

Tone Classification

Giang Le

10/28/2021

Data Import and Preprocessing

Our data contains 12 voice reports from 12 recording sessions.

```
# Read in all the voice reports.
dataFiles <- lapply(Sys.glob("*/channel1/acoustic_measurements_*.csv"), read.csv)

## Add the following categorical predictors.
# Gender: F and M (done)
# Noise type: quiet, 78 or 90 (done)
# Single token or token in a sentence
# Syllable type
# Tone

# Converting to DataFrames

f_1_78 <- as.data.frame(dataFiles[1])
f_1_90 <- as.data.frame(dataFiles[2])
f_1_q <- as.data.frame(dataFiles[3])

m_1_78 <- as.data.frame(dataFiles[4])
m_1_90 <- as.data.frame(dataFiles[5])
m_1_q <- as.data.frame(dataFiles[6])

m_2_78 <- as.data.frame(dataFiles[7])
m_2_90 <- as.data.frame(dataFiles[8])
m_2_q <- as.data.frame(dataFiles[9])

m_3_78 <- as.data.frame(dataFiles[10])
m_3_90 <- as.data.frame(dataFiles[11])
m_3_q <- as.data.frame(dataFiles[12])

# Assigning gender variable (0 for female and 1 for male)
f_1_78$gender = 0
f_1_90$gender = 0
f_1_q$gender = 0

m_1_78$gender = 1
m_1_90$gender = 1
m_1_q$gender = 1

m_2_78$gender = 1
m_2_90$gender = 1
```

```

m_2_q$gender = 1

m_3_78$gender = 1
m_3_90$gender = 1
m_3_q$gender = 1

# Assigning noise level
f_1_78$noise = 78
f_1_90$noise = 90
f_1_q$noise = 0

m_1_78$noise = 78
m_1_90$noise = 90
m_1_q$noise = 0

m_2_78$noise = 78
m_2_90$noise = 90
m_2_q$noise = 0

m_3_78$noise = 78
m_3_90$noise = 90
m_3_q$noise = 0

### Concatenate all dataframes
voice_reports <- rbind(f_1_78, f_1_90, f_1_q,
                      m_1_78, m_1_90, m_1_q,
                      m_2_78, m_2_90, m_2_q,
                      m_3_78, m_3_90, m_3_q)

dim(voice_reports)

## [1] 3786 36
mean(voice_reports$total.duration)

## [1] 0.3208114
sd(voice_reports$total.duration)

## [1] 0.1289975
min(voice_reports$total.duration)

## [1] 0.034
max(voice_reports$total.duration)

## [1] 0.779
summary(voice_reports$total.duration)

##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0340 0.2140 0.3110 0.3208 0.4180 0.7790

# Assigning if the token is single (1) or not (0).
voice_reports$single <- ifelse(grepl("single", voice_reports$sound.name), 1, 0)

# Assign syllable shapes (do later)

```

```

# Assign tone values
voice_reports$tone <- ifelse(grepl("a", voice_reports$sound.name, ignore.case=T), "A1",
                             ifelse(grepl("à", voice_reports$sound.name, ignore.case=T), "A2",
                                       ifelse(grepl("á", voice_reports$sound.name, ignore.case=T), "B1",
                                             ifelse(grepl("â", voice_reports$sound.name, ignore.case=T), "C1",
                                                   ifelse(grepl("ã", voice_reports$sound.name, ignore.case=T), "C2",
                                                         ifelse(grepl("ä", voice_reports$sound.name, ignore.case=T), "B2",
                                                               ifelse(grepl("ê", voice_reports$sound.name, ignore.case=T), "A1",
                                                                     ifelse(grepl("è", voice_reports$sound.name, ignore.case=T), "A2",
                                                                           ifelse(grepl("é", voice_reports$sound.name, ignore.case=T), "B1",
                                                                               ifelse(grepl("ẽ", voice_reports$sound.name, ignore.case=T), "C1",
                                                                                   ifelse(grepl("ë", voice_reports$sound.name, ignore.case=T), "C2",
                                                                                       ifelse(grepl("ê", voice_reports$sound.name, ignore.case=T), "B2",
                                                                                           ifelse(grepl("u", voice_reports$sound.name, ignore.case=T), "A1",
                                                                                               ifelse(grepl("ù", voice_reports$sound.name, ignore.case=T), "A2",
                                                                                                   ifelse(grepl("ú", voice_reports$sound.name, ignore.case=T), "B1",
                                                                                                       ifelse(grepl("û", voice_reports$sound.name, ignore.case=T), "C1",
                                                                                                           ifelse(grepl("ü", voice_reports$sound.name, ignore.case=T), "C2",
                                                                                                               ifelse(grepl("ü", voice_reports$sound.name, ignore.case=T), "B2",
                                                                                                                   ifelse(grepl("ö", voice_reports$sound.name, ignore.case=T), "B2", "NA")))))))))))))))

# Assign phonation types
voice_reports$phonation <- ifelse(grepl("A1", voice_reports$tone, ignore.case=T), "modal",
                                  ifelse(grepl("A2", voice_reports$tone, ignore.case=T), "breathy",
                                          ifelse(grepl("B1", voice_reports$tone, ignore.case=T), "modal",
                                                  ifelse(grepl("B2", voice_reports$tone, ignore.case=T), "creaky",
                                                          ifelse(grepl("C1", voice_reports$tone, ignore.case=T), "creaky",
                                                                  ifelse(grepl("C2", voice_reports$tone, ignore.case=T), "creaky", "NA"))))))))

# Assign creakiness or not
voice_reports$creaky <- ifelse(grepl("creaky", voice_reports$phonation, ignore.case=T), 1, 0)
head(voice_reports, 20)

