Making a voice recording and evaluating the vocal frequency (pitch) and sound pressure level (loudness)

Dr Sean Pert

Evaluating Voice: Pitch, Loudness and Screening for Voice Disorder

- This is a **basic** step-by-step guide.
- This method is for **in-person, clinician-led only**, and not online recordings over the internet.
- Please see the recommended **tutorials** when you have gained confidence.

Pitch and Fundamental Frequency

- The fundamental frequency is the number of times the vocal folds open and close in one second. The number of cycles of opening and closing are referred to as *Hertz* (Hz).
    - The cismale range is *approximately* 100-140 Hz
    - The gender-neutral range is *approximately* 145-175 Hz
    - The cisfemale range is *approximately* 180-220 Hz
        - (Davies, 2015, p. 122)
- There is a great deal of overlap and individual difference between speakers.
- Pitch and frequency are *not* the same. Hirst and De Looze state that "(pitch)… corresponds to the subjective impression as to how voiced sounds, particularly sonorants and vowels, are perceived on a scale going from low to high, as on a musical scale of notes, while… (frequency) corresponds to the physiological parameter of the frequency of vibration of the vocal folds (aka vocal cords) measured in cycles per second (cps… measured in hertz (Hz)." (p. 336).

Voice intensity (Loudness) and Speech Pressure Level (SPL)

- Sound pressure level is expressed in decibels (dB)
- "SPL can be used to estimate the amount of acoustic energy and power radiated out of the mouth" (Švec & Granqvist, 2018, p. 442)
- Usually measured at 30 cm distance from the mouth to microphone.
- Measurements are sensitive to noise and affected by distance from the microphone.
- See Švec & Granqvist, 2018 for a comprehensive tutorial.

Making a voice recording and evaluating the vocal frequency (pitch)
and sound pressure level (loudness)
Dr Sean Pert

- Measures vary depending on the method employed, but broad ranges are listed below.

Table 1. Intensity of speech – types of speech

Type of speech	Vocal intensity range (dB)
Whisper	20-30 dB
Conversational speech	55-70 dB
Scream	80-125 dB

Table 2. Intensity of speech – Conversational Intensity Levels (CIL)
Second and third sentences of the Rainbow Passage recorded at 30 cm from the speaker.

Group	Mean CIL (dB)	Standard deviation	Range (dB)
Both genders	68.65	2.85	60.21 to 77.72
Cis males	69.53	2.77	60.59 to 79.00
Cis females	67.77	2.71	59.82 to 76.44

(N=20 males, 20 females, trial 1 data)
(After Pausewang Gelfer & Ryan Young, 1997)

Equipment
- A modern computer: Mac, PC or Linux computer
- A USB condenser mic, or headset mic with an amplifier with USB connection (if using a traditional mic connector)
- A quiet room.

Software
- **Praat**. Free to download here:
  https://www.fon.hum.uva.nl/praat/
- **Audacity**. Free to download here:
  https://www.audacityteam.org/download/
- **Visit www.speechtherapy.co.uk** and navigate the menu to 'Voice & Communication Therapy' ➔ 'Objective Voice Assessment' for links to all the software and plug ins required for these evaluations.

Suggested recordings
- **Sustained [ɑ:]** *for at least 6 seconds*
- **Rainbow Passage**
  https://www.york.ac.uk/media/languageandlinguistics/documents/currentstudents/linguisticsresources/Standardised-reading.pdf
- Consider asking the service user to sing '**Happy Birthday to You**' to hear pitch changes.

**Making a voice recording and evaluating the vocal frequency (pitch)
and sound pressure level (loudness)**
Dr Sean Pert

Suggested tutorial

- Curtis, J. (2022). *Acoustic Analysis of Voice*.
 https://www.jamescurtisphd.me/tutorials/voice/acoustic-analysis-of-voice [Accessed 3rd April 2023).

Part 1 – Recording a voice sample

Please note that Praat is capable of recording audio at high resolution (CD quality and above), but the program gives no visual feedback that the recording is being made. Audacity has the advantage of showing you the developing wave form on screen as the recording takes place.

