24.910 Topics in Linguistic Theory: Laboratory Phonology
Spring 2007

24.910
Laboratory Phonology
Intonation

Readings:

- ToBI tutorial, 2.2-2.5, 2.8

Assignments:

- Write a short description of your project.
- ToBI transcription exercises (to be posted).

The Phonology of English Intonation

A very brief introduction to English intonation primarily based on Pierrehumbert (1980) (P80) and Beckman and Pierrehumbert (1986) (BP86).

Pierrehumbert, Janet (1980) *The Phonology and Phonetics of English Intonation*. PhD dissertation, MIT.

Beckman, Mary and Janet Pierrehumbert (1986) Intonational structure in Japanese and English. *Phonology Yearbook* 3.

The phonology of intonation

- Three components relevant to the theory of intonation:
  - Intonation (pitch contour)
  - Stress
  - Phrasing

- Collectively can be referred to as prosody.
  - But sometimes ‘intonation’ is used to cover all of these components.
- These components are closely interrelated.

The representation of intonation

L* H* L-L%

He knew the millionaire

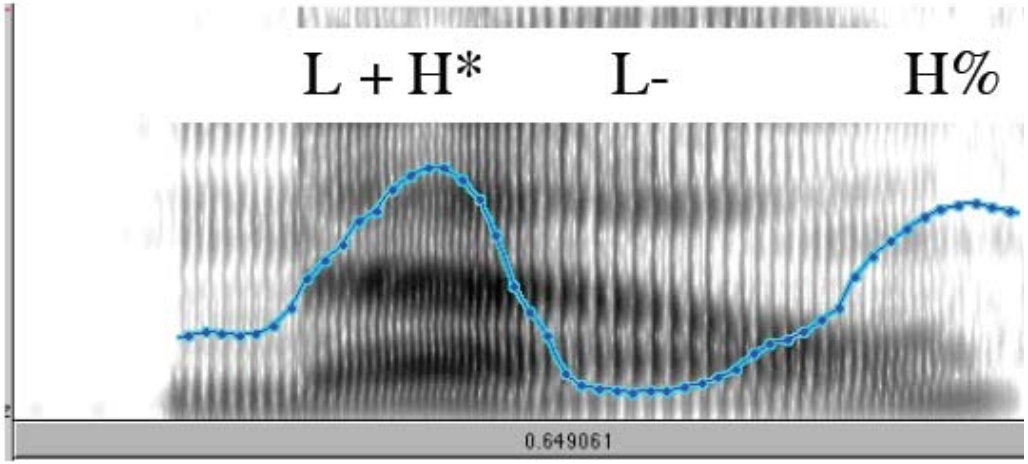
- A string of tones, H(igh), L(ow).
- Three kinds of tones are differentiated, for purposes of tune-text alignment and phonetic interpretation:
  - Pitch accents T* gravitate to stressed syllables
  - Phrase accents T- or T gravitate to (smaller) phrase boundaries
  - Boundary tone T% gravitate to (larger) phrase boundaries

The representation of intonation

Motivation for analyzing intonation contour in terms of a sequence of tones.

- When similar tunes are associated with texts of differing lengths, the consistent properties of the melody are alignments between landmarks in the f0 contour and stressed syllables and the edges of phrases.

A: I hear Nell’s taking a course to be a driving instructor.

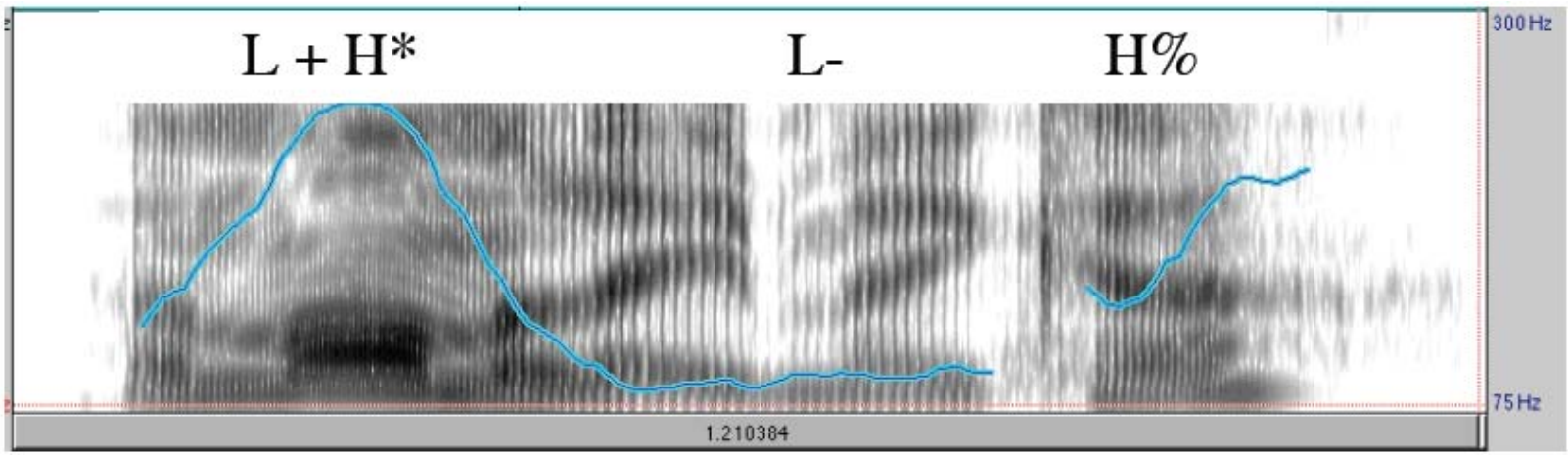


N e ll

Audio: 7_001.wav

- L+H* associates to the main stress,
- L-H% rise occurs at the end of the phrase,
- any interval between is filled by L-.

A: I hear Nell’s got a job as a marmalade maker.



a m ar m a l a de m a k er

Audio: 7_002.wav

Inventory of pitch accents

P80 7 pitch accents	PB86 6 pitch accents	ToBI 5 pitch accents + downstep (!)
H*	H*	H*
L*	L*	L*
H+L*	H+L*	H+!H*
H*+L	H*+L	(H* followed by downstep)
L*+H	L*+H	L*+H
L+H*	L+H*	L+H*
H*+H		

Phrase accents: H-, L-

Boundary tones: H%, L%

Intonation

- Tones are realized primarily in terms of f0 (fundamental frequency).
- Pitch accented syllables are also generally louder and longer than unaccented syllables.

