

The Music of Speech

Melody

Dafydd Gibbon

Mannheim Summer School, June-July 2019

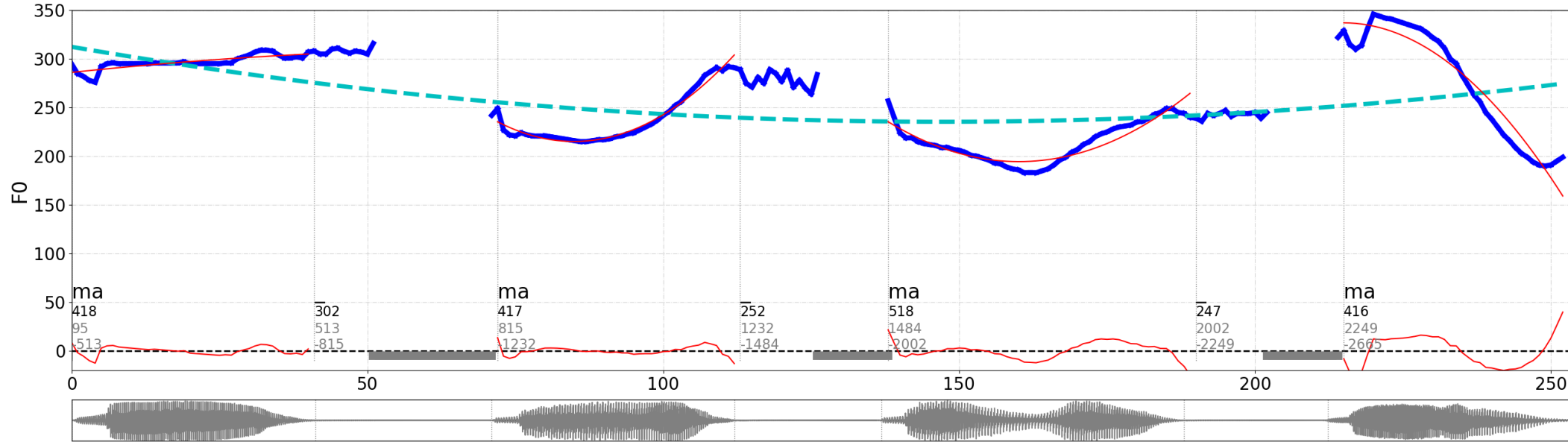
Speech Melody – Tone, accent and intonation

- Local components (the melody of words):
 - English intonation
 - Grammatical stress → phonetic accentuation
 - Lexical stress: pitch accent
 - Phrasal accentuation: phrase accent
 - Discourse focus and emphasis
 - Sino-Tibetan languages:
 - lexical tone
 - accentuation
 - Intonation
 - Niger-Congo languages:
 - Lexical tone
 - Morphological (grammatical) tone
 - Intonation

Lexical tone

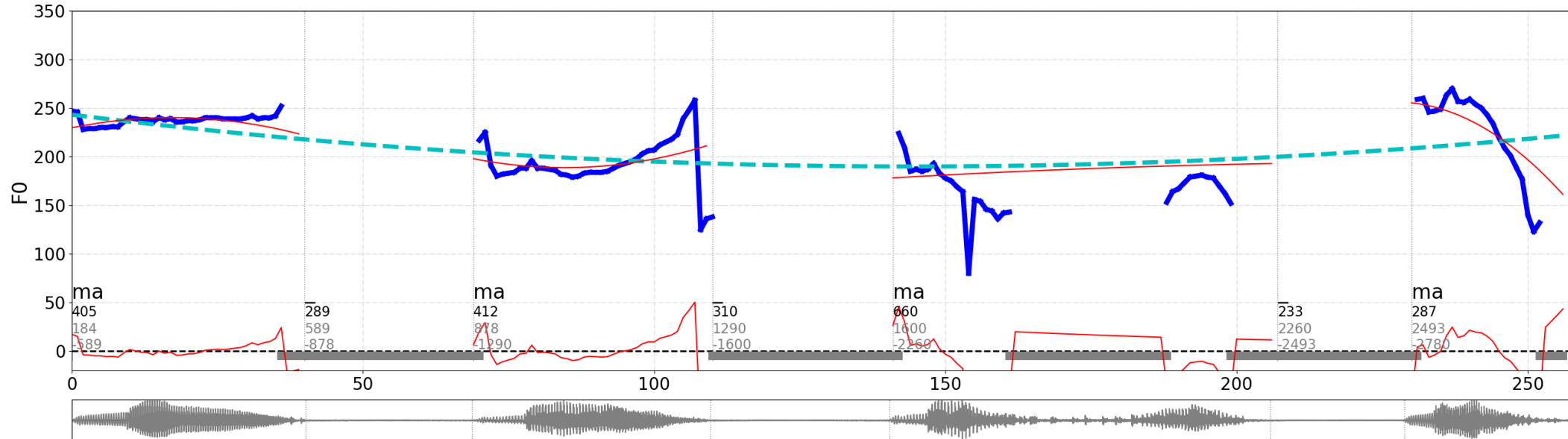
Four lexical tones in Putonghua (Beijing Chinese)

PV 01: "Mandarin_ChinesePods...", tier "Syllables", x-axis 10.0ms, Model: median 1, degree 2, domain "majorIPU"



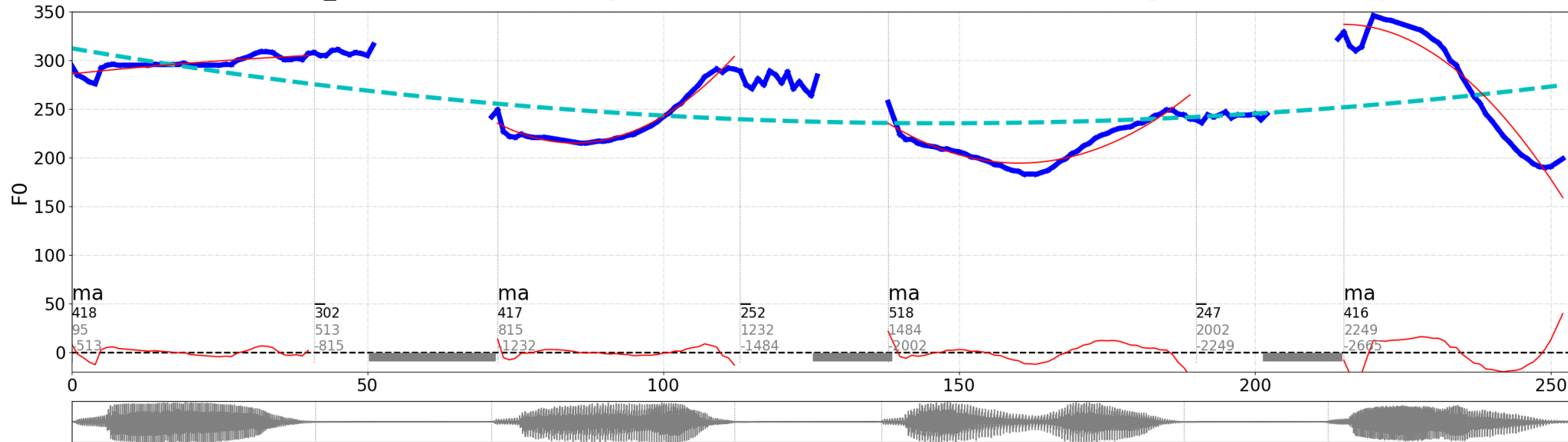
Four lexical tones in Putonghua (Beijing Chinese)

PV 01: "Mandarin_creaky4ma", tier "Syllables", x-axis 10.0ms, Model: median 1, degree 2, domain "majorIPU"

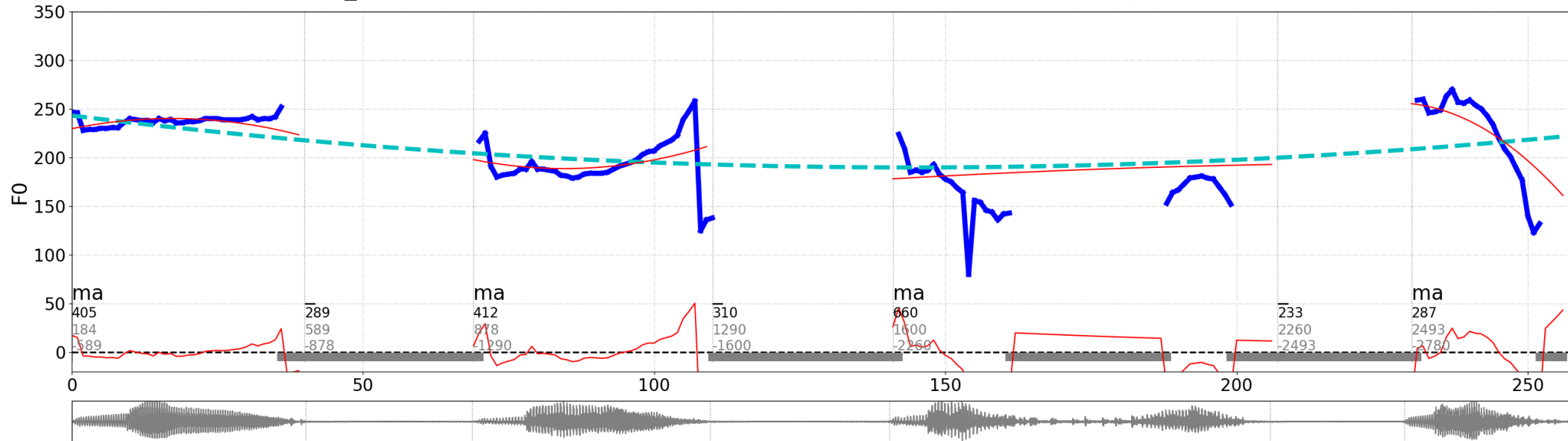


Four lexical tones in Putonghua (Beijing Chinese)

PV 01: "Mandarin_ChinesePods...", tier "Syllables", x-axis 10.0ms, Model: median 1, degree 2, domain "majorIPU"

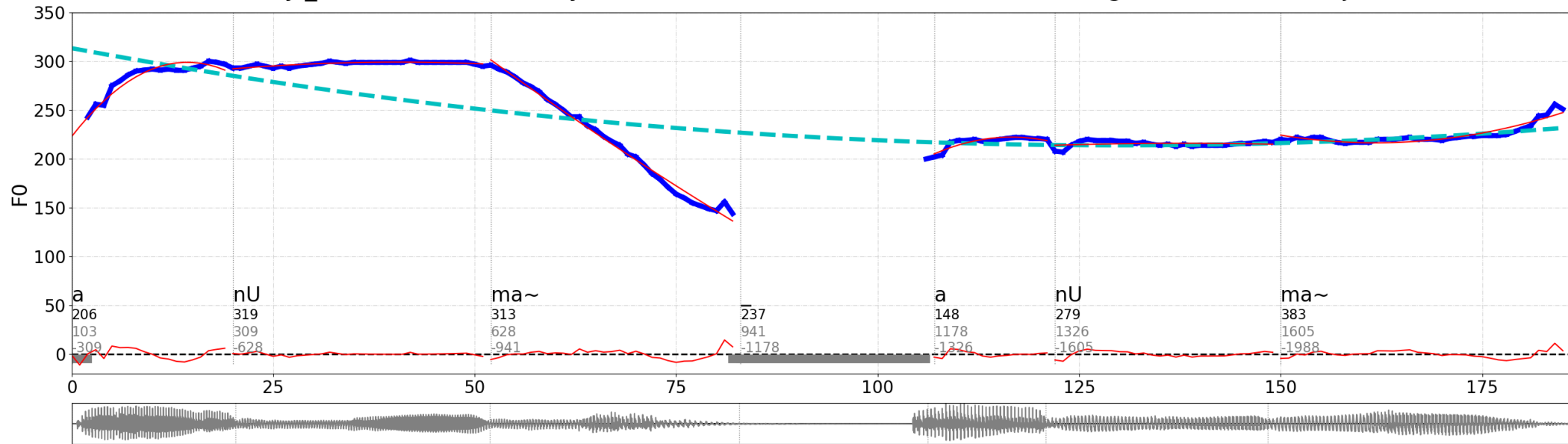


PV 01: "Mandarin_creaky4ma", tier "Syllables", x-axis 10.0ms, Model: median 1, degree 2, domain "majorIPU"



