MASTER COPY
DO NOT DESTROY

TEST DATA AND WAVEFORMS

for

SPEECH PROCESSING UNIT, MODEL SPU-2

by

Wm. Henneberry
and
Paul Grove

THE TECHNICAL MATERIEL CORPORATION
MAMARONECK, N.Y. OTTAWA, ONTARIO

The TMC Model SPU-2 Speech Processing Unit

The Model SPU-2 Speech Processing Unit is a simple, compact, solid state auxiliary unit which can greatly enhance the efficiency of a radio transmitter or receiver system operated with speech input.

Four basic techniques are employed to reduce the peak to average ratio of a speech signal and to maintain a constant level of signal output despite wide variations in input amplitude.

The four basic techniques are:

a) Frequency Band Compression
b) Volume Compression
c) Pre-emphasis
d) Clipping

Each technique is directed toward a principal end: the reduction of the peak to average ratio of the speech waveform. In effect, this allows a transmitter to be operated at a higher average power than would be possible without the SPU-2. In a receiving system, with a weak speech signal in the presence of a high noise level, the SPU-2 can improve the intelligibility by as much as 60%.

The typical unprocessed speech signal has a high peak to average ratio, and a definite "Syllabic Rate." This is clearly shown in Figure 1, which is an oscillogram of the phrase:

"AND WHAT IS SO RARE AS A DAY IN JUNE"

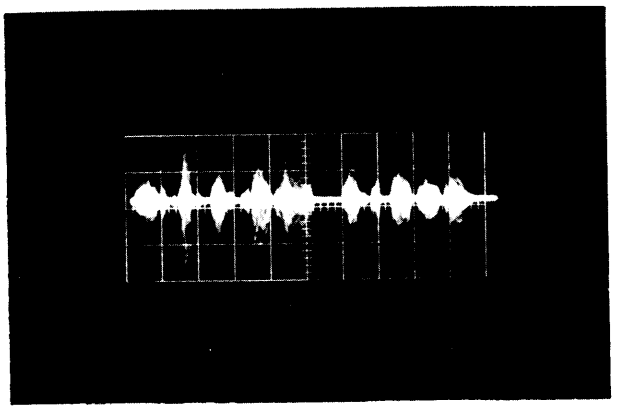


FIGURE 1

In Figure 1, each word is a monosyllable; there are a total of ten syllables, and the time represented is five seconds. This represents a syllabic rate of 2 per second.

Figure 2 is an oscillogram of the phase:

"SHE SELLS SEA SHELLS BY THE SEA SHELL SHORE"

In the case of Figure 2, nine syllables are uttered in a period of nine seconds; this represents a syllabic rate of 1 per second.

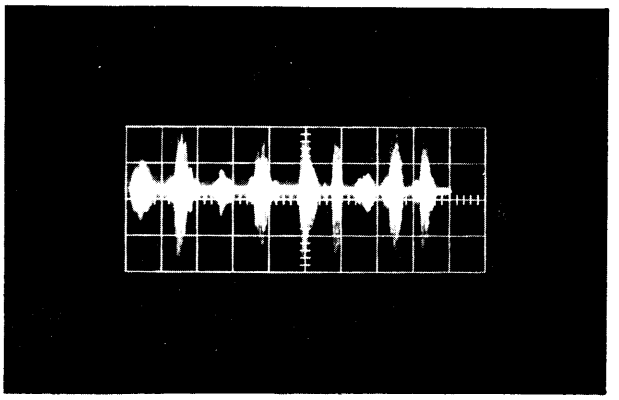


FIGURE 2

Figure 3 is an oscillogram of a longer phrase, uttered at a faster rate than the phrases of Figures 1 and 2. The period represented is five seconds, and there are approximately 25 syllables, indicating a syllabic rate of 5 per second.