Stress

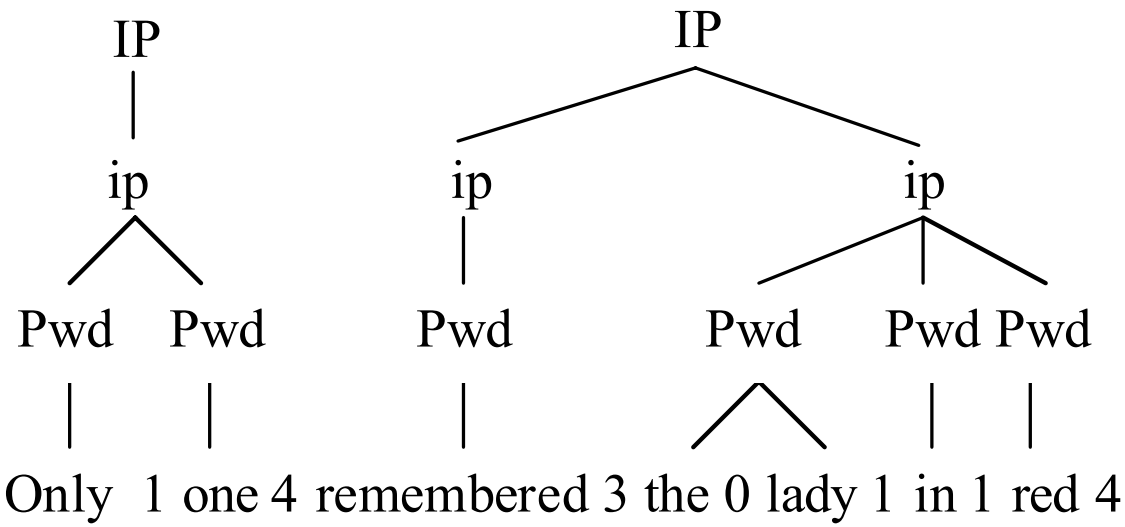
- Stress: ‘relative prominence’ of syllables (Liberman & Prince 1977).
- Lexical stress:

  ìn.to.ná.tion pró.ba.bly

  - Words have a stress pattern.
  - One syllable (primary stress) is more prominent than the rest.
- Precise phonetic correlates are complex.
  - Usually: loudness (duration, intensity), voice quality.
- Most important here: If a word bears a single pitch accent, it is aligned to the primary stressed syllable.
  - Completely unstressed syllables usually cannot bear pitch accents.
- There are also prominence relations between words.
  - Generally correlates with pitch accent placement: the syllables with the strongest stresses have pitch accents.

Phrasing

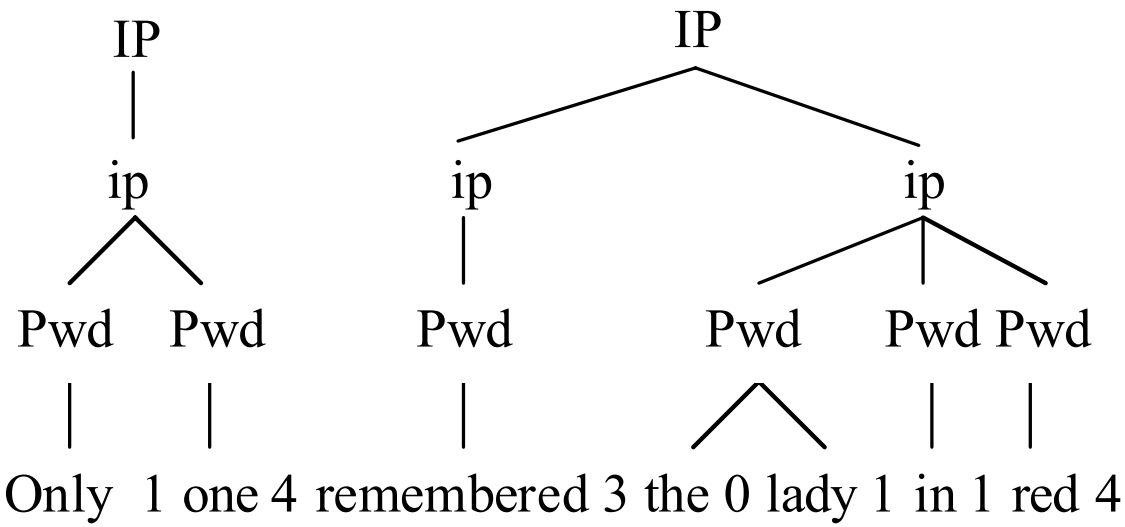
- Utterances are divided into intonational phrases, which are subdivided in intermediate phrases.
- Hierarchical bracketing of the phonological string, usually assumed to be related to, but distinct from, syntax.



Phrasing

Unlike syntactic structure, prosodic structure is argued to be:

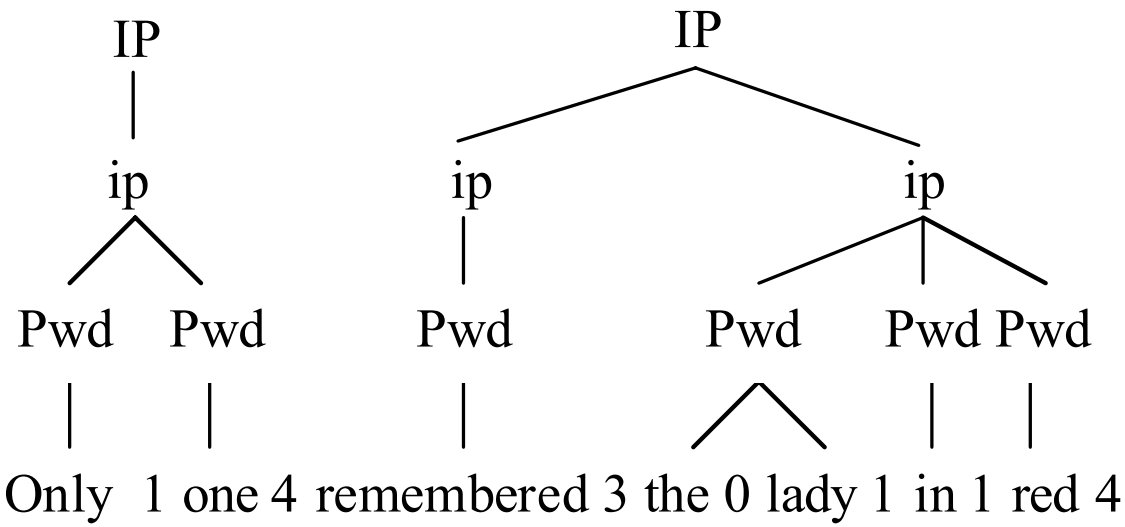
- Exhaustive: a string is fully parsed into constituents of each type.
- Strictly layered: There is a hierarchy of constituents such that each type of constituent only dominates the next type of constituent on the hierarchy



Phrasing

This is a very impoverished form of constituent structure that can be fully specified by placing a symbol between each pair of words indicating the strongest boundary between those words.

- This is how the break index notation represents prosodic structure.



Why phonological constituents?

Why is a phonological constituent structure needed in addition to syntactic structure?

- Syntax-prosody mismatches

This is [the cat [that caught [the rat [that stole [the cheese]]]]]

[This is the cat][that caught the rat][that stole the cheese]

[Sesame St. is brought to you]$_{IP}$[by the children's television workshop]$_{IP}$

[Sesame St. is brought to you by]$_{IP}$[the children's television workshop]$_{IP}$

