

Some explanations for the columns of times, f0 and Oq values in excel.

Example: the token UID 0111 from data of speaker F3: the first syllable /ja²/ in the carrier sentence /ja² mät⁶ xtarget syllablex ʈəŋ³/ 'do you know the word xxx'

Explanation of columns B and C in excel: time points of each item

Figure 1 shows the EGG signal of UID 0111 is segmented in SoundForge with two green vertical bars that mark the beginning and end of this syllable. This segmentation is based on the acoustic signal that is synchronized with the EGG signal in the original file because the transition between the consonant and the rhyme can be much more easily detected visually in the acoustic signal than in the EGG signal.

The time in columns B and C of excel (in seconds) or the start and end columns in Region List (as detailed time in hh:mm:ss) is measured based on the duration from these two green vertical lines to the beginning of the audio file.

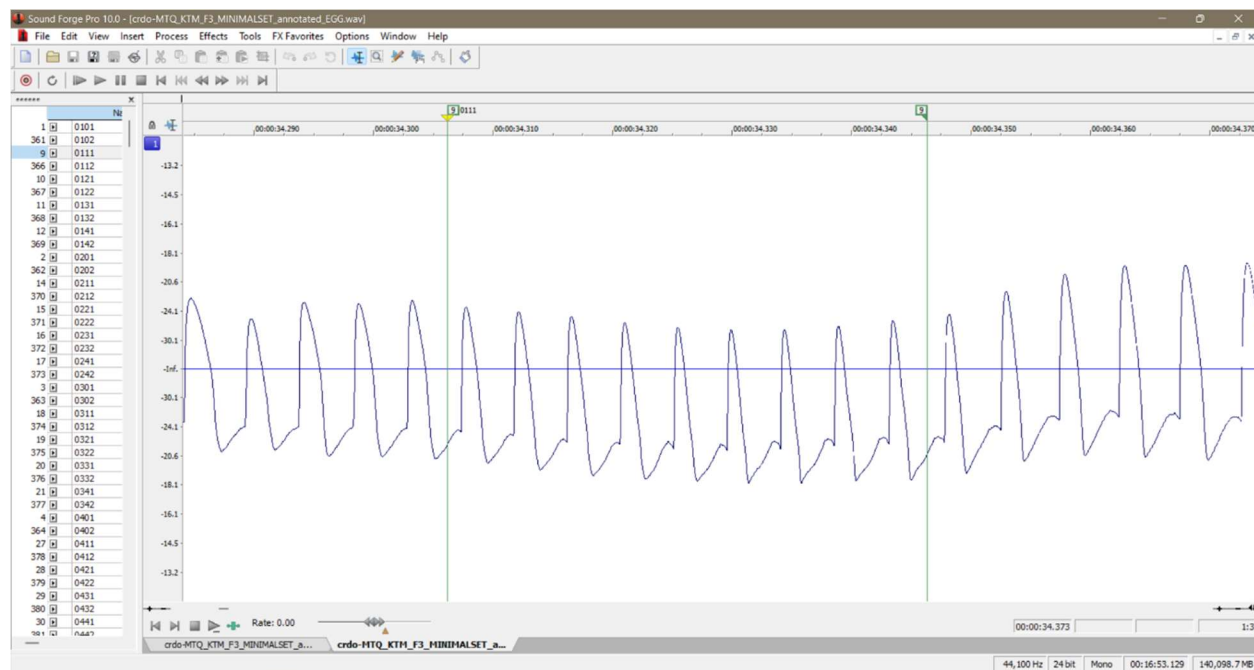


Figure 1: EGG signal in SoundForge with segmentation of the syllable 0111 by two green vertical lines

Explanation of columns D and E in excel: time points of each cycles in the item

When the signal is processed in peakdet token by token, at the turn of token 0111, the EGG signal will be extracted exactly at the location of two segmented green lines in SoundForge as shown in Figure 2. This means that the first green line is at position 0 in the x-scale.