Two lexical tones in Anyi (Niger-Congo)

PV 01: "Anyi_SA-Anouman", tier "Syllables", x-axis 10.0ms, Model: median 1, degree 2, domain "majorIPU"

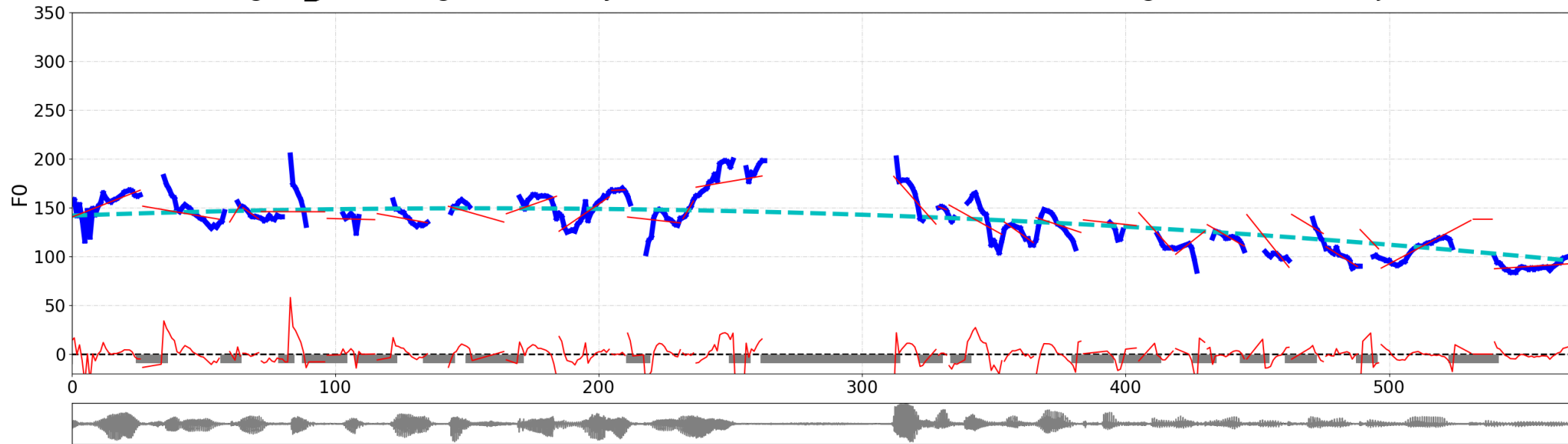


Melody

Intonation

Pitch accent, boundary tone and intonation in English

PV 01: "English_J0104G-Argen...", tier "Syllables", x-axis 10.0ms, Model: median 1, degree 2, domain "majorIPU"



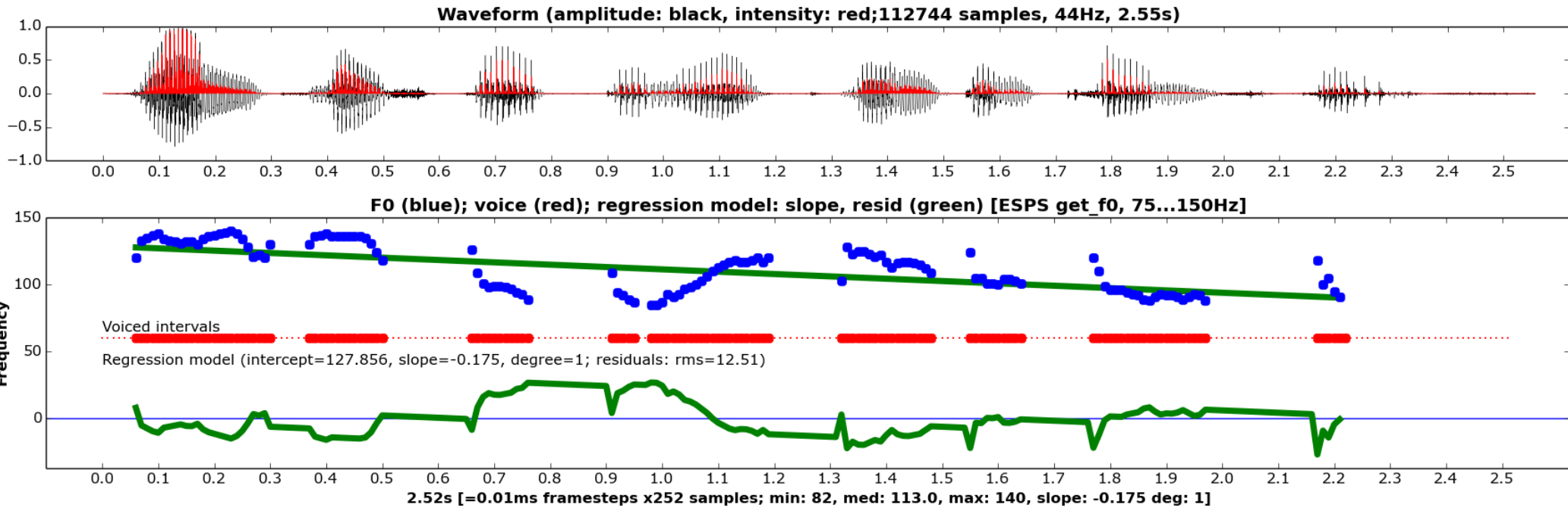
Components of Melody: Intonation, Accentuation

An example from German:

”Und endlich gab der Nordwind den Kampf auf”

(And finally the north wind gave up the contest)

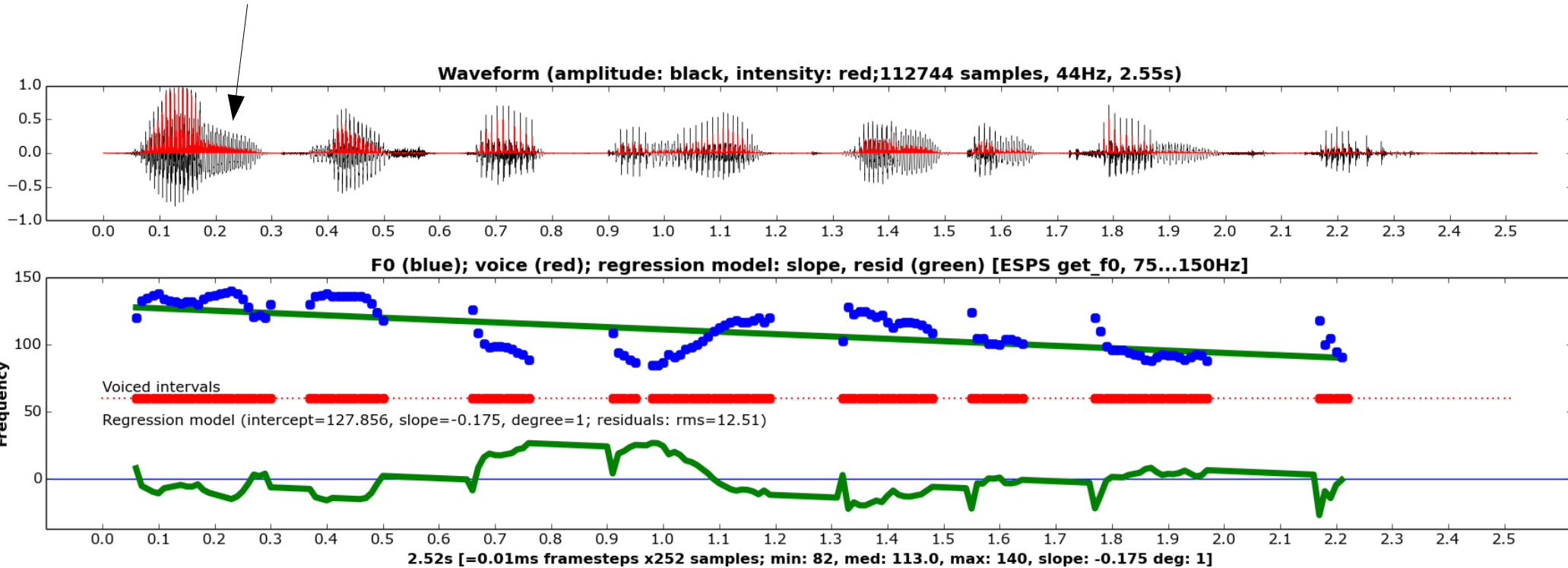
Phonetic correlates: acoustic measurements



Endlich gab der Nordwind den Kampf auf.

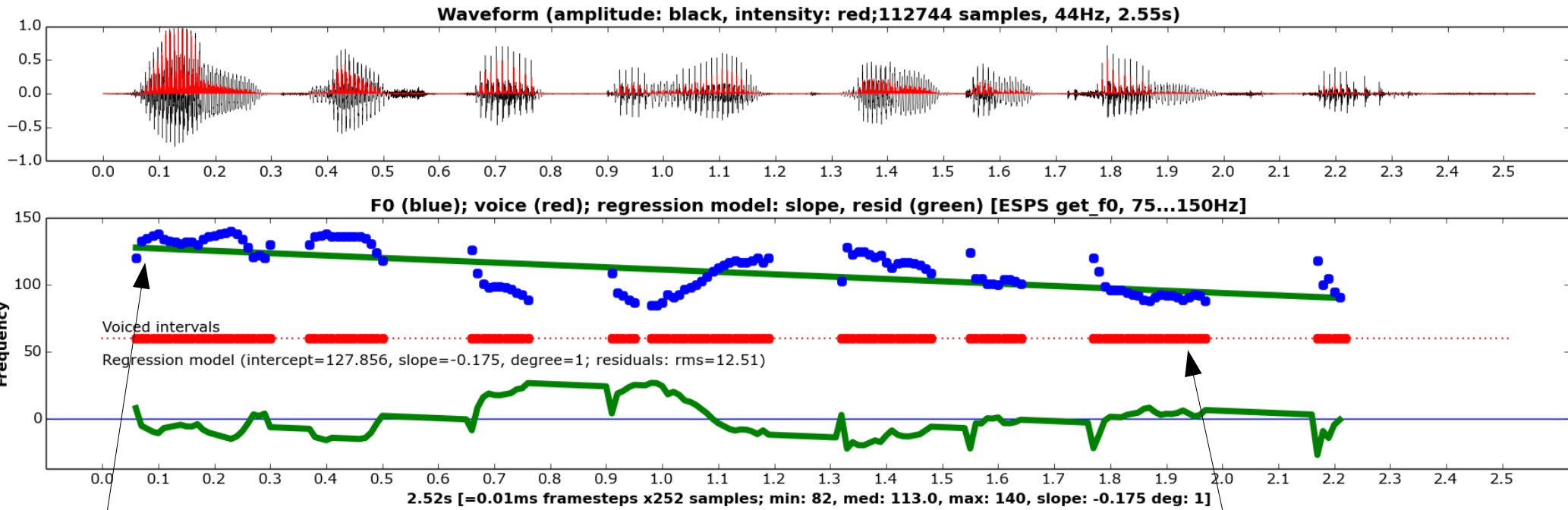
Phonetic correlates: acoustic measurements

amplitude; intensity = $f(\text{amplitude}^2)$



Endlich gab der Nordwind den Kampf auf.

