# PAPERS FROM THE TENTH REGIONAL MEETING

# CHICAGO LINGUISTIC SOCIETY

APRIL 19–21, 1974

EDITED BY:

MICHAEL W. LA GALY

ROBERT A. FOX

ANTHONY BRUCK

CHICAGO LINGUISTIC SOCIETY

GOODSPEED HALL

1050 EAST 59TH STREET

CHICAGO, ILLINOIS 60637

TABLE OF CONTENTS

# An Experiment in Cross-Dialect Vowel Perception

Robert A. Fox
University of Chicago

## INTRODUCTION

Relatively recently, linguists have been concerned with the application of multidimensional scaling techniques to speech perception. Many of these studies have involved the scaling of vowel perception. Multidimensional scaling techniques can aid in discovering the number of perceptual criteria used in discriminating or identifying speech sounds (in our case vowel sounds) and correlating these perceptual criteria (or perceptual dimensions) with articulatory or acoustic properties of those sounds.

As stated by Terbeek and Harshman (1971) however, recent multidimensional scaling studies have dealt ONLY with questions of (1) the number of perceptual dimensions and (2) their physical correlates. These experiments (e.g. Pols et. al. 1969, Singh and Woods, 1970) were limited to attempts to ascertain the perceptual dimensions used by a homogeneous subject population (same language, no mentioned dialectal differences) to discriminate a set of stimulus vowel sounds. Terbeek and Harshman pose the question "To what extent is one's vowel perception related to one's native language? (p. 26)"

The present paper, in an almost corollary question to that investigated by Terbeek and Harshman presents an experiment attempting to ascertain the extent to which one's vowel perception is related to one's dialect. How different (if at all) are the perceptual spaces used in vowels perception as a function of dialect membership?

## EXPERIMENTAL DESIGN

An experiment was designed which required linguistically naive speakers of two different dialect groups to judge the similarity/dissimilarity of vowel diads. These similarity judgements were then analysed to discover the set of perceptual features used by the listeners in their diadic comparisons.

One dialect group consisted of General American speakers (henceforth GA) living in the Hyde Park community of Chicago. The second group consisted of 'Southern' speakers (henceforth S) residing in or around Norman, Oklahoma. The data from the GA group was collected by the author in the phonetics lab at the University of Chicago. The data from the S group was collected at the University of Oklahoma at Norman.[1] Conducting the experiment in the region in which the dialect is primarily spoken avoids major complications arising from code switching phenomena.

Since there would be possible differences in having a subject group judge their own dialect's vowels as opposed to another dialect's vowels, the vowel stimuli consisted of 2 sets of stimulus vowel pairs, one set produced by a speaker of each of the test dialects. These will be referred to as the GA vowel set and the S vowel set. Seven vowels were chosen as stimulus vowels, the particular choice motivated by the desire to present vowel stimuli similar to vowels found in both dialects. The target vowels were / i, e$^{I}$, a, u, ai, o æ /. Figures 1 and 2 give average formant values for the stimulus vowels heard in each set. For the purpose of naturalness, both in production and perception, each vowel was produced in the context / __ t /.

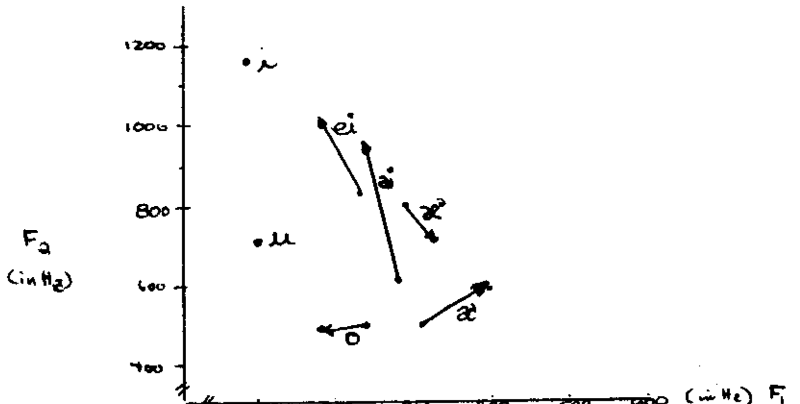


Figure 1. Average Formant values for the S vowel set.