Audacity also allows sections of speech to be highlighted, cut, and pasted into a new sound file. This is helpful if you want to trim onset noise or silence or select a specific spoken utterance (such as sentence 2 of the Rainbow Passage for CPPS analysis, for example – see below).

For this reason, Audacity is recommended for making sound recordings.

1. Launch **Audacity**.

2. Check that the high-quality microphone is selected:
   Audio Setup ➔ Recording Device ➔ *Microphone name,
   such as* RØDE NT-USB Mini
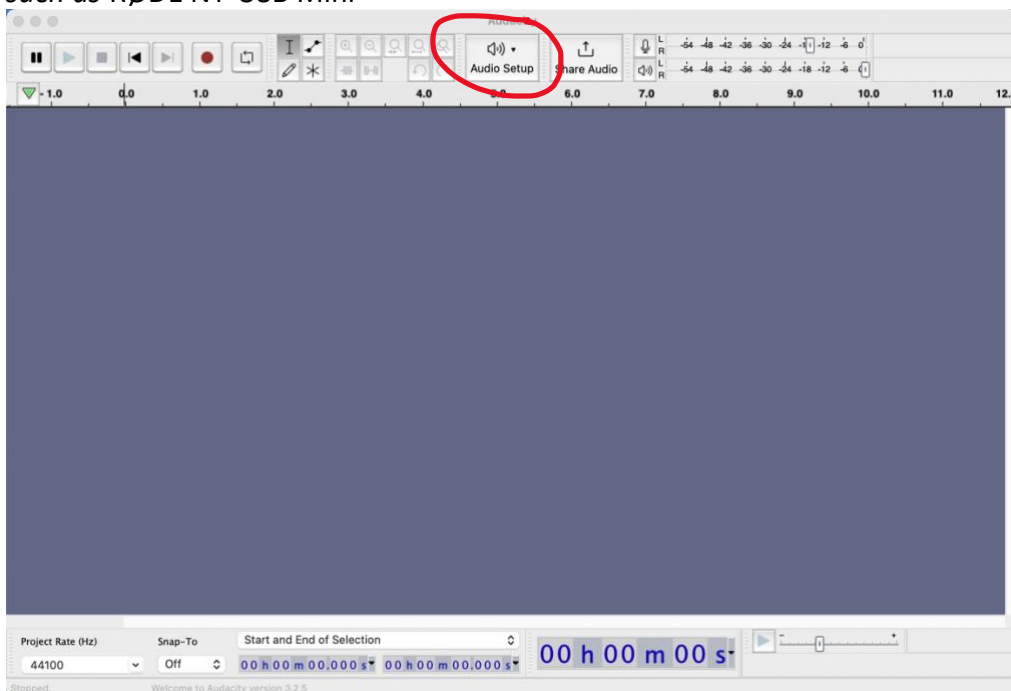
3. The recording should default to Mono, 441000Hz as the program detects your microphone. You can check this in **Audio Setup** and check that **Recording channels ➔ 1 (mono) Recording Channel** is selected. *You only have one vocal tract so mono is the best recording format*.
   In the unlikely event that only a stereo microphone is available, select mono by

**Making a voice recording and evaluating the vocal frequency (pitch)
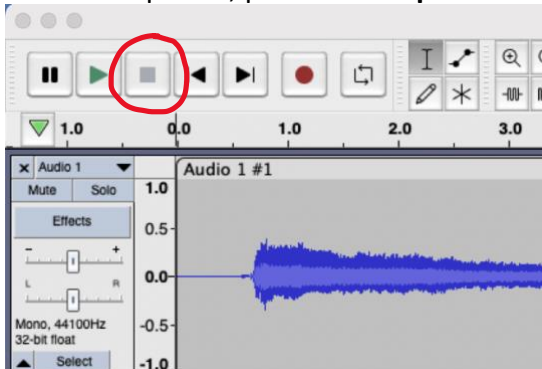and sound pressure level (loudness)**
Dr Sean Pert

choosing Audacity ➜ **Preferences** ➜ **Devices** ➜ **Channels** - Change the number of recording channels to **1 (Mono)**.