Phonetic correlates of phrasing

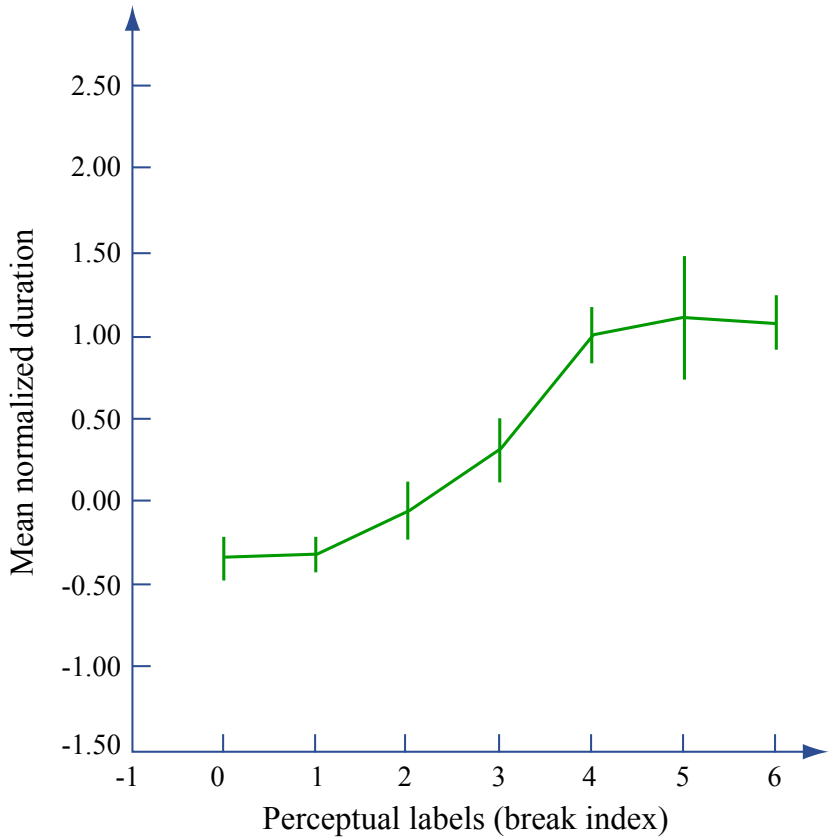
- In many languages, segments are lengthened before constituent boundaries.
- Pauses typically occur at intonational phrase boundaries.

Wightman et al (1992)

- Studied a corpus of sentences read by professional news announcers.
- Prosodic boundaries marked using a 7-point scale of break indices
    - 0 No prosodic break (cliticization)
    - 1 Prosodic word boundary
    - 2 'accentual phrase'
    - 3 Intermediate phrase
    - 4 Intonation phrase
    - 5 Superior major tone group
    - 6 Sentence boundary
- Lengthening measured in terms of normalized segment durations (standard deviations from the mean for that segment, adjusted for estimated speech rate of each sentence).

Phonetic correlates of phrasing

- Lengthening before constituent boundaries, localized to the final syllable rhyme (VC).
- Degree of final lengthening increases with size of prosodic constituent.



Mean normalized duration of the word-final vowel as a function of the break index following the word. The vertical bars correspond to confidence intervals (95% protection level).

Figure by MIT OpenCourseWare. Adapted from Wightman, Colin W., Stefanie Shattuck-Hufnagel, Mari Ostendorf, and Patti J. Price. "Segmental Durations in The Vicinity of Prosodic Phrase Boundaries." *Journal of the Acoustical Society of America* 91 (1992): 1707-1717.

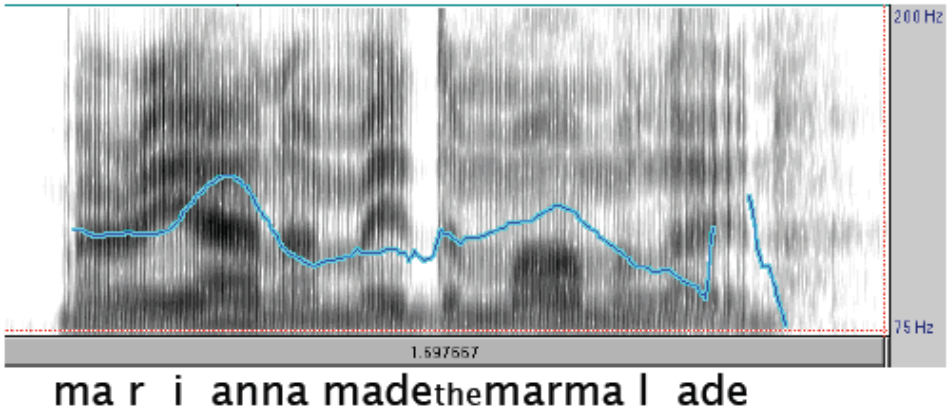
The relationship between phrasing and intonation

ToBI:

- Intermediate phrase must contain one or more pitch accents and a final phrase accent (associated to the ip boundary)
- Intonation phrase may begin, and must end, with a boundary tone (associated to the IP boundary) (as in P80)

E.g. H*L- H*L- L%

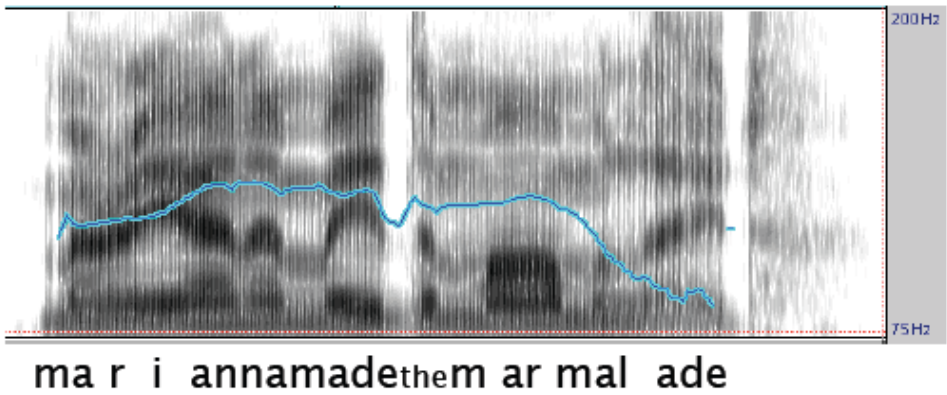
[[‘I’]$_{ip}$[means ‘insert’]$_{ip}$]$_{IP}$



H* L- H* L-L%

[[Marianna][made the marmalade]]

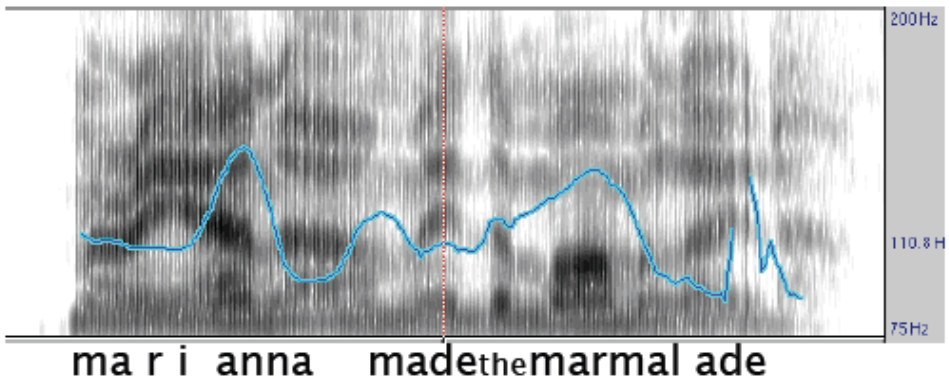
Audio:
7_003.wav



H* H* L-L%

[[Marianna made the marmalade]]

Audio:
7_004.wav



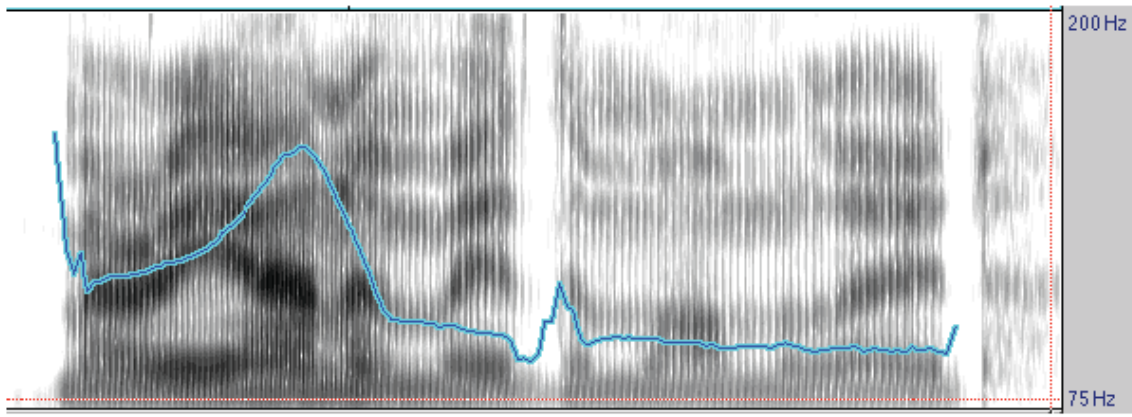
L+H*L-H% H* L-L%

[[Marianna]][made the marmalade]]

Audio:
7_005.wav

The relationship between phrasing and intonation

- Nuclear accent = last accent in an intermediate phrase.
- The phrase tone spreads over the interval between the nuclear accent and the end of the phrase.