```

##	sound.name	total.duration	intensity	spectraltilt	median.F0	mean.F0	sd.F0
## 1	độ	0.229	60.102	-27.852	218.143	200.097	47.803
## 2	độ	0.176	63.003	-27.852	207.166	204.083	6.503
## 3	ẽ	0.444	61.756	-11.698	237.639	187.930	72.74
## 4	ẽ	0.330	59.292	-11.698	223.784	224.516	11.948
## 5	ê	0.430	60.768	-10.194	221.383	220.045	3.625
## 6	ê	0.450	62.075	-10.194	216.449	217.061	4.037
## 7	ê	0.252	61.200	-11.646	215.220	266.343	134.326
## 8	ê	0.173	62.570	-11.646	216.202	204.334	31.707
## 9	ẽ	0.469	60.697	-16.422	165.238	162.472	32.743
## 10	ẽ	0.235	60.548	-16.422	167.355	170.803	14.12
## 11	ẽ	0.419	61.703	-13.377	186.304	203.317	30.96
## 12	ẽ	0.382	60.027	-13.377	180.660	189.969	17.373
## 13	è	0.537	61.288	-11.035	178.111	178.156	7.269
## 14	è	0.450	61.082	-11.035	181.224	180.169	7.256
## 15	TÚT_single	0.175	64.908	-30.200	263.405	262.553	16.382
## 16	TÚT_single	0.183	65.207	-30.200	251.474	253.861	5.784
## 17	TÚT_single	0.194	65.341	-31.444	198.831	195.492	6.208

## 18	TUT_single	0.229	63.707	-31.444	194.444	195.044	6.422
## 19	thề	0.417	61.015	-18.748	91.309	92.451	5.401
## 20	thề	0.426	62.959	-18.748	174.213	176.461	12.694
##	min.F0 max.F0	number.pulses	number.periods	mean.periods	sd.period		
## 1	73.534 262.563	2	0	--undefined--	--undefined--		
## 2	186.197 210.989	27	26	4.914	0.212		
## 3	104.652 266.835	2	0	--undefined--	--undefined--		
## 4	199.758 241.472	69	68	4.447	0.248		
## 5	205.575 224.446	1	0	--undefined--	--undefined--		
## 6	197.153 222.391	92	91	4.599	0.07		
## 7	193.569 599.754	2	0	--undefined--	--undefined--		
## 8	87.511 218.542	27	25	4.709	0.239		
## 9	68.204 203.482	2	0	--undefined--	--undefined--		
## 10	149.351 194.444	36	35	5.886	0.482		
## 11	171.339 279.908	1	0	--undefined--	--undefined--		
## 12	173.984 236.270	69	68	5.232	0.485		
## 13	166.773 191.322	1	0	--undefined--	--undefined--		
## 14	165.546 190.824	78	77	5.551	0.231		
## 15	221.589 341.236	1	0	--undefined--	--undefined--		
## 16	248.457 269.794	40	39	3.925	0.106		
## 17	185.141 201.616	1	0	--undefined--	--undefined--		
## 18	185.762 213.028	40	39	5.146	0.153		
## 19	85.531 112.158	1	0	--undefined--	--undefined--		
## 20	159.266 216.873	71	70	5.695	0.378		
##	fraction.of.locally.unvoiced.frames						
## 1		6.383					
## 2		2.857					
## 3		3.030					
## 4		0.000					
## 5		1.042					
## 6		1.000					
## 7		13.208					
## 8		2.941					
## 9		2.857					
## 10		0.000					
## 11		0.000					
## 12		0.000					
## 13		0.826					
## 14		0.000					
## 15		2.857					
## 16		0.000					
## 17		0.000					
## 18		0.000					
## 19		0.000					
## 20		0.000					
##							
##							
## 1	of locally unvoiced frames: 6.383%	(3 / 47)					0
## 2	of locally unvoiced frames: 2.857%	(1 / 35)					0
## 3	of locally unvoiced frames: 3.030%	(3 / 99)					0
## 4	of locally unvoiced frames: 0	(0 / 72)					0
## 5	of locally unvoiced frames: 1.042%	(1 / 96)					0
## 6	of locally unvoiced frames: 1.000%	(1 / 100)					0
## 7	of locally unvoiced frames: 13.208%	(7 / 53)					0
## 8	of locally unvoiced frames: 2.941%	(1 / 34)					0

```

## 9 of locally unvoiced frames: 2.857% (3 / 105) 0
## 10 of locally unvoiced frames: 0 (0 / 49) 0
## 11 of locally unvoiced frames: 0 (0 / 93) 0
## 12 of locally unvoiced frames: 0 (0 / 84) 0
## 13 of locally unvoiced frames: 0.826% (1 / 121) 0
## 14 of locally unvoiced frames: 0 (0 / 101) 0
## 15 of locally unvoiced frames: 2.857% (1 / 35) 0
## 16 of locally unvoiced frames: 0 (0 / 36) 0
## 17 of locally unvoiced frames: 0 (0 / 39) 0
## 18 of locally unvoiced frames: 0 (0 / 47) 0
## 19 of locally unvoiced frames: 0 (0 / 92) 0
## 20 of locally unvoiced frames: 0 (0 / 95) 0
## degree.of.voice.breaks degree
## 1 0 of voice breaks: 0 (0 seconds / 0.229018 seconds)
## 2 0 of voice breaks: 0 (0 seconds / 0.176274 seconds)
## 3 0 of voice breaks: 0 (0 seconds / 0.444099 seconds)
## 4 0 of voice breaks: 0 (0 seconds / 0.330119 seconds)
## 5 0 of voice breaks: 0 (0 seconds / 0 seconds)
## 6 0 of voice breaks: 0 (0 seconds / 0.449857 seconds)
## 7 0 of voice breaks: 0 (0 seconds / 0.251984 seconds)
## 8 0 of voice breaks: 0 (0 seconds / 0.173127 seconds)
## 9 0 of voice breaks: 0 (0 seconds / 0.469197 seconds)
## 10 0 of voice breaks: 0 (0 seconds / 0.234767 seconds)
## 11 0 of voice breaks: 0 (0 seconds / 0 seconds)
## 12 0 of voice breaks: 0 (0 seconds / 0.381955 seconds)
## 13 0 of voice breaks: 0 (0 seconds / 0 seconds)
## 14 0 of voice breaks: 0 (0 seconds / 0.450367 seconds)
## 15 0 of voice breaks: 0 (0 seconds / 0 seconds)
## 16 0 of voice breaks: 0 (0 seconds / 0.182882 seconds)
## 17 0 of voice breaks: 0 (0 seconds / 0 seconds)
## 18 0 of voice breaks: 0 (0 seconds / 0.229112 seconds)
## 19 0 of voice breaks: 0 (0 seconds / 0 seconds)
## 20 0 of voice breaks: 0 (0 seconds / 0.425752 seconds)
## jitter.local jitter.local.abs jitter.rap jitter.ppq5
## 1 --undefined-- --undefined-- --undefined-- --undefined--
## 2 1.74 85.486 0.746 0.376
## 3 --undefined-- --undefined-- --undefined-- --undefined--
## 4 0.626 27.845 0.241 0.253
## 5 --undefined-- --undefined-- --undefined-- --undefined--
## 6 0.448 20.622 0.222 0.175
## 7 --undefined-- --undefined-- --undefined-- --undefined--
## 8 1.247 58.7 0.273 0.435
## 9 --undefined-- --undefined-- --undefined-- --undefined--
## 10 1.81 106.54 0.853 0.923
## 11 --undefined-- --undefined-- --undefined-- --undefined--
## 12 0.59 30.893 0.195 0.248
## 13 --undefined-- --undefined-- --undefined-- --undefined--
## 14 0.498 27.666 0.214 0.226
## 15 --undefined-- --undefined-- --undefined-- --undefined--
## 16 0.977 38.342 0.563 0.605
## 17 --undefined-- --undefined-- --undefined-- --undefined--
## 18 0.42 21.599 0.182 0.166
## 19 --undefined-- --undefined-- --undefined-- --undefined--
## 20 0.75 42.721 0.365 0.315