4. Press the red **Record** button

5. When completed, press the **Stop** button

6. Highlight the onset and press delete to remove the initial silence, **OR**
   highlight the required middle section (2-3 seconds) of the vowel or highlight the utterance you wish to analyse and then copy ➜ paste into a new file.
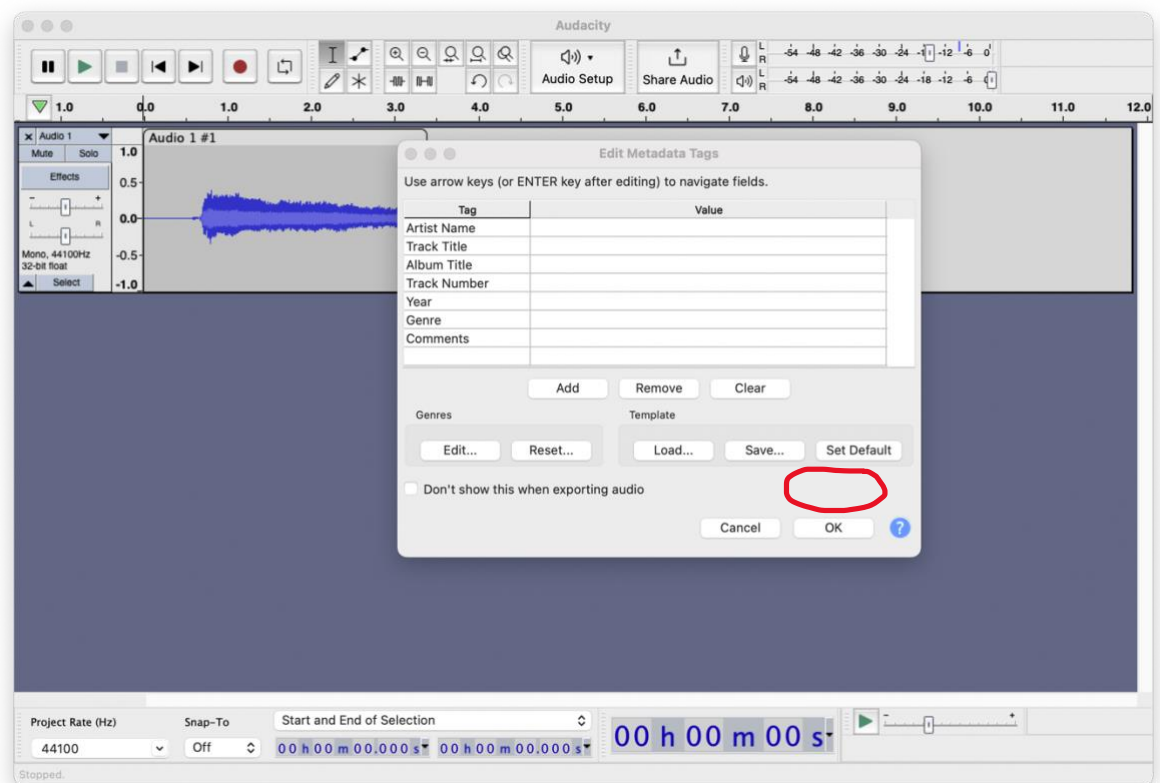

7. **File ➜ Export ➜ Export as WAV**

8. Type in a suitable filename and press **Save**.

Making a voice recording and evaluating the vocal frequency (pitch)
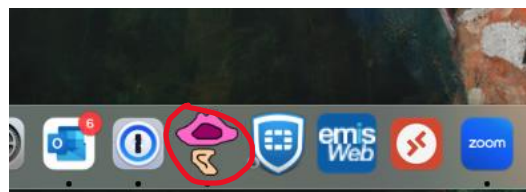**and sound pressure level (loudness)**
Dr Sean Pert

