# **MORTIMER GOLDBERG**

# **The Art of Tape Editing**

Smooth transitions and the undetectable deletion of unwanted material are achieved through expert maneuvering of tape and precise knowledge of just where to break.

HE DEFINITION of *edit* is to prepare for use by compiling, arranging, emending, etc. As applied to recording tape, this article will explore the techniques utilized in the assembling of material into a cohesive unit of *sound*, in the form of speech, music, or both in combination. Music has its rhythm, tempo, melodic variation, speech its cadence and inflections; the musical beat is synonymous with the pacing of speech. An orchestra conductor *edits* music by controlling the orchestra into a continuous and harmonious blend of instruments and a tape editor's function is to combine his material with like precision.

## EDITING OF SPEECH

One of the initial aspects in the editing of speech is to correct the imperfections of speech delivery. These imperfections include, the mistake (fluff), natural hesitations, and the audible thought seeking sounds of "um" and "ah." These are most important in editing for broadcast, where the race against the clock requires maximum skill in editing to correlate all thoughts in an intelligible and natural sequence. The time factor doesn't necessarily apply to the editing speed, but more importantly, to the quantity of material information delivered in a prescribed time.

What is of prime importance in the natural sequence of speech is *pacing*. This naturally varies among individuals. Some speak rapidly, and some slowly. Some enunciate well, while others do not. Regardless of the type of speech pattern encountered, there is one thing that is common to all, and this is the *breath*. All words uttered are done so while

Mortimer Goldberg is technical supervisor with CBS Radio, New York. His distinguished 25-year career includes the editing of numerous documentaries. exhaling; when the air supply is depleted, a breath has to be taken before continuing. This is what natural pacing is dependent upon in each individual speech pattern.

## **BLOCK CUTTING**

For the moment, let us consider *block cutting*. This is the elimination of complete sentences, or thoughts, with no internal editing involved. In order to illustrate the cutting method necessary to preserve the natural pacing, refer to FIGURE 1.

A1 and Z1 are the first and last words, respectively, of sentence 1. B1 is the breath following sentence #1. The same configuration applies to sentences 2, 3, and 4. If we want to keep sentence 1, and cut to sentence 4, eliminating sentences 2 and 3, our sequence would be:

#### (1) Z1 + B1 + A4

This pseudo-algebraic formula indicates that the cut is made just prior to A2, maintaining the breath B1, and joining at the beginning of A4. Using the same method to indicate the *incorrect* means of making the splice, we might come up with

# (2) Z1 + B3 + A4

In formula 2, the cut is made just after Z1, and joined to the breath B3, which precedes sentence 4.

You might ask what difference it makes; a breath is a breath. In some cases, this may be true, but usually, it is not. There are two distinct reasons for it *not* being the case.

Because sentences 2 and 3 have been eliminated, the joining of sentences 1 and 4 is the only thing discernible to the ear. The *breath* that followed sentence 1 is in cadence with the speech delivery to that point, and will flow smoothly to the start of sentence 4. Consider the alternative of using breath 3. Sentences 2 and 3 might have had an increase in verbal exuberance, and hence uttered at a much quicker pace than what preceded. As a result, breath 3 would not be in keeping with the pace of sentence 1, and an unnaturalness would be noticeable.

The second reason is more significant. There is a natural reverberation to the voice that is not readily noticeable

36

when listening to speech, either live or recorded. The degree of this reverberation is dependent upon the surrounding environs. There is less in acoustically treated rooms, such as broadcast or recording studios, and more in hardwall or glass-enclosed areas. The voice of a public speaker in an auditorium, on a public address system, will have a great deal of echo as it reverberates though the auditorium from the loudspeakers. This, of course, is extreme reverberation. However, the sustenance of sound, from one syllable to the other, is ever present in normal conversation.

Referring again to FIGURE 1, Z1's final syllable reverberates into the breath, B1. The joining at the start of A4, as previously stated, will flow in natural sequence. Consider the other possible splice—Z1 + B3 + A4. Breath 3 (B3), contains the reverberation of Z3, and when spliced to combine with A4, in addition to the unnatural pace, will contain the extraneous sound of the Z3 reverberation content.

The difference between the two methods is that the first will be seemingly untouched, whereas the second will be unmistakably rcognized as altered or edited. The recognition is more acute to the trained ear, but still apparent nevertheless to the average listener.

