Language Resources and Evaluation
https://doi.org/10.1007/s10579-024-09789-x

ORIGINAL PAPER

Investigating droplet emission during speech interaction

Francesca Carbone[1,2] · Gilles Bouchet[3] · Alain Ghio[2] · Thierry Legou[2] · Carine André[2] · Muriel Lalain[2] · Caterina Petrone[2] · Antoine Giovanni[2,4]

Accepted: 25 October 2024

Abstract

Conversations (normal speech) or professional interactions (e.g. projected speech in the classroom) have been identified as situations which increase individuals' risk of exposure to respiratory viruses (including SARS-CoV-2) due to the high production of potentially infectious droplets. The few studies addressing this topic contain several methodological and linguistic limitations. This paper describes and validates an original combination of various methods, aimed at providing a global understanding of the complex physiological mechanisms underlying droplet emission during speech production. Twenty-one French speakers produced pseudowords and sentences under different phonetic conditions (e.g. loud vs normal intensity). In Experiment 1, we measured the airflow volume and airflow velocity exhaled from the mouth during speech. In Experiment 2, we measured the airflow velocity exhaled from different positions in the space around the mouth. In Experiment 3, we measured the number and size of expelled droplets. In all experiments, participants were asked to produce pseudowords and sentences under different phonetic conditions in an interactive setting. To validate our methodology, we tested the impact of voice intensity on the physiological measurements. We found that pseudowords and sentences spoken with loud intensity generated increased airflow volume and velocity compared to those spoken with normal intensity. Additionally, the number of droplets was higher for pseudowords spoken with loud intensity compared to normal intensity. From a methodological point of view, our study went beyond previous research by using multiple measures characterising droplet emission during speech. Furthermore, we applied an innovative experimental design, considering droplet emission in an interactive linguistic setting.

Keywords Droplet emission · Aerodynamics · Speech production · Speech physiology · Voice intensity

Extended author information available on the last page of the article

Published online: 03 December 2024

1 Introduction

Individuals are more likely to become infected with respiratory viruses (including SARS-CoV-2 responsible for the Covid-19 disease) the more they are physically exposed to the infectious agents carried by droplets. Their contamination is in fact closely related to the spatial dissemination of droplets emitted by an infected individual. Previous research has identified three key measurable elements characterising the emission of droplets expelled from the respiratory tract (when coughing, sneezing, or talking, for instance): droplet emission rate, droplet size distribution, and the velocity of the exhaled air jet carrying the droplets (e.g. Giovanni et al., 2020; Merghani et al., 2021). Droplet emission rate describes the number of droplets emitted over a specific period (Asadi et al., 2019, 2020). Droplet size distribution is the distribution of droplets depending on their geometric diameters (Morawska et al., 2009; Asadi et al., 2019). Concerning the exhaled air jet carrying the droplets, its displacement in space is typically characterized using two measures: airflow volume and airflow velocity (Merghani et al., 2021). These measures are complementary, offering distinct yet interrelated information about the air jet's movement. Indeed, the airflow volume (or airflow rate), commonly measured in litres per meters (l/m), refers to the quantity or volume of air jet that moves through a specific cross-sectional area in a given amount of time (Gupta et al., 2009, 2010; Merghani et al., 2021). The airflow velocity, commonly measured in meters per seconds (m/s), refers to the speed at which air jet moves in a particular point in the space or environment (Chao et al., 2009; Giovanni et al., 2020; Merghani et al., 2021). Given the non-uniformity of the jet velocity in the space, the airflow velocity measurement is generally made at a single point in the space, corresponding to the centre of the jet emitted from the mouth and reflecting the maximum velocity (Giovanni et al., 2020; Merghani et al., 2021).

Together, the above-mentioned elements are crucial for predicting the trajectory, propagation and spatial distribution of expelled droplets (Giovanni et al., 2020; Merghani et al., 2021). Depending on their size, droplets can assume different velocities and thus variable trajectories once expelled into the air. For instance, large droplets (> 50–100 microns) have ballistic trajectories (Bourouiba et al., 2014; Bourouiba, 2021) and are quickly slowed down by ambient air, falling to the ground within 1–2 s (Wells, 1934). In contrast, intermediate-sized (10–50 microns) and small-sized (< 10 microns) droplets have more unpredictable trajectories as they evaporate quickly and can remain airborne for a long time before settling due to gravity (Bourouiba et al., 2014; Bourouiba, 2021). Due to their unpredictable trajectories and lengthy persistence in the air, small droplets are considered highly contagious and potentially more dangerous than larger-sized droplets (Asadi et al., 2019).

Much previous research has focused on droplet emission during breathing, coughing and sneezing (e.g. Gupta et al., 2009; Tang et al., 2013; Bourouiba et al., 2014; Han et al., 2021), while research into speech droplet emission is still in its early stages. These earlier studies have provided a general understanding of speech droplet emission. Speech seems to release a dramatic number of droplets, significantly more than coughing, sneezing or breathing (Loudon & Roberts, 1967; Papineni & Rosenthal, 1997; Chao et al., 2009; Morawska et al., 2009; Asadi et al., 2019). Speech droplets are mostly small-sized, although intermediate-sized and large-sized droplets can also be expelled while talking (Papineni & Rosenthal, 1997; Morawska et al., 2009; Asadi et al., 2019). These smaller droplets are probably generated during speech from the mucosal layers coating the respiratory tract and from vocal fold adduction and vibration within the larynx (Johnson et al., 2011; Wei & Li, 2016). Furthermore, speech has shown to create a conical, turbulent, jet-like flow and to easily produce directed transport over 2 min short conversations (Abkarian et al., 2020). Due to this high production of small droplets and conical jet-like flow, speech has been determined as more likely to disseminate droplets than sneezing or coughing which mainly carry larger-sized droplets with ballistic trajectories (Stadnytskyi et al., 2021).

1.1 Language-specific properties and droplet emission

Although speech droplet emission is generally understood, only a few recent studies have investigated how this emission is modulated by language-specific phonetic properties. One factor that has been emphasised as crucially influencing speech droplet emission is voice intensity (normal vs loud intensity). A common finding is that loud intensity is positively correlated with the emission rate of droplets (Asadi et al., 2019; Anfinrud et al., 2020; Eiche & Kuster, 2020; Patel et al., 2020; Stadnytskyi et al., 2020). For instance, using an Aerodynamic Particle Sizer (APS), Asadi et al. (2019) showed that the number of droplets produced during loud speech is higher than during normal speech or breathing, regardless of the language used (English, Spanish, Mandarin or Arabic). No significant effect of voice intensity was shown on the size distribution of droplets, with most of the droplets presenting a diameter of nearly 1 micron in both loud and normal intensities. As suggested by Asadi et al. (2019), this increase in the number of droplets may be explained by the physiological mechanisms underlying loud intensity production. This type of production requires a larger exhalation flow rate, i.e. the volume of exhaled air is greater when the intensity is higher, compared to the production of a normal intensity. Consequently, a larger fraction of droplets formed in bronchiolar film rupture may escape from the lungs increasing the concentration of droplets in the exhaled air. Loud intensity also appears to increase the airflow velocity exhaled from the mouth compared to normal intensity (Giovanni et al., 2020). In a preliminary study, Giovanni et al. (2020) analysed the velocity of the exhaled airflow during vocal exercises (e.g. production of vowels with normal and loud intensity) of two e-cigarette users through a propylene glycol cloud. Their findings showed that the speakers produced higher airflow velocities when pronouncing vowels with loud intensity than with normal intensity. In line with the explanation given by Asadi et al. (2019), this acceleration of the airflow may be due to the greater physiological effort needed to expel a larger amount of air involved in the production of a loud voice intensity. From an epidemiological standpoint, these results together suggest that speaking loudly could heighten the extent to which an individual releases infectious pathogen-laden droplets into the air, thereby increasing the risk of transmission to susceptible individuals nearby (Asadi et al., 2019). An airborne infectious disease could thus spread more reliably in a classroom, for instance, where teachers are compelled to speak loudly, as compared to a quiet environment (e.g. a library).

Other factors impacting speech droplet production are the articulation manner of the phonemes (type of consonants and vowels) and voicing (voiced and voiceless sounds). For instance, plosive consonants (such as /t/) have been shown to generate more droplets than fricative ones (e.g. such as /ʃ/; Asadi et al., 2020; Abkarian & Stone, 2020), probably due to the increased physiological effort underlying their production. Similarly, voiced consonants (e.g. /b/) have been shown to lead a greater production of droplets than voiceless consonants (e.g. /p/, Morawska et al., 2009; Asadi et al., 2020), probably because of greater droplet generation within the larynx due to the activation of vocal folds (Morawska et al., 2009).

1.2 Methodological and linguistic limitations of previous studies

Previous literature investigating the impact of specific language properties on droplet emission has shown several methodological limitations. So far, research has mostly considered the impact of speech properties (e.g. loud vs normal intensity) on isolated physiological measures. To our knowledge, no studies have simultaneously measured multiple elements characterising droplet emission (e.g. number and size of droplets, airflow volume and velocity) for the same speech material and speakers. Since such elements are interconnected, the investigation of their relationships is fundamental for a more global understanding of droplet emission. Furthermore, most part of previous studies on speech droplets focused on the number of exhaled droplets (or droplet emission rate, e.g. Asadi et al., 2019), while other measures that are crucial for characterising droplet dissemination (e.g. airflow volume, airflow velocity; Gupta et al., 2010; Giovanni et al., 2020; Merghani et al., 2021) have been overlooked or not adequately considered due to technical constraints. For instance, Giovanni et al. (2020) employed a propylene glycol cloud in their investigation of the airflow velocity exhaled during vocal exercises. This device, however, only delivers 0.3 micron droplets (Bertholon et al., 2013; Giovanni et al., 2020), which does not cover the whole range of droplet sizes expelled through speech. A few other studies have utilised the Particle Image Velocimetry (PIV) technique to measure airflow velocity during speech (e.g. Chao et al., 2009; Kwon et al., 2012; Abkarian et al., 2020; Abkarian & Stone, 2020; Han et al., 2021). The PIV is an optical method of flow visualisation used to obtain instantaneous velocity measurements through the combination of a pulsed laser light sheet and a synchronized CCD camera (Merghani et al., 2021). However, the achievable accuracy of PIV measurements has proved to be limited. The intensity variations that occur in PIV experiments due for instance to the motion of the droplets in the intensity profile of the light sheet or misalignments of the two light pulses limit the obtainable accuracy of PIV measurements, even under otherwise ideal conditions (Nobach & Bodenschatz, 2009).

