TECHNICAL NOTE

*AN APPARATUS AND PROCEDURE FOR
THE PERCEPTUAL STUDY OF SPEECH PAUSES¹*

Speech pause is usually defined in the literature in terms of the silent interval existing between two segmental phonemes (Goldman-Eisler, 1961, and Boomer and Dittman, 1962). Research that sufficiently clarifies the relation of unfilled intervals to their perception as pauses in speech, however, remains to be performed. In fact, most of the inferences regarding the nature and function of speech pauses have been based on subjective observation and judgment of these events as they occur spontaneously, rather than on rigorously controlled experimentation. Methods that use speakers who attempt to control the duration of pauses lack adequate precision. Although tape splicing techniques, such as employed by Boomer and Dittman (1962), do provide a measure of control over silent interval duration, these techniques are ordinarily limited in research applicability and, in addition, are time consuming. The study of speech pauses, therefore, requires the development of a procedure that allows precise control over the duration of intervals of silence as they relate to the perception of speech pauses and, yet, one that is rather straightforward in nature for research purposes.

The present paper describes an apparatus and procedure that provides such capabilities without the disadvantages of previous techniques. In brief, the approach allows the experimenter to maintain control of silent interval duration through recording techniques in which the part of a sentence preceding a pause of interest is recorded on Track A and the remainder of the sentence (that part following the pause) is recorded on Track B. Pause duration is controlled by the subject or the experimenter who may vary the time of onset of the Track B signal with respect to that of the Track A signal by means of a special purpose apparatus (Pause Adjustment Device).

To implement the procedure for providing variable control of pause duration, it is necessary to construct first such a two-track stimulus tape. To this end, sentences are chosen in which a single word boundary is selected as the interval whose duration is to be varied. For instance, if one is interested in pause perception in a stop-phoneme environment, the sentence "He lost contact with reality" may be used and the word bound-

dary between "lost" and "contact" selected as the silent interval to be varied. The sentence is recorded single track at the highest tape speed for the recorder being used.

To obtain the appropriate signal on Tracks A and B respectively, the tape is played back at half speed into a two-channel dubbing recorder operating at the same speed as the playback recorder. Since the resulting signal is slowed by half, it becomes possible to locate the silent interval between "lost" and "contact" through auditory (earphone) monitoring, and manually switch the input from Channel A to B on the dubbing recorder input (see Fig. 1). The sentence is thus dubbed so that Track A contains the portion of the sentence up to and including the word "lost" and Track B contains the remainder of the sentence. The two track tape is monitored on a dual channel storage oscilloscope to determine that the switching actually did take place in the silent interval between the words "lost" and "contact".

If the temporal relations between Tracks A and B signals are left in their originally recorded state, the minimum value that could be achieved in adjusting the duration of a silent interval is that value corresponding to the original duration of that interval plus the duration represented by the physical separation of the two playback heads. To give the subject the capability of adjusting the pause duration between "lost" and "contact" to a value less than that present in the original recording, or to go to absolute zero duration or word overlap if he so desires, one must redub the sentence so that the Track B signal is physically advanced by about 1 sec with respect to the original temporal relations between the two signals. Thus, the listener is uninfluenced by the pause duration that exists in the originally recorded utterance. Figure 2 depicts this re-recording sequence beginning with the original single track recording and ending with the overlapped signals.

To accomplish this, common starting points are marked on both the two-track playback tape and on a fresh dubbing tape with India ink. With the two-track playback tape on one recorder and the dubbing tape on another, both recorders are started simultaneously and the original Track A is copied onto Track A of the dubbing tape. Both tapes are then rewound to the starting points and the playback tape is then advanced another 15 to 20 in. (38 to 51 cm). Again both recorders are started simultaneously and the Track B signal recorded. Since the playback tape had been advanced for this Track B dubbing, this signal is physically advanced by about 1 sec (with the playback recorder oper-

¹This research was supported by NSF Grant GS-2337 and by Grant NB-06459 of the National Institute of Neurological Diseases and Stroke and was part of a doctoral dissertation completed under the direction of Dr. Paul Jensen. Reprints may be obtained from Kenneth F. Ruder, Bureau of Child Research, University of Kansas, Lawrence, Kansas 66044.

Fig. 1. Block diagram of instrumentation used in the generation of a two-track tape. The manual switch is used to switch the input to the dubbing recorder between Channels A and B. The unused track is shorted by this switch to minimize the open circuit noise.

ating at a real time speed of 38 cm/sec) with respect to the original temporal relations between the Track A and B signals.

The recorded sentence in which the Track B signal has been temporally advanced is then made into a tape loop used to record approximately 150 repetitions of the sentence on a single 366-m reel with a 5-sec interval between repetitions. Since the stimulus sentence thus repeats itself every 5 sec, the listener can be presented with as many repetitions of the sentence as needed to make the pause duration adjustment required in the experimental task.

To provide for variable control of the onset of the Track B signal with respect to the cessation of the Track A signal utilizing the two track stimulus tape, a Pause Adjustment Device (PAD) was constructed, which could be used in conjunction with any two-channel tape recorder in which Tracks A and B could be recorded

Fig. 2. Diagram of recording sequence for generating the stimulus tapes. The signal is initially recorded single track (original recording). This recording is then dubbed so that Track A contains the portion of the signal preceding the pause of interest and Track B contains the remainder of the signal (first two-track dubbing). The final stage of the recording sequence consists of advancing the Track B signal in time with regard to the Track A signal (overlapped signals). Direction of tape movement across the heads is left to right.

