SOME ANALYSES OF VOWELS BY SOCIAL GROUP IN THE SURVEY OF VANCOUVER ENGLISH*

John H. Esling

Department of Linguistics
University of Victoria

1. ACOUSTIC ANALYSIS IN DIALECT STUDY

There are three questions in the study of regional and social dialect that are addressed in this acoustic study of vowels. The first is the question of how individual vowels differ in their production from one social or regional variety to another. The second is whether there is a pattern that relates each social group's set of vowels to each other in a systematic way which can therefore be used to differentiate one group from another. The third is to identify the articulatory characteristics that correspond to each group's acoustic pattern, and to evaluate the acoustic parameters which provide the best indication of articulatory differences.

The data for this acoustic study have been drawn from the Survey of Vancouver English carried out by Gregg, *et al*, (1981) at the University of British Columbia. The subjects are 40 female and 40 male natives of Greater Vancouver. All subjects are anglophones who have grown up in Vancouver. They represent the youngest of the three age divisions in the survey, in the range between 16 and 35 years old. Female and male subjects are divided into four social groups of ten subjects each on the basis of social index scores determined in the original survey using the Blishen and McRoberts (1976) social indexing scale combined

----

* I would like to express my appreciation to Professor R.J. Gregg of the University of British Columbia for his encouragement and support of this project, to Dr. S.K. Wong for the development and modification of statistical processing algorithms, and to Jocelyn Clayards for her assiduous attention to measurements. Initial development of acoustic procedures was made possible by grants from the President's Committee on Research and Travel of the University of Victoria. The current project description is a preliminary report of an investigation supported by Social Sciences and Humanities Research Council of Canada grant 410-85-0481, entitled "The Setting Component of Accent in Vancouver".
This is a slightly modified and expanded version of a paper presented at the Annual Meeting of the Linguistic Society of America and American Dialect Society, Seattle, December, 1985.

with other social indicators. Group 1 represents low social index scores, and group 4 represents the highest social index scores.

2. RESEARCH QUESTIONS

1. Are the individual vowels of Vancouver English significantly differentiated across the social groups of the survey?
2. Are long-term average spectra (LTAS) significantly differentiated across the social groups of the Survey of Vancouver English?
3. Do the long-term average spectra reflect patterns of shift from group to group that are represented in the distribution of vowels?
4. Are the first two vowel formants (F1 and F2) comparable acoustic measures to the long-term average spectral representations of the speech of speakers in the survey?
5. Are relationships of vowel clusters across social groups the same for female and male subjects in the survey?
6. What long-term articulatory setting characteristics can be said to describe each of the social groups in the Vancouver survey as suggested by either vowel clustering procedures or LTAS calculations?
7. How do current data and results from the survey help to explain earlier relationships obtained for Vancouver and other English-speaking subjects using vowel clustering and LTAS techniques?
8. Is formant (F1,F2) analysis the most accurate acoustic means of analyzing the relationships between vowel systems of the social groups of a sociolinguistic survey?

3. METHODS OF SPEECH ANALYSIS

To compare vowel clusters across the groups, vocalic nuclei are computed for two tokens of each of ten vowel phonemes for each speaker in the survey, from identical text in reading style. Initial measurements of vowel formants are made using the ILS package on the PDP-11 minicomputer. Subsequent measurements are made using the Micro Speech Lab package developed in the Centre for Speech Technology Research at the University of Victoria on the IBM-PC microcomputer. In the ILS procedure, formant peaks, bandwidths and amplitudes are calculated and printed for each vowel token. The mean first and second formant frequencies are then calculated from the information on the print-out and filed by group for statistical processing. Two speakers from each group were excluded because of poor quality or local interference during recording, leaving eight subjects in each cell with acceptable measurable vocalic nuclei, 32 females and 32 males. The same 64 speakers are therefore included in the LTAS calculations. At the present stage of research, all female and male vowel formant determinations have been made, and LTAS calculations for the female subjects have been completed.