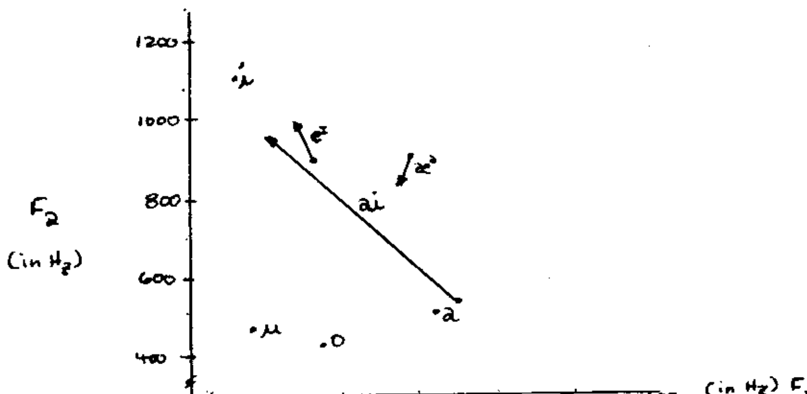


Figure 2. Average Formant values for the GAvowel set.

The S dialect differs from the GA dialect mainly in that it exhibits a much greater degree of diphthongization and centralization. However, though they do differ phonologically -- that difference is not as great as might be hoped for. This fact will become crucial later in discussion of experimental results.

There were 19 subjects, 11 in the GA dialect group and 8 in the S dialect group. Each listener judged 168 vowel pairs (84 pairs for each dialectal vowel set). The subjects expressed their similarity/dissimilarity judgements on a discrete nine point scale:

S . . . . . . . . . . D

(similar to the technique used by Mohr and Wang, 1965). The subjects heard the vowel stimuli over high quality speakers in small groups (the experimental conditions were similar in both Norman, Ok. and Hyde Park ).

The experiment took one half hour to complete, each vowel pair presented 7 1/2 seconds apart. Listeners were instructed to give their first impressions concerning the similarity judgements and four sample vowel pairs were included before the start of the test to acquaint the subjects with the judgement task.

ANALYSIS

The responses obtained from each listener in the similarity/dissimilarity judgements were accumulated into vowel by vowel tables, the cells in each table representing perceptual distance between two vowels. INDSCAL (Carroll and Chang, 1970) was the multidimensional analysis program utilized in this paper, converting these perceptual distance tables into a pattern of points in a perceptual space -- each point representing a vowel.

The number and rotation of dimensions (which represent the perceptual criteria subjects are using in their perceptual judgements) of which the perceptual space is composed is uniquely determined by INDSCAL, and the dimensions extracted are psychologically "real" (this is only subject to the caveat that the dimensions are real for these subjects, in this experiment, judging this set of vowels). INDSCAL is designed to discover both the distances in this perceptual space among each of the vowels and the degree to which each subject relies on each particular perceptual dimension in his/her perceptual judgements.

If a difference exists in the perceptual space of one of the dialects compared with the other, then in an analysis of the combined data this difference should be reflected in the "person loadings" (the weights given by the program for each subject in regard to each dimension).[2] We shall then be concerned with viewing the values of the person loadings or weights as a function of dialect group membership.

