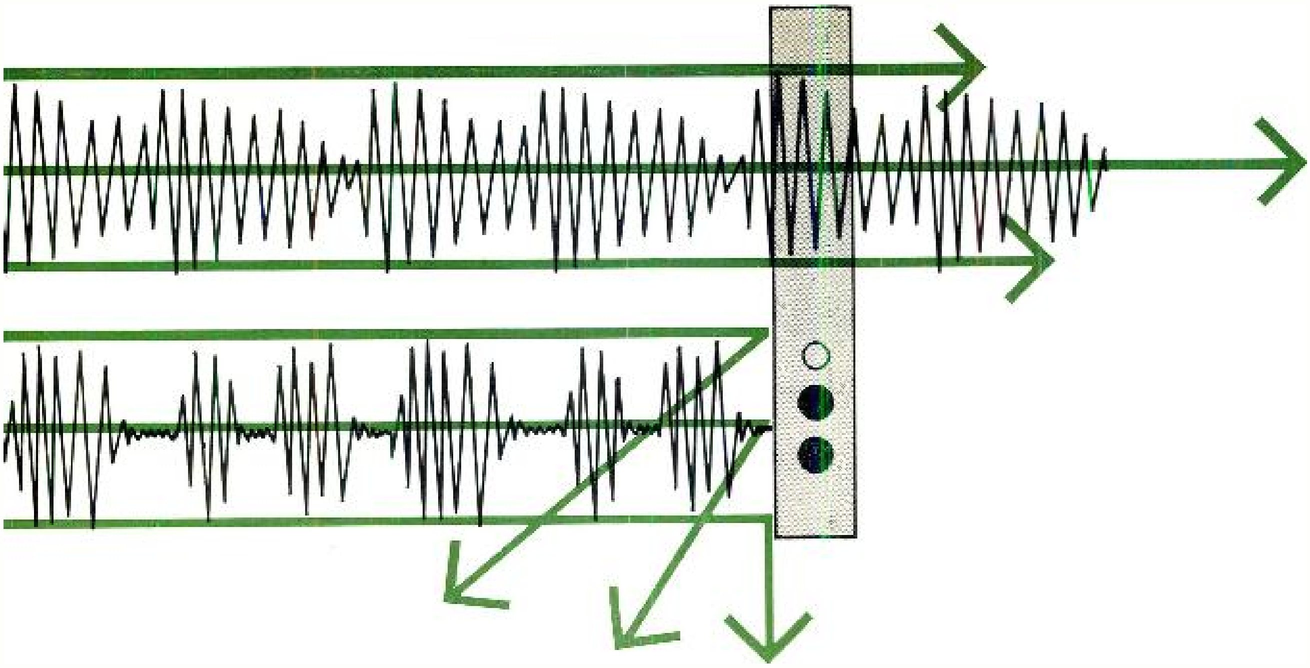


MUSIC/SPEECH DISCRIMINATOR

By FRANK D. GROSS

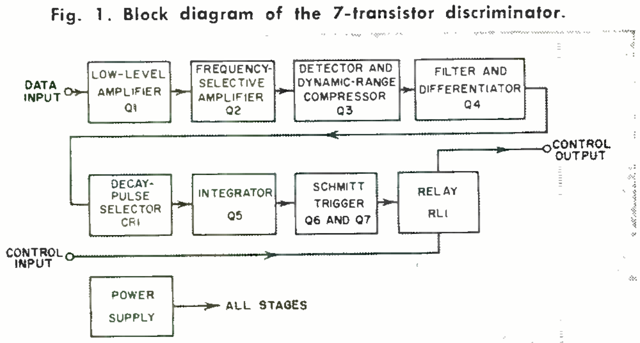
Design and construction of a device that can reject all speech and pass all music. Unit provides continuous background music either for listening or for recording.

CERTAIN radio broadcast commercial-elimination systems leave much to be desired in matters of performance as they are generally limited to a single FM station in any area, they are available on a rental-only basis, and they are generally limited to a specific receiver and audio system.

Several years ago two commercial-elimination systems made their appearance in a technical journal.[1,2] These devices were music-speech discriminators which provided a muting-type of control signal when speech only was present in the program material, and provided a pass-type of control system in the presence of music only. These devices operate well on any high-quality program source, *e.g.*, no singing commercials, and reasonable pauses between musical selections and speech passages. As this is typical of most FM-only stations, and certain AM stations, a wide range of program material operates well with these devices, giving 100% commercial and news elimination.