Previous research also has several limitations from a linguistic point of view. Earlier studies have mostly employed speech material consisting of single words, syllables and isolated phonemes (mostly in English), resulting in low ecological validity. In the few cases where utterances and texts were used, they were usually pronounced outside of a communicative or interactional context (e.g. Asadi et al., 2019; Asadi et al., 2020; Abkarian & Stone, 2020) and were thus unrepresentative of everyday conversations. Moreover, so far only the impact of a few linguistic properties on droplet emission has been considered. For example, previous studies have completely neglected the effects of prosody (e.g. speech melody resulting from activation of the vocal folds) and its interaction with other phonetic properties on droplet emission.

Thus far, the above-mentioned methodological and linguistic limitations make it challenging to accurately predict droplet emission under specific speech conditions. A more holistic approach is required by simultaneously considering multiple key elements characterising droplet emission. In addition, various (segmental and prosodic) speech properties should be manipulated within a linguistic setting aiming at replicating the dynamics of real-life conversations as much as possible.

1.3 Aim and proposal to previous limitations

In this paper, we outline and validate the methodology of three experiments aimed at enhancing our understanding of droplet emission during speech interaction. We sought to overcome both the methodological and linguistic limitations of previous literature.

To our knowledge, this is the first study analysing multiple physiological and aerodynamic measures characterising droplet emission (droplet size, number of droplets, airflow volume, airflow velocity) for the same speech materials (pseudowords, sentences) and speakers under specific speech conditions. Together, these measures represent crucial elements for accurately modelling the trajectory, propagation and spatial distribution of expelled droplets (Giovanni et al., 2020; Merghani et al., 2021). Given their interconnections (Merghani et al., 2021), exploring the relationships between these elements provides deeper insights into the dynamics underlying droplet emission. Additionally, the analysis of complementary measures, such as airflow volume and velocity, ensures a multiparametric correlation, and, thereby, a cross-validation of the results.

Furthermore, we measured airflow velocity during speech production via a more accurate technique than those previously used, i.e. hot-wire anemometry. This technique has never been applied into research on speech droplet emission. It has been widely used by research scientists and engineers in fluid dynamics to measure the velocity and direction of the air (e.g. Comte-Bellot, 1976; Bruun et al., 1988). A hot-wire anemometer consists of a thermal transducer that permits instantaneous flow velocity in metres per second to be calculated from electric voltage measurements (Bruun, 1995). The advantage of the hot-wire anemometer is associated with its very high temporal resolution and excellent frequency response characteristics (Barratt et al., 2016). We measured the airflow velocities at three points into the space. Recording from different positions in the space was motivated by the fact that the air jet exhaled during speech production is not uniform in space or in time.

This non-uniformity results in variable airflow velocities depending on the position in space (Giovanni et al., 2020).

From a linguistic point of view, our study is to our knowledge, the first to investigate the emission of droplets during speech production in an interactive setting, here based on dictations to an interlocutor. Due to the constraints of the measurements, we developed an experimental paradigm using restricted forms of interaction based on partially pre-recorded and partially live interactions. Furthermore, while most previous research has been conducted on English, this is the first study to provide speech droplet emission data for French. Hence, we aimed at contributing to understanding the extent to which previous results can be generalised across languages. The linguistic variables manipulated in this study were aimed at evaluating the impact of voice intensity, segmental and prosodic characteristics (as well as the interaction between these factors) on droplet emission during the production of pseudowords and sentences. We used the voice intensity variable (normal vs loud intensity) as a methodological validation standard for our study. Results concerning the other variables (articulation manner of phonemes, prosody) are not reported here as these aspects fall outside the methodological scope of the current paper.

In line with preliminary findings from Giovanni et al. (2020), we predicted that higher airflow volume and airflow velocity would be found for loud intensity compared to normal intensity. We also expected a greater number of droplets for loud intensity compared to normal intensity, consistent with previous studies on droplet emission rate (e.g. Asadi et al., 2019; Anfinrud et al., 2020). Furthermore, based on findings from Asadi et al. (2019), we expected the size distribution of droplets to be constant across intensity conditions, with most droplets presenting a geometric diameter of nearly 1 micron.

2 Materials and methods

Three experiments were run. In Experiment 1, we measured airflow volume and airflow velocity exhaled during speech. In Experiment 2, we measured the exhaled airflow velocity during speech at different positions in space when exiting the mouth. In Experiment 3, we measured the number and size of droplets expelled during speech. Each experiment included two linguistic tasks. Linguistic Task 1 examined the effects of voice intensity and the articulation manner of phonemes (type of consonants and vowels) on droplet emission. Linguistic Task 2 focussed on the effects of prosody (prosodic focus) and its interaction with voice intensity and the articulation manner of phonemes. We manipulated the position of the most prominent word in the sentence, as prominence leads to higher degrees of physiological effort at the word level (Gussenhoven, 2016), potentially increasing droplet production.

2.1 Participants

Twenty-one female native French speakers aged between eighteen and fifty-five years old (mean age: 25.23, SD 7.59), with no known respiratory or speaking disorders, participated in the three experiments. We chose to focus exclusively on female speakers (excluding male speakers) to minimise the variability of our sample and reinforce the internal validity of our findings. This choice was also motivated by the fact that vocal effort underlying the production of loud voice is particularly found among women (Roy et al., 2004a), probably because they predominantly hold occupations requiring vocal projection (e.g. school teaching; Roy et al., 2004b). All participants were asked to complete a brief survey including age, weight, height, general health status and smoking history as these variables could influence breathing and speech production. The study complied with the Declaration of Helsinki and was approved by the Ethics Committee of Aix-Marseille University (CER AMU 2021-03-11-06). Speakers signed an informed consent form before the experiments.

2.2 Materials

2.2.1 Linguistic Task 1: Pseudowords

Twelve pseudowords were used in this task. They were composed of three syllables contrasted by voicing (voiced vs voiceless consonants, e.g. *pataka*, *bagada*), consonant type (plosives vs fricatives, e.g. *badaga*, *vazaja*) and vowel type (low vowel /a/, back vowel /u/, high-front vowel /i/, e.g. *pataka*, *poutoukou*, *pitiki*). We excluded nasal consonants from the corpus as they involve the emission of airflow from the nose which can affect the measurement of the airflow velocity from the mouth. The full list of pseudowords used in the experiments is shown in Table 1.

2.2.2 Linguistic Task 2: Sentences

Twenty question/answer pairs were used in this task, ten for each prosodic condition (‘focus’ condition, ‘no-focus’ condition). In the ‘focus’ condition, questions induced the participants to reply by focussing on the target words. In the ‘no focus’ condition, questions did not trigger prosodic focus on any of the sentence constituents. Five sentences were employed for each consonant type. Sentences were composed mostly of voiced consonants with the same manner of articulation as the target words (either plosives or fricatives) to limit the influence of consonant type on the release of droplets. Target words were the pseudowords employed in Linguistic Task 1 which were embedded in carrier sentences (e.g. ‘focus condition’: *Est-ce que*

Table 1 Pseudowords used in Linguistic Task 1

Consonants	Vowels	Plosives	Fricatives
Voiceless	/a/	Pataka	Fassacha
	/u/	Poutoukou	Foussouchou
	/i/	Pitiki	Fissichi
Voiced	/a/	Badaga	Vazaja
	/u/	Boudougou	Vouzoujou
	/i/	Bidigui	Viziji

Christian donne la dague? (Non). MONSIEUR BADAGA donne la dague, 'Is Christian giving the dagger away? (No), Mr BADAGA is giving the dagger away'; 'no focus' condition: *Qu'est-ce qu'il se passe? Monsieur Badaga donne la dague*, 'What is going on? Mr Badaga is giving the dagger away'). In the 'focus' condition, speakers were instructed not to utter the '(No)' indicated in parentheses, as its sole purpose was to elicit a more spontaneous emphasis on the target words. The full list of question/answer pairs is shown in Table 2.

2.3 Procedure

Each speaker carried out the three experiments over two days within one week. We divided the experiment into two days to avoid exposing the speakers to a prolonged vocal effort from producing the linguistic material (pseudowords, sentences) with loud intensity. On the first day, the speakers participated in two experiments, while the remaining experiment was carried out on the second day. The order of the experiments was counterbalanced. In Experiment 1 and 2, speakers were instructed to perform the two linguistic tasks (production of the pseudowords and sentences), which were presented in random order. In Experiment 3, speakers were asked to produce only pseudowords (Linguistic Task 1). The participants were not asked to produce the sentences (Linguistic Task 2) due to the difficulty in adapting this task to the APS recordings (see below). Each experimental session presented a duration of about forty minutes.

Before starting the experiments, speakers were told that they would be connected online with an interlocutor who would transcribe their speech productions (pseudowords, sentences) to test a telecommunications system. At the beginning of the experiment, all speakers were shown the same pre-recorded video of a fictitious interlocutor. The fictitious interlocutor was one of the experimenters: a 48-year-old female French native speaker who was not present in the experimental room. We led each participant to believe that the interaction with the fictitious interlocutor was live. The fiction was further played out through the following elements. First, at the end of each repetition, speakers were shown an identification score of the pseudowords/sentences transcribed correctly by the (fictitious) interlocutor. Second, the experimenter fed into the fiction by pretending to talk frequently with the fictitious interlocutor, giving her feedback on the identification score. At the end of the experiments, we asked the participants whether they had understood that the interaction was not real; none of the currently tested participants had caught on to the fact that the interaction did not actually take place.

