XIII. SPEECH COMMUNICATION

Prof. M. Halle	Dr. T. T. Sandel	O. Fujimura*
Prof. K. N. Stevens	Dr. W. S-Y. Wang	H. Fujisaki
Prof. J. B. Dennis	Jane B. Arnold	M. H. L. Hecker
Dr. A. S. House	C. G. Bell	J. M. Heinz
Dr. G. Rosen	P. T. Brady	C. I. Malme

A. SPEECH ANALYSIS†

Work has continued on the "analysis-by-synthesis" programs (1, 2) for the TX-0 computer, with an emphasis on gaining greater precision in the matching of vowel and consonant spectra, greater ease in manipulating test spectra, and a better understanding of various error criteria. Collaterally with the automatic resolution of spectra leading to the development of speech-recognition systems, attention has been given to the problem of semiautomatic analysis for measurement and description of speech sounds.

To facilitate the study of speech spectra and the development of analysis strategies, a general program has been written for direct Flexowriter communication with various computer programs. This study includes various punch-out and read-in programs for storage or retrieval of blocks of speech data. Other investigations include programs for calculating simple functions that have either a pair of complex conjugate poles or zeros, or a real-axis pole or zero; for combining these elemental functions to form more complex spectra; for calculating the outputs of a filter bank of known characteristics when it is presented with a specified input spectrum; and several display programs for viewing speech data and calculated spectra.

These new and modified analysis techniques have been used to determine the positions of poles and zeros associated with the spectra of several vowel, fricative, and nasal sounds. Our preliminary results indicate that the method can be used to locate pole and zero positions with a precision that has not been possible in any analysis methods that have been used heretofore.

C. G. Bell, H. Fujisaki, J. M. Heinz, A. S. House, K. N. Stevens

References

1. C. G. Bell, J. M. Heinz, G. Rosen, and K. N. Stevens, Automatic resolution of speech spectra into elemental spectra, Quarterly Progress Report No. 54, Research Laboratory of Electronics, M.I.T., July 15, 1959, pp. 161-167.

2. C. G. Bell, F. Poza, and K. N. Stevens, Automatic resolution of speech spectra into elemental spectra, Paper A6, Proc. Seminar on Speech Compression and Processing (AFCRC-TR-59-198), December 1959.

---

*On Leave from the Research Institute of Communication Science, University of Electro-Communications, Tokyo, Japan.

†This research was supported in part by the U.S. Air Force (Air Force Cambridge Research Center, Air Research and Development Command) under Contract AF19(604)-6102; and in part by National Science Foundation.

B. DETECTABILITY OF SMALL IRREGULARITIES IN A HARMONIC LINE SPECTRUM*

In a previous report (1), the detectability of small changes in the energy content of a narrow frequency band in a white-noise spectrum was discussed. The filter circuit and psychoacoustic testing methods developed for the noise experiment have been used to determine the differential energy change at a single harmonic frequency in an acoustic line spectrum that would produce a detectable change in timbre.

The psychoacoustic test stimuli were obtained by passing the signal from a pulse generator through a lowpass filter (5-kc cutoff), and then through a filter with a simple resonance and antiresonance. The effect of the second filter was to increase or decrease the amplitude of a single harmonic frequency by a controlled amount. Three types of harmonic line spectra were used in a series of 5 tests. Spectrum 1 was a flat line spectrum produced by impulses having a repetition rate of 500 pps. This type of spectrum was used because its harmonic structure could be controlled accurately. Spectrum 2 was a line spectrum with a 500-cps fundamental, but with higher harmonics attenuated at a rate of 6 db per octave by means of an RC network. This spectrum simulated the output produced by certain musical instruments. Spectrum 3 was a flat line spectrum produced by impulses at a repetition rate of 150 pps. Three additional tests were given to measure the differential intensity change for a pure tone at 300, 1000, and 3000 cps. The stimulus conditions are summarized in Table XIII-1.

Table XIII-1. Summary of tests.

Test Number	Spectrum	Fundamental (cps)	Harmonic (Intensity Varied)	Frequency (cps)
1	1 (flat)	500	2nd	1000
2	1 (flat)	500	6th	3000
3	2 (6-db slope)	500	2nd	1000
4	2 (6-db slope)	500	6th	3000
5	3 (flat)	150	2nd	300
6	pure tone	300		300
7	pure tone	1000		1000
8	pure tone	3000		3000

The intensity changes used in all tests were in steps of 0.5, 1, 2, 4, 6, and 10 db, with

*This work was supported in part by the U.S. Navy (Office of Naval Research) under Contract Nonr-1841(42). Reproduction in whole or in part is permitted for any purpose of the United States Government.

the exception of the tests at 300 cps in which steps of 2, 4, 6, 10, 15, and 20 db were necessary.

The test stimuli were presented to the subjects by means of an Ampex tape recorder and Permoflux PDR-8 earphones. Each recorded test contained 100 stimulus groups in standard ABX presentation. A stimulus group consisted of three tone bursts, each of 500-msec duration with a sound-time fraction of 1/2. In the ABX test, subjects were asked whether the last stimulus sounded more like the first or more like the second; they were given 5 seconds in which to record their decision. Six subjects gave a total of 48 judgments on each intensity change. The reference sound-pressure level at the test frequency in each case was 70 db (re 0.0002 μbar) as determined by a coupler calibration of the earphones. These experimental results are shown in Table XIII-2. The intensity change detectable by an average listener 75 per cent of the time was obtained by averaging the results from each subject. The subjects' test results were also divided into two groups in order to observe the effect of musical training. The data listed under Group 1 in Table XIII-2 are the averaged results from three subjects with no previous musical training. The column under Group 2 contains the averaged results of three subjects, two of whom are amateur violinists and one is an amateur clarinetist. Data from the earlier experiment (1) with the use of white-noise stimuli are included for comparison.