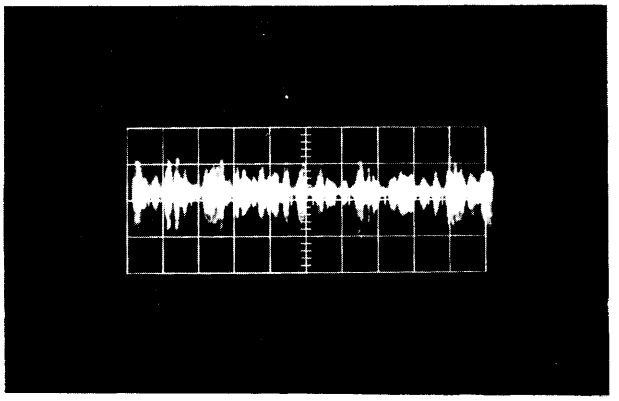


FIGURE 3

An examination of Figures 1, 2 and 3 also discloses that, from instant to instant, the signal excursions range from intense peaks to periods of silence. The average value of these signals, then, is greater than zero but much less than the level of the high peaks.

Single Sideband transmitters are rated in terms of Peak Envelope Power (PEP), and Average Power (Pav), with two equal amplitude, harmonically related tones applied to the system. Under such conditions, the transmitter is operating at full capability and the peak to average ratio is 2:1. When additional tones are added, the in and out of phase addition and subtraction of the various tones raises the peak to average ratio, and, for the same PEP output, the average power must be reduced.

Theoretically, a transmitter rated at 10 KW PEP and 5 KW Pav with two tones, with a peak to average ratio of 2:1, will, when fed with ten equal amplitude tones, be operated at 10 KW PEP and 1 KW Pav with a peak to average ratio of 10:1.

In a practical situation, with a typical unprocessed speech signal input, the situation would be approximately the same.

With proper use of the SPU-2, the peak to average ratio of the previously described transmitter could be reduced to about 5:1; this represents a 3 DB increase in average power.

Because TMC transmitters are conservatively rated, the peak to average ratio of the TMC GPT-10K transmitter can be less than 10:1, and, with the use of the SPU-2, can be as good as 3:1, with insignificant distortion.

Frequency Band Compression is a technique which limits the audio response of a system to the minimum required for a high degree of articulation. The extent to which a communications device is capable of reproducing the original speech meaning is measured in terms of articulation, or intelligibility; naturalness, or "high fidelity" is a secondary consideration.

It has been found that an audio bandwidth from about 200 to 3000 cps is quite sufficient for the transmission and reception of speech with a high degree of articulation. The range of the human voice extends from about 100 to 10,000 cps, a much wider range than that required. In the SPU-2, the frequency band is "compressed", by means of a highly selective filter and specially designed audio transformers, to a range of 200-3000 cps. This eliminates the unnecessary waste of power in a wide frequency coverage which is not required for good articulation.

Volume Compression is a technique whereby signal peaks above a given level are "compressed", or effectively reduced in height. This is accomplished in the SPU-2 by a series of amplifiers and automatic gain control circuits.

The effect of volume compression is to "reduce" high peaks and "raise" low amplitude signals, resulting in a chain of syllables which are more uniform in amplitude than the original syllables. This technique has no significant effect on articulation.

Figures 4 and 5 are oscillograms of the phrase:

"NOW IS THE TIME FOR ALL GOOD MEN TO COME TO THE AID OF THE PARTY I THANK YOU"

Note the high peak to average ratio of Figure 4, photographed before volume compression. Figure 5 was photographed from an oscilloscope presentation "picked off" a point after the volume compression circuits. The effect of volume compression in the SPU-2 is clearly evident.