9. Ignore the 'Edit Metadata Tags' and just click **OK**. The recording will NOT save otherwise.

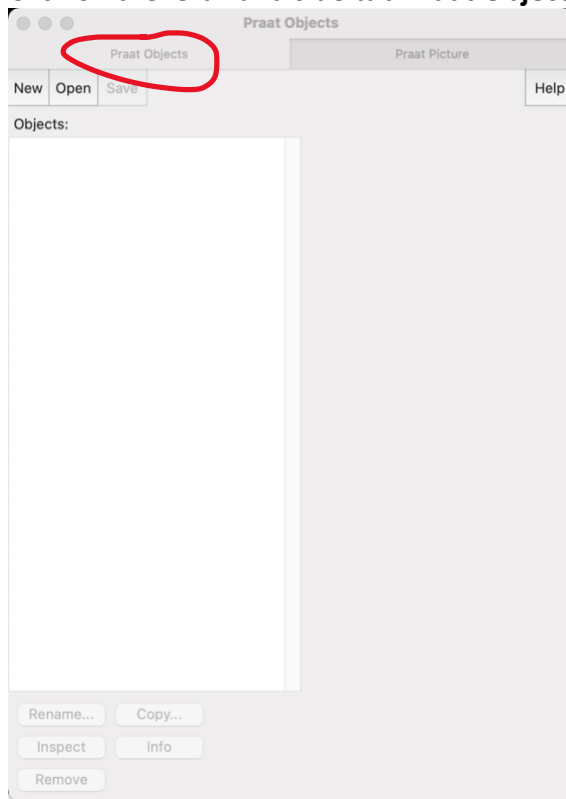## Part 2 – Showing the fundamental frequency (f_0) and loudness (dB)

1. Launch **Praat**. The windows Praat Objects and Praat Picture appear.

Making a voice recording and evaluating the vocal frequency (pitch)
and sound pressure level (loudness)
Dr Sean Pert

2. Click on the left-hand side tab **Praat Objects**.

3. Click **Open ➔ Read from file…** and select the recording you made in Audacity.
4. Click the filename in the windowpane on the left to **highlight** it.
5. Click the button **View & Edit**. A pair of graphs appear showing the waveform at the top and the spectrum below.

**Making a voice recording and evaluating the vocal frequency (pitch)
and sound pressure level (loudness)**
Dr Sean Pert

6. Click and drag the area you wish to analyse. Tip: this is usually just after the start of a vowel such as [ɑ:] to just before the end:
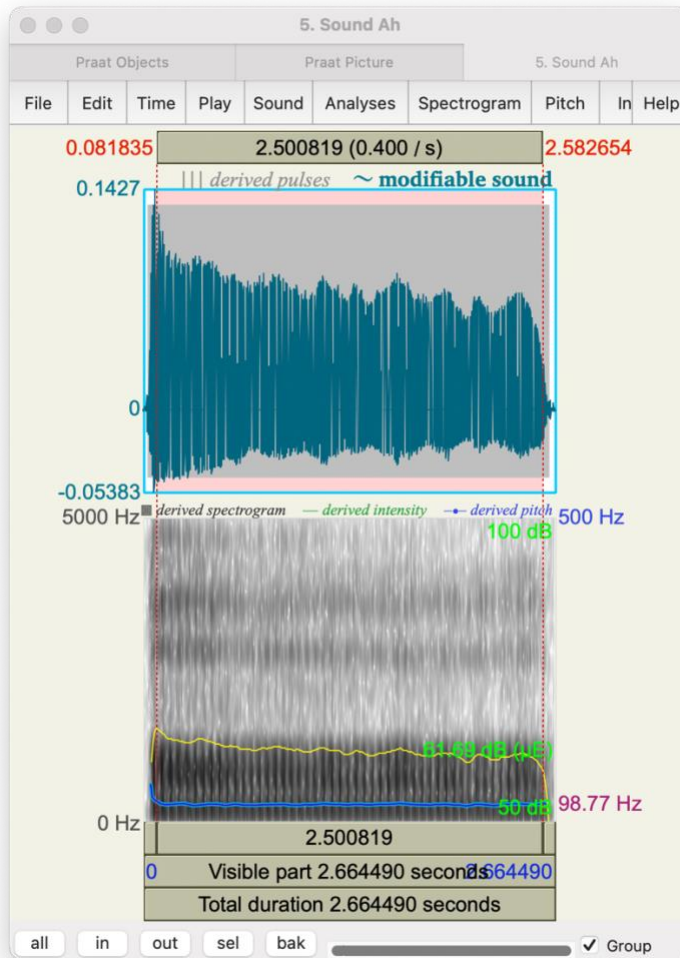
7. Click **Pulses ➔ Voice Report**. A table of pitch results is generated.
Tip: *If you cannot see the* Pulses *tab, make the window larger by dragging a right-hand corner of the Praat window.*

-- Voice report for --

Date: Mon Apr 3 16:05:48 2023

WARNING: some of the following measurements may be imprecise.

