Information sheet for

UCLA 'Production and Perception of Linguistic Voice Quality' Project

Master Spreadsheets:

Acoustic measures and EGG measures

Column Name	***Description of categories***
0	This column contains information about inaccuracies in the measures. 0 end = do not use final third of file 0 = do not use file 0 beg = do not use first third of file 0 mid = do not use middle third of file
Label	Segment label from Praat textgrids
Language	Labels the language of the file in question. The spreadsheet contains information on the following languages: 1. Bo 2. English 3. Gujurati 4. Hmong 5. Luchun 6. Mandarin Chinese 7. Mazatec 8. (Black) Miao 9. Yi 10. Zapotec
Dialect/Village	In some cases, recordings were made in a number of locations, potentially with different dialects. This information is presented in this column. NA = taken from UCLA Phonetics Lab archive or just one recording site/dialect Mazatec = NA Gujurati = NA English = NA Mandarin = NA Luchun = NA Language: Miao Black = Black Miao

	Language: Mandarin CH = China TW = Taiwan Language: Zapotec SJG = San Juan Guelavia Zapotec SMZ = Santiago Matatlan Zapotec Language: Yi and Bo V1 = Village 1 V2 = Village 2 Language: Hmong W = White Hmong
Sex	M = Male F = Female
Speaker #	Speaker Number
Speaker	Sex + Speaker number
Lang_Spk	In order to ensure that each speaker has a unique identifier, use this column. Language_Sex+Speaker number
Phonation	This column contains information about the phonation type of the tokens. M = Modal C = Creaky B = Breathy L = Lax T =Tense
Lphon	Language and Phonation This code contains two parts. The first letter (occasionally including a second lowercase letter) refers to the language name and the second capitalized letter refers to the phonation type. BL = Bo Lax BT = Bo Tense BmB = Black Miao Breathy BmC = Black Miao Creaky

	BmM = Black Miao Modal BmT = Black Miao Tense EM = English Modal GB = Gujarati Breathy GM = Gujarati Modal HB = Hmong Breathy HC = Hmong Creaky HM = Hmong Modal LL = Luchun Lax LT = Luchun Tense MaM = Mandarin Chinese Modal MzB = Mazatec Breathy MzC = Mazatec Creaky MzM = Mazatec Modal YiL = Yi Lax YIT = Yi Tense ZB = Zapotec Breathy ZC = Zapotec Creaky ZM = Zapotec Modal
Vowel	This column lists the vowel type of each file in IPA. Rows with NA are refer to consonants. *Check this when you get back.
Oral/Nasal	This column states whether or not the vowel is oral or nasalized. O = Oral N = Nasalized
Tone from txtgrid	This column contains information using tone labels taken from the textgrids of each individual file. There are certain language-specific conventions to mark tone and this is reflected in this column. (eg. For Mandarin Chinese, the traditional Tone 1, 2, 3 or 4 are used). The user is referred to the Tone Cont. and Tone Cat. columns for the actual tonal contour of each token.

<table>
<tr><td>Tone Cont(our)</td><td>This column contains information about the actual tonal targets of each tone type. The standard IPA range is used, employing the numbers 1 through to 5, where 1 is considered low and 5 high. 3 is considered mid. NAs are used in languages which do not have lexical tone (i.e. English and Gujurati)</td></tr>
<tr><td>Tone</td><td>This column contains information about the tonal category (High, Mid, Low or Big) of the various tone contours. The ‘Big’ category refers to the cases in which there is a large change in pitch between the two or three tonal targets. Tone contours were divided into four categories as follows:

H: 55, 53, 45, 44, 535
M: 24, 33, 35, 31
L: 22, 21, 11, 213, 13
B: 51, 351, 15, 151, 153,</td></tr>
<tr><td>No. Tone</td><td>No. of Tone Targets

This column contains information about the number of tone targets there are for each tone type (1, 2 or 3).</td></tr>
<tr><td>Tphon</td><td>Composite of Tone column and Phonation column</td></tr>
<tr><td>CorV</td><td>Consonant or Vowel

This column states whether the segmented token is a consonant or vowel.

C = Consonant
V = Vowel</td></tr>
<tr><td>Pre_C</td><td>This column provides information about whether the consonant preceding the target vowel is oral or nasal:

O = Oral
N = Nasal
NA = Not applicable (i.e. token extracted is a consonant)</td></tr>
<tr><td>Aspiration</td><td>This column provides information about whether the consonant preceding the target vowel is aspirated or not.

1 = Aspiration of preceding consonant
0 = No aspiration</td></tr>
<tr><td>Duration</td><td>Duration of the labeled interval</td></tr>
</table>

seg_Start & seg_End	Start and end time of labeled interval
Notes on Acoustic measures	All voice quality measures were gathered using Voicesauce (Shue, Keating, Vicenik & Yu 2011). The following measures were gathered: 1. H1, H2, H4 2. A1, A2, A3 3. H1-H2, H2-H4, 4. H1-A1, H1-A2, H1-A3 5. CPP 6. Energy 7. HNR05, HNR15, HNR25, HNR35 8. SHR 9. shrF0, strF0, sF0, pF0 10. sF1, sF2, sF3, sF4 11. pF1, pF2, pF3, pF4 12. sB1, sB2, sB3, sB4 Values for all measures are given as an overall mean, and as a mean over different subsegments of the labeled interval (001-009).
Notes on EGG measures	All EGG measures were gathered using Voicesauce (Shue, Keating, Vicenik & Yu 2011). 1. CQ 2. CQ_H, CQ_PM, CQ_HT, 3. peak_Vel, peak_Vel_Time 4. min_Vel 5. SQ2-SQ1, SQ4-SQ3 6. ratio Values for all measures are given as an overall mean, and as a mean over different subsegments of the labeled interval (001-009).