For long-time spectral analysis, a sample of 45 seconds of continuous speech for each speaker, from the same reading text as used in vowel measurements, is digitized and stored on disk for LTAS processing using a time-series data-capturing program on the PDP-11 minicomputer. One long-term average spectrum is computed for each speaker's voice, using a program on the main-frame computer which accepts only voiced frames of speech while excluding voiceless and low-energy frames. The power spectra of non-overlapping 20-msec windows at 50 Hz resolution and with a pre-emphasis factor of 1 are integrated to obtain the final long-term spectrum. Comparisons between the four groups operate on each of the eight spectra that comprise each group, while visual representations illustrate the average spectrum of each group.

4. STATISTICAL ANALYSIS

To compute the distance between vowel clusters of the contrasting groups, principal components analysis and canonical discriminant analysis are applied for all vowels across groups 1 through 4, and the Mahalanobis distance is computed between each group. This gives the probability with which collections of vowels, both as complete vocalic inventories and as individual vowel phonemes, are differentiated from one another. In addition, a generalized squared distance measure is used to compute, from the entire pool of values, the percentages of single tokens of F1,F2 coordinates that are located nearest the centroids of each of the four groups. Then, using an extension of this same technique, the vocalic inventories of male groups 1 through 4 are compared with previously analyzed model vowel data from reading texts performed by the author to represent contrasting parameters of articulatory setting. Here, test values are assigned to known reference groups in a procedure which yields the percentage of vowels in each group which associate most closely with each of the models (see Esling and Dickson 1985).

These statistics operate on unnormalized first and second formant data, with comparisons performed separately for female and male groups; and produce comparisons across four-group samples of some 600 female and 600 male vowels, and across the four groups for each vowel value one at a time. The same procedure is used to compute the distance between mean LTAS curves of the four female groups. When male LTAS calculations have been completed, they will be compared with LTAS values of the models of articulatory settings.

5. RESULTS FOR FEMALE GROUPS

Female vowel cluster data are significantly differentiated across the four social groups of the survey. All groups show separation at the $p<0.001$ level of significance, except groups 1 and 3 which are differentiated at the $p<0.05$ level, for comparisons of all vowels. Furthermore, a majority of individual vowel phonemes compared one at a time demonstrate similar differentiation across the four female groups. The most coherent and best differentiated groups are groups 2 and 4, illustrated in figures 1 and 2. Linguistic contexts are identical; only speakers vary, according to group affiliation. Figure 3 illustrates this acoustic separation visually on the articulatorily oriented formant chart, with the means of female and male groups. The groups occupy separate corners of the vowel space: group 1, high F1 and low F2; group 2, low F1 and low F2; group 3, low F1 and high F2; group 4, high F1 and high F2.

Female LTAS are not significantly differentiated across the four social divisions of the survey. Group 1 is differentiated from group 2 ($p<0.01$), and group 2 is differentiated significantly from group 4 ($p<0.01$), but other relationships show no significant separation. Figure 4 illustrates group means of the first two peaks of the female LTAS waveforms, and the lack of differentiation in the value of peak 2 that results. Since each speaker is represented by a single LTAS waveform, each cell consists of only eight tokens; whereas in comparisons of vocalic data, each cell consists of approximately 150 observations, with predictably greater reliability.

Female LTAS data do not corroborate vowel cluster distributions, except that groups 2 and 4 are similarly separated by both measures. Because of this difference in results, however, LTAS data appear to reflect other spectral information than what is contained in F1 and F2 of vowel nuclei. Clearly there will be differences, due to the inclusion in LTAS of voiced obstruents, but the direct relationship observed between vowel cluster and LTAS patterns in the speech sample of an urban black dialect of Houston, Texas, English (Esling and Dickson 1985: 166) is not demonstrated here.

6. RESULTS FOR MALE GROUPS

Male vowel cluster values follow the pattern of female values, except that group 1 is not significantly differentiated from group 3 (figure 3). As before, group 2 is the tightest group, and furthest separated from all other groups, particularly from group 4.

Earlier LTAS and vowel comparison experimental data associating extreme tongue-backed settings with working-class

FIGURE 1.