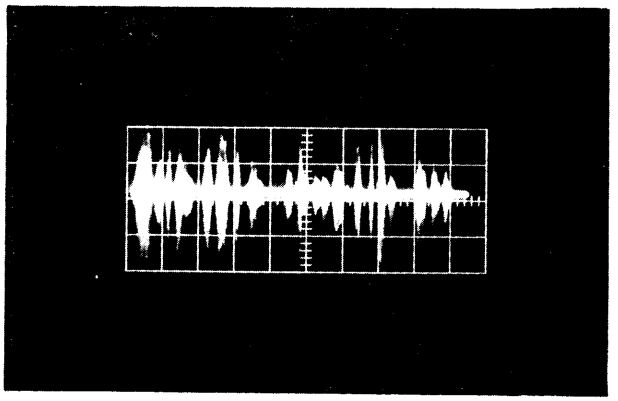


FIGURE
4

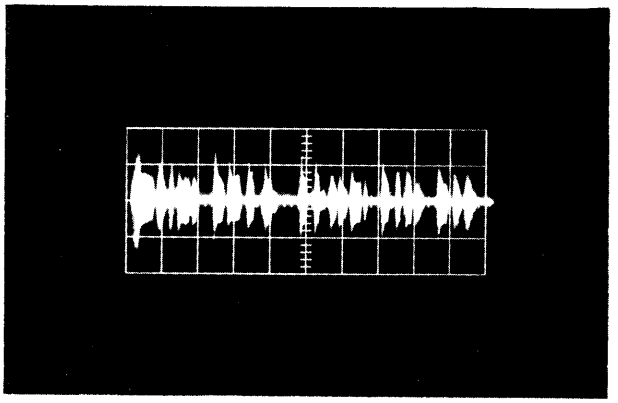


FIGURE
5

The time constants of the AGC circuits in a volume compressor system are extremely important, and are closely associated with the syllabic rate of speech. The syllabic rate has already been roughly defined as the average number of distinct syllables uttered per second. The syllabic rate varies from about .5 cps to about 25 cps, with an average of about 7 cps.

There are two time constants to be considered:

a) the attack time, or speed of response to changing levels.

b) the decay time, which is the time required for the AGC circuits to return to a "quiescent" state after cessation of speech.

Figure 6 is an oscillogram covering a 20 second period. Each vertical line represents 2 seconds of time; each horizontal line measures the amplitude of the AGC voltage deviation from the quiescent value, which, in the case of the SPU-2, is 0 volts, the top line of the presentation.

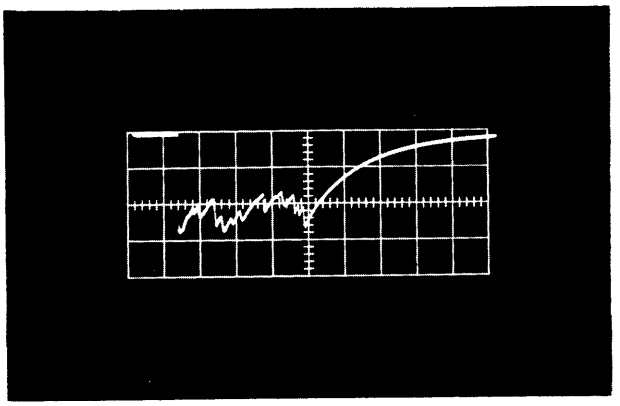


FIGURE 6

At the start of the sweep and for a period of 2.8 seconds, there is no speech input and AGC voltage is zero. At 2.8 seconds, a speech signal is applied to the compressor circuits. Note that at this time AGC voltage becomes almost instantly negative; this negative voltage is applied to the compressor-amplifier circuits to control the gain. As time progresses, the AGC voltage is seen to vary as separate syllables are uttered. The more negative excursions correspond to the higher peaks of the speech signal, and the increased bias reduces these peaks. Conversely, the less negative excursions raise the level of the lower amplitude speech signals. After 10 seconds, (the center of the sweep), the speech signal is stopped and the AGC voltage is seen to decay to zero in the remaining ten seconds.

In the SPU-2, the attack time, or ability of the system to follow recurrent peaks, is based on the average syllabic rate of 7 cps, the period of which is .143 second. The decay time depends on a simple RC circuit; this can be altered quickly and simply to meet the particular needs of the user.

The AGC compressor-amplifier circuits also give the SPU-2 its 40 DB dynamic range. An audio tone at 0 DBM into the device can be varied plus and minus 20 DBM in amplitude and the audio output will remain constant within 2 DB. This is of importance not only in transmitting systems but also in receiving applications where a constant audio output, with widely varying input amplitude, is required.

Pre-emphasis is a technique which compensates for the uneven distribution of power among the lower and higher speech frequencies. The solid line curve of Figure 7 shows the average distribution of power among the speech frequencies of interest.