Each red vertical line in Figure 2 marks the location where Peakdet detected the closure instant, i.e. the beginning of each glottal cycle. The first values in column D are measured by the time between the first red vertical line and the start (at 0). The first value in column E is measured by the duration between the

second red vertical line and the first red line. The second value in column D is repeated by the previous value in column E because the end of a cycle is also the beginning of the following cycles.

In short, by measuring the time between the red vertical lines, which is the closing instant/starting point of each cycle, we obtain information in columns D and E.

Figure 3 is the derivative of the EGG signal (or DEGG signal) which allows us to more easily see the closing instant of each glottal cycle by a positive peak.

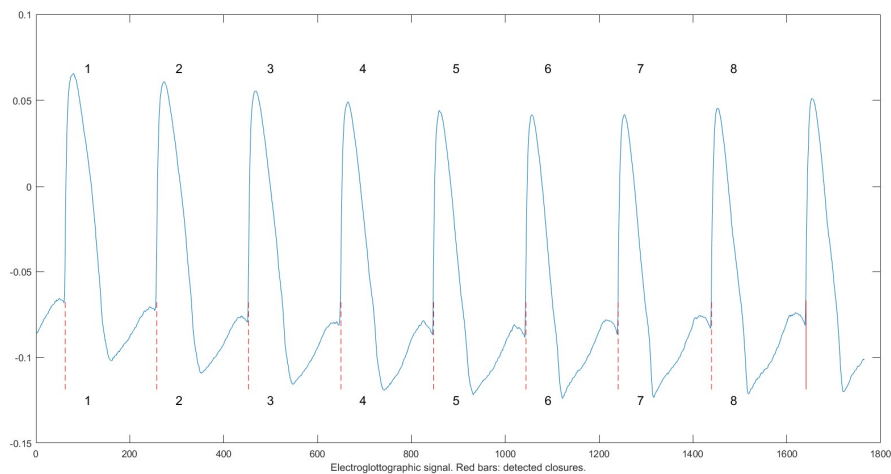


Figure 2: EGG signal displayed in PeakDet for the same syllable 0111 with the demonstration of the beginning of each glottal cycle by a red vertical line at the time of complete closure of each period.

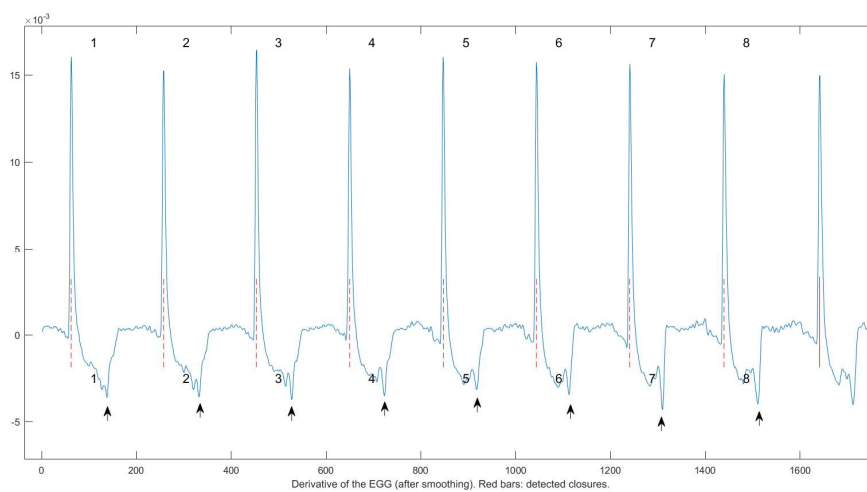


Figure 3: The corresponding derivative of EGG signal (DEGG signal) displayed in PeakDet for the same syllable 0111 with the demonstration of the beginning of each glottal cycle by a red vertical line at positive peak/opening peak (the time of complete closure of each period).

Explanation of f_0 and O_q measurement. (summarized in Figure 5)

Figure 4 shows the f_0 (top) and O_q (bottom) results calculated by Peakdet after automatic detection of the closing peak (positive peak) and the opening peak(s) (negative peak). The f_0 values (column F in excel) are calculated simply by reversing the time of each glottal cycle (i.e., reversing the duration calculated by the time in column E minus the time in column D).

