README file for the Child Subglottal Resonances Database.

In addition to the information presented in this README file, see Yeung, et al (2018) for further information about the corpus. The proper way to cite this corpus is to cite Yeung et al (2018).

Yeung, Gary, Steven M. Lulich, Jinxi Guo, Mitchell S. Sommers, Abeer Alwan (**2018**) *Subglottal resonances of American English speaking children*. Journal of the Acoustical Society of America **144**(6):3437-3449. [link]

The companion Subglottal Resonances Database [link], containing data from adult speakers, is also available and described by Lulich et al (2012):

Lulich, Steven M., John R. Morton, Harish Arsikere, Mitchell S. Sommers, Gary K.-F. Leung, Abeer Alwan (**2012**) *Subglottal resonances of adult male and female native speakers of American English*. Journal of the Acoustical Society of America **132**(4):2592-2602. [link]

Abstract
Subglottal resonances have received increasing attention in recent studies of speech production, perception, and technology. They affect voice production, divide vowels and consonants into discrete categories, affect vowel perception, and are useful in automatic speech recognition. The Child Subglottal Resonances Database is a new speech corpus of simultaneous microphone and (subglottal) accelerometer recordings of 19 male and 9 female child speakers of American English (AE), between 7;6 yr;mo and 17;8 yr;mo of age. The corpus materials consist of 34 monosyllables (14 "hVd" and 20 "CVb" words, where C is [b, d, g], and V includes all AE monophthongs and diphthongs) in a phonetically neutral carrier phrase ("I said a _____ again"), with a median of 6 repetitions of each word by each speaker, resulting in 5,247 individual microphone waveforms, and an equal number of accelerometer waveforms. The target vowel in each utterance has been hand-labeled to indicate the start, stop, and steady-state parts of the vowel. For diphthongs, the steady-state refers to the diphthong nucleus which occurs early in the vowel. Height and age of each speaker is also included. The corpus contains a total of 21,100 files distributed across 56 folders, and the total size of the corpus is approximately 15 GB. The corpus fills a gap in the literature on subglottal acoustics and will be useful for future studies in speech production, perception, and technology. It will be freely available to the speech research community. Work supported in part by NSF grant number 0905250.

Files and Naming Conventions
All files included in the corpus are organized into folders, with two folders for each speaker. The names of these folders have the form <speaker number>_<session number>. For instance, folder 102_1 contains the data from the first session of speaker 102, while folder 102_2 contains the data from the second session of speaker 102. For most speakers, the first session includes the CVb words, while the second session includes the hVd words. For speaker s123, the two sessions were recorded in reverse order. For speaker s124, the first

session was recorded in two takes, and these are distinguished by an "a" or "b" as part of the session number (s124a_1a_1_1_02-Sep-2011_acc.wav). Sixteen additional speakers were not included in the final database due to poorer recording quality or a limited quantity of repetitions. In total, there are 56 folders: 19 folders for the first session of males, 19 for the second session of males, 9 for the first session of females, and 9 for the second session of females.

Each folder for session 1 contains between 308 and 571 files, including 1 *.mat file, equal numbers of *_acc.wav files, *_mic.wav files, *_mic.bmp files, and *_mic.TextGrid files, and 2 additional files called *ACC1.wav and *ACC2.wav (henceforth "*ACC files"). There is a total of 12,263 files among all of the session 1 folders.

- **1 *.mat** file. This is a MATLAB file which contains information about the speaker and the recording session. These files can also be opened using free software, such as Python.

- **_acc.wav** files. These are accelerometer recordings of each repetition of each of the 20 CVb words.

- **_mic.wav** files. These are microphone recordings of each repetition of each of the 20 CVb words. Note that the microphone and accelerometer files were recorded simultaneously.

- **_mic.bmp** files. These are bitmap images of the target vowel in each of the utterances including the locations of the start, end, and steady-state labels.

- **_mic.TextGrid** files. These are the Praat TextGrid files containing the start, end, and steady-state labels.

- **2 *ACC** files. These are accerometer recordings of the sustained vowel [a:] for which excellent signal quality was achieved.

Each folder for session 2 contains between 237 and 389 files, including 1 *.mat file, and equal numbers of *_acc.wav files, *_mic.wav files, *_mic.bmp files, and _mic.TextGrid files. As in the session one folders, the *.mat file is a MATLAB file containing information about the speaker and the recording session, and the *_acc.wav, *_mic.wav, *_mic.bmp, and *_mic.TextGrid files are the recordings themselves, the images of the labeled target vowels, and the Praat TextGrids containing the labels for each repetition of each of the 14 hVd words. There is a total of 8,837 files among all of the session 2 folders.

