, coherent flow structures where sequential plosives create developed, jet-like turbulent flow and complex transport in a conversation (15). Despite these observations, the induced flow patterns and dynamics of generated respiratory aerosols and droplets during plosive vocal activities and at different audio levels remain elusive. Indeed, whether certain infectious diseases are transmitted via aerosols or droplets (16) or even the threshold between aerosols and droplets (17) represent topics that remain under debate. Particle size plays a central role in aerodynamic behavior; the size threshold for aerosols, which remain in the air, can be determined by the background and induced flows. Experimental characterization of such processes is not trivial; measurements often involve high-power illumination synchronized with high-resolution cameras and customized optical settings that require proper and extensive safety precautions and training (18). These methods may be cumbersome as means to quantify fluid-particle interactions during vocal activities over a broad range of individuals, particularly outside of controlled laboratory settings. A promising complementary approach is to exploit multiple, continuous on-body sensing modalities with soft wearable sensors and associated data analytics (19).

Continuous vital signs monitoring with wearable devices has well-established clinical benefits in both hospital and home settings (20). A fully automated soft wearable sensor implements a thin and flexible form factor with deformable electronic circuit and soft device architecture techniques to ensure skin conformability (21), especially important for the highly curved geometry of the human neck. Measured mechano-acoustic (MA) signals span a wide range from quasistatic to the audio band, without confounding

## Significance

Mechanisms for aerosol- and droplet-based transmission of infectious diseases are not well understood, especially those that follow from patterns of speech. Our work explores these processes through studies of flow-particle physics during production of plosive sounds, development of soft electronics platforms for tracking these behaviors, and application of data analytics methods for continuous monitoring. Respiratory particles that interact with dominant vortical structures exhibit unique flow-particle dynamics. Correlations between these flow-particle characteristics and unique biomechanical features of the neck during speech provides an informed strategy for the use of soft, wireless sensors to detect plosive utterances in multiple languages through machine learning approaches.

[1]J.-T.K., W.O., H.H., and H.J. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: lpchamo@illinois.edu or jrogers@northwestern.edu.

effects of ambient and external sounds (22). Recent advances in convolutional neural networks (CNN) and other deep learning methodologies (23, 24) enable accurate pattern recognition of acoustic signals by directly exploiting the time-frequency features without handcrafted features in various applications, such as music genre recognition (25) and cough (26) and singing detection (27). Several wearable devices have been developed for sensing respiratory activities (22, 26), biomechanics (28, 29) and electrograms (30, 31). Still, none has been used to study flow and particle physics during vocal activities in conjunction with optical fluid measurements and numerical simulations. Flexible, skin-mounted bio-integrated electronics with integrated deep learning analytics offer the potential to bridge the gap between our fundamental understanding of flow and particle physics associated with phonation activity and infectious disease transmission through the air.

The presented work explores fluid mechanics in plosive sounds and sets the basis for potential novel applications of plosive sound detection with skin-integrated electronics. In particular, this paper discusses fundamental phenomena associated with particle and induced flow dynamics via particle tracking velocimetry (PTV) and particle image velocimetry (PIV) supported by Euler-Lagrangian direct numerical simulations (DNS) and analytic arguments. Biomechanics aspects of the surface of the neck surface are investigated via three-dimensional digital-image correlation (3D-DIC) to capture unique features of neck vibrations induced by production of plosive sounds. These fluid and biomechanics results form the foundations for a wireless, soft MA device with the capability of plosive sound detection through a convolutional neural network (CNN) approach.

The following sections describe and discuss (i) droplet dynamics and induced flow, formation and impact of dominant vortices for a set of plosive sounds at decibel levels via PTV and PIV, and correlation between audio level (dB) and mass transport; (ii) dynamics of a wide range of particle sizes spanning aerosols and droplets with experimentally validated DNS; (iii) theoretical assessment of the impact of initial condition on particle dynamics and induced flow and thresholds between aerosol and droplet in the Lagrangian frame of reference; (iv) biomechanics across the neck during production of plosive sounds and other vocal/respiratory activities via 3D-DIC; (v) design of optimized, automated, skin-integrated electronic sensors with synchronized pairs of high-bandwidth inertial measurement units (IMUs) across the suprasternal notch (SN) and sternocleidomastoid (SCM) for capturing unique MA signals of plosive sounds; and (vi) CNN of MA signals for detection of plosive sounds in various individuals.

## Results

### Droplet Dynamics and Induced Flow in Plosive Consonants.
The size distribution and dynamics of droplets are experimentally characterized in 80 distinct cycles of four representative BC plosive words ("Peter," "piper," "pick," and "pack") (15) spoken by a 30-year-old healthy male across various audio levels, ranging from 68 to 98 dB using PTV (see *Materials and Methods* and *SI Appendix,* Video S1). Four high-speed cameras synchronized with a 40 W laser are used to characterize the size and particle locations, as illustrated in Fig. 1 *A* and *B*; a representative instant during a plosive sound is shown in Fig. 1*C*. Large droplets of radii $R \gtrsim 100$ µm exhibit ballistic-like behavior with similar velocity magnitude and direction, whereas smaller droplets ($R \lesssim 100$ µm) show irregular velocity vectors. The characteristic Stokes number, Stk $= \frac{2}{9}\frac{\rho u R^2}{\mu L}$,

embodies this phenomenon; here, $\rho$ is the particle density, $u$ is the fluid velocity, $\mu$ is the dynamic viscosity of the air, and $L$ is the reference length given by the equivalent diameter of mouth opening. Large particles with Stk > 1 detach from the flow (9, 32), especially in these highly decelerating flows. PTV-derived particle number and size distribution are illustrated in Fig. 1 *D–F* and *SI Appendix,* Fig. S1. The number of droplets in the BC plosive sounds, shown at various audio levels in Fig. 1*D*, correlates with audio level (dB) regardless of plosive words, similar to previous work on coughing and speech (26). One of the main differences between coughing and BC plosive sound is that the latter contains stop consonants where the lips block the airflow entirely, build pressure, and release flow. This phonetic mechanism generates systematic aerosolization from salivary filaments at the lips that break into speech droplets (33). Depending on the lips' condition, BC plosive sound can generate a significant number of droplets compared with other speech and even more than coughing at a higher audio level due to its unique mechanism (Fig. 1*E*). In addition to the number of droplets, the droplet size distribution represents a key parameter in developing an improved understanding and modeling capability for droplet dispersion induced by plosive consonants. The droplet size distribution is obtained for all measured cycles in the range of $R = 50$ to $150$ µm and fitted with the generalized Pareto distribution (GPD), defined as $f(R) = \frac{1}{\sigma}\left(1 - \left(1 + \frac{k(x-\theta)}{\sigma}\right)^{-1-\frac{1}{k}}\right)$, where $\sigma$, $k$, and $\theta$ are scale, shape, and threshold parameters, respectively, in Fig. 1*F*. It is worth noting that regardless of the type of BC plosive words and audio levels, generated droplets are described well by the GPD in the measured range, similar to the case of coughing (34). The mean traveled distance of droplets, $\overline{x}$, exhibits dependence with $R$ and dB (Fig. 1*G*). As $R$ increases, the maximum $\overline{x}$ increases and associated critical time decreases, as a large droplet carries comparatively larger momentum but experiences higher drag. Systematic increases in $\overline{x}$ occurred with dB. It is worth noting that $\overline{x}$ plateau within length and time scales of $< 10^1$ cm and $< 10^0$ s in the isolated single puffs generated by different plosive consonants. Complementary PIV experiments are performed to reveal the formation and impact of a vortex ring from plosive sounds (Fig. 2 *A* and *B* and *SI Appendix,* Video S2) (33).

