AERODYNAMIC TECHNIQUES FOR PHONETIC FIELDWORK

Didier Demolin

GIPSA-Lab, Université Stendhal Grenoble, France
`Didier.Demolin@gipsa-lab.grenoble-inp.fr`

ABSTRACT

This paper discusses the contribution of aerodynamic techniques to record phonetic data in the field. New tools allow using automatic calibrations procedures as well as quantifying data in a reliable way. This permits addressing research questions such as patterns of variations of phenomena. Analytic methods show how to interpret data from physical principles and integrate those in adequate models.

Keywords: aerodynamics, field methods, analytic methods, description, quantification

1. INTRODUCTION

Phoneticians doing fieldwork are confronted, sooner or later, with the following basic question: how to describe unknown (and sometimes known) speech sounds in a reliable way? This means to understand how they are produced and what are the best primitives, gestures and/or features necessary to describe these sounds. Any phonetician doing fieldwork is now confronted with a second question connected to the first: how to quantify data? The rapid development and miniaturization of research tools and the development of computational techniques makes this enterprise much easier than it was before. Quantifying data allows fieldworkers to address another fundamental issue: what are the patterns of variation of the phenomena which are studied? One important tool to answer those three questions comes from the use of aerodynamic measurements. If the precision of the tools used is accurate enough, this technique allows for the induction of articulatory movements in the vocal tract with great precision. Acoustic cues and features may also be deduced on changes of flow and pressure in the vocal tract. Despite the fact that measuring intraoral pressure can be invasive at times, this is one of the most reliable and powerful techniques phoneticians have to understand the behavior and the production of speech sounds. While the kymograph contributed to the description of sounds from languages from sometimes very remote areas, it could hardly be considered a portable tool, and there has not been any reliable tool systematically used prior to Ladefoged’s work on West African languages in 1968 [3]. Bringing speakers to a laboratory is always possible, but if one wishes to record data from speakers in remote areas, reliable and portable equipment is a necessity. Such tools are now available and can change the way that phoneticians do fieldwork. This leads to new challenges for fieldworkers such as being able to interpret aerodynamic data right in the field and being able to quickly pose new questions based on the interpretation of field data. This also requires reference values (such as maxima, minima and mean for pressure and flow) involved in close or similar phenomena to those studied, along with a basic understanding of the physical principles involved in the aerodynamics of speech. Finally the integration of aerodynamic parameters in models of speech production leads to a better understanding of speech phenomena.

2. CALIBRATION, MEASUREMENT AND QUANTIFICATION OF DATA

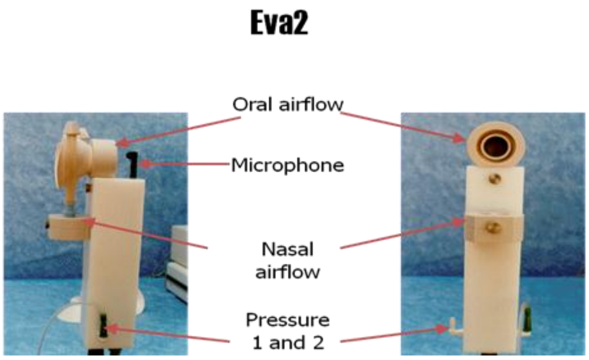
2.1. The portable EVA2 workstation

Several types of equipment to record aerodynamic data in the field are presently available. Whatever the equipment used, there are two important things that must be taken into consideration: calibration and quantification of data. The equipment described here is the EVA2 portable workstation (www.sqlab.fr) which allows for synchronized measurements of aerodynamic and acoustic data with two additional channels. Two airflow channels and two pressure channels are available to measure oral and nasal airflow along with pharyngeal pressure (and if necessary buccal) pressure. The portable EVA2 has a 3 hours battery life which allows researchers to record data in places without electricity. The following section adapted from Ghio and Teston [2] describes the main characteristics of the EVA (Computerised Vocal Assessment, SQLab) system.

