Dept. for Speech, Music and Hearing

Quarterly Progress and Status Report

RASSLAN- a 6-channel closed loop sectioning device

Liljencrants, J.

journal: STL-QPSR
volume: 1
number: 2
year: 1960
pages: 001-003

KTH Computer Science and Communication

I. SPEECH ANALYSIS

A. SPECTRUM SAMPLING INSTRUMENTATION

1. The 51-channel analyzer

Detailed design plans for the 51-channel analyzer were given in the previous progress report. The design of the individual channels of the filterbank is almost completed. At present we are investigating temperature stability and other drift sources and a scheme for tuning the coils and condensers of the filters. The constructional phase will be started in about 3-4 months' time.

J. Liljencrants, U. Rengman

2. RASSLAN - a 6-channel closed loop sectioning device

A provisional instrumentation for spectrum sampling has been constructed, to a large extent utilizing earlier existing units. It has a comparatively low data handling capacity but can do a certain amount of routine work till the 51-channel spectrometer is completed.

The general design and function of the analyzer are illustrated in Fig. I-1.

The speech sample to be analyzed is stored on a tape loop with a revolution time of the order of 2 sec. On the loop is also a triggering signal, at present a 10-kc/s tone burst with a duration of 10 msec. This signal actuates a manually adjustable delay unit, the output of which governs the sampling point. The speech signal is processed in a heterodyne system analogous to that of the proposed 51-channel analyzer. The band-pass analysis is carried out in a bank of six filters with center frequencies from 1 kc/s and upwards. The output of the filters are individually rectified and smoothed. At the arrival of the delayed triggering pulse these signals are simultaneously sampled and stored in memory capacitors. The voltages of these are successively transferred to the recording system with the aid of a