**Computational Results.** Three-dimensional DNS are carried out to model the flow and droplets interaction induced by plosive sounds and complement the experiments on the droplet dynamics smaller than the observable size ($R < 25$ µm) and Lagrangian trajectories. The simulation accounts for two-way coupled physics with respect to mass, heat, and momentum transfer between air and droplets using a Lagrangian point-particle approach. In addition, mass transfer between the droplets and surrounding air due to evaporation is considered with the assumption of incompressibility with a low Mach regime and constant pressure. For the simulation domain, convective and slip boundary conditions are employed for the outlet and the rest of the faces, respectively. Three components are implemented for the model of a human head, namely, outer surface of the head, inner surface of the pathway, and the inlet. No-slip boundary condition is applied for the outer surface. Algebraic wall model boundary condition is employed in the inner surface of the pathway. Dirichlet boundary condition is used to prescribe the uniform velocity of the air for the inlet (see details in *Materials and Methods*). Approximately 500 particles are introduced for the speech duration with a size distribution
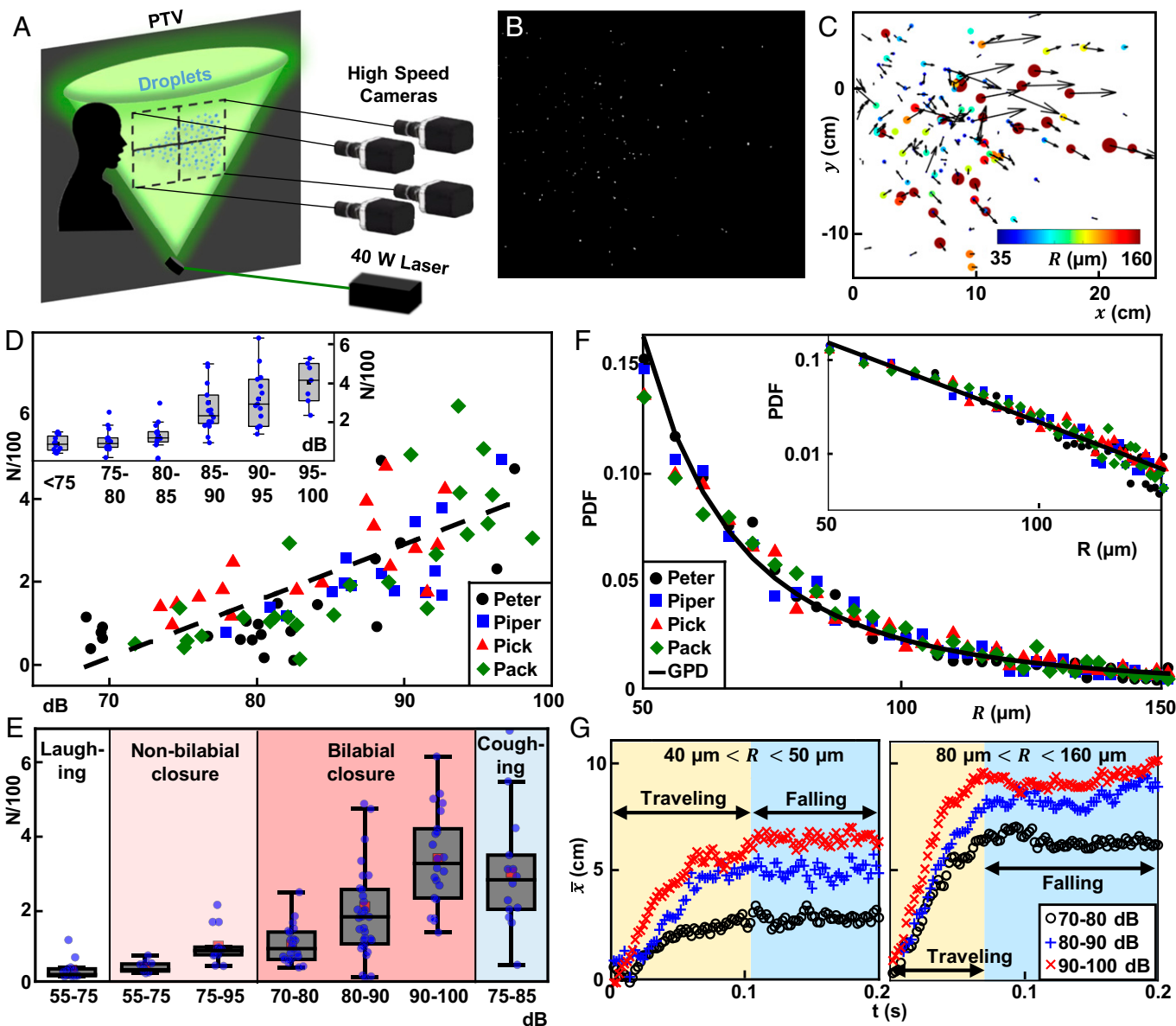
**Fig. 1.** Droplet dynamics of plosive sounds via particle tracking velocimetry (PTV). (*A*) PTV experimental setup. (*B*) Raw image of droplets generated from plosive sounds. (*C*) Detected/tracked droplet via PTV; color and vector denote the radius and velocity of droplets. (*D*) Number of droplets, N, with respect to dB. (*E*) Droplet statistics at various respiratory activities. (*F*) Probability function of droplet radii *R*; inset denotes the PDF in semilong scale and the solid line represents the GPD with $\sigma = 1$, $k = 0.52$, and $\theta = 0$. (*G*) Mean traveled distance of droplets at different *R* and dB.

following the GPD distribution observed in the experiments. Fig. 2*C* demonstrates the particle trajectories with respect to *R* and, as expected, comparatively larger particles travel longer distances with projectile-like motion. Sufficiently small particles are trapped inside the vortex ring structure induced by the sharp shear at the boundary of the jet-like flow and pushed farther than large particles. The vortex ring accompanied by a ballistic mechanism has shown its effectiveness in mass transport in nature (35). It is worth pointing out that the simulations consider the absence of ambient flow to capture essential features induced by plosive sounds. Fig. 2*D* shows a similar instant as in the results from PIV experiments (Fig. 2*B*) to illustrate the specific mechanism of plosive sounds. Notably, large particles' acceleration allows them to escape or be weakly affected by the vortex ring structure. In contrast, the induced flow substantially affects small particles' velocity magnitude and direction. This mechanism is compared and validated with the experiments as shown in Fig. 2*E*. They

exhibit a strong correlation between the audio level and initial gas velocity, $u_g$.

**On a Theoretical Basis for the Particle Dynamics in a Vortex Ring.** Conceptual assessment of validated DNS shed light on the coupled physics between initial gas velocity, $u_g$, and sound level (dB), and associated particle dynamics over a range of sizes. Fig. 2*F* shows the ensemble-averaged velocity in the streamwise direction, $\overline{u} = \langle \frac{1}{T} \int_0^T u\,dt \rangle_R$, with respect to *R*, where T = 0.2 s is the time span. The bulk $\overline{u}$ velocity distributions evidence two distinct regions. One of them is inversely related to $R \lesssim R_c$, suggesting a diffusive regime, whereas the other is proportional to $R \gtrsim R_c$, leading to a ballistic-like regime, where $R_c = \sqrt{\frac{9\mu L S}{2\rho u_g}}$ is the critical particle radius and $S = 0.1$ is the critical Stokes number. Particles in the ballistic regime are modulated by the initial condition and surpass the
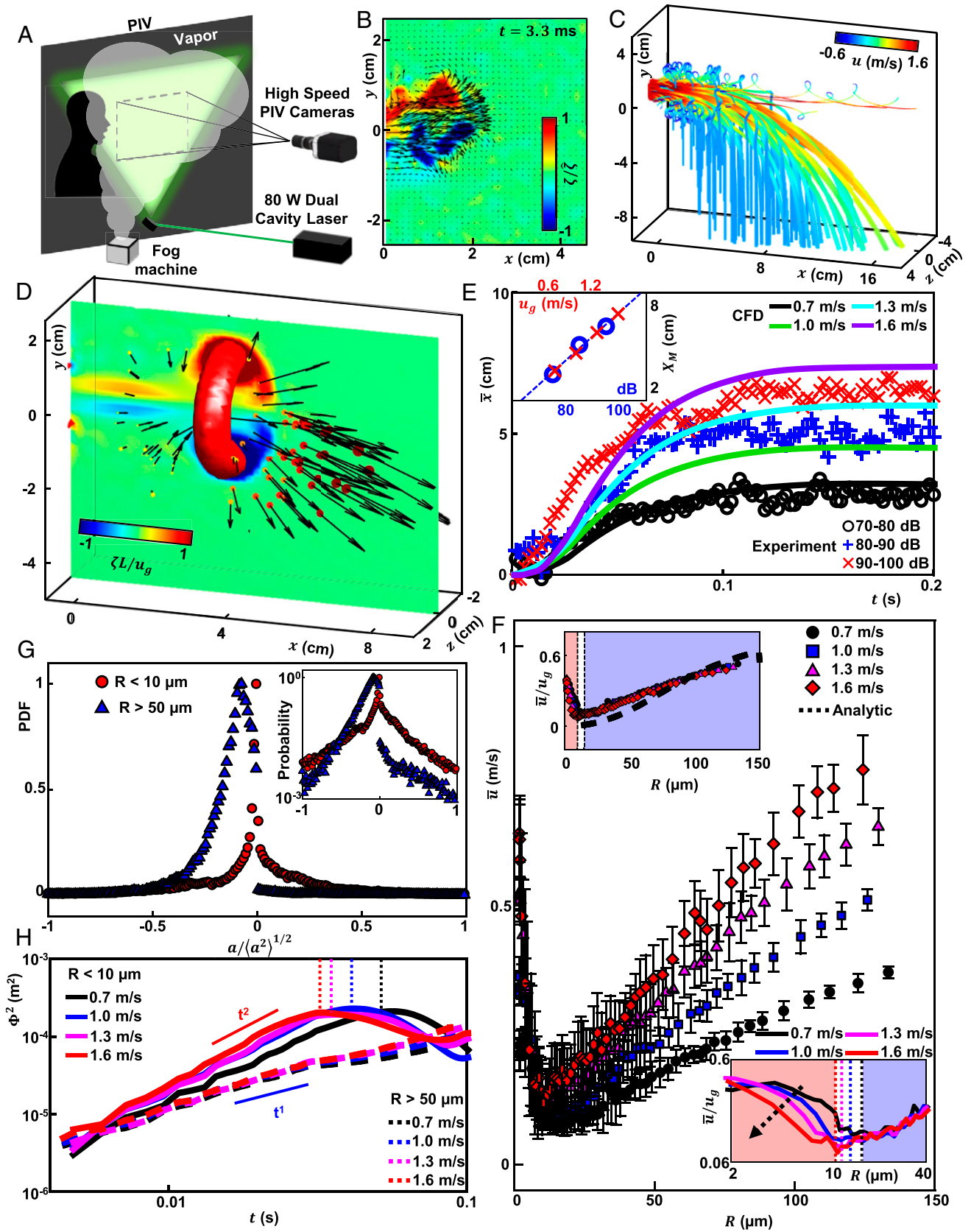
**Fig. 2.** Flow and particle physics of plosive sounds via PIV and computational fluid dynamics (CFD). (*A*) PIV experimental setup. (*B*) Representative instant of the flow field induced by plosive sounds. (*C*) Numerical results of particle trajectories with respect to particle size at the initial gas velocity, $u_g = 1.6$ m/s; color denotes the particle streamwise velocity $u$. (*D*) Representative instant of the flow-particle interaction induced by plosive sounds. (*E*) Validation of the numerical results based on the comparison of particle travel characteristics with PTV results at various dB; the inset shows the correlation of the traveled distance of droplets in the range of 40 μm < $R$ < 50 μm between experimental results with respect to dB and simulation with respect to $u_g$. (*F*) Ensemble-averaged particle velocity $\overline{u}$ with respect to $R$; insets show semilog scale highlighting the ballistic (blue; *Upper Left*) and diffusive (red; *Bottom Right*) regimes; dotted colored vertical lines are the critical radii $R_c$ corresponding to $u_g$. (*G*) Probability density function (PDF) of normalized Lagrangian accelerations, $a/a^{21/2}$, of particles in the diffusive (red circle) and ballistic (blue triangle) regimes. (*H*) Pair dispersion, $\Phi^2$, of particles in diffusive (solid lines) and ballistic (dashed lines) regimes at various $u_g$; dotted colored vertical lines are the critical times $T_c$ corresponding to $u_g$.

induced flow. The bulk velocity in this regime increases with $R$ and $u_g$ can be characterized by

$$\overline{u} = \frac{2R^2 u_p \rho}{9\mu T}\left[1 - \exp\left(-\frac{9\mu}{2R^2\rho}T\right)\right] \quad \text{for} \quad R > R_c = \sqrt{\frac{9\mu LS}{2\rho u_g}},$$

[1]

where $u_p = 0.9 u_g$ is the peak velocity. However, particles in the diffusive regime exhibit an inverse relation with $R$ as they are initially carried by the induced flow. Nondimensional bulk velocity $\overline{u}/u_g$ trends, shown in the top and bottom insets of Fig. 2F, further reveal common features. The top inset reveals a $\overline{u}/u_g$ self-similar behavior that collapses into a scale, as inferred from Eq. **1**. The log-log representation in the bottom inset highlights dynamics in the diffusive regime and critical values separating the two regimes by $R_c$. As $u_g$ increases, $R_c$ decreases due to an increase in the Stk number at a given particle size. Unlike the particle dynamics in the ballistic regime, $\overline{u}/u_g$ does not overlap and exhibits a sharper decrease with respect to $u_g$ due to the distinct interaction between $R$ and the induced flow, as described by the Stk number.

