MULTI-DIMENSIONAL SCALING OF LISTENER RESPONSES TO COMPLEX AUDITORY STIMULI

Z. S. Bond, Donald Fucci, Verna Stockmal, Douglas McColl

Ohio University

ABSTRACT

This study explored the attributes of languages to which listeners attend, using magnitude estimation and multi-dimensional scaling techniques. In magnitude estimation, listeners assign any numerical value to a set of stimuli. In response to the question: How similar is this language to English? fifty college students assigned numerical values to spoken samples of foreign languages. The languages represented Europe, Asia and Africa. Differences between the mean ratings for each language and English were used to construct a proximity matrix which was submitted to MDS analysis. The optimum solution employed three dimensions. The first dimension was interpreted as "familiarity," the second as "speaker affect," and the third as "prosodic pattern." The MDS maps suggest that listeners were using English as a standard of comparison to the acoustic-phonetic properties of other languages. The maps resulted from the relationship between each language and the standard, and speaker and language characteristics which listeners found salient.

1. INTRODUCTION

As complex auditory stimuli, spoken samples of languages differ among themselves on a multitude of properties. Yet, when asked to identify samples of foreign languages or to discriminate between them, listeners are able to perform these tasks at better than chance levels (2, 7, 8, 11). In order to provide such judgments, listeners must be attending to the acoustic-phonetic properties of the language samples, in the phtrase suggested by Muthusamy, et al., their "acoustic signatures."

The purpose of this study is to investigate magnitude estimation as a tool for obtaining insight into the perceptual dimensions listeners use in responding to spoken foreign languages. Magnitude estimation has been found to be effective in other diverse and complex tasks such as scaling the loudness of unfamiliar languages (4) and grammaticality judgments (1). Briefly, in magnitude estimation subjects assign any numerical value to a set of stimuli which differ among themselves on one or more properties. From these numerical values, it is possible to reconstruct the physical or cognitive dimensions on which subjects are making judgments.

In Fucci, McColl, Bond and Stockmal (5), subjects provided magnitude estimation judgments to spoken samples of foreign languages consistently, suggesting that there responses reflected orderliness or patterning in the stimuli. Such magnitude estimation results may be used as input to further analysis of stimulus dimensionality (Coxon, 1982). Extensions of magnitude estimation are also supported by Zahorik (1997), who found equivalence between magnitude estimation and other modes of eliciting subject judgments in psychophysical tasks.

2. PREVIOUS MDS ANALYSES

Two previous studies have used multi-dimensional scaling (MDS) to assess the perceptual structures underlying differences between foreign languages. In both of these, listeners responded using rating scales.

Stockmal, Muljani and Bond (12) report two experiments. In the first, American listeners provided same-different judgments of paired samples of foreign languages as well as a rationale for their judgments. The judgments of same-language pairs were used to construct a proximity matrix which was submitted to MDS. Only six languages were used int he sample, Arabic, Chinese, Indonesian, Japanese, Russian and Spanish. In a two-dimensional solution, the first dimension was related to fundamental frequency excursions or pitch patterns; in this dimension, Chinese and Japanese formed one group, and Arabic Russian and Indonesian formed another. The second dimension was related to speaker rather than language characteristics. In their second experiment, Stockmal, et al. used similarity ratings of paired samples of the same six languages. Ratings on different-language pairs were used as input to MDS. A two-dimensional solution showed groupings consistent with the first MDS analysis.

The second study employing MDS analysis was quite different. Nazzi (9) asked listeners, all native speakers of French, to judge the similarity of paired samples of ten different languages. The languages included in the study were Arabic, Dutch, English, French, Italian, Japanese, Spanish, Swedish, Turkish, and Wolof. The language samples were low-pass filtered at 180 Hz before they were presented to listeners. The matrix obtained from listener ratings yielded a three-dimensional solution. Nazzi interpreted the first dimension as familiarity, in that it separated the European languages from the non-European languages. The second dimension separated languages which exhibited syllable rhythm from stress rhythm, though this classification was clearly apparent only for the European languages. Nazzi suggested that the listeners may have had little opportunity to develop a perceptual representation of unfamiliar languages in the short exposure time they had available.

In this study, we employed the numerical responses obtained in Fucci, et al. (5) as input for MDS analysis to obtain insight into the perceptual dimensions employed by subjects in evaluating the similarity of foreign languages in comparison with a perceptual anchor, English.