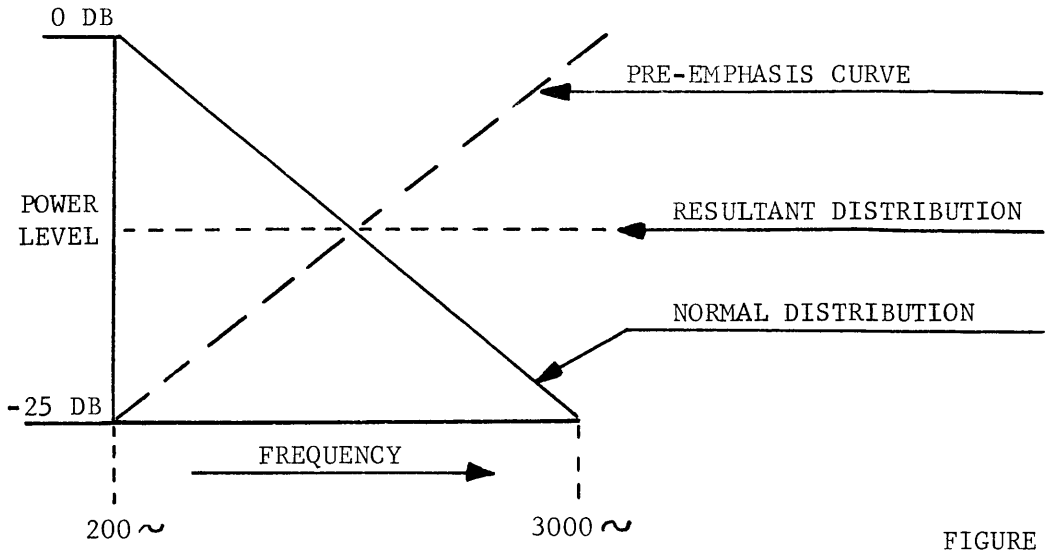


FIGURE 7

Note that the power is concentrated in the low frequencies, and, if a tone at 200 cps is assumed to be at a level of 0 DB, a tone at 3000 cps will be down minus 25 DB. It has been found that the vowel sounds add little to articulation but that the consonants are of great importance in this regard. In general, the vowel sounds are associated with the low frequencies and the consonants with the higher frequencies. A pre-emphasis circuit compensates for the uneven distribution of power by emphasizing the higher frequencies and de-emphasizing the lower frequencies. This is indicated by the dashed curve of Figure 7. This has the effect of creating a more even distribution of power in the speech spectrum, as indicated by the dotted curve of Figure 7.

Clipping is a technique which limits the peak excursions of a signal to a predetermined set value.

Figure 8 is an oscillogram of the single word: "SHIP", taken with a .5 second time base and with no speech processing techniques applied. One would normally consider "SHIP" to be a single syllable word, but the time base is comparatively short, and care was taken to enunciate clearly. The figure clearly shows:

"SHHII--PP"

On a short time base, one can clearly see the complexity of single words.

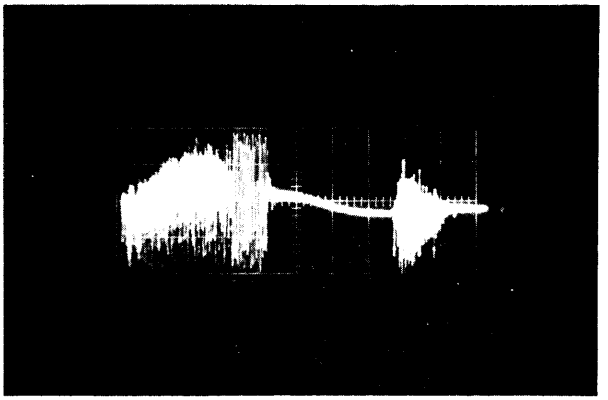


FIGURE 8

Figure 9 is an oscillogram of the same word, on the same time base, after compression, pre-emphasis and clipping. Note that the peaks are limited to a given level. This oscillogram does contain unwanted frequencies, since these are unavoidable created as a result of the clipping action.

