Appendix S3

$C_{<2k}$ interpretation and rationale

The initial rationale for introducing a dedicated metric of low frequency spectrum slope was to assess the degree of nonlinearity in the source, without the confounding influence of SB, which might have its own degree of freedom. Close inspection revealed that very distinct sex differences and training effects appear specifically in the range below 2 kHz. These contrasts are conventionally not resolved as a separate phenomenon, because the related spectral reshaping is not immediately classifiable as a source-only (slope) effect or a filter-only (resonance) effect. The C_{2k} metric supports both interpretations; and evaluating it over the VRP can inform us on which of these two effects is the more relevant. In this Appendix, we therefore examine the specific sensing qualities of this dedicated metric in more detail.

A spectrum centroid metric reports the frequency position of the centre of gravity of the spectral amplitude curve. When there is only one strong spectrum resonance peak, the centroid will generally coincide with the centre frequency of this peak. When the spectrum centroid is applied as a wide-band balancing metric, its variation is similar to that of the previously presented SB metric. Here, the centroid was measured over a limited frequency range only (0...2 kHz). This range includes f_{F1} and f_{F2} and contains the bulk of that spectrum energy which controls the SPL.

A spectrum centroid is usually calculated by dividing the sum of frequency-multiplied spectrum amplitude values by the sum of all spectrum amplitude values. This calculation scheme is important here, as it suggests an alternative interpretation of the centroid. The frequency scaled amplitude matches the spectrum of the derivative of the signal. So, filtering the signal with a +6 dB/octave high-pass filter (to obtain the signal derivative) and then taking the ratio of the output and input amplitudes will also provide a time-domain correlate of the spectrum centroid. This means that if the spectrum centroid changes, while the vocal tract transfer function remains unchanged, then the closing phase of the flow/source waveform must have become steeper (its derivative must have become more negative); a mere signal magnitude change would have affected equally both the signal and its derivative. A comparable proportional dependency is found with the definition of the so-called glottal amplitude quotient (AQ) [42], that weights a peak glottal flow by its peak derivative (the maximum flow declination rate, MFDR). The link is reciprocal: where the AQ characterizes a closure time, the $C_{<2\mathrm{k}}$ metric measures a frequency. Comparably, the AQ remains constant when the AC-flow and its derivative vary in linear proportion [42] (page 134). In the present study, the radiated (high-pass filtered) voice signal is evaluated, rather than the glottal flow signal. The same mechanism remains operative, but the centroid metric will report a higher frequency than for the flow signal, together with an increased sensitivity for the higher derivatives.

Note that, with inverse filtering of flow signals, the frequency content above about 2.5 kHz is typically ignored [34], or discarded by low-pass filtering [30][48]. Any metric that estimates the MFDR from such a band limited flow pulse can thus be expected to scale proportionally with any metric that estimates the spectrum roll-off over the same reduced bandwidth area, like the $C_{<2\mathrm{k}}$ metric. Therefore, the $C_{<2\mathrm{k}}$ metric should behave more like an inverse filtering-based MFDR metric than does the SB metric, which reflects on energy in a frequency range that is typically not included (or disregarded) in the MFDR assessment.

So, for the $C_{<2k}$ metric we foresee two scenarios; on the one hand, the $C_{<2k}$ metric may get stuck on or near to a dominating formant resonance, which then probably carries the SPL. If the $C_{<2k}$ metric seems not to be stuck, it reflects the bias of spectrum energy upwards in frequency *and/or* the relative steepness in the closing glottal flow.

With increasing f_o, the bandwidth over which the $C_{<2k}$ metric is evaluated will decrease, because the starting point of a harmonic series moves towards the 2 kHz bound. In the case of a flat spectral distribution, the $C_{<2k}$ metric will thus have an inbuilt tendency to move progressively upwards with f_o. From the point where there is only one harmonic left to carry all spectrum energy, the $C_{<2k}$ metric equates to f_o. With the vowel /a/ used throughout, there will be little spectrum energy above f_{F2}. f_{F2} can then be expected to be the limit for the highest $C_{<2k}$ frequencies that are seen. This is why, in the figures of the $C_{<2k}$ metric, the color scale stops at 1270 Hz, close to f_{F2}. In the graphs, it is the logarithm of the centroid frequency that is color-mapped. The color thus represents not the linear, but the proportional dependency on f_o and sound pressure. Dividing the $C_{<2k}$ values by f_o could notionally convert this metric to a correlate of the NAQ [49] and also compensate for the above tendency to shadow f_o. However, this transformation was not implemented; because the extra linear gradient on the logarithmic scale that this additional division by f_o would have introduced, tends to conceal the typical contrast that the different voice types exhibit in this regard.