

The Tracheal Resonance Database is a collection of utterances recorded from seven adult female and seven adult male subjects. Subjects were recruited from the students, faculty, and staff of MIT. The subjects' ages ranged from 18-78 years. The speakers reported no voice or hearing problems. Some of the speakers had previous experience being subjects for speech production experiments. Eleven of the subjects were native speakers of American English. One male subject (M1) was a native speaker of Canadian English, while another (M7) was a native speaker of British English. One female subject (F7) was a native speaker of Mandarin Chinese, who moved to the United States at age 10. The data were recorded in 2002 and 2003.

Target syllables were formed by embedding 16 vowels of American English in the syllables hVC and dVC (see table below). Orthographic versions of these target syllables were embedded in the carrier phrase "<syllable>, say <syllable> again".

Symbol	Syllable (hVC)	Syllable (dVC)	Notes
/æ/	had	dad	
/ɔ/	hawed	dawd	
/ɛ/	head	dead	
/ɜː/	heard	dird	"herd" for M7
/i/	heed	deed	
/ɪ/	hid	did	
/ɑ/	hodd	dodd	"hod" for M7
/ow/	hoed	dode	
/u/	hood	dood	
/ʌ/	hud hub	dud dub	
/uw/	who'd	dude	
/ei/	hade	dade	"hayed" for M7
/ai/	hide	dyed	
/oi/	hoid	doid	
/au/	how'd	dowed	
/iu/	hued	duud	

One subject (M4) appeared to include /ɑ/ in his dialect, and that was elicited using the orthographic syllable "hahd". Subject M7, the speaker of British English, produced an additional four syllables: "haird," "hard," "hered," and "hired." Six of the subjects (f1, f6, m3, m4, m5, m6) did not show a distinction between the vowels /ɔ/ and /ɑ/. Subjects f4, f5, f6, and m7 did not produce the dVC syllables. Subject f7 did not record the syllables "hued" and "hoed," and subject m2 did not record the syllable "howd."

The data were recorded in a sound-attenuated chamber at the MIT Speech Communication Group. An accelerometer was glued about one inch above the speaker's sternal notch. A microphone was placed approximately eight inches from the speaker. Five tokens of each sentence were presented to subjects in random order on a computer screen; all subjects produced the utter-

ances in the same order. Both the microphone and accelerometer signals were first amplified by a mixer, and then low-pass filtered at 4.8 kHz. These signals were sampled at 10,000 samples/sec and stored directly to a computer hard drive.

File name convention is

```
<subject_code>_<syllable>_<token_num>_audio_ms.wav  
<subject_code>_<syllable>_<token_num>_acc_ms.wav
```

where

- subject codes are f1-f7 for female subjects or m1-m7 for males
- syllable is as indicated in the table above
- token_num is the token number 1-5
- "audio" indicates the microphone signal and "acc" indicates the accelerometer signal.

(The "ms" in the file names indicates that the file format is Microsoft .wav rather than Klatt .wav, which is commonly used in the Speech Communication Group.)

The data were recorded by graduate student Xuemin Chi and undergraduate student Morgan Sonderegger, and are the basis of Sonderegger's S.B. thesis in Physics, and of an ASA poster by Sonderegger and Chi.

References

Sonderegger, Morgan (2004). Subglottal coupling and vowel space: An investigation in Quantal Theory. S.B. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Sonderegger, Morgan and Xuemin Chi (2004). "Subglottal coupling and vowel space," J. Acoust. Soc. Am., v. 115, p. 2540.