SPEECH FORMANT GENERATOR

A.G. Wempe and A.W. van Maanen

1. INTRODUCTION

According to the widely accepted speech production model, voiced speech can be thought of as the result of a source (the glottal pulse) activating a linear filter network (the vocal tract).
The relatively steep edge of the source waveform causes the source to be a wideband frequency generator so that the frequency contents of the output are 'modelled' by the vocal tract filter function and maxima of this function correspond with maxima (i.e. formants) in the frequency distribution of the output.
In different terms: within each period of the fundamental frequency the steep edge of the glottal pulse excites the vocal tract cavities which produce damped oscillations in their natural frequencies.
During each period of the fundamental frequency the output can be regarded as the composition of a number of damped sinusoids or formants.
In practice, many deviations from this ideal model occur; however, commonly applied analysis and synthesis techniques are based on this concept and parts of speech can be conveniently discribed by using parameters in accordance with this model.
Therefore, a generator producing an output which consists of several damped sinusoids with variable parameters can form a valuable tool in the speech research field.
When designing such a generator two different targets must be distinguished; one of them being the realisation of a generator which produces a vowel signal, sounding as natural as possible, the other being the construction of a device which produces a signal corresponding with the function of a simplified model of the speech production.
We entered the latter concept because of its universal applications such as:
generating simple signals for perception experiments;
testing and judging various signal analysis methods;
practical instruction aid for students, etc.

2. WORKING PRINCIPLE

Formerly, a generator producing damped oscillations as formants was made using the principle of the excitation of active resonating circuits by a sawtooth voltage.

The main disadvantages were:

low damping settings of the formants could easily lead to spontaneous oscillation;

implementation of variable initial phase angle was discarded because of its elaborate technical necessities;

the parameter adjustments were not completely independent of each other;

the formant frequencies couln't be displayed very accurately.

In the speech formant generator to be described next, these disadvantages are overcome by making use of a completely different principle which on the whole can be explained as follows:

In a ROM (read only memory) the binary equivalents of consecutive samples of a sine wave are stored.

The contents of this memory are scanned repetitively so that the sequential outputs form digital samples of a continuous sine wave.

During replay of the memory contents the outputs are converted into corresponding analogue voltages and at the same time multiplied by an exponentially decaying waveform, thus producing a damped sine wave.

The frequency of this 'formant' can easily be varied by altering the memory scanning rate.

This system allows the damping to be variable over a very wide range including zero without instability.

For each formant to be generated, a similar arrangement is made so that all formants can be individually attenuated and added together.

The damped sine waves are all started at the same moment.

Periodic starting of the damped sine waves provides for a continuous 'vowel' sound whereby the F_o can be made variable without exerting influence on the 'formant' parameters.

Controlling the initial phases of the damped sinusoids can be arranged quite easily because of the storage in memories of the sine waves: the scanning of a ROM can be started with the sample which corresponds with the selected phase angle.

If n formants are generated, each period of the output voltage can be defined by the function:

$$V_o = \sum_{k=1}^{n} [A_k.\exp(-\alpha_k t).\sin(\omega_k t+\phi_k)]$$

where A_k is the amplitude of the k^{th} formant

α_k is the damping of the k^{th} formant

ω_k is the angular frequency of the k^{th} formant = 2 .F_k

ϕ_k is the initial phase angle of the k^{th} formant.

3. REALIZATION

3.1. General

The parameters are controllable by means of continuous variable controls. Compared with digital controls, like thumbwheel switches, continuous variable controls render the possibility of adjusting the various parameters in a more convenient way when adjustment by ear is required. The parameter values can be measured electronically and read out on a digital display so that the obtained data can be more accurate than is the case when using dials.

3.2. Formant generation

For each formant a circuit identical to fig.1 is needed and can be described as follows:

The output of an oscillator with variable frequency is fed through a digital input selector to the clock input of a 6-bit binary counter. The first 5 bits of the counter output form the address code for the 32x8 bit ROM, containing 32 consecutive samples of a half sine wave. The digital 8-bit output of the ROM is converted into an analogue value by the multiplying DAC.

As the counter proceeds, the 6^{th} bit of its output changes after each scanning cycle of the 32 samples of the ROM.