3. METHOD

3.1. Subjects.

Fifty college students participated in the study. Subject ages ranged from 18 to 37 years, with a mean age of 21 years. All subjects were native speakers of English and reported normal speech and hearing.

3.2. Materials.

The materials were ten-second spoken samples of each of eleven test languages, as produced by a male speaker. The languages, representing Europe, Asia and Africa, consisted of Akan, Arabic, Chinese, English, French, German, Hebrew, Japanese, Latvian, Russian and Swahili.

3.3. Procedure.

Each listener was tested individually while seated comfortably in a sound treated booth. The listeners heard the language samples through TDH-39 headphones at 65 dB SPL. After receiving instruction about magnitude estimation, the listeners were asked to assign a numerical value to each language sample in two tasks. The first task required the listeners to judge as to how similar each language is to English. In the second task, listeners assigned a value to each sample according to how well they liked it. Further details concerning the procedure are given in Fucci, et al., 1997.

3.4. MDS analysis.

Listener judgments in the first task were used as the basis for MDS analysis. Listener responses were converted to proportions on a ten-point scale. The differences between the mean ratings for each language and English were used to construct a proximity matrix. Since all listeners uniformly gave their highest score to the English sample and since the English served as a perceptual anchor, it was not entered in the MDS analysis.

4. RESULTS

The solution to the matrix was computed by ALSCAL. The optimum solution employed three dimensions. In this solution, RSQ = .921 and Stress = .106, a value considered "satisfactory" according to Jassem and Lobacz (6). The three-dimensional solution also fits their criterion, that the number of dimensions be equal to the integer obtained by dividing the number of scaled items by 4.

Figure 1 shows the languages in a 2-dimensional space formed by the first two dimensions (D1 and D2).

D1 can be interpreted as "familiarity." French and German languages commonly studied in high school and college, are located at one extreme of D1 while Akan and Arabic are at the other extreme. Russian, Hebrew and Japanese are more familiar than Latvian and Swahili. this dimension probably shows the effect of the question used to elicit magnitude estimating judgments.

The second dimension, D2, reflects individual speaker, rather than language characteristics. The speakers who provided the language samples below the horizontal axis read in a dramatic, dynamic manner while the speakers providing the language samples above the axis produced precise and formal readings. the second dimension can be interpreted as "speaker affect."

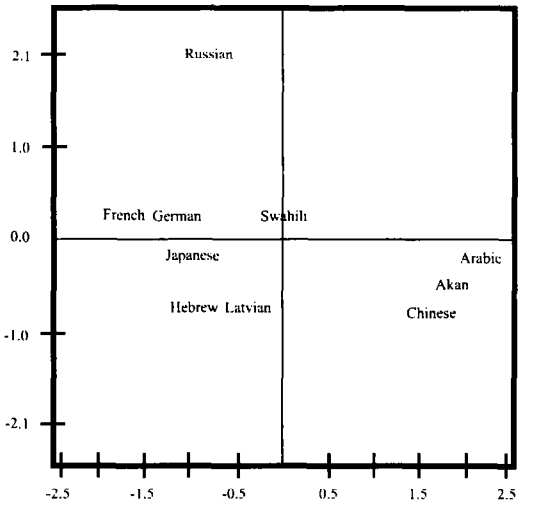

Figure 1: Dimension 1 and dimension 2 of a 3-dimensional analysis for language proximity data based on magnitude estimation judgments. The horizontal dimension represents familiarity. The vertical dimension represents speaker affect.

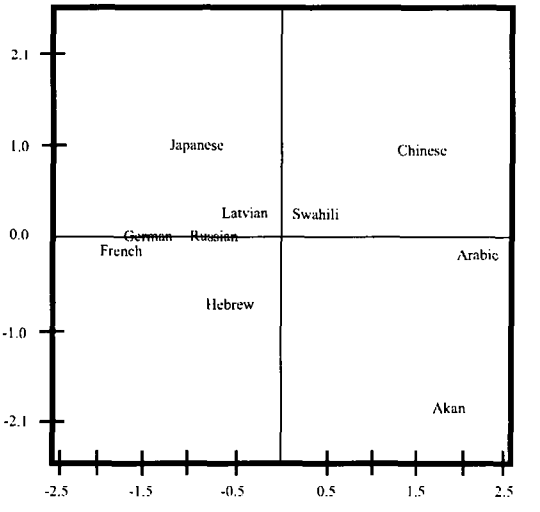

Figure 2: Dimension 1 (horizontal) and dimension 3 (vertical) of a 3-dimensional analysis for language proximity data based on magnitude estimation judgments. The vertical dimension can be interpreted as prosodic pattern.

