LING 576 Acoustic Phonetics
Spring 2009
Rod Casali
Topic number 10: Acoustic Investigation of Tone and Intonation
3-24-09

Reading:
Ladefoged, Peter. 2003. Phonetic data analysis. Malden, Massachusetts and Oxford: Blackwell. Read pp. 75-90.
Bart, Joan L. D. 2001. A field manual of acoustic phonetics. MS, SIL. Read Section 5.2.
Himmelmann, Nikolaus P. and D. Robert Ladd. 2008. Prosodic description: An introduction for fieldworkers. Language Documentation & Conservation 2: 244-274. http://scholarspace.manoa.hawaii.edu/handle/10125/4345. Read Section 5.5. (There is a lot of useful material in the rest of the article also.)
Optional but recommended: Connell, Bruce & D. Robert Ladd. 1990. Aspects of pitch realisation in Yoruba. Phonology 7:1-29.

1. Fundamental frequency and pitch

In general, tone is relevant to *voiced* segments.

The primary perceptual correlate of tone is *pitch*. All else being equal, a high tone will have a higher pitch (i.e., will "sound higher") than a low tone, a mid tone will have intermediate pitch.

Listening examples (from Animere, a Niger-Congo language of Ghana):

- Example 1
- Example 2
- Example 3
- Example 4
- Example 5

The linguistic use of pitch in tone languages is relative. There is no absolute and invariant threshold between phonemic tone levels.

It is important not to overstate the case that tonal use of pitch is relative however. Native speakers of tone languages (at least in some cases that have been investigated) are surprisingly successful at identifying the phonemic value (e.g., high, low, mid etc.) of both contour and level tones even on monosyllabic words pronounced in isolation.

"Thai (Siamese), the national language of Thailand, exhibits five contrastive tones, traditionally labeled mid ( ), low (ˆ), falling (ˆ), high (ˊ), and rising (ˇ)...It has been shown that Thai listeners can easily identify the tones in real-speech monosyllabic words, even in isolation from any phonetic or linguistic context."

-- Gandour (1978:42-43).

How is this possible?

For further consideration: How is pitch used linguistically in ways not involving tone?

The primary acoustic correlate of pitch is fundamental frequency (F_0).

The relationship of pitch to F_0 is non-linear. The ear is more sensitive to changes in F_0 at the low end of the audible frequency range than at the high end.

Example: A change from 300 to 400 Hz will sound like a larger pitch increase than a change from 2300 to 2400 Hz.

Perceived pitch may also depend somewhat on the amplitude of a sound. However, such effects may apply mainly to sinusoidal sound waves and not real speech sounds.

Although pitch correlates closely with F_0, harmonics apparently play a very important role in pitch perception. Listeners attend to harmonics and not just F_0, and infer the fundamental at least in part from the harmonic spectrum.

Evidence for this: If the fundamental frequency is filtered out of a sound wave, this does not seriously affect listeners' identification of the pitch, as long as harmonics (and especially certain harmonics in the middle of the spectrum) are present.

Typical F_0 ranges in normal speech (from Nooteboom 1997:642):

- For adult males: 80 to 200 Hz
- For adult females: 180 to 400 Hz

2. Effects of context

F_0 is heavily influenced by context.

Relevant effects:

- F_0 is raised following voiceless obstruents.
- F_0 is lowered following voiced obstruents.

However, while most attention has been given to voiced *obstruents*, probably because their phonological effects on tone are better documented, the lowering effect of voiced obstruents may not actually differ much from that of voiced *sonorants*.

"The available data suggest that pitch following voiced stops is substantially similar to that following sonorants and that it is the pitch following voiceless stops that is perturbed upwards."

-- Ohala (1978:29)

To some extent, these consonantal effects on F_0 are filtered out or corrected in the course of speech perception (Beckman 1986:128-131).

- In African (and other?) tone languages, utterance-final low tones are generally realized as falling.

This fall is quite audible and may serve to enhance the auditory contrast between low and non-low tones.

Nawuri examples:

- [dʒáŋ] perch (verb)
- [dǯi] eat
- There is a general tendency for high tones in sequence to be realized at a progressively lower F_0 over the course of an utterance

This is known as *declination*.

- Higher vowels have intrinsically higher F_0.

This effect may also be compensated for perceptually.

- High tones tend to be higher than normal when they precede a low tone ("local raising of H before L").

See the Connell & Ladd reading article.

