An Electronic Device to Reduce the Dynamic Range of Speech

By

Eric Michael Hildebrant

Submitted in Partial Fulfillment
of the Requirements for the
Degree of Bachelor of Science
at the
Massachusetts Institute of Technology

May 21, 1982

Signature of Author.

Department of Electrical Engineering,

Certified by Thesis Supervisor,

Accepted by Chairman, Departmental Committee on Theses
An electronic device to reduce the dynamic range of speech was designed, constructed, electrically characterized, and preliminarily evaluated psychoacoustically. The device's signal processing consisted of: 1) filtering incoming speech into bands encompassing the first three formants of speech; 2) infinite peak clipping within each band; 3) filtering the clipped signals to their original ranges; and 4) summing the three filtered/clipped/filtered signals to form the output. Intelligibility tests with listeners who had a simulated restriction of dynamic range showed large differences in favor of the processed speech over the unprocessed, but frequency-equalized (whitened), speech. For phonetically-balanced monosyllables, scores averaged 4% and 36%, and for words in sentences 17% and 98%, for unprocessed and processed speech, respectively. However, a simple clipper system also showed good performance. These preliminary experiments illustrate the problems and potentials of peak clipping as a means of amplitude-range reduction for the severely hearing-impaired.

Thesis Supervisor: Dr. Patrick M. Zurek
Title: Visiting Scientist
TABLE OF CONTENTS

TITLE PAGE----------------------------- 1

ABSTRACT------------------------------ 2

TABLE OF CONTENTS--------------------- 3

LIST OF FIGURES------------------------ 4

INTRODUCTION-------------------------- 7

PREVIOUS EXPERIMENTS------------------ 10

DESIGN RATIONALE FOR THE

THREE-BAND CLIPPER------------------- 26

EXPERIMENTS--------------------------- 29

CONCLUSION---------------------------- 54

APPENDIX----------------------------- 57

REFERENCES--------------------------- 90
LIST OF DIAGRAMS

1. Single Formant Input to the IPC-er --------------- 33
2. Output of IPC-er with a Single Formant Input---- 34
3. 3-Formant Vowel, Input to IPC-er--------------- 35
4. Output of IPC-er with 3-Formant Input---------- 36
5. Output of OF/IPC-er with 3-Formant Input------- 37
6. Input to 3-Band Clipper------------------------ 38
7. Output of 3-Band Clipper----------------------- 39
8. Intelligibility Enhancement of Noisy Speech----- 41
9. Intelligibility of Already Noisy Speech---------- 43
10. Audibility Thresholds---------------------------- 45
11. Long-Term Spectrum of Babble------------------- 47
12. Long-Term Spectrum of Babble------------------- 48
13. Equalized Speech Spectrum---------------------- 49
14. Long-Term Spectrum of 3-Band Clipped Babble--- 50
A1. 3-Band Clipper Block Diagram------------------- 58
A15b. Response of the 3'd Formant----------------- 81
A16. Response of the Peak Clipped 3'd Formant------ 82
A17. Response of the 3'd Formant, Peak Clipped, and
       Filtered-------- 83
A18. Summer Electrical Schematic------------------ 84
A19. Output Response of the 3-Band Clipper-------- 85
A20. Infinite Peak Clipper------------------------- 87
INTRODUCTION

There are two types of practical applications for an electronic device that reduces the dynamic range of speech. One is to help a listener in a noisy environment hear speech that was originally produced in a quiet environment. An example of this is a public address system in a noisy area.

The other application is for an aid to people with sensorineural hearing impairment. Unlike hearing loss of the conductive type, where simple linear amplification is sufficient to overcome an internal attenuation, sensorineural hearing loss is characterized by elevated sound detection thresholds but essentially normal discomfort thresholds. Thus, there is a limited dynamic range for sound reception, a problem that cannot be solved with a linear amplifier.

What is needed is a device that reduces the dynamic range of all incoming sounds, but also allows intelligible perception of speech.

Two major types of amplitude-range reduction are automatic volume control, or amplitude compression, and peak clipping. Amplitude compression is accomplished by circuitry similar to that of the AVC part of a radio receiver. The envelope of the rectified speech signal is used to control the gain of a preceding AC amplifier.
Peak clipping is the other primary method of amplitude-range reduction. Different degrees of peak-clipping are achieved by clipping at different signal levels. In the extreme, infinite peak clipping preserves only the polarity, and none of the amplitude variation, of the original signal. Infinite peak clipping was used exclusively in the present study. The abbreviation "IPC" will be used in the remainder of this paper to stand for the action "infinite peak clip".

One means of achieving an IPC'ed signal is to: 1) Amplify greatly (as much as possible) the input signal; and 2) Present the amplified signal to a Schmitt trigger that has very little hysteresis. This insures that the clipped signal is rectangular (binary-valued, with fast rise and fall times). The IPC'ed signal ideally would change value at the zero-crossings of the input speech signal, have no time lag from input to output, and preserve the polarity of input signal. Practical limitations on an IPC-er are imposed by the noise of the high-gain amplifier. This determines the hysteresis level of the Schmitt trigger stage, which prevents the background noise from producing an output.

Processing speech with an IPC-er yields a binary signal with a peak factor of 1 (ratio of peak voltage to RMS voltage), whenever the input exceeds the hysteresis threshold. This drastically reduces the signal's dynamic range. An IPC-er also limits high-intensity signals instantaneously.
The drawback to speech processed only by an IPC-er (and an adjustable linear amplifier to provide a comfortable listening level) is that it does not have good intelligibility (Licklider, 1946). The processed speech is dominated by low-frequency "rumbles" and "glides", and the quality of the sound is very poor. Some type of processing prior to clipping is necessary to produce intelligible speech. The studies presented in the next section show the historical development of speech-processing schemes involving IPC-ing and their limited application to the hearing-impaired. These findings are also applicable to the processing of "clean" signals for presentation in a noisy environment.

The following section, "Previous Experiments", presents brief but detailed summaries of the important papers that served as background for the present work. The reader may sample from this detailed material, or skip to the subsequent section, "Design Rationale", for a summary of the main points that were taken from the literature and applied to the design of the device under study.
PREVIOUS EXPERIMENTS

"Effects of Amplitude Distortion upon the Intelligibility of Speech", J. C. R. Licklider (1946) JASA 18

This article concerns the reduction in intelligibility caused by center clipping, linear rectification, and peak clipping.

METHOD: Processing consisted of HP filtering, followed by peak clipping.

RESULTS: Amplitude distortion (peak clipping) is more detrimental when non-impulsive noise is mixed with the speech at a point ahead of the non-linear circuit than it is in quiet. When the noise consists of sharp high-amplitude pulses, however, peak clipping eliminates the noise peaks and improves intelligibility.

Noise added to peak-clipped speech tends to cover up some of the effects of distortion.