As an example, consider the file "s102_2_1_2_12-Feb-2011_acc.wav". The initial "s" simply stands for "speaker", and the following numbers "102_2" are identical to the folder name - they are the speaker and session numbers. The next number (1) indicates the utterance number, and the next one indicates the repetition number (2). The date on which the recordings were made is then appended (12-Feb-2011), followed by an indication that this is an accelerometer signal file. Note that some repetitions may have been discarded and repeated, so that the repetition numbers do not always grow by consecutive integers. For instance, for speaker 102, session 2, utterance 4, there is no 1st repetition, i.e. there is no file s102_2_4_1_12-Feb-2011_acc.wav.

For session 1 the utterance numbers correspond to the following CVb words:

Number	Word	Number	Word	Number	Word
1	beeb	8	deb	15	geb
2	beb	9	dob	16	gob
3	bob	10	doob	17	goob
4	bibe	11	dibe	18	gibe
5	bowb	12	dowb	19	gowb
6	boib	13	doib	20	goib
7	deeb	14	geeb		

For session 2 the utterance numbers correspond to the following hVd words:

Number	Word	Number	Word	Number	Word
1	heed	6	hod	11	hide
2	hid	7	hud	12	how'ed
3	hay'ed	8	hode	13	hoid
4	head	9	hood	14	heard
5	had	10	who'd		

The *.mat Information Files

The information files contain information about the speaker and the recording session under four variables, of which the variable speakerInfo is of greatest importance for acoustic analysis of the recordings:

- displayInfo

This variable has three fields indicating the properties of the sound recordings as displayed to the experimenters during live recording:

      fs: The sampling rate. Typically, fs = 16,000.
      bitRate: The bit rate. Typically, bitRate = 16.
      nChannels: The number of channels. This is typically 2.

- recordingInfo

This variable has the same structure as displayInfo, but presents the sampling rate, bit rate, and number of channels of the sound recordings as actually made and saved to file.
      Sampling rate is 48,000 Hz, bit rate is 16, and the number of channels is 2.

- sentenceInfo

This variable contains information about the utterances produced by the speakers:

      wordListFileName: The name and relative location of the text file used to generate the word list to be recorded.
      start: The beginning part of the carrier phrase.
      end: The end part of the carrier phrase.
      wordList: The list of words to be recorded, derived from the text file in *wordListFileName*. This includes both 'I said a' and 'again' as the first two entries, even though they are also separately recorded under the *start* and *end* variables.

presentationID: The ID (utterance number) of each word to be recorded, in random order. For example, in session 1 there are 20 utterances. For speaker s139, the first utterance recorded was "I said a gowb again", with word ID 19.

utteranceOK: Either 1 or 0. Utterances deemed acceptable to the experimenter (i.e. no background noise, no slips of the tongue, etc.) were assigned a value of 1. Utterances deemed unacceptable were assigned a value of 0, and were repeated at the end of the experiment.

repetitionNumber: The indication of how many times a word had been presented to the speaker for recording. Early in each session, the value is usually 1 since most words have been presented only once. Late in each session, the value is usually 5 or 6 since most words have been presented at least 5 times.

- speakerInfo

This variable contains information about the speaker:

ID: The speaker ID.
Session: The session number (either 1 or 2).
Gender: Either 'male' or 'female'
Height: Height in feet and inches (if specified) or centimeters.
Age: Age in months.
nativeLanguage: The native language of the speaker.
testLanguage: The language of the test sentences.

Exceptions

There are two exceptions to the naming conventions, which are also noted above:

- For **speaker s123**, the two sessions were recorded in reverse order. Therefore, for this speaker the data in the session 1 folder are the hVd words, while the data in the session 2 folder are the CVb words.

- For **speaker s124**, the first session was recorded in two takes, and these are distinguished by an "a" or "b" as part of the session number (s124a_1a_1_1_02-Sep-2011_acc.wav).

There is one exception to the location and number of *ACC files:

- For **speaker s109**, there is only one *ACC file, and it is in the session 2 folder instead of the session 1 folder.

Number of Repetitions for Each Word by Speaker

The number of repetitions varies by both word and speaker. Values are presented in the tables below.