Phonetic correlates: acoustic measurements

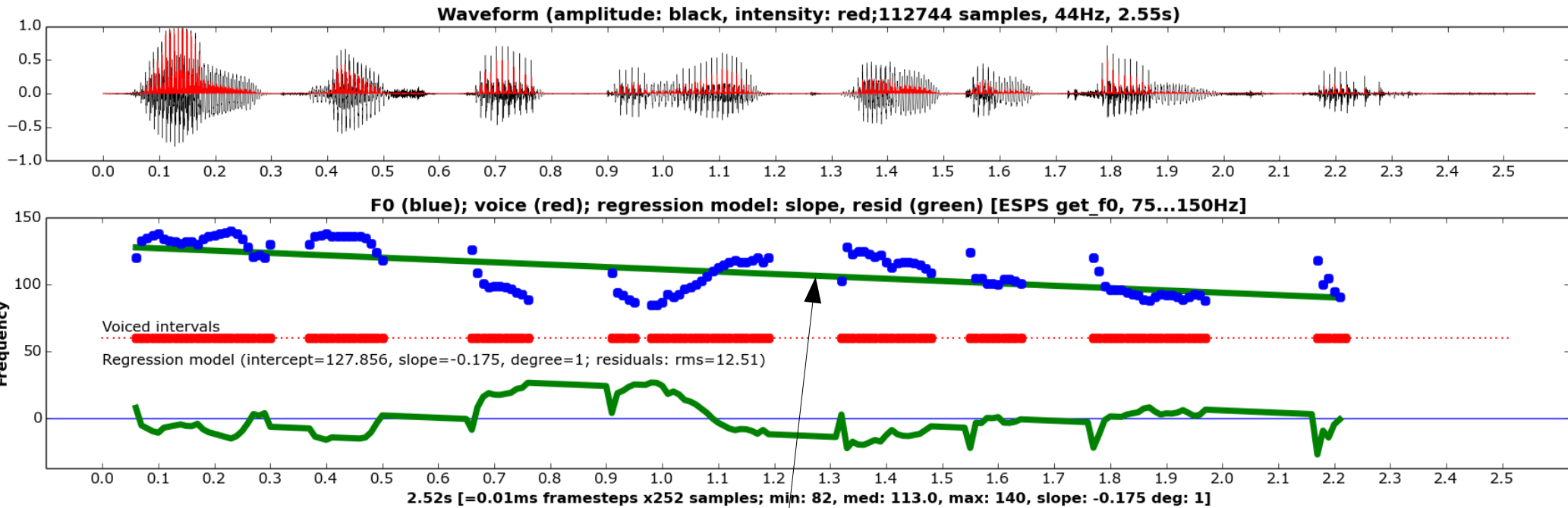


pitch track

voicing

Endlich gab der Nordwind den Kampf auf.

Phonetic correlates: acoustic measurements

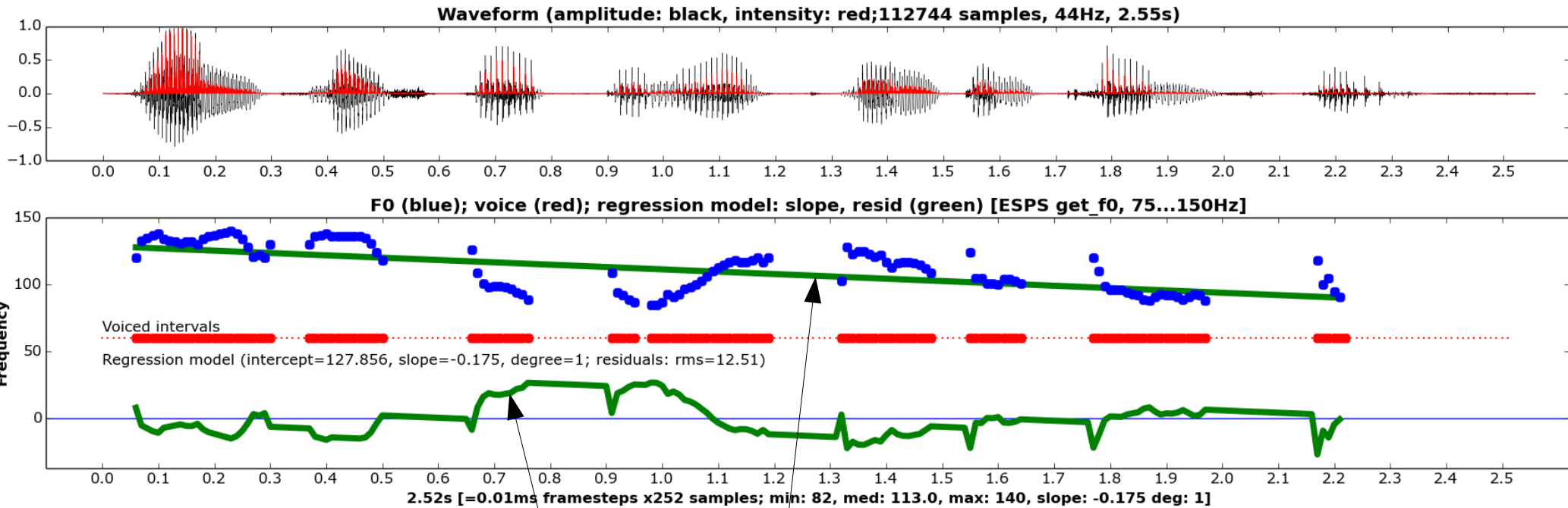


global downtrend model,
declination model

$$\text{smooth}(f_0) - f_0 = \text{microprosody}$$

Endlich gab der Nordwind den Kampf auf.

Phonetic correlates: acoustic measurements

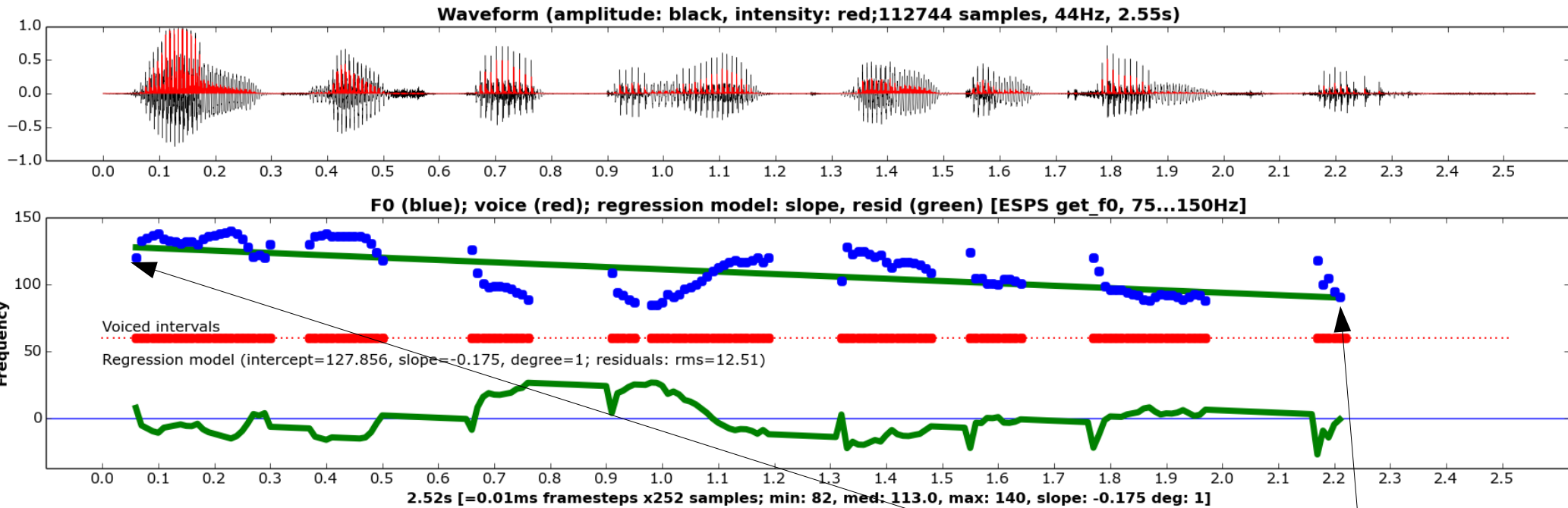


global downtrend model
 declination model
 pitch perturbation model
 microprosody model

$$\text{smooth}(f_0) - f_0 = \text{microprosody}$$

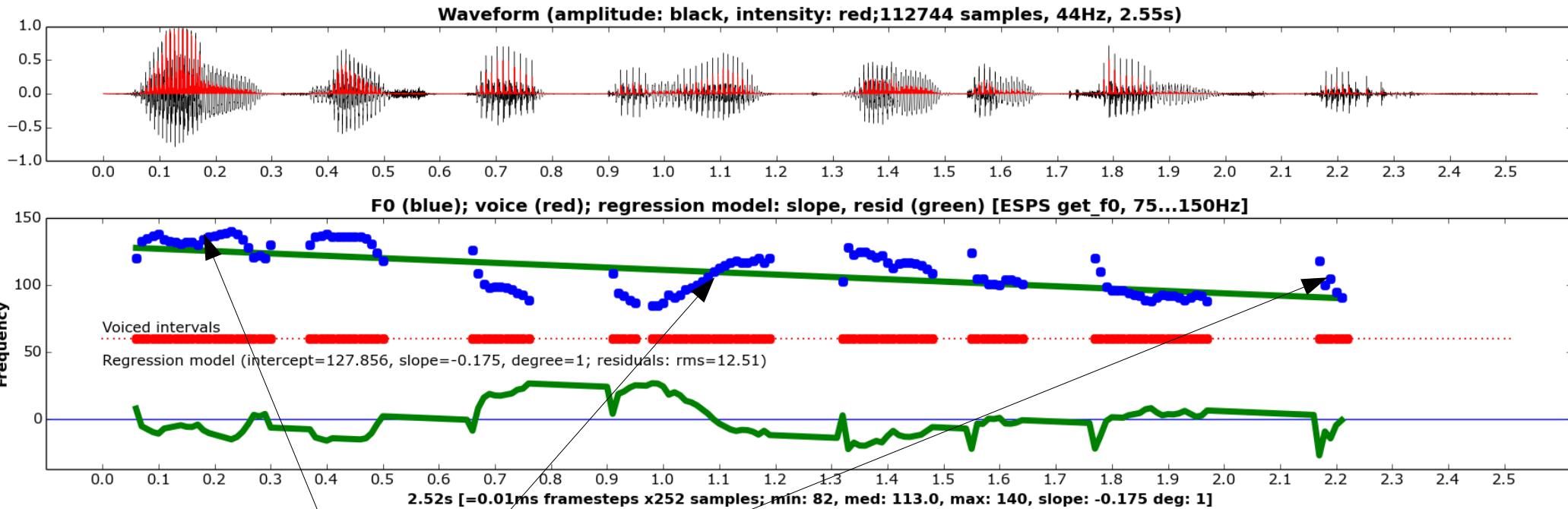
Endlich gab der Nordwind den Kampf auf.