In a quiet environment, the quality and intelligibility of peak-clipped speech is better if the speech spectrum is first tilted by + 6 dB/octave prior to clipping.
"The Intelligibility of Rectangular Speech Waves",

Their aim was to measure the intelligibility of peak-clipped speech and find out why it is more intelligible in a noisy environment.

METHOD: The processing employed a 250 Hz high-pass filter preceding the clipper. Normal and clipped speech were equated for peak sound-pressure levels of 85 dB SPL. Noise was added after the clipper.

RESULTS: Experiments measuring word articulation vs. speech to noise ratio, with white noise introduced at a point following the clipper, showed the clipped speech more intelligible at low (-10 to +3 dB) signal/noise ratios, and less intelligible at higher signal/noise ratios, than unprocessed speech. This crossover was due to intelligibility of clipped speech leveling off at 50% while that of un-clipped speech rose from 50% to 85% as the signal/noise ratio increased from +3 to +25 dB.

The enhanced intelligibility at low S/N ratios is due to the greater power in a clipped speech wave than in a normal speech wave. When their peak amplitudes are equal, there is about 16 dB greater power in a rectangular speech wave than a normal speech wave.
"On the Effect of Frequency and Amplitude Distortion on the Intelligibility of Speech in Noise", I. Pollack (1952) JASA 24

Tests were conducted to determine whether the superiority of peak-clipped speech at low S/N ratios and of unclipped speech at high S/N ratios is a function of the frequency band of the speech signal.

METHOD: Processing consisted of sharp frequency limiting, then infinite peak clipping, then filtering again, and the addition of noise. This signal was presented thru earphones to the listeners.

RESULTS: Consistent with previous findings, intelligibility of broad band (0 to 6700 Hz) speech subjected to peak clipping is considerably less than when only the speech frequencies above 394 Hz are clipped. With no peak clipping, the band 0 to 394 Hz has practically no effect upon intelligibility.

In comparing these two test cases the articulation score for a given sharp-cutoff filtering condition was compared for clipped and un-clipped speech.

The relationship brought out from these tests is that for speech limited to a certain frequency band, the intelligibility of the unclipped speech, relative to that of the infinitely peak-clipped speech, is a function of the S/N ratio and is roughly independent of the frequency band employed.
"On the Power Gained by Clipping Speech in the Audio Band", W. Wathen-Dunn, D. Lipke (1958) JASA 30

This study measured the amplitude distribution of speech, and the effect on speech power of clipping at various levels.

RESULTS: When $E/E_{rms}$ ($E =$ Voltage of a Speech signal) is 5.1, 97 to 98% of the total speech power is contained within that portion (below the 0.1% probability point). So, the peak factor corresponding to 0.1% probability is 5.13, which is 14.2 dB.

24 dB of peak clipping (and re-amplification to the original peak amplitude) yields a power increase of 12 dB. Infinite peak clipping yields a 14.2 dB increase. These power increases are due to a peak-clipped signal not wasting dynamic range on the peaks.
"Intelligibility of Peak-Clipped Speech at High Noise Levels", I. Pollack and J. Pickett (1959) JASA 31

Their aim was to determine the relation of intelligibility to the level of peak clipping when the post-clipped speech power is held equal to the unmodified speech.

METHOD: Masking noise was either a uniform spectrum noise between 250 and 6000 Hz, or a low-frequency noise that fell 12 dB per octave between 250 and 6000 Hz. Overall noise level was 90 or 125 dB SPL. Peak clipping levels were 0, 12, or 24 dB. The over-all speech power after clipping was made equal to that before clipping.

The peak-clipped, power-compensated speech was passed through a band filter (250 to 6000 Hz), mixed with noise, and presented to the listeners. Harvard PB mono-syllabic word tests were used. Plots of % articulation vs. dB of peak clipping (at different S/N ratios) were made.

RESULTS: With uniform spectrum noise it was found that:

1) Intelligibility for equal speech power levels is nearly independent of the degree of peak-clipping over a range of S/N ratios (-10 to 10 dB) and at noise levels of 90 and 125 dB. Intelligibility of clipped speech was slightly superior (averaged over all S/N ratios and noise levels of the tests) by 4.1 %.
2) For a given intelligibility level, higher S/N ratios are required at the high noise level than at the moderate noise level.

3) Above a S/N ratio of 7 dB, at the 125 dB noise level, intelligibility scores higher than 67% could not be obtained.

With the low-frequency noise, an analysis of test variance (articulation vs. dB of peak clipping, with variable S/N levels) failed to reveal peak clipping as a significant variable, nor was the interaction between peak clipping and noise level found to be significant.


Their aim was to find which frequency bands of speech provided maximal intelligibility for hearing-impaired subjects.

The study was motivated by a finding of Martin, et al (1970) that normal relative amplitudes of F1 in a speech signal virtually eliminate perception of F2 transitions in the sensorineural hearing impaired.

METHOD: Subjects with adventitious sensorineural hearing loss were presented with processed speech. The processing consisted of passing the speech through a variable slope HPF.
The HPF used was a RC gaussian type with a -3 dB cutoff at 1600 Hz. Asymptotic attenuation slopes of 12, 18, or 24 dB per octave were obtained by switching in appropriate numbers of RC filter stages, which caused spectral weighting of the 1'st formants to occur. The attenuation range between 500 and 2000 Hz is relatively flat, and a change of slope (number of HPF stages) permits the selection of the relative amount of F1 in the signal.

Sensation levels of 20, 30, and 40 dB relative to the SRT (speech reception threshold) were used for both modified and unmodified speech.

RESULTS: A comparison with un-altered speech at the same RMS level revealed the processed PB words 20% more intelligible.

Results were plots of articulation percentage vs. sensation level measured relative to the SRT. Increased articulation scores resulted from the use of a particular attenuation slope for all sensation levels, although the slope for highest score is not the same from subject to subject (different types of loss for different subjects) nor even for different sensation levels with the same subject. The intelligibility enhancement for any given subject with the "best" slope averages at 20%, however.
"The Influence of First and Second Formants on the Intelligibility of Clipped Speech", I. B. Thomas (1968) JAES 16

The aim was to measure the influence of the the first and second formants on the intelligibility of clipped speech.

METHOD: Speech was filtered so that all but one formant was suppressed prior to infinite peak clipping.

RESULTS: Second-formant clipped speech has an average intelligibility (over 10 subjects) of 71.1% and first-formant clipped speech has an average intelligibility of 7.6%.

Spectrograms of the resulting clipped speech reveal that the behavior of the isolated formant is unaffected by the filtering and clipping process. There are higher bands identifiable as harmonics present along with the original formant, however.

Tests of the intelligibility of second-formant clipped speech were done twice. Comparison of scores, subject by subject, shows a 5.4% learning increase in articulation scores.