Fig. 3. Apparatus to provide variable pause adjustment. The screen shields the instrumentation from the subject's view.

and/or played back separately. As depicted in Fig. 3, two stereo playback head stacks were mounted in a central location near the top edge of a standard aluminum rack panel approximately 20 in. (51 cm) in length. The two head stacks (labelled A and B in Fig. 3) were separated by a distance of about 2 in. (5 cm) in order to allow room for placement of tape guides between them. Track A of the "A" playback head and Track B of the "B" playback head were connected to the tape recorder channel A and B reproduce amplifiers respectively (see Fig. 4) using shielded cable.

A vertical slot 46 cm long and 6 mm wide was cut in the aluminum plate between the two head stacks. A movable tape guide attached to a worm gear assembly (see Fig. 3) protruded through this slot. The worm gear assembly was mounted on the rear of the aluminum plate and was operated by a hand crank (labelled Duration Control Knob in Fig. 3). This worm gear assembly can also (and has been in a later modification of this apparatus) be motor driven by replacing the hand crank with a small dc or reversible ac motor. A number of commercially available motors with variable speed and direction control serve this purpose without modification.

A tape guide was also permanently mounted on each side of each head stack in order to guide the tape across

Fig. 4. Schematic representation of playback head stack coupling to the reproduce amplifiers of the tape recorder.

the heads. Roller bearing tape guides were used in preference to the pin-type guides in order to minimize tape drag.

The two-track stimulus tape, previously described, is threaded across the Channel A playback head, around the movable tape guide, and back across the Channel B playback head to the take-up reel. Since the movable tape guide is attached to the worm gear, its position, relative to the horizontal plane of the playback heads, may be varied up and down the worm gear, thereby increasing or decreasing the effective distance between the two playback heads. Pause (silent interval) duration between these two parts of the sentence may be increased, then, by moving the tape guide away from the playback heads or decreased by moving the tape guide toward the playback heads. During playback, the signals from the two channels are combined so that both Track A and B signals are received in both ears. Pause duration (the time from the end of the Channel A signal to the beginning of the Channel B signal) may be measured by viewing the two channel signal on a storage oscilloscope or by using one of the pause duration analyzers such as described by Goldman-Eisler and Mendoza (1965).

DISCUSSION

Using this apparatus and procedure, it has been possible to compare durations of fluent and hesitation pauses within identical syntactic and phonetic environments (Ruder, 1969) by asking subjects to adjust the duration of the pause between "lost" and "contact" so that (1) the pause duration is considered to be optimal for a fluent, natural, presentation of the sentence and (2) the pause is just perceived as a hesitation. Pause detection thresholds in speech and pause duration as a function of syntactic complexity have also been investigated utilizing this technique (Ruder, 1969). This research thus exemplifies the usefulness of this apparatus and procedure for the systematic and controlled investigation of silent intervals as they affect the perception of pauses in speech.

The Pause Adjustment Device has proven to be quite reliable and easy to maintain. In approximately 200 hr of use, it has required no repair or maintenance other than an occasional oiling of the worm gear.

Although the Pause Adjustment Device (PAD) was designed primarily to study pauses in speech, it is apparent that it has wider applications. It can, in fact, be used in any situation calling for a delay between two signals. Commercial delay lines can give very short delays (in the nanosecond range) but are not as versatile as PAD, which is capable of producing delays of milli-

seconds up to 8 sec (with 46-cm worm gear and a recorder speed of 9.5 cm/sec) with no appreciable distortion of the signal.

PAD can also be modified for use as a delayed feedback recorder by replacing the "A" playback head stack with a record head stack. The minimum delay between record and playback with this device is limited, however, to the head stack separation necessary for the tape guide placement. Commercial delayed feedback machines can thus produce shorter delays between record and playback but are not capable of producing the longer delays of which PAD is capable.

A further advantage of PAD over commercial delay systems is that the delay is accomplished mechanically and can be used with any existing two-channel tape recorder in which channels A and B can be recorded and/or played back separately. The tape recorder itself needs no modification as such. The Pause Adjustment Device is simply mounted in the equipment rack below the tape deck and the two playback heads wired to their respective reproduce amplifiers. The tape recorder can thus be used as a conventional recorder by threading the tape across the tape recorder head stacks or as a delay recorder by threading the tape across the playback head stacks of PAD. This versatility, along with its simplicity of operation and ease of maintenance, further exemplifies the usefulness and practical nature of the Pause Adjustment Device for a variety of signal delay applications and particularly for the study of speech pauses.

REFERENCES

- Boomer, D. and Dittman, A. Hesitation pauses and juncture pauses in speech. *Language and Speech*, 1962, 5, 215-220.
- Goldman-Eisler, F. The distribution of pause durations in speech. *Language and Speech*, 1961, 5, 232-237.
- Goldman-Eisler, F. and Mendoza, R. Automatic pause-time recording, counting, and totalising equipment. 1965. Distributed by: The Clearinghouse for Federal and Scientific Information, U.S. Department of Commerce.
- Ruder, K. F. *Fluent and hesitation pauses as a function of syntactic complexity*. Doctoral dissertation, University of Florida, 1969.

K. F. RUDER
University of Kansas
 P. J. JENSEN
 J. F. BRANDT
University of Florida