Examination of the results shown in Table XIII-2 indicates that:

(a) Without exception, the musically trained subjects were more sensitive to changes in intensity.

(b) A detectable change in second-harmonic intensity was smaller with the sloped spectrum than with the flat spectrum.

(c) At the sixth-harmonic frequency, the barely discernible intensity change, or difference limen (DL), was about the same for both the flat and sloped spectra.

(d) The average difference limen for intensity was smallest at 1000 cps and largest at 300 cps.

(e) The results for the musically trained subjects on tests 1 and 3 (line spectrum) agree closely with those for test 7 (pure tone). The remaining line-spectrum and pure-tone tests do not agree as well for this group, the largest difference showing up at 300 cps.

(f) In general, the subjects without musical training had a smaller DL for intensity in the pure-tone tests than in the line-spectrum tests, the only exception being in test 7b.

(g) The two groups of subjects showed very similar DL's in the tests involving an increase in pure-tone intensity.

(h) The DL values obtained for the musically untrained subjects in the line-spectrum tests are close to those obtained previously with a white-noise spectrum of

Table XIII-2. Summary of experimental results. Data from an earlier experiment in which stimuli with a white-noise spectrum were used are shown for comparison.

Test Number	Test Frequency (cps)	Intensity Difference Limen			
		Group 1 (db)	Group 2 (db)	Average (db)	Standard Deviation (db)
1a (flat spectrum)	1000	5.8	3.3	4.6	±1.6
3a (sloped spectrum)	intensity increase	4.1	1.9	3.0	±1.4
7a (pure tone)		2.8	2.7	2.7	±0.3
noise spectrum				4.2	±0.5
1b (flat spectrum)	1000	6.3	2.5	4.4	±2.3
3b (sloped spectrum)	intensity decrease	4.2	1.9	3.0	±1.4
7b (pure tone)		5.3	3.2	4.3	±1.6
noise spectrum				6.6	±0.9
2a (flat spectrum)	3000	6.0	3.1	4.5	±2.1
4a (sloped spectrum)	intensity increase	6.0	3.4	4.7	±1.3
6a (pure tone)		2.7	2.2	2.5	±0.8
noise spectrum				5.2	±0.6
2b (flat spectrum)	3000	7.0	4.7	5.3	±2.0
4b (sloped spectrum)	intensity decrease	8.7	4.5	6.6	±2.5
6b (pure tone)		4.3	1.4	2.9	±1.6
noise spectrum				8.8	±1.6
5a (flat spectrum)	300	8.2	5.1	6.7	±2.0
8a (pure tone)	intensity increase	2.8	2.6	2.7	±0.5
noise spectrum				9.0	±1.4
5b (flat spectrum)	300	10.5	9.8	10.1	±5.7
8b (pure tone)	intensity decrease	5.3	2.4	3.8	±1.9
noise spectrum				14.3	±2.8

approximately the same bandwidth. (The values listed are the peak heights or notch depths of a simple irregularity in a white-noise spectrum that is detectable 75 per cent of the time by an average subject (1).)

(i) The DL values obtained for the musically trained subjects in the line-spectrum tests are, in general, less than one half as large as those obtained previously with the white-noise spectrum.

About twenty years ago, Loeb (2) reported the results of an experiment similar to the one just described. In his experiment, he used apparatus designed by Nahrgang (3), to achieve independent variation of the intensity of each of seven harmonics of fundamentals ranging from 129 cps to 1740 cps. Differential thresholds for changes in timbre were determined as a function of changes in the intensity of one or more harmonic frequencies. Loeb found that the threshold for timbre change was equal to the differential threshold for intensity of the modified harmonics. His result agrees quite well with the data of Table XIII-2 for the second harmonic at 1000 cps. A possible explanation for the larger differences at 300 and 3000 cps is that Loeb used well-trained subjects and a direct-comparison method of testing.

From the present data, the average DL for intensity of a pure tone was 2.6 db for an increase in intensity, and 3.7 db for a decrease in intensity. These values can be compared with a value of 2.4 db obtained by Dimmick and Olson (4) by the "same or different method" of psychoacoustic testing.

The ratio of the detectable energy change to the total spectrum energy is much smaller in the case of a harmonic line spectrum than it is in the case of a pure tone. The fact that the threshold for timbre change in a complex line spectrum is almost equal to the difference limen for intensity of the modified harmonic is remarkable. It indicates that, for the type of judgment required here, each partial tone is relatively free from the influence of its neighbors – at least, in a harmonic spectrum.

It is known that the amplitude of the vibratory response of the basilar membrane to a pure tone has a broad peak (5). The "sharpening" needed to give the ear its selective characteristics is generally assumed to be provided by the central nervous system. The results given here indicate that the auditory mechanism can indeed perform as a selective filter, and that its selectivity is improved by proper training.

C. I. Malme

References

1. C. I. Malme, Detectability of small irregularities in a broadband noise spectrum, Quarterly Progress Report No. 56, Research Laboratory of Electronics, M.I.T., Jan. 15, 1959, p. 139.

2. E. Loeb, Über die Subjektive Wirkung von Klangfarbenänderungen, Akust. Z. 6, 279-294 (1941).

(References continued on next page)

3. S. Nahrgang, Beobachtungen und Messungen an einem neuen Gerät zur Klangsynthese, Akust. Z. 3, 284-301 (1938).

4. F. L. Dimmick and R. M. Olson, Intensive difference limen in audition, J. Acoust. Soc. Am. 12, 517-525 (1940).

5. G. von Békésy, On the resonance curve and the decay period at various points on the cochlear partition, J. Acoust. Soc. Am. 21, 245-254 (1949).