Dependent on the output of this 6^{th} bit, the analogue output of the DAC is inverted or not inverted by switching an amplifier into the inverting or non-inverting mode.

The result is a sine wave which is sampled 64 times per period.

After low-pass filtering for removal of the high-frequency components due to the sampling process, the output is led through an amplitude

control to an amplifier which provides for a low-impedance output. The reference voltage for the multiplying DAC is derived from an exponential decaying voltage generator producing the function A.exp($-\alpha t$) of which α is made variable.
The generation of this function is based on the principle of discharging a capacitor through a resistor.
The damped sine wave thus generated which is available at the formant output can be started again by a trigger command which resets the counter, the exponential function generator and the variable oscillator so that the next damped sine wave starts from the beginning, i.e. zero phase angle.
The frequency of the 'formant' can be varied by altering the frequency of the variable oscillator (the formant frequency is equal to the variable oscillator frequency devided by 64).
The initial phase angle is made variable as follows:
Prior to each damped sine wave a small amount of time (100 μ sec) is reserved during which the variable oscillator is inhibited.
Part of this time (controllable by a variable one-shot) the digital input selector connects the clock input of the counter with the output of a central oscillator with a fixed frequency (640kHz).
The result is a temporary high frequency scanning of the ROM so that after the reserved time the 'authorized' low frequency scanning, controlled by the variable oscillator, starts at an addres which is selectable by the variable one-shot.
The number of pulses within one 'injection' ranges from 0 to 64 so that the initial phase angle can be varied over a complete period.
During the reserved time of 100 μsec the ROM-output is disabled causing the output voltage to be zero.
This 100 μsec gap in the output affects the spectral contents of the signal only noticably beyond 10 kHz so there is no problem for the frequency area of interest (0-6 kHz).
The cut-off frequency of the low-pass filter is made proportional to the clock frequency so that there is no change of amplitude as the formant frequency is altered.
To perform this dependency the cut-off frequency control and F_k-control are mechanically coupled.

3.3. Timing

A common control circuit (see fig. 2) provides for the triggering pulses for initiating all damped sinusoids simultaneously and for the high frequency generator.

The formant triggering pulses can be generated by a F_o-generator with variable F_o or by hand with the aid of a push-button.

Additionally, triggering from external equipment is possible.

The F_o-generator consists of two mutually coupled one-shots, one of them being variable, the other generating the 100 µsec pulse.

In case of manual or external triggering, the variable one-shot is disabled and only the 100 µsec one-shot is activated.

The 100 µsec pulse is available for triggering an oscilloscope for signal monitoring.

Enabling/disabling the ROM can be externally controlled so that specific parts of the signal can be rejected.

3.4. Outputs

For each formant a connector is installed, providing for individual low-impedance formant output voltages (see fig. 3).

The composite output is obtained by adding 4 formant voltages using a summing amplifier.

Its low-impedance output is available at a line output connector.

For monitoring, a power amplifier and loudspeaker are built-in.

3.5. Display

As the realization of the display unit has been postponed and still has to be constructed, only the main concept is described here.

A common central digital display unit with 4 digits can be used to measure any parameter.

Beside each parameter control a touch button is installed with which the value of the parameter can be connected to the display unit, cancelling the previous connection.

The touch buttons have built-in signalling lights (LED's), showing which parameter is being displayed.

The formant frequencies, phase angles and fundamental frequency can be measured directly by counting the pulses from the various generators during specific intervals of time.
When phase angles are measured, counting is enabled only during the 100μsec intervals; when formant frequencies are measured, counting is enabled only during the remaining intervals.
When, in case of measuring a formant frequency, the effective counting time has reached 1/64 sec (which actually takes slightly more time), the number of counted pulses can be displayed to produce the formant frequency in Hz.
As the F_o-generator is free-running, the counting time for displaying the fundamental frequency in Hz is 1 sec.
An active counting time of 5.625 msec for the phase angle measurement results in a display resolution of .1 degree. (The adjustment resolution however, is 5.6 degrees as the sine-wave is devided into 64 samples).
For measuring the dampings and formant amplitudes, which are accessable only in analogue values, a voltage-to-frequency convertor is needed.
In the exponential function generator a capacitor C which is charged to a voltage A is discharged through a variable resistor R, thus producing a waveform, proportional to A.exp (-t/RC).
The maximum current through the resistor is equal to A/R.
As A is constant, this maximum current is proportional to α.
The value of the current is led to a peak current rectifier which produces a voltage proportional to the peak current.
The voltage-to-frequency convertor enables measuring this voltage by the digital display system.
With a proper scaling factor a calibrated display can be obtained.
The amplitudes can be measured by disconnecting the amplitude control from the formant generator output during the 100 μsec intervals and connecting it with a fixed voltage.
The output of the amplitude control is then coupled with the input of the peak rectifier and further processing occurs as in the case of measuring the damping.
A proper scaling factor provides for a display in mV.

