The Intonator

L.F. Willems

The Intonator, as we have termed this instrument, is a tool for studying the perceptual tolerances of intonation contours. From natural input speech only the intonation is changed according to rules specified by a certain model ('t Hart, 1966).

Use is made of the Vocoder principle, known in analysis-synthesis telephony (Cooper, 1961). Actually, the most difficult part of a vocoder, the pitch extraction, is avoided and at the synthesis side the pitch movements are controlled by a function generator.

Block diagram

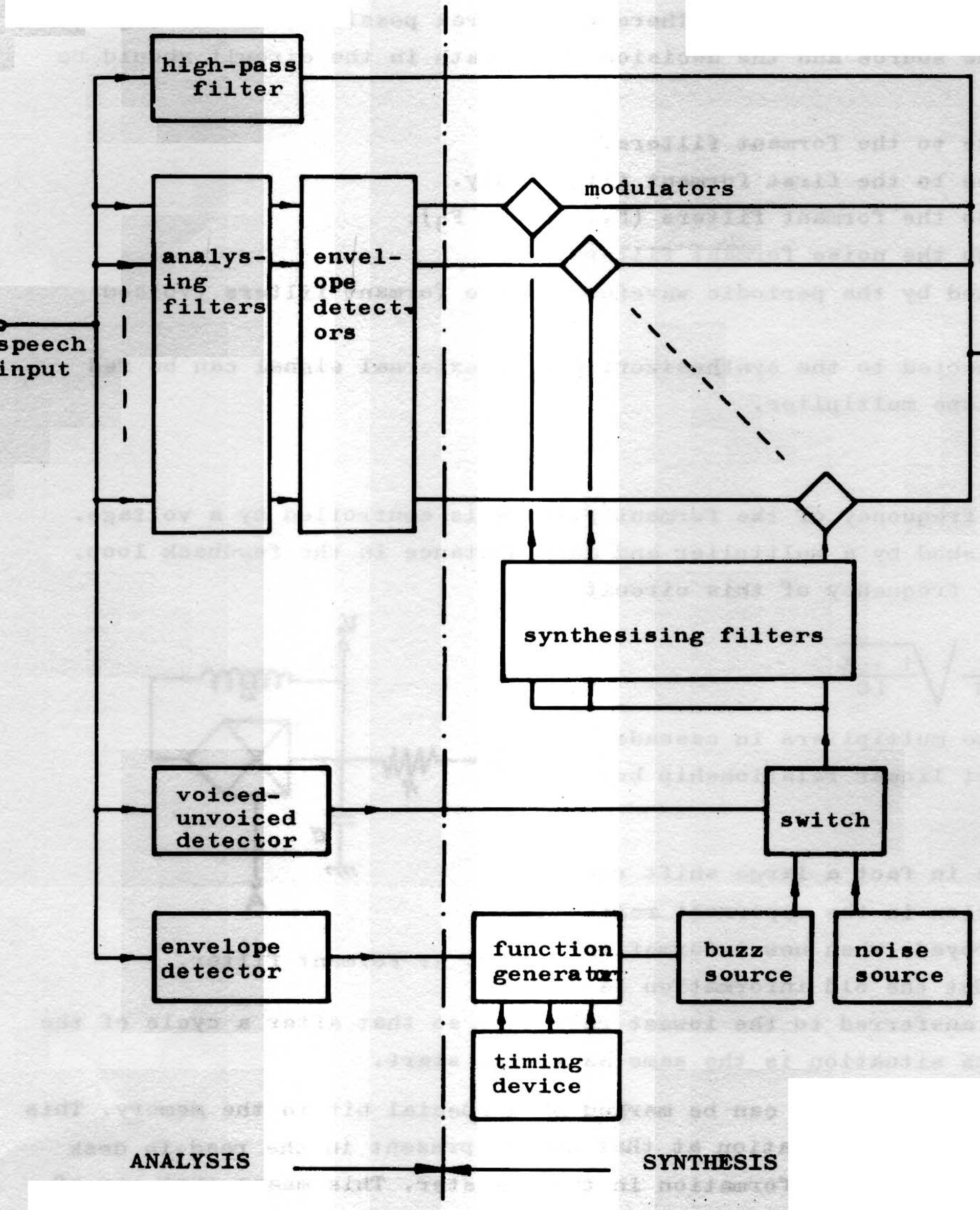

Fig. 1: Block diagram of the Intonator

The sprectrum of the input speech signal is analysed in a set of 25 band-pass filters (simple resonant parallel LRC-circuits), covering a range of 250 to 4,000 Hz. The resonant frequencies of the filters are a full tone apart and the filters have a quality factor Q of 16. The output of each filter is connected to an envelope detector. At the analysis side of the system there is also a voiced-unvoiced detector, using the average number of zero crossings as a criterion (Fant, 1962). At the synthesis side the original spectrum of the input speech is reconstructed. A buzz source or noise source is switched on to the synthesizing filters by the voiced-unvoiced signal. The signals from these filters are multiplied in a set of modulators by the corresponding envelopes. This re-made speech is monotonous. By controlling the periodicity of the buzz source with a special function generator one is able to give the speech output specified pitch movements.

Control of the Fo

The control function of the Fo consists of two elements, called macrointonation and microintonation. They will be described in this order.

