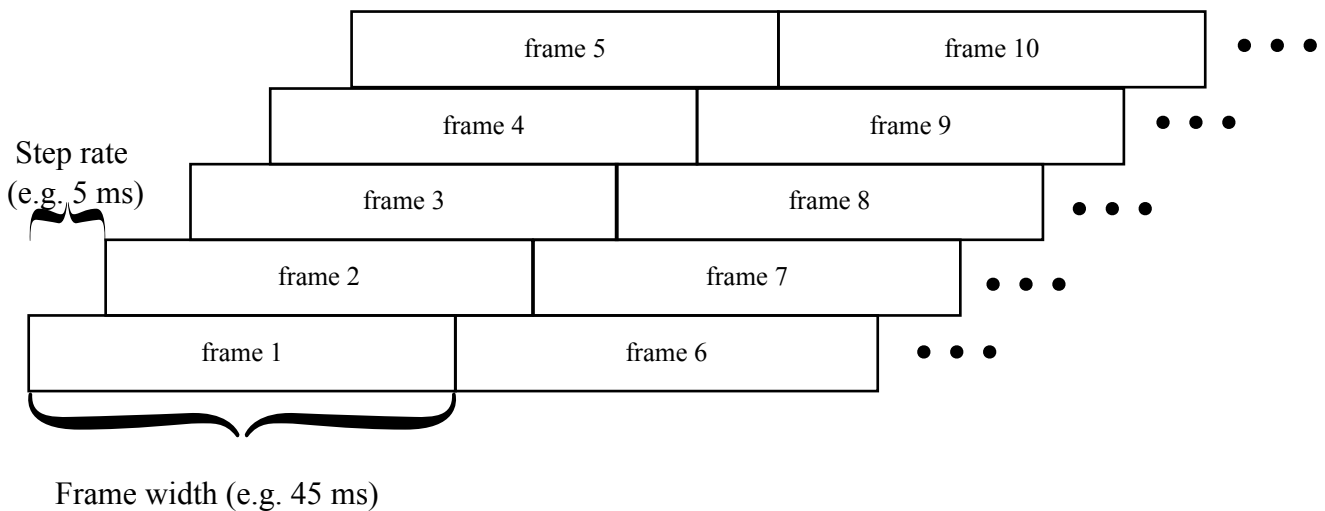


**Disclaimer**

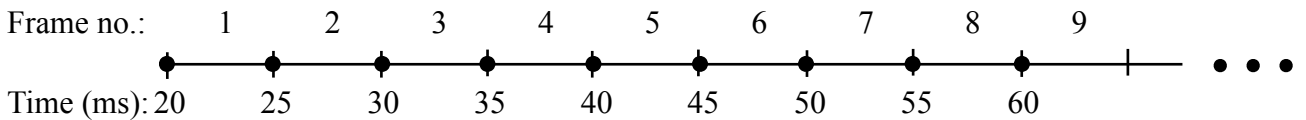
The statements in this file are based on my observations on the behaviour of PRAAT. I have not investigated to source code, which would exactly tell what is going on. If you find inaccurate statements, please mail me at <reetz.phonetics@gmail.com> and I will correct the errors.

**Compare pitch values computed with a script with the interactive display of PRAAT**

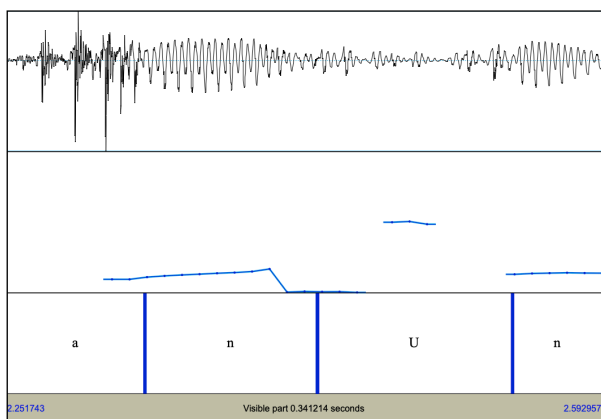
PRAAT, like most other signal processing programs, computes the pitch not for every speech sample (after all, you cannot see a periodicity from one sample), but for a certain time window, called **frame**, around 45 ms wide in PRAAT (the exact time seems to vary depending on some parameters I have not figured out – and I have looked into the source code). These frames move with a certain **step rate**, usually overlapping, across the speech signal.