Most errors on the tests were differentiation problems of the stop consonants.

Loss of intelligibility when wideband speech is clipped can be partially accounted for by the introduction (due to clipping) of distortion products of lower-frequency signals. Also occurring is the direct suppression of second and higher formants by the larger-amplitude first-formant components.
"Enhancement of Speech Intelligibility at High Noise Levels by Filtering and Clipping", I. B. Thomas and R. J. Niederjohn (1968) JAES 16

METHOD: Processed and normal speech were presented in a high level of ambient white noise at the listener's ears. Processing involves high-pass filtering, to attenuate the first formant, followed by infinite peak clipping. The HPF used to process speech before clipping was a four-stage cascaded RC filter with 24 dB/octave asymptotic roll-off, down 3 dB at ~1200 Hz.

The SPL of the noise was maintained at 90 dB.

RESULTS: Average intelligibility of the modified speech without any added noise was 95%, for unmodified speech it was 99%. Noise was added at a constant 90 dB SPL. At a S/N ratio of -5 dB, modified speech yielded an intelligibility score 20% higher than unmodified.

Under high noise conditions, the processed speech's intelligibility is greater than that of normal speech with equal average power.

With the first formant suppressed, the axis crossing of the filtered speech are largely due to the second formant.
"The Intelligibility of Filtered-Clipped Speech in Noise",

Aim: Using their previous paper "Enhancement...", as a start, experiments were performed to determine the optimal cut-off frequency (-3 dB point) and asymptotic slope of the high-pass filter that precedes the infinite peak-clipper.

METHOD:

The band-pass filter for the white noise had a pass-band from 250 Hz to 6800 Hz. The 20 kHz oscillator was amplitude adjusted to produce a 20 kHz signal at the output of the clipper during "no speech" intervals (subjective silence). The HPF was constructed by cascading identical RC filter networks (gaussian filter), buffered between stages by emitter followers. Asymptotic slopes of 6, 12, 18, or 24 dB/octave were considered, and the cut-off frequency varied by changing the capacitor values of all stages. Egan's (1948) PB word lists were used as test material. Both overall level of the noise and the level of the filtered/clipped speech was 90 dB SPL.
RESULTS: In comparing the 4 HPF slopes of 6, 12, 18, and 24 dB/octave, and cut-off frequencies from 400 to 5000 Hz the highest average intelligibility score was obtained for an asymptotic slope of 12 dB/octave, and a -3 dB point of 1100 Hz. This is termed the "Optimal" filter.

The optimal filter (HP) before the IPC has the effect of suppressing low-intelligibility first-formant components which would dominate lower-amplitude, high-frequency components in the output. The optimal filter has the effect of making the IPC-er's output consist mostly of second and some third formant speech signals, which have higher intelligibility than first formant signals. Also, if the IPC's output signals' zero crossings are dominated by first formant components, then harmonics of these components will be present in the second (and higher) formant bands. These harmonics will obscure the identity of the speech sounds presented to the listener.


METHOD: Following is a block diagram of the system:
The noise band-pass filter has cut-off frequencies of 250 and 6800 Hz, and attenuation slopes of 24 dB/octave. The optimal filter and peak clipper were as in the previously described study by Thomas and Niederjohn.

AIM: This experiment, in which noise is added to the speech before processing, is the complement to the previous Thomas and Niederjohn (1968) study.

METHOD: For unmodified speech, noise level (at the summer) was 90 dB SPL. The RMS level of speech was 90, 95, and 100 dB SPL.

For modified (filtered/peak-clipped) speech, the signal to noise ratio at the input to the clipper was set at values of 0, 5, and 10 dB. The SPL at the listener's ear (constant due to clipping) was 90 dB. Listeners heard Harvard PB 50 lists presented binaurally, in phase, through headphones.

RESULTS: At all S/N ratios tested, the modified speech was more intelligible than the unmodified speech. As Lim and Oppenheim (1970) have noted, this result stands out as one of the few successes in the many attempts to enhance the intelligibility of already-noisy speech. Many more sophisticated systems have failed where this simple filtering/clipping scheme has apparently succeeded. Because this finding is not understood, an attempt to replicate it was undertaken. The failure to duplicate Thomas and Ravindran's finding is reported below in the section entitled "Experiments".
"Discrimination of Filtered/Clipped Speech by Hearing Impaired Subjects", I. B. Thomas, D. W. Sparks (1971) JASA 49

AIM: To see how the Optimal Filter/IPC scheme developed in the past papers performs as an aid to the hearing impaired.

METHOD: Speech processed through the "optimal" filter/clipper previously developed was presented to a group of hearing-impaired subjects. The intelligibility of speech so processed was compared to that of speech linearly amplified (uniform frequency-gain characteristics).

Harvard PB word lists were used as test material. Sixteen subjects with a variety of audiometric configurations were tested. Unmodified speech was presented to each subject at SL's from 10 to 40 dB (re their speech reception thresholds (SRT)). Modified speech was presented at the same overall SPL as unmodified speech. The storage oscilloscope method was used to measure the RMS level of speech.

RESULTS: In 13 out of the 17 ears tested, higher intelligibility scores were obtained with modified speech at all SL's. Except for those obtained at 40 dB SL, these differences in intelligibility were highly significant. In two cases, higher scores were obtained for unmodified speech at all SL's. In the last two cases, modified speech was higher at some SL's and unmodified speech higher at others.
Thomas and Sparks give two reasons why this form of speech processing should be helpful to the hearing-impaired. The first comes from the results of Martin and Pickett (1970), which show that for hearing-impaired subjects sensitivity to changes in the frequency of the second formant of speech depends on the relative amplitude of the first formant. If $F_1$ is absent, discrimination ability of hearing-impaired listeners for frequency changes in $F_2$ is not markedly less than that of subjects with normal hearing. When $F_1$ is added at an amplitude about equal to that of $F_2$, the discrimination ability of the subjects for frequency changes of $F_2$ is greatly degraded.

The second reason cited is that of the effects of recruitment (or alternatively, the smaller-than-normal dynamic range). The problem that recruitment creates for the user of a linear-amplifier hearing aid is either an inability to understand speech at low signal intensities, or if the gain of the hearing aid is increased to improve intelligibility the discomfort threshold will often be exceeded by extraneous noises.

Thomas and Spark's signal processing addressed these two problems by greatly suppressing the first formant with the "optimal filter", and supplying a constant SPL to the listener via the peak-clipper.

The two subjects who did not experience increased intelligibility with the modified speech had severe hearing loss in the region of the second formant and relied upon first-formant
"Effects of Whitening and Peak-Clipping on Speech
Intelligibility in the Presence of a Competing Message", L.L.
Young, J. T. Goodman, R. Carhart (1979) AUDIOLOGY 18

AIM: To see if peak clipping can enhance the
intelligibility of a message already corrupted with babble.