FORMANT GENERATOR

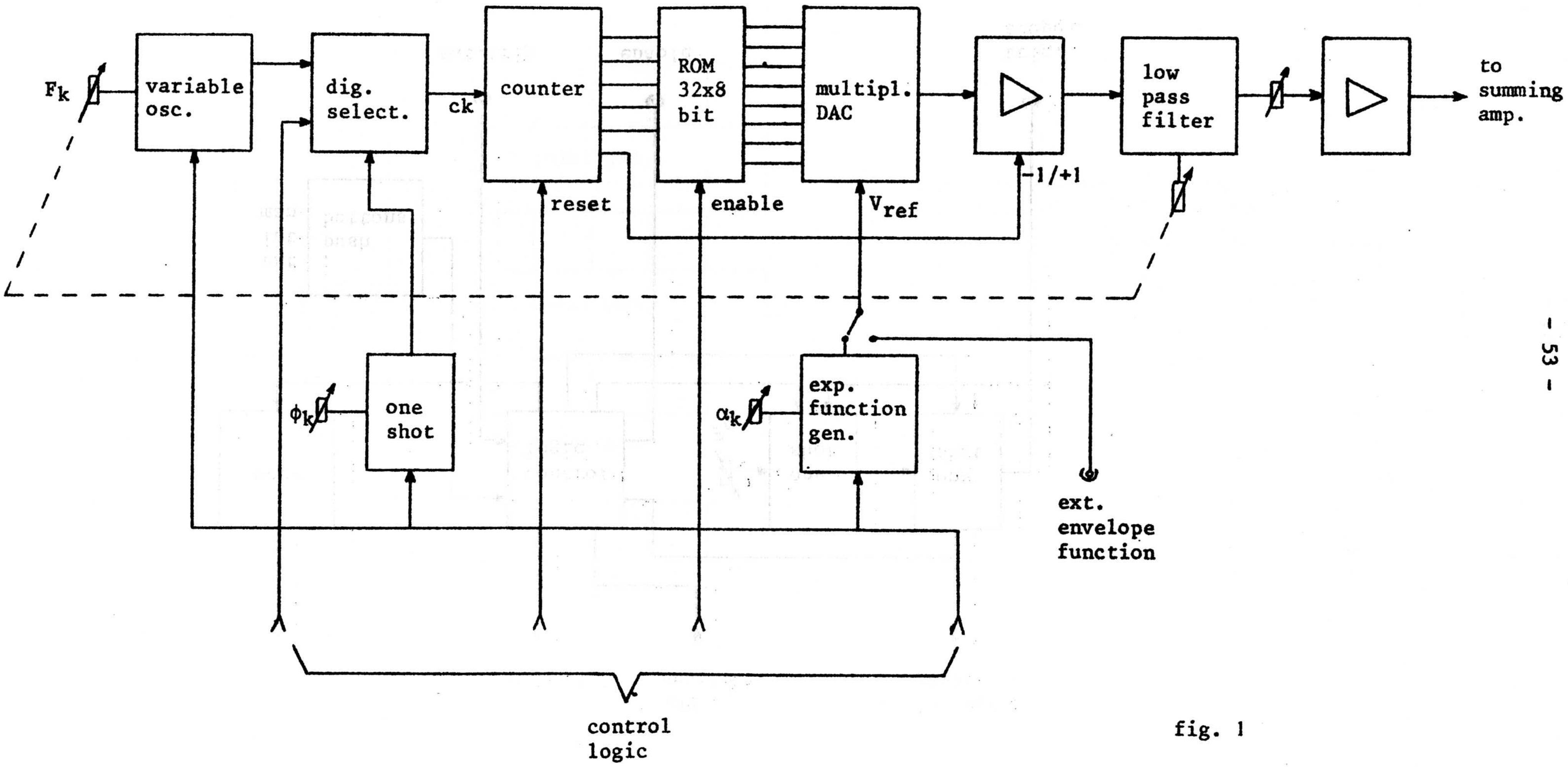

fig. 1

CENTRAL CONTROL UNIT

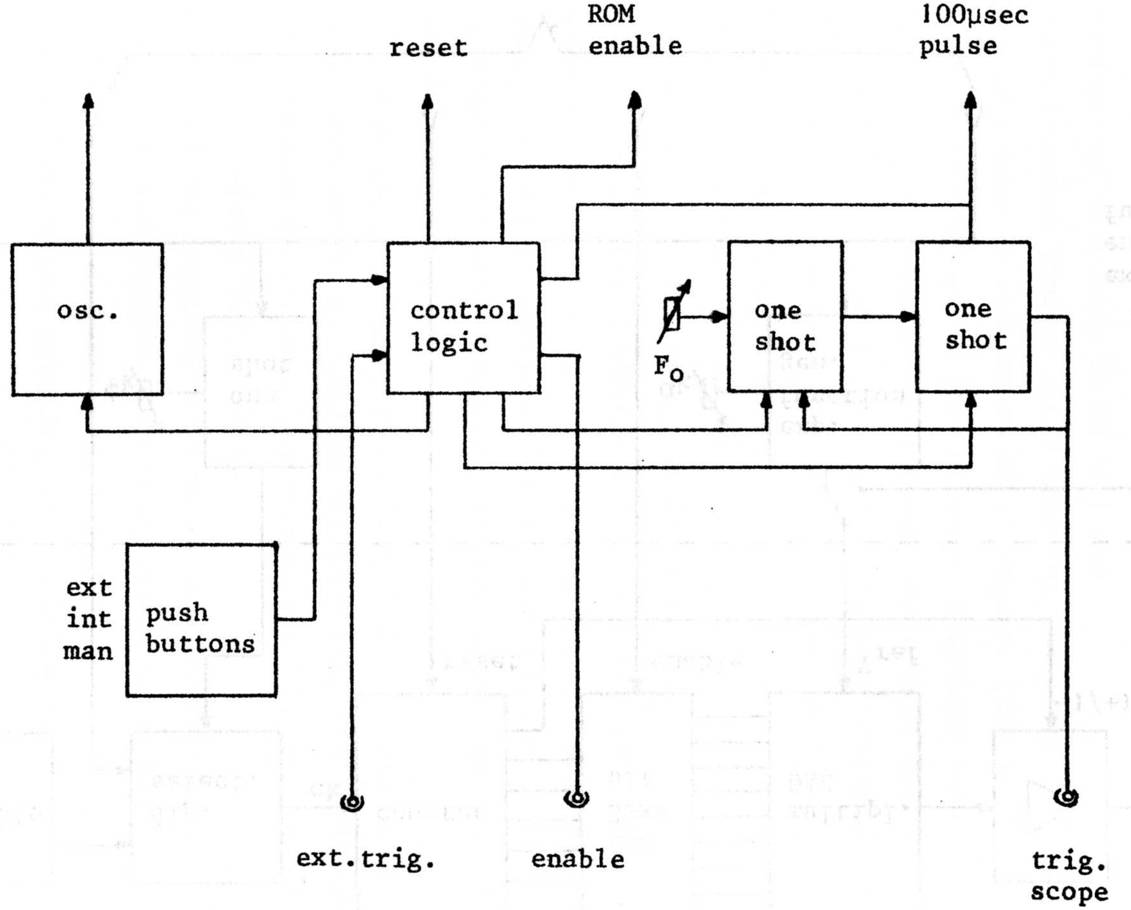