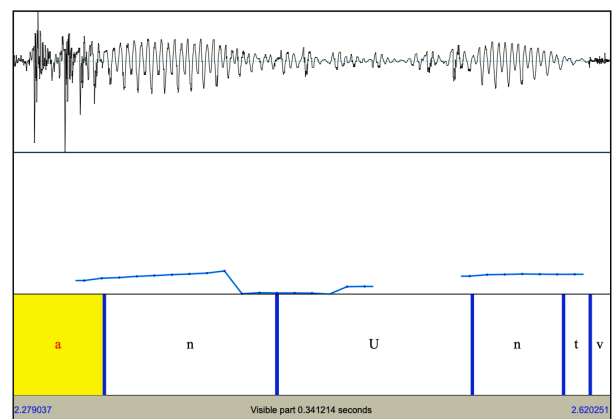
The frame information is positioned at the beginning of a frame in the *View & Edit* of a pitch object:



Additionally, PRAAT computes the pitch in the interactive edit window. Depending on where the window begins (in time of the signal), the pitch analysis might compute different values. Note the difference between the two displays below.



2.251743



2.279037

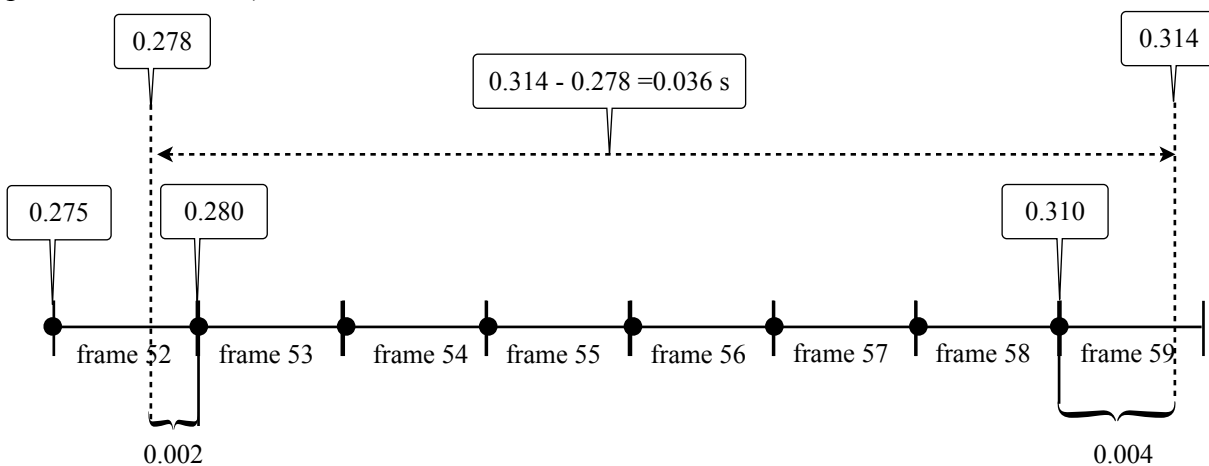
Next to pitch parameters (like lowest and highest pitch frequency to be computed), positioning of the frames, their size and their step rate can change the computed pitch values. Computing the pitch from a *Sound object* with the command *To pitch* will always start at the beginning of the whole file and the results depend on the *Pitch floor*, *Pitch ceiling* and *Step rate* values for a selected algorithm. Consequently, it is not always easy to compare pitch values computed with a script with the interactive display in PRAAT.

### Computing percentage of voicing in an interval in PRAAT

(the necessity for this becomes clear in the next section)

PRAAT can compute pitch mean, standard deviation and quantiles (like the median) of an interval. Actually, PRAAT does not use the beginning and end times of an interval as such, but selects for these times the frames that belong to these times and takes the pitch values of these frames for the time points.