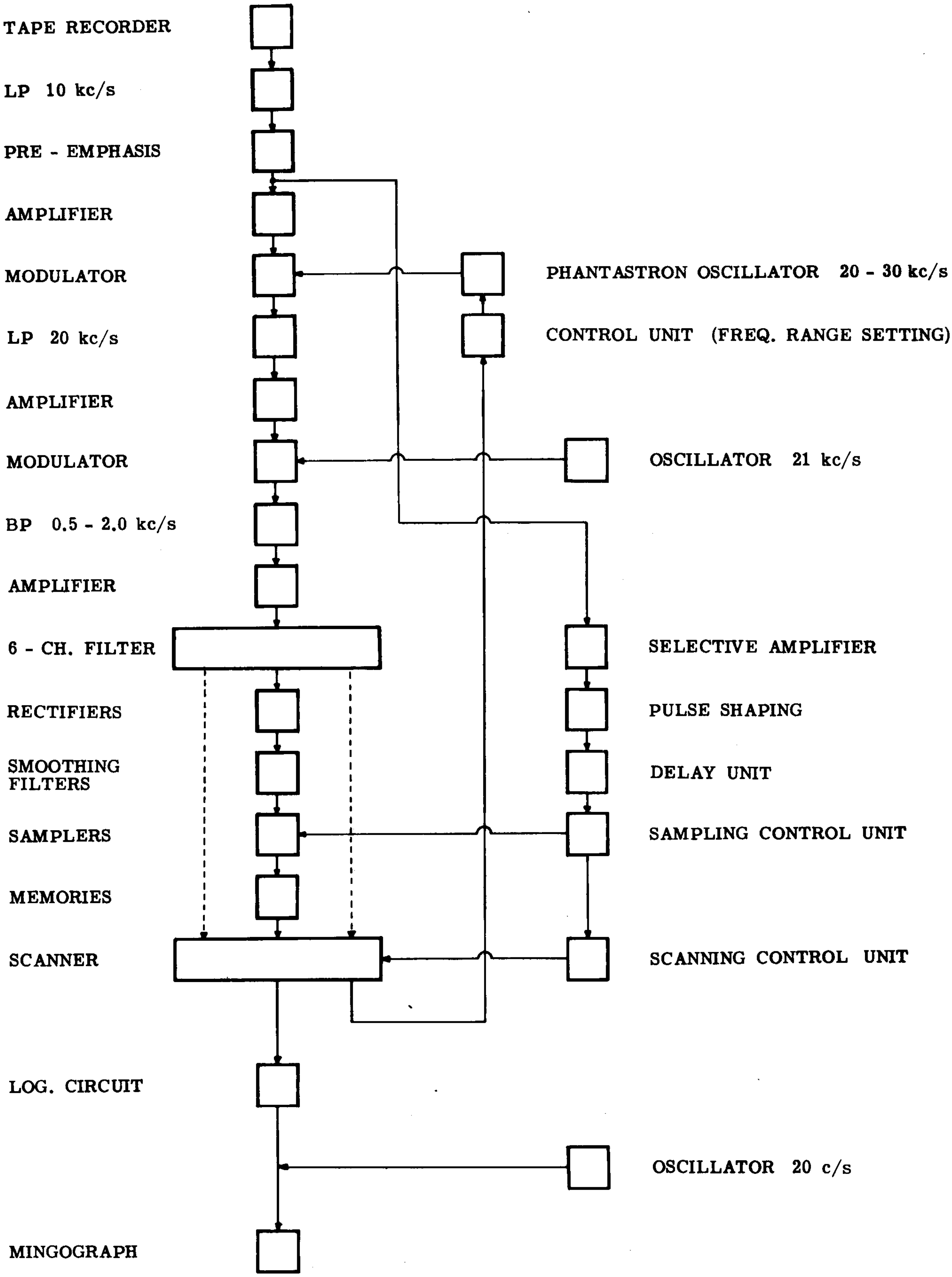

Fig. I-1 Block diagram of the six-channel closed loop spectral sectioning instrumentation.