#### FINDING THE EDIT POINT

The first requirement for locating the edit point is to have the ability to recognize the sound makeup of speech at very slow speed. This speed is not regulated by the electrical movement of the tape, but by manipulating the tape by hand, past the playback head. One method is to maneuver the feed and takeup reels in a to and fro fashion, with the mechanical brakes engaged. A fair amount of coordination is necessary for the rotation of each reel in unison. I have found that it is more effective to hold the takeup reel alone, and put the machine in rewind mode. The mechanical brakes are disengaged, and the hand pressure on the take-up reel prevents the tape from moving to the feed reel. This slight hand pressure enables you to move the tape freely, back and forth in a rocking motion, with one hand and with considerably more control. You are able to seek the identifying sound smoothly, without the squeak associated with the engaged mechanical brakes, when rotated in the stop mode. Release for cutting is made by merely pressing the stop button.

The identification of phonetic sounds, for the purpose of separating syllables, is the key to speech editing. The hard sounds of consonants are more easily recognized than the soft sounds of vowels. The sounds of B, K, and D are percussive in nature, because they are formed by contact— B with closed lips, K with the back of the tongue to the palate, and D with clenched teeth. The percussiveness, that makes the sound easily identifiable, is perceived by rocking the tape past the head at slow speed.

Let's use the word *corporal*, as an example. The start of the word will sound like a click. The P, in like manner, will sound like a puff. The softer intermediate segments of the word will not have as pronounced a definitive identification, as the C and P, but will be recognizable by tonal characteristics.

The vowels, A, E, I, O, U, are the most difficult to identify and separate because of their lack in percussiveness. Sounds of this nature are made with parted lips, and the character of the sound is produced by changing the form of the parted lips. Just voice aloud—a, e, i, o, u. Notice how each sound blends with the next, using limited mouth movement. It is helpful to listen to the material backwards, employing the rocking motion previously mentioned. The comparison of the sound, between forward and reversed motion, often clarifies the division between syllables. A vu meter is also often an invaluable aid. A separation, not discernible to the ear, is often indicated on the



Figure 1. The cutting method used to preserve naturalsounding pacing.

#### vu meter by a dip of the pointer.

It is essential to remember that when editing speech, it may be necessary to alter the *existing correct pronunciation*, for the sake of preserving "pace" and "naturalness." For example, *and* and *but* are two conjunctives used very frequently. In an original recording, the statement may go as follows:

Joe and David were going fishing and he decided, that is, David decided, to ask Bill to go along.

The portion to be eliminated is *He decided, that is,* so that the statement would be:

#### Joe and David were going fishing and David decided to ask Bill to go along.

Cutting at the end of the complete word and, and splicing to David is a true representation of skillful cutting, which includes all components. Although and he flows perfectly, and David has the pacing destroyed by the stuttering effect of two consecutive D sounds. This is what is termed a jump cut. The and should be cut before the d as an, and spliced to David. The pacing is preserved and the result sounds unedited. This situation will always occur when and or but is cut from preceding a word beginning with a vowel, to before a word beginning with a consonant. The described editing techniques in these cases will always prove successful.

Everyone, whether speaking in conversation or being recorded, usually will never consciously mispronounce a word without immediately correcting the error. When editing this mistake, or fluff, using the corrected word completely usually results in an unnatural sound. Mistakes of this kind frequently occur in tongue twisting multi-syllable words, i.e.

## After his application was received, he was given an applointment—appointment—for an interview.

The corrected word is fine in its entirety, but the problem exists in the separation from given an to appointment. There is reverberation in the voice that carries through from syllable to syllable. In this case, the N sound from

Figure 2. Method of threading short loop on a recorder.



db December 1976

AN, blends directly to AP of appointment. Voicing this aloud will confirm the fact. Even a skillful separation after the an, to the corrected appointment, results in an abrupt interruption of the word an. This is due to the remnant of its final N sound, blending into the A of appointment. Leaving the A, of the first word and cutting to the P of the corrected word eliminates the problem. The advantage, in this instance is twofold. Cutting to the percussive P is easier than cutting on the vowel A.