2.2. Technical characteristics

The pneumotacograph (PTG) used in the EVA2 system is based on the principle described by Teston [4, 5]. This device is a PTG with grid (stainless steel wire of 200 µm diameter and a step of mesh of 250µm), reduced in size (to 30 mms diameter and 20 mm length) to optimize its response time and linearity in all articulatory contexts. Thanks to its very sensitive and stable differential pressure transducers (Data Instrument DCXL), it is able to apprehend a flow on the order of 1 cm3/s. The resistance of the grid is 10 Pa by dm3/s (litre/s), i.e. approximately 1% of the intra oral pressure of a normal subject, which does not disturb the normal operation of the vocal tract. Resistance was selected for a level of saturation of the sensor to the value of 10 dm3/s in forced breathing. These values represent a dynamics of 60 dB. To reduce the non-linearity of measurement caused by aerodynamic turbulences produced during speech production, the pressure tap is made in 8 points of the circumference of the measurement pipe and a grid of tranquillization (of negligible resistance) is laid out in front of the pressure taps. The sensor is made of a synthetic material (Polyacetal) which has a very good mechanical resistance to sterilization and UV.

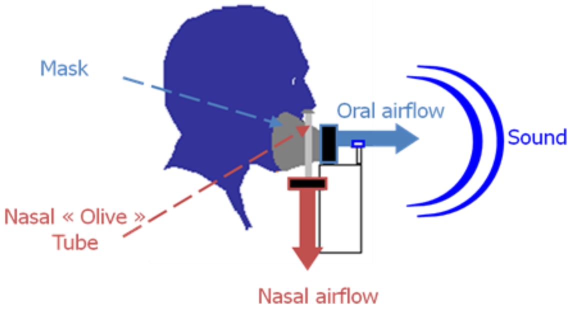
Figure 1: The setting of the EVA2 transducers.



2.3. Adaptation to the anatomy of the speaker

This type of flow meter equips the system EVA2 ™ which is a device designed for the recording and the study of many parameters of the speech and voice production: sound, pitch, intensity, airflows, pressures [6]. The simultaneous recording of the sound and the airflows to the mouth and the nostrils requires the use of a "mouthpiece" which is a mechanical stand of the microphone, flow sensors and flexible silicone mask to form the seal necessary to obtain reliable measurements of oral air flow (Figure 1). The nasal sensor is located under the oral sensor and mounted vertically to capture the natural flow of the air to the nostrils. The flexible silicone masks, supported by the chin allow for normal unrestricted movements of the mandible while providing an excellent seal. They are also easily sterilizable.

Figure 2: Direction of flow and sound within the EVA2 setting.



2.4. Calibration procedures

2.4.1. Pressure measurements

Pressure measurements represent an important domain in medical and scientific instrumentation. In the past, pressure calibration was performed using a U tube with water. Nowadays, the calibration of pressure sensors is usually made by the manufacturer on a unique sacle or ina range of several sacles of fixed dynamics (i.e. without gain variations during use). In the most current applications, pressure sensors are calibrated only once but can be checked on a regular basis for specific and sensitive applications. The EVA2 pressure sensors have two scales, 40 and 200 hPa, calibrated by a precision electronic manometer. The manometer is a THOMMEN type HM28, scale 0-300 hPa, 0,05% FS class within which pressure is generated by a hand precision pump. Some pressure values are important to remind. Reference values are: 1 cm water = 0.9806 hPa, 1 hPa = 1,02 cm water and 1 mb = 1 hPa.

2.4.2. Airflow measurements

The principle of measurement of the flow by a PTG consists of measuring the pressure difference ΔP before and after introducing resistance to the airflow. It varies according to the speed of the fluid, i.e. its volumic flow D. For low values, the flow is laminar. The laws of the fluid mechanics show that, in this case, $\Delta P = R.D$ where R is resistance (Poisseuille equation). For more significant values of flow, it becomes turbulent. The pressure loss ΔP is then proportional to the square of the flow

(Venturi law): $\Delta P = R.\ D^2$. In a general way, the relation between these values is formalized by: $\Delta P = R.\ D\ N$. Precise knowledge of the relation between ΔP and D makes it possible to bring back the measurement of flow to a measurement of differential pressure [2].

2.4.3. Measurements and software

The portable EVA2 is equipped with the *phonedit* software for data recording and processing. It has two features which deserve special mention. The first is the ability to filter the flow and pressure signals, useful for visualizing details of oral and nasal airflow. The second is the possibility of integrating flow to obtain volume. This is particularly useful when comparing subjects or quantifying the volume of air specific to some phenomena.

3. AERODYNAMIC DATA

Languages like Guarani, Kuikuro, Fulfulde and Hendo show that the use of aerodynamic data is particularly useful to describe and understand the way some sounds of these languages are produced.