For more precision, go to "Pitch settings" and choose

the cross-correlation analysis method to optimize for voice research.

Time range of SELECTION

   From 0.081835 to 2.582654 seconds (duration: 2.500819 seconds)

Pitch:

   Median pitch: 98.778 Hz

   Mean pitch: 98.768 Hz

   Standard deviation: 0.740 Hz

   Minimum pitch: 96.805 Hz

Making a voice recording and evaluating the vocal frequency (pitch)
and sound pressure level (loudness)
Dr Sean Pert

Maximum pitch: 103.083 Hz

Pulses:

   Number of pulses: 247

   Number of periods: 246

   Mean period: 10.125833E-3 seconds

   Standard deviation of period: 0.078045E-3 seconds

Voicing:

   Fraction of locally unvoiced frames: 0   (0 / 251)

   Number of voice breaks: 0

   Degree of voice breaks: 0   (0 seconds / 2.500819 seconds)

Jitter:

   Jitter (local): 0.410%

   Jitter (local, absolute): 41.563E-6 seconds

   Jitter (rap): 0.230%

   Jitter (ppq5): 0.200%

   Jitter (ddp): 0.691%

Shimmer:

   Shimmer (local): 3.031%

   Shimmer (local, dB): 0.260 dB

   Shimmer (apq3): 1.696%

   Shimmer (apq5): 1.572%

   Shimmer (apq11): 1.983%

   Shimmer (dda): 5.088%

Harmonicity of the voiced parts only:

   Mean autocorrelation: 0.990479

   Mean noise-to-harmonics ratio: 0.009680

        Mean harmonics-to-noise ratio: 21.866 dB

8. Ignore the warning – the readings are accurate for voice evaluation purposes. If you prefer to get absolutely exact pitch listings, then use the **Pitch** tab and select in turn:
   a. Get pitch
   b. Get minimum pitch
   c. Get maximum pitch
   You will notice that, as long as you have edited out the onset and tail of the vowel, that the listing are the same as the above method. Values will differ if you highlight different parts of a spoken utterance (or the whole of the Rainbow Passsage).

9. With the area of the wave still highlighted click **Intensity ➔ Get intensity**. The intensity is listed as a figure in decibels (db), e.g.,

   61.74602564907431 dB (mean-energy intensity in SELECTION)

10. Copy and paste your results or save the text file and open it in Word to format the file and re-save as a Word and/or PDF document.

**Making a voice recording and evaluating the vocal frequency (pitch)
and sound pressure level (loudness)**
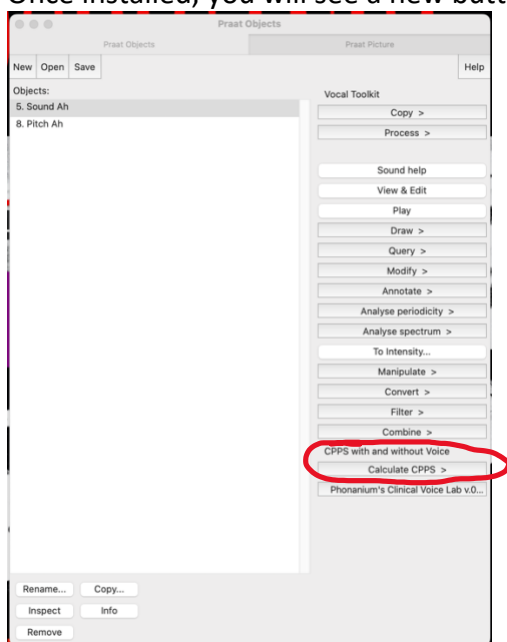Dr Sean Pert

Screening for voice disorder

If you suspect that a service user has a voice disorder, Cepstral Peak Prominence / Cepstral Peak Prominence (Smoothed) (CPP or CPPS) analysis is a quick and easy method of evaluating this with a high level of accuracy.