#### INFLECTION

My experience as tape editor, on numerous CBS Radio documentaries, has been that recorded material destined for the cutting room floor holds the key to a successful editing session. Considerably more recording time is devoted to an interview than the actual program time allotted to it. Consequently, the initial editing consists of block cutting, for the purpose of extracting pertinent subject matter. (In this process, the unwanted material is often allowed to run off on the floor. I always wind this material on to a reel.) This tape is referred to as *out takes*. The remaining material must now be condensed to comply with the allotted allowable time. The correction of fluffs, hesitations, etc., is an aid to this time factor.

The constant use of parenthetical phrases are very time consuming, and do not contribute to the desired program subject matter. It is desirable to eliminate them, but in the process, additional problems are created. For instance:

#### and so the building was scheduled for demolition, because as I said before, the property was sold to the developer.

In the interest of time, *because as I said* etc. is not necessary, since it was stated earlier. *Demolition* should end the statement, but the *inflection* is slightly up, due to the succeeding phrase which previously ended the sentence. The means for correcting the problem comes from an investigation of the *out take* reel. It is only necessary to locate a word with the same suffix *tion* (shun), with a

Figure 3. Extending the loop length and placing it through an empty 7-inch plastic reel.



downward inflection. It may seem to be a needle in the haystack situation, but not so. More often than not, the means of correcting mistakes or inflections by syllable substitution is available. The trick is to listen, recognize what is needed, and with practice, the surgery can be quite successful.

## BACKGROUND SOUNDS

In terms of tape recording, the "background sound" is considered as some action that is occurring while a person is speaking. This could be music, street noise, office machines, etc. There is always background sound present, regardless of conditions, even in the ideally quiet setting of a broadcast studio. If you open one microphone in an empty studio and listen to the monitor speaker, a sound is immediately apparent. There is nothing of significant level indicated on the vu meter, but sound is readily discernible to the ear, even at the monitor's normal volume setting.

Consideration of the most subtle ambient sound is of substantial importance in tape editing. The use of a blank piece of tape for a pause instead of the "quiet studio sound" would result in loss of sound or drop out. In radio broadcasting, numerous recorded inserts are pesented during the course of a newscast. These inserts are termed actualities, which are on-the-spot statements by persons describing situations, or informing the interviewer on matters pertinent to the news item being reported. Invariably, there is a backgound sound behind the statement. Ideally, when the statement ends, the sound is faded out before the newscaster resumes speaking. This is particularly true for telephone recordings, which, because of the nature of the medium, have an inherent sound of "line noise." For radio listening, that fade gives a smooth transition from item to item.

Television production rarcly bothers with this audio consideration, because the ear accepts an abrupt change when the eye views a new scene. In a t.v. newscast, there may be a picture of a battle going on with the associated sound, and at the conclusion of the segment, the scene switches

Figure 4. A loop weighted down on one side of the machine through an empty plastic reel.



38

to the anchor man with a simultaneous absence of the background sound just heard. Seeing the anchor man seated before a microphone, in an entirely different environment, seems perfectly natural to the viewer and the abrupt removal of sound is not particularly noted. Radio listening is entirely different because the ear, when functioning without the eyes, is much more sensitive to audible change. This fact can be easily proven by listening to t.v. audio without watching the picture.

Due to the consolidation of material from numerous sources which are combined through narration, the production of radio documentaries require utmost consideration of background sounds. On-location recordings often are made in different areas for the convenience of persons being interviewed. Different interviews, film clips, etc., pertaining to the sequence, consolidated through editing, have varied background sounds, ranging from very quiet to very noisy. These variations occur not only because of change in ambient sound levels, but because of the editing out of verbal content. The informational material may flow very smoothly, but the sudden changes in ambient sound result in unnatural-sounding sequences. In order to remedy this situation, a common background sound, excerpted from original material, is mixed behind the entire sequence. The nonsequitar backgrounds are thereby melded into a common, and essentially integral part of the presentation.

This same augmented sound may be mixed behind the commentator as he bridges the various topics of the interview, despite the fact that he wasn't present during the original recording. The purpose is not to deceive the listener into the belief that what they are listening to occurred as presented. The commentator invariably qualifies the fact that he was not present. The augmented sound merely contributes to the cohesiveness of the overall product.

