

## Speech Perception – Speech Analysis Project

Map your vowel space. Record tokens of the 15 vowels of English. Using LPC and measurements on the waveform and spectrum, determine F0, F1, F2, F3, and F4 at 3 points in each token plus the voiced duration. Use Praat.

### 1. Recording.

Record 3 tokens of each of the 15 vowels of American English in bVd or hVd context.

These are:

/bid/	/bʌd/	/bud/	/boɪd/
/bɪd/	/bɚd/	/bʊd/	/baɪd/
/bed/		/bod/	/baʊd/
/bɛd/		/bɔd/	
/bæd/		/bad/	

The sampling rate should be in the range of about 10 kHz to 44.1 kHz. They should be recorded in a sentence context: “Yesterday Alice wrote \_\_\_ to me.” Try to say each sentence at the same speaking rate and with the same emphasis.

See the word list at the end for assistance in pronouncing the vowels.

## **2. Analysis**

Generally, it is easier to spot the formants in a wide band spectrogram. For a male speaker, this is a window length of about .004 to .007 sec while for a female it is .002 to .004 sec.

However, don't use the shortest if formants (e.g. F2 and F3 in /i/) appear to merge.

*a. Locations.* The analysis in each token is to be done at three points in the voiced portion. Near the beginning, middle and near the end. Ideally, we want to make our measurements in

the vowel. Due to coarticulation, there is an influence of the initial /b/ and final /d/. Also, since both /b/ and /d/ are voiced, they can not be separated from the vowel. (If you are using hVd, the initial /h/ is voiceless and “easier” to separate from the vowel.)

Determine the beginning and end of the voiced portion (after release and any aspiration of the /b/ to the closure of the /d/). Record this as the voiced duration. The first measurement point is 30 msec after the beginning (release for /b/ or end of aspiration for /h/). The second is half way from beginning to end of the voiced portion. The third is 40 msec before the end (/d/ closure) of the voiced portion.

*b. F0.* The fundamental can be determined either of two ways. Using the waveform, measure the distance (time in sec) from one vocal pulse peak to the corresponding peak in the next pulse. F0 is 1/time.

Using a narrow band spectrum (.040 sec for males, .025 sec for females), measure the distance (frequency) between two adjacent peaks in the spectrum. These adjacent peaks are harmonics. The distance between them is F0.

Note that the software that you are using (e.g. Praat) has a built in means for determining F0 that uses autocorrelation. Use one of the two methods above to cross check it.

*c. F1, F2, F3, & F4.* The formants will be measured using LPC. Praat will set the parameters of this for you. The tracked formants will show up as red dots in the spectrum. Note that this automated formant tracking does make mistakes.

If a value is out of the possible range for one of the formants, check the spectrogram to find the correct value.

There are some formant tracking settings in Praat. For males, look for 5 formants below 5000 Hz. For females, look for 4 formants below 4500 – 5000 Hz. This depends upon vocal tract length: Longer means more formants at lower frequencies.

In some cases, you will not be able to find a formant. If, after using the “tricks”, you can not find a reasonable value, note it as “nm” or “-“ (not measurable).

*d. Tricks.* Occasionally, you will have trouble finding one of the formants at a point in time.

Try moving your analysis window left or right by a few msec or to a neighboring vocal pulse.

Change the length (time) of the analysis window.

Inspect the spectrum for a peak that the LPC is missing.

In spite of your best efforts, this may fail. You may not be able to find a particular formant in a particular token at a particular point in time. The formant (vocal tract resonance) may fall between two harmonics. When two formants are close together, you may not be able to find both.

### **3. Reporting**

For each vowel:

For each token, report the F0, F1, F2, F3, and F4 at each of the 3 locations. Then, determine the average (mean) F0, F1, F2, F3, and F4 for each of the 3 locations across the 3 tokens for each vowel.

Also report the voiced duration for each token and the average voiced duration across the three tokens for each vowel.

## Word list:

- /bɪd/ - bead
- /bɪd/ - bid
- /bed/ - bade (rhymes with maid)
- /bɛd/ - bed
- /bæd/ - bad
- /bʌd/ - bud
- /bɜːd/ - bird
- /buːd/ - boo'd
- /buːd/ - bood (rhymes with wood)
- /bɒd/ - bode (rhymes with road)
- /bɔd/ - baud
- /bɑd/ - bod (rhymes with rod)
- /boɪd/ - boid (rhymes with void)
- /baɪd/ - bide (rhymes with hide)
- /baʊd/ - bowed