Why is an understanding of these contextual effects important to the instrumental investigation and measurement of tone?

3. Investigating tone with speech analysis software

Three ways of measuring F_0:

- From waveform periods.
- From narrow-band spectrograms.
- Using the pitch / F_0 display.

The pitch display is by far the simplest option to use. There may be times however when using one of the other two methods can be useful.

Using pitch displays

Preliminaries:

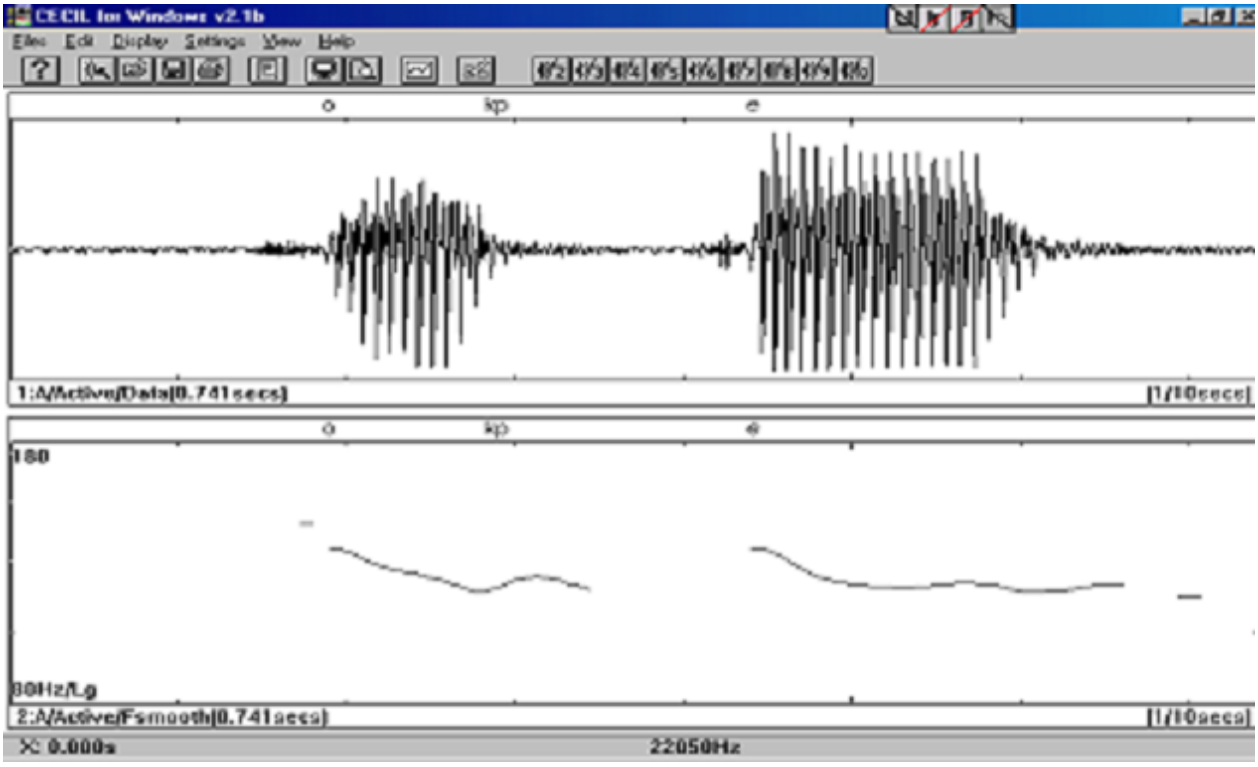
- Choice of display (Speech Analyzer)
- Pitch settings (Speech Analyzer)

Interpretation principles (adapted from lecture notes by Keith Snider):