Inspection of the Lagrangian characteristics of the particle dynamics provides additional insight on particle behaviors as tracers, i.e., aerosols, or inertial particles, i.e., droplets, or in-between state obeying diffusive and ballistic regimes. It includes the probability density function (PDF) of normalized Lagrangian accelerations along the streamwise direction, $a/\langle a^2\rangle^{1/2}$, and pair dispersion, $\Phi^2(t) = \langle[r(t) - r_i]^2\rangle$, where $r(t)$ is the separation between two particles as a function of time, where $r_i$ is the initial separation. Notable characteristics of $a/\langle a^2\rangle^{1/2}$ and $\Phi^2(t)$ of tracer particles in turbulent flows include non-Gaussian PDF of $a/\langle a^2\rangle^{1/2}$ with a heavy-tail distribution, indicating intermittent events (36, 37) and $\Phi^2(t) \propto t^2$ scaling predicted by Batchelor (38). Particle behaviors based on these characteristics can provide information on transmission by quantifying the linkage with the flow at a given length, size, and time scale from a Lagrangian perspective. The $a/\langle a^2\rangle^{1/2}$ PDF in the diffusive regime exhibits a non-Gaussian symmetric distribution with heavy-tail evidencing intermittent events, similar to that found in tracers in fully developed turbulence (36). However, the particle acceleration PDF in the ballistic regime shows a positively skewed distribution with comparatively high kurtosis resembling a skew-normal distribution. The results indicate that particles in this regime mostly face deceleration with reduced effect of flow disturbances. In addition, $R^2(t)$ increases with $t^2$ in the ballistic regime, similar to tracers in developed turbulence (39). A change in the trend occurred at a characteristic strain time scale, $T_c = 2/\omega$, where $\omega$ is the vorticity magnitude. On the other hand, the relative dispersion of particles in the diffusive regime increases linearly in the time span. This behavior occurs with particles in a quiescent fluid undergoing Brownian motions; i.e., the particles are not affected by the induced flow (see details in *SI Appendix, Supplementary Note*).

**Biomechanics Across the Neck during Plosive Consonants.** Fundamental assessments in flow-particle physics associated with critical speech patterns in the context of infectious disease transmission further motivate the development of continuous monitoring technologies for such respiratory activities during natural daily activities in work and home settings. BC plosive sound detection through soft, skin-mounted electronic sensors critically relies on determining optimal physiological mounting locations. Spatiotemporal maps of motions of the neck during various vocal activities determined by 3D-DIC (40) reveal the biomechanics associated with BC plosive sounds (*SI Appendix,* Video S3), using methods similar to those in 3D-PTV of the neck during cardiac and breathing activities (22). A basic schematic and representative images of 3D-DIC measurements are given in Fig. 3 *A* and *B*. Experiments quantify 3D displacements across the entire neck, resolving over 6,000 grid points during representative BC plosive sounds and nonplosive sounds in the 75–85 dB range and during swallowing.

Two physiological locations exhibit distinct mechanisms of plosive sounds. One of these is the SN, a dip between two clavicles and located on the apex of the sternum. The other is the SCM muscle, a two-headed neck muscle above the clavicle. Out-of-plane displacements at the SN, $\Delta z_{SN}$, and the SCM, $\Delta z_{SCM}$, drastically increase shortly before pronouncing the plosive sound "P," as shown in Fig. 3C. The rapid change follows from obstruction of the airstream by the lips, which builds pressure inside the trachea and expands the cartilaginous tube. The quick release of blockage in the airway during the process of plosive pronunciation can be described in terms of the temporal derivatives, $\Delta v_{SN}$ and $\Delta v_{SCM}$ as shown in Fig. 3D. Both $\Delta v_{SN}$ and $\Delta v_{SCM}$ exhibit sharp increases, reaching a local peak velocity $V_p$, and decrease as unique features of plosive sounds in both locations. Three-dimensional maps of relative instants defined as (I), (II), (III), and (IV) in Fig. 3C, where the local maximum deformation in the out-of-plane direction occurs for each plosive sound, capture the similar mechanism across various BC plosive words, as shown in Fig. 3 *E–H*.

The 3D maps of other types of phonetic sounds, such as "will" and "sing," are shown in Fig. 3 *I* and *J*. The deformations of the SN and SCM are the largest in the order of plosive [p] (Fig. 3C), fricative [s] (Fig. 3L), and approximant [w] (Fig. 3K) sounds, consistent with the degree of the blocked airflow. Deformations are relatively small during vowel sounds (e.g., [i] [a]) due to the near-absence of blockage of the airflow relative to that associated with consonant sounds (see *Materials and Methods*). Measurements performed at both the SCM and the SN help to distinguish signals from other activities that also exhibit large deformations at SN such as swallowing, as shown in Fig. 3 *M–P*. The swallowing undergoes a similar displacement magnitude at SN compared with the one found in a plosive sound during the esophageal phase (Fig. 3O). However, such displacements do not occur at the SCM, as the bolus travels downward by peristaltic motions inside the esophagus.