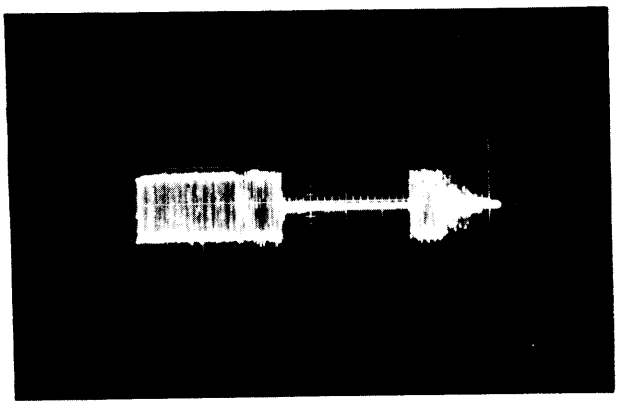


FIGURE
9

Figure 10 is an oscillogram of the same word, on the same time base, after compression, pre-emphasis, clipping and frequency band compression. Undesired frequencies have been removed by following the clipper circuit with a highly selective filter which passes frequencies only in the desired range of 200 to 3000 cps. The effect of the filtering is clearly evident. The effect on the peak to average ratio is also shown.

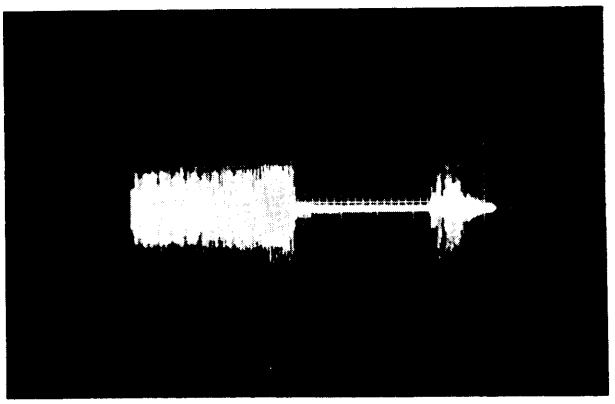


FIGURE
10

In the SPU-2 the clipper circuit may be utilized or bypassed, by means of a front panel switch. This circuit may or may not be desired, depending on the character of the signal itself. The clipper finds its greatest application when the SPU-2 is used in a receiving system and the signal to noise ratio is poor. The clipper cannot improve a "good" signal, but it can increase the intelligibility of a poor one, by as much as 60%. This means that if the index of the intelligibility of a received signal is 40%, the use of the SPU-2 with the clipper circuit "in" can increase the index of intelligibility to 100%.

The "clipping level", in DB, is given by:

$$DB = 20 \log \frac{\text{voltage peak before clipping}}{\text{voltage peak after clipping}}$$

In the SPU-2, the clipping level is 12 DB. This corresponds to a voltage ratio of 4:1; that is, if a voltage peak is 4 volts before clipping, it will be 1 volt after clipping. A clipping level of 12 DB was selected because this level insures the greatest improvement in intelligibility while maintaining the greatest degree of naturalness of the speech signal. It should be pointed out that the compressor circuits ahead of the clipper insure that the clipping level will not exceed 12 DB, with attendant distortion. Actually, severe distortion is not noted until the clipping level becomes greater than 24 DB. The clipper circuit is not ordinarily used with intelligence other than speech.

Figures 11 through 15 are SPU-2 oscillograms of the phrase:

"NAVAGATION IS THE ART OF DETERMINING THE POSITION OF A SHIP AND OF CONDUCTING A VESSEL IN ANY WATERS".

In each case the total time base is 20 seconds; speech is commenced 2 seconds after the start of the sweep, and lasts for 8 seconds. The remaining ten seconds are silent.