Audio:
7_006.wav

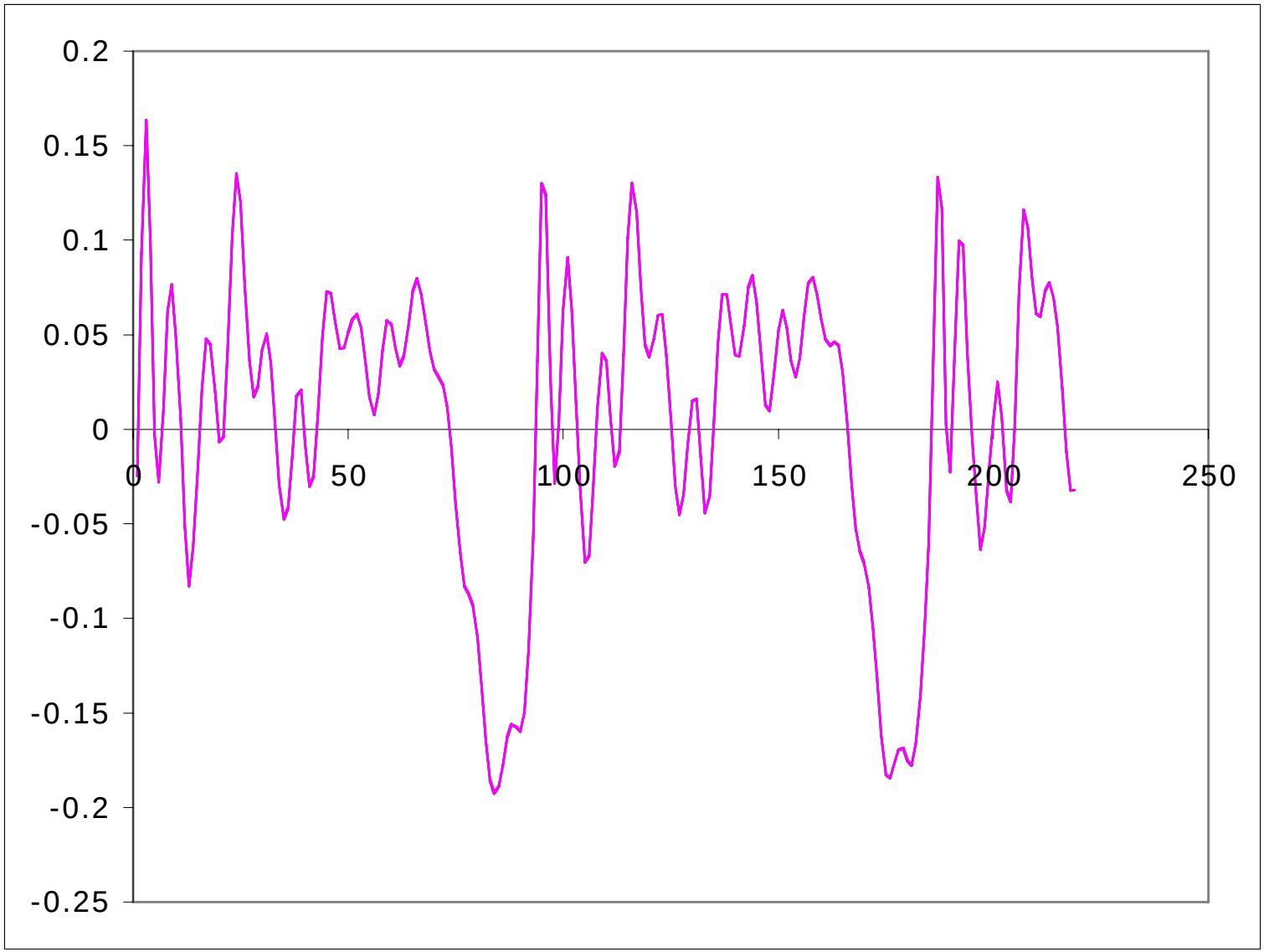
L+H* L-L%

[[Marianna made the marmalade]]

Pitch Tracking

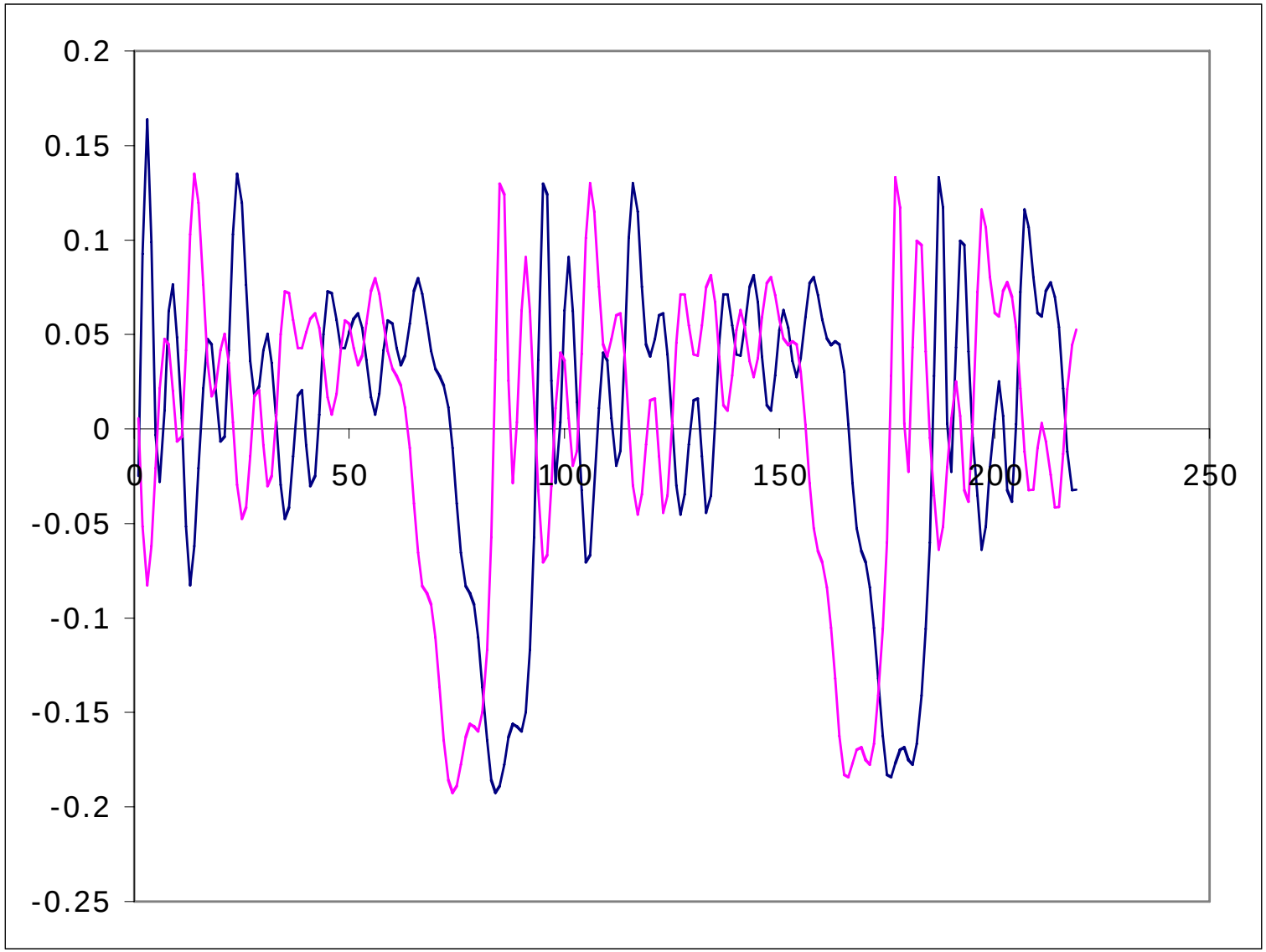
- ToBI transcription is performed with reference to an f0 contour, so it is useful to have some idea of how f0 contours are calculated and interpreted, and what can go wrong at both stages.
- Pitch tracking algorithms:
  - Tentative identification of f0 at regular intervals, usually through picking peaks in an autocorrelation function.
  - Additional processing to select ‘best’ contour, e.g. dynamic programming, subject to smoothness constraints.