RESULTS

For the combined listener response data for the S vowels, INDSCAL indicated that three dimensional solution was optimal. Two of the dimensions correspond to a ± Front offset and ± High. This can best be illustrated by Figure 3, a plot of the plane described by DI by DII. Dimension III does not correspond to any

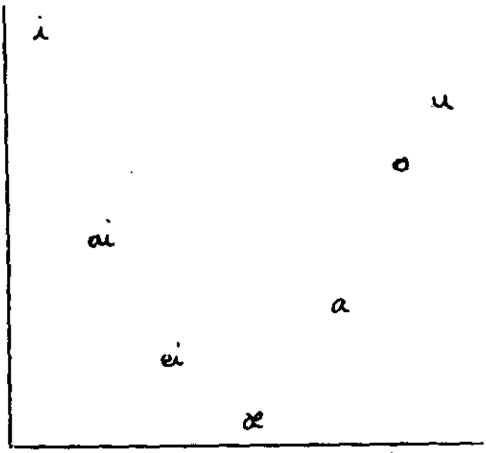


Figure 3. The Vertical axis is DI (± High) and the Horizontal axis is DII (± Front Offset). Please compare to S Formant space, Figure 1.

easily defined feature, but rather separates the diphthong /ai/ from the other six vowels; we shall consider this a ± Back/Front glide dimension, for lack of any further evidence. Dimension III can be seen in Figure 4, shown as an horizontal scale along which the vowels are ordered.

Figure 4. Dimension III (± Back/Front Glide)

The person loadings for each of these three dimensions are given in table 1 and the distribution of values is schematically illustrated in figure 5.

| | DI | DII | DIII | |
|---|---|---|---|---|
| 1. | .778 | .494 | .389 | |
| 2. | .873 | .216 | .438 | |
| 3. | .733 | .480 | .480 | |
| 4. | .507 | .501 | .703 | |
| 5. | .831 | .500 | .243 | |
| 6. | .691 | .525 | .497 | Person loadings for GA group |
| 7. | .771 | .572 | .280 | |
| 8. | .632 | .472 | .614 | |
| 9. | .755 | .462 | .466 | |
| 10. | .659 | .650 | .365 | |
| 11. | .658 | .710 | .251 | |

Table 1. Person loadings for subjects judging S vowel set.

| | DI | DII | DIII |
|---|---|---|---|
| 1. | .673 | .528 | .519 |
| 2. | .780 | .518 | .346 |
| 3. | .685 | .538 | .490 |
| 4. | .731 | .397 | .557 |
| 5. | .678 | .475 | .262 |
| 6. | .814 | .514 | .274 |
| 7. | .551 | .597 | .584 |
| 8. | .814 | .486 | .321 |

Person loadings for S group

Table 1. (continued)

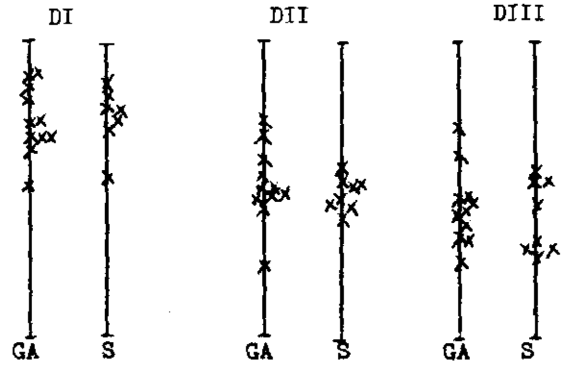


Figure 5. Schematic comparison of the person loadings for each group for each dimension.

No significant differences can be seen between the person weights given to the dimensions by the two dialect groups. Perhaps what is most evident is the range of individual variation among the 19 listeners.

For the combined listener judgement for the GA vowels, evaluation of the perceptual space is not quite as easy. Due to the small number of stimuli used, it is difficult to decide which of the 2, 3, or 4 dimensional solutions is the "correct" one. Since it is the goal of this paper to examine person loadings on the dimensions extracted, and not to make claims about the vowel space itself, we will view the 4 dimensional solution only. The values of the person loadings exhibit similar distribution in all three solutions.