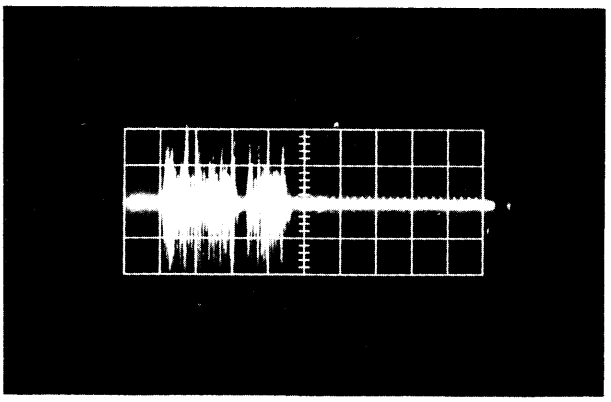


FIGURE 11

Raw, unprocessed speech

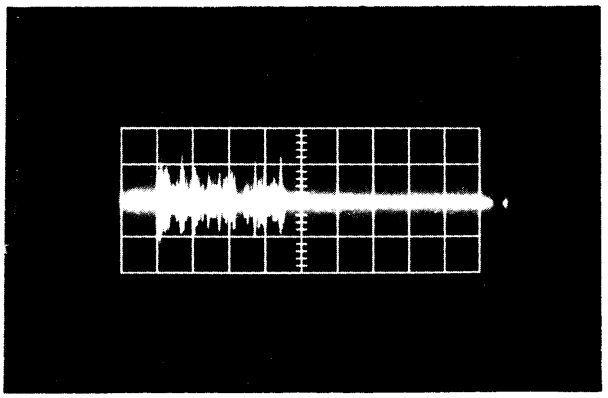


FIGURE 12

After volume compression

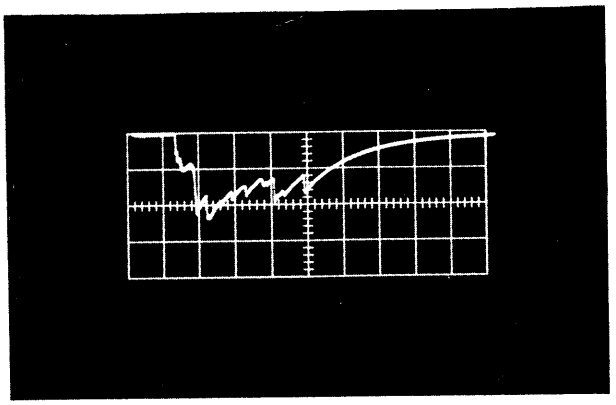


FIGURE 13

AGC characteristic of compressor circuits

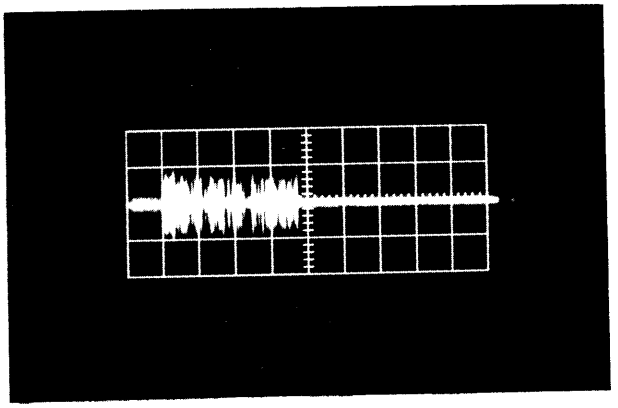


FIGURE 14

After clipping

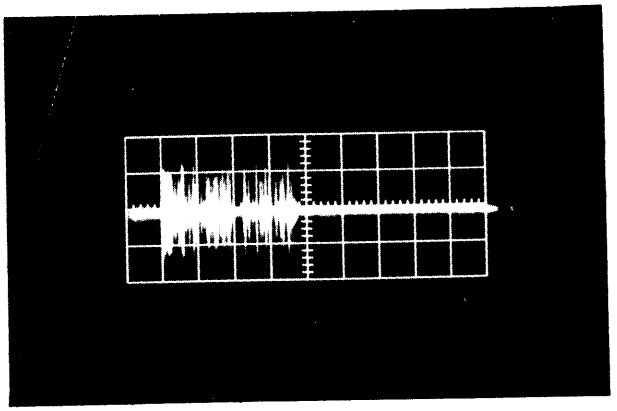


FIGURE 15

After filtering and amplification

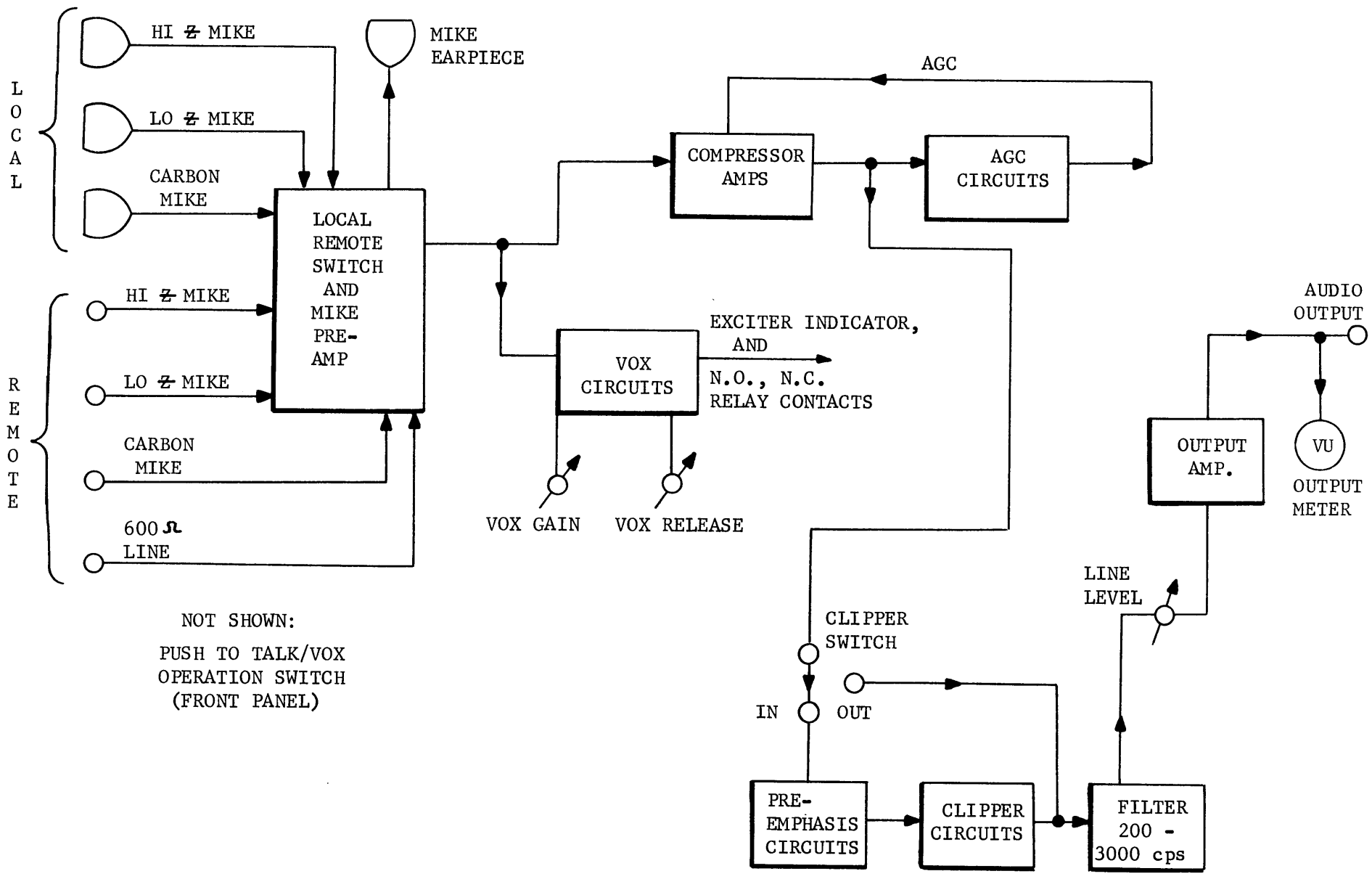


Figure 16. Simplified Block Diagram, SPU-2

A Brief Discussion of the Simplified Block Diagram of the Model SPU-2 Speech Processing Unit (See Figure 16)

a) provisions are made for low impedance, high impedance or carbon microphone inputs from the front panel. (LOCAL INPUT).

b) provisions are made for low impedance, high impedance or carbon microphone inputs, or 600 ohm line input, from a rear terminal strip. (REMOTE INPUT).

c) selection of REMOTE or LOCAL input is made by a front panel selector switch.

d) the microphone input is processed in a preamplifier; the 600 ohm line input, at a nominal 0 DBM level, does not require preamplification.