**Design of Dual-in-Plane (DiP) MA sensor.** The thin, soft physical properties of the wireless device enable continuous measurements of dual MA signals from the skin surface in these regions without discomfort. The DiP device is optimized for plosive sound detection, while retaining sensing capabilities from previous versions designed for cardiopulmonary applications (21). The resulting measurements are immune from cross-talk among multiple individuals and from other effects of ambient sounds, highlighting an essential advantage over other approaches, such as those that rely on microphones. Fig. 4A shows the system relative to a US quarter, with a section of 22 × 65 mm and a thickness under 9 mm, and a weight below 7 g. Fig. 4B and *SI Appendix,* Fig. S3 show an exploded view and layout of the MA device. A silicone enclosure surrounds a flexible printed circuit board (fPCB) and multiple passive and active electrical components, including two identical inertial measurement sensors, a Bluetooth system-on-chip (SoC), and a lithium-polymer battery.
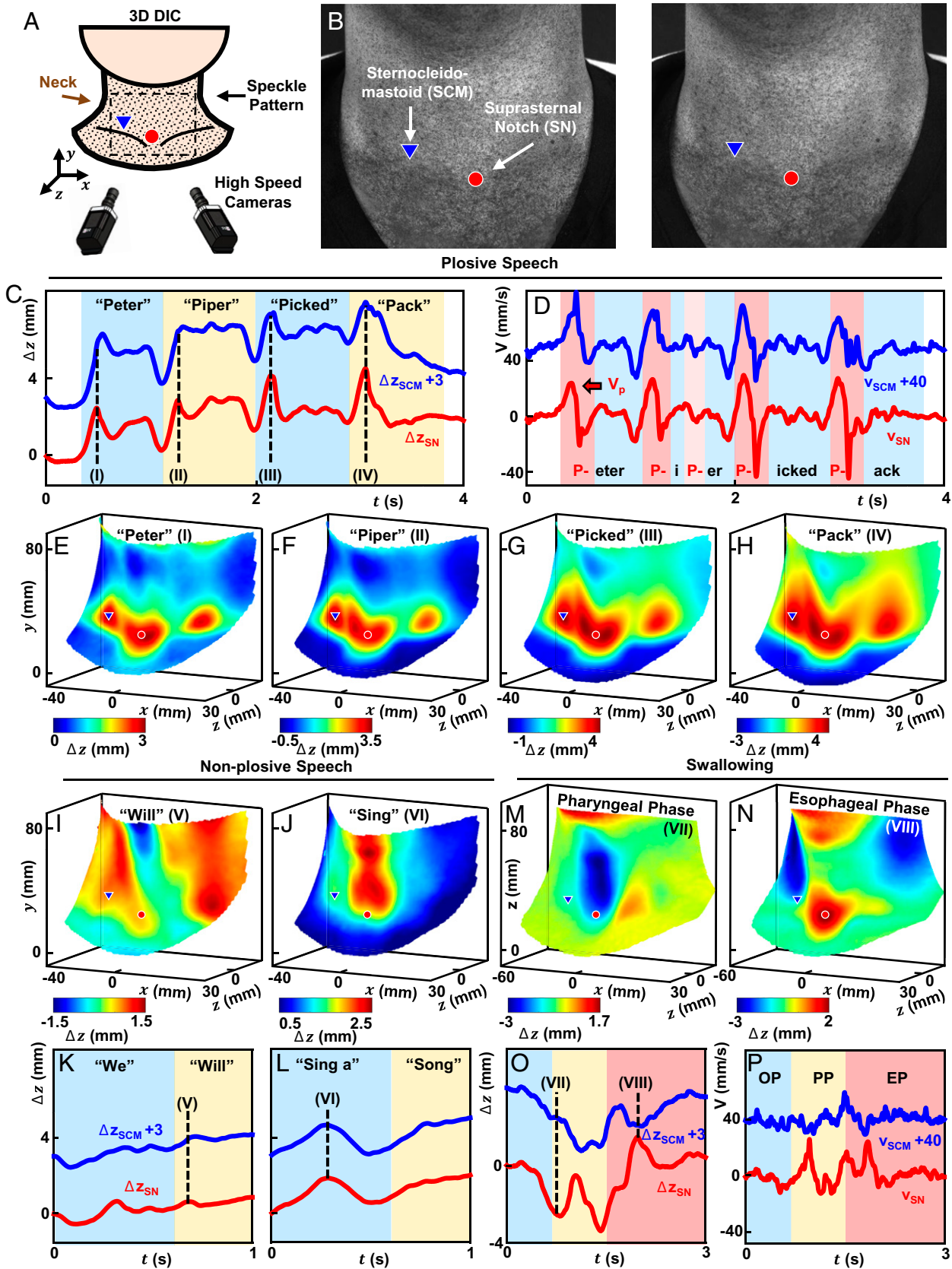
**Fig. 3.** Mechanics of the neck via three-dimensional digital-image correlation (3D-DIC). (*A*) 3D-DIC experimental setup. (*B*) A sample raw image set for stereoscopic vision; red circle and blue triangle symbols denote suprasternal notch (SN) and sternocleidomastoid (SCM). (*C*, *D*) Displacement and velocity profiles of plosive sounds; dotted vertical lines exhibit local maxima. (*E–H*) 3D displacement fields of plosive sounds. (*I*, *J*) 3D displacement fields of nonplosive sounds. (*K*, *L*) Displacement profiles of nonplosive sounds. (*M*, *N*) 3D displacement fields of swallowing. (*O*, *P*) Displacement and velocity profiles of swallowing at the oral (OP), pharyngeal (PP), and esophageal (EP) phases.
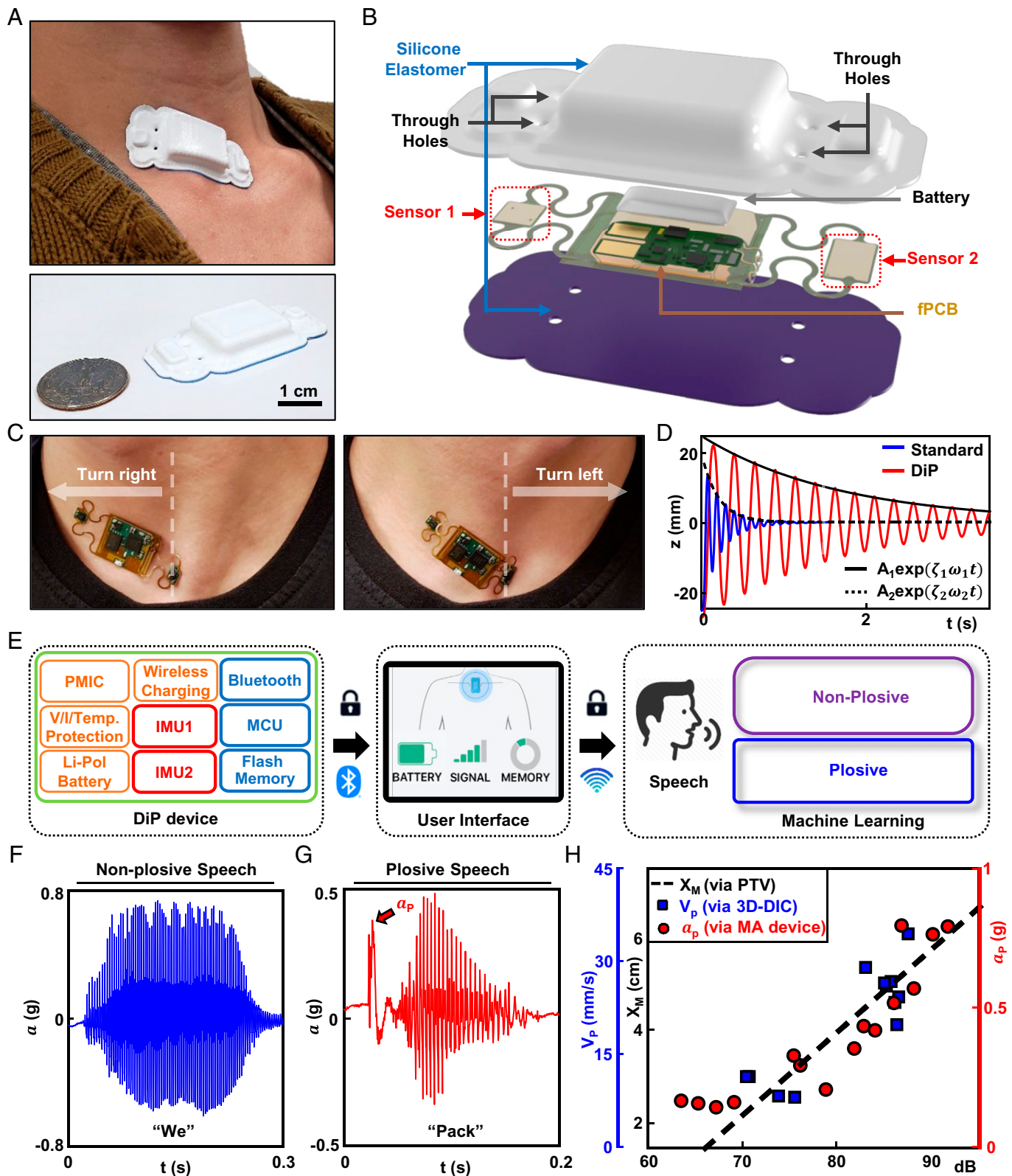
**Fig. 4.** Wireless, skin-interfaced dual-in-plane (DiP) mechano-acoustic platform for plosive sound detection. (*A*) Image of the device interfacing the suprasternal notch (SN) and sternocleidomastoid (SCM) (*Top*); size comparison to a Washington quarter (*Bottom*). (*B*) Exploded schematic illustration of the holey design, including active and passive electrical components, fPCB interconnect schemes, and silicone enclosure architectures. (*C*) Images of the device undergoing various mechanical deformations while worn on the subject: a turning right angle of 90°(*Left*) and a turning left angle of 90°(*Right*). (*D*) Natural frequencies and damping ratios of a DiP device and standard device. (*E*) Block diagram of platform (PMIC and MCU are power management integrated circuit and microcontroller, respectively); a tablet provides the interface to control the device and download the data from the device through Bluetooth. The tablet uploads the data to the cloud platform through a Wi-Fi. The cloud-based machine learning processes the classification of plosive and nonplosive sounds. (*F, G*) MA signals of representative nonplosive and plosive sounds, respectively. (*H*) The fitted linear model of number of droplets $N_L$ via PTV, local peak velocity $V_P$ via 3D-DIC, and local peak acceleration $a_P$ vs. dB.

Serpentine interconnections between these components enable a low equivalent elastic modulus of the integrated devices. Fig. 4C and *SI Appendix,* Fig. S4 show images of a device undergoing various mechanical deformations while worn on a subject during representative motions: stretching (sensor on SCM) and compressing (sensor on SN) under turning right angle of 80°, compressing (sensor on SCM and SN) under turning left angle of 80°, and twisting (sensor on SN) under looking up of 80° motions. Further modifications from previously reported versions of this device (21, 26) customize the system for detecting plosive sounds by adding holes (*SI Appendix,* Fig. S5) through the encapsulating layers. This "holey" design reduces the stresses at the skin interface and provides additional advantages (41); these include improvements in flexibility in comparison with the natural frequency and damping ratio of standard devices (21) as shown in Fig. 4D. The block diagram in Fig. 4E summarizes the system architecture and operation (see *Materials and Methods*).

Local peak acceleration ($a_P$) allows distinguishing nonplosive and plosive sounds as shown in Fig. 4 F and G and *SI Appendix,* Fig. S6. Plosive consonant sounds, such as pronouncing a voiceless "pack" sound, induce a sharp peak (red arrow in Fig. 4G) caused by sudden airflow that follows from releasing the lips, unobserved in nonplosive sounds (Fig. 4F). The mean traveled distance of droplets, $X_M$, from the inset of Fig. 2E, local peak velocity of the skin vibrations at SN, $V_p$, from Fig. 3D, and local peak acceleration of MA signals at SN, $a_p$, from Fig. 3F show positive correlations with respect to dB as shown in Fig. 4H. The results indicate that with a proper calibration between dB, $a_p$ and $X_M$, speech patterns and features associated with the flow-particle characteristics from a wide range of individuals, including those with high potential for disease transmission, could be identified by the sensor.

**CNN Algorithm for Plosive Sound Detection.** The sensor data can classify plosive and nonplosive sounds for a wide range of individuals using machine learning techniques. Fig. 5A shows representative MA signals recorded at the SN and corresponding spectrograms during the articulation of plosive and nonplosive words from a representative English vocabulary (see *Materials and Methods*). MA signals of nonplosive sounds show pronounced harmonic features. Those of plosive sounds exhibit weak but wideband nonharmonic signatures and harmonics. These nonharmonic signatures are closely associated with the quick release of air pressure in the cartilaginous tube. The mechanism produces a rapid change in the tube volume, confirmed by 3D-DIC and MA signals in terms of $V_P$ and $a_P$, respectively. The effect of skin condition especially for older adults, the age group most susceptible to infectious disease, is negligible in detecting plosive speech patterns, as the MA signals show similar temporal and frequency characteristics across different ages (*SI Appendix,* Fig. S7). MA signals from the SCM (*SI Appendix,* Fig. S8) exhibit similar differences between plosive and nonplosive sounds. A CNN can recognize these different features. The algorithm starts with a short-time Fourier transform of synchronized z axis accelerations at the SN and one side of the neck using a Hanning window with a width of 0.4 s moving in time steps of 0.01 s (Fig. 5B). The algorithm considers each windowed signal with a duration of 0.8 s independently in the process of sound classification. A minimum amplitude threshold of −10,000 detects peaks of the logarithm of spectral power integrated across the high-frequency band (>10 Hz) and labels them as speech, with a minimum time interval between peak events of 0.8 s. The spectrogram of the MA signals at the SN (shape: $320 \times 80$) SCM (shape: $320 \times 80$) form a concatenated spectrogram (shape: $320 \times 160$) that serves as the

input of the CNN model (Fig. 5C). The CNN starts with four stages of convolutions with rectified linear unit (ReLU) activation and a kernel size of $3 \times 3$, followed by two layers of fully connected neural networks with ReLU activation and two dropout layers ($P = 0.5$) alternately. The final output of the CNN model has two neurons with Softmax activation, which correspond to the probabilities for classifying the event as a plosive or nonplosive sound.