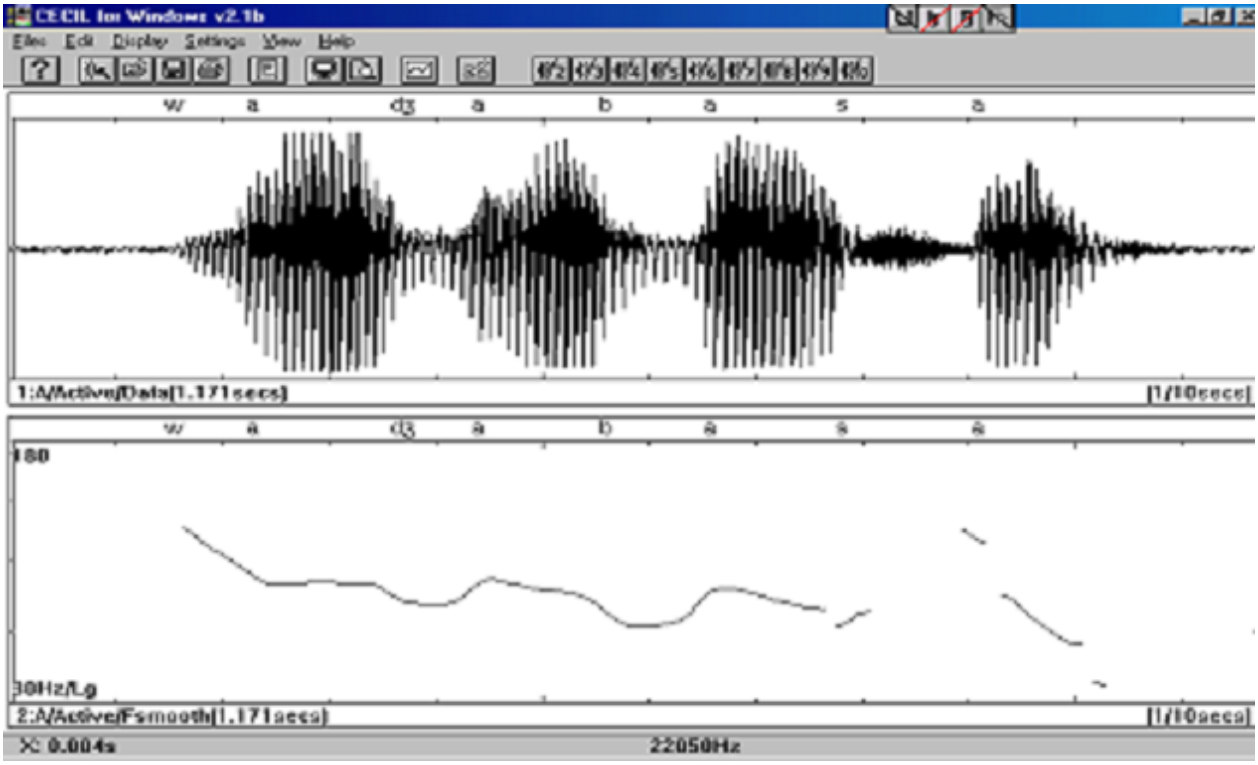
- Ignore irrelevant F_0 effects.

Examples:

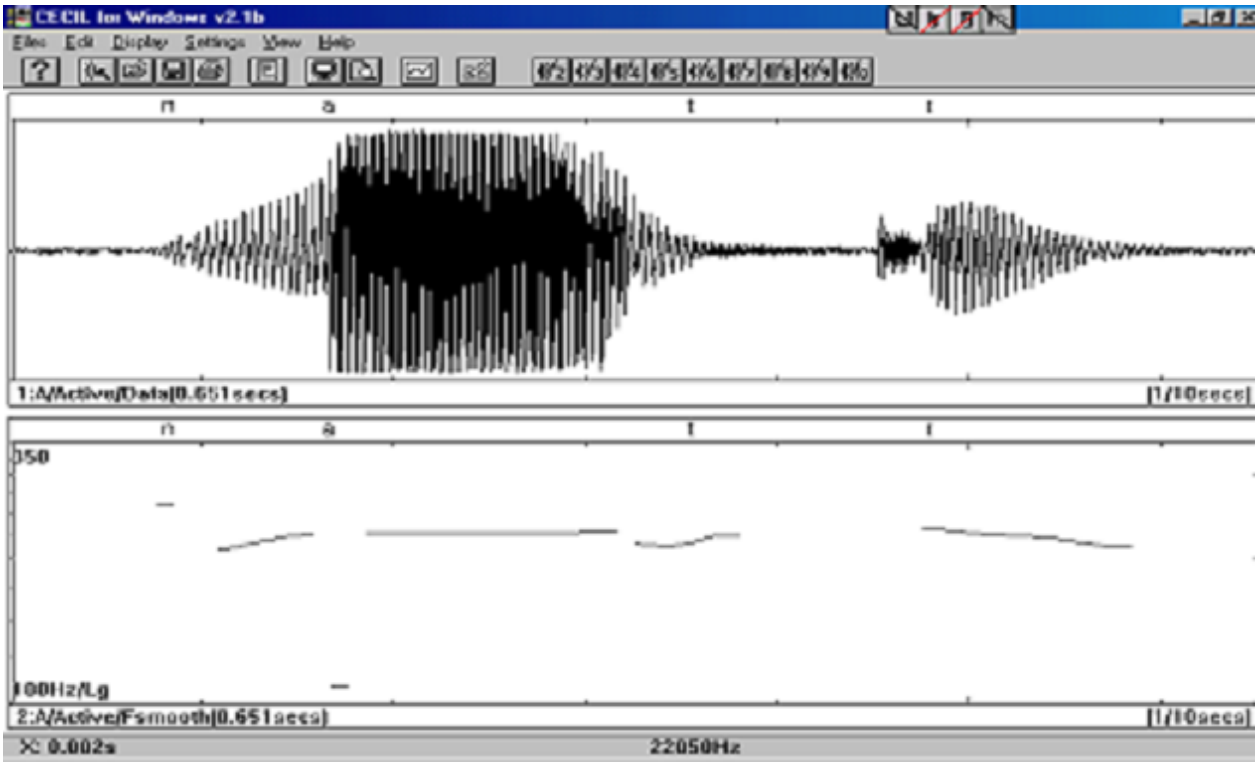
- F_0 during voiceless consonants.