Phonetic correlates: acoustic measurements



boundary tone

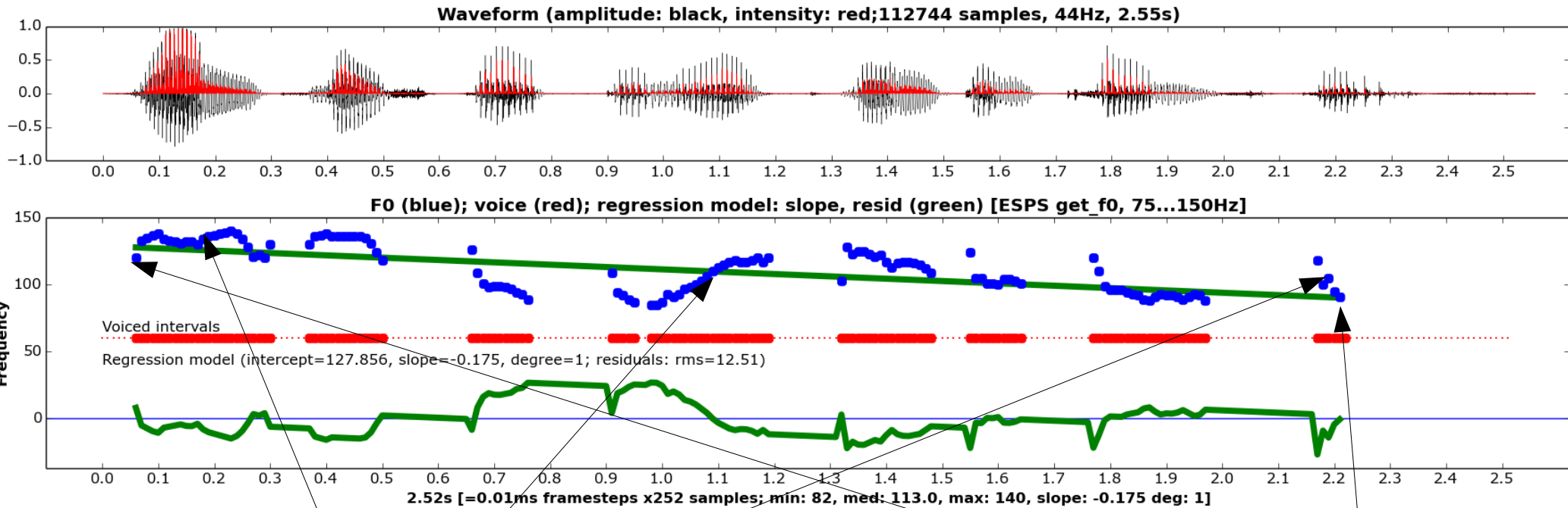
Phonetic correlates: acoustic measurements



local pitch accents

Endlich gab der Nordwind den Kampf auf.

Phonetic correlates: acoustic measurements



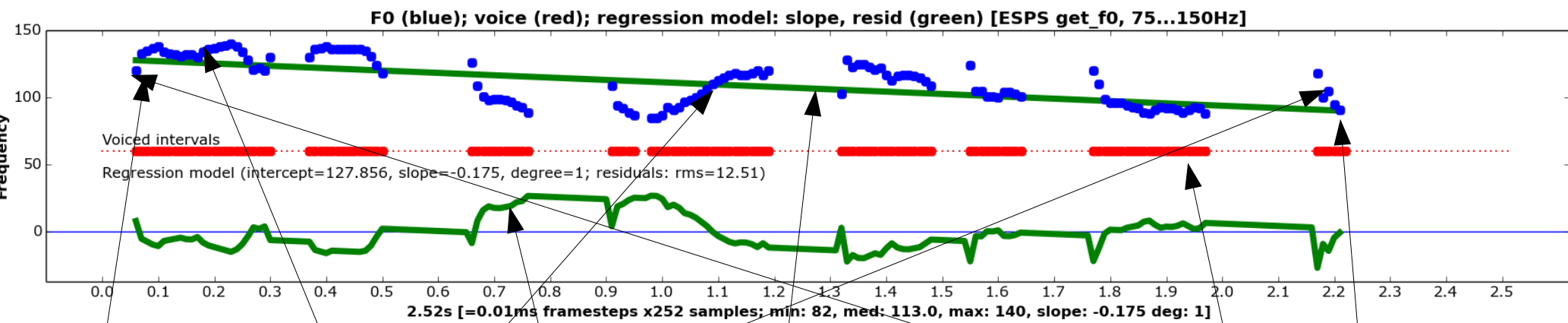
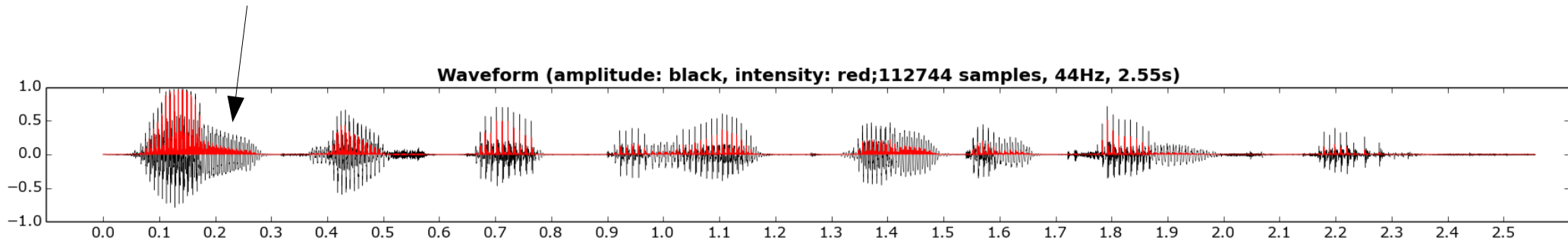
local pitch accents

boundary tone

Endlich gab der Nordwind den Kampf auf.

Phonetic correlates: acoustic measurements

amplitude; intensity = $f(\text{amplitude}^2)$



pitch track
f0

local pitch accents

global downtrend model
declination model

pitch perturbation model
microprosody model

voicing

boundary tone

Endlich gab der Nordwind den Kampf auf.

Modelling intonation:

Fundamental frequency smoothing: global procedures

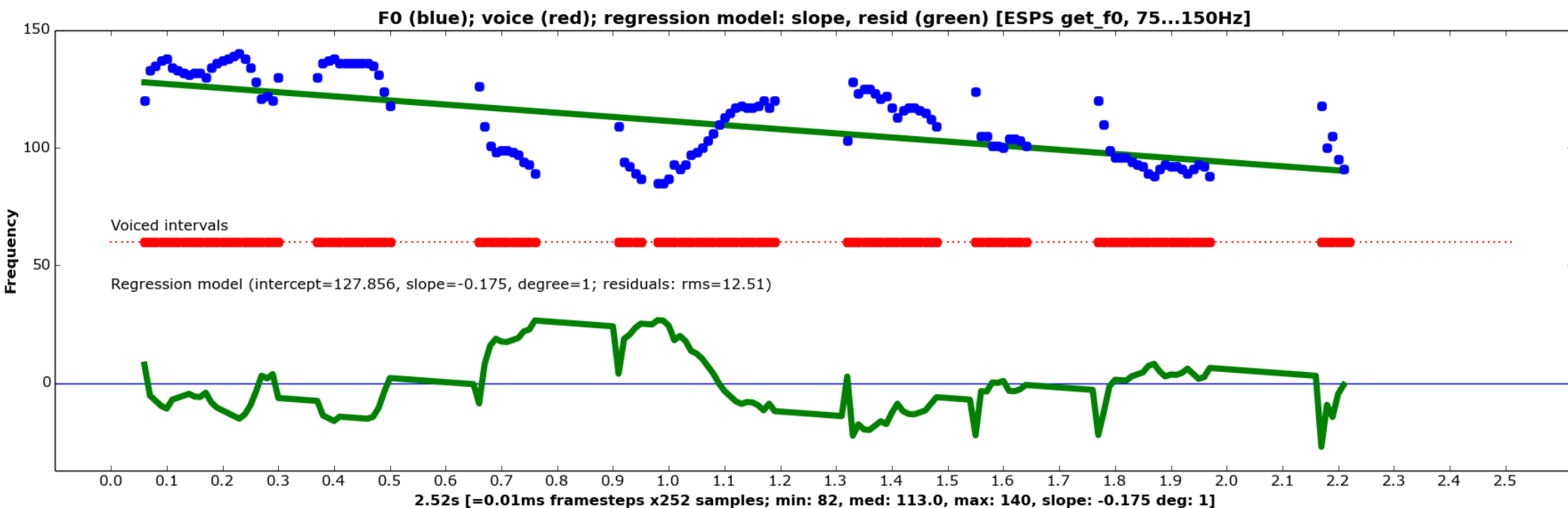
Smoothing: different approaches, different goals

- Smoothing by polynomial regression (degree n):

$$y = a_0 + a_1 \cdot x + a_2 \cdot x^2 + a_3x^3 + \dots + a_nx^n + \varepsilon$$

- Smoothing by linear regression (degree 1)

$$y = a_0 + a_1x + \varepsilon$$



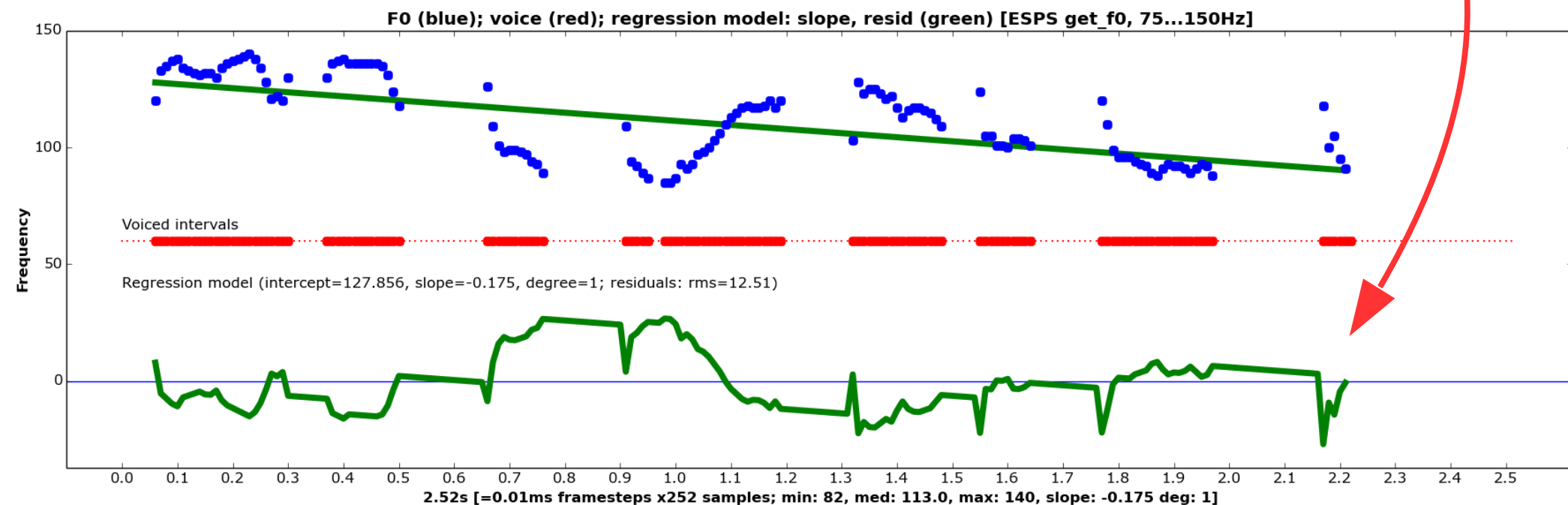
Smoothing: different approaches, different goals