#### THE LOOP

The method used to accomplish this background is the *loop*. Refer to FIGURE 2, which pictures the closed end of tape threaded through a recorder. This particular-sized loop is adequate for a background that has a distinct pattern which, when repeated regularly, would be normal, such as the sound of a trip hammer or pile driver. Access to sound in the clear is often only available in short segments, extracted from pauses during the interview. In order to obtain as much tape as possible to form the loop, it is best to dub the sequence at the highest available speed. If the original recording was made at  $7\frac{1}{2}$  in./sec., then dubbing at 15 in./sec. will give twice as much tape to form an adequate loop size. Even at 15 in./sec., numerous dubbings may be necessary in order to obtain sufficient tape length to form a playable loop.

When sounds are very general, and not as particularly distinctive as the previously mentioned trip hammer, the completion of an adequate loop is much more difficult. The vague murmur of voices is immediately transformed into a constant annoying repetitive pattern when placed in short loop form. To avoid this effect, as many varied sections as possible must be dubbed and edited together, so that close repetitive sections are eliminated. Consequently, considerably more tape is involved to form the loop. FIGURES 3 and 4 indicate how this lengthened loop can be made to play effectively by the weighting action of an empty 7-inch reel. In the majority of cases, the lengthened loop will serve adequately as a background without noticeable repetition.

However, if the nature of the available sound is such that frequent repetition is unavoidable, then further corrective measures are indicated. The lengthened loop is recorded on another tape and then the two mixed together to form the final loop. The only thing necessary in this recording process is to have the two recordings displaced in order to avoid synchronism. This step is invariably adequate in obtaining a suitable background, of authentic origin, for use in a highly edited segment. In order to free a tape recorder for use in the final production, the loop is often transferred to the cartridge. The cartridge is made by deactivating the primary cue so that a 30 or 40 second cartridge will run continuously without stopping.

#### EQUALIZATION & FILTERING

The use of frequency discriminating equipment is generally an aid to the tape editor. On-location recordings are usually performed by reporters utilizing cassette machines. Frequently, in the haste of obtaining an interview, the proper microphone technique is neglected, and numerous sections of the recording have an "off mic" quality. The fact that the person is off mic also results in a voice level lower than normal.

Level adjustment can compensate for the loss, but additional noise is encountered. The use of an equalizer, with a boost in the 2000-3000 Hz range, does wonders for the off mic section. This frequency spectrum is in the voice mid-range, often referred to as the *presence area*. The effect is such that the voice is seemingly brought closer to the microphone, with increased clarity. The pre-emphasis of the equalized section causes its loudness to compare with other non-equalized normal sections of the tape. Invisible intercutting between these sections therefore becomes more feasible.

Quite frequently, due to a malfunctioning recorder, the resultant material sounds muddy or muffled. By *attenuating* the frequencies in the 100-200 Hz range, and again boosting the 2000-3000 Hz range, the clarity of the recording becomes considerably improved to acceptable intelligibility. The elimination of 60 Hz hum, or other steady tone oscillation is accomplished by the use of a notch filter. This device discriminates sharply between discrete frequencies; if it doesn't completely eliminate them, it attenuates them to a more acceptable degree.

#### MUSIC EDITING

The tape editing of music requires the additional skill of recognizing discrete tonal identification, commonly known as perfect pitch. This is not to say that this talent is absolutely essential, but an editor with perfect pitch is able to intercut or blend musical passages more accurately, with fewer attempts.

The needs vary to an extensive degree. In a professional recording session, the one-time recording performance is not unusual, but more often corrections are made without starting from the beginning. The pickup is recorded a few bars prior to where the error occurred. In editing, a convenient point is selected for ease of joining. This may

Figure 5. The normal threading of a tape machine.





Figure 6. The reverse threading of a tape machine.

be a beat, a pause, or whatever effective area is convenient. The only requirement is that the joining be perfect.

More difficult situations occur when alterations of existing music recordings are necessary. This may involve the elimination of a chorus in order to shorten a selection, or excerpting passages for use as musical bridges in a dramatic sequence. Excerpted music is usually employed with a positive start, and then faded out when no longer needed. A more sophisticated treatment will have a natural ending at the desired time. This type of editing results in a newly created musical passage of specific time duration, to satisfy any and all needs.