In the macrointonation function several elementary patterns can be generated. First there is a gradual sloping down during the whole utterance, and this is called the declination. At a certain moment (governed by a manual setting on a timing device, which is a preset counter) a rising function can be generated. At another moment a falling function can be produced. These elementary patterns, rise and delay, can be generated repeatedly. These ramp waveforms are shaped by integration.

The microintonation is derived from the envelope of the input speech (see fig. 2).

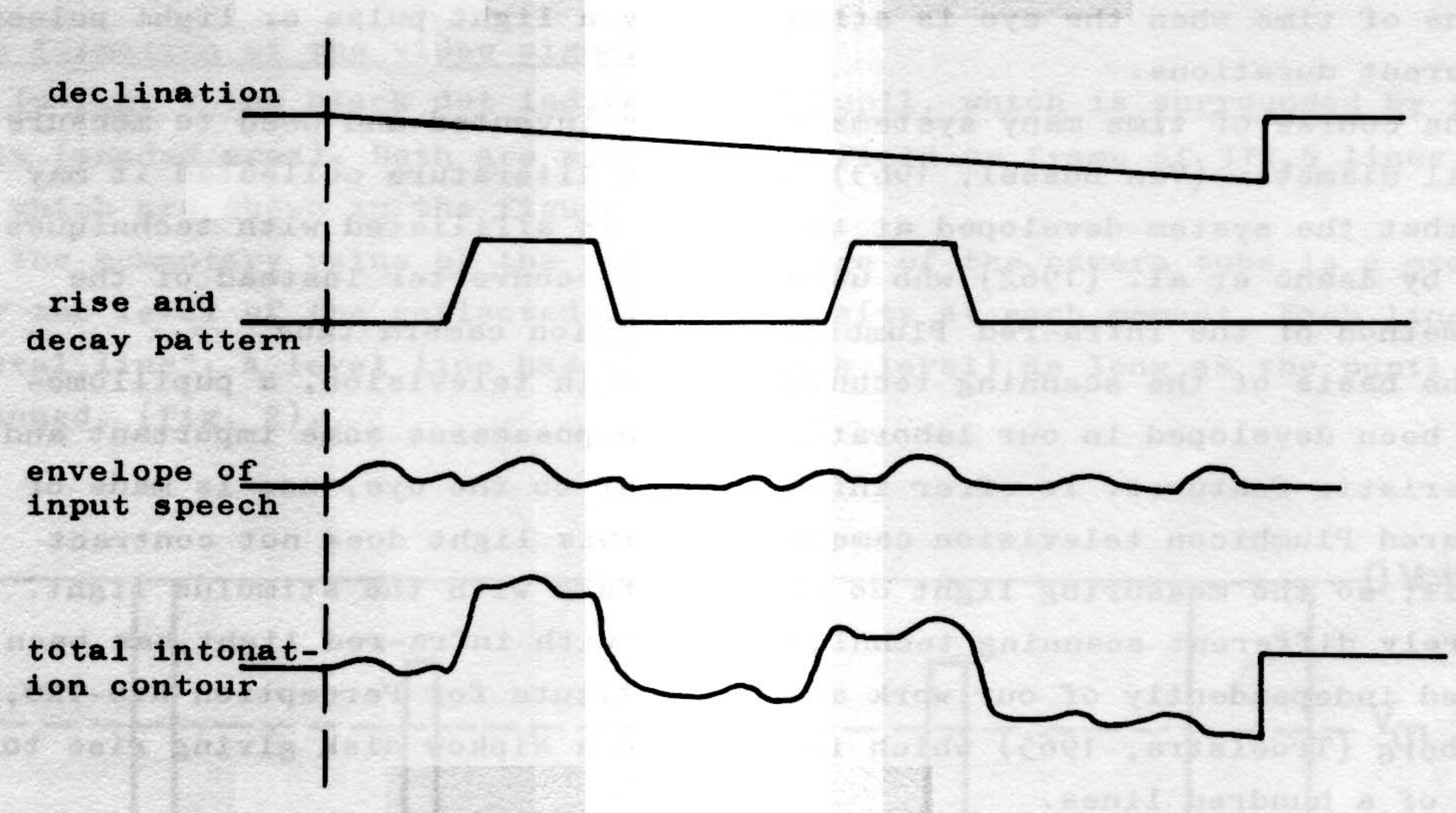

Fig. 2: Intonation patterns

The concept of microintonation is an old one in our laboratory, and was used originally to give some pitch inflection to synthesized words. It was found that naturalness of synthesized utterances improved greatly by controlling the Fo with a certain fraction of the envelope of the speech segments.

A useful trick is applied to improve the intelligibility of consonants. The high-pass (2,500 Hz) filtered part of the input spectrum is added to the output. The spurious pitch information of the input speech in this part of the spectrum is totally masked at the synthesis side by the new pitch information in the lower part of the spectrum. This was found by R.J. Ritsma in his experiments on pitch perception (oral communication).

References

Cooper, F.S. 1962 Speech Synthesizers. Proceedings of the 4th Int. Congr. of Man. Sci. Helsinki 1961. The Hague 1962, p. 3-13.

't Hart, J. 1966 Perceptual Analysis of Dutch Intonation Features. This issue.

Fant, G. 1962 Speech Analysis and Synthesis. Report No. 26, The Royal Institute of Technology Stockholm, Sweden.