3.1. Airflow measurements

Guarani is well known for nasal spans, i.e entire words or parts of sentences where all segments (including voiceless fricatives) can be nasalized. The combination of nasal and oral airflow recordings synchronized with acoustic data allows understanding the timing and synchronization of velum, oral and glottal gestures. Figure 3 shows that the opening of the glottis at the start of the nasalized fricative [ʃ], visible by a peak oral airflow, is made while the velum is still open, which is visible by the peak of nasal airflow. This delays frication noise until there is a sufficient reduction of nasal airflow. This example shows that frication noise's amplitude gradually increases with nasal airflow reduction. Data from several speakers shows similar patterns but amplitude varies considerably from one speaker to the other: oral airflow varies from 101 to 346 dm^3/s while nasal airflow has a much smaller variation (287 to 368 dm^3/s). These values of nasal airflow are quite high when compared to other languages. This leads to the hypothesis that Guarani nasalized fricatives are produced because the velum remains open after the initial glottal opening and the oral constriction.

Figure 3: Audio waveform, oral (1) and nasal airflow (2) for the word [kõʃõ] 'mattress'. Single arrows indicate peak oral airflow at the start of the fricative and peak nasal airflow immediately after. Double arrows show airflow during frication noise and at the constriction release.

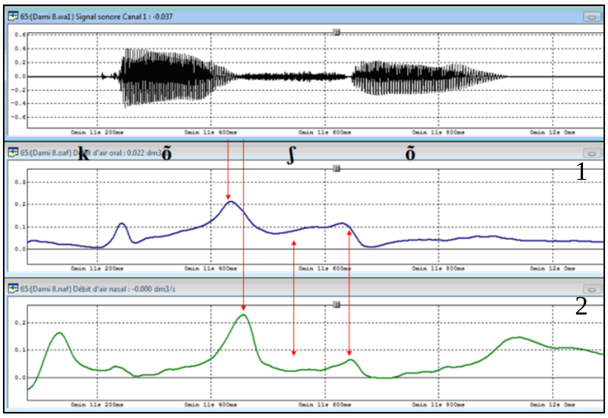
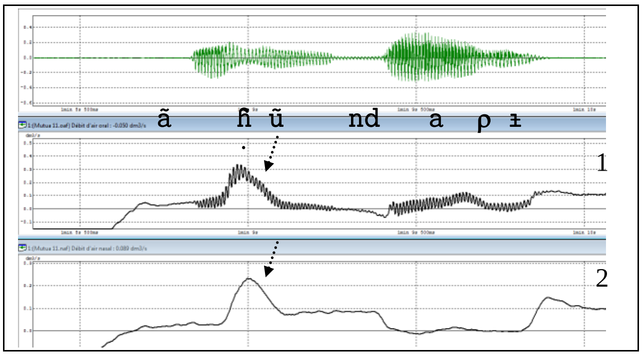


Figure 4 shows a voiced nasalized laryngeal fricative [ɦ̃] in Kuikuro. This sound is produced with quite high oral and nasal airflows. The range is from 112 to 350 dm^3/s for oral airflow and from 287 to 368 dm^3/s for nasal airflow. Examination of synchronized audio and airflows signals reveals that the oral and nasal airflows increase at the same time. This means that the velum opens at the same time that the glottis lets more air going through or that the glottis is vibrating with the interarytenoid space open. As the language also has voiceless (oral and nasal) and voiced nasal laryngeal fricatives [h, ɦ̃, ɦ] this raises the question of the coordination and control of velic and oral gestures. From the data it can be hypothesized that glottal and velum openings are finely coordinated in the case of the nasal segments.

Figure 4: Audio waveform, oral (1) and nasal airflow (2) for the word [aɦ̃ũndapł] 'mattress'. Arrows indicate peak oral airflow at the start of the fricative and peak nasal airflow.



3.2. Airflow and pressure measurements

Figure 5: Audio waveform, intraoral pressure (1) and oral airflow for the word [haɓ8:i0] ‘has/have tied’. The double arrow indicates the end of the larynx lowering and the start of voicing, the arrow on the oral airflow shows ingressive flow after the closure release.