Direct splicing within musical passages is the exception rather than the rule. A pick up on a "hold" note followed by, possibly, a drum beat, could readily be accomplished by a direct splice. More often, it is necessary to take the pickup point from one tape and blend, by mixing, from another tape to a third recording machine. This process requires two identical tapes of the original material, for the purpose of eliminating unwanted passages by selectively blending from one tape to the other. Unwanted portions of the first tape cannot be allowed to continue under the *blend point*. Blank tape is inserted at the end of the desired section, so that there is no possibility of unwanted sound. Very frequently, leader tape is used, so that there is visible indication of the end of the needed recorded portion.

The use of an echo chamber or a reverberation unit offers great advantage when musical separation of sounds is very short at the point of the desired blend. Cutting after the desired note will always sound abrupt, but re-recording this segment through an echo chamber will extend the note sufficiently in order to blend in *tempo* with the dedesired pickup point.

Precise timing is essential at the blend point in order to preserve the tempo. The tape must be in motion, at its proper playback speed, at the *mix point*, to avoid a pitch variation. Calculation of tape time to and from the point of mix can be accomplished by stopwatch timing. However, reaction time must also be considered and the entire process can become cumbersome. A very valuable technique is using the constant speed of the tape recorder transport to time the segment.

FIGURE 5 shows the normal threading of tape in between the capstan and pressure roller, and on to the takeup reel. The capstan revolves in a clockwise direction and controls not only the tape speed, but its direction. Examine the threading of the tape in FIGURE 6. The tape is placed up and around the pressure roller, and goes between it and the capstan from the opposite direction. The clockwise rotation of the capstan moves the tape *toward* the feed reel. In the play mode, with separate feed and takeup motors, there is no mechanical linkage and electrical braking causes each reel to accommodate the tape in whichever direction it moves. We therefore have a method of maintaining the normal position of the tape on the recorder, and playing it backwards at the operating speed.

Applying this technique to the mixing of two tapes requires the following procedure:

Tape 1 contains the first portion of the musical segment. At the point of blend, where no further sound is needed from tape 1, a length of leader tape is spliced in.

Tape 2 is prepared with leader tape prior to the pickup point for the mix with tape 1. Pot settings are determined in advance for proper level into the third tape recorder.

On Tape 1, the desired point of transition is set at the playback head. The pickup point on tape 2 is set at its playback head. Each tape is carefully threaded around the pressure roller and in reverse direction past the capstan, as previously described. Care must be taken that the desired blend point on each tape is at their respective playback heads. When all conditions are satisfied, the play button on each machine is pressed simultaneously. Both machines play backwards at their operating speed. The time duration of this backward play is arbitrary. The usual time is approximately ten to fifteen seconds, mainly for reaching a convenient spot for editing the new mix into the original material. When this is satisfied, both machines are stopped simultaneously. Carefully restore the normal threading of the tape without any change of position of the tape at the respective playback heads.

- (a) The third machine is started in record mode.
- (b) The mixer position for tape 1 is open.
- (c) The mixer position for tape 2 is closed.
- (d) Start tape 1 and tape 2 simultaneously.
- (e) When the leader on tape 2 enters the head housing, open the tape 2 mixer position.
- (f) When the blend occurs, tape 1 will be in leader tape. Close tape 1 mixer position to prevent any audio that is present on the tape after the leader portion from being recorded.

It is evident that the use of leader tape gives a convenient visible indication for the need to control the mixer positions. The most important characteristic in employing this technique is that aside from its placement accuracy, the tape recorders are amply stabilized in speed at the mix point.

## **VOICE & MUSIC SYNCHRONIZATION**

This backtiming technique is also an invaluable aid in mixing voice with music, when it is desired to have the music accent on a particular word or phrase. The use is very significant in the announcement of a program with a musical signature or theme. For example, the theme starts with a trumpet fanfare, followed by a tympani and cymbal crash before continuing into the melody of the signature. The announcement is to start after the cymbal crash. The music tape is cued on its machine at that point, and the start of the announcement cued at the playback head of its machine. The simultaneous backtiming is as previously described. No internal editing is required because the whole musical sequence is required from the beginning. It is only necessary to have enough blank tape at the start of the reels so that the tape won't run off the reel before the completion of the backtiming process.

The applications of this technique are numerous, and any situation involving critical mixing of two tapes will be completed with very satisfactory results.

The recording and editing of tape is a fascinating thing, whether it be for vocation or avocation, but its degree of excellence is only attained through interest and practice.

40