VANCOUVER VOWELS, FEMALE (1–4)

GROUP=2

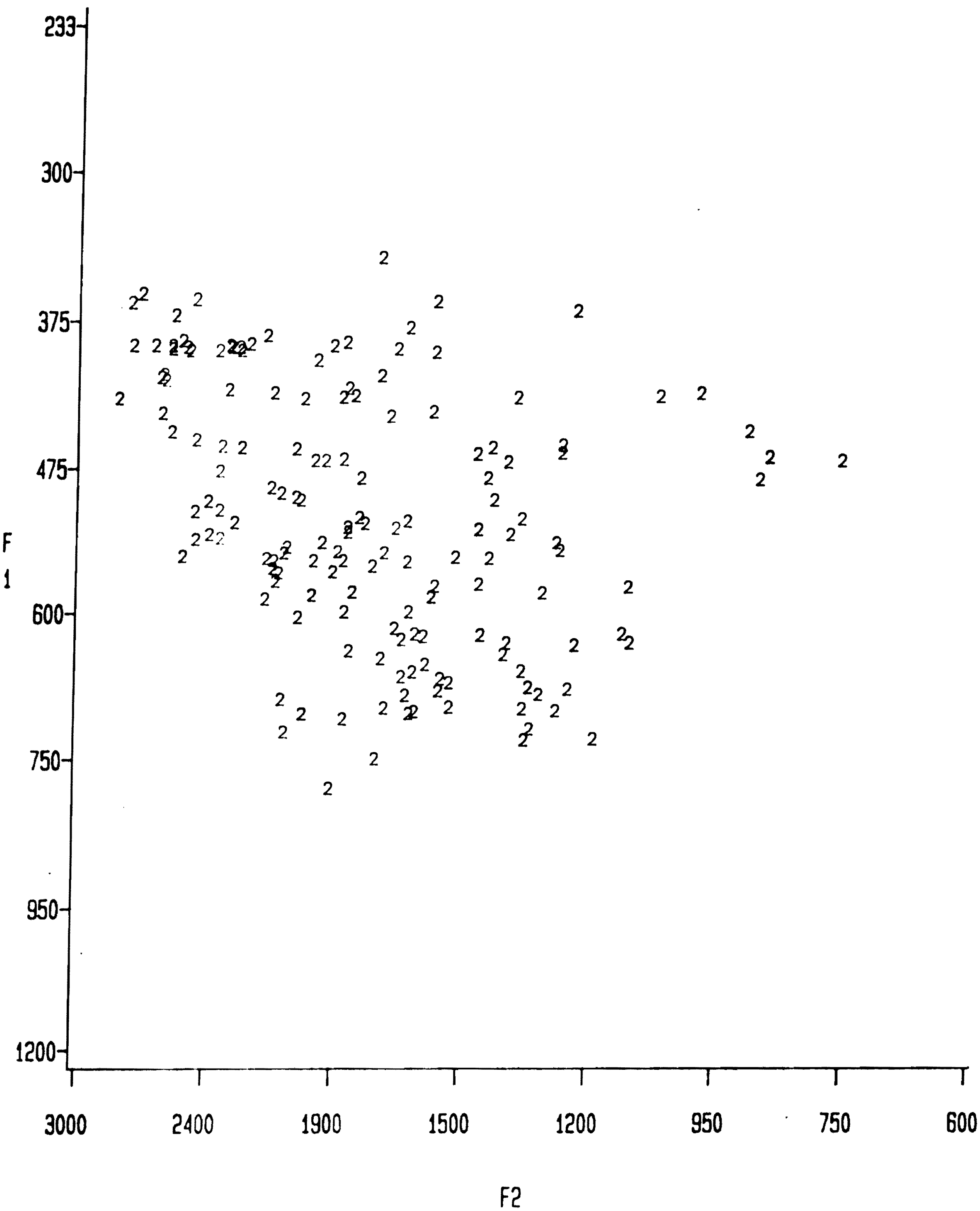


FIGURE 2.

