6.542J, 24.966J, HST.712J LABORATORY ON THE PHYSIOLOGY, ACOUSTICS, AND PERCEPTION OF SPEECH
Fall 2005

Kenneth N. Stevens Joseph S. Perkell

Stefanie Shattuck-Hufnagel

Lab 11 *10/25/05*

Acoustic correlates of prosodic structure: Part 1

Reference

Klatt, D.H. (1976), Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. **J. Acoust. Soc. Amer. 59**, 1208-1221

Procedure

A: Durations of phrase-final vs. non-phrase-final words and syllables. In the directory called prosodylab1 you will find a set of utterances recorded by four speakers (es, hh, rm and sk). These utterances were designed to highlight the effect of an intonational phrase boundary on the durations of target words. Each utterance is made up of two intonational phrases, and each speaker produced the boundary in two different places: utterances with AB in their filename have the form 'Please say A or B, and C will [verb]', and those with BC have the form 'Please say A, or B and C will [verb]', where the comma indicates an intonational phrase boundary. As a result, each target word A and B occurs in both phrase-final and non-phrase-final position.

Each speaker produced two utterances of each of the following sentences:

> Please say Maine or Duke, and Rice will play. (AB version)
> Please say Maine, or Duke and Rice will play. (BC version)
> Please say Thomas or Kenneth, and Lucas will stay. (AB version)
> Please say Thomas, or Kenneth and Lucas will stay. (BC version)

Speakers SK, RM and ES also produced a reiterant imitation of each sentence, in which they used repetitions of the syllable 'ma' to imitate the original utterance.

Normal speech. For at least two of the four speakers, use xkl to estimate the duration of each segment of the last syllable of the two target words A and B in each utterance, e.g. the /m/, /ei/ and /n/ of *Maine*, etc. (Select speakers ES and SK if you plan to do the optional measures of reiterant utterances,

described below.) Segment durations are estimated by determining the time between relevant acoustic landmarks; give some thought to which acoustic landmarks will provide the most reliable measures, particularly for stop consonants. You can ignore target word C, which was included as a control for variation in speaking rate. For *Thomas*, also measure the overall duration of the first syllable (you do not need to measure the duration of each segment separately for these word-initial syllables).

The questions you are addressing are:

1) Is the final syllable of a word longer when it occurs at the end of an intonational phrase?
2) If so, does this boundary-related lengthening affect the syllable onset as well as the rhyme?
3) Does it affect unstressed final syllables as well as stressed ones (i.e. *–mas* vs. *Maine*)?
4) Does it affect the first syllable of a two-syllable word (i.e. *Tho-*), or only the phrase-final syllable directly before the boundary (*-mas*)?

Optional: Duration at intonation phrase boundaries in reiterant speech

The files for speakers ES and SK include their reiterant imitations after the normal utterance, repeating the metrical and intonational pattern of the utterance using sequences of the syllable *ma*. It is not obvious what kind of a representation is guiding the production of these imitations, and an understanding of what aspects of the original utterance are preserved will help to clarify this point. To determine whether reiterant speech preserves pre-boundary lengthening, compare the duration of the syllable 'ma' in pre-boundary vs. non-pre-boundary position. (For some speakers and for some regions of utterances, the reiterated *ma* syllables may be too difficult to measure. Where is this most likely to occur?)

The question you are addressing is:

Is the pattern of duration lengthening at phrase boundaries that is observed for normal speech also observed in reiterant speech?

B. Placement of phrase-level prominence (pitch accent)

In the subdirectory earlyaccent in prosodylab2, you will find a set of directories that contain utterances from the BU FM Radio News Corpus. Each of the 6 subdirectories contains a reading of the same news paragraph by a different professional radio news broadcaster: f1a, f2b, f3a, m1b, m2b and m3b. All of the speech files except that produced by f3a have been labeled for their prosodic structure (intonational phrase boundaries, marked by 3's and 4's) and phrase-level prosodic prominences (marked by *s), so they can be used to test the hypothesis that the final (i.e. nuclear) pitch accent of a phrase occurs on the main-stress syllable of its word, while a pre-nuclear accent can occur on an earlier full-vowel syllable of its word, especially if it is the first accent in the phrase. The paragraph contains 9 words which are candidates for this kind of early accent, or 'stress shift': *nineteen, seventy-six, Democratic, campaign, de-politicize, nineteen, seventy-one, Massachusetts,* and *association*. For the 5 speakers whose prosody is labeled, determine (for each of the 9 target words) a) whether it is the first pitch-accented word in the intonational phrase, the last pitch-accented word, a phrase-medial accented word, or unaccented, b) if it is accented, whether the pitch accent is on the main-stress syllable, an earlier stressed syllable, or both. Because the phrasing and accentuation may differ for different speakers, you will need to examine each utterance. (Do you agree with the labeling for pitch accents and for intonational phrasing?)

The questions you are addressing are:

Is main-stress pitch accent more likely for an accent that is the last (i.e. nuclear) accent in its intonational phrase?
Is early pitch accent more likely for an accent that is the first accent in its intonational phrase?
Do you find any double-accented words? Why might this occur?

To display the wave forms, pitch tracks and prosodic labels using xwaves, cd to the directory of a speaker (e.g. f1a), and type

xwaves filename.16k

to display the wave form. In the Multidimensional Signal Display window, find the line marked Input File and click (left) just after the filename to activate. Delete the .16k and type in .f0 to display the f0 track. In the same window, click (left) on the middle button in the bottom right corner, which is marked xlabel, to bring up a label window. Click in the upper left corner to activate, and type in the filename (e.g. f1acjrlp2) followed by .breaks to display the break index labels. Delete .breaks and type in .tones to display the tones, and delete .tones and type in .words to display the words.

C. Irregular pitch periods as a marker of intonation phrase boundaries. For either the *Maine-Thomas* utterances or the FM Radio News paragraphs, select two speakers to analyse for the locations of regions of irregular pitch periods. Using xwaves to display both the wave form and the estimated F0 track, determine a) the location of irregular pitch regions, b) for each region, whether it occurs at a phrase onset, at a phrase offset, or phrase-medially and c) whether it is on a pitch-accented syllable, and d) whether it occurs in a region which is generally low in F0. Are there differences between individual speakers? What kinds of problems do you encounter in making these judgments?

The questions you are addressing are:

1) Do regions of irregular pitch periods tend to occur at the edges of intonational phrases?
2) Do irregular pitch periods also tend to occur at the beginning of pitch-accented vowel-initial syllables?
3) Do speakers differ in the distribution of their regions of irregular pitch periods?

Optional: Labelling pitch accents and intonational phrase boundaries. To test your intuitions about the location of phrase-level prominences and constituent boundaries, you can label speaker f3a's paragraph, using a simplified system of * for pitch accent and] for phrase boundary. Follow the instructions above to bring up .tone, .break and .word windows where you can record your labels, but BE SURE TO use a filename that is specific to your labelling group (or to you, if you are labelling solo) so that each set of labels will be saved separately. The .break and .tone label files will be blank when they come up, so you can enter your own labels by positioning the cursor where you want it with the mouse, and then typing in the * or [ ] from the keyboard. Closing xwaves by clicking on the QUIT button in the MSD window will automatically save these label files, which is why it is especially important to select a label filename that is specific to your labelling group.

How does this speaker's behavior compare with the other speakers, for the distribution of Early Pitch Accent and irregular pitch periods? What kinds of challenges did you encounter in labelling the pitch accents (phrase level prominences) and intonational phrase boundaries? Did you gain any insights from this experience?