The training uses data collected from 10 volunteers, each consisting of ~2,500 articulations of plosive words and 2,500 articulations of nonplosive words from a representative English vocabulary (see *Materials and Methods*). A standard method to validate the generalization performance of a machine learning model relies on a leave-one-out strategy, where one leaves a subject out of the training set (nine subjects for training) and then tests the trained model on this subject. Iterations apply this approach to each of the 10 subjects. Each training set consists of a random collection of 80% of the labeled events from the nine subjects, thereby leaving the remaining 20% for validation. The training uses an Adam optimization algorithm. Fig. 5D shows the averaged confusion matrix of 10 leave-one-out testing cycles. The model achieves accuracies of $0.87 \pm 0.03$ for nonplosive words and $0.87 \pm 0.04$ for plosive words. Fig. 5E shows the overall classification accuracies on each subject using a model trained on the other nine subjects. The overall accuracy is above 0.85 for nine subjects and 0.83 for one subject, with MA spectrograms from SN and SCM as input features. In comparison, models trained only on the spectrograms from SN or only on those from the side show lower accuracies of about 0.8. Fig. 5F presents the receiver operating characteristic curves for each subject. The high area under the curve >0.89 for all subjects indicates that the model achieves a good balance between sensitivity and specificity. Plosive sound detection in another language, Korean as an example, exhibits similar results to English's, suggesting that the approach captures the essential physics of plosive sound regardless of language (*SI Appendix,* Fig. S8).

## Discussion

This paper presents the results of experimental and numerical investigations of the flow-particle physics induced by production of plosive sounds. The associated dynamics embody vortex ring, tracer- and inertial particle interactions. We identify diffusive and ballistic particle regimes, where particles exhibit a self-similarity behavior predominantly governed by initial conditions of the speech in the ballistic regime. In contrast, particles in the diffusive regime are trapped inside the vortex ring and show a strong dependence on background flow structures. The unique biomechanics of the neck surface during production of plosive sounds motivate the development of a customized, soft wearable device for plosive sound detection. Using the CNN method, we demonstrate effective plosive sound detection in English and Korean. This approach may serve as a complementary method to identify unique speech patterns associated with the flow-particle physics in droplet/aerosol transmission over a broad range of individuals in both home and clinical settings.

## Materials and Methods

**PTV on Tracking Droplets.** Particle dynamics of plosive sounds generated by a healthy male (31 y old) were quantified by PTV. Plosive words including "Peter," "piper," "packed," and "pack" were repeated for 22, 15, 17, and 26 times at various audio levels ranging from 65 to 99 dB, respectively (measured by Decibel × calibrated by SD-4023 sound level meter and R8090 Sound Level Calibrator). PTV experiments were recorded by four synchronized 2MP Emergent HT-2000M
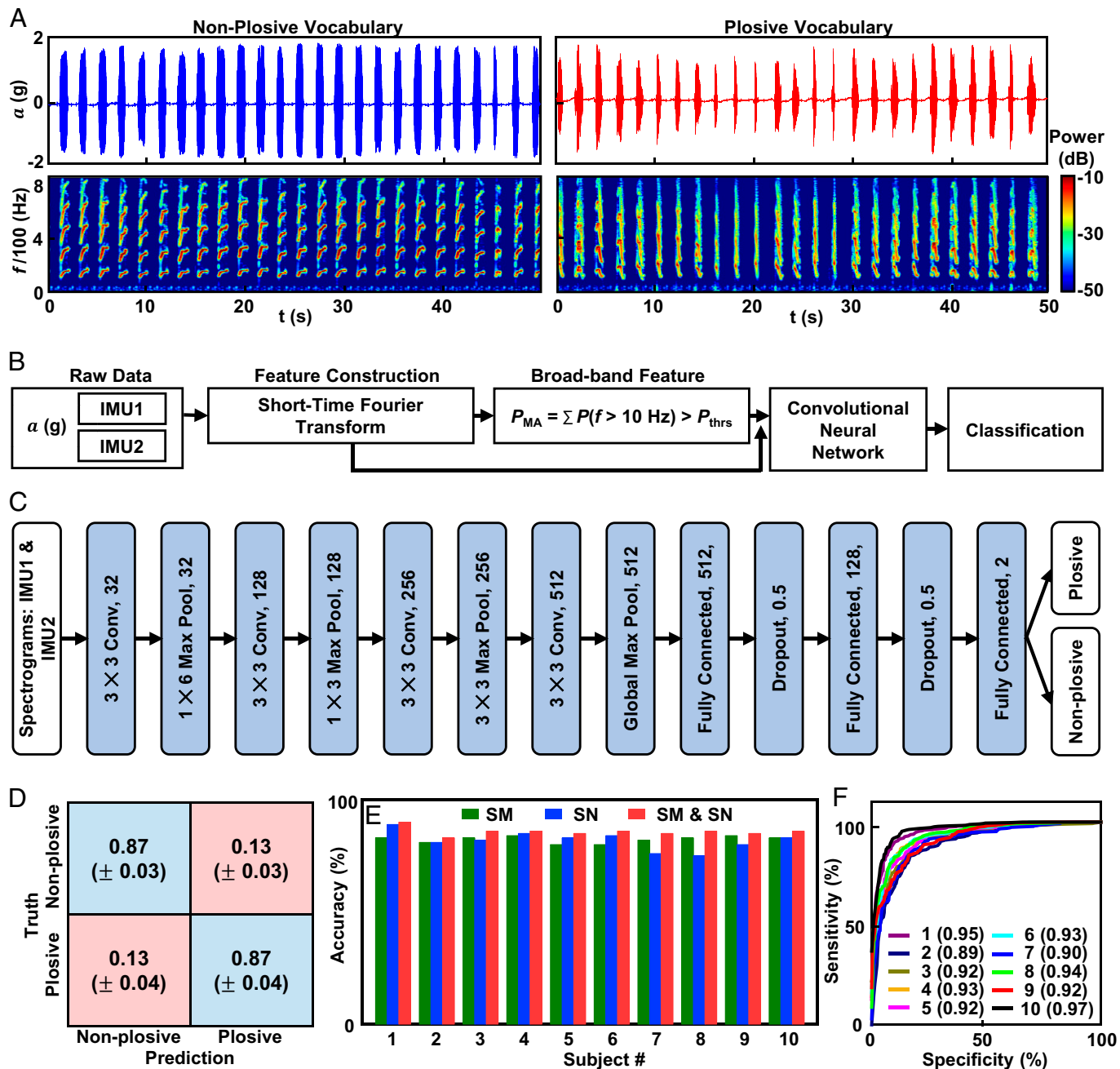
**Fig. 5.** Machine learning–based classification of plosive and nonplosive sounds from SN (IMU1). (*A*) Spectrograms of plosive and nonplosive words indicate different frequency characteristics. (*B*) Data processing flow. (*C*) Architecture of the convolutional neural network. (*D*) Confusion matrix of the neural network. (*E*) Overall accuracy of the neural network. (*F*) Receiver operational curve of the neural network on different subjects.

cameras with a 50 mm F1.4 manual-focus Kowa lens at the frame rate of 500 fps. The investigation volume for tracking particles was $\sim30 \times \sim20 \times \sim10$ cm³ illuminated by a synchronized Nd:YLF (Neodymium-doped yttrium lithium fluoride) laser with pulse energies of 25 mJ (527-49-M, Terra). Preprocessing, calibration, tracking, and postprocessing were processed by the PTV code developed by the RETEG group at UIUC (42). Image sequences were preprocessed by subtracting the background noise and enhancing the contract, and particles were detected in a subpixel level with the area estimation. The scattering cross-section of a detected particle, refractive index of particle as well as the surrounding medium, and wavelength of the light source were implemented to calculate the actual radius of detected particles based on the Mie scattering theory (43, 44). The minimum radius of particles measured in this work was $\sim25$ μm. Detected particles were tracked using the Hungarian algorithm and linked by performing a five-frame gap closing to produce longer trajectories. Velocity and Lagrangian acceleration were filtered and computed using fourth-order B splines (45).

**PIV on Plosive Flow Fields.** PIV experiments were implemented to understand flow dynamics of plosive sound produced by a human. A male adult spoke the "pa" sound in front of a closed acrylic box with a hole, on which he put his mouth. The cross-section of the box was $20 \times 20$ cm square, and the length was 40 cm. Before each experiment, the acrylic box was filled with green, dyed water droplets generated from a nebulizer, which served as tracer particles. The PIV system includes a Nd:YLF laser with pulse energies of 50 mJ (527-80-M, Terra), a digital camera ($2,560 \times 1,600$ pixels, CMOS Phantom Miro 340), a synchronizer, and a PIV control software (Insight 4G software, TSI). A 1 mm thick laser sheet illuminated a middle plane of the box. The field of view covered 130 cm in a streamwise direction from the mouth and 190 cm in a cross-flow direction. With a sampling frequency of 300 Hz, 1,000 image pairs were captured, where delta $t$ between the images for one pair was 50 μs. Two-dimensional velocity fields were obtained from the image pairs, using a recursive cross-correlation method with the final window size $20 \times 20$ pixel. The corresponding velocity

field consists of a total of 12,504 velocity vectors with a spacing of $\Delta x = \Delta y = 1.4$ mm.