- F_0 dip during voiced consonants.



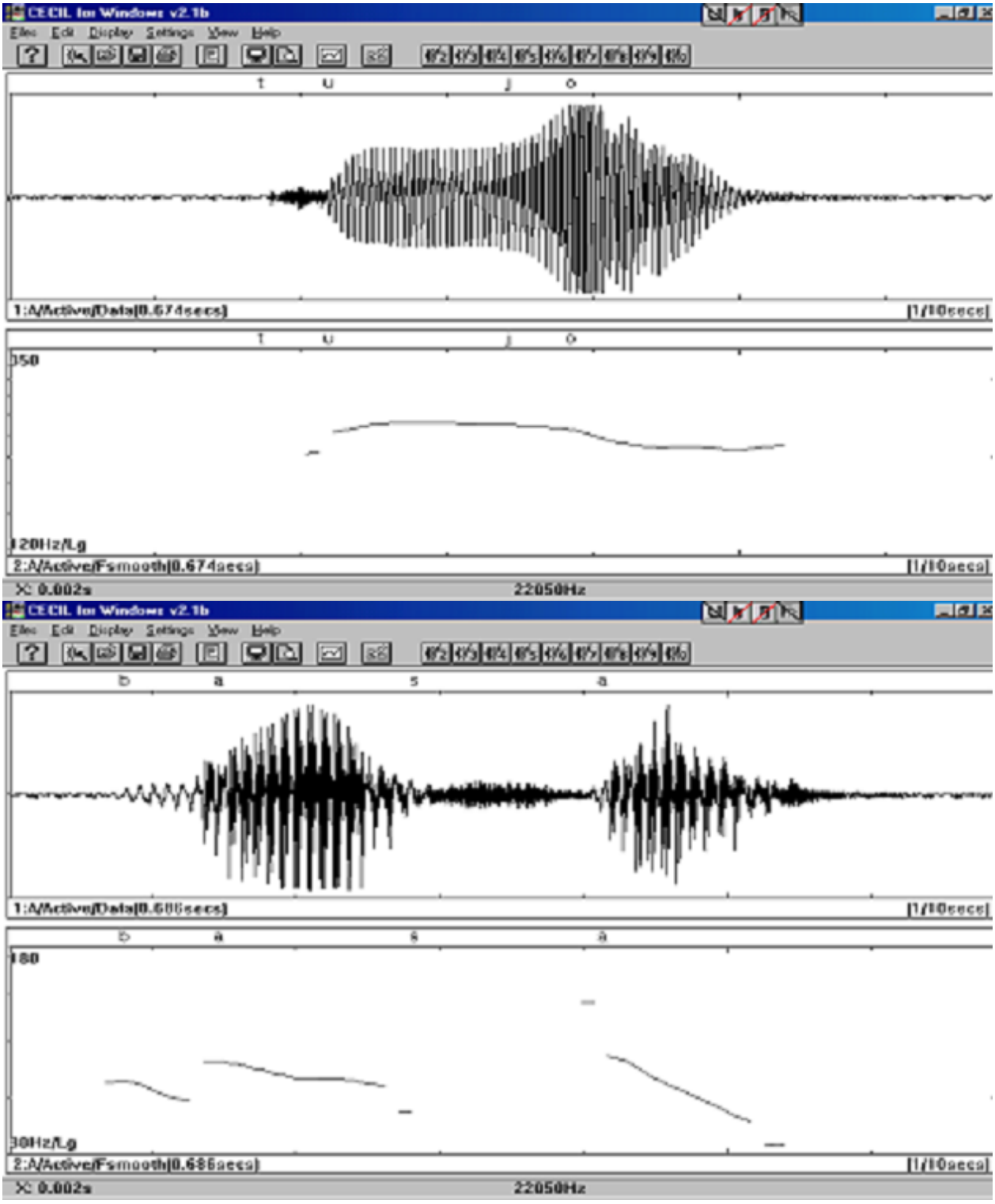
- F_0 changes during *onset* consonants.