2.3.1 Linguistic Task 1: pseudowords

Speakers were instructed to dictate the pseudowords to the fictitious interlocutor. Speakers were asked to produce them with two levels of intensity (normal vs loud). Voice intensity level was controlled by the speakers themselves by means of a Standalone Feedback Vumeter. They were instructed to produce an intensity which reached the yellow band of the Vumeter, corresponding to either loud or normal intensity. Speakers were asked to keep their voice intensity constant as best they could. Each intensity level corresponded to a separate block, with the blocks presented in random order. Within each block, pseudowords were presented randomly. They were each spoken five times consecutively with a neutral prosody. Speakers had to pronounce the pseudowords in sync with their appearance on the screen. The pseudowords were presented in a ‘karaoke’ format, with the five repetitions displayed consecutively at a uniform pace. This presentation modality helped minimising individual differences in speech rate. Each block contained three repetitions of the complete list of stimuli. Speakers were given a short break at the end of each repetition to take a sip of water to stay hydrated. A training session preceded the experiment. For Experiment 1 and 2,

Table 2 Question/answer pairs used in Linguistic Task 2

Focus condition		
Plosives	*Est-ce que Robert donne la dague ?* ‘Is Robert giving the dagger away?’	*(Non). MONSIEUR BADAGA donne la dague*. ‘(No). MR BADAGA is giving the dagger away’.
	Est-ce que Michel déduit la dette ? ‘Is Michel deducting the debt?’	*(Non). MONSIEUR BADAGA déduit la dette.* ‘(No). MR BADAGA is deducting the debt’.
	Est-ce que David goûte le gâteau ? ‘Is David tasting the cake?’	*(Non). MONSIEUR BADAGA goûte le gâteau.* ‘(No). MR BADAGA is tasting the cake’.
	Est-ce que Xavier gobe les bonbons ? ‘Is Xavier gobbling up the sweets?’	*(Non). MONSIEUR BADAGA gobe les bonbons.* ‘(No). MR BADAGA is gobbling up the sweets’.
	Est-ce que Christian boit la bibine ? ‘Is Christian drinking the booze?’	*(Non). MONSIEUR BADAGA boit la bibine.* ‘(No). MR BADAGA is drinking the booze’.
Fricatives	*Est-ce que Robert vise le voleur ?* ‘Is Robert targeting the thief?’	*(Non). MONSIEUR VAZAJA vise le voleur.* ‘(No). MR VAZAJA is targeting the thief’.
	Est-ce que Michel venge le voisin ? ‘Is Michel avenging the neighbour?’	*(Non). MONSIEUR VAZAJA venge le voisin.* ‘(No). MR VAZAJA is avenging the neighbour’.
	Est-ce que Xavier jette la valise ? ‘Is Xavier throwing away the suitcase?’	*(Non). MONSIEUR VAZAJA jette la valise.* ‘(No). MR VAZAJA is throwing away the suitcase’.
	Est-ce que David vole le vélo ? ‘Is David stealing the bike?’	*(Non). MONSIEUR VAZAJA vole le vélo.* ‘(No). MR VAZAJA is stealing the bike’.
	Est-ce que Christian joue le joker ? ‘Is Christian playing the joker?’	*(Non). MONSIEUR VAZAJA joue le joker.* ‘(No). MR VAZAJA is playing the joker’.
No focus condition		
Consonant group	Question	Answer

Table 2 (continued)

Focus condition		
Plosives	*Qu'est-ce qu'il se passe ?* 'What is going on?'	*Monsieur Badaga donne la dague* 'Mr Badaga is giving the dagger away'
		Monsieur Badaga déduit la dette 'Mr Badaga is deducting the debt'
		Monsieur Badaga goûte le gâteau 'Mr Badaga is tasting the cake'
		Monsieur Badaga gobe les bonbons 'Mr Badaga is gobbling up the sweets'
		Monsieur Badaga boit la bibine 'Mr Badaga is drinking the booze'
Fricatives		*Monsieur Vazaja vise le voleur* 'Mr Vazaja is targeting the thief'
		Monsieur Vazaja venge le voisin 'Mr Vazaja is avenging the neighbour'
		Monsieur Vazaja jette la valise 'Mr Vazaja is throwing away the suitcase'
		Monsieur Vazaja vole le vélo 'Mr Vazaja is stealing the bike'
		Monsieur Vazaja joue le joker 'Mr Vazaja is playing the joker'

72 pseudowords were collected for each participant [12 pseudowords x 2 intensity levels (normal vs loud) x 3 repetitions]. In Experiment 3, a modified version of this task was performed. First, speakers were told to quietly breathe through their mouth over one minute. Then, they were asked to pronounce the pseudowords without pauses consecutively for seven seconds. Of the seven seconds, we recorded the five central ones. This modification was necessary to avoid the influence of expiration (during pauses) on the number and size of speech droplets. Moreover, breathing recordings were used as baseline to normalise speech droplet measurements. By doing so, we accounted for individual differences and neutralised the effect of droplets (extraneous to phonation) that might be present in the air.

2.3.2 Linguistic Task 2: sentences

An elicitation procedure was designed to obtain variation in sentence-level prominence naturalistically using a question-answer paradigm. Speakers were instructed to answer pre-recorded questions which induced them to reply by focussing on the target words ('focus' condition) or not ('no focus' condition). For the 'focus' condition, target words were indicated in uppercase on the screen. As in Linguistic Task 1, they pronounced the sentences with two levels of voice intensity (normal vs loud) within two separate blocks, presented in random order. Speakers had to pronounce the sentences in sync with their appearance on the screen. Each block contained three repetitions in which trials were presented randomly. Speakers were given a short break after each repetition to take a sip of water. The experiment was preceded by a training session. In total, we collected 120 sentences for each participant [10 sentences x 2 focus conditions ('focus' vs 'no focus') x 2 intensity levels (normal vs loud) x 3 repetitions]. The production of sentences (Linguistic Task 2) was excluded from Experiment 3 due to the speakers' difficulty in pronouncing whole sentences without inhaling or exhaling for several seconds.

2.4 Setting and apparatus

2.4.1 General apparatus

Several workstations were used to conduct the experiments. In all three experiments, an Interactional Workstation was used to visually display the linguistic material (pseudowords and question/answer pairs) on a screen. The material was presented by means of PERCEVAL software (André et al., 2003) and displayed to the participants on a screen connected by a VGA cable to the computer. Audio recordings were made with a Neumann TLM 102 Cardioid Condenser microphone connected to an acoustic input of the Data Acquisition Workstation. The sound was recorded with a sampling at 25 kH−16bits. The speakers wore professional-grade headphones to listen to the pre-recorded questions (see Linguistic Task 2). A Standalone Feedback Vumeter was used through a Samsung Tablet (Android Application 'Audio Level', Trajkovski Labs) to control speakers' level of voice intensity. The Vumeter was placed next to the screen to allow participants to check their level of voice intensity while performing the linguistic tasks. To induce a change in vocal effort, the gain of the tablet microphone was adjusted: low gain to induce loud speech vs high gain to induce normal speech. In Experiment 2, we also used a Video Controller composed of two web cameras to film the speakers' mouths to ensure that they constantly maintained the same position during data acquisition.

In Experiment 1 and 2, an Anemometer Controller and a Data Acquisition Workstation were used for data recordings and acquisition. The Anemometer Controller was a StreamLine Pro CTA system (Dantec Dynamics). This device was connected to the probes (Dantec 55P11 dimensional probes), a thermocouple, a Multifunction I/O Device National Instruments via BNC and a computer equipped with StreamLine software via a RS232 protocol. The Anemometer Controller outputs were plugged into the auxiliary inputs of the Data Acquisition Workstation. The Data Acquisition Workstation corresponded to the EVA2 workstation (SQLab-LPL, Aix-en-Provence, France, Ghio et al., 2012; Fig. 1, Data Acquisition Workstation) and was used for simultaneous data acquisition (wave sound, airflow volume and velocity). Data were recorded, displayed and analysed with the Phonedit software developed by the LPL (https://lpldev.lpl-aix.fr/phonedit/). Airflow velocity values obtained in volts through the Anemometer Controller were converted to velocity in m/s by applying the King's Law (King, 1914; see e.g. Bruun et al., 1988, Brunn, 1995): E^2 (volts) = A + B. Velocity^n, where A, B and n were obtained empirically by using a calibrated system. In our case, we found that A = 1.3905, B = 0.9535 and n = 0.496 (in line with n = 0.44 in Brunn, 1988). The velocity value in m/s was obtained by the reciprocal function:

$$\text{Velocity(m/s)} = \exp\left(\frac{\ln\left(\frac{E^2 - A}{B}\right)}{n}\right)$$

The airflow volume and velocity were recorded with a sampling frequency equal to 6250 Hz. In Experiment 3, data recording and acquisition were obtained directly through the Aerodynamic Particle Sizer (APS, TSI model 3221) and the TSI Aerosol Instrument Manager Software for APS (5) [Aerodynamic Particle Size Analyzer 3300, Instruction Manual (Serial Number = 145) (1983), TSI Inc., St. Paul, MN]. Before each session participants quietly breathe through their mouth over one minute. The number of droplets recorded during breathing was used as baseline for the normalisation of the number of droplets produced for each pseudowords using the following formula:

$$\frac{\text{no. of droplets (speech) - no. of droplets (breathing)}}{\text{no. of droplets (breathing)}} * 100,$$

where *no. of droplets (speech)* corresponded to the number of droplets produced for each pseudoword (spoken for 5 s), while *no. of droplets (breathing)* corresponded to the number of droplets produced during breathing over 5 s. This latter was calculated

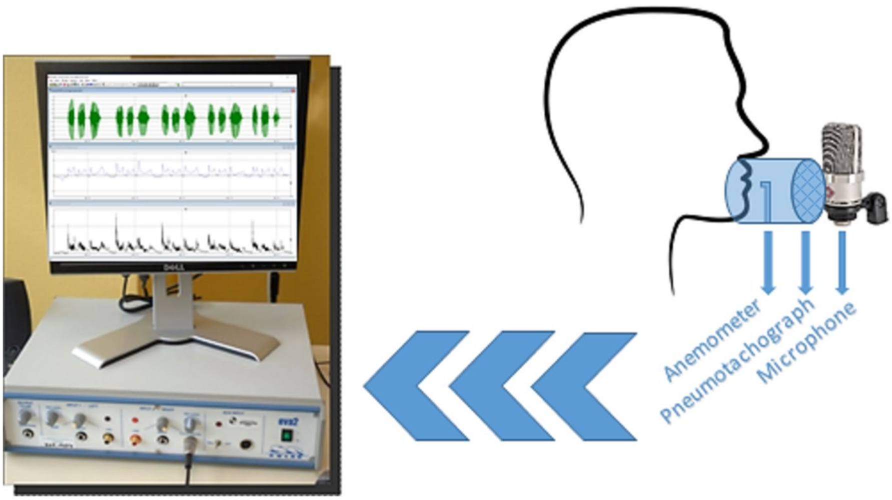

Fig. 1 Schematic representation of Experiment 1. Data were recorded simultaneously through built-in features of the EVA2 workstation: airflow volume (l/s) was recorded by a pneumotachograph; airflow velocity (m/s) by an anemometer; wave signal and SPL intensity by a microphone

from the one-minute breathing recordings through this formula: (no. of droplets during one-minute breathing/60 s) * 5 s. Details about the experimental set-up of each experiment are given in the following paragraphs.