fig. 2

SIMPLIFIED DIAGRAM OF THE COMPLETE SYSTEM

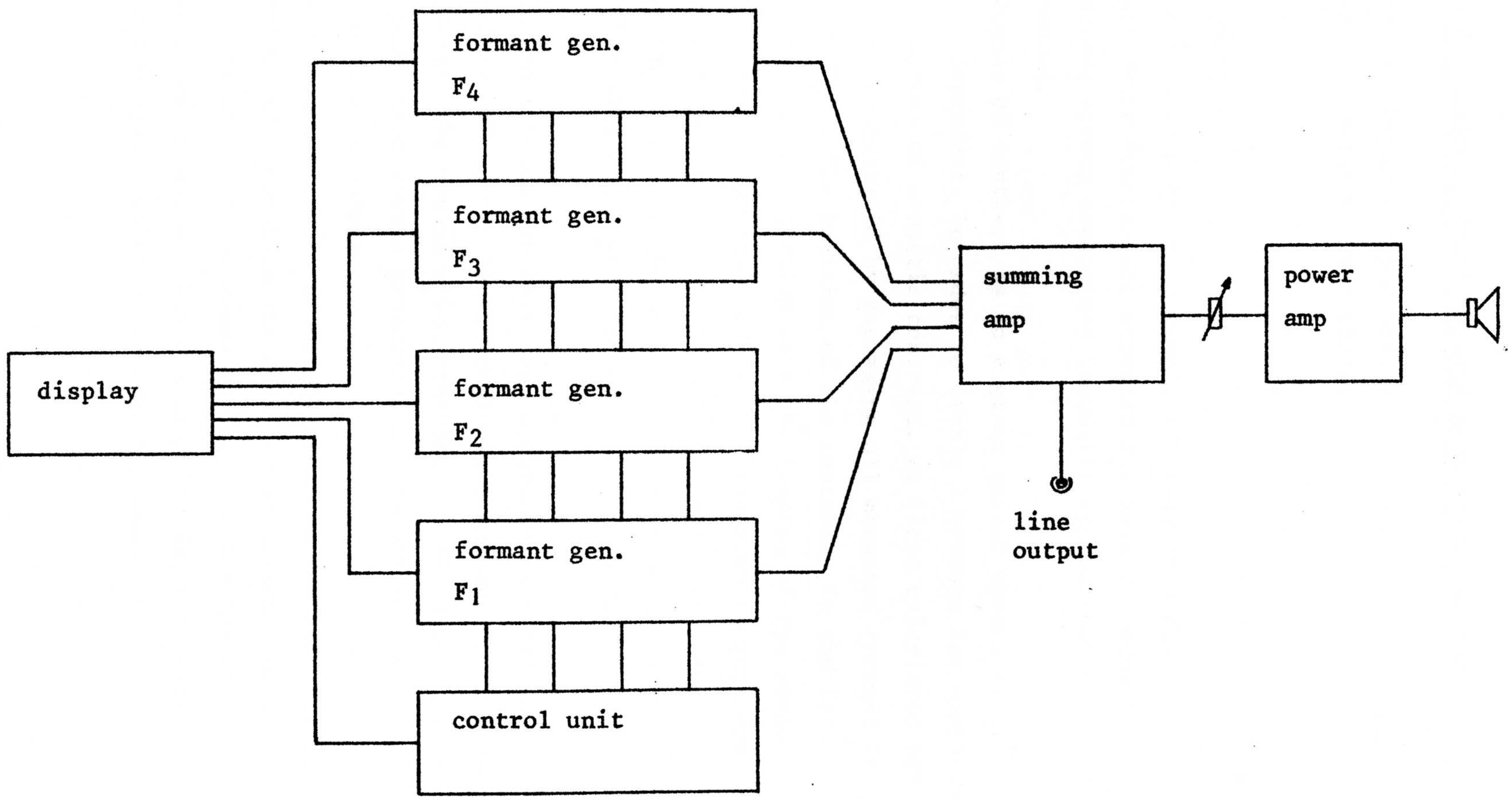

fig. 3

PROCEEDINGS

FROM THE INSTITUTE OF PHONETIC SCIENCES
OF THE UNIVERSITY OF AMSTERDAM

6

AMSTERDAM 1981