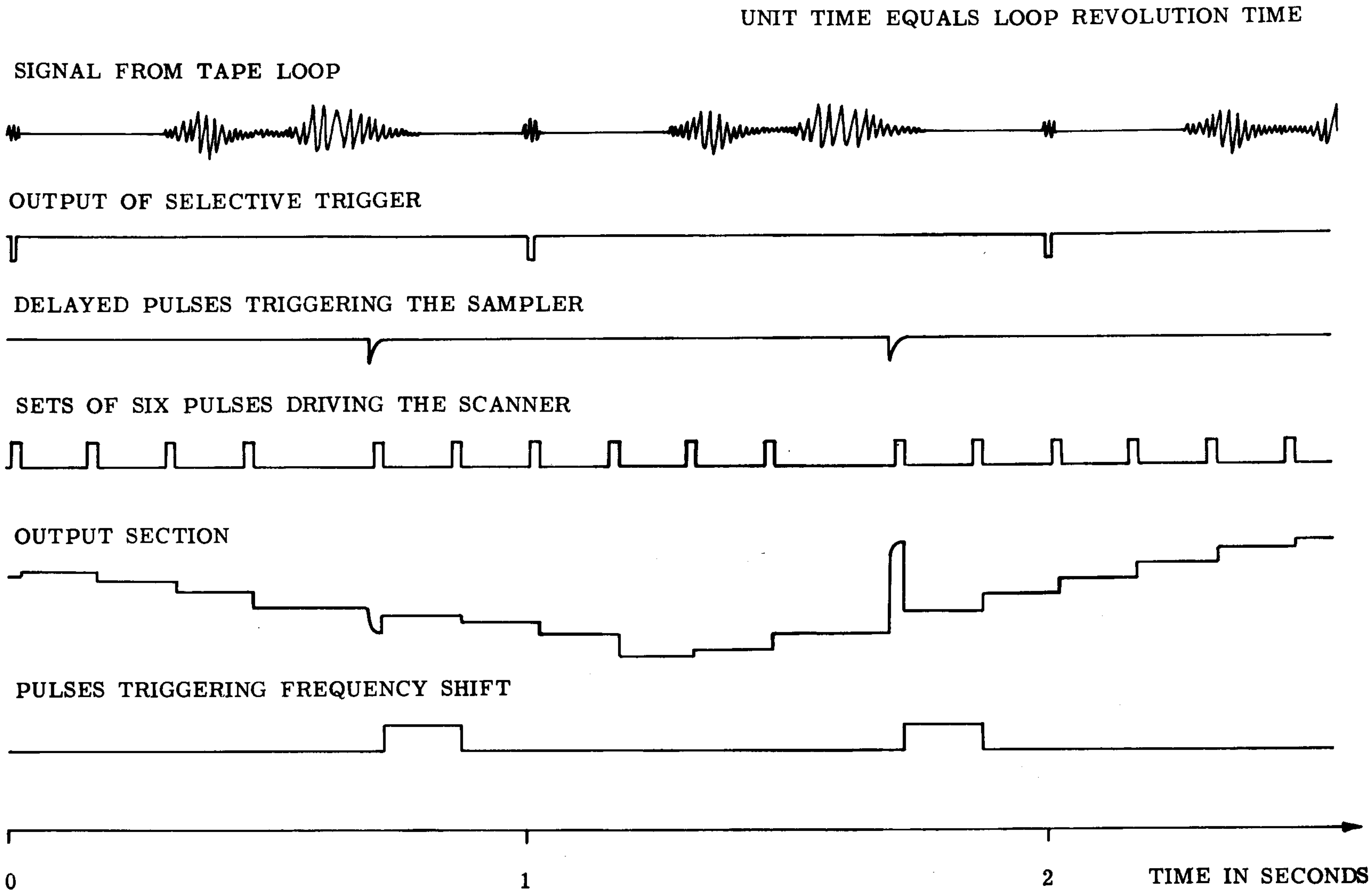

Fig. I-2 Sampling diagram for the six-channel spectrograph.

rotating telephone switch with a revolution time equal to that of the tape loop. When the synchronous sampling is carried out a 30-position telephone switch is put into action. This switch moves on one step selecting a carrier control voltage appropriate for the next cycle of analysis. For each revolution the frequency range of analysis is thus shifted by a constant amount. A complete spectrum section is obtained after 30 revolutions of the tape loop, a procedure that will take about 60 sec. The output of the scanner is converted to a logarithmic measure with the aid of a copper monoxide diode with forward current. The recording is finally carried out on a Mingograph recorder operating at a low paper speed (5 to 20 mm/s). In order to give a more clear picture the recorded signal is superimposed on a constant, low-frequency sine wave. The section consists of 30 adjacent sets of 6 columns representing the spectral energy in 180 bands with equal spacing and width.

The six analyzing filters have adjustable center frequencies and bandwidths corresponding to spacing of 15, 25, 50, and 100 c/s and bandwidths of 31, 62, 125, and 250 c/s, respectively. With the four mentioned spacings a total frequency range of 2.7, 4.5, 9.0, and 18.0 kc/s, respectively, will be covered. The heterodyne system, however, can manage frequencies below 10 kc/s only. Higher frequency ranges may be reached by means of reducing the playback speed of the sounds on the loop. The absolute position of the frequency range of the analysis may be chosen arbitrarily.

The smoothing filters have integration times variable in octave steps from 10 to 160 msec. All filters have the STL standard minimum overshoot characteristics.

As illustrated in Fig. I-2 there will occur an extra spike between every sixth column in the output section owing to the recharging of the memory capacitors. These spikes are helpful for determining the frequency scale. The dynamic range is somewhat wider than 40 dB.

In Fig. I-3 and I-4 are shown some typical analysis records from RASSLAN. It should be noted that these sections are taken from signals with high-frequency pre-emphasis (pole at $-2\pi \cdot 5000$ r/s, zero