2.4.2 Experiment 1

The goal of this experiment was to simultaneously measure the exhaled airflow volume (l/s) and the airflow velocity (m/s) during speech production. The representation of the experiment is reported in Fig. 1.

The airflow volume (l/s) was recorded through a pneumotachograph with a stainless-steel wire (Ghio & Teston, 2004). This device's reduced size (three cm diameter and two cm length) optimises its response time and linearity in all the articulatory contexts. Associated with highly sensitive and stable differential pressure transducers (Honeywell DCXL), the device can capture a flow of the order of one cm3/s. The resistance of the grid was 10 Pa by dm3/s, or approximately 1% of the intra-oral pressure of a normal subject, so as not to disturb the functionality of the vocal tract. The pressure tap was made in eight points of the circumference of the measurement pipe, and a grid of tranquillisation (negligible in resistance) was laid out in front of the pressure taps. This reduced the non-linearity of measurement caused by aerodynamic turbulence produced during speech production. The pneumotachograph, which is built into the EVA2 workstation, provided directly calibrated data using SESANE Phonedit software (Ghio et al., 2012). This device required speakers to wear a soft silicone rubber mask (Ghio & Teston, 2004) during the linguistic tasks to prevent air leakage without hindering articulatory movements.

The airflow velocity at the outlet of the mouth was measured using a hot-wire anemometer (Jørgensen, 2002) providing good time resolution (Bruun, 1995). The device contains a wire that is electrically heated and cooled by the passage of the airflow. The cooling determines a change in the resistance of the wire and provide measurements of the airflow velocity. This device, operated through the StreamLine Pro CTA system (Dantec Dynamics), was embedded into the pneumotachograph. Specifically, the anemometer consisted of a miniature wire probe with one sensor with a five μm diameter (Dantec 55P11) forming the heating element. The probe consisted of 1.25 mm-long plated tungsten wire sensors suspended between two straight prongs. The temperature of the exhaled air was measured simultaneously using a thermocouple, as the operating principle is based on cooling by forced convection of a small tungsten wire heated by the passage of an electric current. This device was placed at the outlet of the mouth parallel to the centre of the mouth so that the maximum airflow velocity was recorded. The output of the StreamLine Pro CTA system was an analogue electrical signal connected to an auxiliary input in the EVA2 workstation.

Simultaneously, we also recorded the wave signal and SPL intensity in dB (independent of the recording gain). The wave signal was recorded with a Neumann TLM 102 microphone placed 15 cm away from the corner of the speaker's lip. The microphone was placed on the side so as not to disturb the airflow measurements. The EVA2 workstation had a built-in sonometer that allowed us to compare calibrated

SPL intensities between different recordings independently from the recording gain level.

The outcome variables from Experiment 1 were peak values for the SPL intensity (dB), the airflow volume (l/s), and airflow velocity (volt). Airflow velocities in volts were transformed in m/s prior to statistical analysis.

2.4.3 Experiment 2

The goal of this experiment was to simultaneously measure the exhaled airflow velocity (m/s) at different positions in the space around the mouth during speech production. The representation of the set-up for Experiment 2 is reported in Fig. 2.

The airflow velocity at the outlet of the mouth was measured using three hot-wire anemometers with one-sensor probes (Dantec 55P11) operated through the StreamLine Pro CTA system (Dantec Dynamics, see Experiment 1) and connected to the EVA2 workstation. They were separately fitted on a Dantec probe holder, arranged vertically and parallel to the speaker's mouth. Anemometer 1 was placed above the speaker's lips, anemometer 2 parallel to the centre of the lips and anemometer 3 below the lips. The vertical distance was 1.5 cm between anemometer 1 and 2, and 2 cm between anemometer 2 and 3. The horizontal distance between the three anemometers and the speaker's lips was 5.75 cm. This arrangement was thought to cover the whole cone of airflow emitted from the mouth during speech (Abkarian et al., 2020; Giovanni et al., 2020).

The participant's position was controlled visually by the experimenter through two webcams focussing on the mouth in front and sagittal positions. A headrest helped the speakers to maintain the same head position, limiting their movements. Unlike Experiment 1, in Experiment 2, speakers were not required to wear a mask to measure exhaled airflow velocity without compressing the mouth. As in Experiment 1, the wave signal and the SPL intensity in dB (independent of the recording gain) were recorded.

Fig. 2 Schematic representation of Experiment 2. Data were recorded simultaneously through built-in features of the EVA2 workstation: airflow velocity exhaled in three positions from the mouth were recorded by the three anemometers; wave signal and SPL intensity by a microphone

The outcome variables from Experiment 2 were peak values for the SPL intensity (dB) and airflow velocities (volt) recorded by anemometer 1, anemometer 2 and anemometer 3. As in Experiment 1, airflow velocities in volts were converted to m/s.

2.4.4 Experiment 3

The goal of this experiment was to measure the size and number of droplets expelled during speech. The recordings of droplets were made through an Aerodynamic Particle Sizer (APS, TSI model 3321), a state-of-art instrument (Asadi et al., 2019, 2020). The representation of the set-up for Experiment 3 is reported in Fig. 3.

Data were recorded using TSI Aerosol Instrument Manager Software for APS [Aerodynamic Particle Size Analyzer 3300, Instruction Manual (Serial Number = 145) (1983), TSI Inc., St. Paul, MN]. The APS operates at a total flow rate of 5 L/min (sheath flow rate ≅ 4 L/min, sample flow rate ≅ 1 L/min) and measures the size distribution of particles larger than 0.5 µm. It also detects the presence of particles between 0.37 µm and 0.5 µm, however, without providing precise size measurements. For this reason, particles with a geometric diameter between 0.37 µm and 0.5 µm were not included in the analyses (Asadi et al., 2019).

During the experiment, participants sat in front of the APS and spoke into a metallic funnel (diameter = 14 cm) that was connected to the APS sampling inlet via a conductive silicon tube (distance between funnel hole and APS inlet = 12 cm, tube inner diameter = 2 cm). A tennis ball supported by a stand was placed behind the participants' heads to ensure that they kept their heads in the same position during the linguistic tasks. A screen showing the stimuli was placed behind the APS to guide participants in accomplishing the tasks. Speech production was recorded by a voice recorder (Zoom H1n) situated directly on the side of the APS.

We conducted the experiments in a controlled environment. The air of the room was filtered to reduce the number of particles by means of an air filtering system composed of a High-Efficiency Particulate Air (HEPA) filter attached to a ventilator. Control experiments were conducted to ascertain whether the air in the experimental room was effectively purified by the implemented system. Using the APS, we measured the

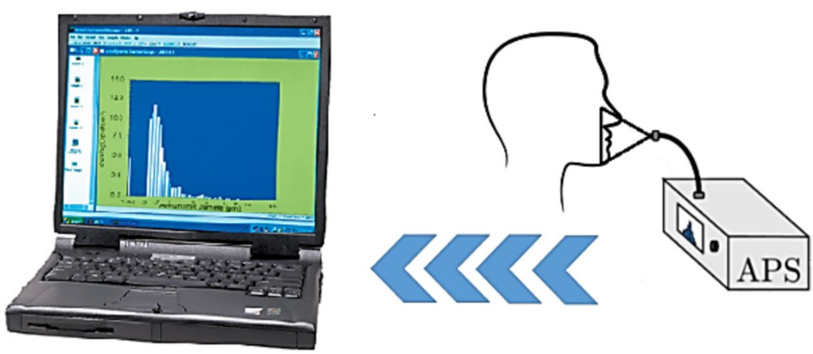

Fig. 3 Schematic representation of Experiment 3. Data were recorded by the Aerodynamic Particle Sizer (APS) and TSI Aerosol Instrument Manager Software

particle concentration in the room for one minute before and after activation of the air filtering system. Before the activation of the air filtering system, the number of particles measured by the APS in the experimental room was 20.606. After three hours of system operation, the number of particles decreased significantly to 206 (99% reduction). Before each experimental session, the air filtration system was activated for three hours and remained in constant operation throughout the entire duration of the experiment.

The outcome variables from Experiment 3 were the normalised number of droplets split by droplet size (fifty-one levels from 0.5 to 20 μm).

2.4.5 Segmentation of speech material and data extraction

During Experiments 1 and 2, acoustic (wave signal; SPL intensity) and physiological measures (airflow volume in l/m and airflow velocities in m/s) were continuously recorded. The speech material was segmented in two steps. First, the data were automatically segmented in SPPAS (SPeech Phonetization Alignment and Syllabification; Bigi, 2015) using its tools for silence detection, TextGrid creation and manual orthographic transcription (IPU Transcribe). Transcribed TextGrid intervals corresponded to individual pseudowords (spoken 5 times consecutively), and sentences separated by silence. Tokens were typically less than 1 s in length. The automatic segmentations were then corrected manually. A total of 7584 tokens were obtained from Experiment 1 and Experiment 2 (3756 tokens from Experiment 1: 1356 pseudowords, 2400 sentences; 3828 tokens from Experiment 2: 1428 pseudowords, 2400 sentences). The intended total was 8064 tokens, including 3024 pseudowords (21 speakers × 12 pseudowords × 2 intensity levels × 3 repetitions × 2 experiments) and 5040 sentences (21 speakers × 20 sentences × 2 intensity levels × 3 repetitions × 2 experiments). However, 480 tokens were missing due to data loss (e.g. rupture of hot wires during the experiments). Scripts were applied to extract the physiological measurements (airflow volume peaks in l/m, airflow velocity peaks in m/s) corresponding to the tokens. Concerning Experiment 3, the APS allowed us to record the number and size of droplets separately for each pseudoword spoken for 5 s. No data loss was observed, and 1512 pseudowords were collected.