```

##	shimmer.local	shimmer.local.db	shimmer.apq3	shimmer.apq5			
## 1	--undefined--	--undefined--	--undefined--	--undefined--			
## 2	4.275	0.442	1.14	1.762			
## 3	--undefined--	--undefined--	--undefined--	--undefined--			
## 4	1.327	0.117	0.543	0.625			
## 5	--undefined--	--undefined--	--undefined--	--undefined--			
## 6	1.751	0.163	0.865	0.784			
## 7	--undefined--	--undefined--	--undefined--	--undefined--			
## 8	2.244	0.195	0.794	1.168			
## 9	--undefined--	--undefined--	--undefined--	--undefined--			
## 10	4.857	0.634	1.579	1.607			
## 11	--undefined--	--undefined--	--undefined--	--undefined--			
## 12	1.927	0.175	0.551	0.642			
## 13	--undefined--	--undefined--	--undefined--	--undefined--			
## 14	2.16	0.214	0.461	0.605			
## 15	--undefined--	--undefined--	--undefined--	--undefined--			
## 16	2.131	0.196	0.824	1.047			
## 17	--undefined--	--undefined--	--undefined--	--undefined--			
## 18	3.799	0.344	2.123	1.871			
## 19	--undefined--	--undefined--	--undefined--	--undefined--			
## 20	1.68	0.179	0.653	0.665			
##	shimmer.apq11	mean.autocorr	mean.NHR	mean.HNR	F1	F2	F3
## 1	--undefined--	0.878	0.179	12.707	483.544	892.193	3164.547
## 2	4.392	0.962	0.053	18.742	483.544	892.193	3164.547
## 3	--undefined--	0.902	0.140	15.992	579.219	2386.550	3061.336
## 4	1.38	0.979	0.023	20.417	579.219	2386.550	3061.336
## 5	--undefined--	0.974	0.052	25.633	459.312	2321.976	2681.153
## 6	0.979	0.979	0.038	26.296	459.312	2321.976	2681.153
## 7	--undefined--	0.882	0.184	12.837	589.323	2276.287	2893.363
## 8	1.602	0.894	0.156	15.351	589.323	2276.287	2893.363
## 9	--undefined--	0.938	0.079	15.077	492.646	2382.580	2887.993
## 10	2.254	0.966	0.038	16.928	492.646	2382.580	2887.993
## 11	--undefined--	0.975	0.031	19.640	469.019	2381.116	2864.526
## 12	1.09	0.988	0.012	21.111	469.019	2381.116	2864.526
## 13	--undefined--	0.987	0.018	25.883	496.146	2391.518	2735.378
## 14	1.299	0.985	0.022	23.938	496.146	2391.518	2735.378
## 15	--undefined--	0.939	0.105	20.641	428.646	820.776	3166.460
## 16	1.703	0.984	0.018	23.193	428.646	820.776	3166.460
## 17	--undefined--	0.959	0.056	20.837	399.714	935.860	2910.953
## 18	1.922	0.983	0.023	24.161	399.714	935.860	2910.953
## 19	--undefined--	0.975	0.026	18.324	502.106	2248.255	2971.067
## 20	1.156	0.988	0.013	21.172	502.106	2248.255	2971.067
##	F4	gender	noise	single	tone	phonation	creaky
## 1	3701.005	0	78	0	B2	creaky	1
## 2	3701.005	0	78	0	B2	creaky	1
## 3	4024.072	0	78	0	C2	creaky	1
## 4	4024.072	0	78	0	C2	creaky	1
## 5	3764.261	0	78	0	A1	modal	0
## 6	3764.261	0	78	0	A1	modal	0
## 7	3616.906	0	78	0	B2	creaky	1
## 8	3616.906	0	78	0	B2	creaky	1
## 9	4283.656	0	78	0	C1	creaky	1
## 10	4283.656	0	78	0	C1	creaky	1
## 11	3900.203	0	78	0	B1	modal	0

```
## 12 3900.203      0    78      0   B1      modal      0
## 13 3869.799      0    78      0   A2      breathy     0
## 14 3869.799      0    78      0   A2      breathy     0
## 15 3803.949      0    78      1   B1      modal      0
## 16 3803.949      0    78      1   B1      modal      0
## 17 3651.814      0    78      1   B2      creaky      1
## 18 3651.814      0    78      1   B2      creaky      1
## 19 4094.997      0    78      0   A2      breathy     0
## 20 4094.997      0    78      0   A2      breathy     0
```