e) the speech input selected is processed in the compressor-amplifier-AGC circuits, and the volume compressed output is applied to the CLIPPER IN OUT switch. The AGC circuits provide the compressor-amplifier circuits with a bias which changes with signal levels.

f) the signal from the REMOTE LOCAL switch is also applied to the VOX circuits; these circuits activate and de-activate a relay with both normally closed and normally open contacts, for operation of a remote transmitter or exciter unit. A front panel EXCITER lamp operates in conjunction with these circuits.

g) not shown on the block diagram is a PTT/VOX front panel switch; the relay of paragraph (f) above may be operated by either the VOX circuits or the PUSH TO TALK button of a local microphone, as desired. The front panel VOX GAIN control is adjusted so that the instant audio is applied, the EXCITER relay will activate. The VOX RELEASE front panel control provides a variable delay, after cessation of speech, before the EXCITER relay de-activates.

h) the CLIPPER IN OUT switch allows the volume compressed signal to be routed through or around the pre-emphasis-clipper circuits, as desired.

i) beyond the clipper circuits, a highly selective filter is inserted to limit the bandpass to audio frequencies in the range 200 to 3000 cps.

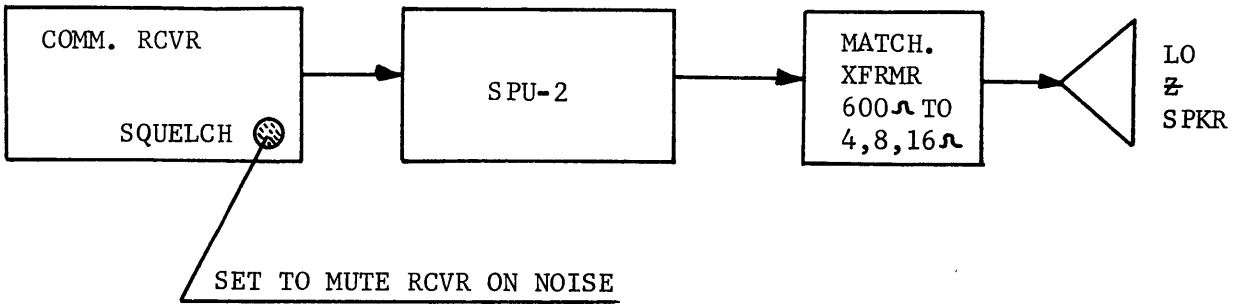
j) the output of the filter is applied to a final amplifier with associated LINE LEVEL control.

k) the output circuit is fitted with a true VU meter, providing a constant front panel indication of signal output.

l) an OUTPUT selector switch allows the output to be connected to either USB terminals, LSB terminals, or both. This allows use of the SPU-2 with either an SSB or AM transmitter.

Some Suggested Uses for the Model SPU-2

1. There are many instances when a user wishes to monitor a receiver output with a speaker, but does not want noise in the speaker when no intelligence is being received. In addition, it is usual that the receiver output varies over a wide range, particularly when the receiver is monitoring a net. In this case, the 600 ohm output of the receiver is connected to the 600 ohm input of the SPU-2, and the 600 ohm output of the SPU-2 is connected to the speaker via a matching transformer. The SQUELCH control on the associated receiver is adjusted to mute the receiver when no signal is being received. This is necessary because, since the SPU-2 has a 40 DB dynamic range, it requires a "clear channel" signal. The PTT/VOX switch on the SPU-2 is set to VOX.



The SPU-2 will now activate the instant the receiver delivers an output, and will maintain that output even though the receiver output varies over a plus or minus 20 DB range. When the receiver is again muted, the SPU-2 will de-activate. An arrangement might also be connected by the user in conjunction with the EXCITER relay, whereby this relay could light a lamp or activate a buzzer at a point remote from the receiver room.

2. There are cases when a user must receive a speech signal and re-transmit it to another point. The received signal might have a very low signal to noise ratio, and be barely intelligible. In this case, the SPU-2 can be used as a link between receiver and transmitter; it will increase the intelligibility of the poorly received signal and process it for application to the transmitter. The transmitter may also be operated at a higher average power than would be possible without the SPU-2. In addition, the VOX circuits of the SPU-2 can be used to activate and de-activate the transmitter for further savings in power.