Speech Aerodynamics Database, Tools and Visualisation

Shi Yu[1], Clara Ponchard[1], Roland Trouville[1], Sergio Hassid[2], Didier Demolin[1]

[1]Laboratoire de Phonétique et Phonologie, CNRS-UMR 7018, Université Sorbonne Nouvelle Paris 3, Paris, France
[2]Hôpital Erasme, Université Libre de Bruxelles, Belgium

{shi.yu, clara.ponchard, didier.demolin}@sorbonne-nouvelle.fr, sergio.hassid@ulb.ac.be, roland.trouville@gmail.com

Introduction

Aerodynamic processes underlie the characteristics of the acoustic signal of speech sounds. Empirical data on aerodynamic parameters of speech production remain very limited, mainly due to the difficulties of its acquisition. Aerodynamic data are important for the understanding of sound patterns in languages. The difficulties consist of internally the nature of aerodynamic processes and, externally the requirements of clinical and medical assistance. This database is one of the very few databases of the aerodynamics of speech production available today to the scientific community. Therefore it offers a unique opportunity for speech scientists to study the aerodynamics of speech production.

Goals:
- Record a substantial amount of simultaneous data to study speech aerodynamics
- Obtain reference values of the aerodynamics of speech production for both female and male speakers in different languages

Methodology

Simultaneous and synchronous recordings of 5 types of signal
- Subglottal Pressure **Ps** (unit: hPa)
- Intraoral Pressure **Po** (unit: hPa)
- Oral Airflow **Qo** (unit: dm3/s)
- Nasal Airflow **Qn** (unit: dm3/s)
- Speech Acoustic Signal + Fundamental frequency (fo) + Intensity

Subglottal pressure (Ps) : the pressure below the focal folds which drives their vibrations.
- Measured directly by tracheal puncture.

Intraoral pressure (Po) : the pressure inside the oral cavity
- Measured with a flexible plastic tube inserted through the nasal cavity to the oropharynx.

Oral airflow (Qo)
- Measured with a flexible silicone rubber mask covering the mouth.

Nasal airflow (Qn)
- Measured through an olive inserted in one or two nostrils.

All measurements are acquired with Physiologia workstation (Teston and Galindo, 1990) and processed with Phonedit software (Ghio, 2002)

Audio data are segmented and annotated with Praat (Boersma, 2001)

Acknowledgements

This work is supported by grants ARC 1996-2000 of the Belgian Government and by ArtSpeech (Pho- netic Articulatory Synthesis, DS0707–2015) from the French National Research Agency and by the French Investissements d'Avenir - Labex EFL program (ANR- 10-LABX-0083).

Database Description

This database contain materials of 3 languages: English, French, and Amharic, cross-language comparison is possible with a subset of this database

English

- Logatome

Logatome consists of the form (/CVCV / or / V V/), they are constructed by combining three vowels [i], [a], [u] and stop consonants [p], [t], [k], [b], [d], [g]. All words are produced with carrier phrases : "say __ again.", "__ say __ again." or "Say __ __ __"

- Variation of vocalic and consonantal features

Feature variations include vowel height, oral/nasal consonant opposition, voiced/unvoiced stops, fricatives and affricates, sonorant, and consonant clusters.

Target words are summarised in the table below

	Stops	Fricatives	Affricates	Sonorants	Clusters
initial	peel, bean [i] pin, bin [ɪ] pain, bane [e] pan, ban [æ]	fear, veer [ɪ] fin, vim [ɪ] feign, vain [e] fan, van [æ]		wean [i], yule [u] win [ɪ], year [ɪ], wane [e], yell [ɛ] wan [a], yawl [ɔ]	spall [sp] slam [sl] small [sm]
initial	team, deem [i] tame, dame [æ] tam, dam [æ]	seem, zeem [i] same, zane [e] sam, zam [æ]	cheer, jeer [ɪ] chain, jane [e] chan, jam [æ]	real, lean [i] rain, lame [e] ram, lam [æ]	
initial		thin [θɪ], dheen [ði] thain [θe], then [ðe] thumb [θʌ], dhawn [ðɔ]		mean [i], kneel [i] mint [ɪ] main [e], name [e] man [æ], nom [a]	
final	rape [ep] rabe [eb] late [et] laid [ed]	lafe [ef], laugh [æf], rave [ev], lave [æv], shelf [f], shelve [v] lace [es] raise [ez]	rache [etʃ] rage [edʒ]		rest [st], raised [zd] rent [nt], rend [nd] lance [ns], lapse [ps], false [ls]
intervocalic	ripple [p] ribbon [b]	riffle [f] riven [v] whistle [s] wizen [z]		rimmle [m] whittle [r], riddle [r]	
pretonic	repel [p], rebel [b] entire [t], endure [d] retire [t], reduce [d]	resource [s] resign [z]		wimple [mp], wimble [mb] wintle [nt], windle [nd] remain [m]	