Checking

```
# How many values are of each category
length(voice_reports$tone[voice_reports$tone == "A1"])
```

```
## [1] 574
```

```
## [1] 574
length(voice_reports$tone[voice_reports$tone == "A2"])
```

```
## [1] 575
```

```
## [1] 575
length(voice_reports$tone[voice_reports$tone == "B1"])
```

```
## [1] 719
```

```
## [1] 719
length(voice_reports$tone[voice_reports$tone == "B2"])
```

```
## [1] 768
```

```
## [1] 768
length(voice_reports$tone[voice_reports$tone == "C1"])
```

```
## [1] 575
```

```
## [1] 575
length(voice_reports$tone[voice_reports$tone == "C2"])
```

```
## [1] 575
```

```
## [1] 575
length(voice_reports$tone[voice_reports$tone == "NA"])
```

```
## [1] 0
```

```
## [1] 0
```

Convert categorical values to factors

```
## Not sure if this is necessary for variables already binarily coded.
voice_reports$gender <- as.factor(voice_reports$gender)
voice_reports$noise <- as.factor(voice_reports$noise)
voice_reports$tone <- as.factor(voice_reports$tone)
voice_reports$single <- as.factor(voice_reports$single)
voice_reports$phonation <- as.factor(voice_reports$phonation)
voice_reports$creaky <- as.factor(voice_reports$creaky)
```