**Computational Fluid Dynamics.** Three-dimensional numerical simulations are carried out by solving the compressible Navier-Stokes equations for the gas phase, while the droplets are modeled as Lagrangian point particles. The air and droplets are two-way coupled with respect to mass, heat, and momentum transfer. The governing equations for mass, momentum, energy, and gas species are Favre-filtered conservation and are given as follows:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{u}) = S_Y,$$

$$\frac{\partial \rho \vec{u}}{\partial t} + \nabla \cdot (\rho \vec{u} \vec{u}) = -\nabla p + \nabla \cdot \tau + S_u,$$

$$\frac{\partial \rho E}{\partial t} + \nabla \cdot (\rho \vec{u} H) = \nabla \cdot (\tau \cdot u) - \nabla \cdot (\rho \alpha \nabla h) + S_q,$$

$$\frac{\partial \rho Y}{\partial t} + \nabla \cdot \left(\rho \vec{u} Y\right) = -\nabla \cdot j + S_Y,$$

where $\vec{u}$, $\rho$, $p$, and $\tau$ are the velocity vector, mixture density, pressure, and viscous stress tensor, respectively, for the air. The specific total energy density is defined as $E = e + \frac{1}{2}\vec{u} \cdot \vec{u}$ as the sum of specific internal energy, $e$, and specific kinetic energy. $H = E + p/\rho$ is the stagnation enthalpy, $\alpha$ is the mixture thermal diffusivity, and $h$ is the mixture enthalpy. $Y$ and $j$ are the mass fraction and diffusive species flux, respectively, associated with gaseous aerosols. $S_Y$, $S_u$, $S_q$ denote the mass, momentum, and heat transferred from the dispersed aerosol to the gas phase, respectively. The air obeys an ideal gas equation of state, and the governing equations are solved using a second-order finite volume discretization on unstructured grids. The numerical scheme is constructed with low numerical dissipation, making it amenable for turbulent transport and dispersion simulations.

Droplets are assumed as discrete single-component spherical particles whose volumetric loading is small, and their density is much larger than that of the air. The evaporation of droplets assumes a quiescent surrounding air, or equivalently, the air evolves at a faster time scale compared with that of evaporation, and the system has a constant pressure. The governing equations to the dynamics of droplets are (46):

$$\frac{dX_i}{dt} = v_i,$$

$$\frac{dv_i}{dt} = \left(\frac{f_1}{\tau_d}\right)(u_i - v_i) + g_i,$$

$$\frac{dT_d}{dt} = -\frac{6Nu}{\rho_L C_L D^2} \frac{\lambda}{C_p} \left[\Delta h^* - L_v \ln\left(\frac{1 - Y_\infty}{1 - Y_{vs}}\right)\right],$$

$$\frac{dm_d}{dt} = -\pi D Sh \frac{\lambda}{C_p} \ln\left(\frac{1 - Y_\infty}{1 - Y_{vs}}\right),$$

where $X_i$, $v_i$, $T_d$, and $m_d$ denote position, velocity, temperature, and mass of a single droplet. $u_i$ and $T_G$ are the local velocity and temperature of the carrier air, $g_i$ is gravitational acceleration, and $L_v$ refers to the latent heat of vaporization. The parameter $f_1$ is the correction to Stokes drag for droplet motion and evaporation and to heat transfer due to evaporation, where $f_1$ is defined as $f_1 = 1.0 + 0.15 Re_d^{0.687} + 0.0175 Re_d / (1 + 4.25 \times 10^4 Re_d^{-1.16})$ (47). The time constant for Stokes flow $\tau_d$ is defined as $\tau_d = \frac{\rho_d D_d^2}{18 \mu_G}$, where $D_d$ is the diameter of a single aerosol and $Nu = \frac{\alpha D_d}{\lambda}$ and $Sh = \frac{h_m D_d}{D}$ are the Nusselt (Nu) and Sherwood (Sh) numbers, where $\alpha$ is convective heat transfer coefficient and $\lambda$ is the thermal conductivity of the air. $h_m$ is convective mass transfer coefficient and $D$ is mass diffusivity of air. $\Delta h^* = \beta(h_\infty - h_v(T_s, Y_s)) / (\exp(\beta) - 1)$, where $\beta$ is defined as $\beta = \ln((1 - Y_\infty)/(1 - Y_{vs}))$.

The $\mu$, $\rho$, $\lambda$, $C_P$, $h$, and $Y$ indicate viscosity, density, thermal conductivity, heat capacity at constant pressure, enthalpy, and the mass fraction, respectively, with the subscripts $i$, $d$, $s$, vs., $\infty$, and $L$ indicating vector components, droplet

properties, properties at droplet surface, saturated vapor state, far-field properties, and liquid (droplet) properties. The code implementation is validated by comparing the evolution of the temperature and radius of droplets to the work of Wong and Lin (48). For more details of numerical methods, the readers are referred to previous works (49, 50).

The computational domain is a pipe of radius 0.5 m and 1.15 m in length ($x \in [-0.15, 1]$). The initial ambient condition is a stagnant flow with a temperature and pressure of 298.15 K and 101,325 Pa. Mass fraction of the water vapor is set as 0.014. As a human head model, a small sphere of radius 0.075 m is included at $x = -0.075$ m at the conduit center with a pathway of air and droplets that lead to the mouth, modeled as a pipe of diameter 2 cm. The inlet of the simulations is located at the other end of this modeled pathway, which coincides with the center of the sphere. Thus, the air jet and droplets are ejected into ambient air through the mouth at the origin. Inlet velocity of plosive sounds is modeled as a rectangular function with a constant value of 0.7, 1.0, 1.3, and 1.6 m/s for 0.05 s, which correspond to the cases shown in Fig. 2*E* and *SI Appendix*, Video S3. Parameters are chosen to yield a similar group dynamic of droplets produced in the experiments.

Additionally, we assume the gas is fully saturated. Droplets are injected throughout the same time duration with the 90% of the background gas velocity. The size distribution of injected droplets are set to follow the log-normal distribution provided by the experiment (e.g., see Fig. 1*F*). The total number of droplets introduced for the entire speech duration is ~400 to ~500. This number of droplets is similar to the experimental results (Fig. 1*D*) and ensures the convergence of the group dynamic statistics.

Algebraic wall model boundary condition is used at the inner wall of the pathway. It assumes an equilibrium turbulent boundary layer and a velocity profile following the law of the wall, leading to a nonlinear equation for the wall stress. Given the flow variables from the near-wall cell, such as velocity, density, and viscosity, the wall stress is solved using Newton's method applied to the inner surface of the pathway. Moreover, the no-slip boundary condition is used for the surface of the head. In contrast, the slip boundary condition is employed for all boundary faces except the face at the downstream location. The Navier-Stokes characteristic boundary condition is used for subsonic nonreflecting outflow.

We use grids of hexagonal close-packed topology. The grid is partially refined near the pathway wall and near the mouth to capture the viscous layer and vortex formation. The refined grid in the former region is 0.6 and 0.3 mm for the latter. Also, an additional refinement zone of grid size 2.5 mm is added for the evolution of the jet. Then, the grid is coarsened to 4 cm toward the edges of the domain resulting in 5.4 million elements. A slice of the domain illustrating meshing is displayed in *SI Appendix*, Fig. S10.

**3D-DIC on Plosive Biomechanics of the Neck Surface.** Neck vibrations of a healthy male (31 y old) during production of plosive and nonplosive sounds as well as swallowing motions are captured by 3D-DIC. The measurement used two of those cameras used in the PTV experiments operating at a frame rate of 500 fps. The image sequences were processed with the open-source 3D-DIC software, MultiDIC (40). The investigation area was $10 \times 9 \times 6$ cm$^3$, covering the entire neck. The DIC subset radius and spacing were 20 and 10 pixels, resolving over 6,600 grids.

**Design and Fabrication of DiP MA Sensor.** For the design of DiP MA sensor, Autodesk EAGLE (version 9.6.0) was used to create a stretchable and bendable MA sensor's fPCB schematic diagram and board layout. The use of serpentine-shaped outlines connecting separate four islands (main body, charging coil, two IMUs) ensured a high degree of mechanical decoupling. A summary of the bill of materials for the device includes a BLE SoC (nRF52840, Nordic Semiconductor), PMIC (power management integrated circuit; BQ25120, Texas Instruments), 8Gb NAND (NOT-AND) flash memory (MT29F8G01, Mm), 3.7V lithium-polymer battery (75 mAh), voltage and current protection IC (integrated circuit) for Li-Polymer battery (BQ2970, Texas Instruments), and two IMUs embedded temperature sensing unit (LSM6DSL, STMicroelectronics). SEGGER Embedded Studio (Release 4.20, SEGGER Microcontroller GmbH) downloaded customized firmware to BLE (Bluetooth Low Energy) IC, then the fPCB with soldered components was folded and connected with the battery. A freeform prototyping machine (Roland MDX 540) created aluminum molds for the top and the bottom layers, made of silicone

elastomer (Silbione-4420, each 300 mm thick, mixed at a ratio of 1:1 base to curing agent) cured at 95 °C in an oven for 20 min. Placing the fPCB between these top and bottom layers, filling the gaps with a soft silicone material (Eco-Flex 0030, 1:2 ratio) and then curing at 60 °C in an oven for 30 min completed the fabrication.