Let's assume a step rate of 5 ms, then the frames in the figure below begins at  $20 + 51 \times 5 \text{ ms} = 0.275 \text{ s}$  (remember, that the first frame will begin at 20 ms and pitch values are reported at the start positions of frames).



Let's assume that only frames 52, 53 and 59 are voiced, frames 54 to 58 are unvoiced. To find the associated frame number, PRAAT offers a function *Get frame number from time* (remember, start time of the first frame might vary; i.e., one cannot compute from a given time the frame number knowing the step rate). In our example, the function *Get frame number from time* will give 'frame 52' for the beginning of the interval, and 'frame 59' for the end of it. The percentage of voicing is then (computed in time):

$$([\text{end time of frame 52}] - [\text{left edge of interval}] + \\ ([\text{number of voiced frames from 53 to 58 (including these)}] \times [\text{step rate}]) + \\ ([\text{right edge of interval}] - [\text{beginning time of frame 59}]))$$

or, in our example, with one frame (53) fully within interval (computation is done in ms to make it easier to read):

$$(280 - 278) + (1 \times 5) + (314 - 310) = 3 + 5 + 4 = 11$$

The total duration of the interval is  $(314 - 278) = 36$

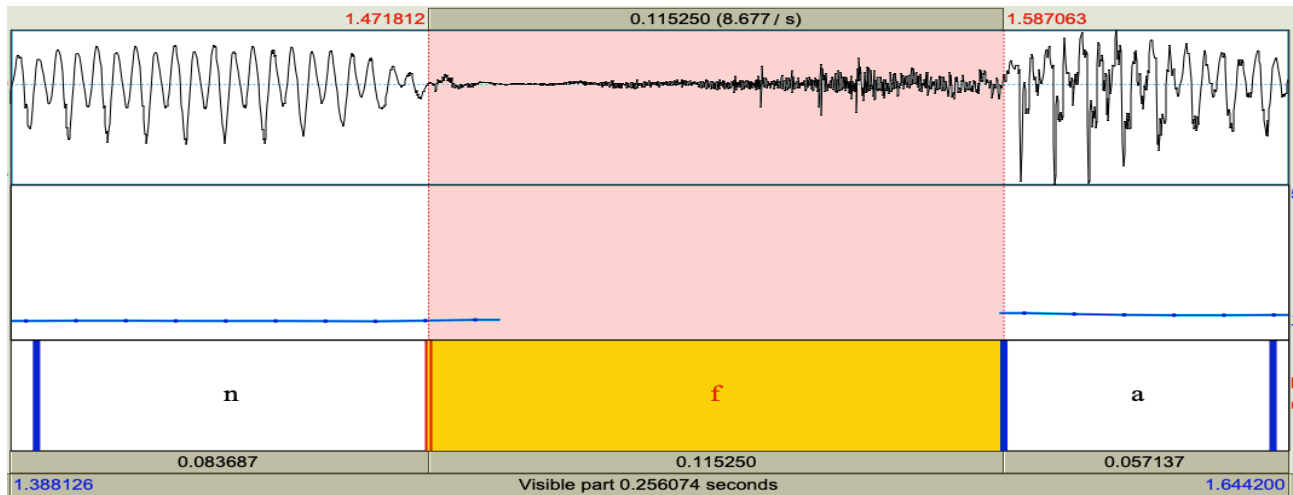
The percentage of voiced duration is then  $(11 / 36) \times 100 = 30.56\%$

that is, about 31% of the interval is voiced.

If we would have counted only the voiced frame of the interval, we would have frames 52, 53 and 59 as voiced, 54 to 58 unvoiced, i.e.  $(3 \text{ voiced} / 5 \text{ unvoiced}) \times 100 = 60\%$  voicing of the interval.