- Smoothing by polynomial regression (degree n):

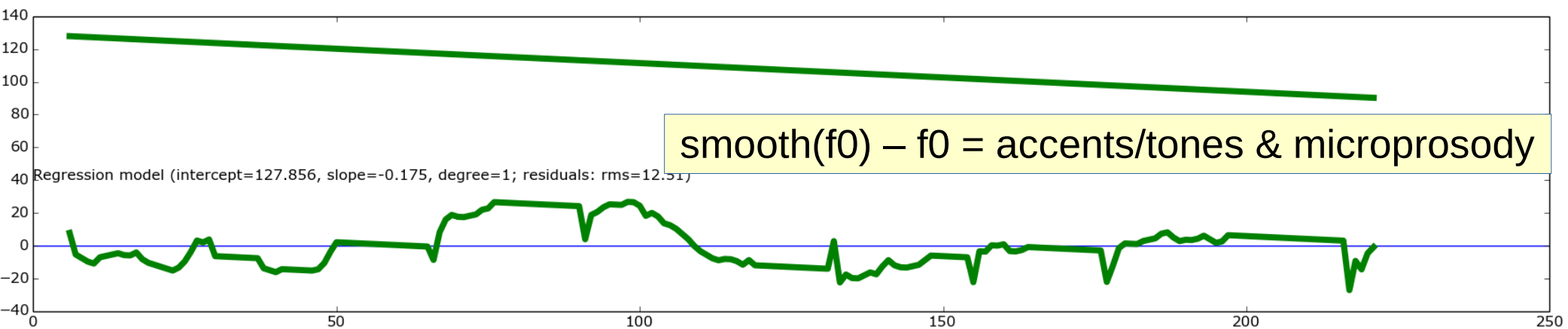
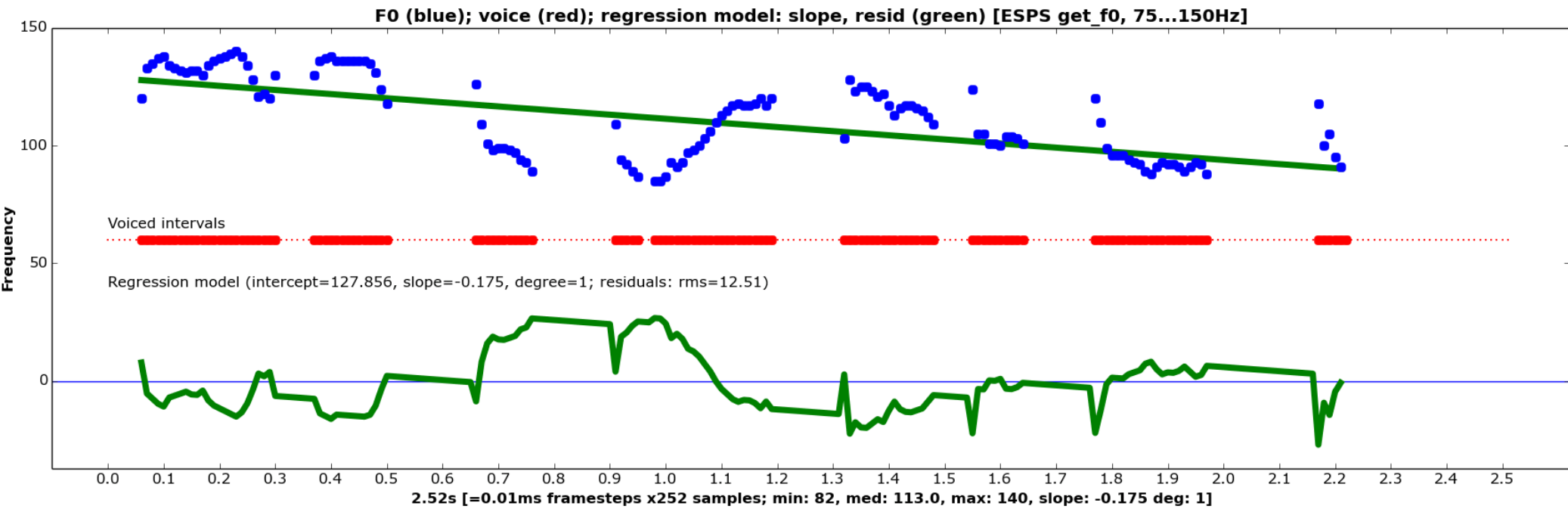
$$y = a_0 + a_1 \cdot x + a_2 \cdot x^2 + a_3 x^3 + \dots + a_n x^n + \varepsilon$$

- Smoothing by linear regression (degree 1)

$$y = a_0 + a_1 x + \varepsilon$$

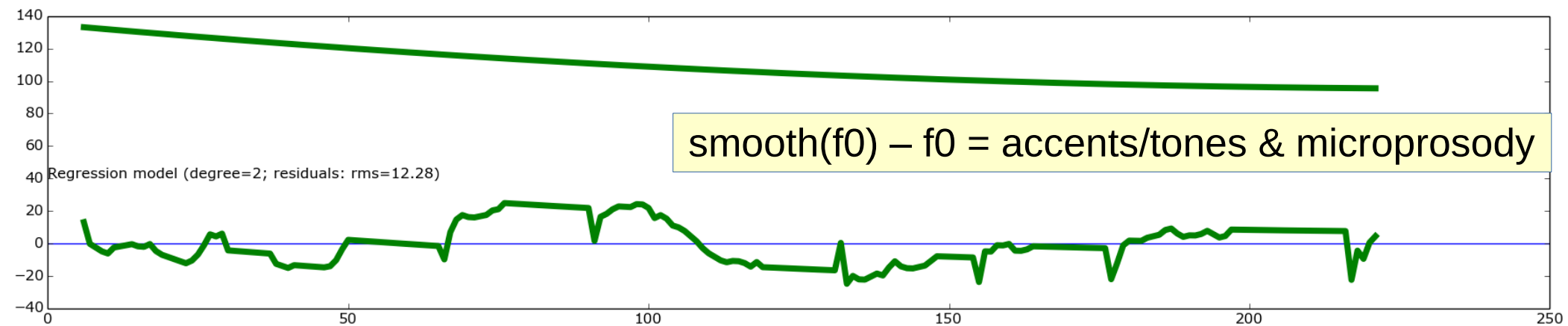
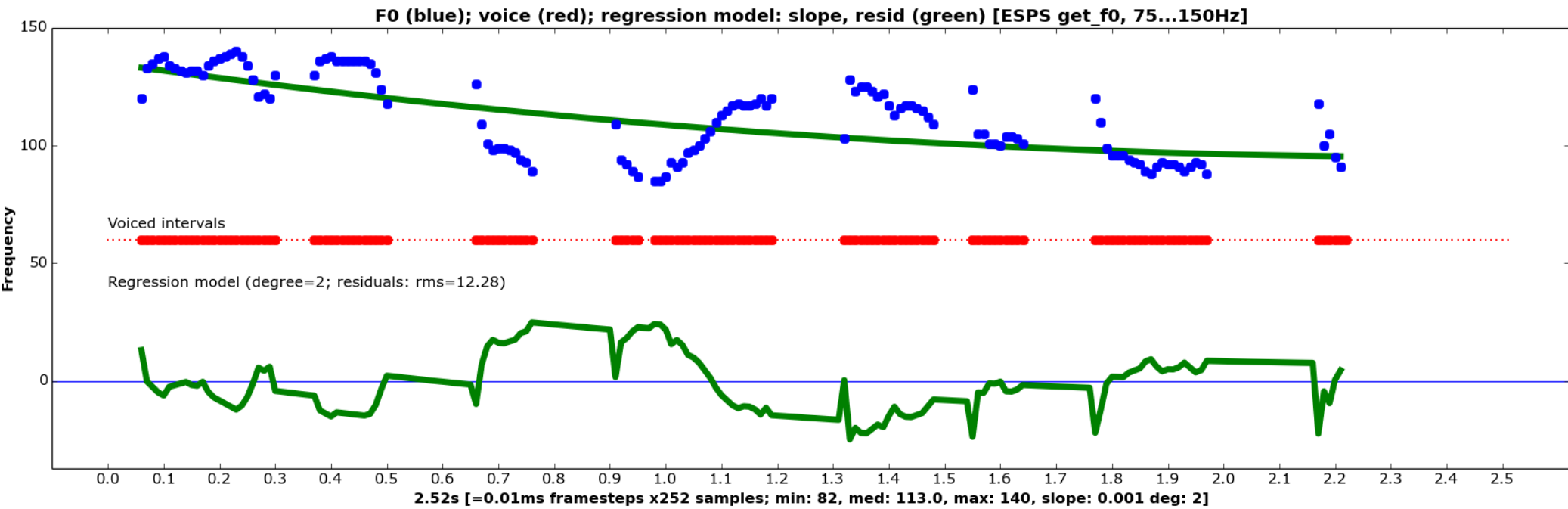