autocorrelation



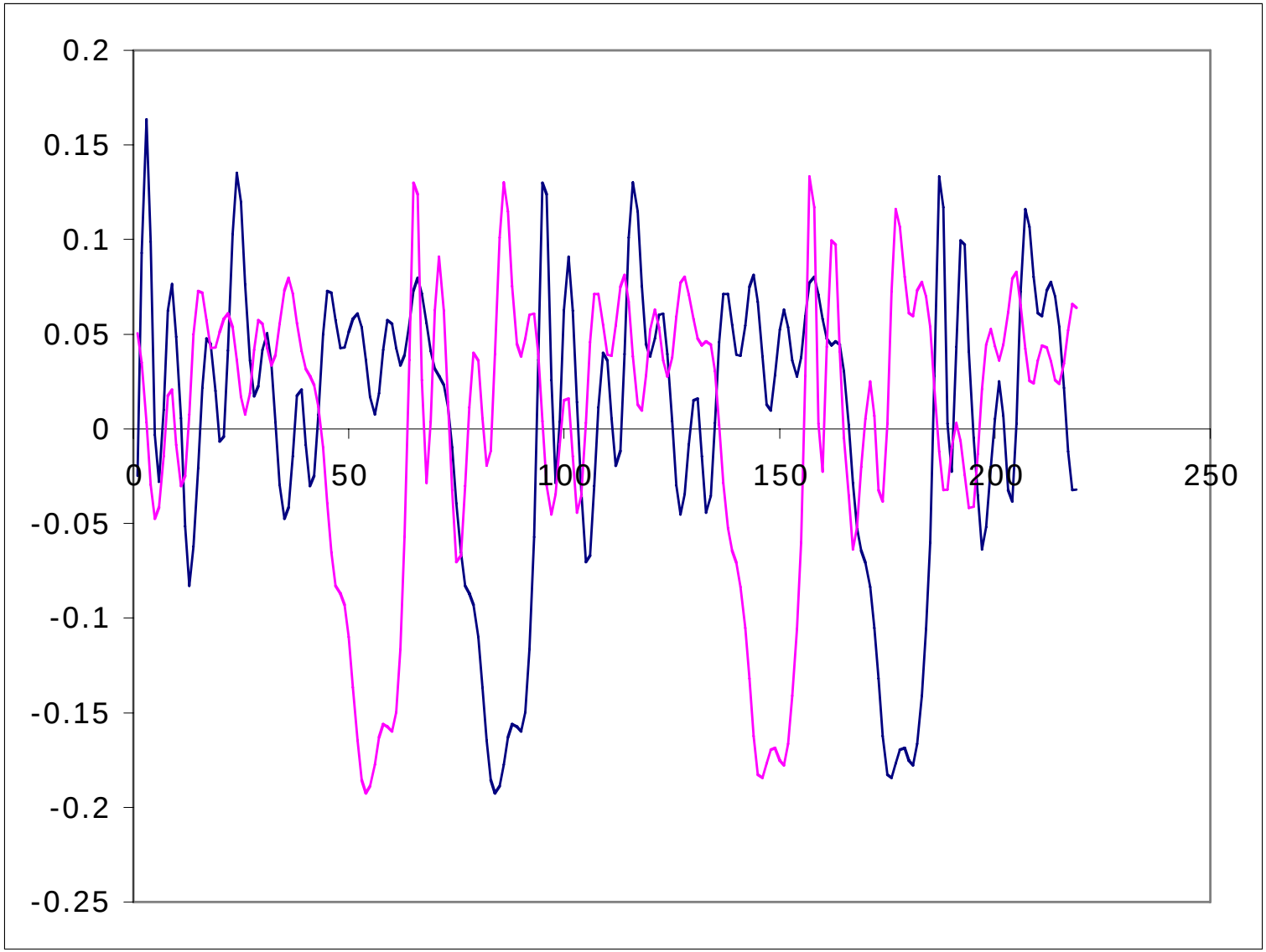
Lag = 0

autocorrelation



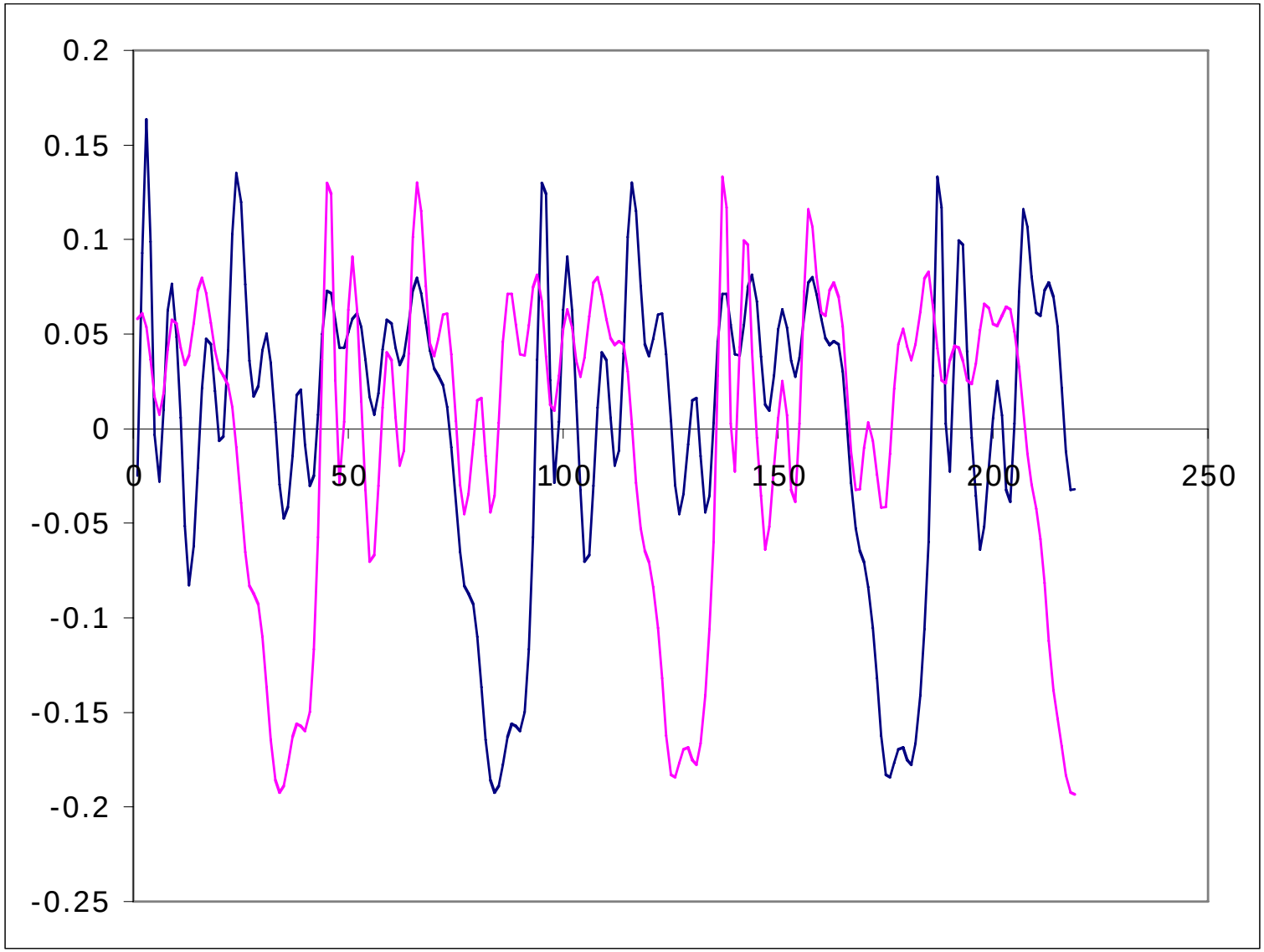
Lag = 10 samples

autocorrelation



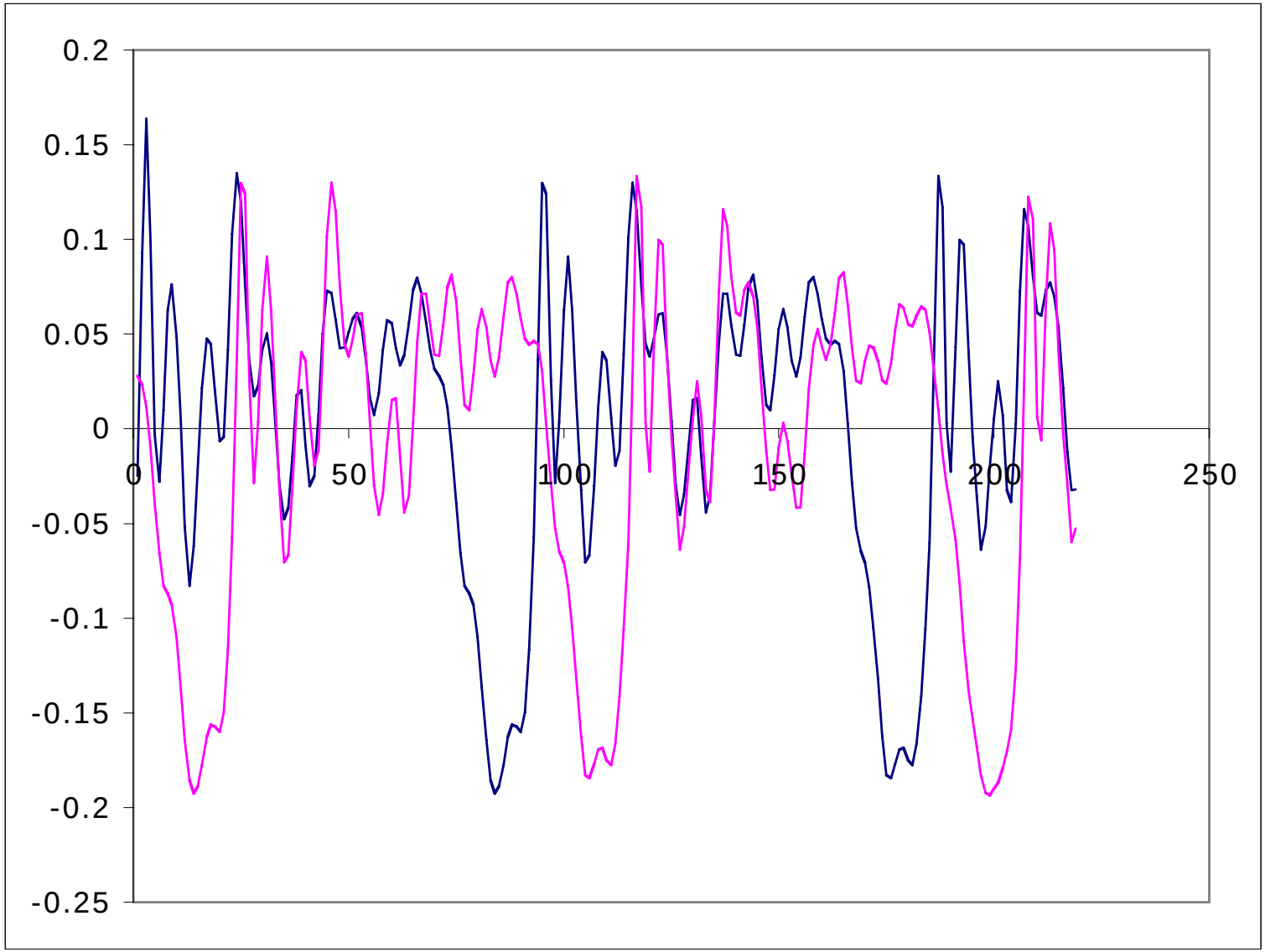
Lag = 30

autocorrelation



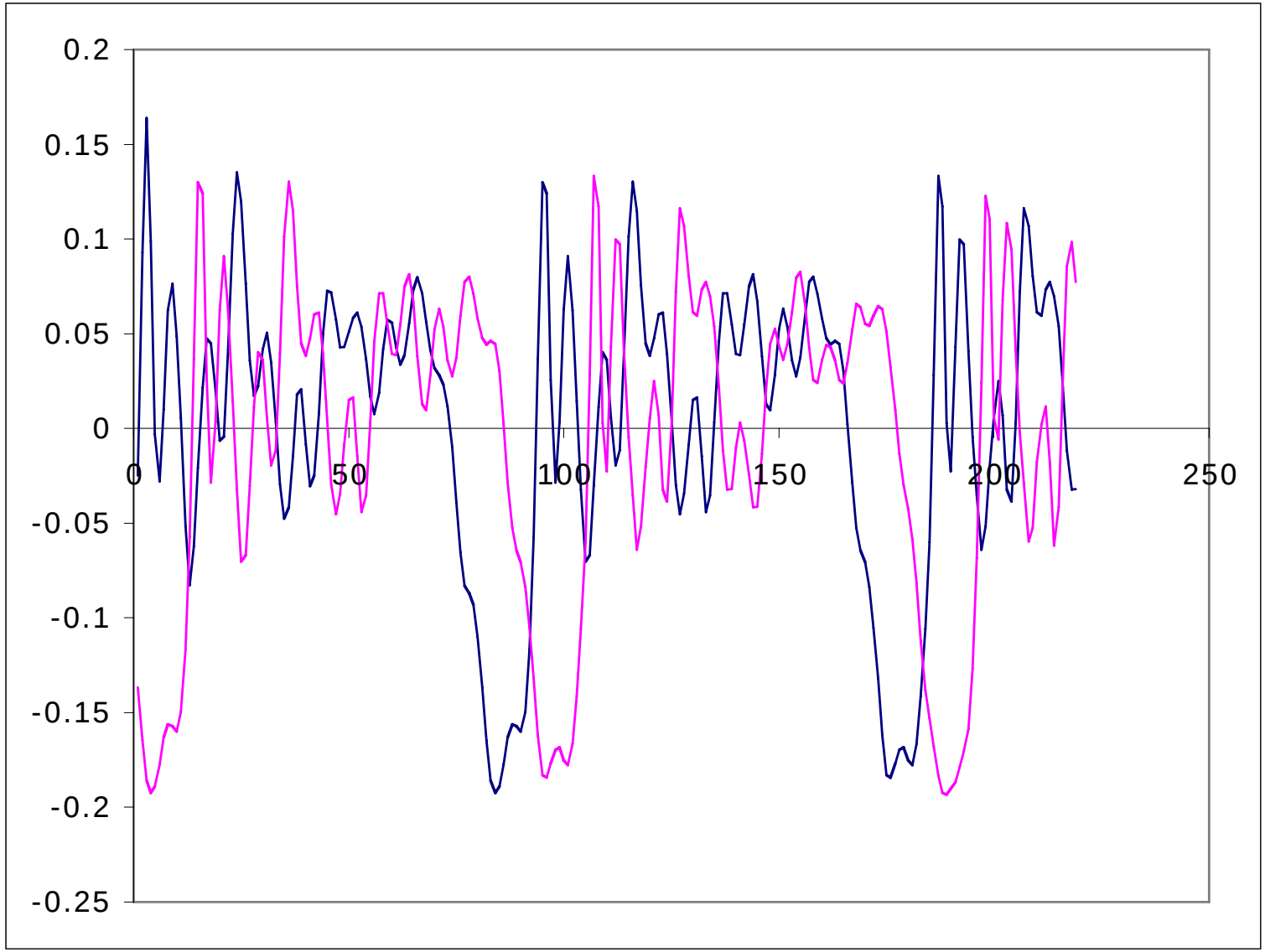
Lag = 50

autocorrelation