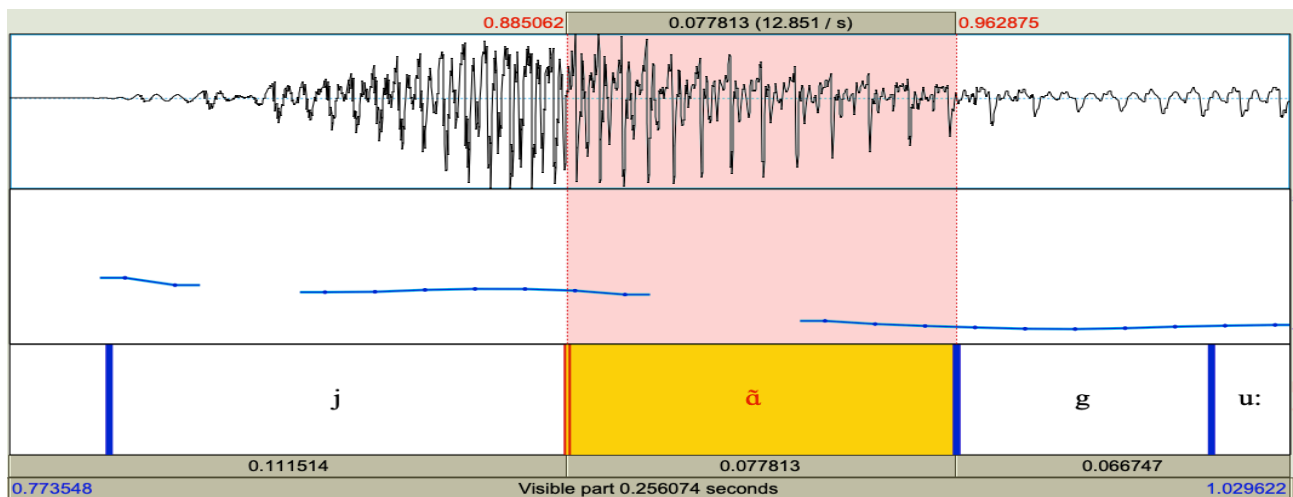
### Computing mean (median, quantiles) of an interval in PRAAT

PRAAT can compute mean, standard deviation and quantiles (like the median) of an interval. But it can be the case that only one sample of speech that belongs to a voiced frame falls within an interval. But this one sample (which is ‘voiced’ because it belongs to a voiced frame) will be taken to compute the average pitch for the whole interval (in the figure and list below there are samples (and frames) at both edges of the interval that contribute to the computation):