After constructing and testing one of these devices, the author decided that a solid-state music-speech discriminator (MSD) could be built with certain advantages over the previous all-tube systems, specifically with regard to size, flexibility, separation, temperature, and over-all system cost. The basic function of the two devices is similar, but the solid-state version achieves its operation through significantly different circuitry. The device is self-powered and is readily adaptable to a variety of audio systems. There are many applications for such a device. The MSD can provide continuous background music in such locations as doctors' and dentists' offices and retail stores, without being limited to one program source or a monthly rental fee. Attached to a home hi-fi center, the MSD can perform the same function. Used backwards, the MSD can be pressed into service as a "superpage" unit with certain plant audio systems or serve as a news monitor. The device may also be used to automatically "borrow" music while tape recording, without the customary editing and splicing. This device is a seven-transistor unit, the size of two books, and can be home constructed at a total cost of a little less than $30.00.

Fig. 1. Block diagram of the 7-transistor discriminator.



Theory of Operation

The nature of music and speech differ markedly in one respect: their decay rates. Decay, or the fall time in most music, is very gradual, indefinite, overlapping, and poorly damped while speech syllable (phoneme) endings are sharp, distinct, highly damped, and repetitive. Fig. 1 is a block diagram of the system that provides a control signal based upon this difference. The low-level audio signals from the program source are impedance-matched in the low-level amplifier stage, Q1. This is a conventional stage, needed because of drive requirements. The frequency-selective amplifier, Q2, has a frequency response that is limited to passing only mid-range signals (350-1400 cps). This stage amplifies speech signals with telephone quality and eliminates hum and sibilant sounds, bass instruments, hiss, noise, and the higher frequency musical components. These components would tend to mask or confuse the mid-range decay informa-

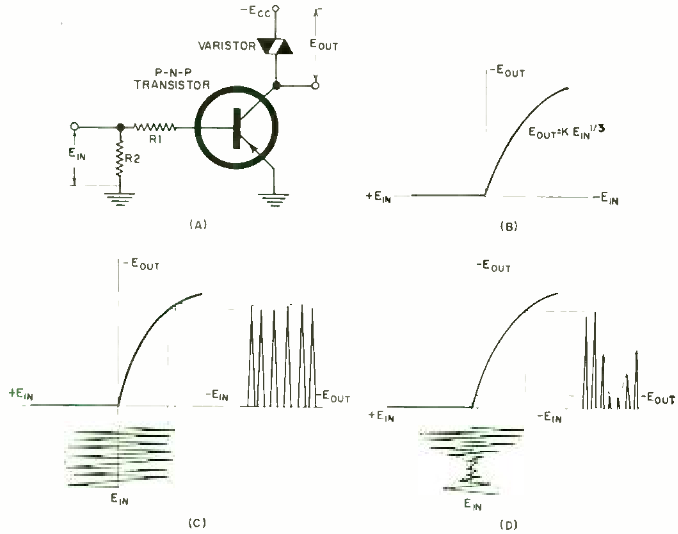


Fig. 2. (A) Detector and dynamic-range compressor circuit. Positive signals back-bias the stage and are lost. Negative signals are non-linearly amplified due to varistor characteristic. (B) Transfer characteristics of the stage. (C) Small amplitude variations, typical of music, are considerably reduced in amplitude. (D) Large amplitude variations, typical of speech decay, are considerably increased in size.

tion. This stage also eliminates any multiplex or commercial-killing ultrasonic signals that may be present. The third stage is a detector and dynamic-range compressor, consisting of an unbiased transistor and a varistor collector load. This stage removes the bottom half of the audio signal from further processing and amplifies signals in such a manner that the minor signal variations are reduced considerably, while the major signal variations are considerably amplified. This operation is explained and detailed in Fig. 2.