2.5 Statistical analyses

Linear mixed models were performed by means of the R-environment (R Development Core Team, 2012, version 3.5.1, package R *lme4 version 1.1–31*, package *ggplot2* version 3.4.0; Bates et al., 2015; Wickham, 2016). Prior to statistical analysis, airflow velocity and volume values were log-transformed to achieve a normal distribution. Data were plotted without logarithmic transformation, however, to be shown in the original scale.

2.5.1 Experiment 1

We entered as dependent variables (DVs) in separate analyses: (i) SPL intensity peaks (dB); (ii) airflow volume peaks (l/s); and (iii) airflow velocity peaks (m/s).

Separate analyses were performed for pseudowords (Linguistic Task 1) and sentences (Linguistic Task 2). For each model, we tested the fixed factors INTENSITY (loud/normal). We also included REPETITION (3) as a control variable. For Linguistic Task 1, random intercepts for SPEAKER (1–23) and PSEUDOWORD (1–12) were included, as well as by-speaker and by-pseudoword random slopes for INTENSITY. For Linguistic Task 2, random intercepts for LISTENER (1–23) and SENTENCE (1–20) were included, as well as by-listener and by-sentence random slopes for INTENSITY.

2.5.2 Experiment 2

We entered as DVs in separate analyses: (i) SPL intensity peaks (dB); (ii) airflow velocity peaks (m/s) recorded from anemometer 1; (iii) airflow velocity peak recordings from anemometer 2; and (iv) airflow velocity peak recordings from anemometer 3. As in Experiment 1, separate analyses were performed for pseudowords (Linguistic Task 1) and sentences (Linguistic Task 2). The same fixed and random factors as well as random slops as for Experiment 1 were applied.

2.5.3 Experiment 3

We entered the raw number of droplets as DVs. INTENSITY (two levels: loud/normal), SIZE of droplets (μs, continuous) as well as their interaction were the fixed factors. We also included REPETITION (3) as a control variable. Random intercepts for SPEAKER (1–23) and PSEUDOWORD (1–12) were included, as well as by-speaker and by-pseudoword random slopes for INTENSITY.

3 Results

3.1 Experiment 1

Figure 4 illustrates the results for the SPL intensity (dB), airflow volume (l/s) and airflow velocities (m/s) according to voice intensity (loud, normal) for pseudowords and sentences. Loud intensity was associated with increased values compared to normal intensity. In particular, pseudowords and sentences presented higher SPL intensity when spoken with loud intensity than when spoken with normal intensity [pseudowords spoken with loud vs normal intensity: ($\beta=9.09$, $SE=0.95$, $t=9.52$, $p<0.001$); sentences spoken with loud vs normal intensity: ($\beta=6.44$, $SE=0.59$, $t=10.81$, $p<0.001$)], as well as a higher airflow volume [pseudowords spoken with loud vs normal intensity: ($\beta=0.01$, $SE=0.01$, $t=4.06$, $p<0.001$); sentences spoken with loud vs normal intensity: ($\beta=0.01$, $SE=0.01$, $t=4.36$, $p<0.001$)] and a higher airflow velocity [pseudowords spoken with loud vs normal intensity: ($\beta=0.01$, $SE=0.01$, $t=4.51$, $p<0.001$); sentences spoken with loud vs normal intensity: ($\beta=0.15$, $SE=0.03$, $t=4.48$, $p<0.001$)]. The full output of the statistical models for the SPL intensity, airflow volume and airflow velocity are included in Appendices I-II.

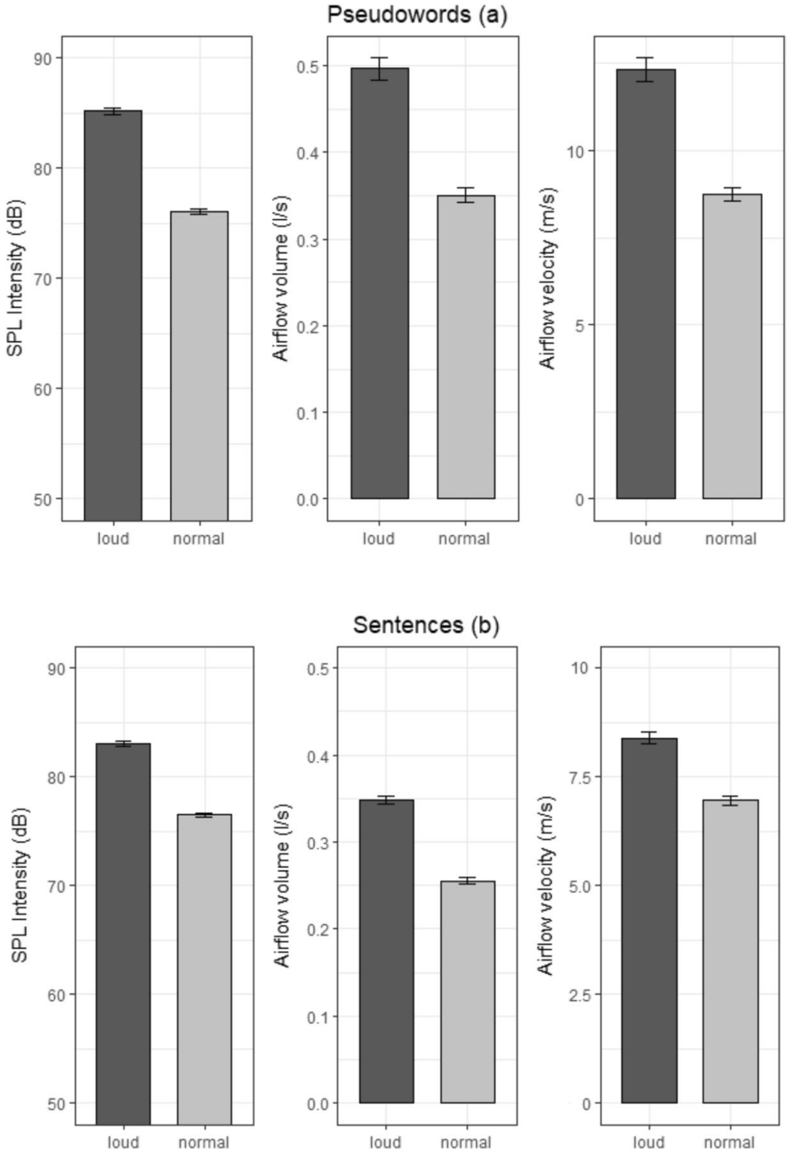

Fig. 4 Means and standard errors for the SPL intensity (dB), airflow volume (l/s) and airflow velocity (m/s) split by voice intensity (loud/normal) for pseudowords plot (**a**) and sentences plot (**b**)

3.2 Experiment 2

Figure 5 illustrates the results for the SPL intensity (dB) and the airflow velocity (m/s) recorded by the three anemometers (anemometer 1, anemometer 2, anemometer 3) according to voice intensity (loud, normal) for pseudowords (a) and sentences

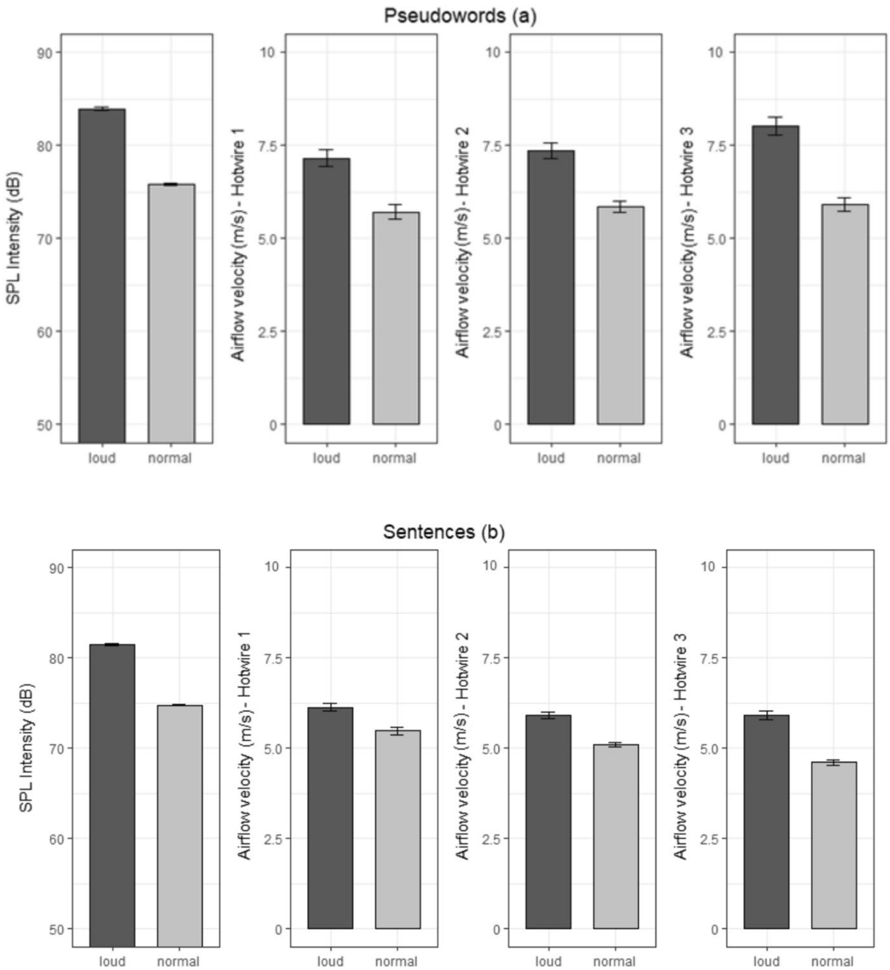

Fig. 5 Means and standard deviations for the SPL intensity (dB), airflow velocity (m/s) recorded by anemometer 1, anemometer 2 and anemometer 3 split by voice intensity (loud/normal) for pseudowords plot (**a**) and sentences plot (**b**)