These four dimensions are shown in Figure 6 (represented as horizontal scales). Dimension I is best interpreted as ± Round, while Dimensions II and III are similar to those seen in the S vowel set results, ± Front Offset and ± Back/Front Glide, respectively. Dimension IV defies an easy label and no satisfactory interpretation is in evidence. The plane defined by DIII and DIV (see Figure 7) is perhaps more interpretable in terms of clustering (± High Front Offset), but I shall leave this to the reader's own speculation.

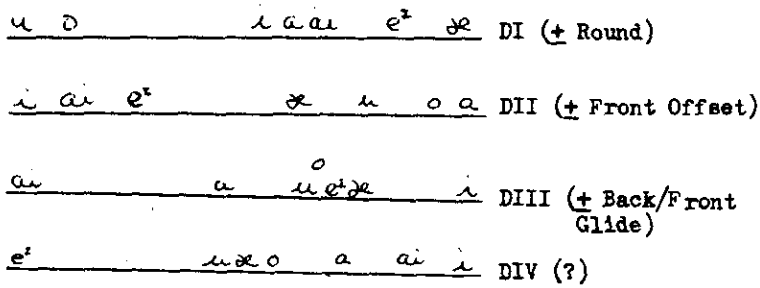


Figure 6. The four-dimensional solution for the GA vowel set judgement results.

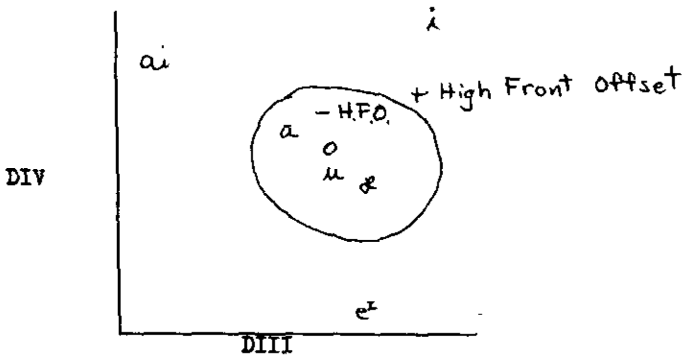


Figure 7. The DIII by DIV plane, best interpreted perhaps as + High Front Offset.

The person loadings for these four dimensions are given in table 2, and are schematically illustrated in Figure 8. Similar

| | DI | DII | DIII | DIV | |
|---|---|---|---|---|---|
| 1. | .664 | .516 | .247 | .486 | |
| 2. | .356 | .192 | .656 | .638 | |
| 3. | .764 | .573 | .228 | .195 | |
| 4. | .568 | .436 | .578 | .388 | |
| 5. | .581 | .547 | .370 | .473 | |
| 6. | .521 | .332 | .432 | .658 | Person loadings |
| 7. | .346 | .831 | .251 | .350 | for GA subjects |
| 8. | .567 | .666 | .492 | .147 | |
| 9. | .663 | .671 | .371 | .176i* | |
| 10. | .650 | .464 | .572 | .420 | |
| 11. | .718 | .537 | .382 | .228 | |
| | | | | | |
| 1. | .401 | .732 | .264 | .475 | |
| 2. | .488 | .557 | .481 | .467 | |

Table 2. Person loadings for subjects judging GA vowel set.

| | | | | |
|---|---|---|---|---|
| 3. | .597 | .518 | .353 | .503 |
| 4. | .697 | .479 | .399 | .353 |
| 5. | .494 | .749 | .346 | .287 |
| 6. | .689 | .558 | .465 | .061i* |
| 7. | .556 | .496 | .529 | .410 |
| 8. | .568 | .418 | .463 | .537 |

Person loadings for S subjects.

Table 2. (continued)
*These represent imaginary person loadings and indicate a partial breakdown of the spacial model. In general it shows that a particular subject is not utilizing that particular dimension at all.

