# TALKPOWER: a guide to speech processing

**Whether it's a pile-up or just random noise that is spoiling reception of your signal at the far end, there are just two ways to improve your signal. One is to stick more ERP in the appropriate direction. The other is to use your existing ERP more effectively. Any other method, such as a change of mode or frequency, needs the cooperation of the other operator. If he can't hear you, he can't cooperate with you.**

By
Richard Lamont G4DYA

## Bandwidth and noise

Any electronic system for recording or transmitting information, of any sort, has two fundamental limitations on it. One is bandwidth, the other is noise. There is not a lot we can do about the bandwidth of amateur speech transmissions; we are stuck with using about two or three kilohertz in the case of SSB. As for noise, whether random, 'white noise' (hiss) or man-made interference, there are a number of things that can be done at the transmitting end to improve the signal at the receiver output. One way is to improve the signal-to-noise ratio of the RF chain, from modulation in the transmitter to demodulation in the receiver, by increasing transmitter power or by building better aerials.

The only other way of sending out a better signal is to change the speech waveform, if possible, so that it is more intelligible for a given peak-to-peak level. Fortunately, speech has a number of quirks that allow this to be done.

## Speech waveforms

If, as I imagine most *HRT* readers know, you look at a speech waveform on an oscilloscope one of its most striking characteristics is the existence of high-amplitude peaks with relatively little in between. These peaks tend to be produced by vowel sounds, with consonants forming only brief, low-amplitude transients. But it is the consonants that carry most of the information. If the level of the consonants can be boosted, without increasing the existing peaks, then the signal will be much easier to understand at the receiving end.

In fact, the peak-to-average ratio of speech is typically about 14dB, and can be as high as 20dB. The consonants may be 30dB below peak level. The exact figure depends on the many factors that make up the sound of an individual's voice — sex, age, language, accent, personality, mood etc. (Even for a given individual the ratio can vary a lot — people's voices go up and down a lot in both level and pitch when they are excited, much less if they are relaxed.)

As well as being peaky, speech waveforms can be asymmetrical, with positive and negative peaks differing by up to 8dB.

The other important characteristic of speech is its frequency

**Speech before (top trace) and after clipping (lower trace). Note how the low level transients are boosted in between the large peaks. This doesn't happen with compression**