Session 1: Number of repetitions for each word by speaker. Note that for speaker s123, the corresponding data are in the Session 2 folder.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
s102	6	6	6	6	7	7	6	6	7	5	6	7	6	6	6	5	6	6	6	6
s103	7	8	7	6	8	6	6	8	7	6	6	9	9	8	7	8	6	8	6	6
s104	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
s107	6	6	6	6	5	6	6	6	7	7	6	6	6	6	5	6	6	6	5	6
s109	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
s111	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
s113	5	5	6	3	3	5	4	4	5	3	5	6	2	4	4	6	6	6	5	4
s116	6	6	6	6	6	6	6	6	7	6	6	6	6	6	5	6	6	6	6	5
s118	6	6	6	5	6	5	6	6	6	6	6	6	6	6	6	6	6	6	6	6
s119	6	5	5	4	5	5	5	6	6	3	5	6	6	6	7	6	4	6	7	5
s120	4	3	5	4	1	5	5	5	4	3	6	5	3	3	5	5	4	4	3	5
s121	6	5	5	6	3	5	5	6	6	5	6	5	6	5	2	5	5	5	3	5
s123	4	4	4	4	4	4	4	4	3	3	4	4	4	4	4	4	4	4	4	4
s124	4	3	5	3	5	3	3	5	5	2	4	3	5	4	0	4	5	5	2	6
s125	6	5	6	6	4	3	6	6	5	6	6	5	5	6	6	5	6	5	6	5
s126	6	7	6	6	6	6	6	6	6	6	6	6	6	5	6	6	6	7	6	6
s127	6	6	6	6	6	6	6	5	6	7	6	6	6	6	6	6	6	6	5	6
s129	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
s130	6	2	5	5	5	6	6	5	4	4	6	6	6	6	4	5	6	2	5	5
s132	6	6	6	6	5	6	6	6	4	3	6	4	5	6	6	2	5	5	6	5
s133	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	3	6
s135	5	6	6	6	3	5	6	5	6	6	6	4	6	6	6	6	4	4	5	5
s136	4	5	5	5	6	4	6	5	3	4	6	1	6	7	3	6	4	6	5	6
s139	6	5	6	6	3	5	6	4	5	6	6	6	7	6	5	6	6	6	5	6
s140	4	1	6	6	2	1	5	6	6	2	5	5	1	3	6	5	3	5	4	2
s141	6	6	6	6	6	6	6	6	5	5	6	6	6	6	6	6	6	5	5	6
s143	6	5	6	6	5	6	6	6	6	6	6	6	5	6	6	5	6	6	6	6
s144	6	6	6	6	6	6	6	6	5	6	6	5	6	5	6	6	6	7	6	5

Session 2: Number of repetitions for each word by speaker. Note that for speaker s123, the corresponding data are in the Session 1 folder.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
s102	6	6	5	6	6	6	5	6	6	6	6	6	7	7
s103	6	6	7	7	5	6	6	10	6	7	7	9	7	8
s104	6	6	6	6	7	6	7	6	6	6	6	6	6	6
s107	6	6	6	6	7	5	6	5	6	6	6	6	7	6
s109	6	6	6	5	6	6	6	6	6	6	6	7	6	6
s111	6	6	6	6	6	6	6	6	6	7	6	6	5	7
s113	6	1	5	5	6	1	4	3	5	5	5	4	5	6
s116	6	5	6	6	6	5	6	5	5	6	6	6	4	6
s118	6	6	6	6	6	6	6	6	6	6	6	5	6	6
s119	6	6	5	4	6	2	6	6	6	6	6	4	4	5
s120	6	5	6	6	4	4	7	6	5	6	6	6	5	6
s121	6	6	5	6	6	2	6	3	6	6	4	5	6	5
s123	6	6	6	6	6	5	5	6	6	5	5	6	6	6
s124	5	5	5	4	6	1	4	3	4	6	6	6	2	6
s125	5	6	6	4	6	6	6	6	6	6	6	6	4	6
s126	6	6	6	5	6	7	5	6	6	6	7	6	6	6
s127	7	6	6	6	6	6	6	6	6	6	6	6	6	6
s129	6	6	6	6	6	6	6	6	6	6	6	6	6	6
s130	6	6	6	5	6	4	5	5	5	6	6	6	6	6
s132	4	5	6	5	6	6	1	4	6	6	6	6	4	6
s133	6	6	6	6	6	6	6	4	6	7	6	6	6	6
s135	6	5	6	6	6	7	6	5	5	6	4	6	6	6
s136	4	5	6	5	6	4	5	5	4	5	6	6	6	0
s139	6	6	6	6	5	6	6	6	6	6	6	6	6	6
s140	6	0	5	5	2	1	6	4	4	6	6	3	5	6
s141	6	6	6	5	6	5	6	6	8	6	6	7	5	6
s143	6	5	6	5	6	5	5	5	6	6	6	5	6	6
s144	6	6	6	6	6	5	6	6	5	6	6	6	6	6