Global linear regression contour



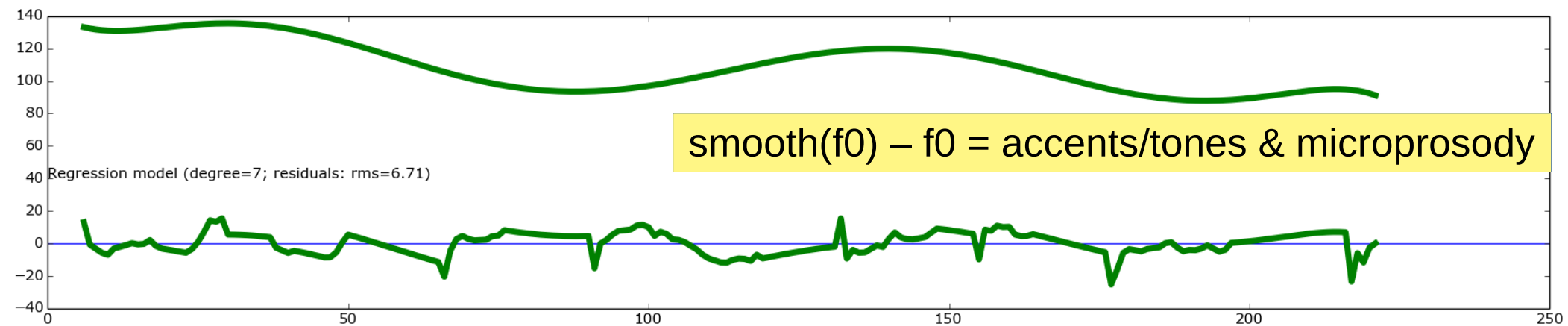
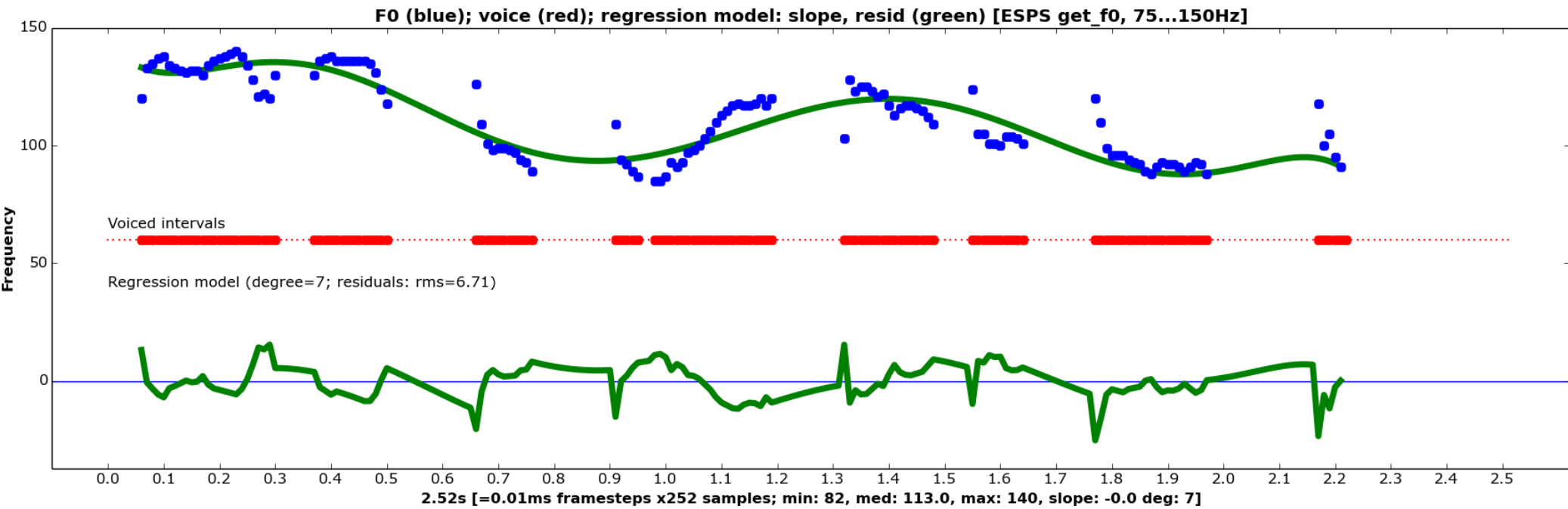
Endlich gab der Nordwind den Kampf auf.

Global quadratic regression contour



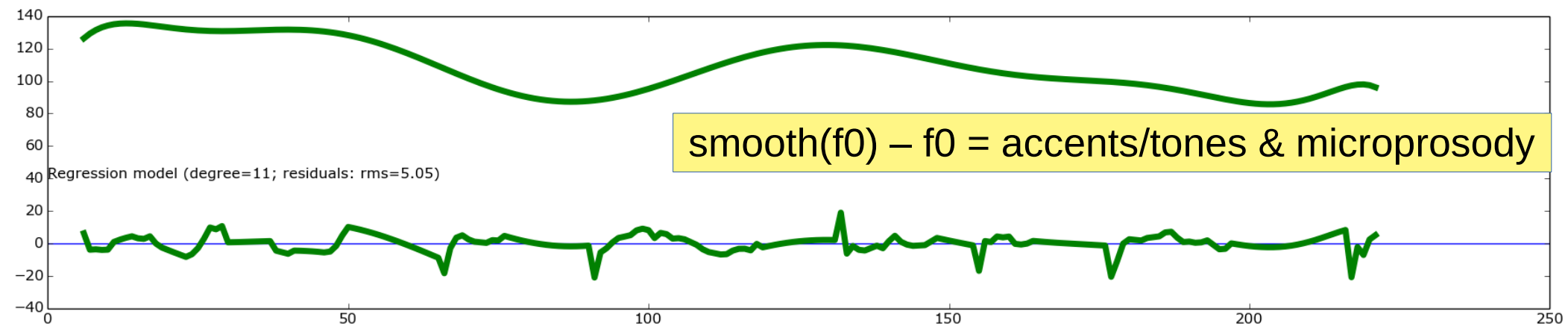
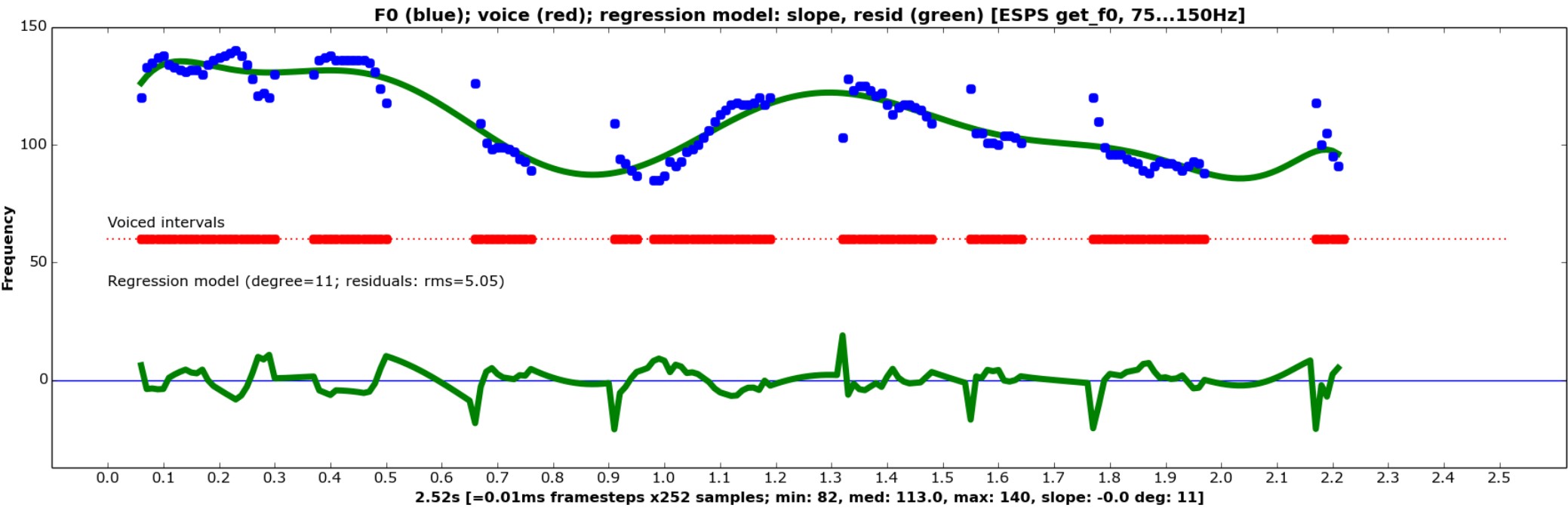
Endlich gab der Nordwind den Kampf auf.

Global regression contour, degree 7



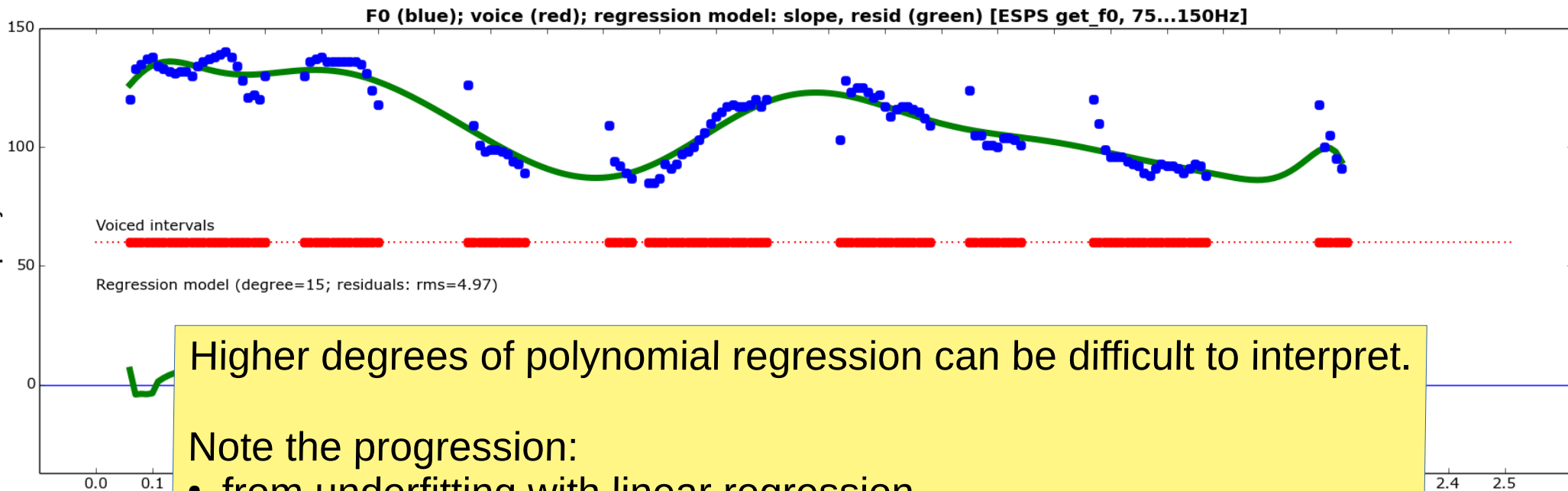
Endlich gab der Nordwind den Kampf auf.

Global regression contour, degree 11



Endlich gab der Nordwind den Kampf auf.

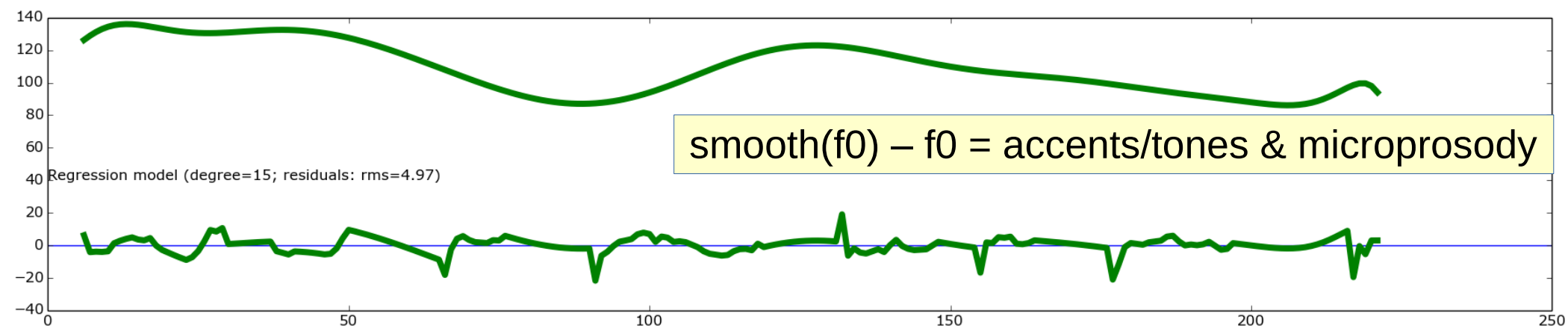
Global regression contour, degree 15



Higher degrees of polynomial regression can be difficult to interpret.