(b). Pseudowords and sentences presented a higher SPL intensity when spoken with loud intensity than when spoken with normal intensity [pseudowords spoken with loud intensity vs pseudowords spoken with normal intensity: (β=8.04, SE=0.57, t=14.03, p=<0.001); sentences spoken with loud intensity vs sentences spoken with normal intensity: (β=6.76, SE=0.61, t=11.03, p<0.001)], as well as a higher airflow velocity recorded by anemometer 1 [pseudowords spoken with loud intensity vs pseudowords spoken with normal intensity: (β=0.19, SE=0.05, t=4.14, p<0.001); sentences spoken with loud intensity vs sentences spoken with normal intensity: (β=0.09, SE=0.05, t=2.14, p=0.0457)], by anemometer 2 [(pseudowords spoken with loud intensity vs pseudowords spoken with normal intensity: (β=0.19, SE=0.04, t=4.77, p<0.001);

sentences spoken with loud intensity vs sentences spoken with normal intensity: (β=0.09, SE=0.03, t=2.91, p=0.0086)] and by anemometer 3 [pseudowords spoken with loud intensity vs pseudowords spoken with normal intensity: (β=0.01, SE=0.01, t=4.56, p<0.001); sentences spoken with loud intensity vs sentences spoken with normal intensity (β=0.01, SE=0.01, t=5.02, p<0.001)]. The output of the statistical models for the SPL intensity, airflow volume and airflow velocities recorded by the three anemometers are included in Appendices III-IV.

3.3 Experiment 3

Figure 6 reports the results for the (normalised) number of droplets distributed by size for pseudowords spoken with loud and normal intensity. Pseudowords showed an increased number of droplets when spoken with loud intensity than when spoken with normal intensity (β=95.2, SE=24.19, t=3.93, p<0.001). In addition, we analysed the interaction between the number of droplets, droplets size and voice intensity level. Pseudowords spoken with loud intensity presented an increased number of droplets of different sizes (β=3.64, SE=0.77, t=4.74, p<0.001). However, the size distribution of droplets was similar in both loud and normal intensity conditions, with most of the droplets presenting a geometric diameter of nearly 1 micron (see Fig. 6). The complete output of the statistical model is reported in Appendix V.

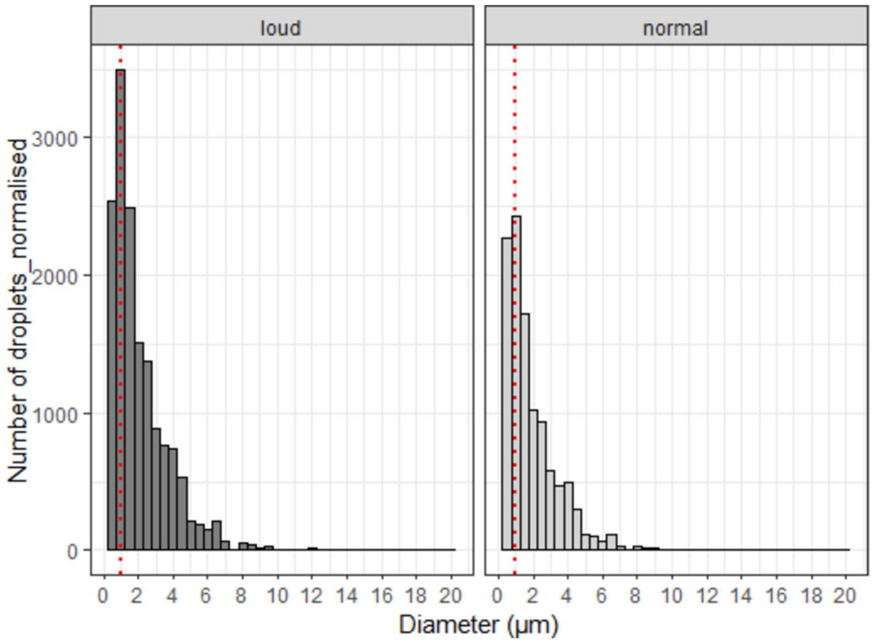

Fig. 6 Histogram of the number of droplets (y-axis) as function of droplet size (x-axis) split by voice intensity (loud/normal). The dashed lines (in red) correspond to the mean values

4 Conclusions and future work

In this paper, we described the methodology employed in three experiments aimed at investigating airflow and droplet emission during interactive speech. We used as a methodological validation standard for our study physiological measurements (airflow volume, airflow velocity, number and size of droplets) for pseudowords and sentences produced with loud vs normal intensity. We found increased physiological measurements when pseudowords and sentences were produced with loud intensity compared to normal intensity. These findings were consistent with those of previous studies (Asadi et al., 2019; Giovanni et al., 2020), hence confirming the robustness of the methods employed.

From a methodological point of view, we went beyond previous research by employing multiple measures of airflow and droplet emission, thus allowing for future development of a more complete understanding of speech droplet dissemination. Furthermore, we introduced the hot-wire anemometry technique into research on speech droplet emission. This technique has been successfully employed in the field of fluid dynamics (Bruun, 1995), and its accurate velocity measurements show promise in helping us to understand airflow dynamics involved in speech production. The technique could thus be used in substitution to previously employed techniques (i.e. e-cigarettes; Giovanni et al., 2020; Chao et al., 2009) that provide less accurate velocity measurements. From a linguistic point of view, we applied an innovative experimental design considering droplet emission both at the word and sentence level in an interactive setting more representative of conversations in everyday life.

Thanks to its innovative methodology and design, our study will provide reference data on droplet emission according to voice intensity, phoneme characteristics and prosodic conditions. We expect significant differences in the parameters considered (airflow volume, velocities at different points in space, droplet size and number) depending on the type of phonemes. For instance, the production of plosives such as /p/ should involve a burst-like release of air (Akbarian et al., 2020; Asadi et al., 2020) that might lead to increased airflow volume, increased airflow velocities at various spatial points and, a greater number and size of droplets. In contrast, fricatives, characterized by continuous airflow through a narrow constriction (Proctor et al., 2010), may produce a more uniform distribution of airflow, leading to less spatial variation in velocity and a smaller number and size of droplets compared to plosives. Furthermore, we aim to investigate the effects of prosodic focus and its interaction with voice intensity and phoneme characteristics on droplet emission. Phonetically, prosodic focus in French is signalled by multiple parameters, such as fundamental frequency (f0, the main acoustic cue of intonation), intensity, segmental duration, voice quality, and the spatial/temporal extent of articulatory movements (De Jong et al., 1993; Loevenbruck, 1999; Dohen & Lœvenbruck, 2004). As a consequence, prosodic focus can lead to an increase in physiological effort, potentially increasing droplet emission. This effect could be further amplified when speaking in a loud voice, or when producing phonemes that require significant physiological effort, such as high vowels and plosives.

In addition to the above-mentioned features, future research should also consider the impact of the intrinsic durational properties of vowels and consonants. Indeed, the duration of phonemes could affect how airflow is sustained over time, potentially shaping droplet production dynamics. For instance, in closed syllables in French (where a consonant terminates the syllable), the low vowel /a/ has been found to be over 70% longer than high vowels (/i/, /u/; O'Shaughnessy, 1981). Similarly, in monosyllabic words, fricatives tend to have longer durations than plosives (O'Shaughnessy, 1981). Longer durations of vowels (e.g. /a/) or consonants (e.g. fricatives) may facilitate more continuous airflow, potentially leading to a sustained emission of droplets. In contrast, shorter high vowels (e.g. /i/) or more abrupt consonants (e.g. plosives) may generate sharp bursts of airflow, resulting in different droplet patterns, such as increase in droplet size or concentration. Understanding these durational effects could provide crucial insights into the aerodynamic and acoustic mechanisms underlying droplet production, with significant implications for improving speech simulation models.

From an epidemiologic point of view, our data may be used by simulation studies to accurately predict the transport, dispersion and evaporation of droplets emitted under specific speech conditions. This knowledge is essential to adapt protective devices (physical distance, masks) to limit the potential spread of respiratory viruses during speech interaction. Our study will also bring new insights into fluid mechanics in speech, thus permitting the inference of speech production mechanisms in a non-invasive manner. Earlier research on speech physiology mostly focussed on the anatomic and articulatory mechanisms of speech production inside the vocal tract, while those developed outside the vocal tract remain poorly investigated. Understanding them is crucial as they can provide insights into vocal efficiency and functioning and can serve as a clinical measure during diagnosis and assessment.

Appendix

See Tables 3, 4, 5, 6, and 7.