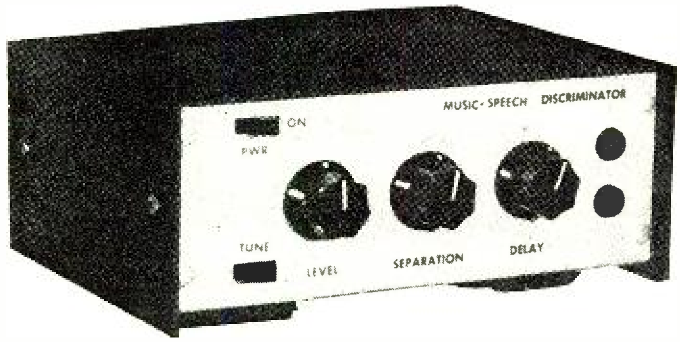
The net effect of this operation is to significantly reduce the gradual musical variations, while greatly enhancing the speech attack and decay variations. The amount of compression is on the order of the cube root of the input signal variations. This circuit performs considerably better than the drawdown limiter or logarithmic amplifier used in previous tube units.

The fourth stage is a filter and differentiator. The filter removes all audio frequencies and leaves only the audio envelope with its attack and decay information. The differentiator then extracts the attack and decay information, which appears as a varying series of pulses. This is simply an *RC* high-pass network. Diode *CR*1 is the pulse selector, which rejects all attack pulses, and all decay pulses below a critical adjustable magnitude.

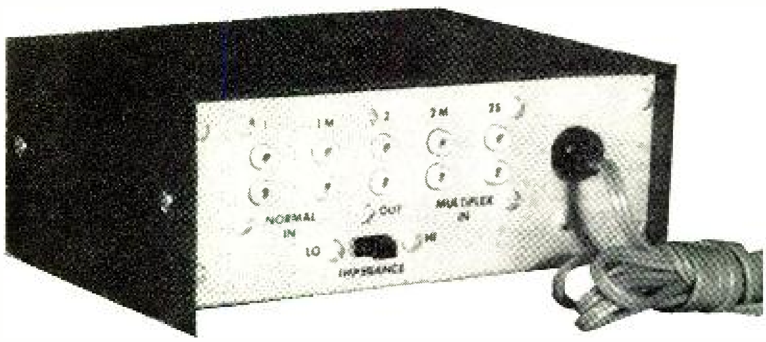
At this point in the circuit, speech consists of a series of sharp repetitive decay pulses, while music consists only of a random pulse or two of considerably smaller magnitude and much wider spacing. The actual music-speech discrimination has been accomplished. All that remains is the conversion of these decay pulses into a useful "on-off" type of control signal.

These decay pulses are caught by the integrator $Q5$, which rapidly charges a large capacitor, but turns off between decay pulses to force the capacitor to gradually discharge over a second path. This second path is the Schmitt trigger $Q6$ and $Q7$. When the average capacitor voltage exceeds a critical value, the control relay is energized. This capacitor also provides a delay after the last decay pulse to permit complete muting during a speech passage between words and phrases. The Schmitt trigger prevents relay chatter or marginal operation by always presenting a very high or a very low relay voltage. The operation of these stages is indicated by the waveforms of Fig. 3.

There are six controls in the circuit: three variable and three switched. A "Level" control establishes the level of the audio input signals. A "Separation" control determines the minimum level of decay information sent to the integrator, while the "Delay" control adusts the time delay of the Schmitt trigger, variable from .7 to 2 seconds. A "Tune" switch defeats the MSD while tuning the receiver. The remaining two switches are the "Power On-Off" and the "Impedance" selecting switch for the input audio connection.