VANCOUVER VOWELS, FEMALE (1–4)

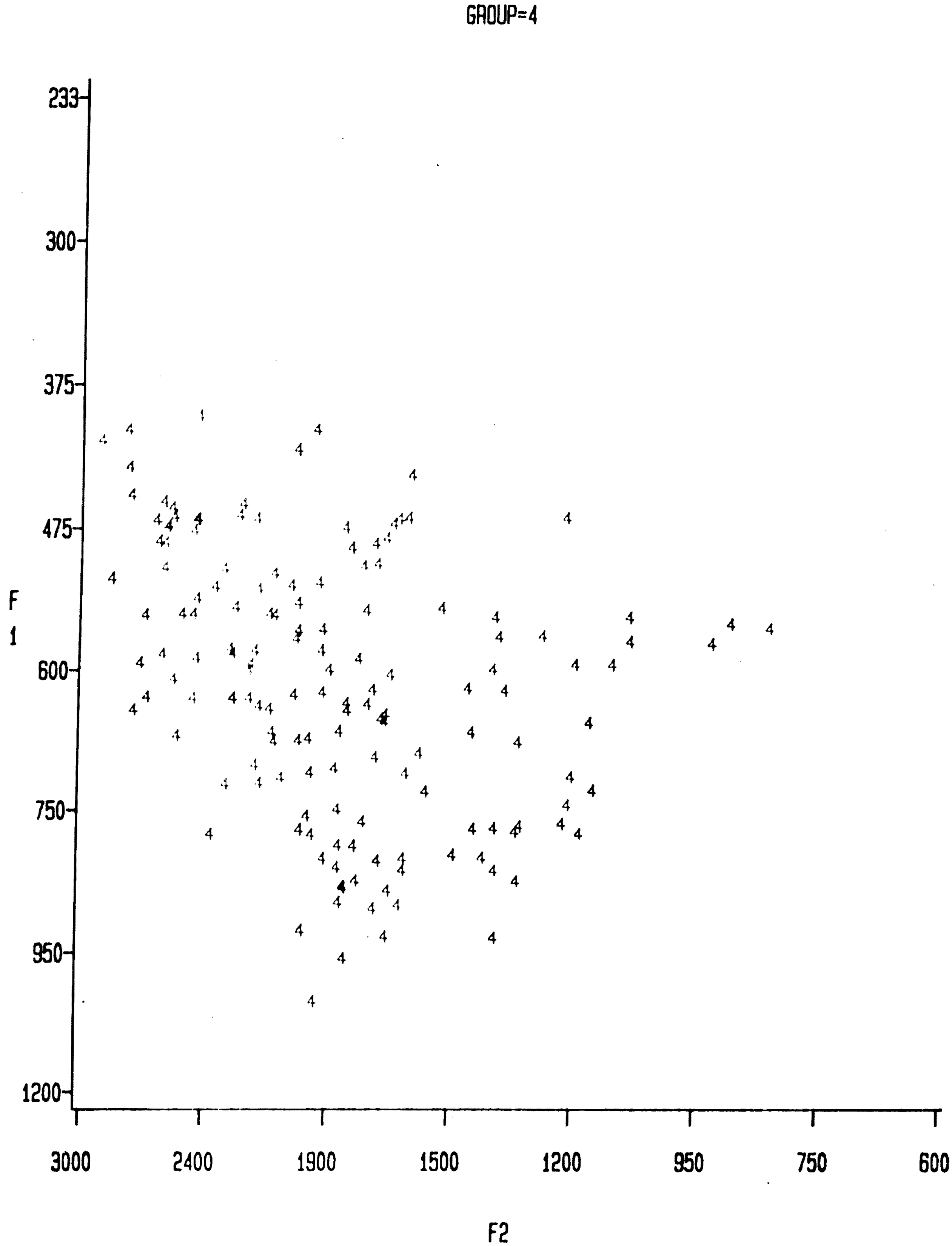


FIGURE 3.