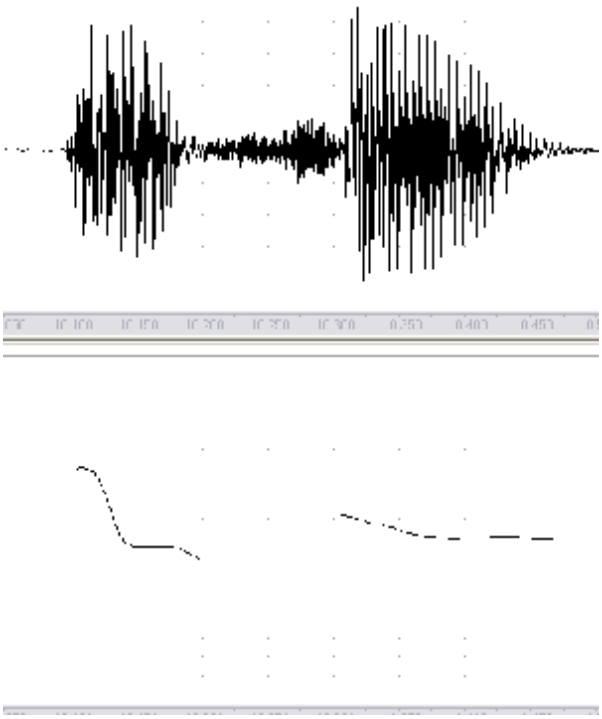
- Minor utterance-initial rises or falls.

"An utterance that begins with Hi tone and has an onset (either voiced or voiceless) will typically begin with a short rise up to the level of the unperturbed tone...Similarly, an utterance that begins with Lo tone and has an onset (either voiced or voiceless) will typically begin with a short fall down to the level of the unperturbed tone."

--Keith Snider, lecture notes.



- F_0 values that cannot possibly fall within your speaker's pitch range.
- Any other F_0 behavior that common sense, your knowledge of tonal behavior, or your ears tell you cannot possibly be correct.