Table 3 Appendix I (experiment 1 - pseudowords). Outputs of the linear mixed models estimated for the following DVs: SPL intensity (dB); airflow volume (l/s) and airflow velocity (s/m)

	Estimate	*SE*	*z*	*p*
SPL Intensity				
(Intercept)	75.8722	1.4283	53.120	< 2e-16***
Intensity_loud	9.0855	0.9540	9.524	1.75e-08***
Repetition2	0.1837	0.1482	1.239	0.215555
Repetition3	0.5222	0.1497	3.488	0.000504***
Airflow volume				
(Intercept)	2.832e-01	3.184e-02	8.894	2.04e-09***
Intensity_loud	9.856e-02	2.425e-02	4.064	0.000625***
Repetition2	4.029e-03	6.504e-03	0.619	0.535702
Repetition3	9.942e-03	6.568e-03	1.514	0.130348
Airflow velocity				
(Intercept)	2.131e+00	1.016e-01	20.977	5.16e-15***
Intensity_loud	2.778e-01	6.161e-02	4.510	0.000164***
Repetition2	4.122e-03	2.620e-02	0.157	0.875027
Repetition3	2.606e-02	2.646e-02	0.985	0.324877

For each model, the fixed effects were the INTENSITY (loud, normal) and REPETITION (3). The reference level (intercept) corresponds to values for pseudowords spoken with normal intensity. The statistical formula was DV ~ INTENSITY + REPETITION + (1 + INTENSITY|SPEAKER) + (1 + INTENSITY|PSEUDOWORD)

Table 4 Appendix II (experiment 1 - sentences). Outputs of the linear mixed models estimated for the following DVs: SPL intensity (dB); airflow volume (l/s), and airflow velocity (s/m) recorded by anemometer 1, anemometer 2 and anemometer 3

	Estimate	*SE*	*z*	*p*
SPL Intensity				
(Intercept)	77.26061	1.50052	51.489	< 2e-16***
Intensity_loud	6.44004	0.59555	10.814	1.07e-09***
Repetition2	0.10594	0.09577	1.106	0.269
Repetition3	−0.04419	0.09577	−0.461	0.645
Airflow volume				
(Intercept)	2.287e-01	2.083e-02	10.979	6.12e-10***
Intensity_loud	6.944e-02	1.593e-02	4.359	0.000336***
Repetition2	7.645e-03	2.616e-03	2.922	0.003509**
Repetition3	1.154e-02	2.616e-03	4.411	1.07e-05***
Airflow velocity				
(Intercept)	2.00751	0.05444	36.877	< 2e-16***
Intensity_loud	0.15159	0.03384	4.480	0.000257***
Repetition2	−0.01632	0.02053	−0.795	0.426635
Repetition3	−0.01913	0.02053	−0.932	0.351670

For each model, the fixed factors were INTENSITY (loud, normal) and REPETITION (3). The reference level (intercept) corresponds to values for sentences spoken with normal intensity. The statistical formula was DV ~ INTENSITY + REPETITION + (1 + INTENSITY|SPEAKER) + (1 + INTENSITY|SENTENCE)

Table 5 Appendix III (experiment 2 – pseudowords). Outputs of the linear mixed models estimated for the following DVs: SPL intensity (dB) and airflow velocity (m/s) recorded by anemometer 1, anemometer 2 and anemometer 3

	Estimate	*SE*	*z*	*p*
SPL Intensity				
(Intercept)	75.6444	0.7087	106.733	< 2e-16***
Intensity_loud	8.0406	0.5730	14.034	1.13e-11***
Repetition2	0.3115	0.1222	2.549	0.0109*
Repetition3	0.3034	0.1211	2.506	0.0123*
Airflow velocity – anemometer 1				
(Intercept)	1.63050	0.18093	9.012	1.98e-07***
Intensity_loud	0.19295	0.04662	4.139	0.000498***
Repetition2	−0.04201	0.03400	−1.236	0.216841
Repetition3	−0.06475	0.03372	−1.920	0.055002
Airflow velocity – anemometer 2				
(Intercept)	1.72145	0.15892	10.832	2.33e-08***
Intensity_loud	0.18776	0.03936	4.771	0.000128***
Repetition2	−0.04080	0.03139	−1.299	0.193996
Repetition3	-0.04108	0.03113	−1.320	0.187201
Airflow velocity – anemometer 3				
(Intercept)	1.704e + 00	1.195e-01	14.259	4.91e-14***
Intensity_loud	2.551e-01	5.596e-02	4.560	0.000192***
Repetition2	7.882e-04	3.151e-02	0.025	0.980046
Repetition3	2.604e-02	3.124e-02	0.834	0.404687

For each model, the fixed effects were the INTENSITY (loud, normal) and REPETITION (3). The reference level (intercept) corresponds to values for pseudowords spoken with normal intensity. The statistical formula was DV ~ INTENSITY + REPETITION + (1 + INTENSITY|SPEAKER) + (1 + INTENSITY|PSEUDOWORD)

Table 6 Appendix IV (experiment 2 – sentences). Outputs of the linear mixed models estimated for the following DVs: SPL intensity (dB); airflow velocity from anemometer 1, anemometer 2 and anemometer 3

	Estimate	*SE*	*z*	*p*
SPL Intensity				
(Intercept)	75.03652	0.70523	106.399	< 2e-6***
Intensity_loud	6.75815	0.61265	11.031	4.32e-10***
Repetition2	−0.19782	0.08562	−2.310	0.0209*
Repetition3	−0.45161	0.08562	−5.275	1.45e-07***
Airflow velocity – anemometer 1				
(Intercept)	1.74022	0.09437	18.441	< 2e-16***
Intensity_loud	0.09875	0.04622	2.137	0.0457*
Repetition2	−0.01673	0.01984	−0.843	0.3994
Repetition3	−0.03409	0.01985	−1.718	0.0859
Airflow velocity – anemometer 2				
(Intercept)	1.74332	0.05522	31.570	< 2e-16***
Intensity_loud	0.09915	0.03408	2.909	0.00863**
Repetition2	−0.02123	0.01582	−1.342	0.17977
Repetition3	0.01184	0.01582	0.749	0.45418
Airflow velocity – anemometer 3				
(Intercept)	1.595e + 00	7.626e-02	20.910	< 2e-16***
Intensity_loud	2.032e-01	4.044e-02	5.024	7.32e-05***
Repetition2	2.400e-03	1.805e-02	0.133	0.894
Repetition3	2.478e-02	1.805e-02	1.373	0.170

For each model, the fixed effects were the INTENSITY (loud, normal) and REPETITION (3). The reference level (intercept) corresponds to values for sentences spoken with normal intensity. The statistical formula was DV ~ INTENSITY + REPETITION + (1 + INTENSITY|SPEAKER) + (1 + INTENSITY|SENTENCE)

Table 7 Appendix V. output of the linear mixed model estimated for number of droplets with the fixed effects INTENSITY (loud, normal), SIZE (0.5–20 µs) and their interaction as well as REPETITION (3)

	Estimate	SE	z	*p*
(Intercept)	116.1886	42.2584	2.749	0.012132*
Intensity_loud	95.2306	24.1918	3.936	0.000476***
Size	10.2047	0.5431	18.790	< 2e-16***
Size: intensity_loud	3.6435	0.7680	4.744	< 2e-16***
Repetition2	5.2531	4.9453	1.062	0.288126
Repetition3	12.2032	4.9453	2.468	0.013602*

The reference level (intercept) corresponds to values for number of droplets to pseudowords spoken with normal intensity. The statistical formula is number of droplets_normalised ~ INTENSITY * SIZE + REPETITION + (1 + INTENSITY | SPEAKER) + (1 + INTENSITY| PSEUDOWORDS)

Acknowledgements The authors would like to thank all the participants in this study. The authors would also like to thank the CEP staff (www.lpl-aix.fr/~cep) for assisting in the experiments, and Federica Procino for her help in data collection.

Author contributions All authors contributed to the conceptualisation and realisation of the entire study. The first draft of the article was written by F.C. (the first author). Critical review, comments and revisions of the first draft were made by all authors.

Funding This study was supported by an ANR-20-COV7-0006 grant from the French National Research Agency in 2020.

Data availability The data supporting this study are not publicly available at this time as they are still being used for ongoing analyses and further publications.

Declarations

Conflict of interest The authors declare no competing interests.

References

Abkarian, M., & Stone, H. A. (2020). Stretching and break-up of saliva filaments during speech: A route for pathogen aerosolization and its potential mitigation. *Physical Review Fluids, 5*(10), 102301. https://doi.org/10.1103/PhysRevFluids.5.102301

Abkarian, M., Mendez, S., Xue, N., Yang, F., & Stone, H. A. (2020). Speech can produce jet-like transport relevant to asymptomatic spreading of virus. *Proceedings of the National Academy of Sciences, 117*(41), 25237–25245. https://doi.org/10.1073/pnas.201215611

André, C., Ghio, A., Cavé, C., & Teston, B. (2003). PERCEVAL: a Computer-Driven System for Experimentation on Auditory and Visual Perception. *Proceedings of the International Congress of Phonetic Sciences* (ICPhS), 1421–1424, Barcelona, Spain.https://doi.org/10.48550/arXiv.0705.4415

Anfinrud, P., Stadnytskyi, V., Bax, C. E., & Bax, A. (2020). Visualizing speech-generated oral fluid droplets with laser light scattering. *New England Journal of Medicine, 382*(21), 2061–2063. https://doi.org/10.1056/NEJMc2007800

Asadi, S., Wexler, A. S., Cappa, C. D., Barreda, S., Bouvier, N. M., & Ristenpart, W. D. (2019). Aerosol emission and superemission during human speech increase with voice loudness. *Scientific Reports, 9*(1), 1–10. https://doi.org/10.1038/s41598-019-38808-z

Asadi, S., Wexler, A. S., Cappa, C. D., Barreda, S., Bouvier, N. M., & Ristenpart, W. D. (2020). Effect of voicing and articulation manner on aerosol particle emission during human speech. *PloS ONE, 15*(1), e0227699. https://doi.org/10.1371/journal.pone.0227699

Barratt, D., Nguepnang, M. A., & Atkins, M. D. (2016). Fast-response velocity and shear stress measurements. In K. Tongbeum & J. S. Seung Jin (Eds.), *Application of Thermo-Fluidic Measurement Techniques.* Oxford: Butterworth Heinemann. https://doi.org/10.1016/B978-0-12-809731-1.00004-6

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bertholon, J. F., Becquemin, M. H., Roy, M., Roy, F., Ledur, D., Maesano, I. A., & Dautzenberg, B. (2013). Comparaison de l'aérosol de la cigarette électronique à celui des cigarettes ordinaires et de la chicha. *Revue Des Maladies Respiratoires, 30*(9), 752–757. https://doi.org/10.1016/j.rmr.2013.03.003

Bigi, B. (2015). SPPAS - multi-lingual approaches to the automatic annotation of speech. *Phonetician - International Society of Phonetic Sciences, 111-112 / 2015-I–II*, 54–69. ISSN 0741-6164.

Bourouiba, L. (2021). The fluid dynamics of disease transmission. *Annual Review of Fluid Mechanics, 53*, 473–508. https://doi.org/10.1146/annurev-fluid-060220-113712

Bourouiba, L., Dehandschoewercker, E., & Bush, J. W. (2014). Violent expiratory events: On coughing and sneezing. *Journal of Fluid Mechanics, 745*, 537–563. https://doi.org/10.1017/jfm.2014.88

Bruun, H. H. (1995). *Hotwire anemometry, principles and signal analysis*. Oxford University Press.