The O_q results are calculated by 2x2 methods: (i) maxima on unsmoothed signal (column G in excel), (ii) maxima on smoothed signal (column H in excel), (iii) barycentre of peaks on unsmoothed signal (column I in excel), (iv) barycentre of peaks on smoothed signal (column G in excel). The opening peak(s) of each glottal cycle are not marked in the EGG and DEGG figures but can be retrieved by using the formula in Figure 5 if glottal cycle duration and oq values are available.

These calculations of the O_q values can be checked visually from the figure of the DEGG signal. After verifying the opening peak(s) in the DEGG signal to see which method reflects the most reasonable opening phase, I select the best method to store its results in the new column (which is put in the K column in the excel file) and suppress the values at the vague cycles with no precise opening peak (by setting the value(s) to 0). Imprecise opening peaks are the cases where: (i) there is not any clear peak present, (ii) there is more than one peak present and they are far from each other and it is not clear which one should be the real opening peak since none of them is a clear prominent peak.

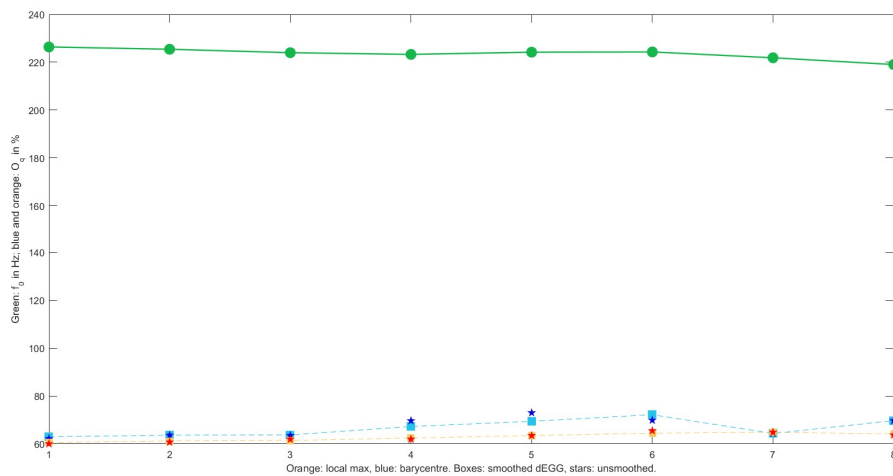


Figure 4: The results of f_0 and O_q measured (4 methods) by PeakDet

F₀ and O_q: two parameters calculated from EGG signal

F₀ (Fundamental frequency)

- Have **Pitch** as their perceptual counterpart

- High F₀ ---- High pitch
- Low F₀ ---- Low pitch

O_q (Open quotient)

- Relate to **phonation types**:

- Breathy ↑ - High O_q ---- flow phonation
- Modal - Medium O_q ---- modal phonation
- Creak ↓ - low O_q ---- pressed phonation

- Formula: (unit: Hz)

$$F_0 = \frac{1}{T}$$

(T: glottal cycle Duration)

- Formula: (unit: %)

$$O_q = \frac{(\text{Open phase})}{(\text{Open phase} + \text{Closed phase})}$$

(= T: glottal cycle duration)

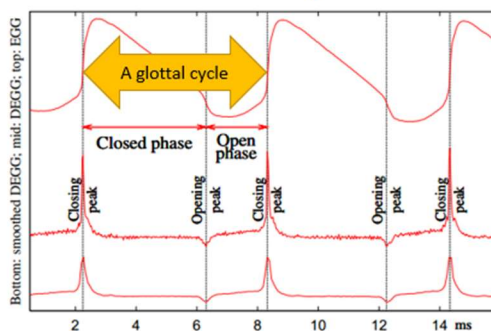


Figure 5: The visualization of the calculation of f₀ and O_q based on the EGG signal