METHOD: Young, et al compared the intelligibility of
unmodified, whitened, and whitened/clipped (30 dB of clipping)
speech across S/N ratios with babble as noise introduced before
speech processing. "Whitening" refers to speech which has been
shaped such that its long-term frequency spectrum could be
considered flat, or "white". This was done using a one-third
octave multifilter (G. R. 1925). However, signals below 250
Hz were attenuated severely. The "babble" was composed of 5
talkers, each reading a separate passage.

The words were presented to listeners at an average SPL of
85 dB. The listeners had passed a 20 dB HL pure-tone screening
test at octave frequencies from 0.125 to 8 kHz. Signal to noise
ratios of -12, -8, 0, 8, and 12 dB were used. Test material was
the Lehiste-Peterson word lists.
RESULTS: Unmodified speech and whitened speech had virtually equal intelligibility, while whitened/clipped speech was much less intelligible (30% vs. 70% at 0 dB signal to noise).

The authors assert that Thomas's HPF/IPC would not yield higher intelligibility scores than their whitening/30 dB peak clipping system. Whitening the speech, they claim, would probably reduce the masking effect that F1 has on F2 in the sensorineural hearing impaired.
DESIGN RATIONALE FOR THE "THREE-BAND CLIPPER"

From the background papers described above, from other papers, and from common knowledge come the following observations pertinent to the design of the system that was constructed and tested. Concerning normal speech, two facts are important. First, it is known that speech signals have a peak factor of 14 dB (Wathen-Dunn and Lipke, 1958). Second, consonants, which contribute much more to intelligibility than vowels, are about 25 dB lower in intensity than vowels (Fletcher, 1953). Assuming that the RMS measurement of speech is primarily a measure of vowel RMS, we come to the conclusion that there is a range of almost 40 dB between the peak amplitude of vowels and the RMS amplitude of weak consonants.

Studies concerning sensorineural hearing-impaired listeners show that they have a severely reduced dynamic range between the threshold of sensation and the threshold of discomfort.

The peak amplitude of an IPC-ed signal is equal to its RMS voltage, for a "peak factor" of 1. Infinite-peak-clipped signals have the same output amplitude value, regardless of the input signal's amplitude (within practical input dynamic ranges). Except for very simple inputs, there is not yet a mathematical summary that would predict the clipped spectrum from the input spectrum. However, some qualitative rules can be applied. One rule is that if two signals (they need not be sine waves) are mixed and input to a IPC-er, and there is a ratio of 'x' dB
between their amplitudes, the ratio of those signal's amplitudes in the output will increase from 'x' to 'x' + 6 dB, as 'x' varies from 0 to "infinite" dB. That is, smaller signals are suppressed. Another rule is that the effects depend on the amplitudes but not the frequencies of the signals involved.

These observations are integrated in the following rationale leading to the design of the 3-band clipper. First, we recognize that the basic problem of a severely-hearing-impaired listener is that the range between the high amplitude peaks of vowels and the low-amplitude consonants is much larger than the dynamic range between audibility and discomfort thresholds. Thus, as the level is increased, vowels begin to become uncomfortable before a consonant's energy is fully audible. Because consonants are so important for intelligibility the problem is a serious one.

From Licklider's classic studies we know that infinite clipping is a means for drastically reducing the amplitude range of speech that simultaneously does not destroy intelligibility. In particular, if the clipping is preceded by an attenuation of the low-frequency components, intelligibility can be quite good.

Several studies (e.g. Licklider, Bindra and Pollack, 1948, and Thomas and Neiderjohn 1968) have shown that filtering and clipping of speech can enhance intelligibility substantially when noise is added after processing (especially when processed and unprocessed speech are equated in peak amplitude). This condition can be viewed as simulating the reduced dynamic range
of the hearing-impaired. Thus, there is promise for using this
approach with the hearing impaired. Thomas and Sparks(1971)
attempted to evaluate such a system with hearing-impaired
listeners, comparing the intelligibility of processed speech to
unprocessed speech. Their results favor the processed speech,
but the appropriate control of whitening the spectrum of the
unprocessed speech was not performed.

It was believed that improvements could be made upon the
highpass filter/clipper system used by Thomas and Sparks (1971)
and Thomas and Niederjohn (1968). The basic difference in the
present approach was to keep separate the frequency regions
characteristic of the first three formants of speech. It was
thought that the spectral formant peaks, which are known to carry
valuable information for the identity of speech sounds, could be
maintained by isolating the formant regions prior to clipping.
Further, the out-of-band distortion could be filtered out after
clipping and prior to re-assembly. The price for this frequency
specificity was thought to be only a slight increase in peak
factor due to filtering and summing of the three signals.

With such a system it was reasoned that the largest
component in a formant region would be maintained at about the
same output level, independent of the input level, and
uninfluenced by components in other frequency regions. The
expectation was that this arrangement would allow intelligibility
at least as good as the single high-pass filter/IPC-er system
described in the literature.
EXPERIMENTS

Preliminary experiments were performed to evaluate the intelligibility of speech processed by various methods. Two basic equipment configurations were used:

\[ \text{Signal} \rightarrow \Sigma \rightarrow \text{PROCESS} \]

\[ \text{Noise} \rightarrow \Sigma \]

These correspond to the situations in which either the speaker or listener is in a noisy environment, with speech processing in-between. Situation "A" would model a listener to a 'P. A.' system in a noisy work area, and situation "B" models a listener using a hearing aid in a noisy environment (cocktail party, etc.). In both cases, intelligibility is measured and parameters such as S/N ratio, or the sound pressure level (SPL) varied to assess trends and significance of factors concerning the particular processing method's effect on intelligibility.
Intelligibility. Listeners wearing binaural headphones (TDH49) were seated in a sound-proof chamber (Industrial Acoustics Co. 10-2060). Previously recorded lists of phonetically-balanced (PB) monosyllabic words (Egan, 1948) or Harvard sentences (IEEE, 1969) served as test material.

The speaker of the PB words was Phil Herman; speaker of the sentences was David Ackroyd. The monosyllables were presented with no carrier phrase. The listeners scored their own answer sheets and so were informed of their performance and the nature of their errors.

Two male listeners served as subjects: the author (EH), age 28, and the supervisor (PZ), age 32. EH has a slight hearing loss above 4 kHz, but normal thresholds at other frequencies. PZ has clinically normal hearing.

Long-term RMS voltage of speech. A method similar to that employed by Thomas and Sparks (1971) was used to determine the average RMS voltage level of speech. The speech signal is stored on the screen of a storage oscilloscope for an entire list of 50 words (or its equivalent for babble, etc). The vertical width of the stored band, excluding rare peaks that occur about one-in-a-thousand, is read out and taken to be the peak-to-peak value. This value is divided by 10 (20 dB, 14 for the peak factor for speech, and 6 dB for one half the peak-to-peak value) to obtain the RMS value of the speech. This measurement was done
with a list (No.1) and checked with a few other lists. Inter-list differences were less than 2 dB, and so were taken to have the same long-term level.