**DiP MA Sensor for Plosive Sound Detection.** All of the participants provided written/verbal consent prior to their participation in this research study. Study procedures were approved by the Northwestern University Institutional Review Board (IRB#: STU00209682). All study-related procedures were carried out in accordance with the standards listed in the Declaration of Helsinki, 1964. During the study, participants wore an MA device at the SN and SCM (Fig. 4*A*). In the case of participants, a clinician/research staff assisted in placing the sensor. A tablet provides the interface to control the MA device wirelessly, allowing downloading of the sensor data through Bluetooth. The tablet uploads the data to the cloud platform through a Wi-Fi interface. *SI Appendix*, Fig. S11 shows representative results of *z* axis MA signals in static (sitting, orange-shaded area) and dynamic (walking, green-shaded area) conditions. A sampling rate of 1.6 kHz successfully captures high-fidelity signatures of vocal cord vibrations manifested through the SN and SCM while pronouncing various words, including plosive and nonplosive consonants. Data from the IMU1 (black) and IMU2 (red), collected from the SN and SCM, respectively, during three different phrases, 1: "Peter piper picked a peak," 2: "Sing a song of six pence," and 3: "We will beat the coronavirus," at four different voice levels (1 is low and 4 is high), are shown in *SI Appendix*, Fig. S8. The oscillatory behaviors for the DiP and single device are captured by the high-speed tracking method (*SI Appendix*, Video S4) (51). The DiP device exhibits 55.0% and 73.3% decreases in natural frequency and damping ratios, respectively, compared with those of a single IMU device. The improvement of its flexibility allows for measuring independent, isolated MA signals in both locations of SN and SCM, minimizing crosstalk while accommodating physiological neck motions.

**Plosive Sound Classification.** Classification of sounds is based on the size of the air passage during the production of the sound. First, vowels and consonants correspond to sounds with a free air passage or blocked or turbulent airflow (14). In phonetic usage, consonants can be classified by the degree of airflow blockage, from large to small: stops [p, t, k, b, d, g, m, n, ng], fricatives [f, v, th, s, z], approximants [w, r], and lateral [l]. Here, the stops represent complete closure of articulators, which blocks the airstream through the mouth or nose. Two types of stops include oral stops (i.e., plosives) [p, t, k, b, d, g], in which airflow is perfectly blocked before speech, and nasal stops [m, n, ng], in which airflow through the oral cavity stops while nasal airflow continues. Moreover, oral stops can be classified into more specific categories according to the phase of articulation: BC [p, b], in which the two lips block the airflow, and the others [t, k, d, g], in which the tongue tip or body blocks the airflow.

**Machine Learning Algorithm.**
*Representative vocabulary for nonplosive words.* English: we, will, well, wall, worry, why, who, where, wow, you, your, raw, rare, he, her, how, high, hire, here, hurry, may, me, my, mine, mining, man, mom, mean, meaning, many, money, mini, minimum, main, moon, name, nanny, nine, none, noun, noon, no, nun, wear, wealth, way, rain, rich, ring, really, relieve, health, hush, hair, fine, five, fall, false, favor, flavor, fever, three, church, sign, sing, song, same, say, sell, sale, slow, vanish, vary, value, jar, jam, join, zoo, zoom, learn, less, lime, lemon, long, line.

Korean: 우리(uri; we, us), 라면(ramyeon; instant noodle), 하늘(haneul; sky), 승진(seungjin; promotion), 주사(jusa; injection), 사자(saja; a lion), 장신(jangsin; tall), 명상(myeongsang; meditation), 해(hae; year), 재(jae; ashes), 새(sae; bird), 신사(sinsa; gentleman), 주말(jumal; weekend), 주일(ju-il; weekday), 놀이(nol-I; play), 해산물(haesanmul; seafood), 논산(nonsan; Nonsan), 전주(jeonju; Jeonju), 울산(ulsan; Ulsan), 청주(cheongju; rice wine), 서울(seoul; Seoul), 사수(sasu; shooter), 모래(molae; sand), 레몬(lemon; lemon), 멜론(mellon; melon), 정치(jeongchi; politics), 맛살(mas-sal; taste), 호수(hosu; lake), 망치(mangchi; hammer), 노루(nolu; Roe deer), 나무(namu; tree), 수치(suchi; shame), 스물(seumul; twenty), 스님(seunim; monk), 소모(somo; consumption), 전체(jeonche; all), 하루(halu; one day), 하산(hassan; climbing down), 해소(haeso; solution), 후추(huchu; pepper),

호소(hoso; appeal), 한숨(hansoom; sigh), 마루(malu; floor), 마흔(maheun; forty), 마련(malyeon; prepared), 명칭(myeongching; designation), 명소(myeongso; sights), 망상(mangsang; delusion), 상추(sangchu; lettuce), 상해(sanghae; wound), 나라(nala; country), 남해(namhae; Namhae), 서해(seohae; West Sea), 선생(seonsaeng; teacher), 성찰(seongchal; reflection), 재해(jaehae; disaster), 함수(hamsu; function), 명시(myeongsi; Express), 추상(chusang; abstract), 상자(sangja; box), 한정(hanjeong; limit), 마음(ma-eum; mind), 사람(salam; person), 사정(sajeong; circumstance), 주제(juje; subject), 낭만(nangman; romance), 사랑(sarang; love), 정체(jeongche; identity), 점수(jeomsu; score), 사물(samul; objects), 사망(samang; dead), 채무(chaemu; financial obligation), 현재(hyeonjae; now), 현장(hyeonjang; scene), 만찬(manchan; feast), 모자(moja; hat), 주체(juche; subject), 멸시(myeolsi; contempt), 추수(chusu; harvest), 매미(maemi; cicada), 차례(chalye; order), 차마(chama; but), 호재(hojae; good news), 소리(soli; sound), 소설(soseol; novel), 소수(sosu; decimal).

*Representative vocabulary for plosive words.* English: pipe, pay, pie, pop, pup, pep, pupa, papa, peep, puppy, peppy, pebble, bay, baby, boy, buy, bye, by, bee, beep, Bob, Bobby, play, person, pillow, pass, pin, praise, pain, poor, plosive, please, poison, pump, plan, ball, bench, beach, blue, belly, below, beef, bean, burn, bunny, bear, balloon, bar, brush, busy, base, key, kit, kid, kick, cat, cow, coke, cake, tie, toy, ticket, date, day, dye, dad, dog, dig, guy, get, gate, good, tell, kill, dizzy, grow, tiny, tire, time, team, tall, toss, toast, trunk, terminal, kind, king, kiss, candle, car, common, come, dinner, doll, drill, dawn, dark, dollar, ghost, gift, girl, guess, gas, green, Greek, glove.

Korean: 폐포(pyepo; alveoli), 포부(pobu; aspiration), 부부(bubu; couple), 피부(pibu; skin), 패배(paebae; defeat), 파(pa; onion), 피(pi; blood), 파이(pai; pie), 표피(pyopi; epidermis), 밥(bab; rice), 법(beob; method), 벼(byeo; rice plant), 표어(pyoeo; watchword), 폭발(pogbal; explosion), 배부(baebu; distribution), 비파(bipa; loquat), 펍(peob; pub), 부피(bupi; volume), 배포(baepo; distribute), 팝(pab; pop), 바보(babo; dumb), 뽀빠이(ppoppai; Popeye), 도끼(dokki; ax), 토끼(tokki; rabbit), 가격(gagyeog; price), 기도(gido; prayer), 고기(gogi; meat), 대구(daegu; Dae-gu), 크기(keugi; size), 도둑(dodug; theif), 도덕(dodeog; moral), 투기(tugi; speculation), 태극기(taegeuggi; Korean flag), 도구(dogu; tool), 독(dog; poison), 깨(kkae; sesame), 가게(gagae; store), 가구(gagu; furniture), 그때(geuttae; then), 카드(kadeu; card), 때(ttae; time), 떡국(tteoggug; rice cake soup), 턱(teog; chin), 코끼리(kokkili; elephant), 개구리(gaeguli; frog), 개미(gaemi; ant), 도랑(dolang; channel), 도미(domi; Sea bream), 대화(daehwa; conversation), 기린(gilin; giraffe), 토론(tolon; debate), 통로(tonglo; passage), 타락(talag; corruption), 통화(tonghwa; currency), 고향(gohyang; hometown), 건전지(geonjeonji; battery), 도화지(dohwaji; drawing paper), 가을(ga-eul; autumn), 겨울(gyeoul; winter), 계란(gyelan; egg), 꾸러미(kkuleomi; package), 고라니(golani; elk), 고난(gonan; hardship), 고명(gomyeong; garnish), 도난(donan; theft), 대학(daehag; university), 다리미(dalimi; iron), 다리(dali; bridge), 대전(daejeon; Daejeon), 기사(gisa; report), 도사(dosa; Taoist), 동대문(dongdaemun; Dong Dae Mun), 트럭(teuleog; truck), 카메라(kamela; camera), 가족(gajok; family), 탈락(tallag; leaving out), 땅거미(ttang-geomi; dusk).

***CNN.*** The CNN starts with four stages in the following order: 32-channel $3 \times 3$ convolution, $1 \times 6$ max pooling, 128-channel $3 \times 3$ convolution, $1 \times 3$ max pooling, 256-channel $3 \times 3$ convolution, $3 \times 3$ max pooling, 512-channel $3 \times 3$ convolution, global max pooling. The model subsequently consists of a fully connected neural network with 512 neurons at the input and a dropout layer ($P = 0.5$) followed by a fully connected neural network with 128 neurons at the input and a dropout layer ($P = 0.5$). At the final output are two neurons representing the probabilities of the two classes. All layers use the ReLU activation. The CNN uses an Adam optimizer for training. The training process follows a leave-one-out strategy, where one leaves a subject out of the training set (nine remaining subjects for training) and then tests the trained model on this subject. Each training set applies a fivefold cross-validation procedure. This approach iterates through each of the 10 subjects. All analyses used Python 3.0 with SciPy and TensorFlow packages.