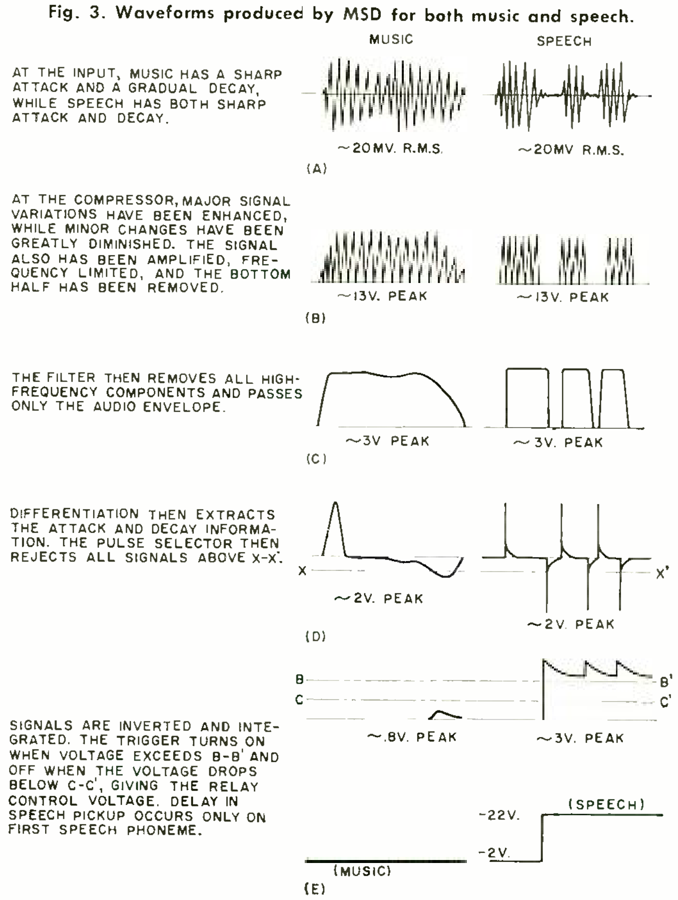


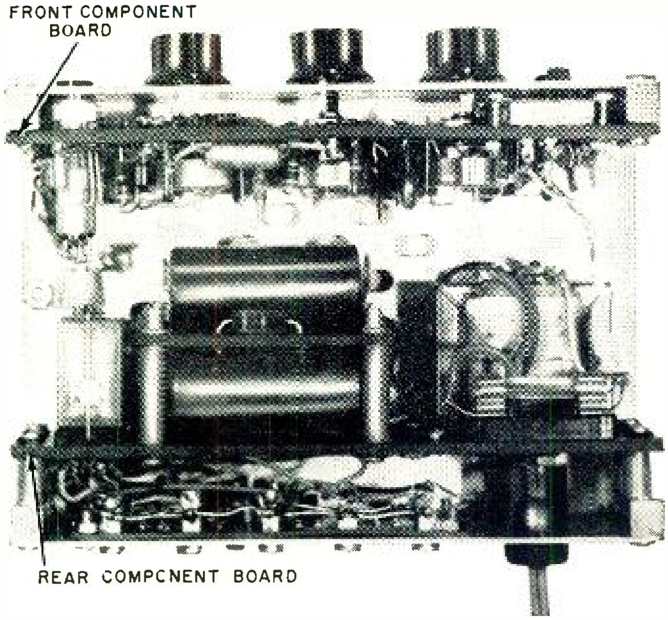
Front view of the MSD unit. The green and red indicator lamps at extreme right show presence of music or speech.



Rear view showing interconnecting jacks and impedance switch.

Fig. 3. Waveforms produced by MSD for both music and speech.





Top view with cover removed. Note use of two circuit boards.

The circuit diagram of the unit is shown in Fig. 4. All parts are conventional and low in cost. The transistors are industrial silicon types and are readily available.