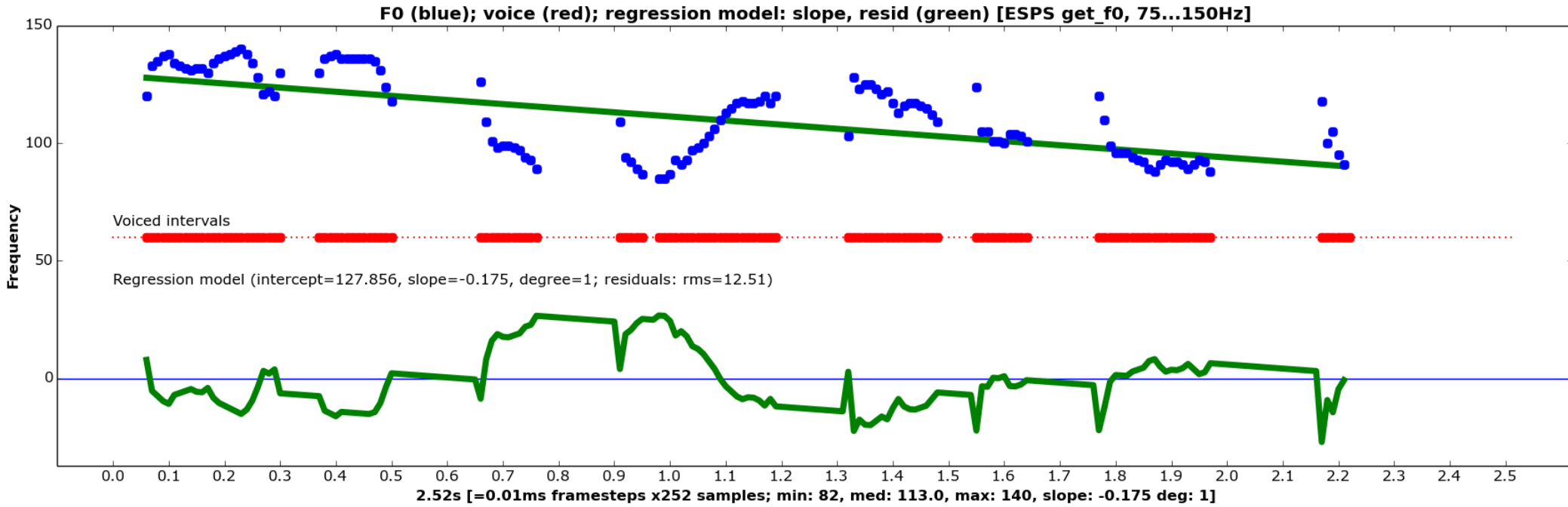
Note the progression:

- from underfitting with linear regression
- to overfitting with higher degrees polynomial regression



Endlich gab der Nordwind den Kampf auf.

Global regression contours, up to degree 20

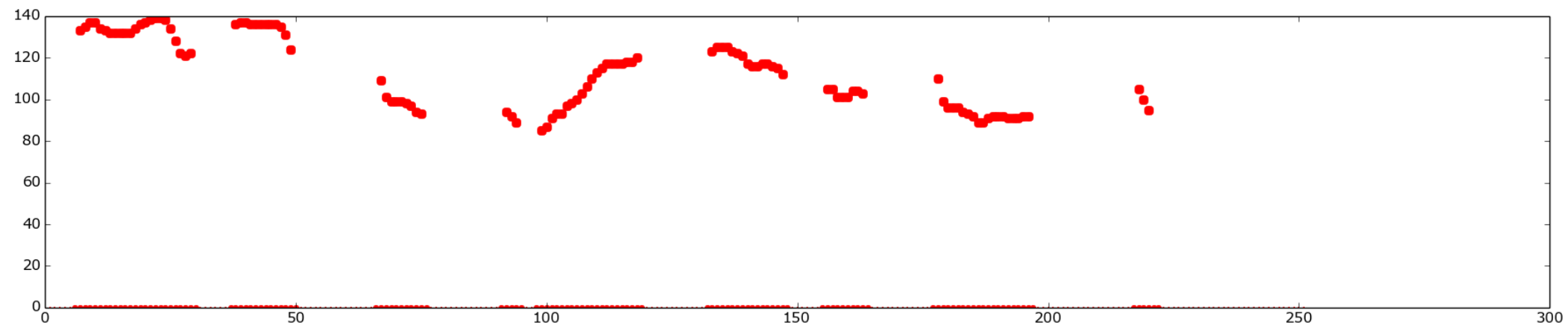


Endlich gab der Nordwind den Kampf auf.

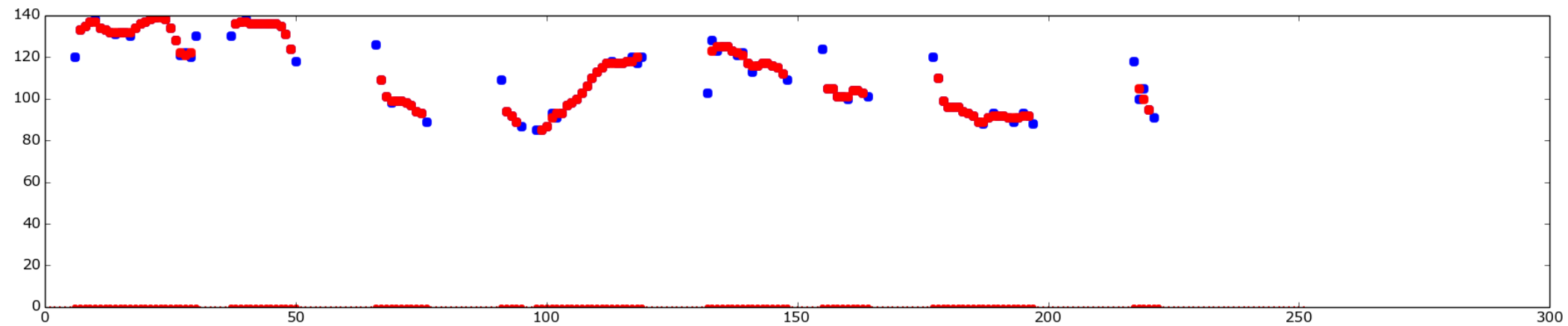
Modelling accentuation:

Fundamental frequency smoothing: local procedures

Simple median filter (scope: 3), often used

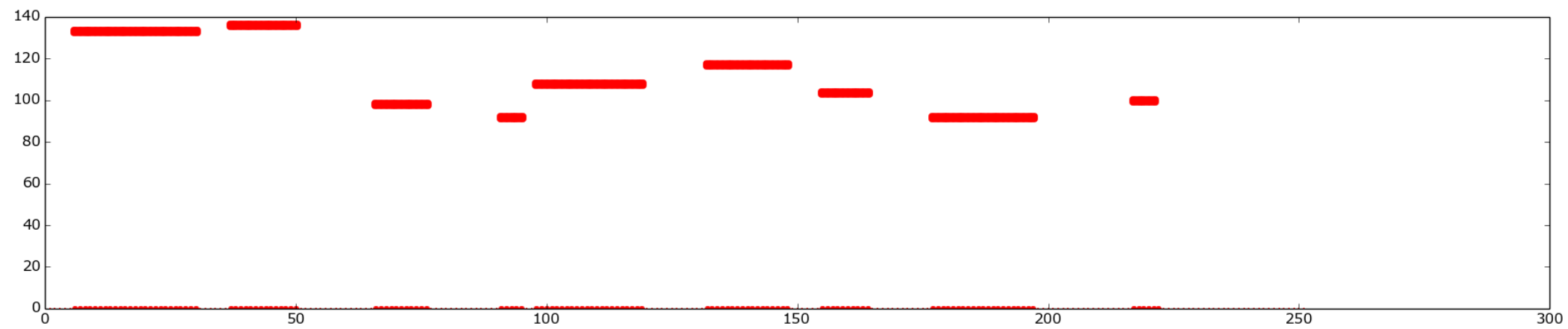


Endlich gab der Nordwind den Kampf auf.

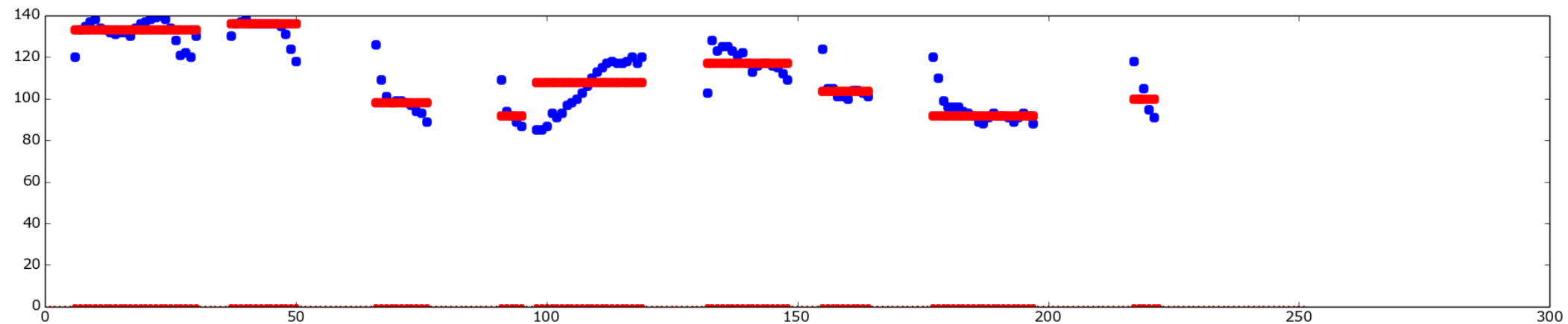


Each F0 value is normalised to the median F0 value of its immediate neighbours

Simple local median levelling filter – robotic!

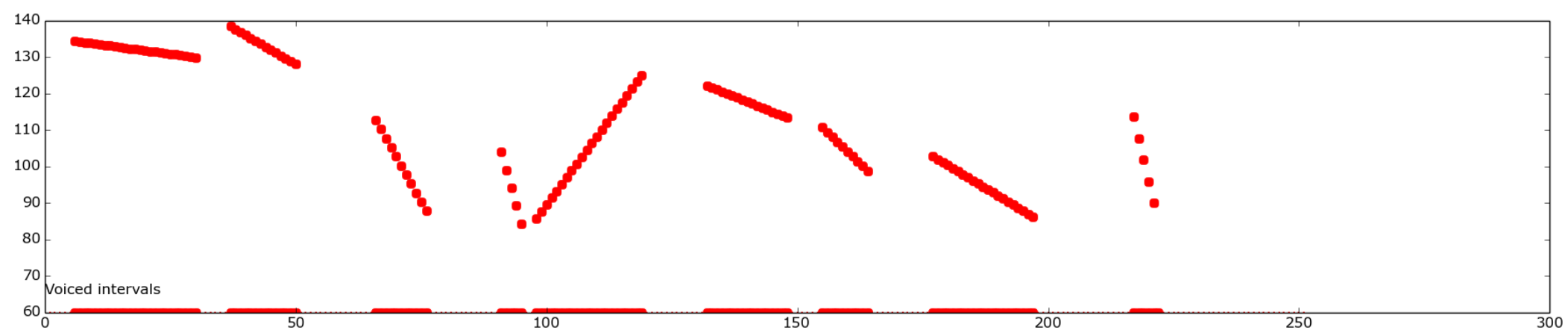


Endlich gab der Nordwind den Kampf auf.