**Data Availability.** All study data are included in the article and/or supporting information.

Author affiliations: aQuerrey Simpson Institute for Bioelectronics, Northwestern University, Evanston, IL 60208; bCenter for Turbulence Research, Stanford University, Stanford, CA 94305; cDepartment of Electrical and Computer Engineering, University of California - Davis, Davis, CA 95616; dDepartment of Mechanical Science and Engineering, University of Illinois, Urbana, IL 61801; eCascade Technologies Inc., Palo Alto, CA 94303; fBionics Research Center of Biomedical Research Institute, Korea Institute of Science and Technology, Seoul 02792, Republic of Korea; gDepartment of Mechanical Engineering and Materials Science, Pratt School of Engineering, Duke University, Durham, NC 27708; hDepartment of Chemical and Biomolecular Engineering, University of Illinois, Urbana, IL 61801; and iDepartment of Dermatology, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611

1. L. Bourouiba, Fluid dynamics of respiratory infectious diseases. *Annu. Rev. Biomed. Eng.* **23**, 547–577 (2021).
2. L. Bourouiba, The fluid dynamics of disease transmission. *Annu. Rev. Fluid Mech.* **53**, 473–508 (2021).
3. G. Bagheri, B. Thiede, B. Hejazi, O. Schlenczek, E. Bodenschatz, An upper bound on one-to-one exposure to infectious human respiratory particles. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2110117118 (2021).
4. A. Giri et al., Colliding respiratory jets as a mechanism of air exchange and pathogen transport during conversations. *J. Fluid Mech.* **930**, R1 (2022).
5. L. Bourouiba, Turbulent gas clouds and respiratory pathogen emissions: Potential implications for reducing transmission of COVID-19. *JAMA* **323**, 1837–1838 (2020).
6. F. Yang, A. A. Pahlavan, S. Mendez, M. Abkarian, H. A. Stone, Towards improved social distancing guidelines: Space and time dependence of virus transmission from speech-driven aerosol transport between two individuals. *Phys. Rev. Fluids* **5**, 122501 (2020).
7. M. Z. Bazant, J. W. Bush, A guideline to limit indoor airborne transmission of COVID-19. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2018995118 (2021).
8. A. Foster, M. Kinzel, Estimating COVID-19 exposure in a classroom setting: A comparison between mathematical and numerical models. *Phys Fluids (1994)* **33**, 021904 (2021).
9. K. L. Chong et al., Extended lifetime of respiratory droplets in a turbulent vapor puff and its implications on airborne disease transmission. *Phys. Rev. Lett.* **126**, 034502 (2021).
10. C. S. Ng et al., Growth of respiratory droplets in cold and humid air. *Phys. Rev. Fluids* **6**, 054303 (2021).
11. L. Bourouiba, E. Dehandschoewercker, J. W. Bush, Violent expiratory events: On coughing and sneezing. *J. Fluid Mech.* **745**, 537–563 (2014).
12. G. Bagheri et al., Exhaled particles from nanometre to millimetre and their origin in the human respiratory tract. *medRxiv* [Preprint] (2021). https://www.medrxiv.org/content/10.1101/2021.10.01.21264333v1. Accessed 3 October 2021.
13. G. Johnson et al., Modality of human expired aerosol size distributions. *J. Aerosol Sci.* **42**, 839–851 (2011).
14. M. J. Ball, J. Rahilly, *Phonetics: The Science of Speech* (Routledge, 2014).
15. M. Abkarian, S. Mendez, N. Xue, F. Yang, H. A. Stone, Speech can produce jet-like transport relevant to asymptomatic spreading of virus. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 25237–25245 (2020).
16. T. Greenhalgh et al., Ten scientific reasons in support of airborne transmission of SARS-CoV-2. *Lancet* **397**, 1603–1605 (2021).
17. K. Randall, E. T. Ewing, L. C. Marr, J. L. Jimenez, L. Bourouiba, How did we get here: What are droplets and aerosols and how far do they go? A historical perspective on the transmission of respiratory infectious diseases. *Interface Focus* **11**, 20210049 (2021).
18. B. H. Kim et al., Three-dimensional electronic microfliers inspired by wind-dispersed seeds. *Nature* **597**, 503–510 (2021).
19. H. Jeong, J. A. Rogers, S. Xu, Continuous on-body sensing for the COVID-19 pandemic: Gaps and opportunities. *Sci. Adv.* **6**, eabd4794 (2020).
20. C. L. Downey, S. Chapman, R. Randell, J. M. Brown, D. G. Jayne, The impact of continuous versus intermittent vital signs monitoring in hospitals: A systematic review and narrative synthesis. *Int. J. Nurs. Stud.* **84**, 19–27 (2018).
21. K. Lee et al., Mechano-acoustic sensing of physiological processes and body motions via a soft wireless device placed at the suprasternal notch. *Nat. Biomed. Eng.* **4**, 148–158 (2020).
22. H. Jeong et al., Differential cardiopulmonary monitoring system for artifact-canceled physiological tracking of athletes, workers, and COVID-19 patients. *Sci. Adv.* **7**, eabg3092 (2021).
23. A. Krizhevsky, I. Sutskever, G. E. Hinton, ImageNet classification with deep convolutional neural networks. *Commun. ACM* **60**, 84–90 (2017).
24. Y. LeCun, Y. Bengio, G. Hinton, Deep learning. *Nature* **521**, 436–444 (2015).
25. Y. M. Costa, L. S. Oliveira, C. N. Silla Jr., An evaluation of convolutional neural networks for music classification using spectrograms. *Appl. Soft Comput.* **52**, 28–38 (2017).
26. X. Ni et al., Automated, multiparametric monitoring of respiratory biomarkers and vital signs in clinical and home settings for COVID-19 patients. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2026610118 (2021).
27. S. Leglaive, R. Hennequin, R. Badeau, Singing voice detection with deep recurrent neural networks in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE, 2015), pp 121–125.
28. G. C. Ozmen et al., Detection of meniscal tear effects on tibial vibration using passive knee sound measurements. *IEEE Trans. Biomed. Eng.* **68**, 2241–2250 (2021).
29. S. An et al., AdaptNet: Human activity recognition via bilateral domain adaptation using semi-supervised deep translation networks. *IEEE Sens. J.* **21**, 20398–20411 (2021).
30. J. H. Koo et al., Wearable electrocardiogram monitor using carbon nanotube electronics and color-tunable organic light-emitting diodes. *ACS Nano* **11**, 10032–10041 (2017).
31. H. Joo et al., Soft implantable drug delivery device integrated wirelessly with wearable devices to treat fatal seizures. *Sci. Adv.* **7**, eabd4639 (2021).
32. S. Kalenko, A. Liberzon, Particle-turbulence interaction of high Stokes number irregular shape particles in accelerating flow: A rocket-engine model. *Int. J. Multiph. Flow* **133**, 103451 (2021).
33. M. Abkarian, H. A. Stone, Stretching and break-up of saliva filaments during speech: A route for pathogen aerosolization and its potential mitigation. *Phys. Rev. Fluids* **5**, 102301 (2020).
34. S. Balachandar, S. Zaleski, A. Soldati, G. Ahmadi, L. Bourouiba, Host-to-host airborne transmission as a multiphase flow problem for science-based social distance guidelines. *Int. J. Multiph. Flow* **132**, 103439 (2020).
35. M. Song, H. Kim, D. Kim, Vortex ring formation coupled with a translating bluff body. *J. Fluid Mech.* **903**, A8 (2020).
36. A. La Porta, G. A. Voth, A. M. Crawford, J. Alexander, E. Bodenschatz, Fluid particle accelerations in fully developed turbulence. *Nature* **409**, 1017–1019 (2001).
37. J.-T. Kim, J. Shen, S. L. DiMarco, Y. Jin, L. P. Chamorro, Lagrangian acceleration in Rayleigh-Bénard convection at various aspect ratios. *Phys. Rev. Fluids* **3**, 113502 (2018).
38. G. Batchelor, The application of the similarity theory of turbulence to atmospheric diffusion. *Q. J. R. Meteorol. Soc.* **76**, 133–146 (1950).
39. M. Bourgoin, N. T. Ouellette, H. Xu, J. Berg, E. Bodenschatz, The role of pair dispersion in turbulent flow. *Science* **311**, 835–838 (2006).
40. D. Solav et al., An open-source toolbox for multi-view 3D digital image correlation. *IEEE Access* **6**, 30520–30535 (2018).
41. S. S. Kwak et al., Skin-integrated devices with soft, holey architectures for wireless physiological monitoring, with applications in the neonatal intensive care unit. *Adv. Mater.* **33**, e2103974 (2021).
42. J.-T. Kim, J. Nam, S. Shen, C. Lee, L. P. Chamorro, On the dynamics of air bubbles in Rayleigh–Bénard convection. *J. Fluid Mech.* **891**, A7 (2020).
43. C. F. Bohren, D. R. Huffman, *Absorption and Scattering of Light by Small Particles* (John Wiley & Sons, 2008).
44. J. Schaefer, S.-C. Lee, A. Kienle, Calculation of the near fields for the scattering of electromagnetic waves by multiple infinite cylinders at perpendicular incidence. *J. Quant. Spectrosc. Radiat. Transf.* **113**, 2113–2123 (2012).
45. J.-T. Kim, L. P. Chamorro, Lagrangian description of the unsteady flow induced by a single pulse of a jellyfish. *Phys. Rev. Fluids* **4**, 064605 (2019).
46. R. Miller, K. Harstad, J. Bellan, Evaluation of equilibrium and non-equilibrium evaporation models for many-droplet gas-liquid flow simulations. *Int. J. Multiph. Flow* **24**, 1025–1055 (1998).
47. C. Crowe, M. Sommerfeld, Y. Tsuji, *Multiphase Flows with droplets and particles* ( $\bar{Z}$ , 1998).
48. S.-C. Wong, A.-C. Lin, Internal temperature distributions of droplets vaporizing in high-temperature convective flows. *J. Fluid Mech.* **237**, 671–687 (1992).
49. L. Fu, M. Karp, S. T. Bose, P. Moin, J. Urzay, Shock-induced heating and transition to turbulence in a hypersonic boundary layer. *J. Fluid Mech.* **909**, A8 (2021).
50. G. A. Bres et al., Large-eddy simulations of co-annular turbulent jet using a Voronoi-based mesh generation framework in *2018 AIAA/CEAS Aeroacoustics Conference* (2018), pp. 3302.
51. C. Liu et al., Wireless, skin-interfaced devices for pediatric critical care: Application to continuous, noninvasive blood pressure monitoring. *Adv. Healthc. Mater.* **10**, e2100383 (2021).