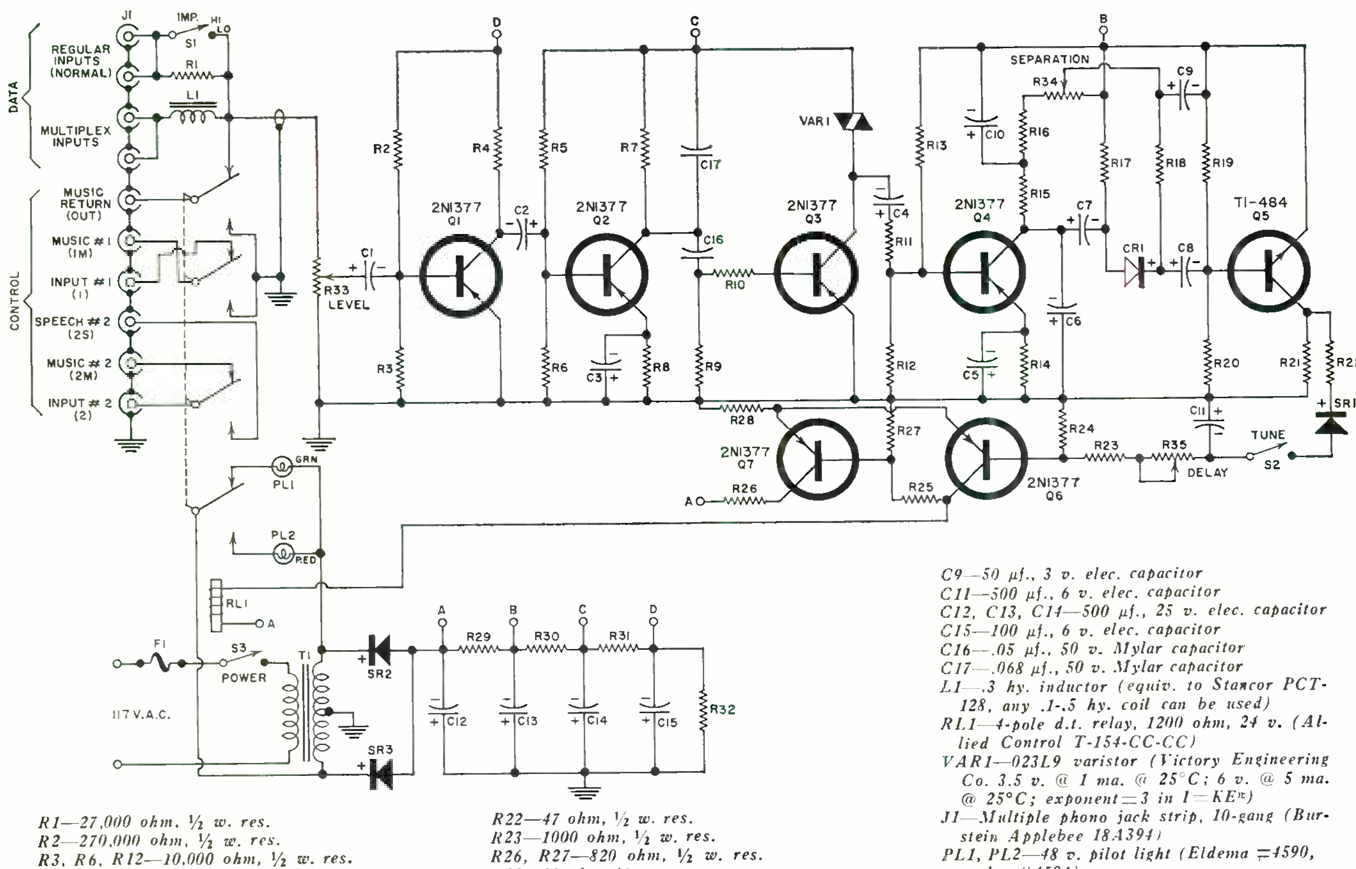
The flexibility of the circuit comes about from its variety of possible inputs and control functions. A dual regular input is provided. This permits return of the incoming audio from the tuner to the amplifier or preamplifier. An impedance switch matches the input to the MSD and has a "Lo" (5000-ohm) position for use after a cathode-follower or low-impedance line, and a "Hi" (33,000-ohm) position for use after a plate-coupled output. A dual multiplex input is also provided, which will use a low-impedance multiplex line without degrading the multiplex signals. *L*1, a .3-henry cup-core inductor, provides pre-emphasis equalization for the MSD only. The multiplex line is unaffected by the series compensation.

Four basic control functions are provided by the relay: (1) The input signal appears at an output jack during musical passages; (2) an isolated input is switched to one output jack during speech and a second output jack during musical passages; (3) an isolated input is switched to an output jack during musical passages; and (4) dual pilot lamps light red for speech and green for music. All four functions are completely independent and isolated from each other. All the input/output connections are accomplished by a ten-gang multiple phono jack—a surplus item. By proper use of this jack bay, the MSD will work into virtually any audio system.

The internal power supply is conventional, using two silicon rectifiers in a center-tapped configuration to provide d.c. circuit voltages of 25, 20, 15, and 6 volts to the various stages

(*Continued on page* 73)

Fig. 4. Schematic diagram and complete parts list for the music-speech discriminator constructed by the author.