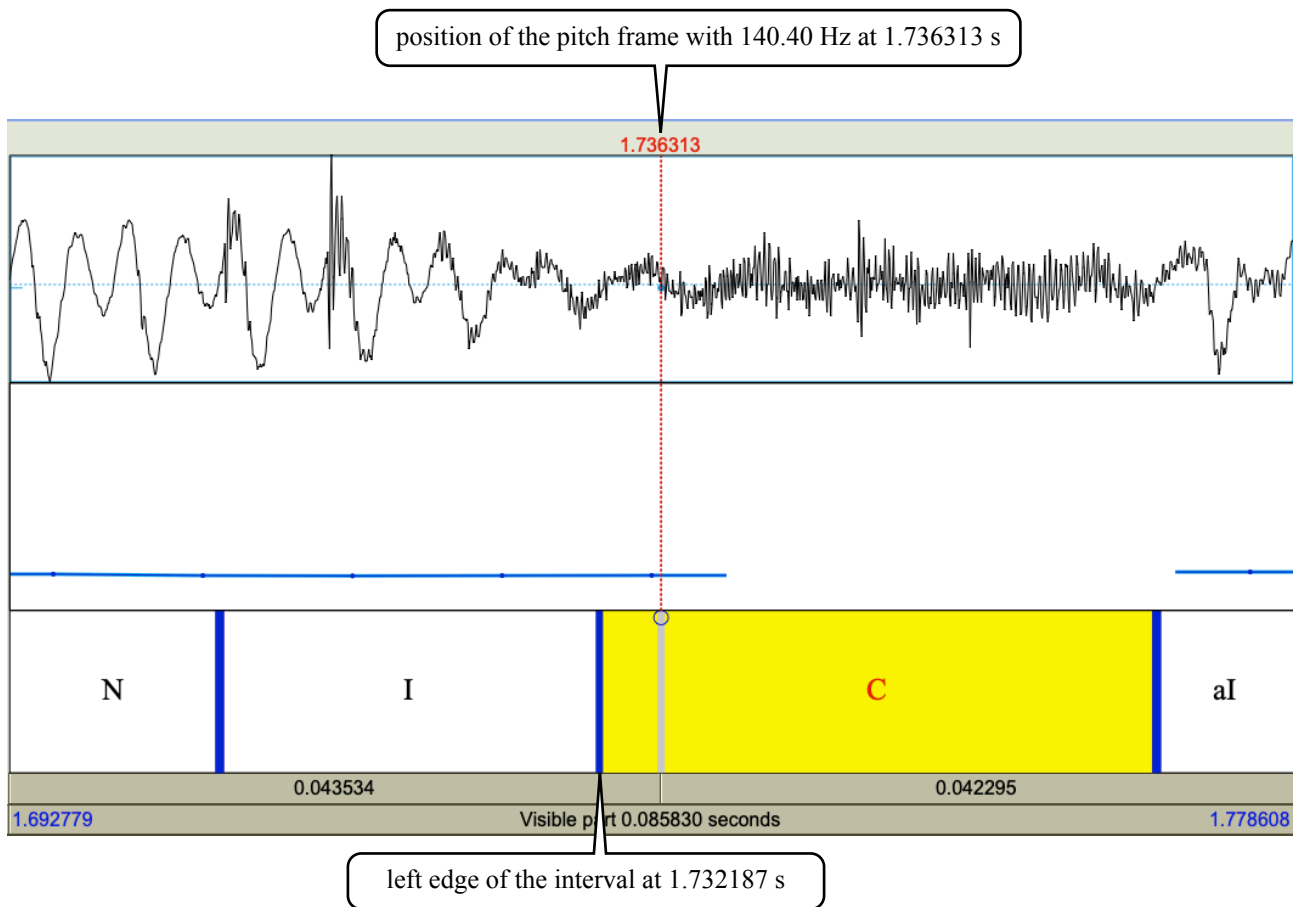
File	Label	Start(s)	Duration(ms)	Center(Hz)	Mean(Hz)	StDev(Hz)	Median(Hz)	Voiced(%)
g071a000	f	1.4718	115.3	NA	136.38	12.95	140.37	9

In this example, there are a bit more than one voiced frame at the beginning of the interval with the label [ f ], and a fraction of a frame at the end of it. Still the listing reports a mean and standard deviation and a median. The reason is that all frames without a pitch value (i.e., unvoiced frames) are treated as **missing values** (in R-notation NA = Not Available). Consequently, these frames are not included in the computation of the mean or the determination of the median. (Setting the pitch in these frames to ‘0’ would be not correct since there is ‘no pitch’ – and not a pitch with the frequency ‘0’.) Only the missing pitch for the center location of the interval could be an indicator that the interval is not fully voiced. To help to identify these cases, the scripts *Pitch.praat* and *Pitch\_contour.praat* compute the percentage of the voicing in an interval (see previous section) to allow an estimate whether the interval is really voiced or not. Note, however, that also a voiced interval can have several missing values (i.e. voiceless frames) because the pitch algorithm failed to determine the pitch.



File	Label	Start(s)	Duration(ms)	Center(Hz)	Mean(Hz)	StDev(Hz)	Median(Hz)	Voiced(%)
g071a000	ã	0.8851	77.8	NA	162.51	43.27	137.06	61

Additionally, PRAAT seems to compute the mean and standard deviation from all stretches of speech that are voiced (in a similar manner as explained in the previous section), but counts only whole frames to determine quantiles (e.g. median, minima and maxima, etc.):



File	Label	Start(s)	Dur(ms)	Center(Hz)	Mean(Hz)	StDev(Hz)	Min(Hz)	Median(Hz)	Max(Hz)	Voiced(%)
g071	C	1.7322	37.2	NA	140.39	NA	140.40	140.40	140.40	25

(The fact that the mean is reported with 140.39 Hz and the quantiles with 140.40 Hz is most likely a consequence of the computation of the fraction of a frame from the internal pitch value of 140.39555838849725 Hz; rounding this value to 2 decimals gives 140.40 Hz whereas the computation of the mean lead to a slightly lower value due to rounding errors.)