The RMS value of the Random Noise (white) was obtained with a Ballantine RMS Voltmeter, or an HP Spectrum Analyzer (Hewlett-Packard 3582A) when a VRMS/√Hz figure was desired.

**Sound Pressure level.** All SPL's are re. 0.0002 dynes/cm.² Our headphones had been previously calibrated, and known to produce a signal at 110 dB SPL at the listener's ear when 1V is applied. The headphones were assumed to be acceptably linear over the voltage range employed.

**Experiment I**

Measuring the Spectra of Clipped Vowels

As stated above, the rationale for the 3-band clipper comes largely from the desire to process the formant regions separately so that formant positions are maintained in their frequency range, and their peaks automatically equalized in amplitude. In this series of measurements is illustrated this approach through spectral measurements of differently-processed vowel sounds. A steady-state vowel was electronically synthesized using a Bell Telephone "Speech Synthesis" kit. This device consists of a square-wave buzz source followed by three cascaded formant resonators.
The effect of a clipper on a single-formant vowel is shown in Diagrams 1 and 2. Note that the resonance has been sharpened by the clipper. This effect is expected from the rule-of-thumb stated above about the larger components suppressing the smaller components in a clipper. The effect of an IPC-er on a three-formant vowel is shown in Diagrams 3 and 4. Here it is seen that the first formant is preserved but the second and third formants are obscured. Diagram 5 shows the effect of IPC'ing on the three-formant spectrum after it has passed through a 2-pole, 1100-Hz HP filter (Thomas and Niederjohn's "Optimal Filter" described in Section 2). The effect of this filter is to attenuate the first formant so as to bring the amplitude of the first and second formants more into balance. As a result, in the clipped spectrum in Diagram 5 both the 1'st and 2'nd formants are maintained. These examples should illustrate that, if the goal is to preserve formant locations, it is desirable first to separate the formant regions by filtering, clip each band, and then filter out the out-of-band components.

The action of the three-band clipper is seen in Diagrams 6 and 7. The formant peaks are maintained in frequency (approximately), and equalized in amplitude. Note that in the first formant region the maximum component has shifted to the next lower harmonic.
$c = 0.05 \mu s$

1st resonance

"Single Formant Input to IPC."

Diagram 1

10 dB

Input

Page 33

5 dB

Frequency

0

1 kHz

2 kHz

3 kHz

4 kHz

5 kHz
1st resonance

output of IPC.

"Output of IPC with a Single Formant Input"

Diagram 2

Page 34

10 dB

15 dB
"3-formant Vowel, Input to IPC."

Diagram 3

Input (from Synthesizer)
Diagram 4

"Output of IPC with 3-formant Input"

Output of I.P. Clipper w/o Eq. filter

frequency (Hz) →

0 1 kHz 2 kHz 3 kHz 4 kHz 5 kHz

10dB

5dB
Output of clipper with opt. filter

Output of "optimal Filter"/IPC with 3-formant Input

Diagram 5

Frequency (kHz) →

Page 37
"Output of 3-band Clipper"
(Vowel Synthesizer is Input)

Diagram 7

Frequency (Hz) ->

dBV ->
Experiment II

Replication of Thomas and Ravindran's Experiment

METHOD: The experimental procedure of Thomas and Ravindran (1974), and described in the Previous Experiments section, was followed as closely as possible.

RESULTS: Diagram 8 shows the intelligibility (%) vs. S/N ratio (dB) for curves representing the results of Thomas and Ravindran, and the present study. The results are very different between the different experimenters. Thomas and Ravindran showed the processed speech more intelligible than unprocessed, and we have showed the un-processed speech more intelligible than processed, at all S/N ratios investigated. If our measurement of speech RMS voltage was inaccurate, then our data could be shifted horizontally to compensate. However, the fundamental differences between Thomas and Ravindran's results and our own would still not be resolved.

Experiment III

Intelligibility of 3-Band-Clipped Speech that is Already Corrupted with Babble.

Block Diagram

[Diagram of a block diagram showing a tape signal input, a babble input with an attenuator, and a 3-band clipper output.]

[Diagram not described further in the text.]
Replication of I.B. Thomas's
"Intelligibility Enhancement of Already Noisy Speech Signals"
JAES 22 (1974)

Diagram 8

METHOD: The formant-amplitude-control knobs were adjusted informally by PZ for best intelligibility and quality. The other listener (EH) found those settings acceptable. The settings are First Formant 56 (-16 dB), Second Formant 50 (-16 dB), Third Formant 65 (-22 dB). The input gain knob was set so that the "peak indicator" LED came on very infrequently.

Both the signal and babble's RMS voltages were measured at the input to the summer by the storage-oscilloscope method to establish a zero dB S/N ratio. The HP attenuator was then used to change the S/N ratio.

Phonetically-balanced (PB 50) word lists were used (Egan, 1948). Long-term SPL of the speech signal at the earphones was 80 dB for all tests of this experiment.

RESULTS: Diagram 9 gives the intelligibility results. Each point ('X' or 'O') represents the score of a 50-word test (one list) for either EH or PZ. The modified speech is always less intelligible than unmodified. An increase of about 10 dB of the S/N ratio is necessary for equal intelligibility of modified and unmodified speech over the intelligibility range of 20 to 65%. Both curves in the diagram have roughly the same shape.
Intelligibility of Already Noisy Speech—
with and without modification.
Noise = Babble

Diagram 9
P.Z. = O
E.H. = X
Experiment IV

Comparison of Processing Schemes for Reception by Listeners with Limited Dynamic Ranges.

Block Diagram

It was desired to test the various speech-processing schemes with listeners who have small dynamic ranges. However, for this preliminary testing we chose not to test hearing-impaired subjects. Further, a masking noise could not be used to elevate the detection thresholds of normal-hearing subjects and reduce the dynamic range to 20 dB because such a noise would be painfully intense. Thus, the discomfort threshold was simulated artificially with a visual indicator that would light when the signal exceeded a specified level.

METHOD: Threshold elevation in the normal-hearing listeners was produced with a random white noise mask. Diagram 10 gives pure-tone (generated with the MAICO audiometer) detection thresholds in headphone voltage (measured with the HP spectrum analyzer) vs. frequency for both PZ and EH. These detection thresholds were performed with "quiet" presentation (pure tone only), and "noise masked" with a noise spectrum level of
Diagram 10

Masked with 30 dB S.P.L. / \sqrt{\text{H}_3}
white noise

Thresholds in quiet
Better of the ears

Audibility Thresholds vs. Earphone Voltage

of E.H. - x
P.Z. - o

Earphone Voltage (dBV)

frequency (\text{H}_3)
-80dBV Hz, which converts to 30 dB SPL. This masking noise elevated our detection thresholds as shown in the diagram.