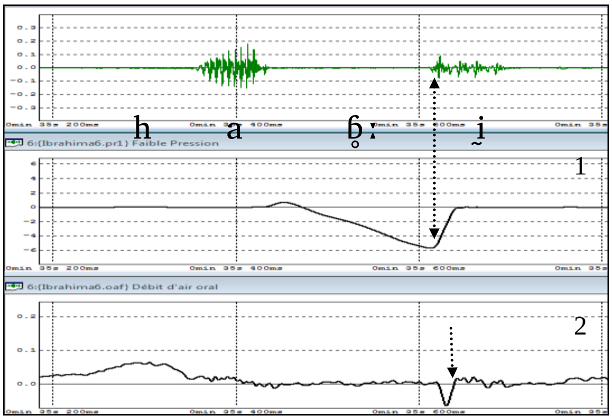
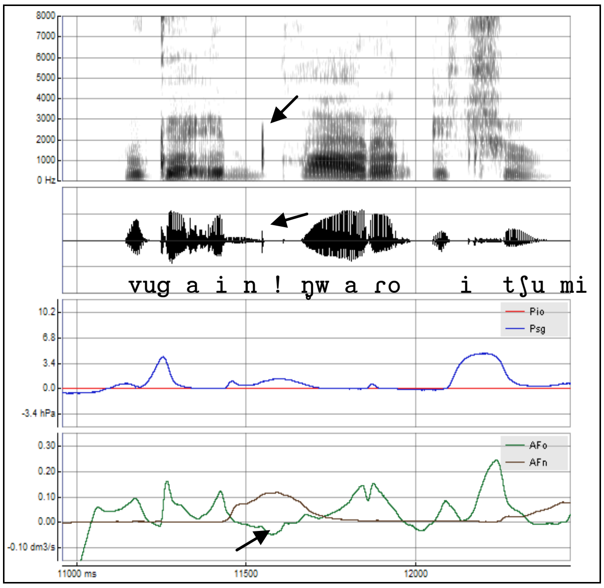


Figure 6: Spectrogram, audio waveform, intraoral pressure (Pio), oral (AFo) and nasal (AFn) airflow of the short sentence [vuga in!ŋgwaɾo itʃumi] ‘say weapon ten times’. Arrows on the spectrogram and audio waveform indicates the click burst after the alveolar nasal [n]. The fall in Afo, indicated by the arrow, reflects the backward movement of the tongue which goes from an alveolar to a velar place of articulation. Pio is not negative because it is measured behind the velum which is also in contact with the tongue dorsum.



Fulfulde has implosive consonants in its phonemic inventory. The combination of oral airflow, intraoral pressure (Pio) and acoustic signal measures show that Fulfulde implosives can be voiceless. This is clearly visible at Figure 5 showing that the larynx is lowered with a closed glottis. Indeed Pio becomes negative with no voicing. The only detectable voicing is at the end just before Pio rises before the closure release. This short prevoicing is characteristic of voiceless implosives [1].

Rwanda show that the combination of oral and nasal airflows with intraoral pressure and the acoustic signal permits to show that click bursts can emerge between two nasal consonants the first being front and the second back. This is visible at Figure 6 where a burst appears between the alveolar nasal [n] and the velar [ŋ].

4. CONCLUSION

The necessity to quantify observed phenomena is part of the scientific endeavor. The examples given above show that in addition to this, there is a necessity for phoneticians to obtain sets of quantitative data in order to understand patterns of variation. Advent of new research tools to measure aerodynamic phenomena in speech contributes to such a goal.

5. REFERENCES

[1] Demolin, D. 1995. The phonetics and phonology of glottalized consonants in Lendu. *Papers in Laboratory Phonology* 4. Cambridge University Press, 368-385.

[2] Ghio, A., Teston, B. 2004. Evaluation of the acoustic and aerodynamic constraints of a pneumotachograph for speech and voice studies. *Proceedings of International Conference on Voice Physiology and Biomechanics*, Marseille: Univ. Méditerranée, 55-58.

[3] Ladefoged, P. 1968. *A Phonetic Study of West African Languages*. Cambridge: Cambridge University Press.

[4] Teston, B. 1987. Etude d’un aérophonomètre de grande dynamique et faible constant de temps. *16è, JEP, GALF*, Hammamet, 132-135.

[5] Teston, B. 1993. A new airflow design for the investigation of speech production. *Eurospeech*. ISCA, Berlin, 3, 405-409.

[6] Teston, B., Galindo, B. 1995. A new diagnosis and rehabilitation for speech and voice pathologies. *Eurospeech*. ESCA, Madrid. 4, 1883-1886.