XII. SPEECH COMMUNICATION[*]

Prof. K. N. Stevens
Prof. M. Halle
Prof. J. B. Dennis
Prof. J. M. Heinz

Dr. A. S. House
Jane B. Arnold
W. L. Henke

A. P. Paul
T. E. Taylor
S. S. Reisman
E. C. Whitman

A. ANALYSIS OF ERROR SCORES DERIVED FROM AN AUTOMATIC VOWEL REDUCTION SCHEME

In an evaluation of the performance of an automatic procedure for reducing vowel spectra to descriptions in terms of the time-variant vocal-tract resonances and source characteristics, the behavior of the squared-error score has been examined in some detail. This examination provides some insight into the spectrum-matching accuracy of the automatic program and, it is hoped, may lead also to better understanding of the operational meaning of an adequate match in terms of error score.

The automatic program under investigation has already been described,[1,2] and its accuracy in matching spectra has been discussed[3,4] in terms of differences in pole specifications between vowel samples analyzed by the automatic procedures and by manual procedures. The manual reduction methods are described in detail in a paper by Bell et al.[5] The speech materials are those reported by House,[6] consisting of eight vowels in 14 consonantal contexts spoken by three talkers.

Three contiguous spectral samples located medially in each of the vowels were analyzed manually and automatically, and for each such analysis there corresponds a squared-error score. The mean-error scores associated with each vowel and talker are displayed in Table XII-1 according to spectrum-matching method. Inspection of these scores indicates that the magnitude of the mean error associated with a given vowel, for example, is reasonably stable from talker to talker, and that the average error score of each talker is highly similar. There is, however, a statistically significant difference between the magnitude of the mean-error scores of the two methods, manual and automatic, with the latter providing lower scores. This tendency is quite consistent when scores for individual vowels and talkers are studied. For both the manual and automatic analyses, the vowel means display significant differences, with higher scores tending to be associated with the back, rounded vowels. This tendency may be aggravated when rounded vowels are in an environment provided by labial consonants. In general, however, consonantal context does not provide a significant source of error-score variation.

The mean-error scores per filter for the two methods of analysis are shown in Table XII-2. This presentation emphasizes the consistency with which the automatic

---

Table XII-1. Mean-error scores (in db^2) for spectrum matches to various vowels, arranged by methods and talkers. Scores are computed over filters 1-24 corresponding to a frequency range of 100-3050 cps. Each entry is based on three centrally located samples taken from vowels in 14 consonantal contexts.

Method	Talkers	Vowels								Mean
		ɩ	I	ɛ	æ	ɑ	ʌ	U	ʊ	
Automatic	1	45	34	29	33	30	37	66	55	41
	2	65	42	41	36	46	43	65	56	49
	3	40	35	31	28	34	38	59	65	41
	Mean	50	37	33	32	37	39	63	59	44
Manual	1	49	48	48	48	40	47	80	62	53
	2	49	54	54	46	42	48	53	55	50
	3	44	44	47	39	46	46	59	66	49
	Mean	47	48	49	45	43	47	64	61	51

Table XII-2. Mean error scores (in db^2) per filter for spectrum matches to various vowels, arranged by methods. Scores are computed over filters 1-24 corresponding to a frequency range of 100-3050 cps. Each entry is based on three centrally located samples taken from vowels in 14 consonantal contexts as spoken by three talkers.

Method	Vowels								Mean
	ɩ	I	ɛ	æ	ɑ	ʌ	U	ʊ	
Automatic	2.1	1.6	1.4	1.4	1.5	1.6	2.6	2.5	1.8
Manual	2.0	2.0	2.1	1.9	1.9	2.0	2.7	2.5	2.1

procedure provides lower scores. According to this table the fit of the vowel data is, on the average, accurate to better than 2 db in each of 24 frequency bands.

A. S. House, A. P. Paul

References

1. A. P. Paul, Strategy for the Automatic Matching of Speech Spectra, S. M. Thesis, Department of Electrical Engineering, M. I. T., August 1961.

2. A. P. Paul, Automatic matching of vowel spectra, Quarterly Progress Report No. 63, Research Laboratory of Electronics, M. I. T., October 15, 1961, p. 137.

3. A. P. Paul, An automatic spectrum matching program, Quarterly Progress Report No. 66, Research Laboratory of Electronics, M. I. T., July 15, 1962, pp. 275-278.

4. A. S. House, A. P. Paul, K. N. Stevens, and Jane B. Arnold, Acoustical description of syllabic nuclei: Data derived by automatic analysis procedures, Proc. Fourth International Congress on Acoustics, Copenhagen, August 21-28, 1962 (Paper G22).

5. C. G. Bell, H. Fujisaki, J. M. Heinz, K. N. Stevens, and A. S. House, Reduction of speech spectra by analysis-by-synthesis techniques, J. Acoust. Soc. Am. 33, 1725-1736 (1961).

6. A. S. House, On vowel duration in English, J. Acoust. Soc. Am. 33, 1174-1178 (1961).