Summary of current data

```
summary(voice_reports)
```

```
##      sound.name      total.duration      intensity      spectraltilt
## Length:3786      Min.      :0.0340      Min.      :46.69      Min.      : -44.880
## Class :character  1st Qu.:0.2140      1st Qu.:60.91      1st Qu.: -26.676
## Mode  :character  Median :0.3110      Median :64.94      Median : -17.795
##                               Mean  :0.3208      Mean  :64.74      Mean  : -19.761
##                               3rd Qu.:0.4180      3rd Qu.:69.47      3rd Qu.: -13.379
##                               Max.   :0.7790      Max.   :80.76      Max.    :  2.476
##      median.F0      mean.F0      sd.F0      min.F0
## Min.      : 64.1      Min.      : 65.06      Length:3786      Min.      : 53.19
## 1st Qu.:118.3      1st Qu.:120.72      Class :character  1st Qu.: 89.42
## Median :143.2      Median :146.21      Mode  :character  Median :115.19
## Mean   :154.8      Mean   :158.63                               Mean   :123.80
## 3rd Qu.:178.8      3rd Qu.:185.94                               3rd Qu.:150.39
## Max.   :572.5      Max.   :435.42                               Max.   :299.23
##      max.F0      number.pulses      number.periods      mean.periods
## Min.      : 68.75      Min.      : 1.00      Min.      : 0.00      Length:3786
## 1st Qu.:143.49      1st Qu.: 1.00      1st Qu.: 0.00      Class :character
## Median :172.48      Median : 5.00      Median : 3.50      Mode  :character
## Mean   :206.09      Mean   :23.35      Mean   :22.10
## 3rd Qu.:234.74      3rd Qu.:39.00      3rd Qu.:37.75
## Max.   :599.75      Max.   :151.00      Max.   :150.00
##      sd.period      fraction.of.locally.unvoiced.frames      fraction
## Length:3786      Min.      : 0.000      Length:3786
## Class :character  1st Qu.: 0.000      Class :character
## Mode  :character  Median : 0.000      Mode  :character
##                               Mean   : 2.443
##                               3rd Qu.: 2.041
##                               Max.   :50.000
##      number.of.voice.breaks      degree.of.voice.breaks      degree
## Min.      :0.00000      Min.      : 0.000      Length:3786
## 1st Qu.:0.00000      1st Qu.: 0.000      Class :character
## Median :0.00000      Median : 0.000      Mode  :character
## Mean   :0.08928      Mean   : 1.859
## 3rd Qu.:0.00000      3rd Qu.: 0.000
## Max.   :3.00000      Max.   :62.660
##      jitter.local      jitter.local.abs      jitter.rap      jitter.ppq5
## Length:3786      Length:3786      Length:3786      Length:3786
## Class :character  Class :character  Class :character  Class :character
## Mode  :character  Mode  :character  Mode  :character  Mode  :character
##
##
##
##      shimmer.local      shimmer.local.db      shimmer.apq3      shimmer.apq5
## Length:3786      Length:3786      Length:3786      Length:3786
## Class :character  Class :character  Class :character  Class :character
## Mode  :character  Mode  :character  Mode  :character  Mode  :character
##
##
##
##      shimmer.apq11      mean.autocorr      mean.NHR      mean.HNR
```



```
## Length:3786      Min.   :0.5900   Min.   :0.0007   Min.   : 1.714
## Class :character  1st Qu.:0.8780   1st Qu.:0.0230   1st Qu.:11.672
## Mode  :character  Median :0.9480   Median :0.0650   Median :16.133
##                  Mean   :0.9217   Mean   :0.1166   Mean   :16.232
##                  3rd Qu.:0.9790   3rd Qu.:0.1750   3rd Qu.:20.543
##                  Max.   :0.9990   Max.   :0.7660   Max.   :34.233
##      F1           F2           F3           F4           gender
## Min.   : 201.0   Min.   : 514.6   Min.   :1860   Min.   :2873   0: 947
## 1st Qu.: 380.1   1st Qu.: 907.9   1st Qu.:2598   1st Qu.:3526   1:2839
## Median : 472.7   Median :1584.1   Median :2722   Median :3695
## Mean   : 552.5   Mean   :1498.5   Mean   :2745   Mean   :3734
## 3rd Qu.: 767.6   3rd Qu.:1958.7   3rd Qu.:2866   3rd Qu.:3888
## Max.   :1192.0   Max.   :2805.6   Max.   :3458   Max.   :4950
## noise    single    tone      phonation    creaky
## 0 :1260    0:1893    A1:574    breathy: 575    0:1868
## 78:1263    1:1893    A2:575    creaky :1918    1:1918
## 90:1263                    B1:719    modal   :1293
##                                B2:768
##                                C1:575
##                                C2:575
```

Logistic Regression on Gender

```
head(voice_reports)
```

```
##      sound.name total.duration intensity spectraltilt median.F0 mean.F0 sd.F0
## 1      đò        0.229    60.102      -27.852    218.143 200.097 47.803
## 2      đò        0.176    63.003      -27.852    207.166 204.083 6.503
## 3      ẽ        0.444    61.756      -11.698    237.639 187.930 72.74
## 4      ẽ        0.330    59.292      -11.698    223.784 224.516 11.948
## 5      ê        0.430    60.768      -10.194    221.383 220.045 3.625
## 6      ê        0.450    62.075      -10.194    216.449 217.061 4.037
##      min.F0 max.F0 number.pulses number.periods mean.periods sd.period
## 1  73.534 262.563           2           0 --undefined-- --undefined--
## 2 186.197 210.989          27          26         4.914         0.212
## 3 104.652 266.835           2           0 --undefined-- --undefined--
## 4 199.758 241.472          69          68         4.447         0.248
## 5 205.575 224.446           1           0 --undefined-- --undefined--
## 6 197.153 222.391          92          91         4.599         0.07
##      fraction.of.locally.unvoiced.frames
## 1                                6.383
## 2                                2.857
## 3                                3.030
## 4                                0.000
## 5                                1.042
## 6                                1.000
##
##      fraction number.of.voice.breaks
## 1 of locally unvoiced frames: 6.383% (3 / 47)          0
## 2 of locally unvoiced frames: 2.857% (1 / 35)          0
## 3 of locally unvoiced frames: 3.030% (3 / 99)          0
## 4 of locally unvoiced frames: 0 (0 / 72)              0
## 5 of locally unvoiced frames: 1.042% (1 / 96)          0
## 6 of locally unvoiced frames: 1.000% (1 / 100)         0
##      degree.of.voice.breaks                                degree
```

```

## 1          0 of voice breaks: 0   (0 seconds / 0.229018 seconds)
## 2          0 of voice breaks: 0   (0 seconds / 0.176274 seconds)
## 3          0 of voice breaks: 0   (0 seconds / 0.444099 seconds)
## 4          0 of voice breaks: 0   (0 seconds / 0.330119 seconds)
## 5          0 of voice breaks: 0   (0 seconds / 0 seconds)
## 6          0 of voice breaks: 0   (0 seconds / 0.449857 seconds)
##      jitter.local jitter.local.abs      jitter.rap      jitter.ppq5
## 1 --undefined-- --undefined-- --undefined-- --undefined--
## 2          1.74          85.486          0.746          0.376
## 3 --undefined-- --undefined-- --undefined-- --undefined--
## 4          0.626          27.845          0.241          0.253
## 5 --undefined-- --undefined-- --undefined-- --undefined--
## 6          0.448          20.622          0.222          0.175
##      shimmer.local shimmer.local.db      shimmer.apq3      shimmer.apq5
## 1 --undefined-- --undefined-- --undefined-- --undefined--
## 2          4.275          0.442          1.14          1.762
## 3 --undefined-- --undefined-- --undefined-- --undefined--
## 4          1.327          0.117          0.543          0.625
## 5 --undefined-- --undefined-- --undefined-- --undefined--
## 6          1.751          0.163          0.865          0.784
##      shimmer.apq11 mean.autocorr mean.NHR mean.HNR      F1      F2      F3
## 1 --undefined--          0.878      0.179      12.707 483.544 892.193 3164.547
## 2          4.392          0.962      0.053      18.742 483.544 892.193 3164.547
## 3 --undefined--          0.902      0.140      15.992 579.219 2386.550 3061.336
## 4          1.38          0.979      0.023      20.417 579.219 2386.550 3061.336
## 5 --undefined--          0.974      0.052      25.633 459.312 2321.976 2681.153
## 6          0.979          0.979      0.038      26.296 459.312 2321.976 2681.153
##      F4 gender noise single tone phonation creaky
## 1 3701.005      0      78      0 B2      creaky      1
## 2 3701.005      0      78      0 B2      creaky      1
## 3 4024.072      0      78      0 C2      creaky      1
## 4 4024.072      0      78      0 C2      creaky      1
## 5 3764.261      0      78      0 A1      modal      0
## 6 3764.261      0      78      0 A1      modal      0