VANCOUVER VOWELS (1–4)

FEMALE AND MALE VOCALIC MEANS

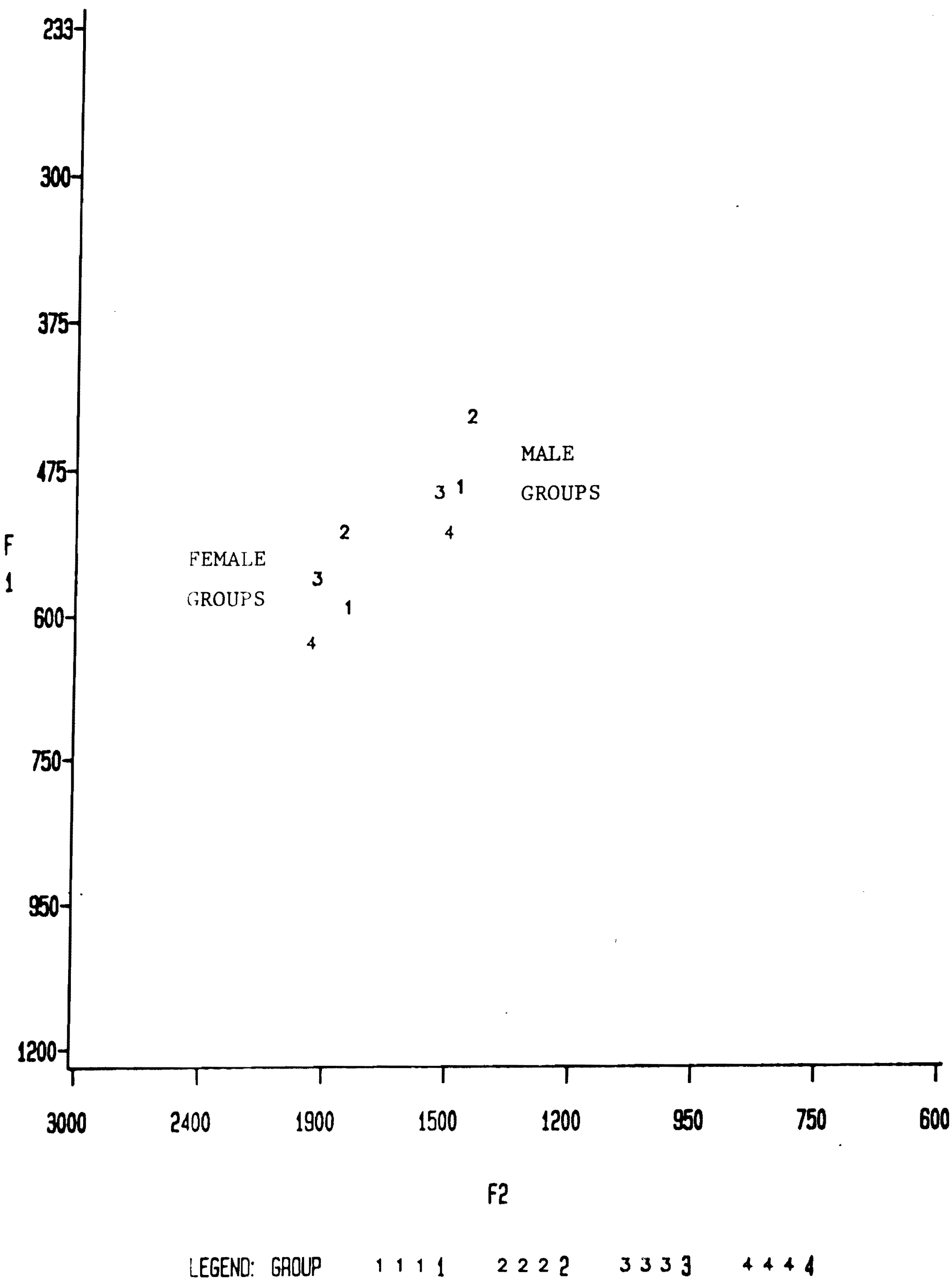
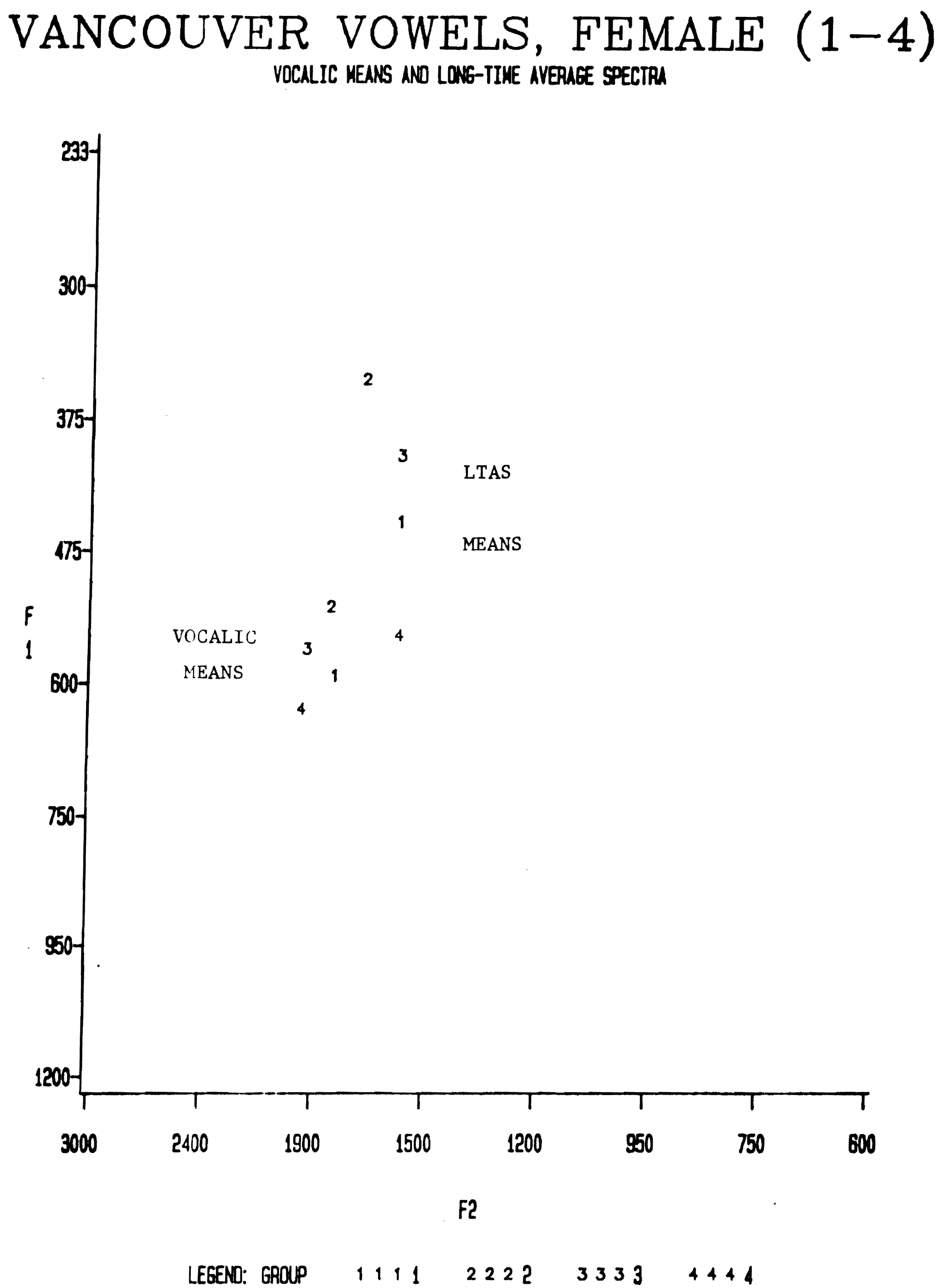