Sound file: OfE.wav

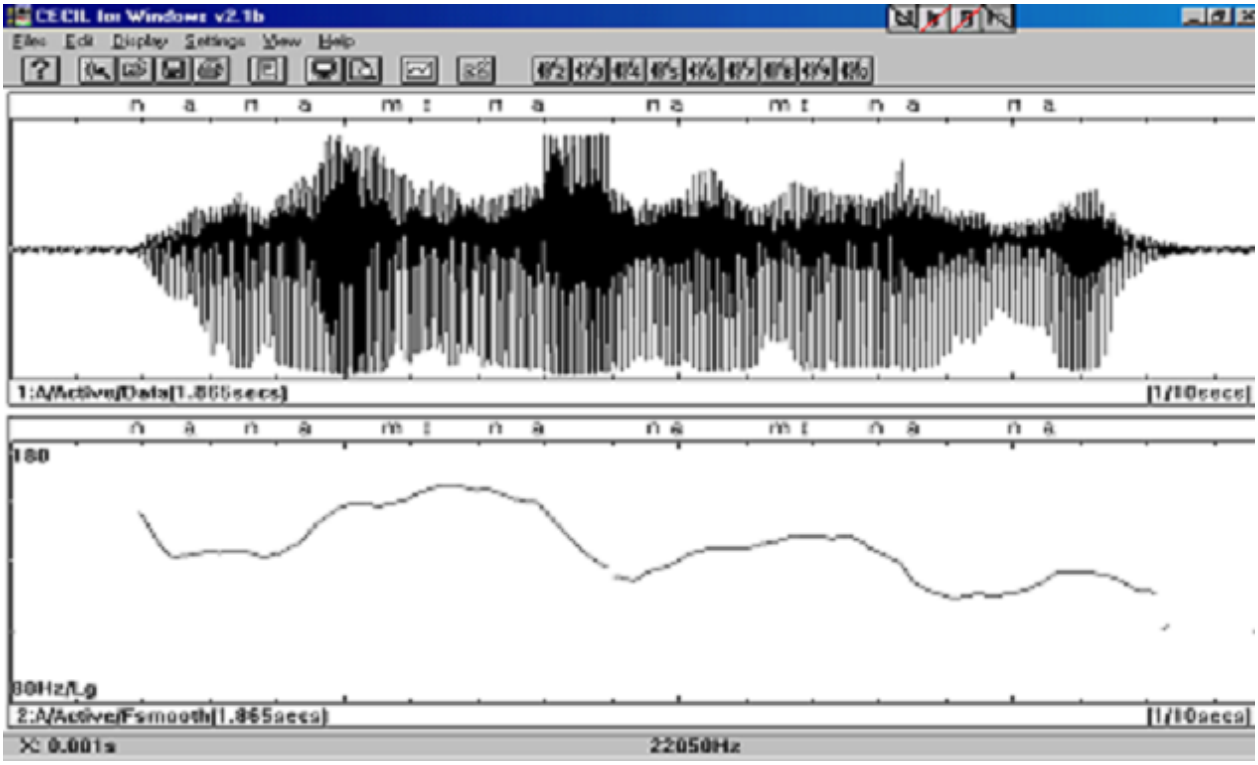
- Take into account the effects of context (cf. Section 3 above).
- Be skeptical of interpreting F_0 contours as rising or falling tones if they are not audible (especially if the magnitude of the change is small).
- Seek to identify the *target* F_0 of a syllable.

Principles for determining target F_0:

- A general tip: Consult the spectrogram display for help in identifying segment boundaries (e.g. between nasal consonants and vowels) where this is difficult in the waveform.

Example: Animere [wo] "you"

- If the F_0 is level over most/all of the syllable, take the value at this level portion as the F_0 of the syllable.
- Often, there will be a "shoulder" or "shelf" in the F_0 curve of a syllable, which will be preceded and/or followed by a rise or fall (e.g., due to perturbing effects of consonants). In this case, take the F_0 of the shoulder / shelf as the F_0 value for the syllable.



- Where there is no shoulder and the F_0 contour changes constantly throughout the syllable, use either of the following strategies:

- Measure F_0 at the right edge of the voiced portion of the syllable, ignoring obvious consonant perturbations (Keith Snider, lecture notes).
- Measure F_0 at the syllable's energy peak (Connel & Ladd 1990).

4. Alternatives to instrumental measurement of F0

As useful as instrumental investigation of tone can be, it is important not to neglect other aids and techniques that can help with tone analysis. Use of speech analysis software is not a panacea.

Other sources of insight:

- Training one's ears.
- Comparative listening.
- Whistling (by native speakers).

It's possible to look at a whistled tone pattern also.

This is generally not possible with the F_0 display in Speech Analyzer, but it is possible using a spectrogram. The medium band setting works well for this purpose.

- Native speaker intuition (Kutsch Lojenga 1996).

References

Beckman, Mary E. 1986. Stress and non-stress accent. Dordrecht: Foris Publications.

Connell, Bruce & D. Robert Ladd. 1990. Aspects of pitch realisation in Yoruba. Phonology 7: 1-29.

Gandour, Jackson T. 1978. The perception of tone. In Tone: A Linguistic Survey, ed. by Victoria A. Fromkin, 41-76. New York: Academic Press.

Kutsch Lojenga, Constance. 1996. Participatory research in linguistics. Notes on Linguistics 72:13-27.

Nooteboom, Sieb. 1997. Prosody of speech: Melody and rhythm. The Handbook of the Phonetic Sciences, ed. by. William J. Hardcastle and John Laver. Oxford: Blackwell.

Ohala, John J. 1978. Production of tone. In Tone: A Linguistic Survey, ed. by Victoria A. Fromkin, 5-39. New York: Academic Press.