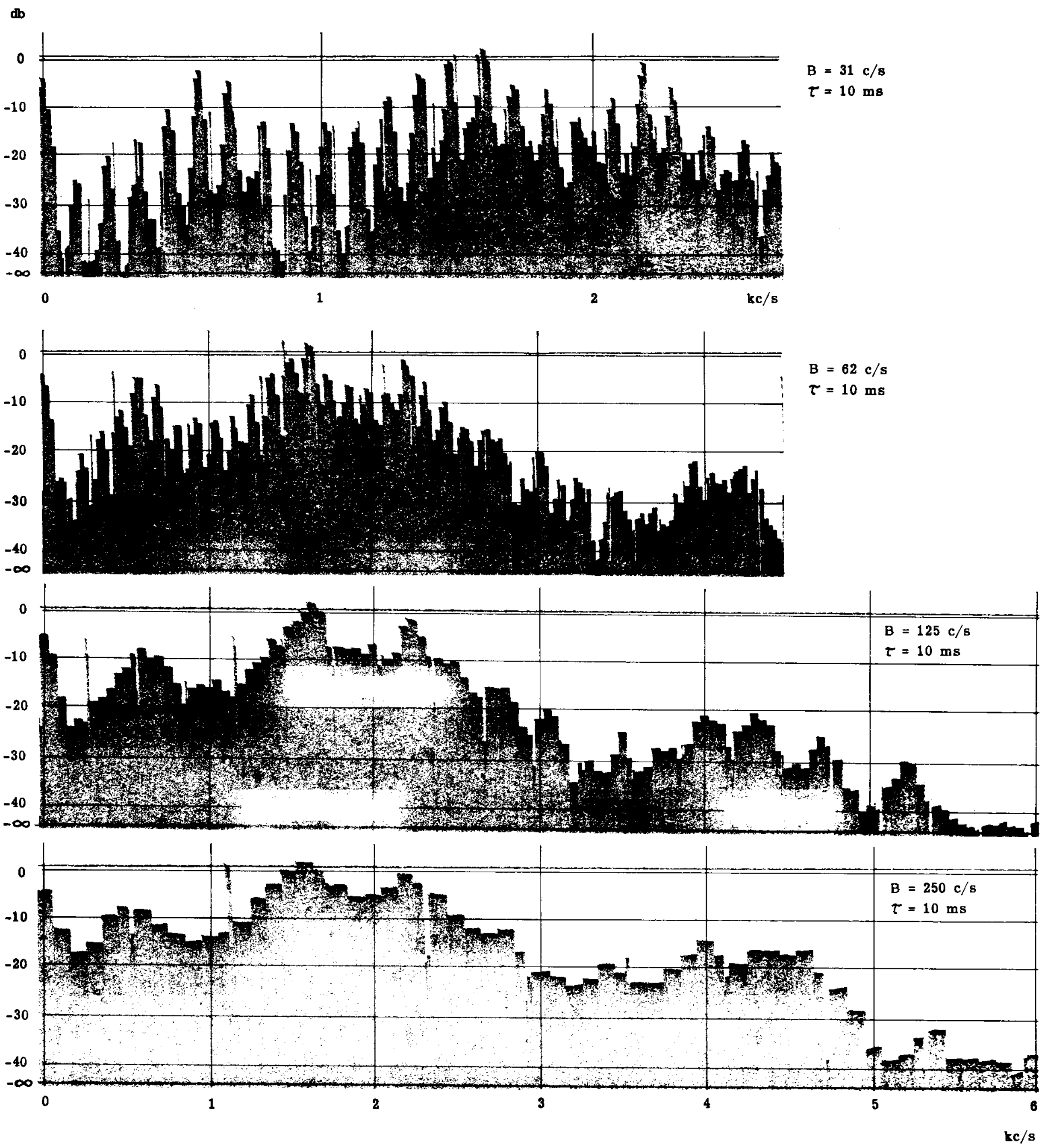

Fig. 1—3 Examples of spectral sections of a vowel [æ] produced with the six-channel spectrograph.