*R*1—27,000 ohm, ½ w. res.
*R*2—270,000 ohm, ½ w. res.
*R*3, *R*6, *R*12—10,000 ohm, ½ w. res.
*R*4, *R*7, *R*15—4700 ohm, ½ w. res.
*R*5—150,000 ohm, ½ w. res.
*R*8, *R*14—470 ohm, ½ w. res.
*R*9—3900 ohm, ½ w. res.
*R*10—18,000 ohm, ½ w. res.
*R*11, *R*19—22,000 ohm, ½ w. res.
*R*13—180,000 ohm, ½ w. res.
*R*16—750 ohm, ½ w. res.
*R*17—12,000 ohm, ½ w. res.
*R*18—47,000 ohm, ½ w. res.
*R*20—680,000 ohm, ½ w. res.
*R*21, *R*24, *R*25—5100 ohm, ½ w. res.
*R*22—47 ohm, ½ w. res.
*R*23—1000 ohm, ½ w. res.
*R*26, *R*27—820 ohm, ½ w. res.
*R*28—22 ohm, ½ w. res.
*R*29—180 ohm, ½ w. res.
*R*30—270 ohm, ½ w. res.
*R*31—6200 ohm, ½ w. res.
*R*32—2700 ohm, ½ w. res.
*R*33—5000 ohm linear-taper pot
*R*34—750 ohm, linear-taper pot
*R*35—2000 ohm linear-taper pot
*C*1, *C*8—10 μf., 25 v. elec. capacitor
*C*2—1 μf., 25 v. elec. capacitor
*C*3, *C*5, *C*10—50 μf., 6 v. elec. capacitor
*C*4, *C*6, *C*7—4 μf., 25 v. elec. capacitor
*C*9—50 μf., 3 v. elec. capacitor
*C*11—500 μf., 6 v. elec. capacitor
*C*12, *C*13, *C*14—500 μf., 25 v. elec. capacitor
*C*15—100 μf., 6 v. elec. capacitor
*C*16—.05 μf., 50 v. Mylar capacitor
*C*17—.068 μf., 50 v. Mylar capacitor
*L*1—.3 hy. inductor (equiv. to Stancor PCT-128, any .1-.5 hy. coil can be used)
*RL*1—4-pole d.t. relay, 1200 ohm, 24 v. (Allied Control T-154-CC-CC)
*VAR*1—023L9 varistor (Victory Engineering Co. 3.5 v. @ 1 ma. @ 25°C; 6 v. @ 5 ma. @ 25°C; exponent = 3 in $I = KE^n$)
*J*1—Multiple phono jack strip, 10-gang (Burstein Applebee 18A394)
*PL*1, *PL*2—48 v. pilot light (Eldema #4590, socket #4594)
*F*1—½ amp., 250 v. fuse (AGC-½)
*S*1, *S*2, *S*3—S.p.s.t., 1 amp slide switch
*T*1—42 v. c.t. @ .5 amp power trans. (Femco TRP-2712 or larger equiv. Stancor TP-1 with two 18-volt secs. in series)
*CR*1—1N34A diode
*SR*1, *SR*2, *SR*3—500 ma., 600 p.i.v. power diode (1N2071 or equiv.)
*Q*1, *Q*2, *Q*3, *Q*4, *Q*6, *Q*7—2N1377 transistor (Texas Instrument)
*Q*5—T1-484 transistor (Texas Instrument)

Music/Speech Discriminator

(*Continued from page 38*)

through an *RC* type of filter network.

The low-level amplifier, *Q*1, is also conventional and is a common-emitter *p-n-p* transistor stage. The frequency-selective amplifier, *Q*2, is similar to the initial stage with the exception of the response-altering capacitors (*C*2, *C*16, *C*17).

The operation of *Q*3, the dynamic-range limiter and detector, is indicated in Fig. 2. A varistor does not directly obey Ohm's Law but instead has a voltage-current relation of:

$$I = KE^n$$

where *n* in this case is approximately three. Reversing this relation gives:

$$E = K'I^{1/3}$$

and in this particular circuit $E = E_{out}$ and $I = I_{collector}$. But collector current is equal to base current times the gain of the transistor, or: $I_c = H_{fe}I_b$ and I_b is determined mainly by the input voltage and *R*10. Combining all constants and substituting gives:

$$E_{out} = K'' E_{in}^{1/3}$$

Thus, although all signals are amplified considerably, the amplitude variation at the output of this stage is noticeably less than the amplitude variation at the input. This stage will operate properly only when driven hard. Audio peaks should produce a varistor voltage of around 13 volts.

The filter stage, *Q*4, is once again similar to the initial stage, with the exception of filter components *R*15 and *C*6. A portion of the collector load is bypassed by *C*10 to provide a variable two-volt bias supply for the pulse selector. The filter has a cut-off frequency of 12 cps. Differentiation is accomplished by *C*7-*R*17, providing a 50-millisecond time constant and extracting all attack and decay information. The pulse selector is a biased diode that will only pass those decay signals above a variable (0-2 volt) threshold level. All attack information is removed by this biased diode.