To simulate a reduced dynamic range, a "discomfort threshold" was imposed artificially. This threshold was an LED that would light whenever the wideband speech signal exceeded a certain voltage. For the tests performed, this "pain" indicator was set to light when a 1 kHz pure-tone was at either -45dBV or -35dBV (as measured on the HP spectrum analyzer). This represents useable dynamic ranges of about 15 and 25 dB, respectively.

With this artificial reduction of dynamic range, it would be unfair to compare the intelligibility of speech processed by the 3-band clipper to that of unmodified speech. Because of the slope of the long-term speech spectrum only the low-frequency portion of the spectrum would fit into the dynamic range. An appropriate comparison would be to speech that has been spectrally tailored to fit in the listener's dynamic range. Since detection thresholds are nearly constant, and the "pain" threshold is constant, the necessary tailoring amounts to a whitening of the long-term speech spectrum. The frequency-gain characteristic that whitens speech was determined by measuring the spectrum of continuous babble and adjusting the settings of a G.R. 1925 one-third-octave multifilter (Diagrams 11, 12, and 13). The same settings of the multifilter were used to whiten speech for the intelligibility tests. This "whitened" babble can be compared with Diagram 14, the long-term spectrum of
"Long Term Spectrum of Babble"

Diagram II
"Long term Spectrum of Babble"

Diagram 12
"Equalized Speech Spectrum" of long term babble, using the G.R. #925 multifilter.

Diagram 13
3-band-clipped babble.

The third system that was compared was the "Optimal Filter/Infinite Peak Clipper" (OF/IPC) described by Thomas and Niederjohn (1968), and described here in Section 2.

The masking noise RMS voltage (which caused the detection threshold shift) remained constant throughout the experiments. When different devices were chosen for speech processing (one-third-octave multifilter, 3-band clipper, or OF/IPC), different signal output voltages would be presented to the summer. To compensate for this output voltage level difference, an attenuator located in the sound chamber was adjusted so that the LED pain indicator came on only very infrequently. This adjustment was done using continuous discourse speech material, was not allowed to be altered during the testing sessions by the listeners, and has no effect on the pre-set voltage threshold that the pain indicator is sensitive to. This adjustment was meant to simulate a hearing aid user's adjustment of output level to prevent pain.

Intelligibility test material was either PB 50 word lists (Egan, 1948), or Harvard sentences (IEEE, 1969). The listeners were EH and PZ. Each entry in Table I presents results with either 50 PB monosyllabic words, or 10 Harvard sentences. The 15 dB range was not employed with the PB words because nothing was intelligible after adjusting the overall level for infrequent triggerings of the "pain indicator". This was true with either
whitened speech or 3-band clipped speech. The OF/IPC system was not investigated with PB words.

Table I

PB words

<table>
<thead>
<tr>
<th>Dynamic Range</th>
<th>Whitened</th>
<th>3 band-clipped</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PZ</td>
<td>EH</td>
</tr>
<tr>
<td>25 dB</td>
<td>4%</td>
<td>4%</td>
</tr>
</tbody>
</table>

Harvard Sentences

<table>
<thead>
<tr>
<th>Dynamic Range</th>
<th>Whitened</th>
<th>3-band Clipped</th>
<th>OF/IPC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PZ</td>
<td>EH</td>
<td>PZ</td>
</tr>
<tr>
<td>25 dB</td>
<td>18%</td>
<td>16%</td>
<td>98%</td>
</tr>
<tr>
<td>15 dB</td>
<td>0%</td>
<td>0%</td>
<td>6%</td>
</tr>
</tbody>
</table>

The results clearly indicate that speech processed by either the OF/IPC or the 3-band clipper is much more intelligible than whitened speech. When the usable dynamic range was 15 dB, OF/IPC was superior to 3-band clipping. This is apparently due to the OF/IPC having a peak factor of 1, and 3-band clipper signals
having a peak factor of about 4 dB (measured on the storage oscilloscope). The larger peak factor is due to filtering after clipping in the 3 bands. These three filtered signals are then added together, which also increases the peak factor.
There are many types of processing schemes that could achieve amplitude range reduction of speech. It is important to keep in mind that the relevant "figure of merit" for such a system is the intelligibility of the processed speech relative to the reduction in amplitude range.

This study was concerned with two types of signal processing that were shown to effect a reduction in the amplitude range of speech without severe loss of intelligibility. The two signal processing systems were the OF/IPC system of Thomas and Neiderjohn (1968), and the 3-band clipper developed here. Speech intelligibility with these two processing schemes was compared to that of either unmodified, or frequency-equalized (whitened) speech.

In the first phase of the study the effects on intelligibility of noise at the inputs to the various systems were compared. This study was unable to reproduce the results of Thomas and Ravindran (1974), which showed enhancement of intelligibility of already noisy speech. With this system, listeners in prior studies (Licklider and Pollack 1948, Thomas and Ravindran 1974) were able to achieve close to 100% intelligibility of processed speech with high (>25 dB) S/N ratios. However, we (PZ, EH) were able to achieve only about 75% intelligibility with high S/N ratios. So, perhaps the reason why
we were unable to reproduce Thomas and Ravindran's (1974) results is that we are not yet well enough trained with IPC'ed speech. To resolve the inconsistencies brought out here, further testing of that system should be performed to see if people can learn to recognize the HP/IPC'ed speech better with listening experience.

We were able to produce results similar to those reported in the literature with experiments that had noise added to the processed speech. In this case, the intelligibility of processed speech was vastly greater than that of unprocessed speech. The processing schemes effective here were the OF/IPC'er and the 3-band clipper. With a listening dynamic range of 25 dB, the two systems allowed about the same level of intelligibility. When the range was lowered to 15 dB, the OF/IPC'er yielded greater intelligibility than the 3-band clipper. This is apparently due to the larger peak-factor of speech processed by the 3-band clipper compared with the 0 dB peak factor of speech processed by the OF/IPC.

In the Thomas and Sparks (1971) evaluation, their HPF/IPC'er system was found to be of great value to several of their hearing-impaired listeners. Two listeners with only low-frequency hearing produced lower scores with the clipper system than linear amplification (uniform gain across frequency). Perhaps the 3-band clipper could be of assistance here, since it has 1st formant speech waveforms present in the output.
Because of the positive results shown by both the 3-band clipper and the simpler OF/IPC-er with simulated reduction of dynamic range, future experiments should include evaluation of intelligibility with listeners who have severe sensorineural hearing loss and very narrow dynamic ranges.
APPENDIX: ELECTRICAL DETAILS OF THE 3-BAND CLIPPER.

