A SYSTEM THAT TRANSFORMS THE SPEECH SPECTRUM FOR THE PARTIALLY DEAF

Ferial T. El-Mokadem * and Robert W. Newcomb **

* Fairchild Space and Electronics Company, Germantown, Maryland 20767
** University of Maryland, College Park, Maryland 20742

ABSTRACT

This paper contributes to the area of special alteration of the speech signal to overcome high-frequency hearing loss.

A system is proposed and described which transforms the speech spectrum into a lower frequency range while at the same time providing good representation of its information content. The design approach is based on analyzing the speech signal into its harmonic components, dividing these by a constant, and then recombining the new components into a band - compressed signal, while preserving the amplitude spectrum of the original speech.

I. INTRODUCTION

Several attempts to develop special aids for audiometric correction of the high-frequency hearing loss have been recorded in the literature [1-4] . In these systems, the compensation of high frequency loss of hearing has been achieved by special alterations of the speech signal in order to make use of any low-frequency residual hearing. Two categories of alterations appear to be reasonable. In one category, only elements in the speech signal, which are impossible or very difficult to perceive with the type of hearing loss in question, are translated to low-frequency signals. In this case, the frequency transposition is meant especially to treat the fricative consonants which contain mainly high frequencies. Increasing distortion due to frequency overlap of the vowel phonemes and frequency-translated fricative phonemes is very likely to exist in such systems. In the second category, the whole speech spectrum is transposed into a lower frequency range, thus a simultaneous compression is being applied to both voiced and unvoiced speech spectra. The effect of such an operation is to convert normal speech into a reduced bandwidth synthetic speech. Intelligibility of synthetic speech permits frequency compression by factors of two or three. However, the perception of a synthetic speech associated with reduction of bandwidth requires training. This has been pointed out by Pimanow's hearing aid [4] which operates on the principle of the Vocoder, a device which reconstructs speech rather than transforming it as we propose .

Aiming to develop a device that may help the deaf of high-frequency loss, we presently propose the new "PLL frequency transformer". This system performs a bandwidth compression over the whole speech spectrum. The compression is realized by dividing the individual harmonic components of the speech signal by a constant factor. The magnitude of the compression factor can be any desired integer or rational number (an inherent limitation to compressing voiced speech by a factor of 4 has been indicated by Knorr [1]). In this system, as will be shown next, the monolithic phase-locked loop (PLL) has been utilized to operate both as selective band-pass filters, thus separating the harmonic components of speech signal, and frequency selective frequency dividers to obtain the desired frequency transformation. Therefore, no pitch extraction apparatus is needed for the compression operation since the compressed speech is produced directly from the original one.

II. THE PLL FREQUENCY TRANSFORMER

It is well known that the spectrum of a voiced sound can be described in the form of N carriers which are separated by a frequency interval approximately equal to the pitch frequency, that is, at least 100 Hz [5,6]. Moreover, in continuous speech, each of these carriers is modulated in both amplitude and frequency with a rate that ranges from 20 to 40 Hz. It is precisely this description using N carriers that forms the basis of the PLL Frequency Transformer (PLL-FT). Figure 1 shows a schematic diagram of the PLL-FT. In this system we assumed 38 analyzing channels, each being 90 Hz - wide, thus covering a 100-3500 Hz - input speech. In general, the number of channels and frequency-band covered by each may vary with the intended media of application. For instance, in the case of high pitched female voices, a lesser number of channels with larger bandwidth covered by each can be adopted in the system design.