```

```

logit_gender = glm(gender ~ mean.F0 + total.duration + intensity + mean.HNR, family = "binomial", data = voice_reports)
summary(logit_gender)

```

```

##
## Call:
## glm(formula = gender ~ mean.F0 + total.duration + intensity +
##      mean.HNR, family = "binomial", data = voice_reports)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -3.0389  -0.0104   0.1600   0.4089   3.6153
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -10.326029  0.682930 -15.120  <2e-16 ***
## mean.F0       -0.037658  0.001378 -27.328  <2e-16 ***
## total.duration -4.282263  0.485318  -8.824  <2e-16 ***
## intensity      0.333715  0.013984  23.865  <2e-16 ***
## mean.HNR      -0.103735  0.010847  -9.563  <2e-16 ***
## ---

```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 4259.1  on 3785  degrees of freedom
## Residual deviance: 2030.6  on 3781  degrees of freedom
## AIC: 2040.6
##
## Number of Fisher Scoring iterations: 6
```

Logistic Regression on Creaky

```
logit_creaky = glm(creaky ~ mean.F0 + total.duration + intensity + spectraltilt + number.pulses + mean.HNR, family = "binomial", data = voice_reports)
```

```
##
## Call:
## glm(formula = creaky ~ mean.F0 + total.duration + intensity +
##      spectraltilt + number.pulses + mean.HNR, family = "binomial",
##      data = voice_reports)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -3.3908  -0.5292   0.0479   0.5794   3.0066
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    3.201521    0.513323   6.237 4.46e-10 ***
## mean.F0         0.018511    0.001155  16.027 < 2e-16 ***
## total.duration -1.491809    0.393382  -3.792 0.000149 ***
## intensity      -0.003770    0.007247  -0.520 0.602949
## spectraltilt   -0.110813    0.006679 -16.592 < 2e-16 ***
## number.pulses   0.002771    0.001741   1.591 0.111564
## mean.HNR       -0.470632    0.015475 -30.413 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5247.9  on 3785  degrees of freedom
## Residual deviance: 2897.3  on 3779  degrees of freedom
## AIC: 2911.3
##
## Number of Fisher Scoring iterations: 6
```

Multinomial Regression to predict the Noise Level.

```
## Use the multinom function from the nnet package (Ref: https://stats.idre.ucla.edu/r/dae/multinomial-)
library("nnet")
# Use the 78 noise level as the reference level
voice_reports$noise2 <- relevel(voice_reports$noise, ref = "78")
multinom_noise <- multinom(noise2 ~ mean.F0 + total.duration + intensity + spectraltilt, data=voice_reports)
```

```
## # weights: 18 (10 variable)
## initial value 4159.346125
## iter 10 value 2581.712398
## iter 20 value 2500.771170
## final value 2500.770889
## converged
```

```
summary(multinom_noise)
```

```
## Call:
## multinom(formula = noise2 ~ mean.F0 + total.duration + intensity +
##          spectraltilt, data = voice_reports)
##
## Coefficients:
##      (Intercept)      mean.F0 total.duration  intensity spectraltilt
## 0      23.98251 -0.013331533      -5.795162 -0.3457656 -0.06686426
## 90     -18.31800  0.006286322       2.800752  0.2429637  0.01129598
##
## Std. Errors:
##      (Intercept)      mean.F0 total.duration  intensity spectraltilt
## 0      0.9973863 0.0013025075      0.4981253 0.01486610 0.006887869
## 90      0.8839808 0.0008520673      0.3566817 0.01183026 0.005179322
##
## Residual Deviance: 5001.542
## AIC: 5021.542
```

```
# The result in general supports our predictions regarding the relationship
# between relative noise levels
# and F0, duration, intensity, etc.
```

```
# For instance,
# A one-unit increase in mean F0 is associated with the decrease in the
# log odds of quiet vs. 78 noise level in the amount of 0.0133
# A one-unit increase in mean F0 is associated with the increase in the
# log odds of 90 noise vs. 78 noise in the amount of 0.006
```

```
# A one-unit increase in duration is associated with the decrease in the
# log odds of quiet vs. 78 noise level in the amount of 5.795
# A one-unit increase in duration is associated with the increase in the
# log odds of 90 noise vs. 78 noise in the amount of 2.80
```