The functional block diagram is shown in diagram A1. A linear input pre-amp provides adjustable gain to accommodate different input devices. Monitoring the output of the pre-amp is an LED "peak indicator" to indicate when the pre-amp is clipping and thus introducing unwanted distortion.

The pre-amp's output is applied to filters that separate the input signal into the first three formant regions of speech (1'st band = 200 to 900 Hz; 2'nd band = 900 to 2800 Hz; 3'd = 2.8 kHz to 6 kHz). Audio taper controls allow adjustment of the input to each of the three bands. Ten-turn potentiometers control the cutoff frequencies of the filters.

The outputs of the initial filters are sent to an infinite peak-clipper, consisting of an AC amplifier (gain = 12), followed by a Schmitt-trigger (Hysteresis 'dead zone' = 74 mV), which insures that the clipper's output is either high or low (i.e. squares off the signal).

Each IPC-er's output is filtered to eliminate out-of-band distortion products.

Finally, a summing device adds the three filtered/IPC-ed/filtered signals together. Output impedance of the summer is 30 ohms.
All of the output signals present at Banana jacks (test points labeled 1 thru 9 on the block and electrical schematic diagrams) are short-circuit proof, in the sense that no damage will be done to the circuit.

Performance Specifications

<table>
<thead>
<tr>
<th></th>
<th>Min.</th>
<th>Typ.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive Supply</td>
<td>8</td>
<td>9</td>
<td>11 Volts</td>
</tr>
<tr>
<td>Negative Supply</td>
<td>-8</td>
<td>-9</td>
<td>-11 Volts</td>
</tr>
<tr>
<td>Clock Supply</td>
<td>4</td>
<td>5</td>
<td>5.3 Volts</td>
</tr>
<tr>
<td>Current-Positive Supply</td>
<td></td>
<td></td>
<td>100 mA</td>
</tr>
<tr>
<td>Negative Supply</td>
<td></td>
<td></td>
<td>100 mA</td>
</tr>
<tr>
<td>Clock</td>
<td></td>
<td></td>
<td>100 mA</td>
</tr>
<tr>
<td>Input Impedance</td>
<td></td>
<td></td>
<td>10k ohms</td>
</tr>
<tr>
<td>Output Impedance</td>
<td></td>
<td></td>
<td>30 ohms</td>
</tr>
<tr>
<td>Output Impedance at test points</td>
<td></td>
<td></td>
<td>5k ohms</td>
</tr>
<tr>
<td>Output Current</td>
<td></td>
<td></td>
<td>18 mA typical</td>
</tr>
<tr>
<td>Output Signal</td>
<td></td>
<td></td>
<td>1 VRMS</td>
</tr>
<tr>
<td>Output Noise</td>
<td></td>
<td></td>
<td>3mV Max.</td>
</tr>
</tbody>
</table>

Input Dynamic range (with gain control
at a fixed setting) 57 dB

Band Frequency Ranges

1'st 200 -- 900 Hz

2'nd 900 -- 2800 Hz

3'd 2800 -- 6000 Hz

Output Signal Peak Factor ~4 dB

Component Selection Criteria

Because sharp definition of formant bands is important, "EG and G Reticon" Mos-monolithic switched-capacitor filters (data sheet enclosed pg. ) were used as formant-band filters. The R5609 (seven-pole, six-zero elliptic low-pass), and R5611 (five pole Chebyshev) were used in series to form a band-pass filter. Thus, the cutoff frequencies of the pass-band can be independently adjusted.

The Reticon switched-capacitor filters have the draw-back of clock residue (25 to 100 mV rectangular pulses riding on the signal), and DC offset (~100 mV). In this application, clock residue is removed by a single-pole RC low-pass filter, and the DC offset by AC coupling.
The other critical active component is the HA -2605 (Harris semiconductor) — a "Wide Band, High Impedance Operational Amplifier" (bipolar monolithic). It is internally compensated, and the following specifications show its suitability in this application:

1) Input Impedance 500 Mohms

2) Gain (DC) 150,000

3) Slew Rate 7V/μS

4) Settling time (to 0.1%, large signal: 1.5μS)

5) Output short circuit protected

6) Power: 90 mW total, with supplies at 15V.

These features allow its use as a low-power high-quality audio amplifier. Its input impedance and gain allow flexible choice of feedback component values and ease of interfacing (cascading) for multistage design.

Analysis of Individual Circuit Blocks

I. Clocks

Diagram A2 shows one of four identical clocks necessary to drive the Reticon filters. The clock should have a sharp-edged output with fast rise/fall times to provide precise definition of
Clock Oscillator — using the SN74123N

Diagram A2

1) Circled Numbers are Pin Numbers
2) Potentiometers are 10K, 10 turn Beckman Cermet.
3) Clock A = 2700 pF  Nominal Frequency 90 kHz  R 4.3K
   Clock B = 750 pF  280 kHz  4.3K
   Clock C = 510 pF  450 kHz  4.3K
   Clock D = 160 pF  1.4 MHz  2.1K

page 62
the -3 dB corner frequency. A TTL level voltage swing is adequate. A duty cycle of 50% was easy to implement, and desirable when high frequency clocks are required. For these reasons, the SN74123n—"Dual Retriggerable Monostable Multivibrator" was used. The resistor R on the clock output is to protect the Reticon filter when the supplies are off and the clock signal is on.

II. "Peak Indicator", or AC Voltage Threshold Detector (diagram A3)

This device is used to visually indicate when a selected AC voltage has been reached and maintained for a short length of time. The input signal is selectively attenuated, and applied to a Schmitt-Trigger (with ~4V hysteresis). The Schmitt-trigger output is AC coupled, and rectified (with a 100 ohm resistor to protect the diode from high current surges). The rectified voltage is stored on the 1 uF capacitor, and measured with a transistor amplifier / LED indicator. In operation, if the Schmitt-Trigger's threshold is exceeded, the output rectangular wave will be rectified, stored, and used to drive the visual indicator. It can be made more sensitive by lowering the 2 kohm resistor in the Schmitt trigger.
AC Voltage Threshold Detector
The input signal for this device is normally taken at the output of the 2-pole low-pass filter on the Pre-Amp schematic.

III. Pre-Amp (diagram A4)

This amplifier has an input audio taper attenuator (see diagram A5 for the calibration curve) for input gain control, to allow easy interfacing with various external devices, and to adjust so that the distortion LED is not illuminating. It is DC coupled to an AC amplifier (gain = 76) with gain down 3 dB at 160 Hz. Next is a 2-pole butterworth active HPF (down 3 dB at 200 Hz). This is followed by a 2-pole butterworth LPF, down 3 dB at 6 kHz). These filters act to set the system’s high and low frequency pass band points (the lower frequency edge of the first band, and the high frequency edge of the 3’d band). Later in the system, the other edges of the 3-formant bands will be defined by Reticon switched-capacitor filters.