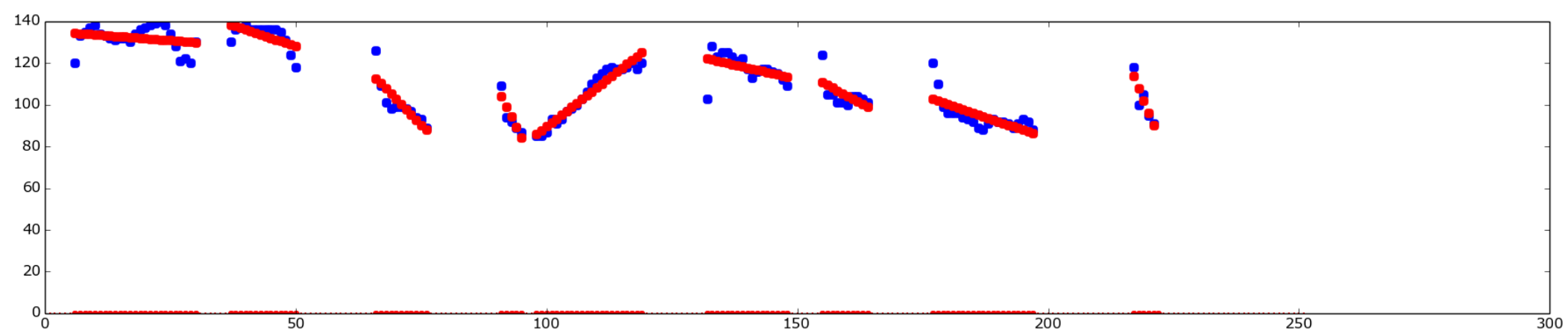


Each F0 value in a sequence is normalised to the median F0 value for the sequence

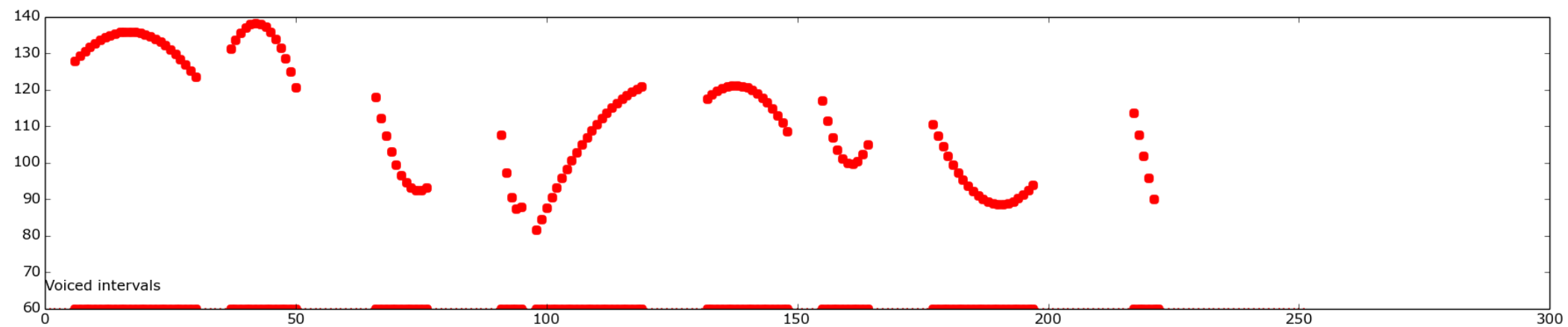
Local voicing regression contours, degree 1



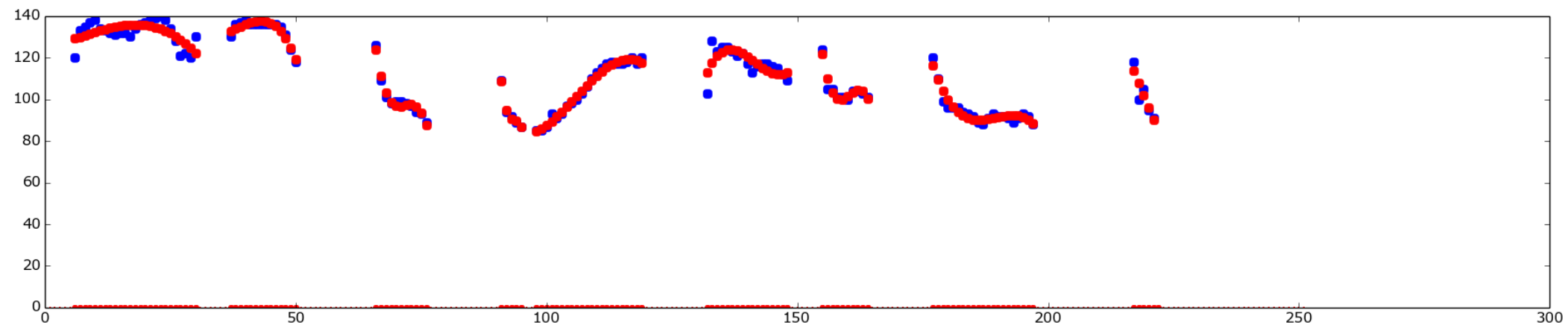
Endlich gab der Nordwind den Kampf auf.



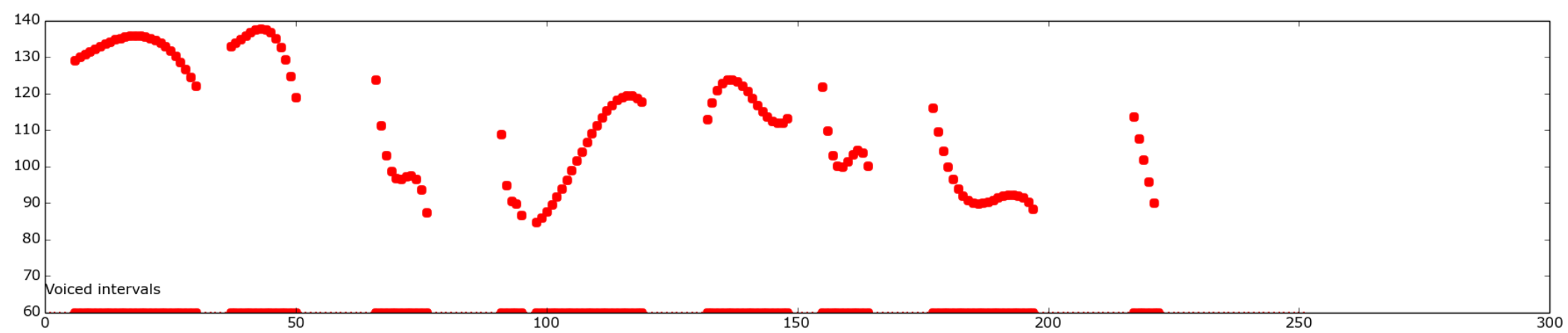
Local voicing regression contours, degree 2



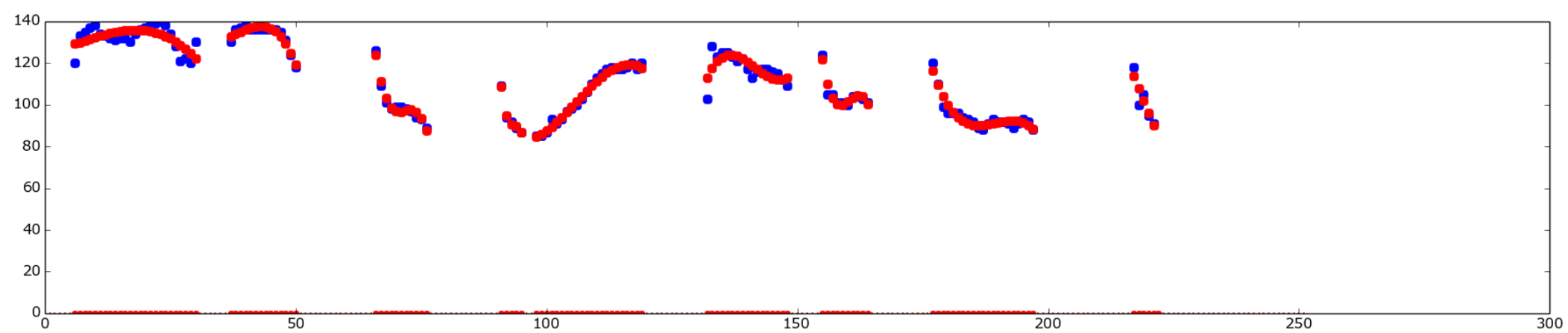
Endlich gab der Nordwind den Kampf auf.



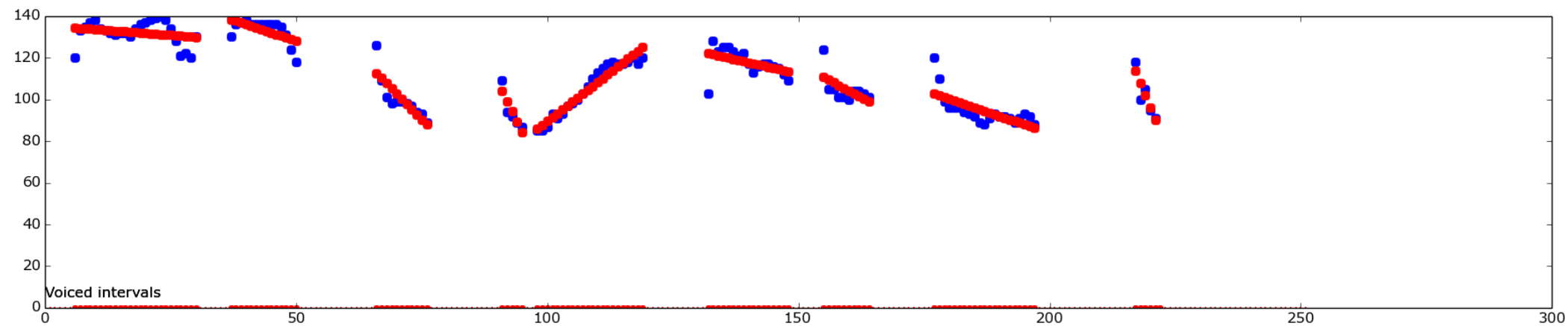
Local voicing regression contours, degree 3



Endlich gab der Nordwind den Kampf auf.



Local voicing regression contours (1...5)



Endlich gab der Nordwind den Kampf auf.

Higher degrees of polynomial regression are necessary with longer and more complex utterances, but can be difficult to interpret.

Note the progression:

- from underfitting with linear regression
- to overfitting with higher degrees polynomial regression

Summary and conclusion

- Different languages have different kinds of melody
 - Accent and intonation languages like English
 - Tone and intonation languages like Chinese
- The differences are in
 - Linguistic structure
 - Phonetic shape
- The melodies can be modelled
 - Globally (for intonation)
 - Locally (for accent and tone)