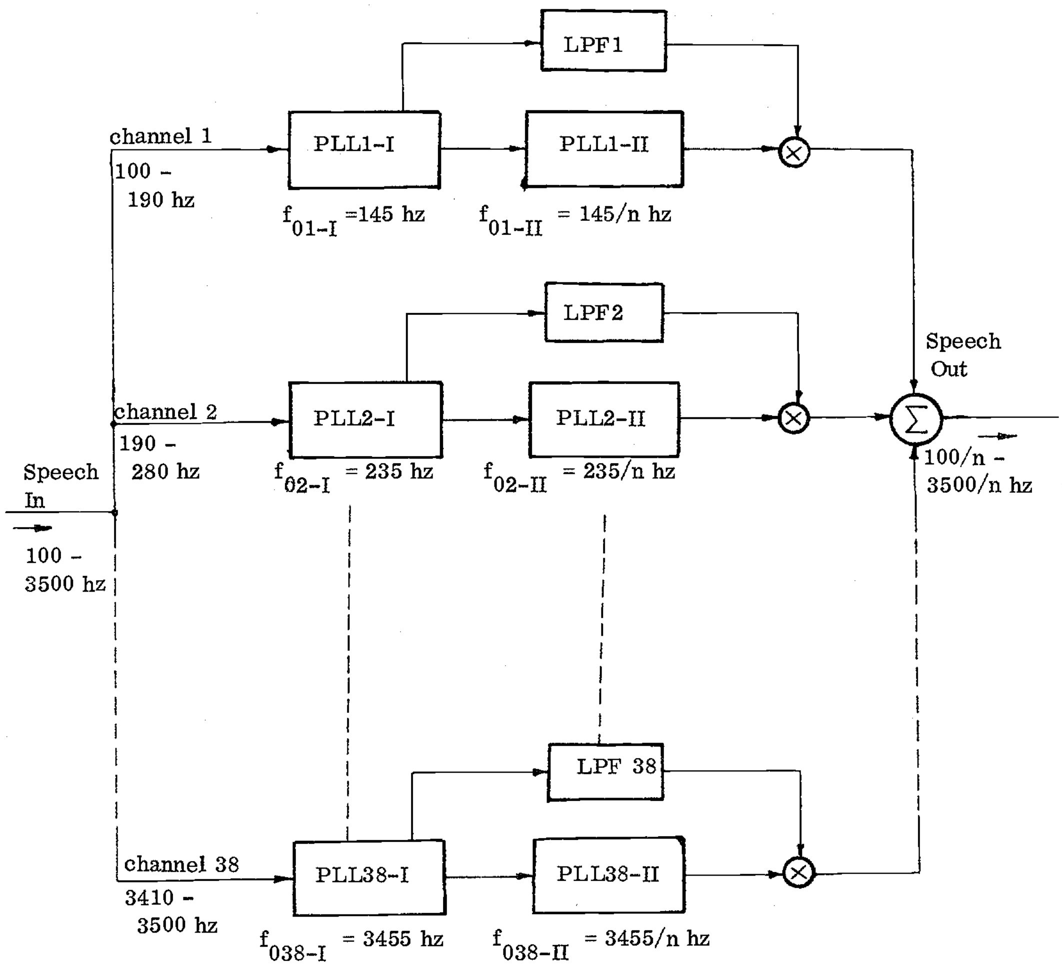

Fig. 1. Block diagram of the PLL frequency transformer.

When the system is excited by an input voiced speech from the microphone, the first PLL in each channel acts as a frequency selective network, thus locking onto one of the voiced speech harmonics that falls within its capture range. Then automatically tracks the variation (within the loop's lock range) which results due to changes in the frequency of the excitation waveform. The output of the first PLL is passed through a second phase-locked loop (PLL) in the same channel which divides its frequency by a certain constant n. The free running frequency f_{oII} of the second loop and its capture range are set to correspond to the nth submultiple of the frequency f_{oI} and capture range of the first PLL. An additional output signal of the first PLL conveying the amplitude of the incoming signal locked by the loop is filtered and then used to modulate the output signal amplitude of the second PLL. A low-pass filter (LPF) of width 20-40 Hz will be sufficient to pass undisturbed the amplitude variations of the speech harmonics.

Consequently, the output of the multiplier in each channel will have an amplitude proportional to the amplitude of the incoming locked speech component and 1/n of its frequency. The outputs of all the channels are then added together yielding an output speech signal which has a bandwidth 1/n the original speech bandwidth, but the relationships among the relative amplitudes of the speech harmonics are retained after the frequency division.

When the input speech is unvoiced, as for the plosive and fricative consonants, the spectral representation is a continuous spectrum, such as is obtained from measurements of random noise. In this case, many in-band signals may be present at the input of each analyzing phase-locked loop of the PLL frequency transformer system. However, each PLL will still be able to lock onto one of these in-band signals, which is most likely the one having the largest amplitude assuming the other in-band signals having much smaller amplitudes. Their existence may degrade the operation of the loop as a frequency-selective circuit by decreasing its lock range. In case more than one in-band input signal is present and which have comparable amplitudes, they can interfere with the locking process by forming with the loop VCO signal some beat frequencies, these being of the same magnitude as the locking beat note; these may thus drive the loop out of lock. In any event, the system output will still be a set of carriers. each now randomly modulated in amplitude and in frequency

III. DISCUSSION

A new system has been described for compressing the bandwidth of speech. Although the system design approach is compatible with some other system [7], the method of implementation using phase-locked loop systems is a new approach. This provides powerful tools in the design of such stringent narrow-speech analyzer filters as well as offering simple and stable frequency dividers. In fact, the main advantage that uniquely characterizes the new PLL frequency transformer is its ability to track automatically the individual speech harmonics and to reproduce correctly (while being compressed at the output stage) their frequency changes.
Experimental results to date on a 2 channel model of the system has been encouraging [8].

IV. REFERENCES

1. S.G.Knorr: "A Hearing Aid for Subjects with Extreme High-Frequency Losses", IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-24, Dec. 1976, 473-480.

2. A. Risberg, "A Bibliography on Recording of Speech for the Deaf and Hard of Hearing," Quarterly Progress Report (Speech Transm. Lab., Stockholm, 76-88.

3. D. Ling, "Three Experiments on Frequency Transposition," Proceedings of the Conference on Speech-Analyzing Aids for the Deaf, American Annals of the Deaf, Vol. 113, No. 2 (March 1968), 283-294.

4. L. Pimonow, "L'Application de la Parole Synthetique dans la Correction Auditive," Acoustica 12 (1962), 285-290; English Translation in the J.Aud. Research, Vol. 3 (1963), 73-82.

5. H. Fletcher, "Speech and Hearing in Communication," Van Nostrand, New York, 1953.

6. J. L. Flanagan, "Speech Analysis Synthesis and Perception," New York: Academic, 1965.

7. M. R. Shroeder, B. F. Logan and A. J. Prestigiacomo, "New Methods for Speech Analysis-Synthesis and Bandwidth Compression," Proc. Stockholm Speech Comm-Seminar, R.I.T., Stockholm, Sept. 1962.

8. F. El-Mokadem, "Theory of Subharmonic Generation With Application to Hearing-Aid Circuitry; Ph.D. Dissertation, University of Maryland, 1976.