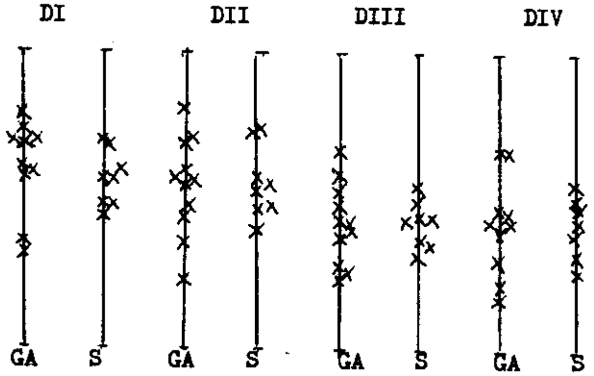


Figure 8. Schematic comparison of the person loadings for each group for each dimension.

to the results from the previous vowel stimuli set, there is no significant group difference, while individual variation is quite prevalent.

DISCUSSION

I shall not discuss the dimensions which were extracted in the experiment <u>per se</u>, for we will be most concerned with the person loadings. However, I would like to mention, that though some of the dimensions may seem a bit unusual, it is the case that no one has adequately investigated perceptual spaces using both monothongs and diphthongs as the vowel stimuli -- as a result, no one really knows what perceptual dimensions would be extracted. This type of experimentation is very needed and would be very elucidating.

This experiment was designed to find out if the perceptual space of listeners is, in part, determined by dialect group membership. This question is based on the hypothesis that the phonological structure of one's native speech partially determines his/her perceptual space. Subsequently, if the speech of two groups of listeners differ phonologically, then there will be a concomitant difference in their respective perceptual spaces.

Our failure to find significant difference in the person loadings need not be interpreted as proof that this hypothesis is incorrect; indeed it would be difficult to reach that conclusion in light of the results presented in Terbeek and Harshman (1971). How then should we view our results?

If the difference in the perceptual space is determined by the degree of phonological difference between the speech of two groups of speakers -- then one would logically expect a continuum of perceptual space differences based on the degree of phonological differences in the language between the two goups. In regard to this experiment, one would expect the relatively small amount of phonological variation form one dialect to another to produce equally small differences in the perceptual space utilized by each dialect group (reflected in our experiment in the person loadings). As pointed out in the paper, however, there is a great deal of individual variation in these person loadings -- individual variation which would tend to obscure any small differences in thee person loadings which we would otherwise expect.

To further test this hypothesis, one needs to conduct a similar experiment using dialects exhibiting greater phonological difference than the ones used in this experiment, great enough to be seen above the "noise" produced by individual variation.

* I would like to thank Dale Terbeek for his helpful comments on all phases of this experiment.

FOOTNOTES

1. I would like to thank Ken McGraw, Department of Psychology, University of Oklahoma for collecting the S group data.
2. Difference in person loadings are not the only way in which two subject groups can differ in regard to their perceptual spaces. One group's perceptual space may differ from another's in regard to number, content, or rotation of dimensions. The method of analysis used in this paper is designed to find if a difference exists -- not to necessarily discover the nature of that difference.

BIBLIOGRAPHY

1. Carroll, J. D. and J.J. Chang, "Analysis of individual differences in multidimensional scaling via an N-way generalization of 'Eckart-Young' decomposition," Psychometrika 35, 283-319, 1970.
2. Mohr, B., and W. Wang, "Perceptual Distance and the Specifiction of Phonological Features," Phonetica 18:31-45, 1968.
3. Pols, L.C.W., van der Kamp, L.J.Th. and R. Plomp, "Perceptual and Physical Space of Vowel Sounds," J. Acoust. Soc. Amer. 458-467, 1969.
4. Singh, S. and G. Woods, "Perceptual Structure of 12 American Vowels," J. Acoust. Soc. Amer., 49.6, 1861-1866, 1970.
5. Terbeek, D. and R.A. Harshman, "Cross-Language Differences in the Perception of Natural Vowel Sounds," UCLA Working Papers in Phonetics 19, June 1971.