Bruun, H. H., Khan, M. A., Al-Kayiem, H. H., & Fardad, A. A. (1988). Velocity calibration relationships for hot-wire anemometry. *Journal of Physics E: Scientific Instruments, 21*(2), 225. https://doi.org/10.1088/0022-3735/21/2/020

Chao, C. Y. H., Wan, M. P., Morawska, L., Johnson, G. R., Ristovski, Z. D., Hargreaves, M., Mengersen, K., Corbett, S., Li, Y., Xie, X., & Katoshevski, D. (2009). Characterization of expiration air jets and droplet size distributions immediately at the mouth opening. *Journal of Aerosol Science, 40*(2), 122–133. https://doi.org/10.1016/j.jaerosci.2008.10.003

Comte-Bellot, G. (1976). Hot-wire anemometry. *Annual Review of Fluid Mechanics, 8*(1), 209–231. https://doi.org/10.1146/annurev.fl.08.010176.001233

De Jong, K., Beckman, M. E., & Edwards, J. (1993). The interplay between prosodic structure and coarticulation. *Language and Speech, 36*(2–3), 197–212. https://doi.org/10.1177/002383099303600305

Dohen, M., & Lœvenbruck, H. (2004). Pre-focal rephrasing, focal enhancement and post-focal deaccentuation in French, *Proceedings of the 8th International Conference on Spoken Language Processing*, (pp. 1313–1316), Jeju, Korea.

Eiche, T., & Kuster, M. (2020). Aerosol release by healthy people during speaking: Possible contribution to the transmission of SARS-CoV-2. *International Journal of Environmental Research and Public Health, 17*(23), 9088. https://doi.org/10.3390/ijerph17239088

Ghio, A., & Teston, B. (2004). Evaluation of We used Sesane _ software the acoustic and aerodynamic constraints of a pneumotachograph for speech and voice studies. *Proceedings of the International Conference on Voice Physiology and Biomechanics*, 55–58, Université de la Méditerranée, Marseille, France.

Ghio, A., Pouchoulin, G., Teston, B., Pinto, S., Fredouille, C., De Looze, C., Robert, D., Viallet, F., & Giovanni, A. (2012). How to manage sound, physiological and clinical data of 2500 dysphonic and dysarthric speakers? *Speech Communication, 54*(5), 664–679. https://doi.org/10.1016/j.specom.2011.04.002

Giovanni, A., Radulesco, T., Bouchet, G., Mattei, A., Révis, J., Bogdanski, E., & Michel, J. (2020). Transmission of droplet-conveyed infectious agents such as SARS-CoV-2 by speech and vocal exercises during speech therapy: Preliminary experiment concerning airflow velocity. *European Archives of Oto-Rhino-Laryngology, 278*, 1687–1692. https://doi.org/10.1007/s00405-020-06200-7

Gupta, J. K., Lin, C. H., & Chen, Q. (2009). Flow dynamics and characterization of a cough. *Indoor Air, 19*(6), 517–525. https://doi.org/10.1111/j.1600-0668.2009.00619.x

Gupta, J. K., Lin, C. H., & Chen, Q. (2010). Characterizing exhaled airflow from breathing and talking. *Indoor Air, 20*(1), 31–39. https://doi.org/10.1111/j.1600-0668.2009.00623.x

Gussenhoven, C. (2016). Foundations of intonational meaning: Anatomical and physiological factors. *Topics in Cognitive Science, 8*(2), 425–434. https://doi.org/10.1111/tops.12197

Han, M., Ooka, R., Kikumoto, H., Oh, W., Bu, Y., & Hu, S. (2021). Experimental measurements of airflow features and velocity distribution exhaled from sneeze and speech using particle image velocimetry. *Building and Environment, 205*, 108293. https://doi.org/10.1016/j.buildenv.2021.108293

Johnson, G. R., Morawska, L., Ristovski, Z. D., Hargreaves, M., Mengersen, K., Chao, C. Y. H., Wan, M. P., Li, Y., Xie, X., Katoshevski, D., & Corbett, S. (2011). Modality of human expired aerosol size distributions. *Journal of Aerosol Science, 42*, 839–851. https://doi.org/10.1016/j.jaerosci.2011.07.009

Jørgensen, F., E. ——How to measure turbulence with hot-wire anemometers, Publication no.: 9040U6151. Date 2002-02-01. Dantec Dynamics A/S, & Box, P. O. (2002). 121, Tonsbakken 16–18, DK-2740 Skovlunde, Denmark.

King, L. V. (1914). On the convection of heat from small cylinders in a stream of fluid: Determination of the convection constants of small platinum wires with applications to hot-wire anemometry.

Philosophical Transactions of the Royal Society of London Series A Containing Papers of a Mathematical or Physical Character, 214(509–522), 373–432. https://doi.org/10.1098/rsta.1914.0023

Kwon, S. B., Park, J., Jang, J., Cho, Y., Park, D. S., Kim, C., Bae, G. N., & Jang, A. (2012). Study on the initial velocity distribution of exhaled air from coughing and speaking. *Chemosphere, 87*, 1260–1264. https://doi.org/10.1016/j.chemosphere.2012.01.032

Loevenbruck, H. (1999). Articulatory effects of contrastive emphasis on the Accentual phrase in French. *The Journal of the Acoustical Society of America, 106*(4_Supplement), 2153–2153. https://doi.org/10.1121/1.427370

Loudon, R. G., & Roberts, R. M. (1967). Droplet expulsion from the respiratory tract. *American Review of Respiratory Disease, 95*(3), 435–442.

Merghani, K. M. M., Sagot, B., Gehin, E., Da, G., & Motzkus, C. (2021). A review on the applied techniques of exhaled airflow and droplets characterization. *Indoor Air, 31*(1), 7–25. https://doi.org/10.1111/ina.12770

Morawska, L. J. G. R., Johnson, G. R., Ristovski, Z. D., Hargreaves, M., Mengersen, K., Corbett, S., Chao, C. Y. H., Li, Y., & Katoshevski, D. (2009). Size distribution and sites of origin of droplets expelled from the human respiratory tract during expiratory activities. *Journal of Aerosol Science, 40*(3), 256–269. https://doi.org/10.1016/j.jaerosci.2008.11.002

Nobach, H., & Bodenschatz, E. (2009). Limitations of accuracy in PIV due to individual variations of particle image intensities. *Experiments in Fluids, 47*, 27–38. https://doi.org/10.1007/s00348-009-0627-4

O'Shaughnessy, D. (1981). A study of French vowel and consonant durations. *Journal of Phonetics, 9*(4), 385–406. https://doi.org/10.1016/S0095-4470(19)31016-2

Papineni, R. S., & Rosenthal, F. S. (1997). The size distribution of droplets in the exhaled breath of healthy human subjects. *Journal of Aerosol Medicine, 10*(2), 105–116. https://doi.org/10.1089/jam.1997.10.105

Patel, R. K., Shackelford, I. A., Priddy, M. C., & Kopechek, J. A. (2020). Effect of speech volume on respiratory emission of oral bacteria as a potential indicator of pathogen transmissibility risk. *The Journal of the Acoustical Society of America, 148*(4), 2322–2326. https://doi.org/10.1121/10.0002278

Proctor, M. I., Shadle, C. H., & Iskarous, K. (2010). Pharyngeal articulation in the production of voiced and voiceless fricatives. *The Journal of the Acoustical Society of America, 127*(3), 1507–1518.

Roy, N., Merrill, R. M., Gray, S. D., & Smith, E. M. (2004a). Voice disorders in the general population: Prevalence, risk factors, and occupational impact. *The Laryngoscope, 115*(11), 1988–1995. https://doi.org/10.1097/01.mlg.0000179174.32345.41

Roy, N., Merrill, R. M., Thibeault, S., Parsa, R. A., Gray, S. D., & Smith, E. M. (2004b). Prevalence of voice disorders in teachers and the general population. *Journal of Speech and Hearing Research, 47*(2), 281–293. https://doi.org/10.1044/1092-4388(2004/023)

Stadnytskyi, V., Bax, C. E., Bax, A., & Anfinrud, P. (2020). The airborne lifetime of small speech droplets and their potential importance in SARS-CoV-2 transmission. *Proceedings of the National Academy of Sciences, 117*(22), 11875–11877. https://doi.org/10.1073/pnas.2006874117

Stadnytskyi, V., Anfinrud, P., & Bax, A. (2021). Breathing, speaking, coughing or sneezing: What drives transmission of SARS-CoV-2? *Journal of Internal Medicine, 290*, 1010–1027. https://doi.org/10.1111/joim.13326

Tang, J. W., Nicolle, A. D., Klettner, C. A., Pantelic, J., Wang, L., Suhaimi, A. B., Tan, A. Y. L., Ong, G. W. X., Su, R., Sekhar, C., Cheong, D. D. W., & Tham, K. W. (2013). Airflow dynamics of human jets: Sneezing and breathing-potential sources of infectious aerosols. *PloS One, 8*(4), e59970. https://doi.org/10.1371/journal.pone.0059970

Wei, J., & Li, Y. (2016). Airborne spread of infectious agents in the indoor environment. *American Journal of Infection Control, 44*(9), S102–S108. https://doi.org/10.1016/j.ajic.2016.06.003

Wells, W. F. (1934). On air-borne infection. Study II. Droplets and droplet nuclei. *American Journal of Hygiene, 20*, 611–618.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org

Authors and Affiliations

Francesca Carbone[1,2] · Gilles Bouchet[3] · Alain Ghio[2] · Thierry Legou[2] · Carine André[2] · Muriel Lalain[2] · Caterina Petrone[2] · Antoine Giovanni[2,4]

✉ Francesca Carbone
f.carbone@kent.ac.uk

1 School of Psychology, University of Kent, Canterbury, UK

2 Aix Marseille Univ, CNRS, LPL, Aix-en-Provence, France

3 Aix Marseille Univ, CNRS, IUSTI, Marseille, France

4 ENT-HNS Department, Aix Marseille Univ, La Conception University Hospital, Marseille, France