The integrator stage, *Q*5, is an *n-p-n* switch gated by the decay pulses. When on, this stage rapidly charges the integrator capacitor *C*11. When off, *SR*1 is back-biased, and the capacitor cannot discharge by this path. *C*11 can only discharge through *R*35, *R*23, *R*24, *Q*6, and *R*28. Current flow through *R*23 trips the Schmitt trigger and energizes the output relay. A detailed analysis of this stage is available elsewhere.[3] The trigger turns on with an input of 1.3 volts and turns off when input voltage drops below 1 volt.

Construction & Operation

The unit was designed in a package consisting of a U-shaped aluminum

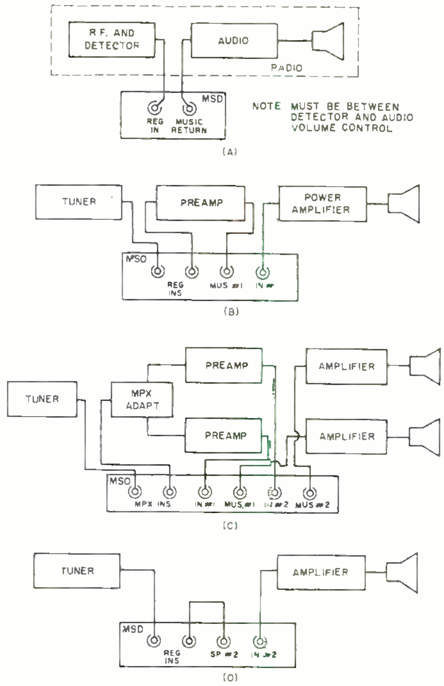


Fig. 5. System interconnections. (A) Connections to radio only. (B) Normal hi-fi or plant audio. (C) Stereo multiplex connections. (D) News-only monitor connections.

chassis and an overlapping 20-gauge steel cover. All circuitry except the input jacks and $L1$ are mounted on the two twin printed boards that face each other. The MSD measures 7″ x 5″ x 2½″ and is finished in brushed aluminum and heavy grey alligator wrinkle. The photographs indicate parts locations and the assembly technique.

Home duplication of this device should present no problems provided certain obvious modifications are made. It is recommended that a larger package be used, perhaps a "Mini-box" and conventional terminal boards, instead of the punched chassis and printed wiring used. No component location is critical as long as leads remain sensibly short and suitable shielding is provided for the low-level circuitry.

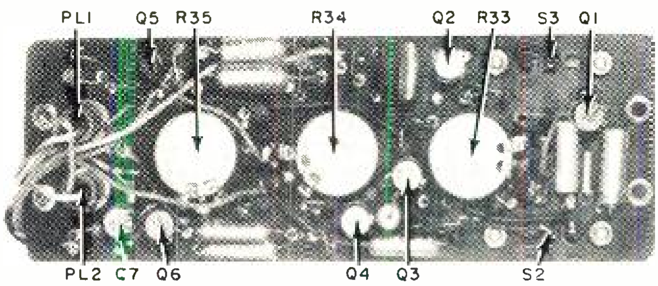
The MSD may be system-connected in many different ways, some of which are shown in Fig. 5.

After installation, the MSD is switched to the "Tune" position until suitable program material is selected. All three controls are placed in their minimum positions. The level control is brought up to a setting such that the MSD triggers on a musical selection and is then advanced slightly beyond this point. Next, the separation control is advanced until the MSD does not trigger on music. The delay control is then adjusted during the first speech passage to eliminate any chopping. The unit is now correctly adjusted. As long as program material does not change radically, the MSD will operate properly without resetting controls.

This device works surprisingly well on high-quality FM stations, rejecting all speech and passing all music. The only music that will ever fool the MSD are certain guitar, bass, harp, or piano percussion solos; these are quite rare and are flatly rejected as if speech. Most all singers are passed, as are most forms of music. Speech passages overriding a weak musical background will be passed or rejected depending upon control settings. Control settings are a bit more critical on the average AM station giving somewhat reduced reliability, resulting in an occasional chopped passage or the passage of a singing commercial.

REFERENCES

1. Roth, J.: "Automatic Music-Speech Discriminator Silences Speaker," Electronics, August 1955.
2. Ives, R. L.: "Music Pulse Analyzer Rejects Voice Signals," Electronics, April 1957.
3. Millman & Taub: "Pulse and Digital Circuits," McGraw Hill Book Co., New York, 1956. ▲



The front component board has transistors and controls mounted.

The rear circuit board has the power components mounted on it.