FIGURE 4.


VANCOUVER VOWELS, FEMALE (1–4)
VOCALIC MEANS AND LONG-TIME AVERAGE SPECTRA
233
300
375
475
600
750
950
1200
F1
LTAS
MEANS
VOCALIC
MEANS
3000
2400
1900
1500
1200
950
750
600
F2
LEGEND: GROUP 1 1 1 1 2 2 2 2 3 3 3 3 4 4 4 4

Vancouver speech can be demonstrated in the vowel cluster values for male groups 1 and 2, as distinct from values for groups 3 and 4 which imply articulatory fronting. It was impossible to differentiate earlier Vancouver data, which had been gathered from males in group 2 and group 1, from the LTAS or vowel cluster values of the Houston sample (Esling and Dickson 1985: 159-160). The Houston speakers are 14 community college students, all black and natives of the city, recorded in reading style in 1984. Further testing demonstrates that the Houston sample vowel data are indeed related to Vancouver groups 1 and 2 ($p>0.82$; 68% by generalized squared distance), while combined groups 1 and 2 remain distinct acoustically from combined Vancouver groups 3 and 4 ($p<0.0001$). In contrast with preliminary tests, the Houston data are significantly differentiated from Vancouver groups 3 and 4 in this extended comparison ($p<0.003$; 32% by squared distance).

7. ARTICULATORY INTERPRETATION

Articulatory interpretations of the acoustically differentiated female social groups are based initially on the parallel acoustic separation of the F1,F2 values of four supralaryngeal models evaluated in previous research (Esling and Dickson 1985: 163-166). The orientation of group 1 vowels resembles the low F2 of laryngo-pharyngalized tongue retraction. The low F1 of group 2 vowels suggests retraction with tongue raising as in velarization. High F2 in group 3 suggests fronting and raising as for palatalization. High F1 and F2 values in group 4 resemble the pattern of nasal voice.

To quantify these associations, sample data are compared with each phonetically performed model. Generalized squared distance assignments (of unknown values into known groups) relate male vowel cluster data and model data as follows (each row equals 100%):

group 1: L 23%; V 48%; P 19%; N 10%
group 2: L 14%; V 77%; P 07%; N 01%
group 3: L 24%; V 44%; P 23%; N 09%
group 4: L 28%; V 37%; P 19%; N 17%

where L represents laryngo-pharyngalized; V, velarized; P, palatalized; and N, nasalized voice. These associations are only performed for the male subjects at this stage since the model data represent male fundamental frequency and vocal tract shape, and the survey data are not normalized. The distributions suggest the same pattern observed for female vowel clusters, in particular between groups 2, where most velarized assignments are made, and 4, where most nasalized assignments are made. For the single vowel phoneme /ʌ/, the difference in the respective test-reference assignment of velarized and nasalized models to groups 2 and 4 is even more pronounced. Highest association of velarized vowels is with group 2 (85%), and highest association of nasal vowels is with group 4 (50%). These results will be further tested for validity using normalized comparisons of combined female and male data with similarly normalized reference models of contrasting articulatory settings (see Hindle 1978).

8. MEASUREMENT OF FORMANTS

A word on technique of formant frequency measurement would be appropriate here. Monsen and Engebretson (1983: 89), comparing spectrographic techniques with linear prediction analysis of formant frequencies, have found that "for fundamental frequencies between 100 and 300 Hz, both methods are accurate to within approximately ±60 Hz for both first and second formants". The statistical procedures described above should be adequate to differentiate values within finely detailed enough frequency envelopes, but the question of how those original values were derived is both of greater relevance and of greater interest in helping to explain why formant frequencies differ and why they may be hard to measure. Monsen and Engebretson rightly point out that formant frequencies can be obscured by masking from the fundamental or broadening of bandwidths.

> To put it another way, all vowels are not created equal. It may be easier or harder to accurately recover the resonances of the vocal tract in the vowel sound wave depending on objective factors such as the fundamental frequency, the degree of nasalization of the vowel, or the position of the articulators (1983: 96).