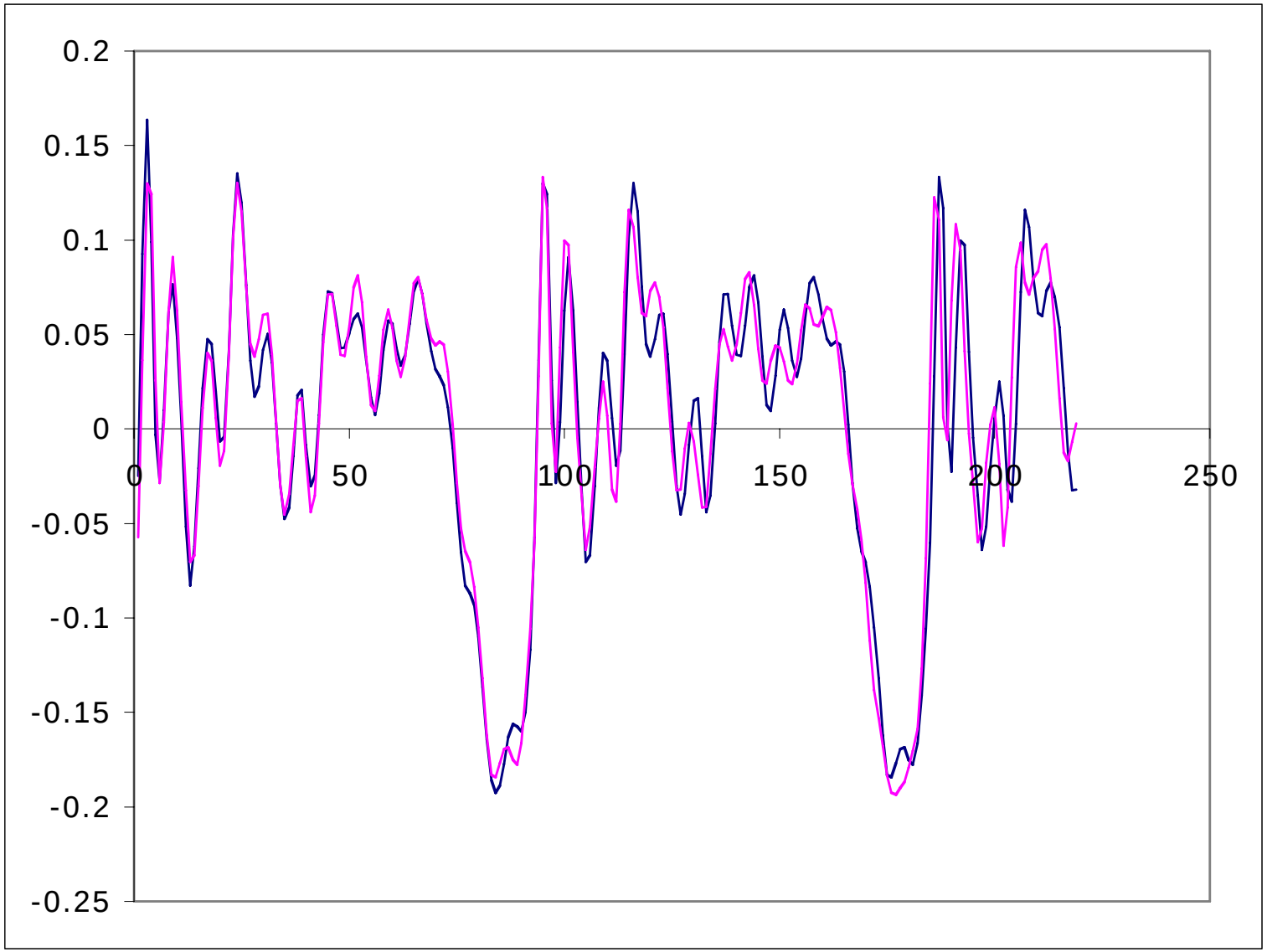
Lag = 70

autocorrelation



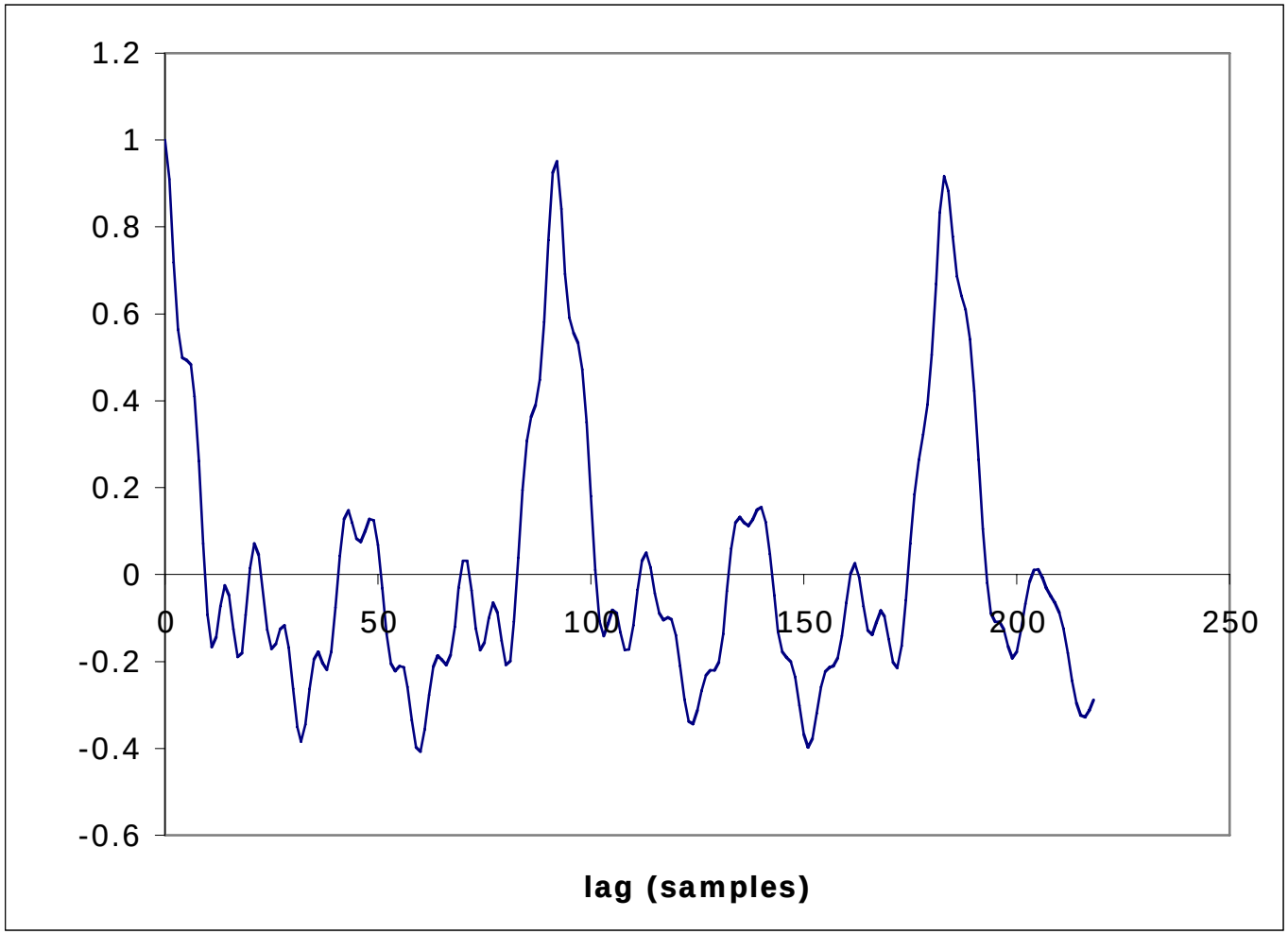
Lag = 80

autocorrelation

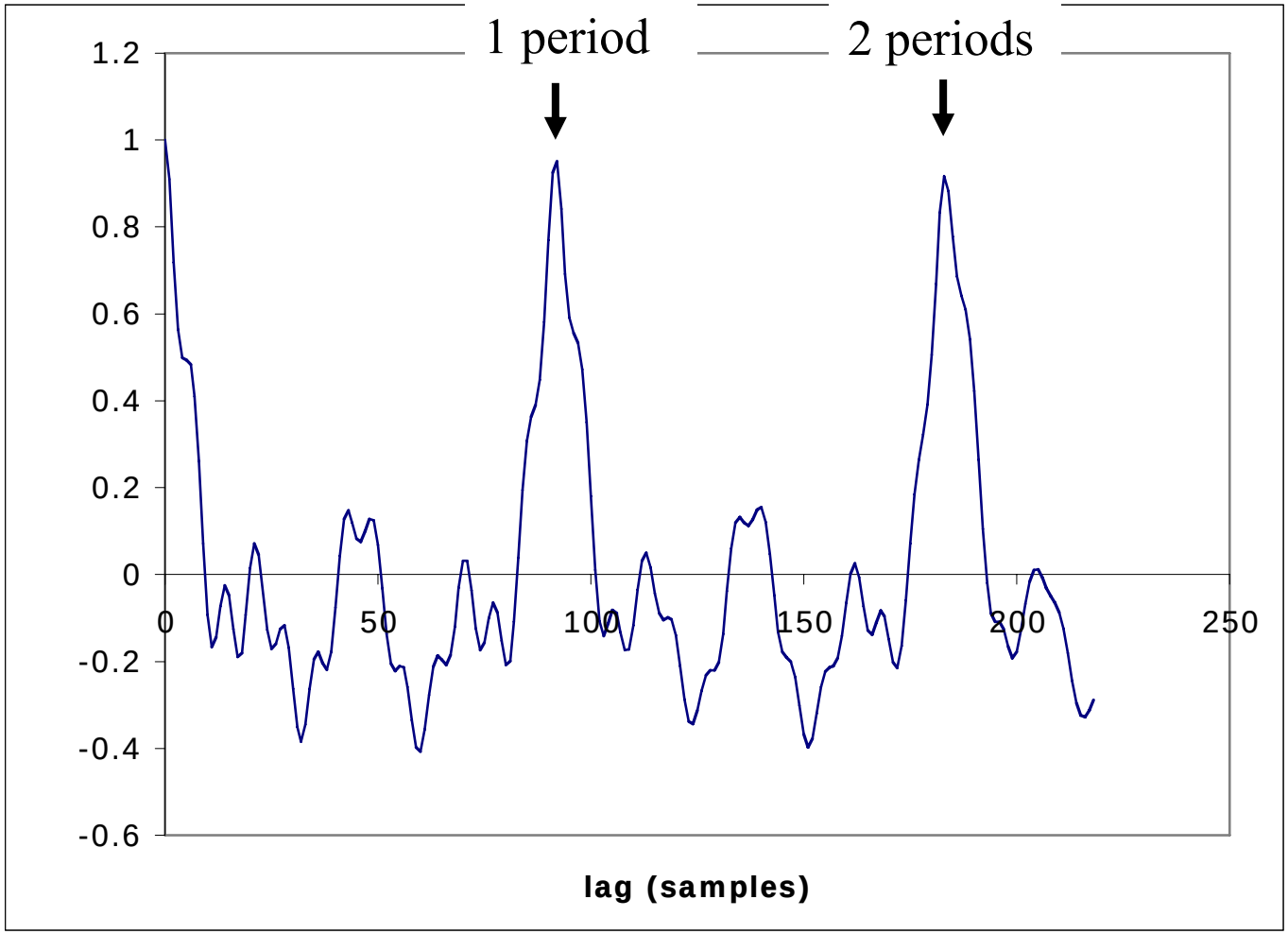


Lag = 92 samples

Autocorrelation function



Autocorrelation function



Pitch Tracking

Common errors:

- Pitch halving: two periods treated as one.
- Pitch doubling: one period is treated as two.
- Failure to detect periodicity.

Interpreting f0 Tracks

- No f0 during voiceless sounds.
- Consonants can ‘perturb’ f0.
  - f0 is usually raised after obstruents, more after voiceless obstruent.
  - f0 can fall before obstruents.
- Creaky voice - often involves very long pitch periods, and fluctuating period lengths.
  - Pitch trackers often fail to detect periodicity.
  - Creaky voice is common at the ends of phrases.

Pitch Tracking in PRAAT

Parameters:

- Pitch halving: two periods treated as one.
- Pitch doubling: one period is treated as two.
- Failure to detect periodicity.