```
# A one-unit increase in intensity is associated with the decrease in the
# log odds of quiet vs. 78 noise level in the amount of 0.35
# A one-unit increase in intensity is associated with the increase in the
# log odds of 90 noise vs. 78 noise in the amount of 0.24
```

```
## Giang to double check this result
# A one-unit increase in spectraltilt is associated with the decrease in the
# log odds of quiet vs. 78 noise level in the amount of 0.066
# A one-unit increase in spectraltilt is associated with the increase in the
# log odds of 90 noise vs. 78 noise in the amount of 0.011
```

Clean up undefined values to prepare for Classification

Method 1: Simply drop values that are undefined in jitter and shimmer variables

```
voice_reports_clean <- voice_reports[!(voice_reports$jitter.local==" --undefined-- " | voice_reports$shimmer.local==" --undefined-- ")]
```

Convert two variables to numeric

```
voice_reports_clean$jitter.local <- as.numeric(voice_reports_clean$jitter.local)
```

```
voice_reports_clean$shimmer.local <- as.numeric(voice_reports_clean$shimmer.local)
```

```
summary(voice_reports_clean)
```

```
##      sound.name      total.duration      intensity      spectraltilt
## Length:1894      Min.   :0.0430      Min.   :47.04      Min.   : -44.880
## Class :character  1st Qu.:0.2092      1st Qu.:60.63      1st Qu.: -26.650
## Mode  :character  Median :0.2980      Median :64.76      Median : -17.773
##                               Mean  :0.3164      Mean  :64.52      Mean  : -19.752
##                               3rd Qu.:0.4100      3rd Qu.:69.17      3rd Qu.: -13.379
##                               Max.   :0.7790      Max.   :80.55      Max.   :  2.476
##      median.F0      mean.F0      sd.F0      min.F0
## Min.   : 64.1      Min.   : 65.06      Length:1894      Min.   : 53.19
## 1st Qu.:117.1      1st Qu.:120.01      Class :character  1st Qu.: 89.66
## Median :142.0      Median :144.38      Mode  :character  Median :115.21
## Mean   :152.8      Mean   :156.54                               Mean   :123.25
## 3rd Qu.:177.5      3rd Qu.:184.03                               3rd Qu.:149.36
## Max.   :570.8      Max.   :435.42                               Max.   :281.04
##      max.F0      number.pulses      number.periods      mean.periods
## Min.   : 68.75      Min.   : 4.00      Min.   : 3.00      Length:1894
## 1st Qu.:142.09      1st Qu.: 25.00      1st Qu.: 24.00      Class :character
## Median :169.10      Median : 39.00      Median : 37.50      Mode  :character
## Mean   :202.04      Mean   : 45.37      Mean   : 44.17
## 3rd Qu.:230.61      3rd Qu.: 62.00      3rd Qu.: 61.00
## Max.   :598.15      Max.   :151.00      Max.   :150.00
##      sd.period      fraction.of.locally.unvoiced.frames      fraction
## Length:1894      Min.   : 0.000      Length:1894
## Class :character  1st Qu.: 0.000      Class :character
## Mode  :character  Median : 0.000      Mode  :character
##                               Mean   : 2.325
##                               3rd Qu.: 1.923
##                               Max.   :47.945
##      number.of.voice.breaks      degree.of.voice.breaks      degree
## Min.   :0.000      Min.   : 0.000      Length:1894
## 1st Qu.:0.000      1st Qu.: 0.000      Class :character
## Median :0.000      Median : 0.000      Mode  :character
## Mean   :0.104      Mean   : 1.272
## 3rd Qu.:0.000      3rd Qu.: 0.000
## Max.   :3.000      Max.   :45.979
##      jitter.local      jitter.local.abs      jitter.rap      jitter.ppq5
## Min.   : 0.1560      Length:1894      Length:1894      Length:1894
## 1st Qu.: 0.8492      Class :character  Class :character  Class :character
## Median : 1.5280      Mode  :character  Mode  :character  Mode  :character
## Mean   : 2.3710
## 3rd Qu.: 3.0100
## Max.   :19.6740
##      shimmer.local      shimmer.local.db      shimmer.apq3      shimmer.apq5
```

```
## Min. : 0.891 Length:1894 Length:1894 Length:1894
## 1st Qu.: 3.456 Class :character Class :character Class :character
## Median : 5.298 Mode :character Mode :character Mode :character
## Mean : 6.645
## 3rd Qu.: 8.218
## Max. :62.405
## shimmer.apq11 mean.autocorr mean.NHR mean.HNR
## Length:1894 Min. :0.6020 Min. :0.0007 Min. : 1.991
## Class :character 1st Qu.:0.8820 1st Qu.:0.0220 1st Qu.:11.755
## Mode :character Median :0.9500 Median :0.0630 Median :16.238
## Mean :0.9232 Mean :0.1144 Mean :16.349
## 3rd Qu.:0.9800 3rd Qu.:0.1727 3rd Qu.:20.741
## Max. :0.9990 Max. :0.7550 Max. :34.233
## F1 F2 F3 F4 gender
## Min. : 201.0 Min. : 514.6 Min. :1860 Min. :2873 0: 473
## 1st Qu.: 380.2 1st Qu.: 907.9 1st Qu.:2599 1st Qu.:3526 1:1421
## Median : 473.0 Median :1584.1 Median :2722 Median :3694
## Mean : 552.9 Mean :1498.7 Mean :2744 Mean :3734
## 3rd Qu.: 768.4 3rd Qu.:1958.7 3rd Qu.:2865 3rd Qu.:3888
## Max. :1192.0 Max. :2805.6 Max. :3458 Max. :4950
## noise single tone phonation creaky noise2
## 0 :631 0:947 A1:287 breathy:288 0:934 78:631
## 78:631 1:947 A2:288 creaky :960 1:960 0 :631
## 90:632 B1:359 modal :646 90:632
## B2:384
## C1:288
## C2:288
```

Classification using SMV (ref <https://medium.com/@ODSC/build-a-multi-class-support-vector-machine-in-r-abcd4b7dab6>)