Murton et al. (2020) reported an accuracy rate of 94.5% in identifying the presence of a voice disorder. This confirmed findings by Saunder et al. (2017) who found that employing Praat, "…CPPS values alone have shown to be a reliable and valid way to measure dysphonia severity and are highly predictive of voice disorder status." (p. 565).

Murray et al. (2022) comment that "Cepstral peak prominence (CPP) is the recommended acoustic measure of dysphonia, replacing previously relied upon perturbation measures (jitter and shimmer)."

This method requires you to download and install the **Cepstral Peak Prominence - Praat with and without voice detection** plug-in by Elizabeth Heller Murray here: https://osf.io/t5hrv/ There are videos available on how to install the plug ins on the website.

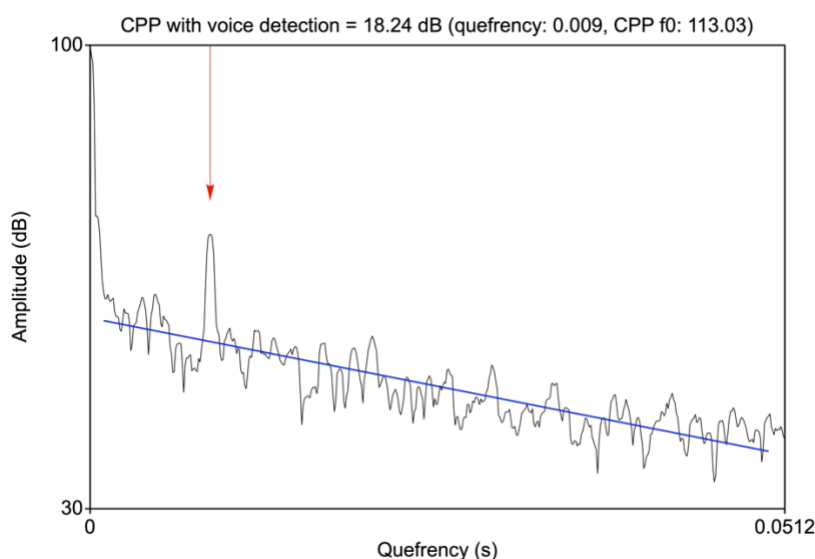Once installed, you will see a new button appear on the Praat menu:

**Making a voice recording and evaluating the vocal frequency (pitch)
and sound pressure level (loudness)**
Dr Sean Pert

To calculate Cepstral Peak Prominence – Smoothed (CPPS) values

1. Highlight the sound file for analysis.
    a. This should be a **Sustained [ɑ:]** *for at least 6 seconds*, **OR**
    b. The second utterance of the Rainbow Passage (edit this from your recording in Audacity, highlight the utterance and paste into a new file, then export to Praat as detailed above),

2. Click the **Calculate CPPS >** button.

3. Select **Single file**…

4. The **Run script: select CPP peak search range** window appears.
    The default values of:
    Minimum CPP peak search(default 60): 60
    Maximum CPP peak search(default 330): 330
    These values don't need adjusting – click **OK**.

5. **Wait!** It can take up to 30-45 second to generate a figure due to the complex mathematics involved. Your computer hasn't crashed!

    a. You will see a graph with the **CPPS peak**, and CPPS value figure highlighted with a red arrow.
    b. To see the numeric figures, click the next tab along in Praat labelled **Praat Info**.

6. An example graph and info for sustained **[ɑ:] vowel**. Note that the value of **18.24 dB** (and the CPPS_full value is also above the cut-off point at 17.71)is higher than the cut-off point of **14.45 dB** and therefore no voice disorder is indicated in this case.

