

iafpa
25



york

july 2016

25th Annual Conference of the International Association for Forensic Phonetics and Acoustics

York, UK
24th – 27th July 2016

Abstracts



UNIVERSITY *of York*

J P French *Associates*
Forensic speech and acoustics laboratory

Vowel sounds produced with varying production parameters: Conceptualisation and realisation of a database

Dieter Maurer¹, Thayabaran Kathiresan², Heidi Suter¹, Volker Dellwo²

¹Institute for the Performing Arts and Film, Zurich University of the Arts, Switzerland

dieter.maurer@zhdk.ch, heidi.suter@zhdk.ch

²Phonetics Laboratory, Department of Comparative Linguistics, University of Zurich, Switzerland

thayabaran.kathiresan@uzh.ch, volker.dellwo@uzh.ch

With regard to vowel sounds, there is no extensive, empirical database that documents systematic variation of basic production parameters. However, we consider such a corpus to be a prerequisite for a deeper understanding of both phoneme- and speaker-dependent acoustic characteristics. Currently, we are working on a corresponding database for Standard German vowels. In the following, the concept and status of realisation of this database, limited to the investigation of untrained speakers, is presented.

Concept

Untrained speakers (children, women, men) are selected according to three qualitative criteria: (1) a large vocal range (two octaves for adults and 1.5 octave for children at minimum), (2) the ability to reproduce sounds on a given f_0 , and (3) a clear spontaneous vowel articulation for a range of f_0 of one octave at minimum.—Each selected speaker produces the sounds of the long German vowels /i-y-e-ø-ε-a-o-u/ and varies basic production parameters such as f_0 (C-major scale up and down the entire vocal range; each pitch is presented to the speaker as electronic piano sound), vocal effort (medium, low, high), phoneme context (V for all f_0 and all vocal efforts, sVsV for middle and high f_0 and medium vocal effort only) and, in addition to the voiced sounds, phonation mode (breathy phonation in V context, whisper phonation in V and sVsV context). Each speaker also reads a phonetically balanced text. If, during a recording session, the speaker or examiner believes that a sound could possibly be improved, additional productions are recorded for the corresponding production condition.—All sounds are digitally recorded in a quiet room and with a constant speaker-microphone distance of 30 cm. The microphone input gain is adjusted according to the vocal effort investigated. In order to subsequently determine the actual sound pressure level, for each recording session, a 1 kHz sinus wave is recorded with a -20 dB gain using a specific calibration tool. Each sound in the database is annotated with standard information on speaker and indications on recorded vowel sounds. In order to provide additional indicative information, the database also features the results of acoustic analysis for the vowel nuclei (pitch contour, spectrum, LPC filter curve, spectrogram and formant tracks) and the results of a listening test (assignment of the perceived vowel quality) involving five professionally trained singers and speakers.

Actual status

Up to now, we have recorded eight adults and eight children (gender balanced). With the exception of one child, all speakers produced the sounds over a vocal range of two octaves or more. Considering that there is a vowel sound for each production condition (sound selection according to the highest identification score of the listening test), in average, c. 500 systematic recordings are available for each single speaker, or c. 8000 recordings for all 16 speakers.

Additional note

In parallel, we are also recording the sounds of professional singers and speakers, including additional variations of production parameters, such as production style, phoneme context in minimal pairs, and creak phonation.