```
library(e1071)
set.seed(777)
n <- nrow(voice_reports_clean)
ntrain <- round(n*0.75) # 75% for training set
tindex <- sample(n, ntrain)
train <- voice_reports_clean[tindex,c("total.duration", "intensity",
                                       "spectraltilt", "mean.F0", "jitter.local",
                                       "shimmer.local", "mean.HNR", "gender",
                                       "noise", "F1", "F2", "tone")]
test <- voice_reports_clean[-tindex,c("total.duration", "intensity",
                                       "spectraltilt", "mean.F0", "jitter.local",
                                       "shimmer.local", "mean.HNR", "gender",
                                       "noise", "F1", "F2", "tone")]

# Some factors cause any error probably due to not having the same levels between train and test?
svm_model <- svm(tone ~ total.duration + intensity + spectraltilt + mean.F0 + jitter.local
                + shimmer.local + mean.HNR + gender + noise + F1 + F2, data=train,
                method="C-classification", kernal="radial",
                gamma=0.1, cost=10)

summary(svm_model)

##
```

```
## Call:
## svm(formula = tone ~ total.duration + intensity + spectraltilt +
##      mean.F0 + jitter.local + shimmer.local + mean.HNR + gender +
##      noise + F1 + F2, data = train, method = "C-classification", kernal = "radial",
##      gamma = 0.1, cost = 10)
##
##
## Parameters:
##   SVM-Type:  C-classification
##   SVM-Kernel: radial
##         cost: 10
##
## Number of Support Vectors: 1049
##
## ( 152 224 177 123 159 214 )
##
##
## Number of Classes: 6
##
## Levels:
##  A1 A2 B1 B2 C1 C2

prediction <- predict(svm_model, test)
confusion <- table(test$stone, prediction)
confusion
```

```
##      prediction
##      A1 A2 B1 B2 C1 C2
##  A1 55  5  9  0  0  0
##  A2  3 61 11  1  1  0
##  B1  4  9 68  5  6  7
##  B2  2  0  5 71  9 16
##  C1  1  1  6  8 51  2
##  C2  0  0  4  9  2 42
```

Analysis of the F0 contours extracted from MatLab

```
F0test <- read.table("f-1-78/channel1/F0-test", sep = " ")
F0test
```

```
##      V1 V2 V3      V4 V5      V6      V7      V8      V9      V10
## 1 A_single.wav  0  0 330.7472  0 217.4336 217.4336 217.4336 217.4336 219.518
##      V11      V12      V13      V14      V15      V16      V17      V18      V19
## 1 228.7008 231.918 233.4128 235.268 235.268 235.9507 235.9507 235.9507 235.7079
##      V20      V21      V22      V23      V24      V25      V26      V27
## 1 235.7079 235.5496 235.5496 234.9233 234.0742 233.3375 232.6173 232.0362
##      V28      V29      V30      V31      V32      V33      V34      V35
## 1 229.7338 229.4299 229.4299 228.7012 227.6814 227.6814 227.6533 227.6533
##      V36      V37      V38      V39      V40      V41      V42      V43      V44
## 1 227.4748 227.4748 227.4748 225.99 225.99 225.9108 225.9108 225.9108 225.9108
##      V45      V46      V47      V48      V49      V50      V51      V52
## 1 225.9108 225.9108 224.4451 224.4451 224.4451 223.5466 223.313 223.313
##      V53      V54      V55      V56      V57      V58      V59      V60
## 1 223.4968 223.4968 223.4968 223.4968 223.6848 223.6848 222.3647 222.3647
##      V61      V62      V63      V64      V65      V66      V67      V68
```

##	1	222.8896	222.8896	222.2789	222.2789	222.2789	222.2789	222.2789	222.2789
##		V69	V70	V71	V72	V73	V74	V75	V76
##	1	223.5903	224.143	224.143	223.5903	223.5903	222.9754	222.9754	222.9754
##		V77	V78	V79	V80	V81	V82	V83	V84
##	1	222.9754	223.3358	223.7282	223.7282	223.7282	223.7282	223.7282	222.8693
##		V85	V86	V87	V88	V89	V90	V91	V92
##	1	221.9893	221.9893	221.9893	221.7172	221.3941	221.3941	219.2201	219.2201
##		V93	V94	V95	V96	V97	V98	V99	V100
##	1	219.2201	219.2201	219.2201	220.9215	220.9215	221.4581	221.4581	221.6208
##		V101	V102	V103	V104	V105	V106	V107	V108
##	1	223.0844	223.8789	225.949	228.527	231.6212	233.5022	239.6262	239.6262
##		V109	V110	V111	V112	V113	V114	V115	V116
##	1	239.6262	239.4289	234.491	226.3376	222.9532	221.7991	218.3516	217.1759
##		V117	V118	V119	V120	V121	V122	V123	V124
##	1	215.6452	209.9376	208.0533	203.8661	0	0	0	0
									NA