CPP with voice detection = 18.24 dB (quefrency: 0.009, CPP f0: 113.03)

CPP_full (NO voice detection)   17.706870782972032

**Making a voice recording and evaluating the vocal frequency (pitch)
and sound pressure level (loudness)**
Dr Sean Pert

CPP_voice (WITH voice detection)   18.236719850232117

7. CPPS values: You will see two figures (if the graph is displayed, click the next tab along in Praat labelled **Praat Info**).
    a. If you have evaluated a speech sample using a sustained **[ɑ:] vowel**, then note down the '**CPP_full (NO voice detection)** figure (or highlight and copy and paste).
    b. If you have evaluated the second utterance of the Rainbow Passage, then note down the **CPP_voice (WITH voice detection),** (or highlight and copy and paste).
        Voice detection figures are for when there are voiceless segments or pauses in the speech recording, such as when connected speech is produced. The CPPS figure is therefore higher than without voice detection. You will notice that a new sound file is generated with only voiced segments and the voiceless segments removed.

8. You may wish to use **Praat Picture**, **File ➔ Save as PDF…** to export the graph with the CPPS peak value noted with voice detection.

**Making a voice recording and evaluating the vocal frequency (pitch)
and sound pressure level (loudness)**
Dr Sean Pert

9. Compare with the Voice disorder cut-off points:

Table 1. CPPS Cut-off points for Praat analysis

Speech sample	CPPS value should be **at or above this value**. *If the Praat CPPS value is LOWER, then there is a high probability that a voice disorder is present*
sustained **[ɑ:] vowel**	**14.45 dB**
Second sentence of the Rainbow Passage	**9.33 dB**

Values are WITH voice detection
(there should be no significant difference for sustained vowel as this is completely voiced)
(After Murton et al., 2020).

10. Include this information in your case notes and in any referral to Ear, Nose and Throat/Voice Pathology Speech and Language Therapy.

**Making a voice recording and evaluating the vocal frequency (pitch)
and sound pressure level (loudness)**
Dr Sean Pert

References

Davies, S., Papp, V. G., & Antoni, C. (2015). Voice and Communication Change for Gender Nonconforming Individuals: Giving Voice to the Person Inside. *International Journal of Transgenderism*, *16*(3), 117-159. https://doi.org/10.1080/15532739.2015.1075931

Hirst, D., & Looze, C. D. (2021). Fundamental Frequency and Pitch. In R.-A. Knight & J. Setter (Eds.), *The Cambridge Handbook of Phonetics* (pp. 336-361). https://doi.org/10.1017/9781108644198.014

Heller Murray, E. S., Chao, A., & Colletti, L. (2022). A Practical Guide to Calculating Cepstral Peak Prominence in Praat. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2022.09.002

Murton, O., Hillman, R., & Mehta, D. (2020). Cepstral Peak Prominence Values for Clinical Voice Evaluation. *American Journal of Speech-Language Pathology*, *29*(3), 1596-1607. https://doi.org/10.1044/2020_AJSLP-20-00001

Geifer, M. P., & Young, S. R. (1997). Comparisons of intensity measures and their stability in male and female speakers. *Journal of Voice*, *11*(2), 178-186. https://doi.org/10.1016/S0892-1997(97)80076-8

Sauder, C., Bretl, M., & Eadie, T. (2017). Predicting Voice Disorder Status From Smoothed Measures of Cepstral Peak Prominence Using Praat and Analysis of Dysphonia in Speech and Voice (ADSV). *Journal of Voice*, *31*(5), 557-566. https://doi.org/10.1016/j.jvoice.2017.01.006

Švec, J. G., & Granqvist, S. (2018). Tutorial and Guidelines on Measurement of Sound Pressure Level in Voice and Speech. *Journal of Speech, Language, and Hearing Research*, *61*(3), 441-461. https://doi.org/10.1044/2017_JSLHR-S-17-0095

Disclaimer

Although the methods described above are thought to be appropriate and accurately described, the author will not accept any responsibility for clinical decisions made using the above methods and tools. Objective measurements are not a replacement for clinical judgement and such measures should form part of your voice evaluation, alongside a case history, subjective voice evaluation measures and other aspects. Please see:

- https://www.rcslt.org/speech-and-language-therapy/clinical-information/voice/
- https://www.asha.org/practice-portal/clinical-topics/voice-disorders/