To control the amplitude in each band are three audio-taper attenuators that feed the pre-amps’s output signal to the three channel’s. See diagram A5 for the calibration curve of the control.
Diagram A4

Input

10K Audio

1μF 1K 75K

27K .02

56K

10K 10K .002 .001

100K Audio Taper

a b c

Pre-Amp

CHANNEL
Diagram A5

2nd Formant Knob

3rd Formant Knob

"Input Gain" Knob

1st Formant Knob

Page 67
IV. 1'st Channel (diagram A6)

Pre-amp signal 'a' is input to a Reticon LPF, which defines the higher-frequency edge of the first formant. Test point 1, the "First Formant Output", is taken at the output of the anti-clock residue filter (fn = 1234 Hz). See diagram A7 for the amplitude vs. frequency response curve of this test point, relative to the input of the device. The R5609 also has DC offset, which is removed by the AC coupling, and the (gain = 12) AC amplifier. The amplifier's output is DC coupled to the Schmitt trigger. (I would recommend AC coupling here, to totally eliminate the DC output offset of the AC amplifier). With nominal supplies, the hysteresis "dead zone" is 70 mV. The output of the Schmitt trigger is test point two, whose amplitude vs. frequency curve, relative to the input of the total device, is shown in diagram A8.

The rectangular Schmitt trigger output is passed thru an anti-aliasing LPF before being filtered by a Reticon LPF, to remove the higher harmonic distortion products of the rectangular waves. Test point 3's amplitude vs. frequency response relative to the input of the total device is given in diagram A9. The post- IPC filtering of distortion products below 200 Hz was not considered necessary because of their relatively low amplitude, and low contribution to articulation.
"Response of the 1st Formant test point 1"

Diagram A7  page 70
"Response of the Peak-Clipped 1st Formant, test point 2"

Diagram A8 page 71
"Response of the 1st Formant, Peak-Clipped, and filtered Test Point 3"
V. 2'nd Channel (diagram A10)

This band has Reticon high-, and low-pass filters to set the range of the 2'nd formant. Anti-clock residue filters have a natural frequency of 3700 Hz. The AC amplifier / Schmitt trigger circuitry is identical to that of channel 1. Reticon high-, and low-pass filters are used on the IPC'ed output signal to remove higher harmonics and distortion products that are outside of the second-formant frequency range. Test points 4, 5, and 6 are the "2'nd formant", "2'nd formant Peak Clipped", and "2'nd Formant PC/filtered" signals. Their amplitude vs. frequency curves relative to the pre-amp's input are given on diagrams A11, A12, and A13.

VI. 3'd Channel (diagram A14)

An input Reticon HPF is used to define the lower-frequency edge of the 3'd Formant. The AC amplifier / Schmitt trigger circuitry is identical to that of channel 1. A Reticon HPF after the IPC-er is used to prevent low-frequency distortion products, caused by the IPC-er, from appearing in the second-formant frequency range. Higher harmonics (above 6 kHz) are removed with a single-pole RC anti-clock filter \((f_N = 7400 \text{ Hz})\), and the single-pole anti-aliasing filter \((f_N = 7400 \text{ Hz})\). Test points 7, 8, and 9 represent the 3'd formant, 3'd formant IPC-ed, and 3'd formant IPC / filtered. Their amplitude vs. frequency curves,
"Response of the 2\textsuperscript{nd} Formant
Test point 4"

Diagram All

Page 75
"Response of the Peak-Clipped 2nd Formant, Test Point 5"
"Response of the 2\textsuperscript{nd} Formant,
Peak Clipped, and filtered
Test point 6"

Diagram A13

page 77
relative to the pre-amp's input are given on diagrams A15, A16, and A17.

VII Summer (diagram A18)

This circuit adds the signals from the three bands, and divides the sum by 3 (to prevent clipping in the op-amp summer).

The amplitude vs. frequency curve for the complete three-band-clipper (output/input) is given in diagram A19. The "dips" at 900 and 2800 Hz are apparently due to out-of-phase addition in the crossover regions.

The overall downward slope in the spectrum is due to anti-clock filters (single pole) on the Reticon's output. This effect could be reduced (at the expense of greater clock noise) by raising the fN of each filter. Or, an active filter (multi-pole) could be used instead, with a higher fN.

Recommendations to improve circuit:

1) As previously mentioned in the discussion of the 1'st channel, AC coupling is recommended between the IPC-er's amplifier and peak clipper.
Response of the 3'd Formant
Test point 7" (0-5 kHz)
"Response of the 3rd Formant
Test Point 7" (0 - 25 kHz)

Diagram A15-b Page 81
"Response of the Peak-Clipped 3'd Formant, Test Point 8"

Diagram A16

Page 82
"Response of the 3rd Formant, Peak Clipped, and Filtered Test Point"
Diagram A18

Summer A18

Audio Output

page 84
"Output Response of the 3-band Clipper"

Diagram A19

Page 85
2) The value of "R" (the clock signal feed-out resistor) on the clock schematic (diagram) should be lowered to make the clock signal more rectangular shaped (as opposed to exponential). The resistor is present to protect the Reticon filter from damaging current which would occur if the plus 9 and minus 9 volt power supplies are turned-off before the clock power supply. I think that these (high-valued) resistors are causing the spurious tones, clicks, and whistles sometimes present at the output. They could be lowered to ~400 ohms and still give some protection to the filters.

3) External offset (output voltage) nulling of the op amps in the IPC-er could be done to obtain smaller hysteresis, and consequently, greater input dynamic range.

VIII. "Infinite Peak Clipper" and I. B. Thomas and R. J. Niederjohn's "optimal filter" (diagram A20)

Also constructed was a battery-operated Infinite Peak Clipper to be used separately, or in conjunction with Thomas's "optimal filter". The IPC has a 10 kohm input impedance, and a sensitivity control to vary the gain of the signal (0 to 151) before it is made rectangular by a Schmitt trigger with a hysteresis "dead zone" of 88 mV. Current drain is about 3 mA from each 9-V battery, quiescent, to about 11 mA drain operating into a short circuit. Output impedance is 1 kohm.
"Optimal Filter" (use with 10k load)

"Infinite Peak Clipper"

Diagram A20  page 87
The "optimal filter" is a 2-pole Gaussian HPF with a -3 dB frequency of 1100 Hz, and an asymptotic slope of 12 dB/octave. It is used by plugging it into the input of the IPC-er, which provides the necessary 10 kohm load. The optimal filter has an input impedance of 500 ohms. The amplitude vs. frequency curve of this filter is given on diagram A21.
"Response of I.B. Thomas's Optimal Filter"

Diagram page 89

frequency (Hz) →
REFERENCES