Figure 2 shows the languages in a 2-dimensional space formed by D1 and D3. The third dimension appears to reflect judgments based on prosody, that is, on rhythm in combination with fundamental frequency patterns. Chinese and Japanese lie at one extreme. Both of these languages are characterized by fundamental frequency excursions over syllables. At the other extreme is Akan, characterized by syllable rhythm and by prominent vowel nasalization. The remaining languages lie in a mid-range, though they cluster only partially into stress vs. syllable rhythm types. Latvian clusters with Swahili, perhaps because both languages employ syllable rhythm and vowel quantity. German, Russian, Arabic and Hebrew employ stress rhythm. French is out of place on this dimension, in that it is considered to employ syllable rhythm. In spite of some difficulties, the third dimension of the MDS solution can be interpreted as "prosodic pattern."

5. DISCUSSION

Of the dimensions which might have served as the basis for listener judgments, the three which emerged in this study are familiarity, speaker affect, and prosodic pattern. Familiarity and rhythmic class -- essentially equivalent to prosodic pattern -- were also suggested by Nazzi (9). Individual speaker characteristics emerged in the MDS analysis reported by Stockmal, et al. (12). The solutions provided by MDS seem to be relatively robust, even when many aspects of the experimental situation are changed.

The way the task was defined for the listeners probably influenced the aspects of the language samples to which they attended. The listeners were asked to judge how similar each language was to English rather than judging the similarity of pairs of languages. In a sense, the task could be analyzed as an AX task, with English serving as the standard or anchor. This question focused the listeners' attention to the familiarity of the language samples. Since magnitude estimation readily permits listener ratings in response to different questions, it may prove an excellent techniqure for obtaining listener judgments which can then be used in further analyses.

6. REFERENCES

1. Bard, E. G., Robertson, D. & Sorace, A. "Magnitude estimation of linguistic acceptability", *Language, Vol. 72,* 1996, p 32-68.

2. Bond, Z. S. & Fokes, J. "Identifying foreign languages", *Proceedings of the XIIth International Congress of Phonetic Sciences, Vol. 2,* 1991, p 198-201.

3. Coxon, A. P. M . *The Users Guide to Multidimensional Scaling,* Heinemann Educational Books, Exeter, New Hampshire, 1982.

4. Fucci, D., Bettagere, R., Gonzales, M. D. & Reynolds, M. E. "Language familiarity in magnitude-estimation scaling of loudness by young adults", *Perceptual and Motor Skills, Vol. 80,* 1995, p 419-423.

5. Fucci, D., McColl, D., Bond, Z. & Stockmal, V. "Perceptual dimensions underlying judgment of languages", *Perceptual and Motor Skills, Vol. 85,* 1997, p 1468-1470.

6. Jassem, W. & Lobacz, P. "Multidimensional scaling and its applications in a perceptual analysis of Polish consonants", *Journal of Quantitative Linguistics, Vol. 2,* 1995, p 105-124.

7. Lorch, M. & Meara, P. "How people listen to languages they don't know", *Language Sciences, Vol. 11,* 1989, p 343-353.

8. Muthusamy, Y. K., Barnard, E. & Cole, R. A., "Automatic language identification: A review/tutorial", *IEEE Magazine, October,* 1994, p 1-16.

9. Nazzi, T. "Du rythme dans l'acquisition et le traitment de la parole", *Ph.D. Dissertation, Ecole des Hautes Etudes et Sciences Sociales,* 1997, Paris.

10. Stockmal, V. "Discrimination of unknown foreign languages in spoken utterances: A developmental study", *A. Thesis, Ohio University,* 1995, Athens, OH.

11. Stockmal, V., Muljani, D. & Bond, Z. S. "Can children identify samples of foreign languages as same or different?", *Language Sciences, Vol. 16,* 1994, p 237-252.

12. Stockmal, V., Muljani, D. & Bond, Z. S. "Perceptual features of unknown foreign languages as revealed by multi-dimensional scaling", *Proceedings of the International Conference on Spoken Language Processing, Vol. 4,* 1996, p 1748-1751.

13. Zahorik, Pavel, "Scaling perceived distance or virtual sound source?", *Paper presented at Acoustical Society of America,* State College, PA, 1997.