The ILS peak-picking routine used here is observed to encounter masking problems of just this sort. Since the object of this study is to isolate those contributions of vocal tract resonance that are external to the individual vowels themselves, its results can help identify which articulatory configurations will affect otherwise identical vowels in a given way. Group 1 vowels produce greatest loss of second formant. This results in a smaller number of tokens that are acceptable for inclusion, and (perhaps not incidentally) wider deviation of the tokens that remain. Group 2 is the easiest group to measure, with all peaks and bandwidths clearly distinguishable, and has correspondingly the most coherent set of values. Group 3 is also not difficult to measure, but group 4 begins to demonstrate the appearance of an intermediate peak and widening bandwidths in all vowels for the largest number of speakers, both male and female. This secondary, usually higher amplitude peak overlaps in bandwidth with peak 1, and has been averaged into the computation of F1 since it is distinctly not associated with F2. This phenomenon occurs only rarely in other groups, and when it does the voice demonstrates pronounced nasality. It seems likely, therefore, that a generalized low back position of the articulators in group 1, evident in the F1,F2 values of remaining vowels, causes a decreasing peak 2 to merge with an increasing peak 1 for many tokens. The fronted and nasalized setting of group 4, implied by the damped but increased values of peak 1 due to the combined calculation, and the slightly higher values of peak 2, would not be apparent if these somewhat spectrally confusing tokens had to be eliminated.

9. FURTHER RESEARCH

In summary, these results enable us to verify that the relationship found in preliminary research between Vancouver vowels and vowels from the Houston, Texas, sample is due to the predominance of group 1 and 2 subjects in the Vancouver test sample, and their acoustic similarity to the Houston sample. Vancouver group 3 and 4 subjects, on the other hand, demonstrate differentiated (fronted or nasalized) acoustic results from groups 1 and 2, as well as from the Houston (velarized) values. Groups 2 and 4 are clearly separated, with group 2 closely resembling the Houston speech and group 4 most differentiated from the Houston sample.

Consistent vowel clustering values for female data can be related to tests of phonetically modelled vowels as follows: 1, laryngo-pharyngalized; 2, velarized; 3, palatalized; 4, nasalized. These can now be posited as tentative articulatory explanations for testing in revised research procedures. As not all vowels respond to a particular background setting in the same way, each individual vowel set will be compared to the four models one at a time.

The relationships presented using the methods described here will be tested in continuing work with refined techniques, including: (a) integration and comparison of female and male values using log-mean normalization procedures, and comparison with similarly normalized phonetic models; (b) comparison of each vowel set with the vowel values of the various model articulatory settings; (c) inclusion of diphthong information in calculations of generalized group vowel clustering, to test for similar patterns; (d) LPC synthesis and test-reference matching where original vocalic values are modified synthetically for comparison with the original values of vowels from contrasting groups.

REFERENCES

Blishen, Bernard R. (1967). A Socio-economic Index for Occupations in Canada. *Canadian Review of Sociology and Anthropology* 4: 41-53.

Blishen, Bernard R. and Hugh A. McRoberts. (1976). A Revised Socioeconomic Index for Occupations in Canada. *Canadian Review of Sociology and Anthropology* 13: 71-79.

Esling, John H. and B. Craig Dickson. (1985). Acoustical Procedures for Articulatory Setting Analysis in Accent. Pp. 155-170 in *Papers from the Fifth International Conference on Methods in Dialectology*. H.J. Warkentyne, ed. Victoria: University of Victoria.

Gregg, R.J., M. Murdoch, E. Hasebe-Ludt, and G. deWolf. (1981). An Urban Dialect Survey of the English Spoken in Vancouver. Pp. 41-65 in *Papers from the Fourth International Conference on Methods in Dialectology*. H.J. Warkentyne, ed. Victoria: University of Victoria.

Hindle, Donald. (1978). Approaches to Vowel Normalization in the Study of Natural Speech. Pp. 161-171 in *Linguistic Variation: Models and Methods*. David Sankoff, ed. New York: Academic Press.

Monsen, Randall B. and A. Maynard Engebretson. (1983). The Accuracy of Formant Frequency Measurements: a Comparison of Spectrographic Analysis and Linear Prediction. *Journal of Speech and Hearing Research* 26: 89-97.