- Word length and Stress position

Within a carrier phrase, target word length is modulated by adding suffixes successively, stress position is carried by lexical information so that aerodynamic process can be studied in an ecological context. For example:
pit, pity, pitying, pityingly

Amharic

- Simple consonants
- Labialised consonants
- Geminated consonants
- Glottalised consonants
- Ejectives

French

- Vowel & Logatome

Most phonetic features of French vowels and consonants are obtained, including nasals, sonorant, voiced/unvoiced stop and fricatives

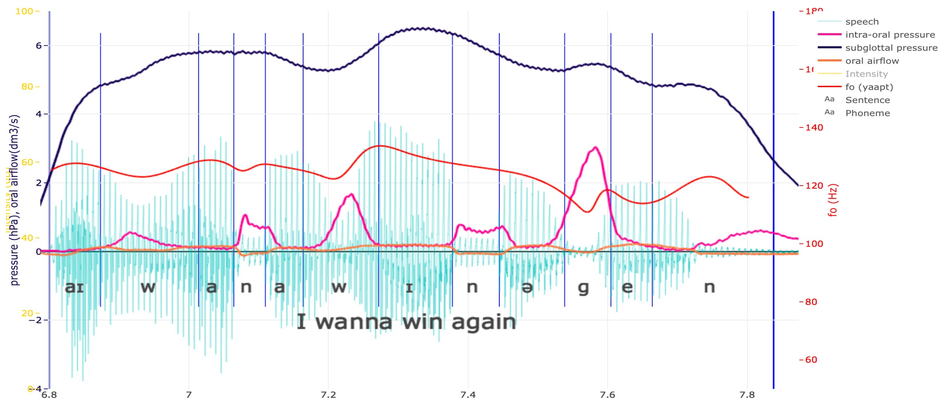
- Sentences

C'est une chanson triste.	C'est une chanson qui m'attriste.
C'est une maison grise.	C'est une maison qui me grise.
La démonstration du président de l'assemblée nationale m'a convaincu de la gravité de la situation.	La démonstration du président m'a semblé convaincante.
Cette nouvelle théorie linguistique provoque si j'ai bien compris, une nouvelle polémique.	Tu aurais tout intérêt si tu tiens à conserver son amitié à lui téléphoner plus souvent.
Tu vois cette maison ? C'est la maison que j'aimerais visiter.	Ce n'est pas le château, c'est la maison que j'aimerais visiter.
Mais non tu n'as rien compris, c'est la maison que j'aimerais visiter, pas le château.	Elle est complètement débile cette histoire de passeport volé.
Cette histoire de passeport volé elle est complètement débile	Je n'aime pas les films violents que l'on passe à la télé le samedi après-midi et toi?
Anne-Marie dit lui de venir tout de suite.	Dis-lui de venir tout de suite Anne-Marie.
Elle n'est pas venue Anne-Marie.	Selon moi, elle prendra le train ou elle ne viendra pas.
Elle viendra ou elle ne viendra pas ?	Il a une version écrite ou une version orale de ce document ?

Variations of sentence length, syntactic structure, modality are included within these sentences

Interface

A website allow all data to be downloaded: https://corpus.ilpga.fr/aerodynamics/index.php
Data I/O and conversion tool and visualisation interface are developed using Python and Plotly.



Conclusion

This speech aerodynamic database provides simultaneous and synchronous recordings of acoustic and aerodynamic measurements of speech production in three languages, accompanied by web interface for download and data visualisation. All data are provided with segmentation and annotation.
This database provides precious aerodynamic measurements on speech production, containing in particular direct Ps measurement. It will allow a substantial development of studies on the aerodynamics of speech production.

References

- Boersma, P. (2001). Praat, a system for doing phonetics by computer. Glot. Int., 5(9):341–345.
- Demolin, D., Trouville, R., Wang, R., and Signorello, R. (2017). Oral and nasal vowels effects on subglottal pressure. The Journal of the Acoustical Society of America, 142(4):2582–2582.
- Ghio, A. (2002). Phonedit: Multiparametric speech analysis. LPL (Aix-en-Provence, France) www.lpl.univ-aix.fr.
- Teston, B. and Galindo, B. (1990). The physiologia system: Description and technical specifications.