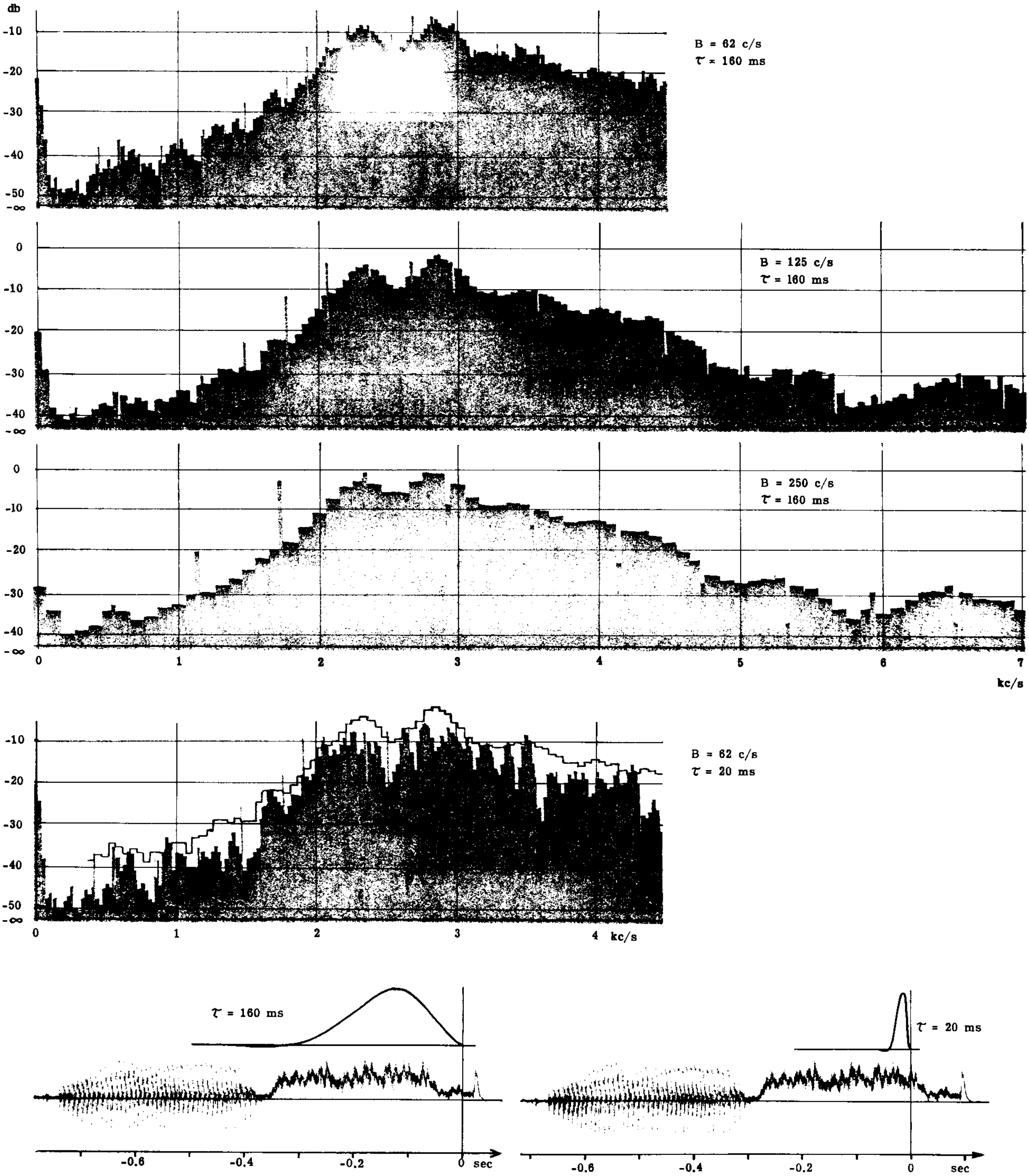

Fig. 1—4 Examples of spectral sections of a fricative [ʃ] produced with the six-channel analyzer.

at $-2\pi \cdot 200$ r/s).

The narrow-band analysis of the vowel (B=31 c/s and 62 c/s) displays an accurate harmonic structure the envelope of which contains more details than the wide-band spectrum (B=250 c/s). The analysis of the fricative, Fig. I-4, illustrates the influence of the quotient between analysis and smoothing filter bandwidths upon the random deviations from the mean envelope, see section A-3. Below the spectral section there is included a duplex oscillogram of the speech wave time function and a memory function for the smoothing filter, the time base of which is located at the sampling point.

J. Liljencrants

3. Analysis of random signals

The spectral analysis of random noise signals involves a band limitation of a band-pass filter of width B followed by a rectification and smoothing through a low-pass filter having the integration time τ.* The DC voltage thus obtained (V) has a random ripple superimposed on it, the RMS value of which may be called σ_V.

It is known [(1)] that

$$\frac{\sigma_V}{V} = k \cdot \frac{1}{\sqrt{B\tau}} \qquad \text{(Eq. I-1)}$$

See for instance ref. (1) and the previous progress report. If the signal is plotted on a logarithmic scale and provided $\sigma_V : V$ is reasonably small this formula can be rewritten:

$$\sigma_y = k \cdot 20 \cdot \log_{10} e \cdot \frac{1}{\sqrt{B\tau}} \qquad \text{(Eq. I-2)}$$

which is the RMS value of random striations in dB. The conversion

* For definition of τ see ref. (2). In terms of equivalent bandwidth $\tau = 1/2F$, where F is close to the 6 dB cutoff frequency of the LP-filter.